Integrative analysis of the influence of aluminum stress on the expression of the metabolic phenotype of cultivated rice (*Oryza sativa* L.)

ómicas

Chrystian Camilo Sosa Arango B.Sc

UNIVERSIDAD DEL QUINDÍO

El conocimiento es de todos
Minciencias

February 2023

Pontificia Universidad
JAVERIANA
Cali

Rice cultivation as a key element in food security

- Rice is the most widely used crop for caloric intake in Latin America and Asia
- Rice is an important part of sustainable development strategy

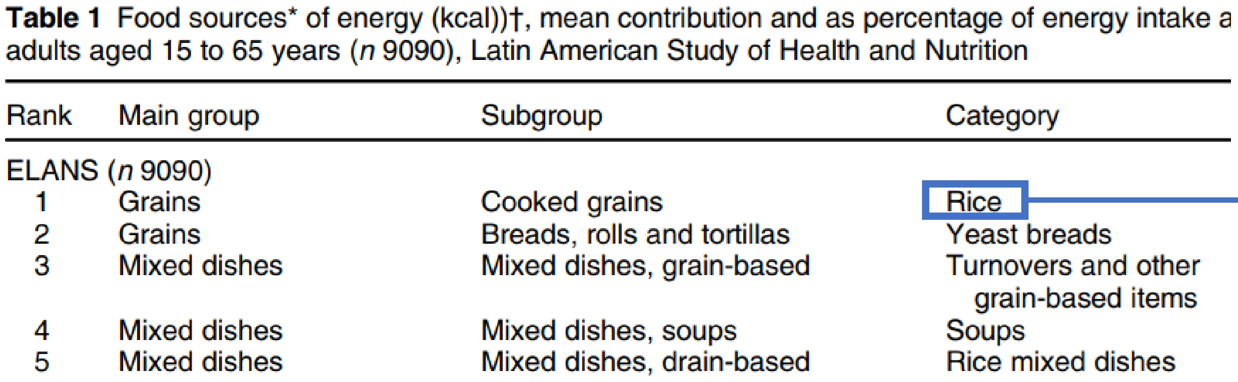
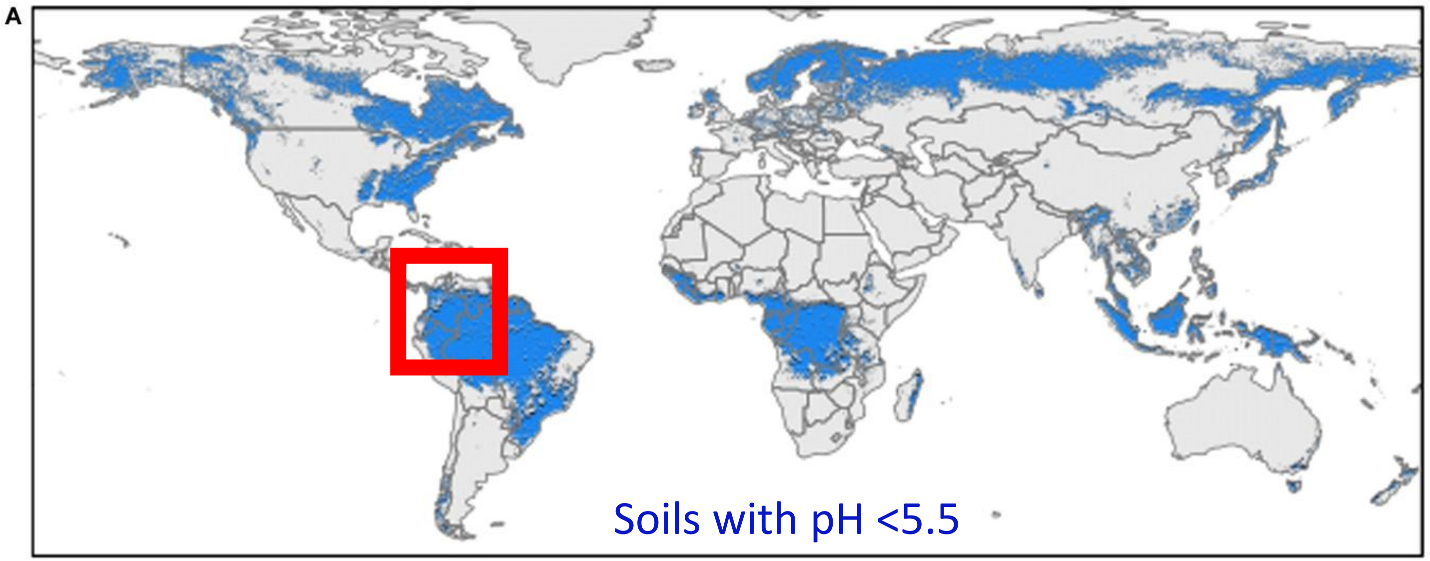


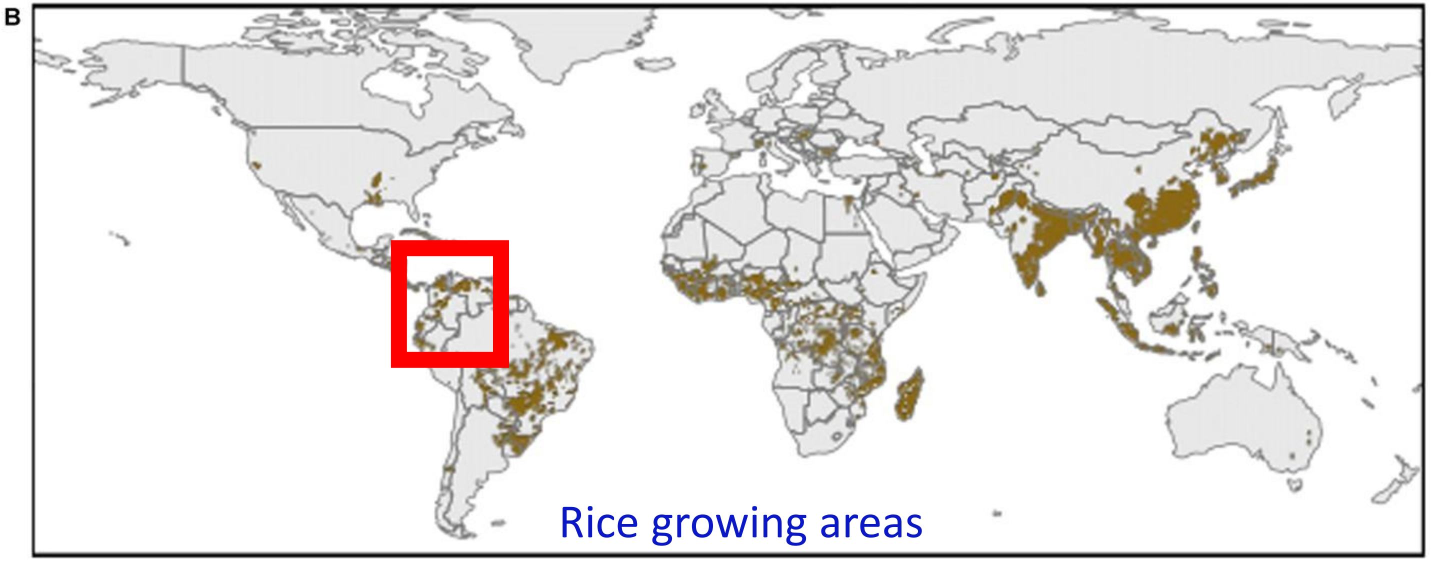
Table 1 Food sources* of energy (kcal))†, mean contribution and as percentage of energy intake a adults aged 15 to 65 years (*n* 9090), Latin American Study of Health and Nutrition

Rank	Main group	Subgroup	Category
ELANS (*n* 9090)			
1	Grains	Cooked grains	Rice
2	Grains	Breads, rolls and tortillas	Yeast breads
3	Mixed dishes	Mixed dishes, grain-based	Turnovers and other grain-based items
4	Mixed dishes	Mixed dishes, soups	Soups
5	Mixed dishes	Mixed dishes, drain-based	Rice mixed dishes

Fisberg et al., (2021); https://bit.ly/3GaLu0K; https://bit.ly/3EkWQir



Soils with pH <5.5



Rice growing areas

Gallo-Franco et al., (2020)

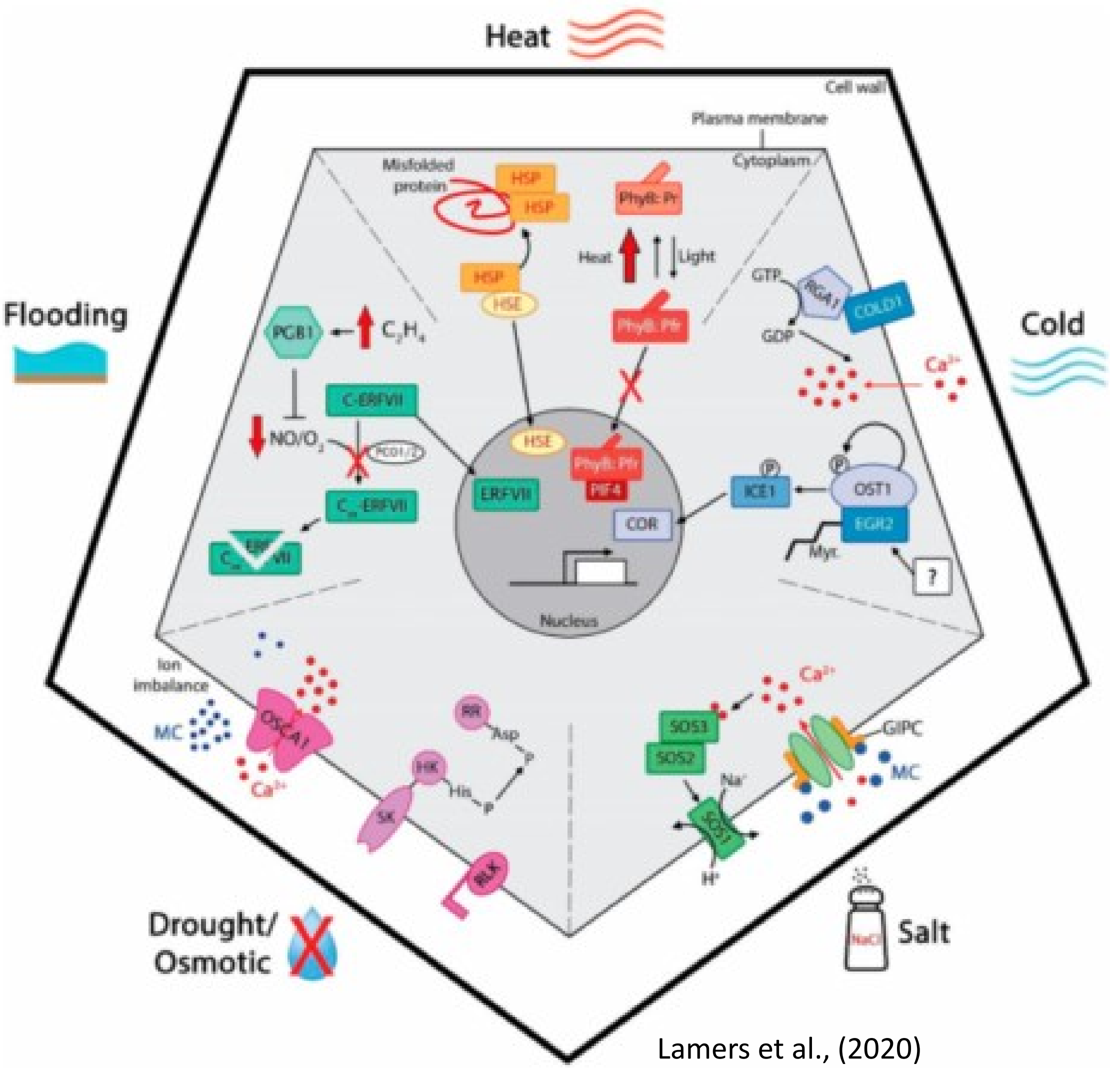
The problem of acidic soils

- pH <5
- About 50% of the arable land is in acid soils.
- 85% of acid soils are located in the torrid region of South America
- Aluminum is the most abundant metal on earth crust

Main types of abiotic stresses exist in plants

- Temperature
- Cold
- Salt
- Drought
- Flooding
- Heavy metals

The physiological response is complex!



Lamers et al., (2020)

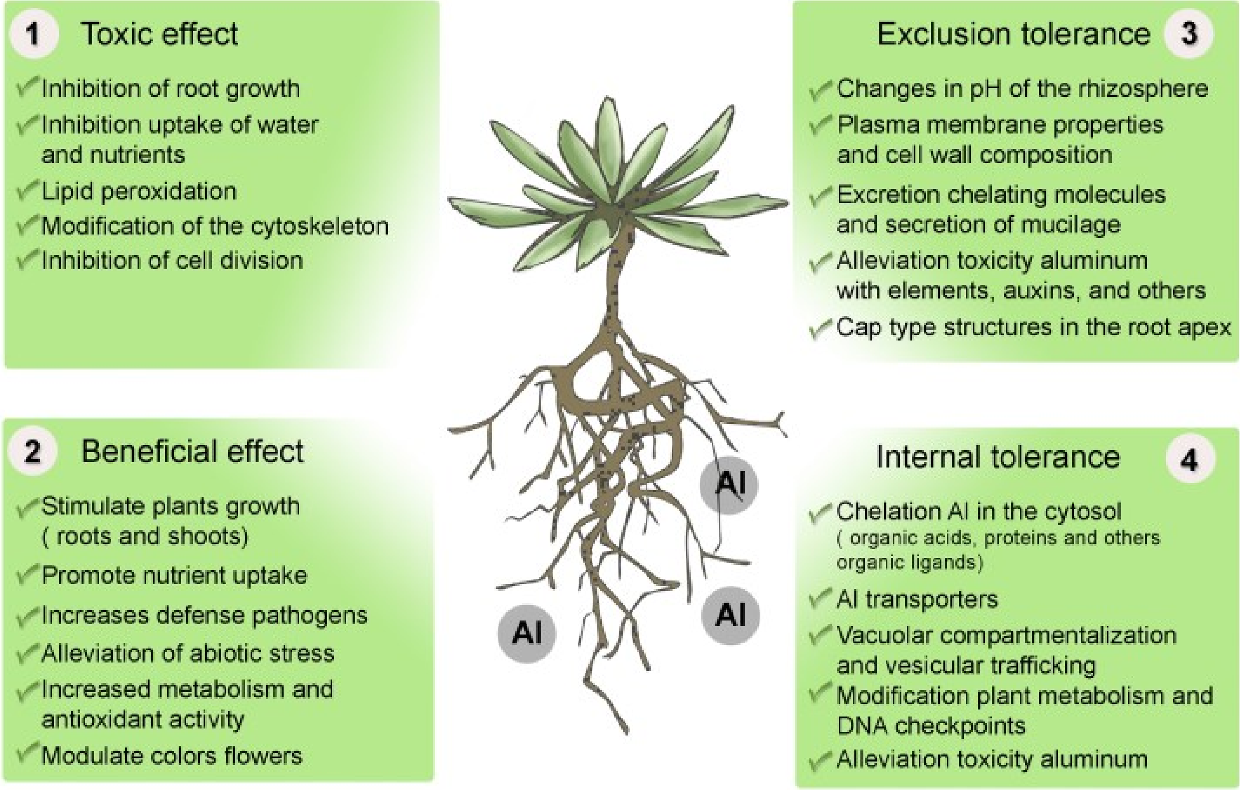
Plants have mechanisms that allow them to tolerate aluminotoxic conditions!

- Toxic effects
- Beneficial effects
- External exclusion mechanisms
- Internal mechanisms

Al^{3+}
$Al(OH)^{2+}$
$Al(OH)^{4-}$
$[Al(H_2O)_6]^{3+}$

Phytotoxic chemical species

Rice can be affected by aluminum!



1 Toxic effect

- ✓ Inhibition of root growth
- ✓ Inhibition uptake of water and nutrients
- ✓ Lipid peroxidation
- ✓ Modification of the cytoskeleton
- ✓ Inhibition of cell division

2 Beneficial effect

- ✓ Stimulate plants growth (roots and shoots)
- ✓ Promote nutrient uptake
- ✓ Increases defense pathogens
- ✓ Alleviation of abiotic stress
- ✓ Increased metabolism and antioxidant activity
- ✓ Modulate colors flowers

Exclusion tolerance 3

- ✓ Changes in pH of the rhizosphere
- ✓ Plasma membrane properties and cell wall composition
- ✓ Excretion chelating molecules and secretion of mucilage
- ✓ Alleviation toxicity aluminum with elements, auxins, and others
- ✓ Cap type structures in the root apex

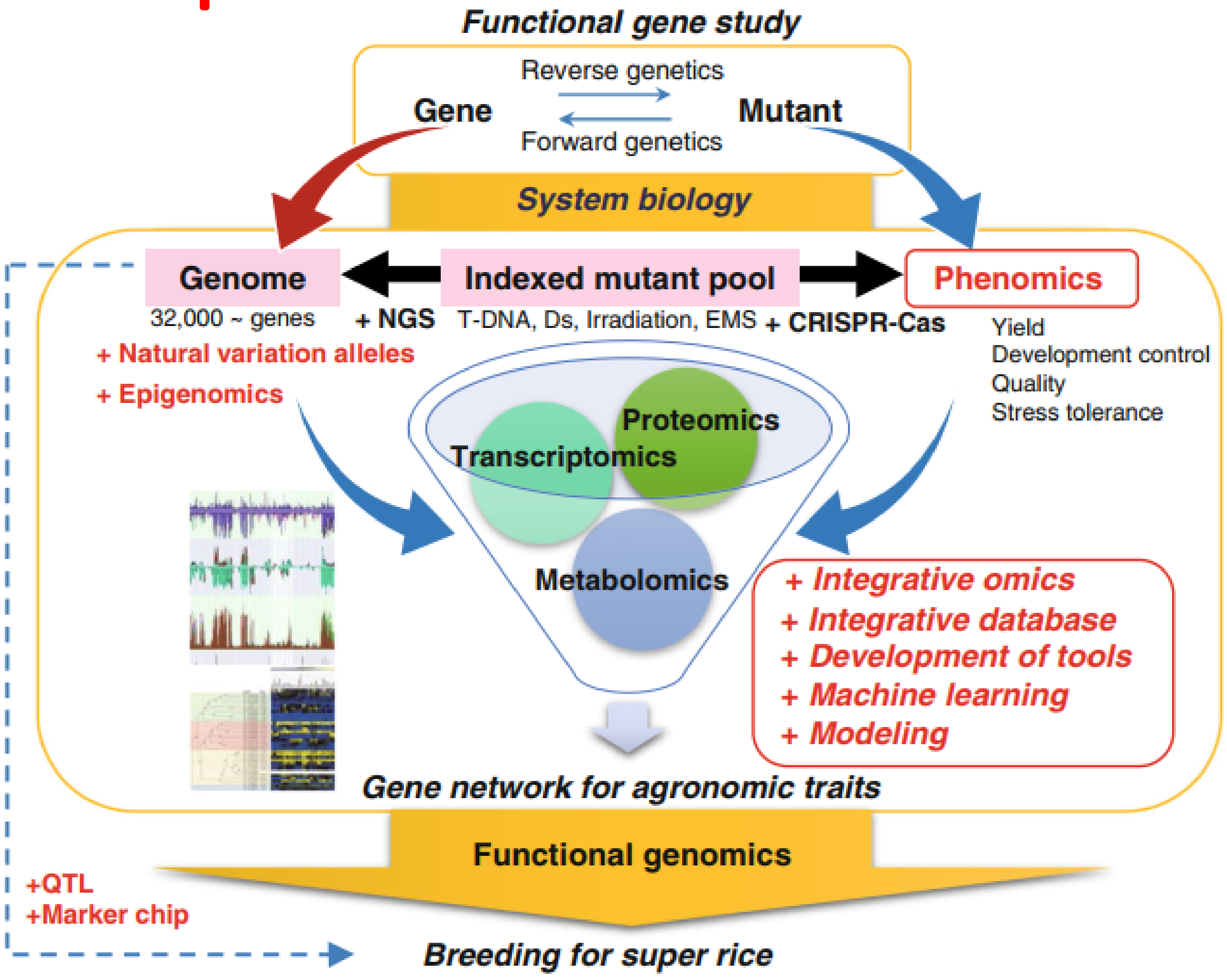
Internal tolerance 4

- ✓ Chelation Al in the cytosol (organic acids, proteins and others organic ligands)
- ✓ Al transporters
- ✓ Vacuolar compartmentalization and vesicular trafficking
- ✓ Modification plant metabolism and DNA checkpoints
- ✓ Alleviation toxicity aluminum

Famoso et al., (2010); Bojórquez Quintal et al., (2017)

Rice is a model species for plant stress studies!

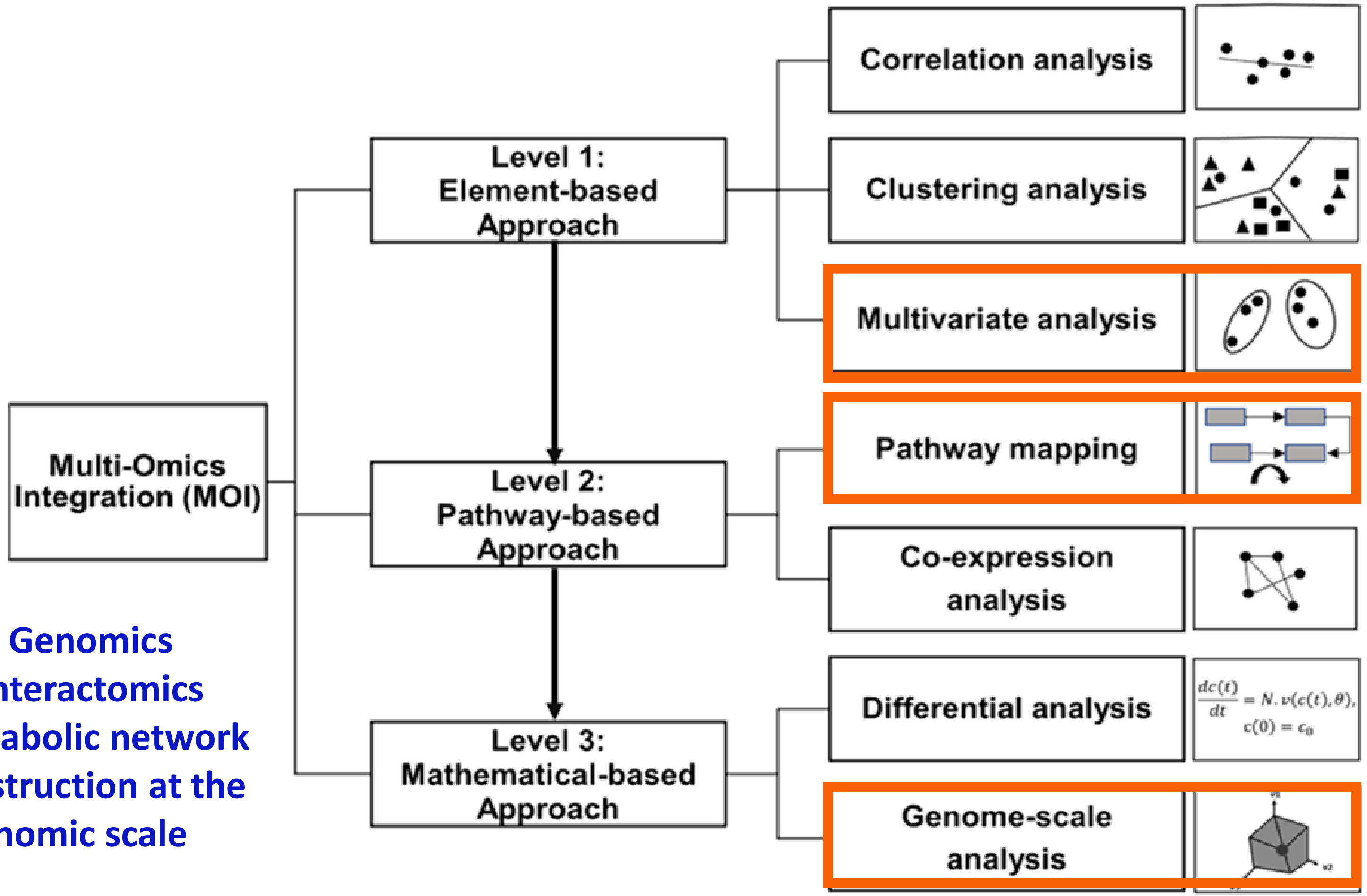
- 3000 rice accessions has been sequenced
- Information for about 10000 coding proteins is available
- Available information for aluminum stress conditions
- There is a genetics core that act by identified stress types!



Hong et al., (2019)

What do we know about the interaction of rice with aluminum?

- There is a molecular signaling network in different types of stresses
- Genes identified via omics technologies
- **It is unknown how the interaction between biological processes, proteins, and metabolites related to aluminum tolerance takes place.**



Multi-Omics Integration (MOI)

Level 1: Element-based Approach

Level 2: Pathway-based Approach

Level 3: Mathematical-based Approach

Correlation analysis

Clustering analysis

Multivariate analysis

Pathway mapping

Co-expression analysis

Differential analysis

$\frac{dc(t)}{dt} = N.v(c(t),\theta)$, $c(0) = c_0$

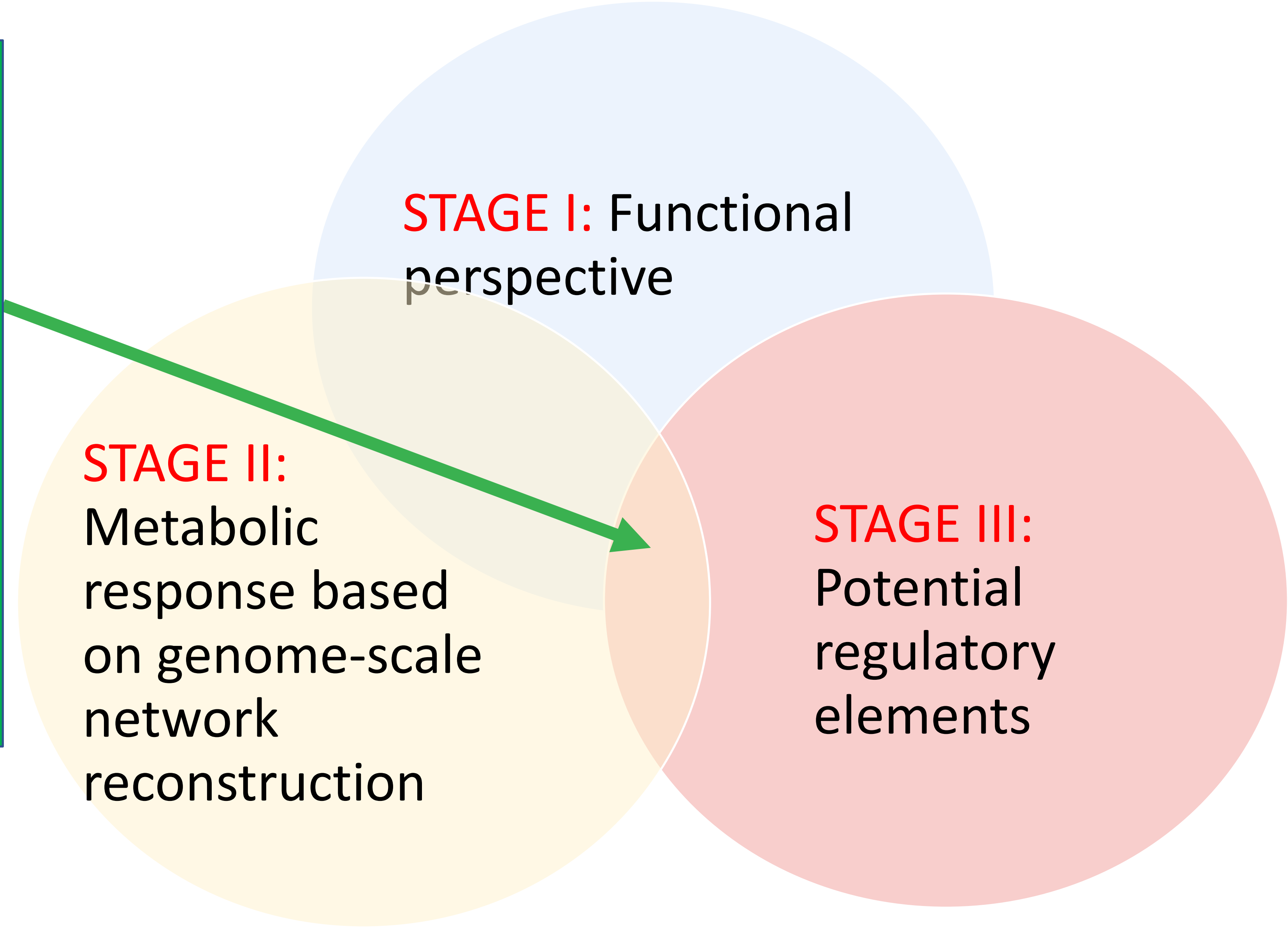
Genome-scale analysis

1. Genomics
2. Interactomics
3. Metabolic network reconstruction at the genomic scale

Jamil et al., (2020)

Understand the aluminum effect over the **rice cellular response** to aluminotoxic conditions

To **model the aluminum effect** over the cellular metabolism of cultivated rice (*Oryza sativa* L.),

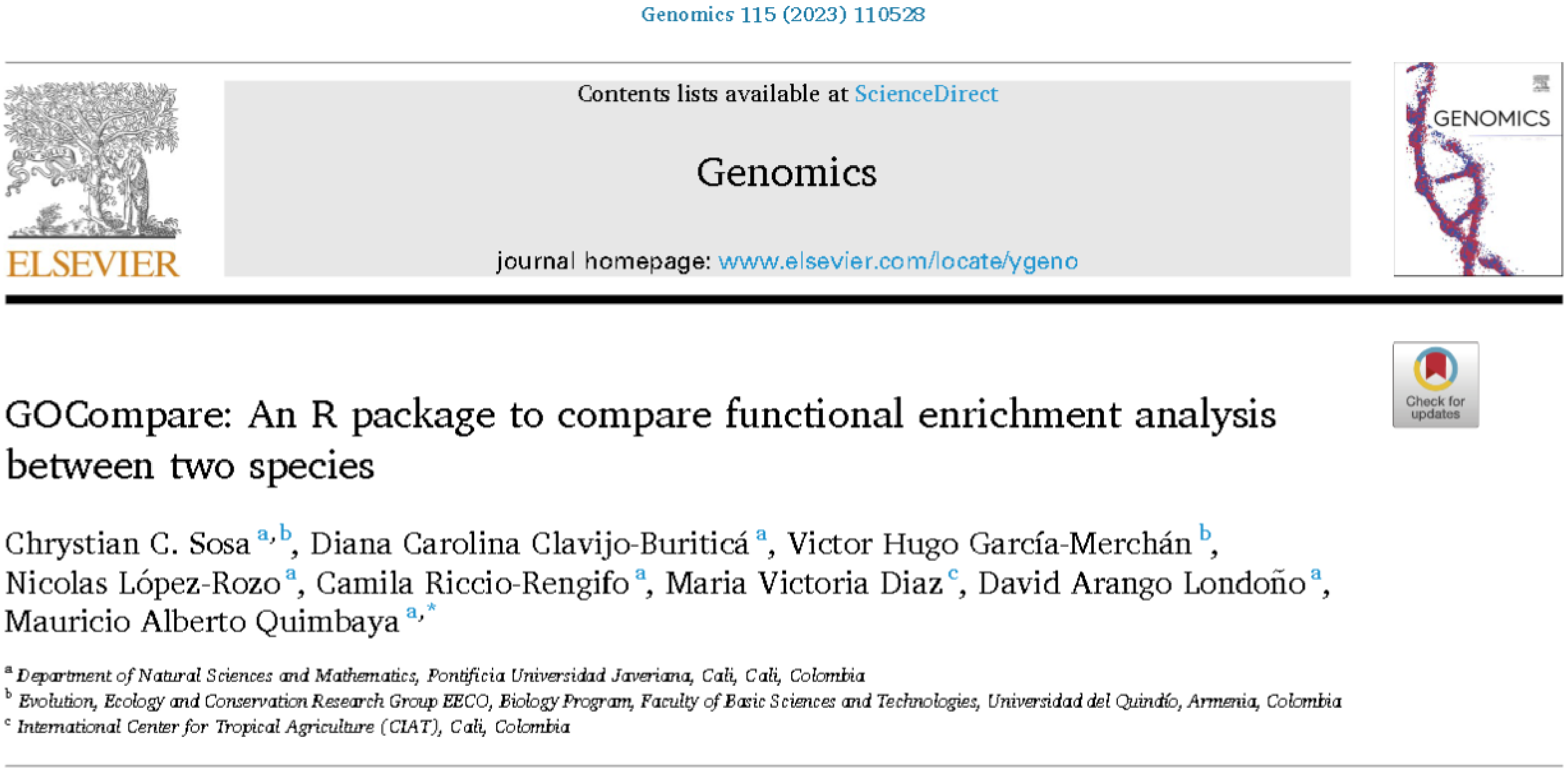


STAGE I: Functional perspective

STAGE II: Metabolic response based on genome-scale network reconstruction

STAGE III: Potential regulatory elements

STAGE I: Functional perspective
(Graph theory for comparative genomics)



Genomics 115 (2023) 110528

Contents lists available at ScienceDirect

Genomics

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

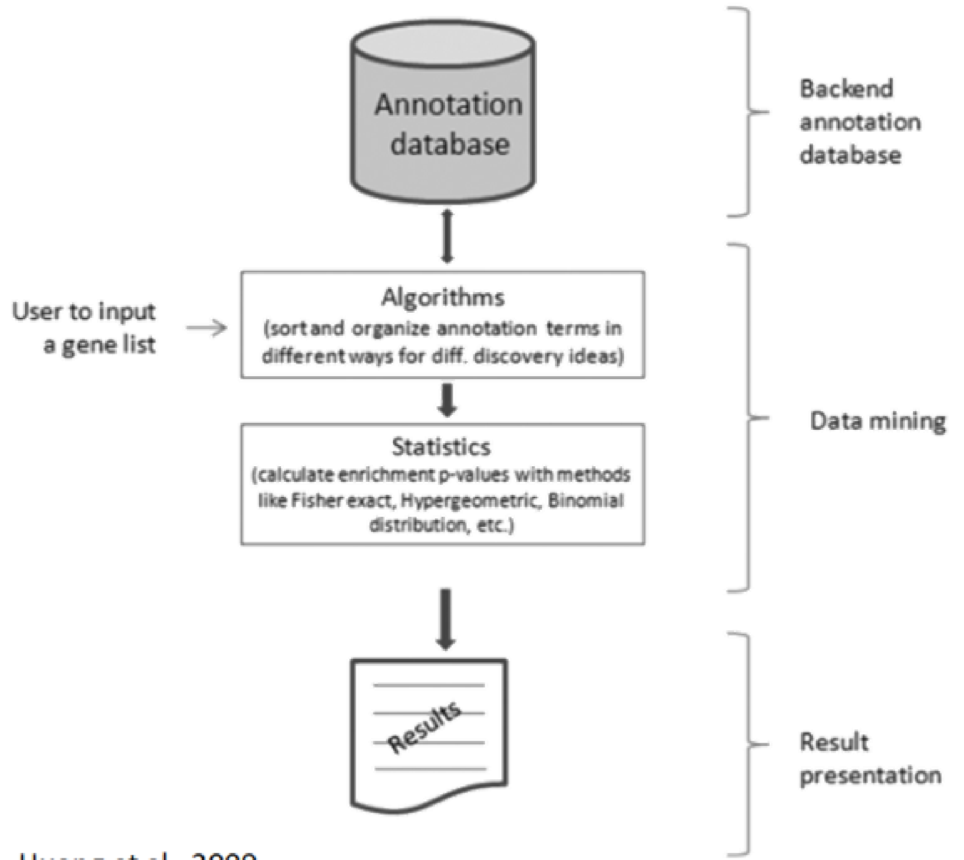
GOCompare: An R package to compare functional enrichment analysis between two species

Chrystian C. Sosa[a,b], Diana Carolina Clavijo-Buriticá[a], Victor Hugo García-Merchán[b], Nicolas López-Rozo[a], Camila Riccio-Rengifo[a], Maria Victoria Diaz[c], David Arango Londoño[a], Mauricio Alberto Quimbaya[a,*]

[a] Department of Natural Sciences and Mathematics, Pontificia Universidad Javeriana, Cali, Cali, Colombia
[b] Evolution, Ecology and Conservation Research Group EECO, Biology Program, Faculty of Basic Sciences and Technologies, Universidad del Quindío, Armenia, Colombia
[c] International Center for Tropical Agriculture (CIAT), Cali, Colombia

Functional enrichment analysis

How an enrichment analysis works?



Annotation database

User to input a gene list

Algorithms (sort and organize annotation terms in different ways for diff. discovery ideas)

Statistics (calculate enrichment p-values with methods like Fisher exact, Hypergeometric, Binomial distribution, etc.)

Results

Backend annotation database

Data mining

Result presentation

Huang et al., 2009

Multiple gene lists enrichment comparison is descriptive!

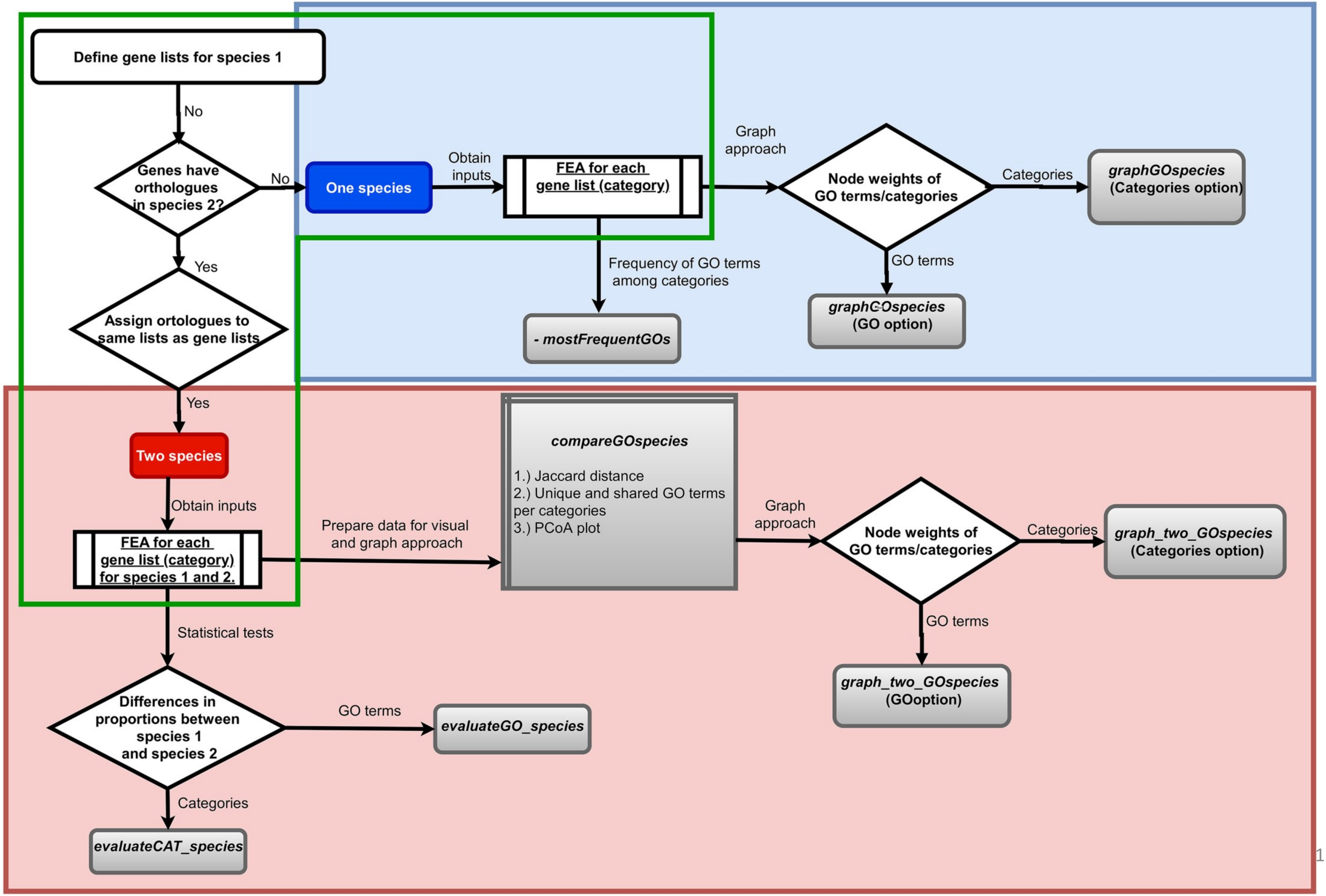
How multiple gene lists are compared?

The idea

- Use undirected weighted graphs to identify statistically relevant GO terms for categories (gene lists) and species

https://github.com/ccsosa/GOCompare

- **Advantages**
    1. Avoid the use of complex algorithms (memory and DAG dependent)
    2. Visualization of GO terms interactions
    3. Detect differences in genome annotations
    4. Numerical outcomes easily readable in Cytoscape
    5. Statistical testing for GO terms and gene lists composition
    6. Reproducibility



Define gene lists for species 1

No

Genes have orthologues in species 2?

No

One species

Obtain inputs

FEA for each gene list (category)

Graph approach

Node weights of GO terms/categories

Categories

***graphGOspecies* (Categories option)**

GO terms

***graphGOspecies* (GO option)**

Frequency of GO terms among categories

- mostFrequentGOs

Yes

Assign ortologues to same lists as gene lists

Yes

Two species

Obtain inputs

FEA for each gene list (category) for species 1 and 2.

Prepare data for visual and graph approach

compareGOspecies

1.) Jaccard distance
2.) Unique and shared GO terms per categories
3.) PCoA plot

Graph approach

Node weights of GO terms/categories

Categories

***graph_two_GOspecies* (Categories option)**

GO terms

***graph_two_GOspecies* (GOoption)**

Statistical tests

Differences in proportions between species 1 and species 2

GO terms

evaluateGO_species

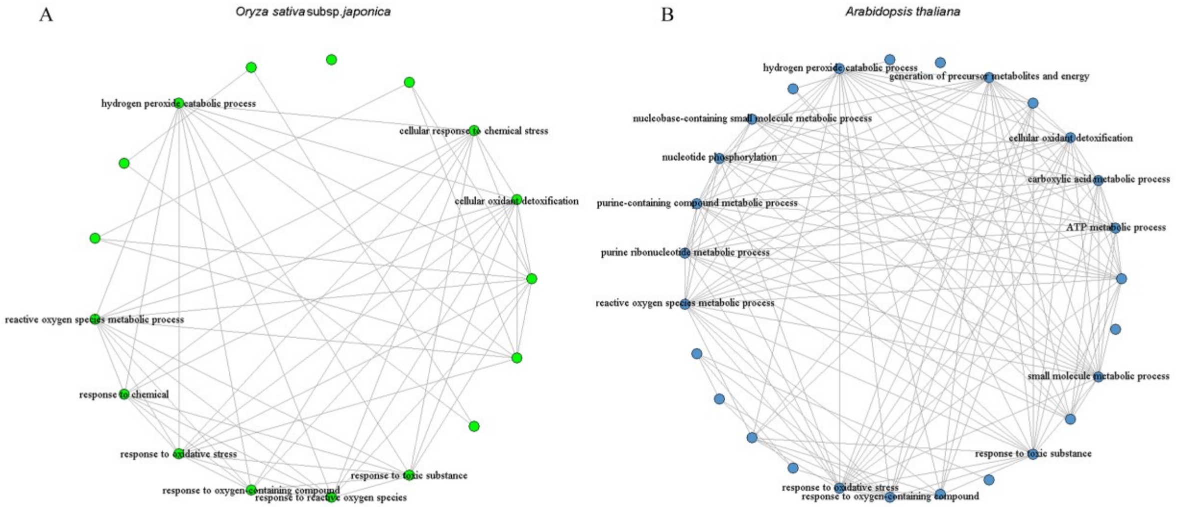
Categories

evaluateCAT_species

- Categories (gene lists) similarity

$$w(e) = \frac{|BPu \cap BPv|}{|BP|}$$

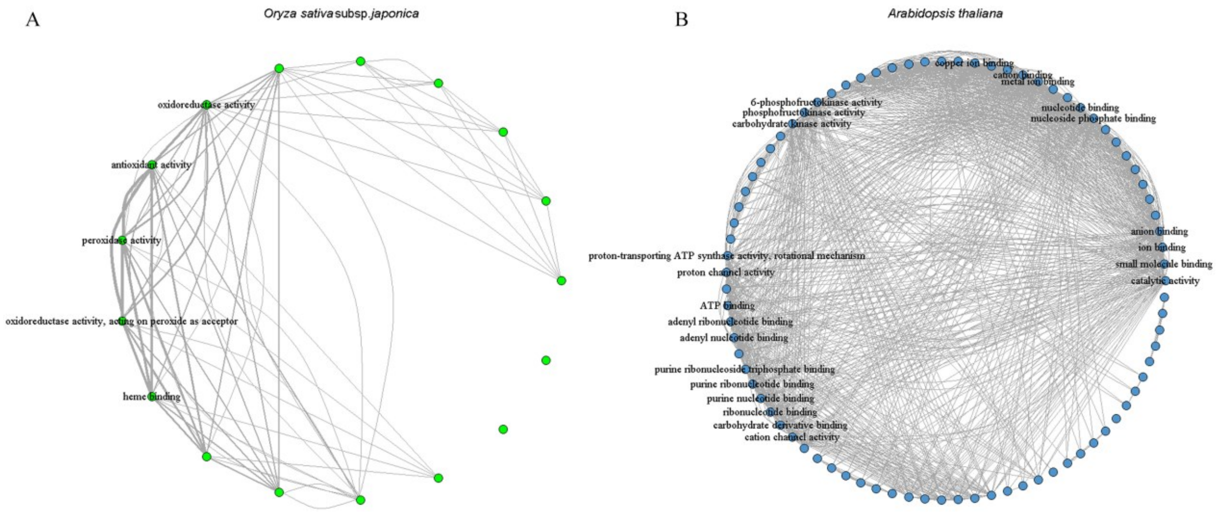
$$K_w(u) = \sum_{v \in V} w(u, v)$$



- GO Terms frequency

$$w'(e') = \frac{|Cu' \cap Cv'|}{|BP|}$$

$$K'_w(u') = \sum_{v' \in V'} w'(u', v')$$



- **Comparison between species**

$$\widehat{K}_w(u) = \sum_{v \in V_1} w(u,v) + \sum_{x \in V_2} w(u,x)$$

- **High values within categories with highly frequent GO terms co-occurring**

- **A Combined approach for categories:**
  - **Shared:**

$$s(e) = \frac{|N_1 \cap N_2|}{|N_1 \cup N_2|}$$

$$K_s(u) = \sum_{v \in (V_1 \cup V_2)} s(u,v)$$

  - **Combined: (Shared + Categories weight)**

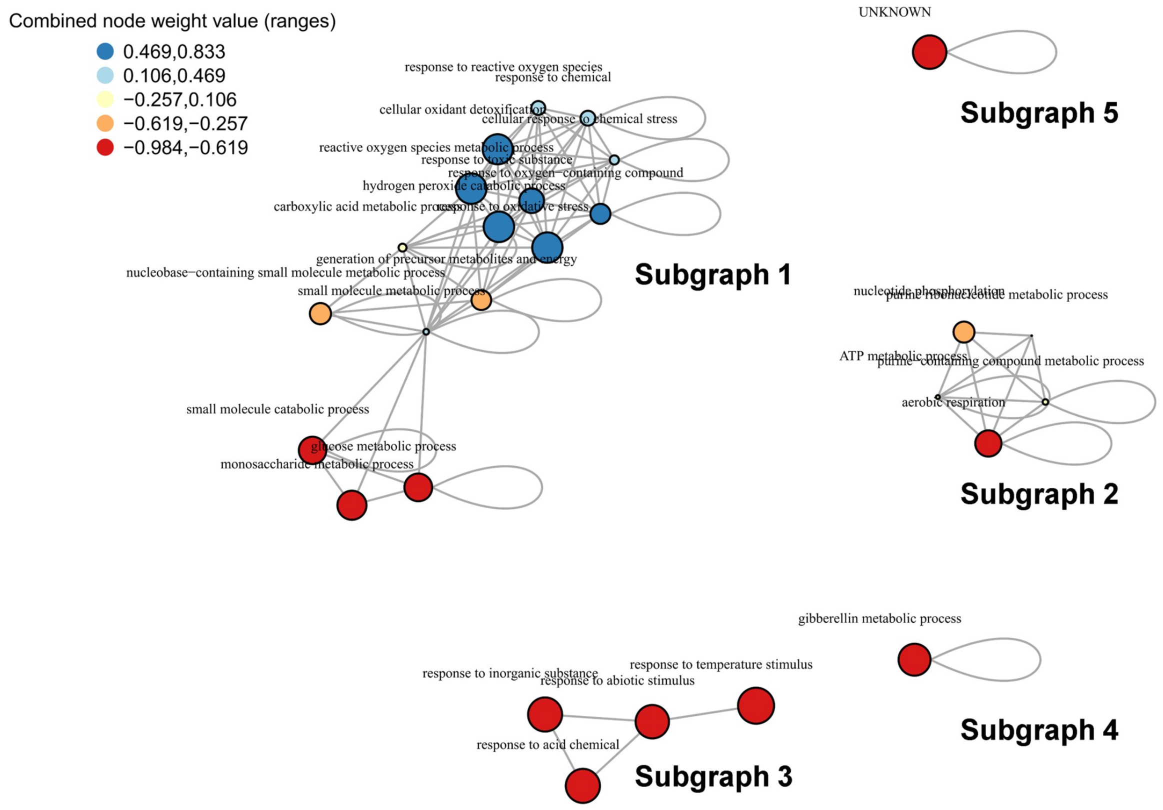
$$min\_max(y) = \frac{y - min(y)}{max(y) - min(y)}$$

$$K_c(u) = min\_max(\widehat{K}_w(u)) + min\_max(K_s(u)) - 1$$

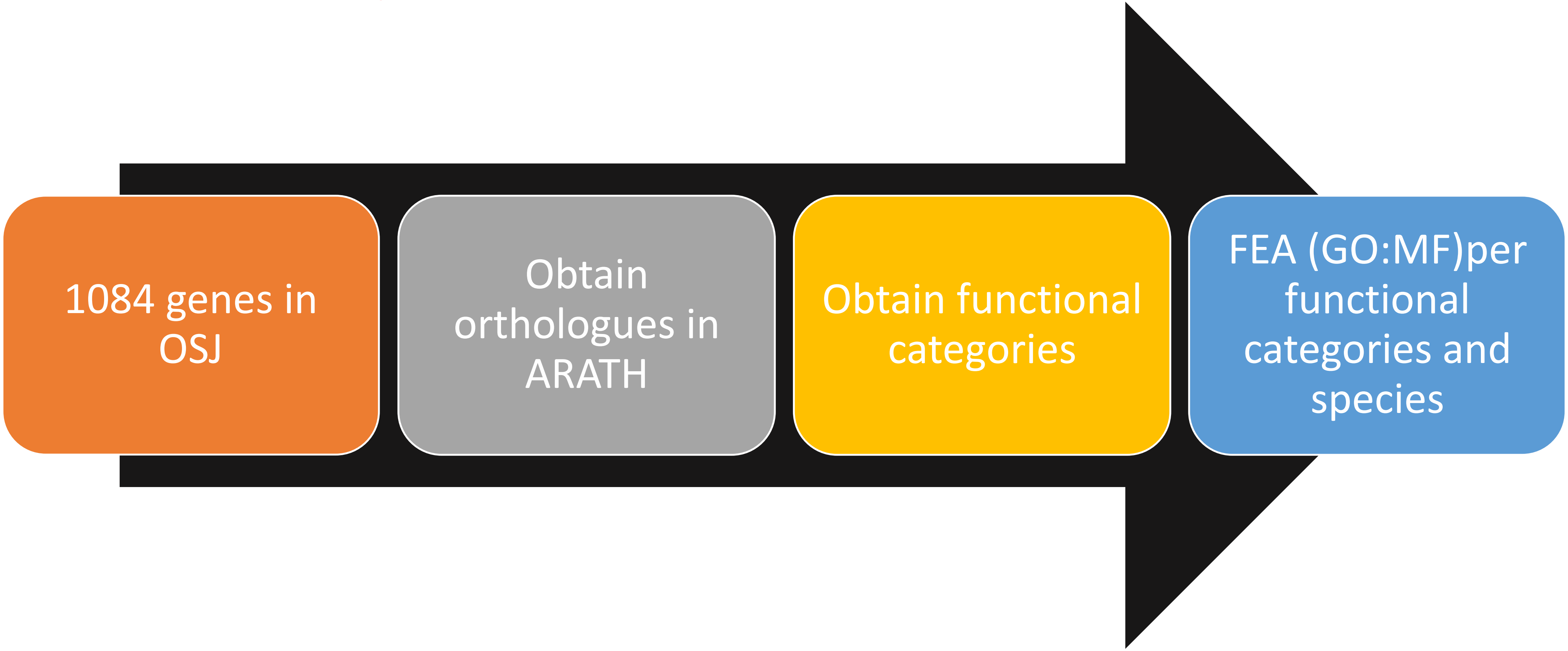
- **An index with values from -1 (unique for a species) to 1 (Shared between species)**

Towards a functional perspective comparative genomics

Functional category	GO weight	Shared GO weight	Combined weight
Reactive oxygen species metabolic process	0.636	7.989	0.8303
Response to reactive oxygen species	0.36	7.6	0.3701
Response to temperature stimulus	0.012	0	−1



¿How to detect divergences in aluminum tolerance mechanisms between *Oryza sativa* subsp. japonica and *Arabidopsis thaliana*?



1084 genes in OSJ

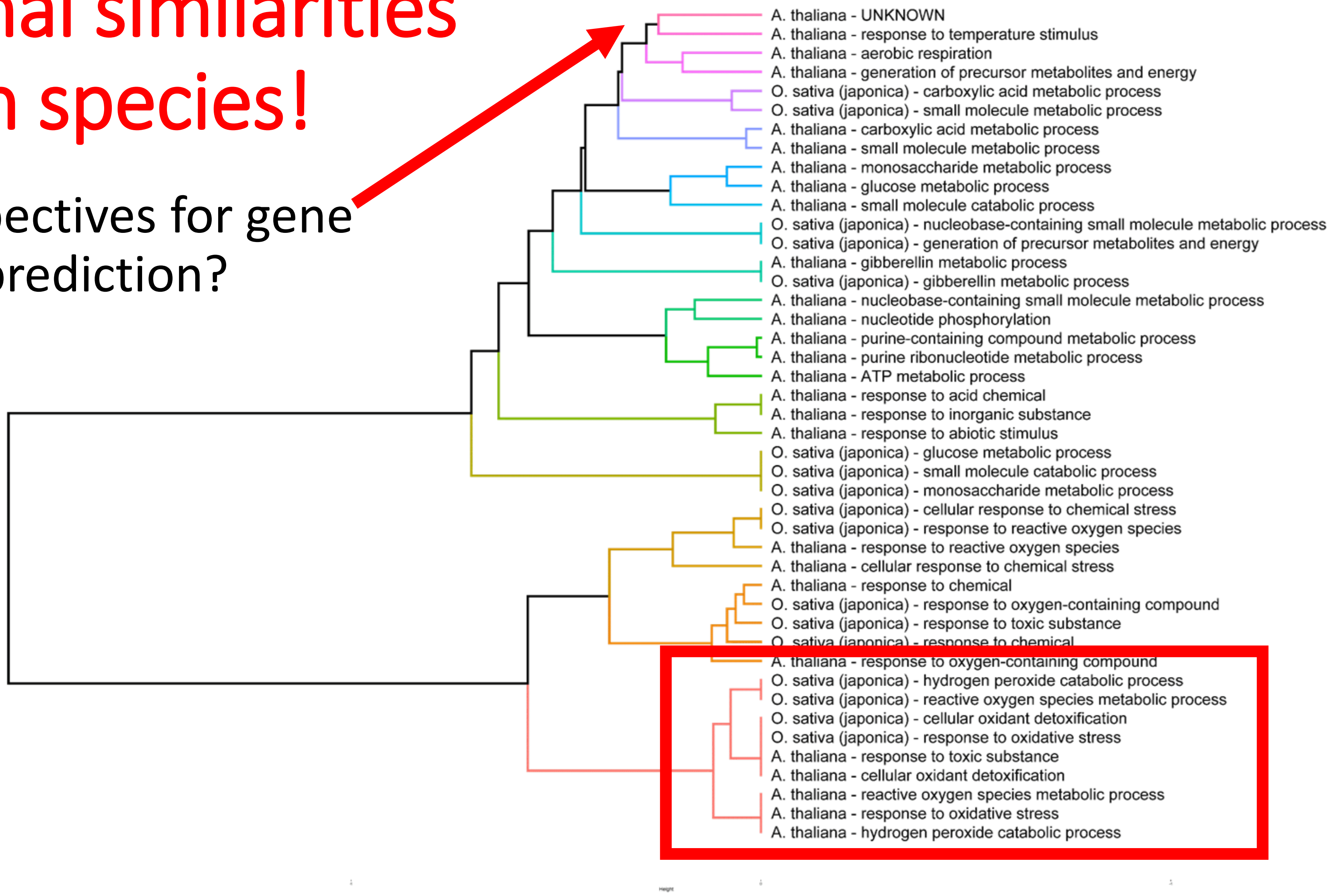
Obtain orthologues in ARATH

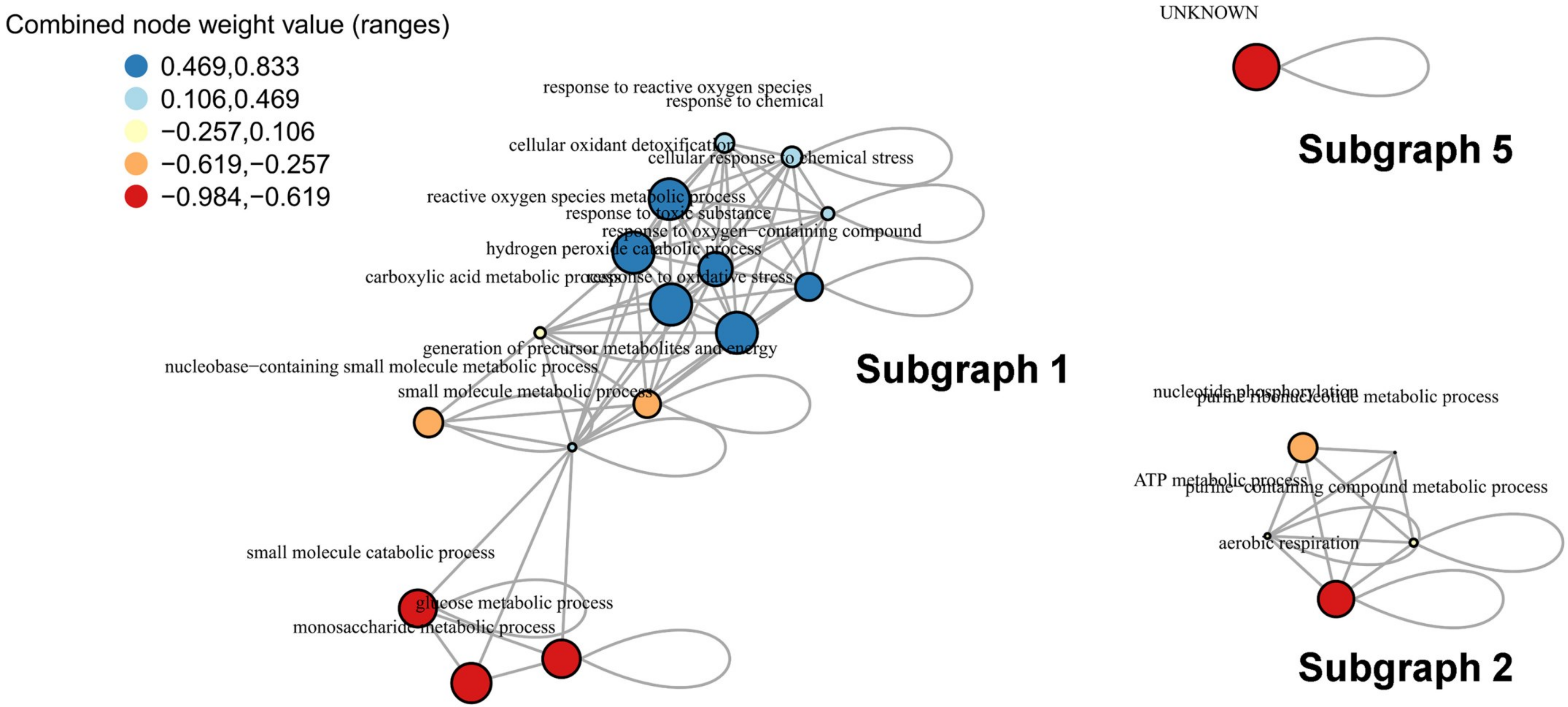
Obtain functional categories

FEA (GO:MF)per functional categories and species

Functional similarities between species!

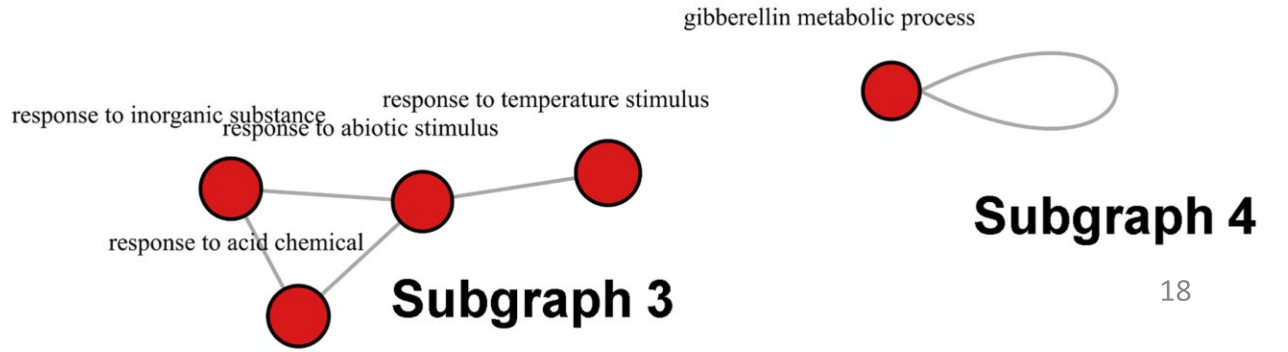
Future perspectives for gene annotation prediction?



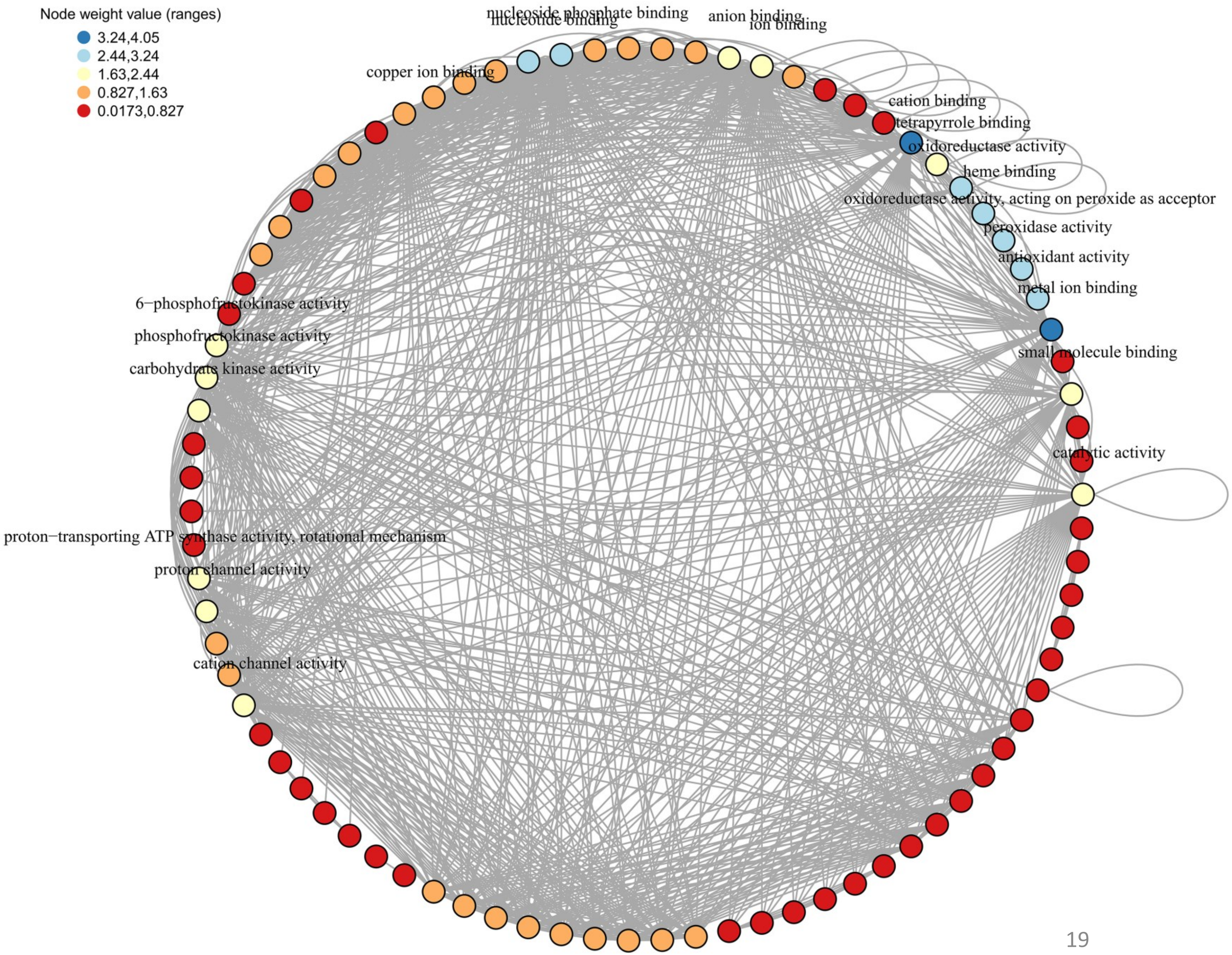


Categories similarity

Conserved ROS mechanisms between OSJ and ARA!



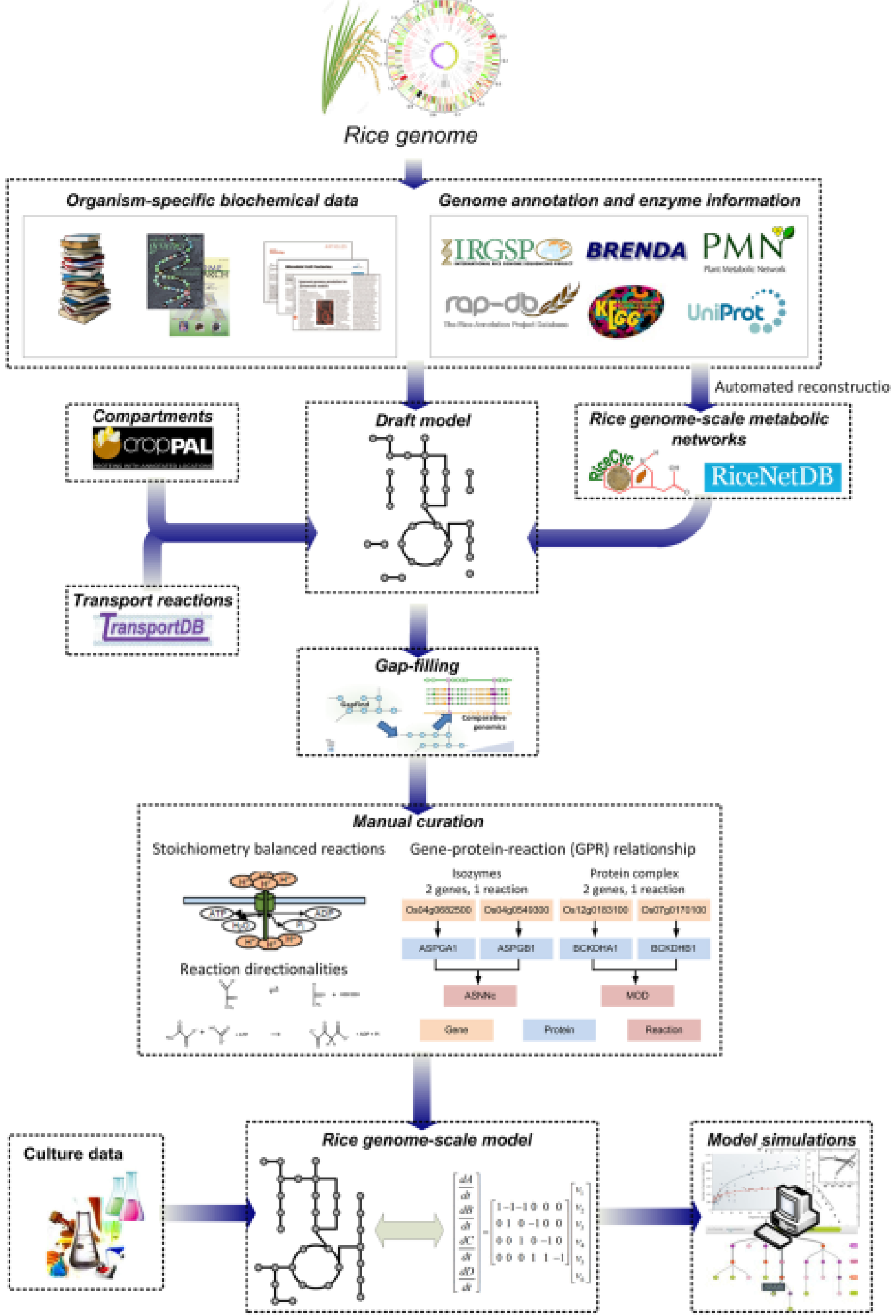
Most frequent GO molecular functions



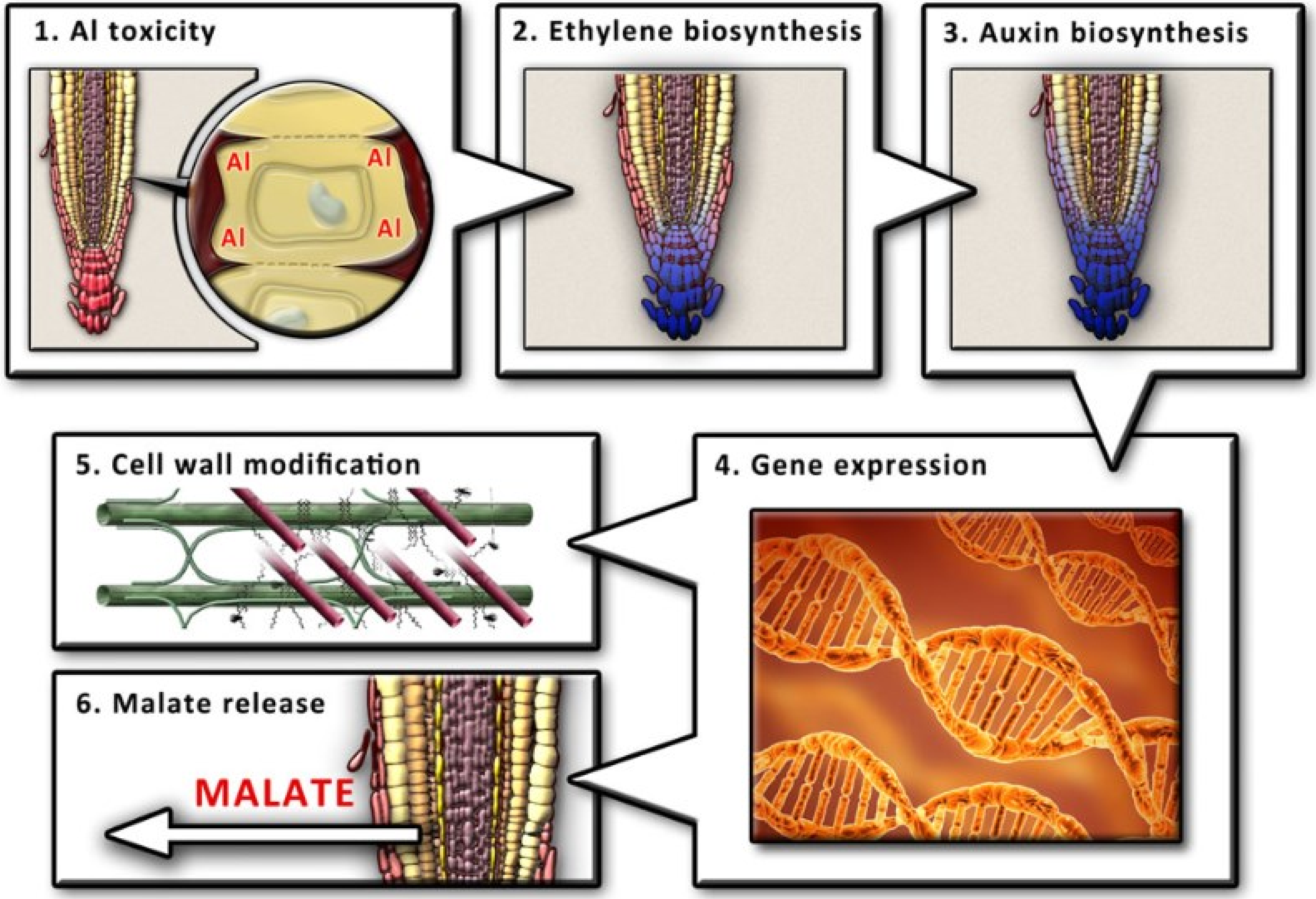
Oxidoreduction process
Binding processes

STAGE II: Metabolic response based on genome-scale network reconstruction

- Manual curation of the iOS2164 model
- 228 new reactions detected from 1084 genes and aluminum:
  - 41 reactions associated with aluminum
  - 187 reactions associated with genes that were associated with aluminum studies in rice.

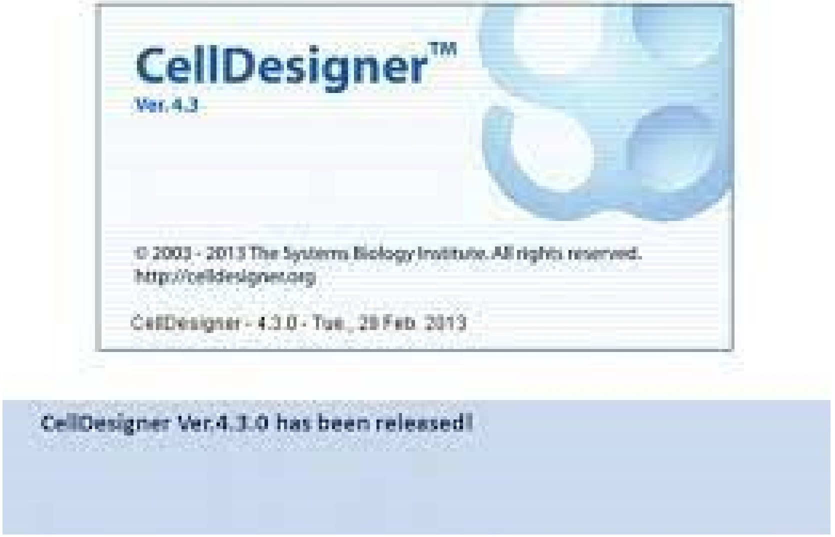


Lakshmanan et al., (2015)



ONGOING: 74 candidate genes

- **39 high priority**
- **16 possible targets**


CellDesigner™
Ver. 4.3

© 2003 - 2013 The Systems Biology Institute. All rights reserved.
http://celldesigner.org

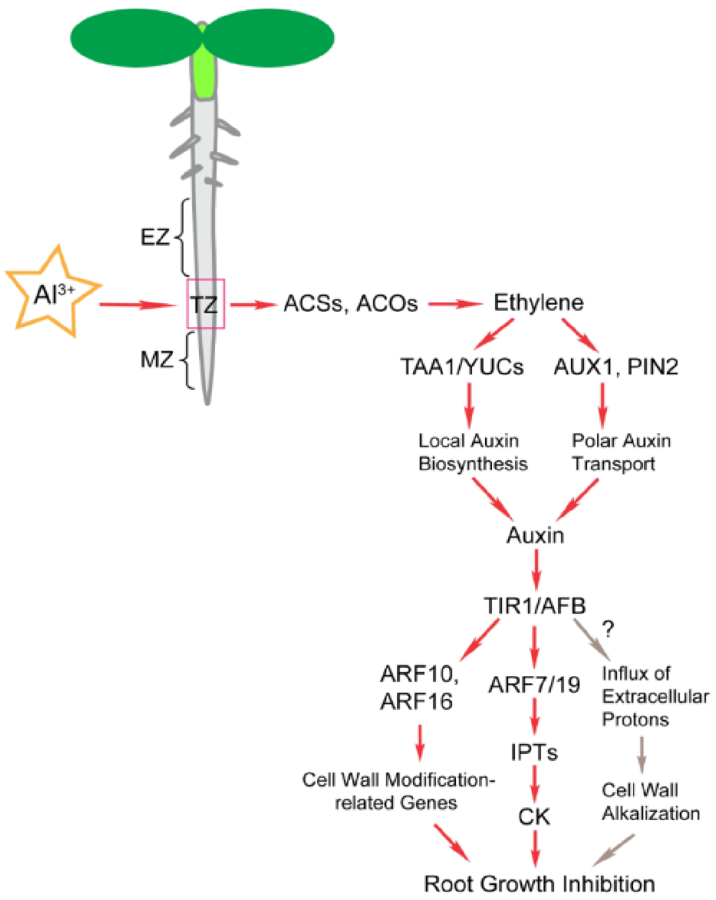
CellDesigner - 4.3.0 - Tue., 28 Feb. 2013

CellDesigner Ver.4.3.0 has been released!

How aluminum affects the root growth response? → Auxins and ethylene biosynthesis

Kopittke, (2016)

The Al in rice root model vs the root regulatory mechanisms in rice



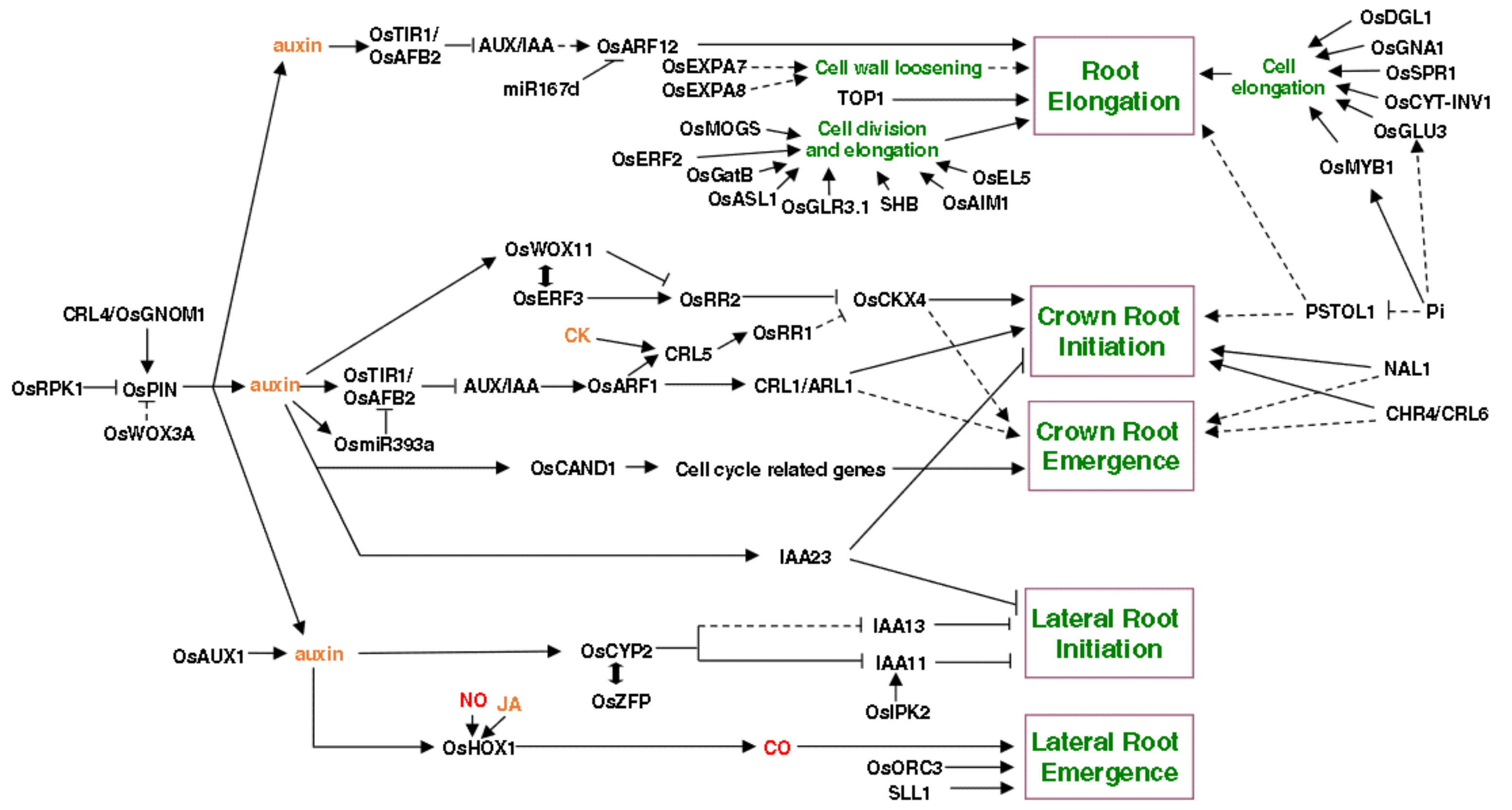
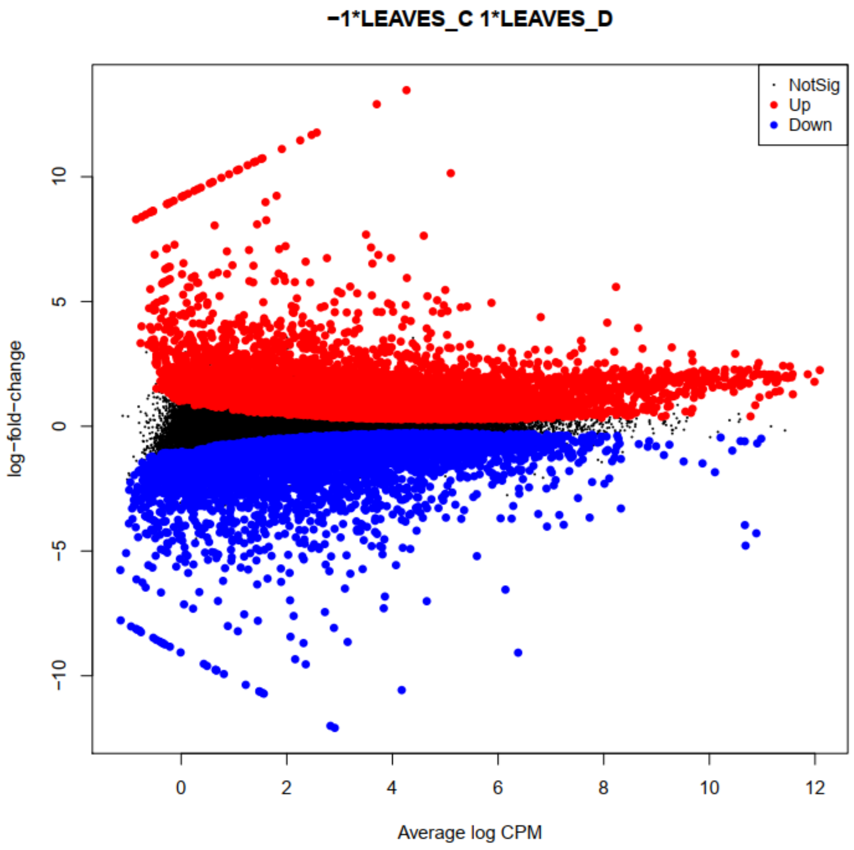
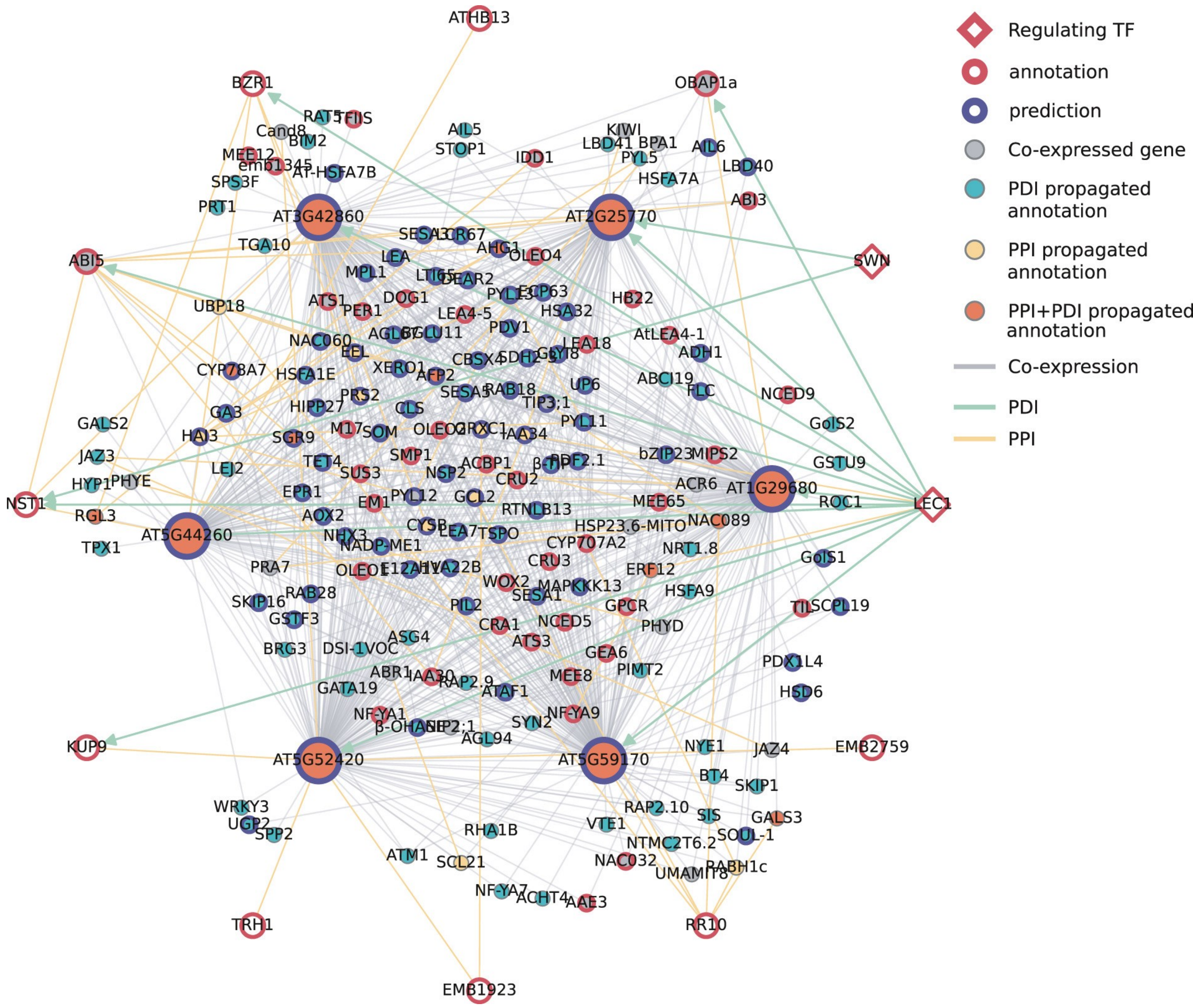


Fig. 1 The molecular regulatory mechanisms of root development in rice. Arrows represent positive regulatory actions. Lines ending in a flat head indicate a negative regulatory action. Dashed lines represent interactions that have not been experimentally confirmed. Double-headed arrows indicate that two proteins interact. Text color code: genes or protein, black; hormones, yellow; signals, red; biological processes, green

Liu et al, (2022); Meng et al., (2019)

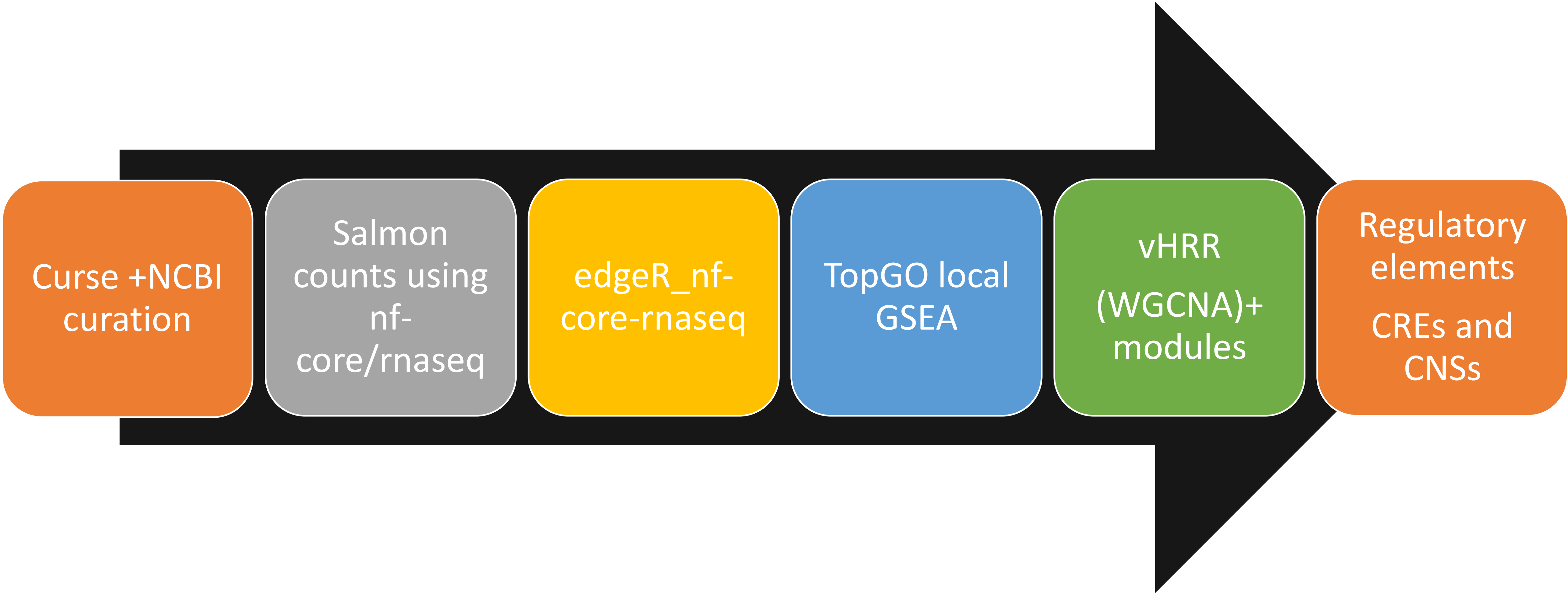
STAGE III: Potential regulatory elements





Depuydt &Vandepoele (2020)

The workflow



Curse +NCBI curation

Salmon counts using nf-core/rnaseq

edgeR_nf-core-rnaseq

TopGO local GSEA

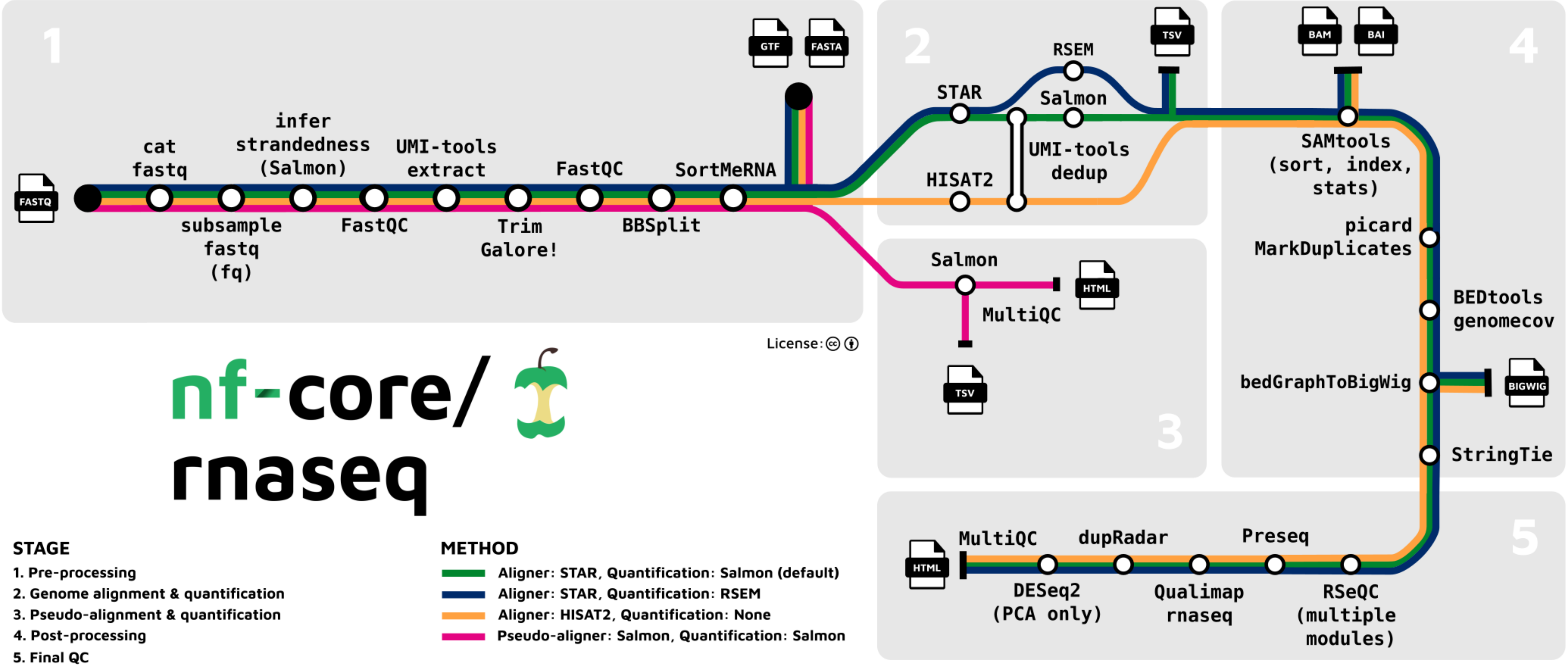
vHRR

(WGCNA)+ modules

Regulatory elements

CREs and CNSs

Original nf-core/rnaseq workflow

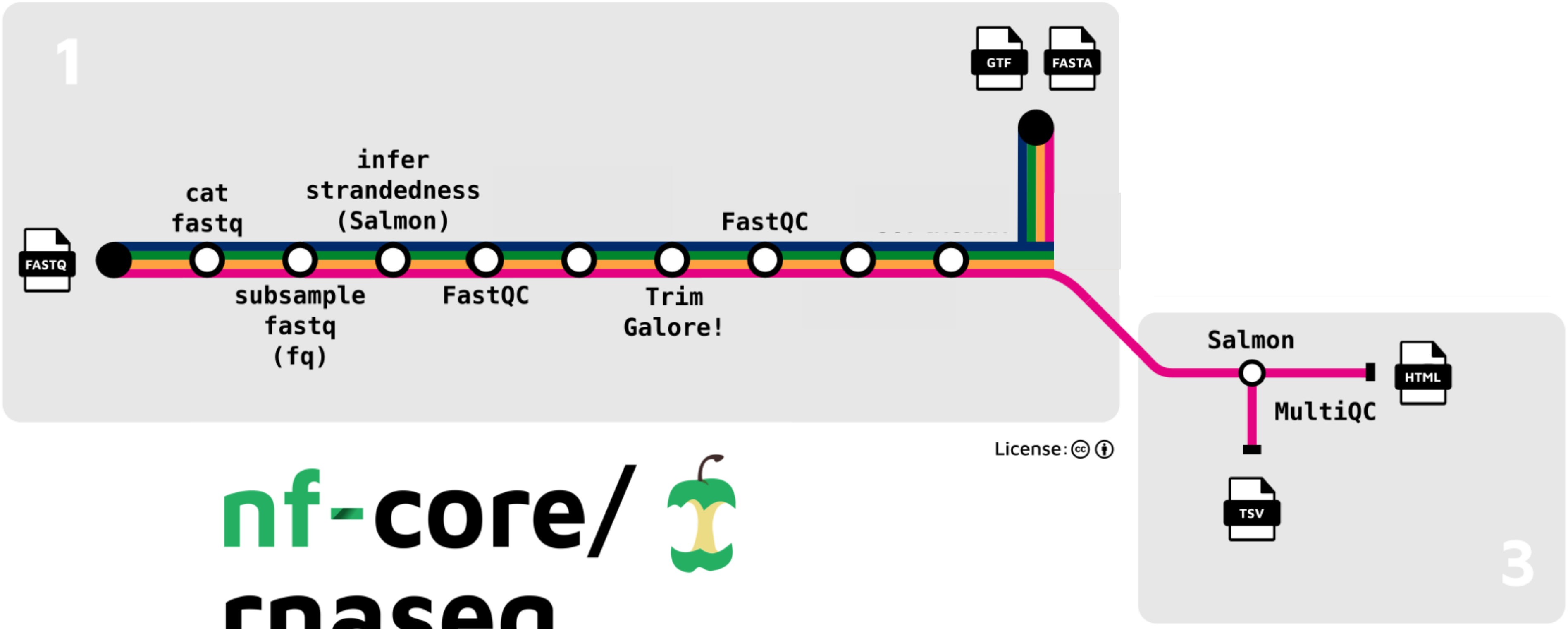


STAGE
1. Pre-processing
2. Genome alignment & quantification
3. Pseudo-alignment & quantification
4. Post-processing
5. Final QC

METHOD
- 🟩 Aligner: STAR, Quantification: Salmon (default)
- 🟦 Aligner: STAR, Quantification: RSEM
- 🟧 Aligner: HISAT2, Quantification: None
- 🟪 Pseudo-aligner: Salmon, Quantification: Salmon

https://nf-co.re/rnaseq

Bulk RNASeq experiments



nf-core/ rnaseq

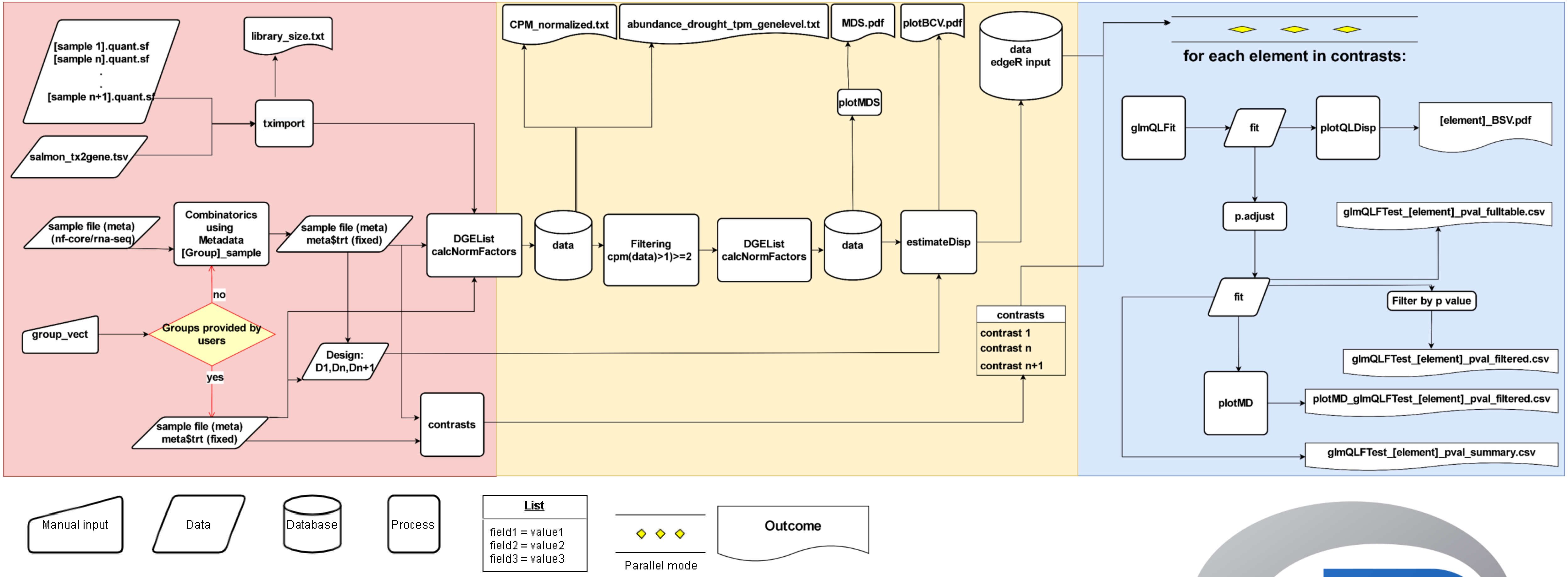
DEG step

STAGE	METHOD
1. Pre-processing	
3. Pseudo-alignment & quantification	
	🟥 Pseudo-aligner: Salmon, Quantification: Salmon

- **The exploratory step requires aligners use**

- **A differential expressed genes calculation step is missing!**

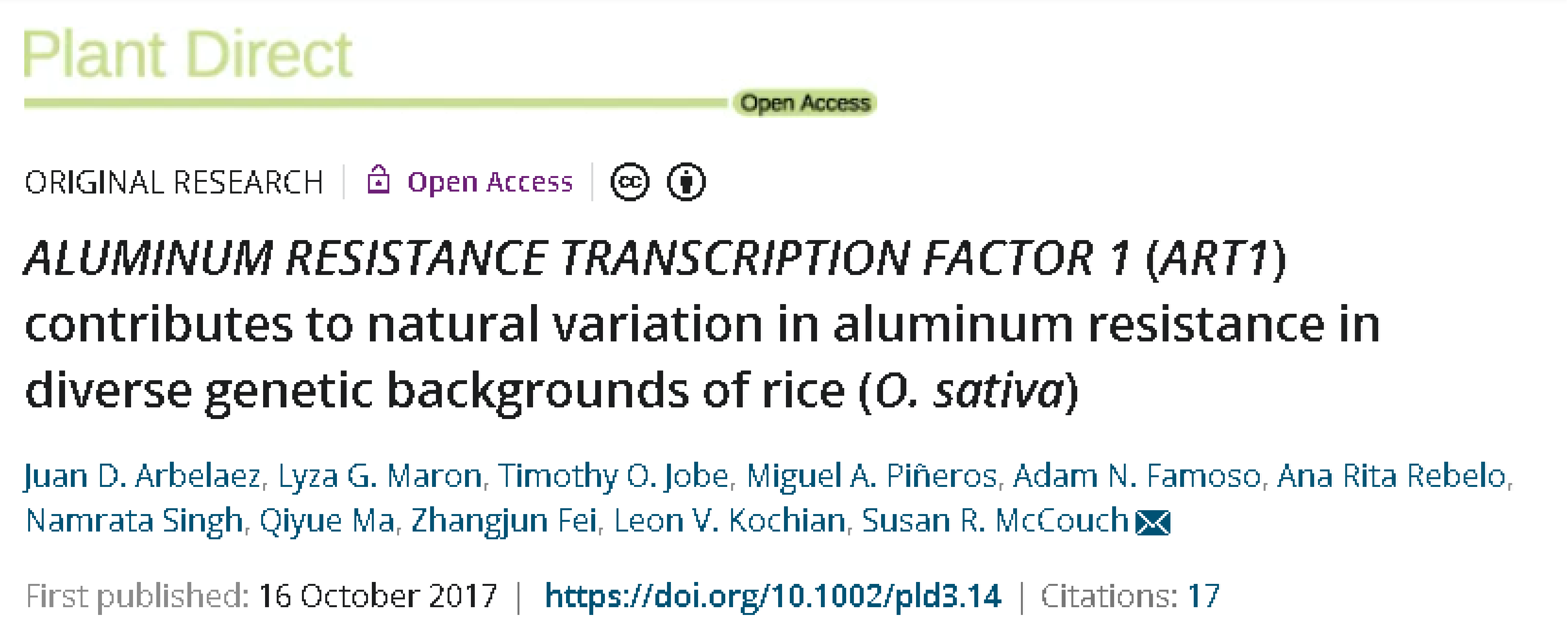
https://nf-co.re/rnaseq



- Red: Data preprocessing
- Yellow: EdgeR preprocessing
- Blue EdgeR paralleling processing

https://github.com/ccsosa/edgeR_nf-core-rnaseq

Studies:



Plant Direct

ORIGINAL RESEARCH | Open Access

ALUMINUM RESISTANCE TRANSCRIPTION FACTOR 1 (ART1) contributes to natural variation in aluminum resistance in diverse genetic backgrounds of rice (*O. sativa*)

Juan D. Arbelaez, Lyza G. Maron, Timothy O. Jobe, Miguel A. Piñeros, Adam N. Famoso, Ana Rita Rebelo, Namrata Singh, Qiyue Ma, Zhangjun Fei, Leon V. Kochian, Susan R. McCouch

First published: 16 October 2017 | https://doi.org/10.1002/pld3.14 | Citations: 17

Characterization of aluminum stress tolerance mechanisms through transcriptional analysis of the contrasting rice species *Oryza sativa* and *Oryza glumaepatula*

IN PREPARATION

- **28 samples in two studies**

Variety	Aluminum	Control
Azucena	7 (3 P1)	7 (3 P1)
AZU[IR64121] (near-isogenic line)	5	3
BGI	3	3

Thanks…

ómicas

ccsosaa@javerianacali.edu.co