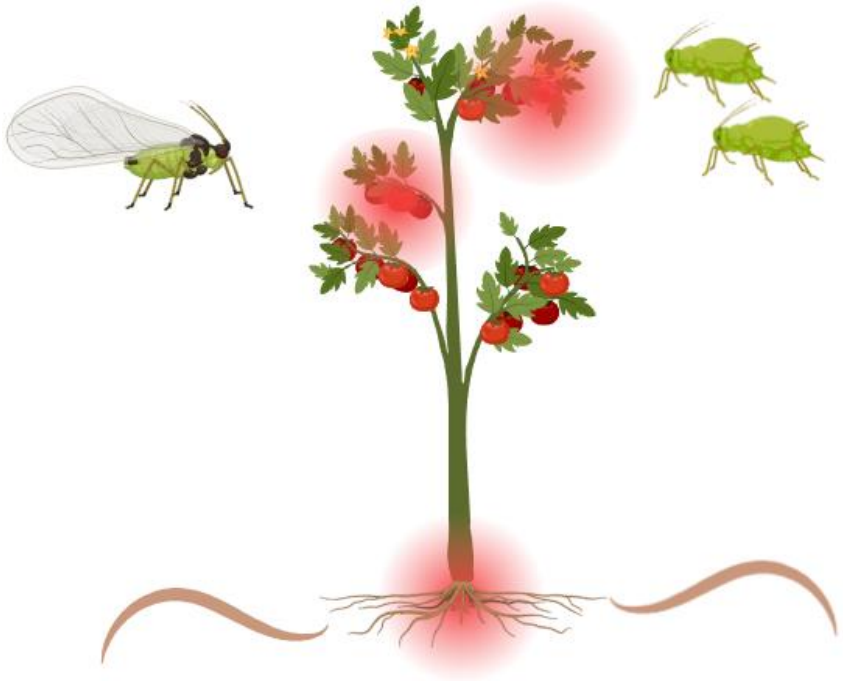
From knowledge graph to topological data analysis: a novel framework to analyze gene regulatory networks for tomato-multi-pathogen interactions

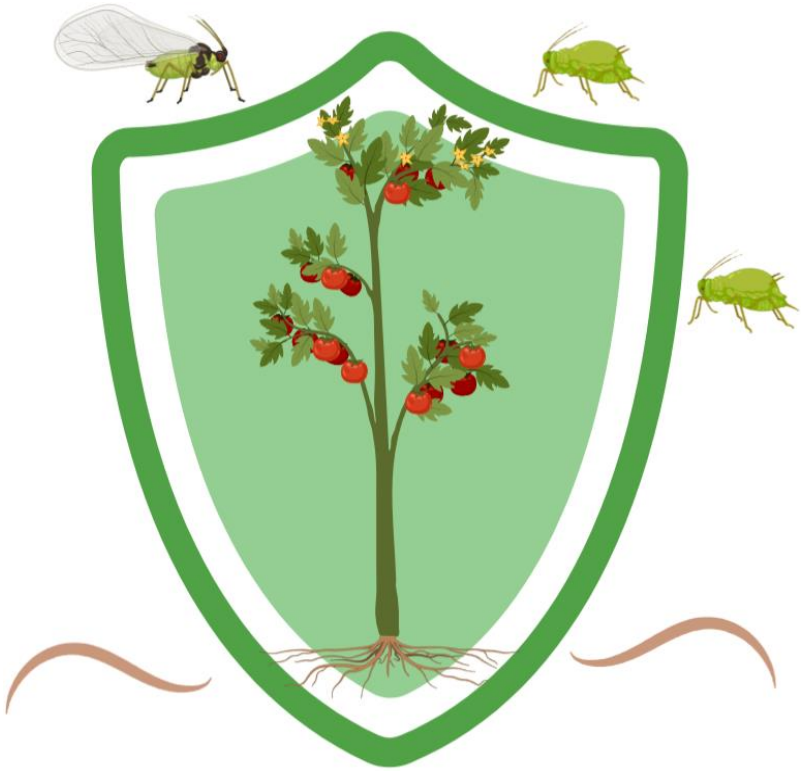
Maxime MULTARI
(SMILE Team - Institut Sophia Agrobiotech)
maxime.multari@inrae.fr

Second year PhD Student supervised by:
Dr. Silvia BOTTINI
Dr. Stéphanie JAUBERT

Plant Defense

- **Plant-pathogen interaction**

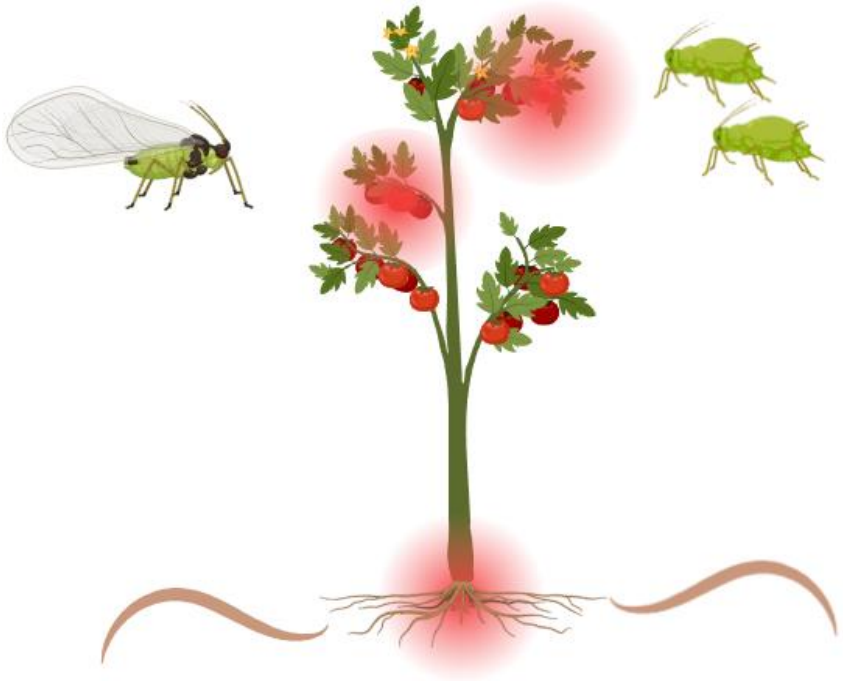




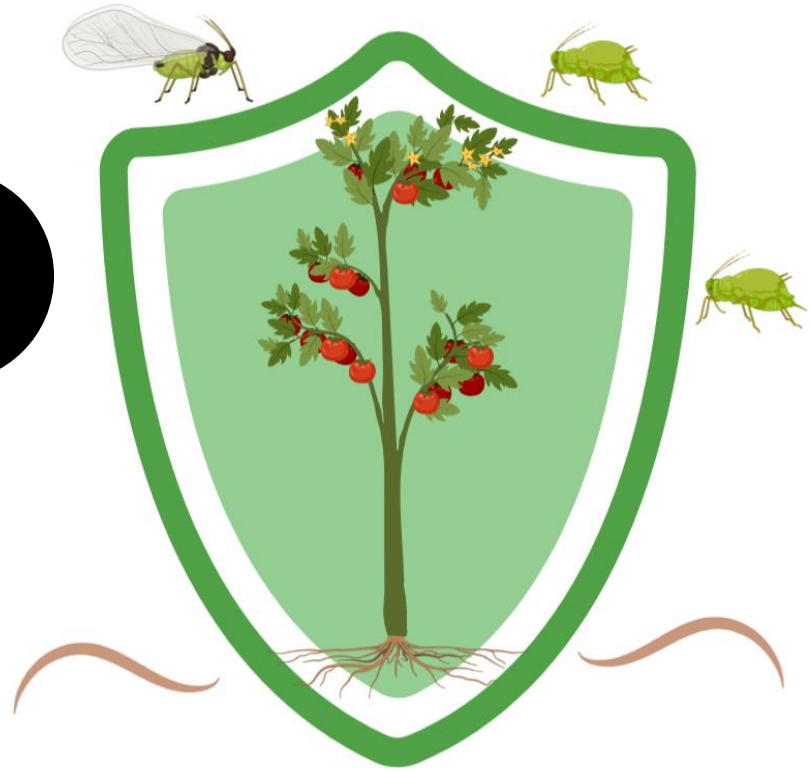
Created with BioRender.com

Plant Defense

- **Plant-pathogen interaction**



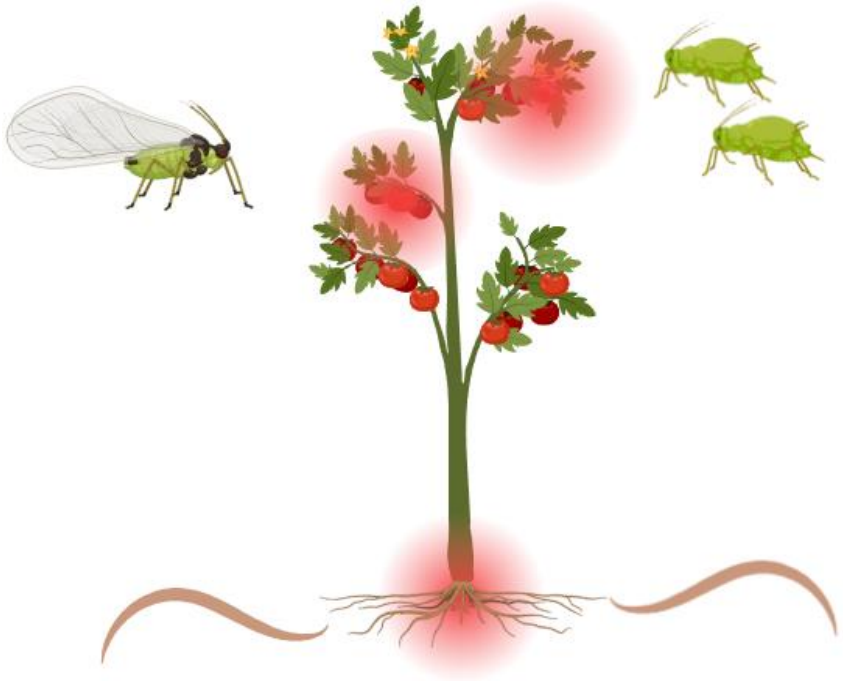
How can plants deal with all those heterogeneous pests' attacks?



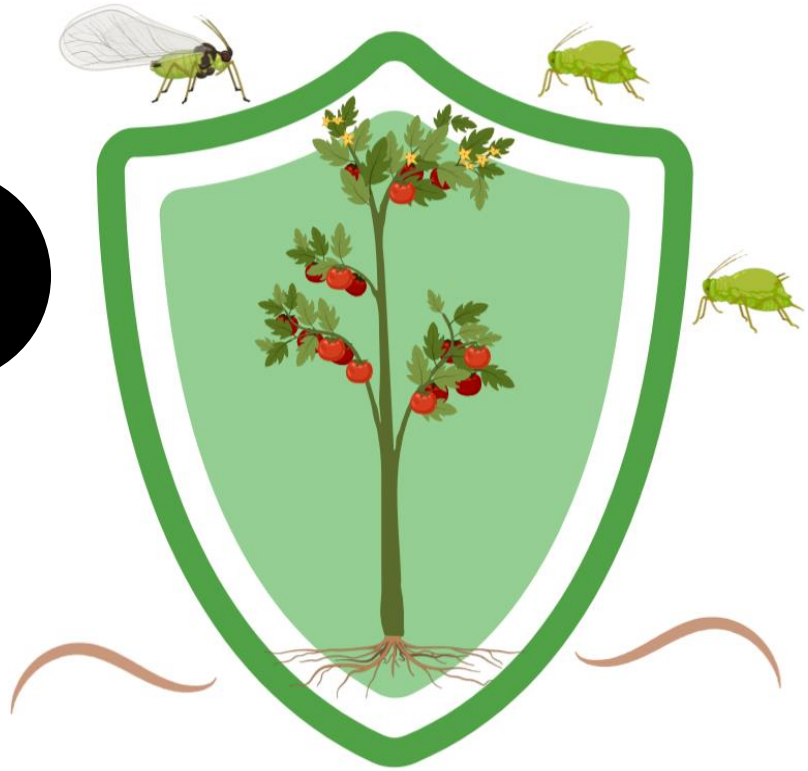
Created with BioRender.com

Plant Defense

- **Plant-pathogen interaction**



How can plants deal with all those heterogeneous pests' attacks?



- **Better understand the plant responses**
  - **<u>Specific and common plant response signatures</u>**

Created with BioRender.com

Tomato as model species

- Solanum lycopercisum

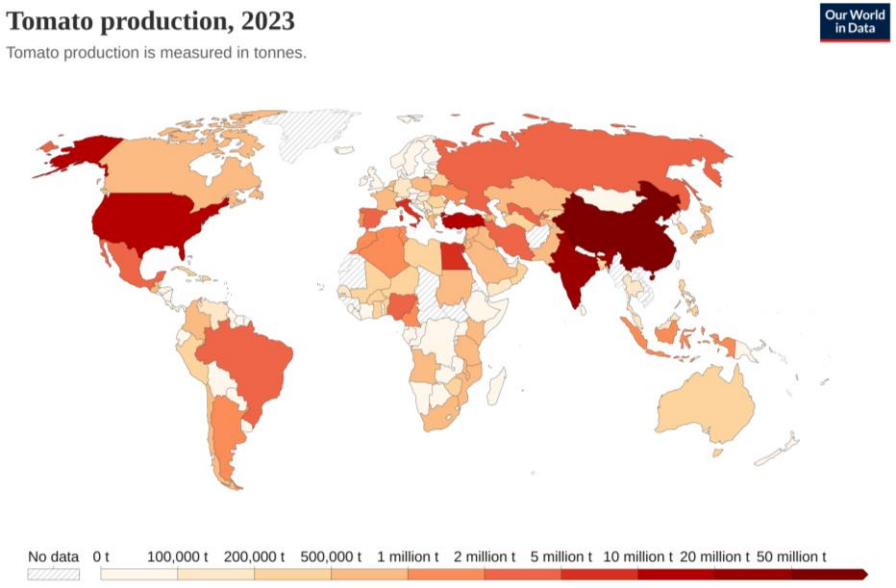


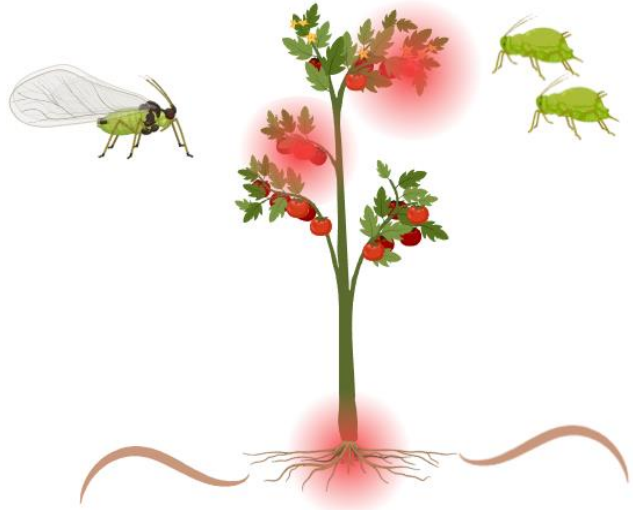
INRAE

- Global agro-economic interest
- More than 200 diseases caused by pathogens
- Enough knowledge spread across databases

Model species for a plant of agro-economic interest

- Employed at ISA



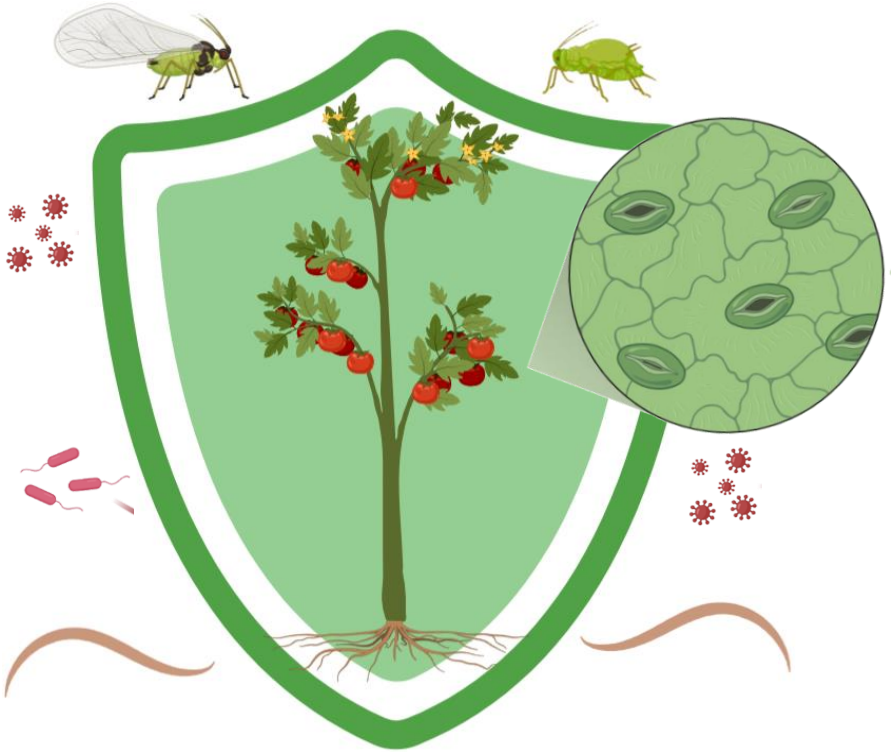


Created with BioRender.com

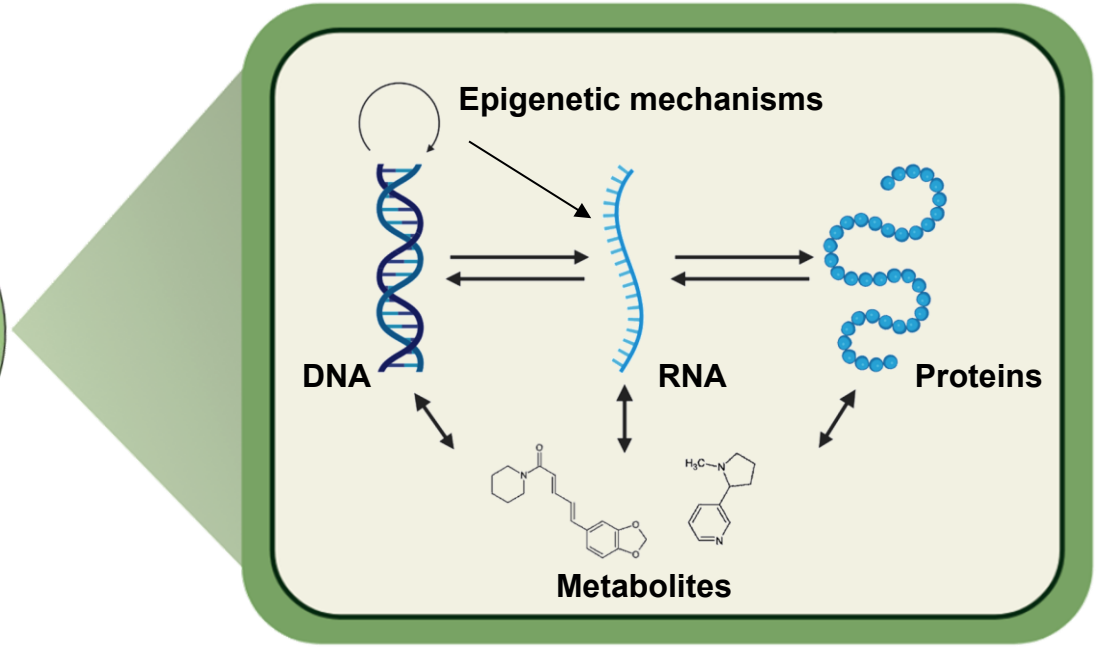
Singh, V.K., Singh, A.K., and Kumar, A. (2017). Disease management of tomato through PGPB: current trends and future perspective. 3 Biotech *7*, 255. https://doi.org/10.1007/s13205-017-0896-1

Omics Data to understand the plant response

- **Plant response to biotic stresses**
- **Focus on *Solanum lycopersicum***



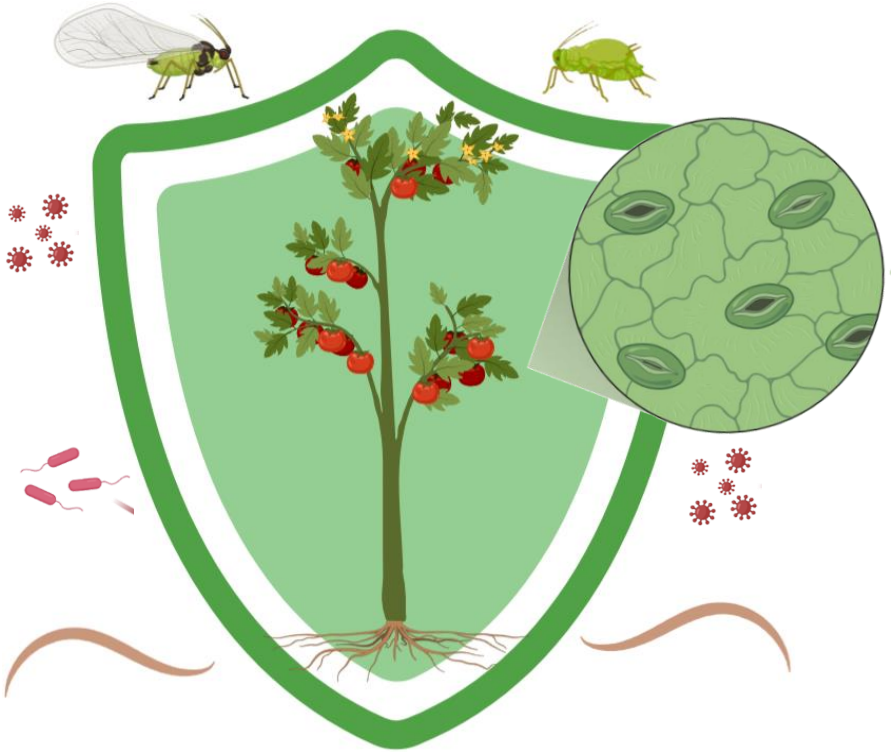
Created with BioRender.com



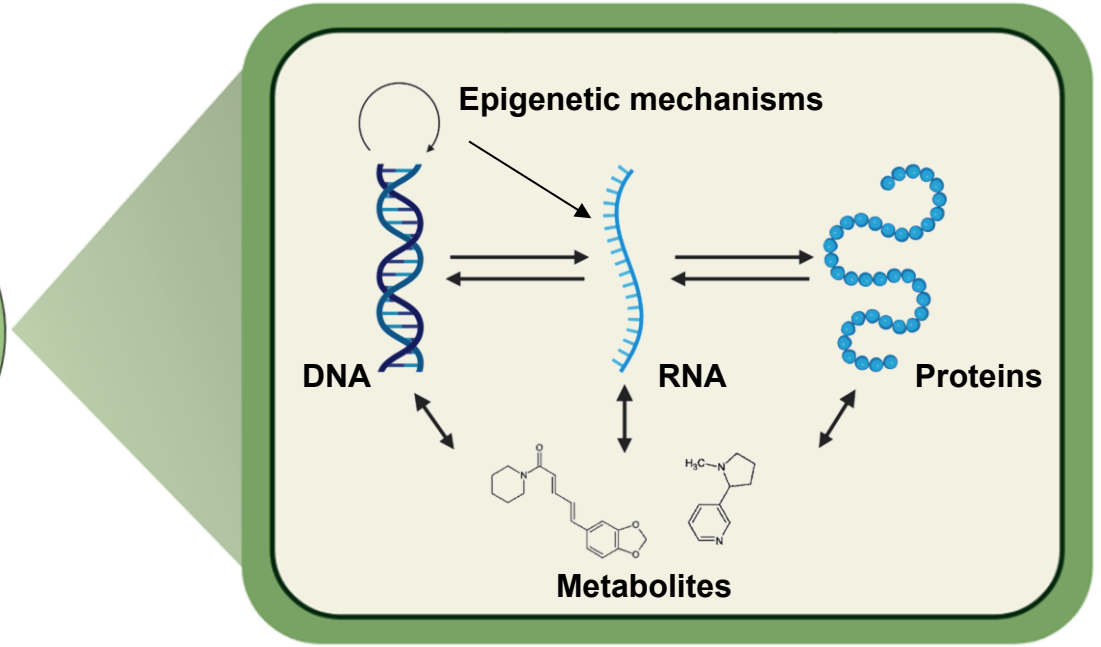
- Usage of Omics technologies: Genomics, Transcriptomics, …

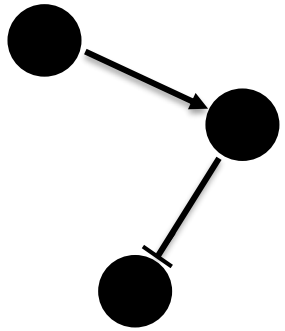
Omics Data to understand the plant response

- **Plant response to biotic stresses**
- **Focus on *Solanum lycopersicum***



Created with BioRender.com



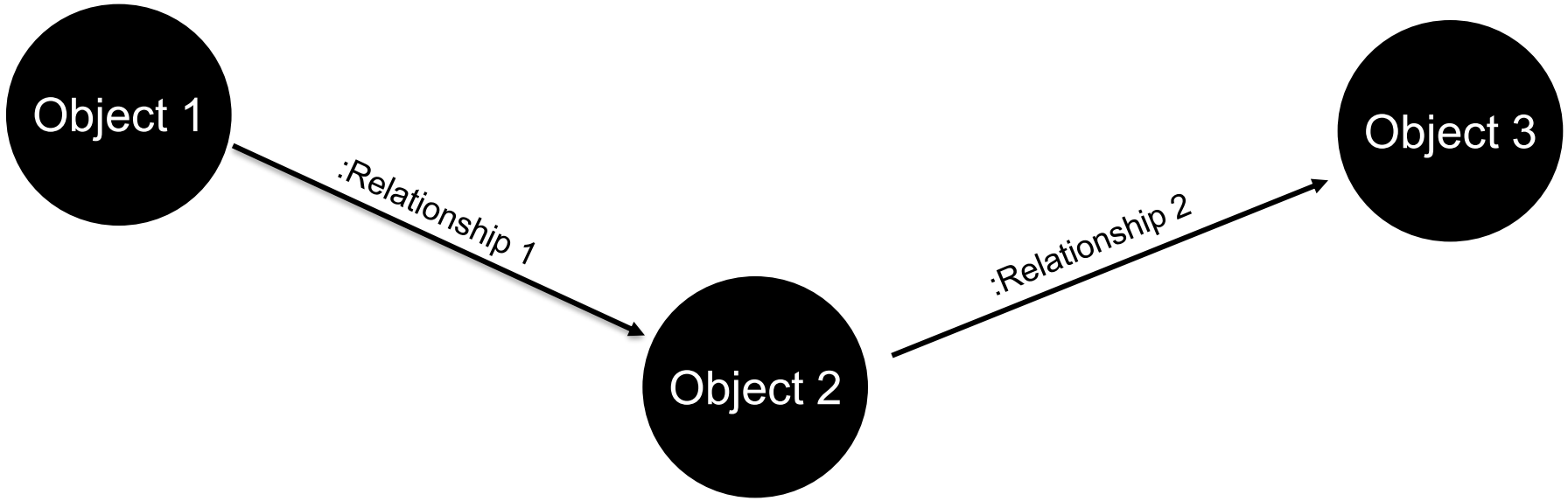


Regulatory mechanisms

- Usage of Omics technologies: Genomics, Transcriptomics, …

Knowledge Graph

- **Representing the knowledge**

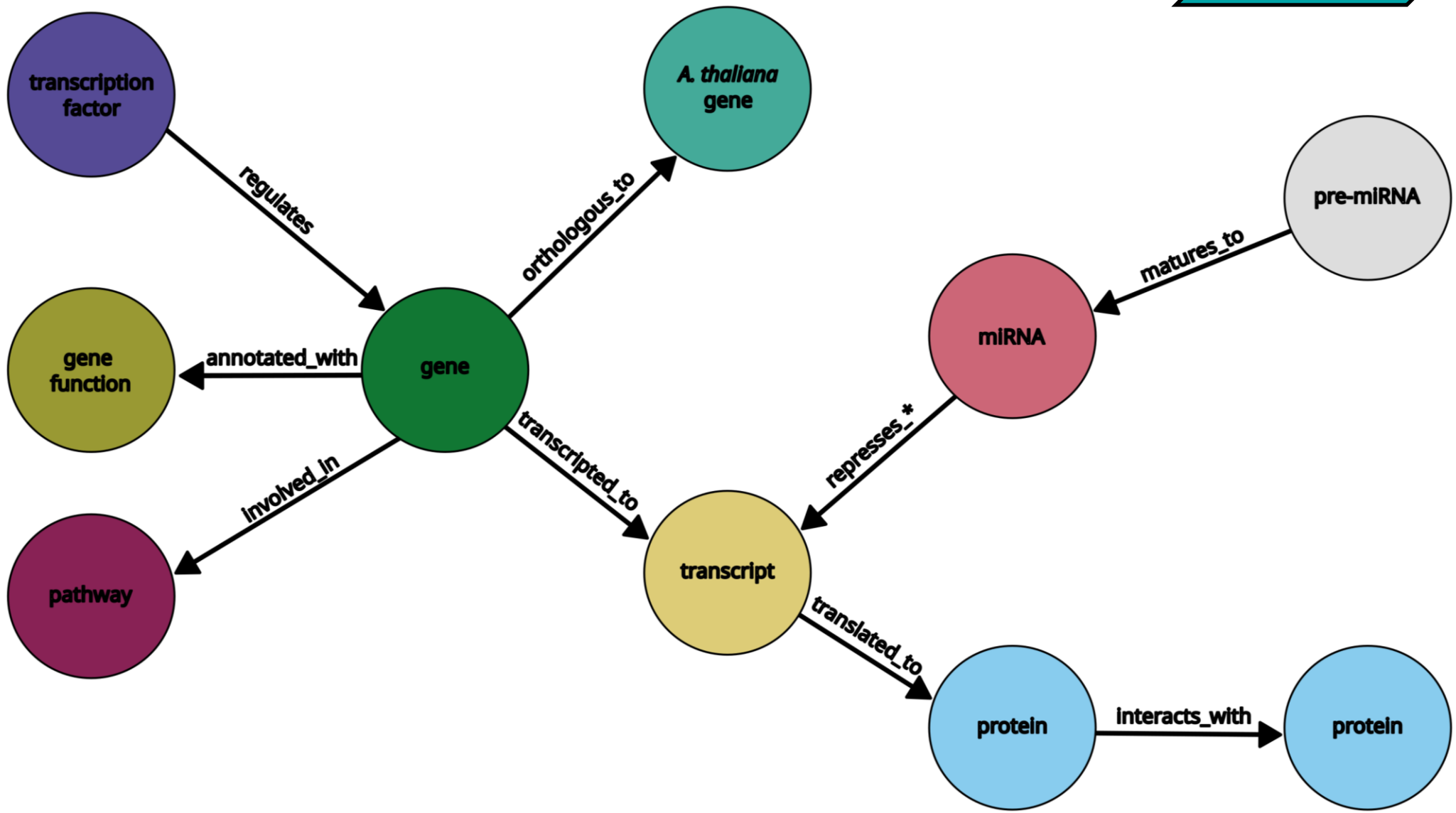


TomTom: Tomato Knowledge Graph

Labeled Property Graph

TomTom

From 10+ databases



TomTom: Tomato Knowledge Graph

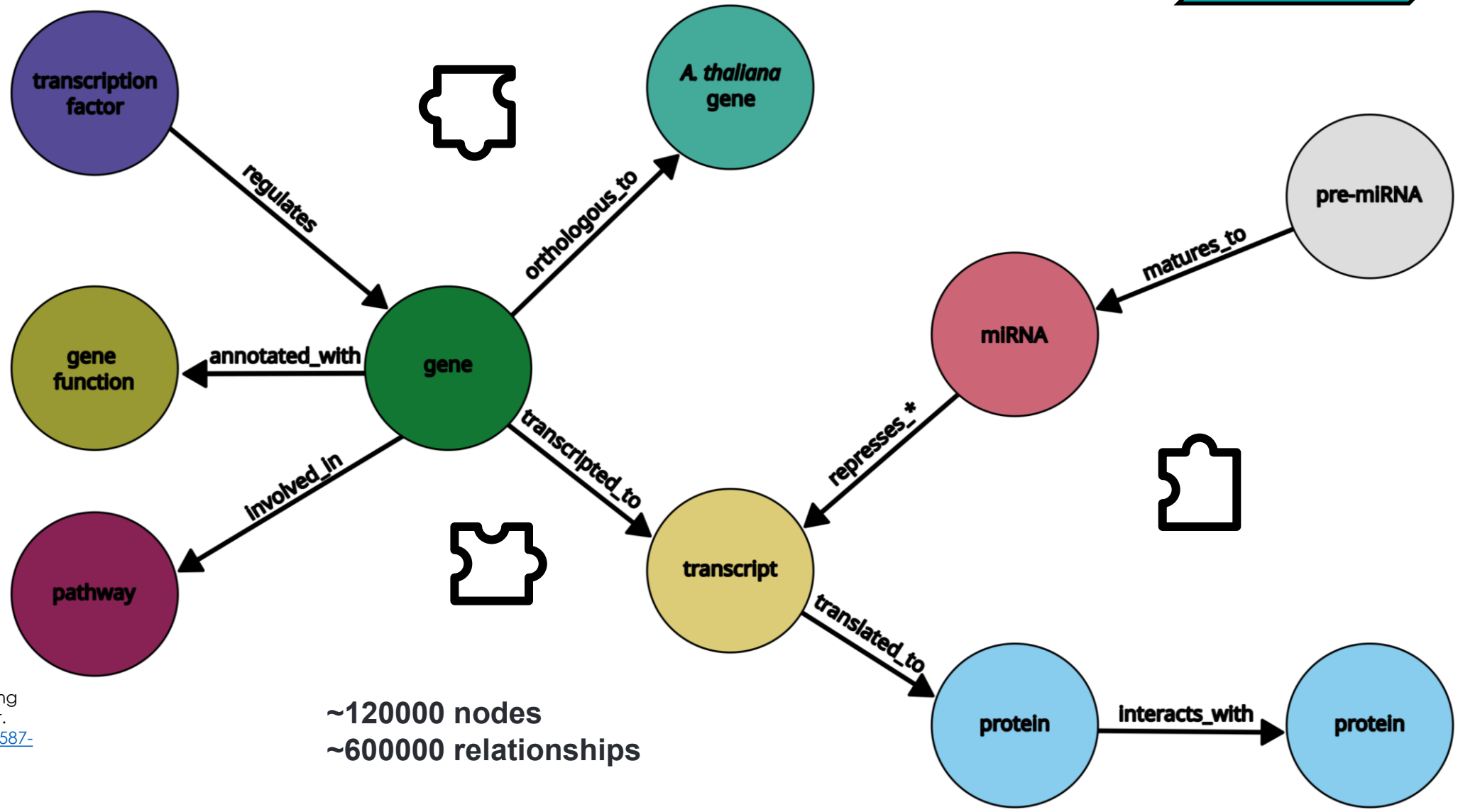
Labeled Property Graph

TomTom

From 10+ databases

csv

API



biocypher
a unifying framework for biomedical knowledge graphs

Lobentanzer, S., et al. (2023). Democratizing knowledge representation with BioCypher. Nat Biotechnol *41*, 1056–1059. 10.1038/s41587-023-01848-y.

~120000 nodes
~600000 relationships

TomTom: User Interface

TomTom

Welcome | Search | Network

Search in TomTom

To search some gene/protein you must use the organized locus name like Solyc08g005440

Search for a Node

Solyc08g005440

Select Molecular Level:

Gene × Transcript × Protein ×

Select Gene Relationships:

TF regulation × Pathways × Arabidopsis orth… ×

Select Transcript Relationships:

miRNA repressio… × miRNA repressio… × miRNA repressio… × miRNA precursors ×

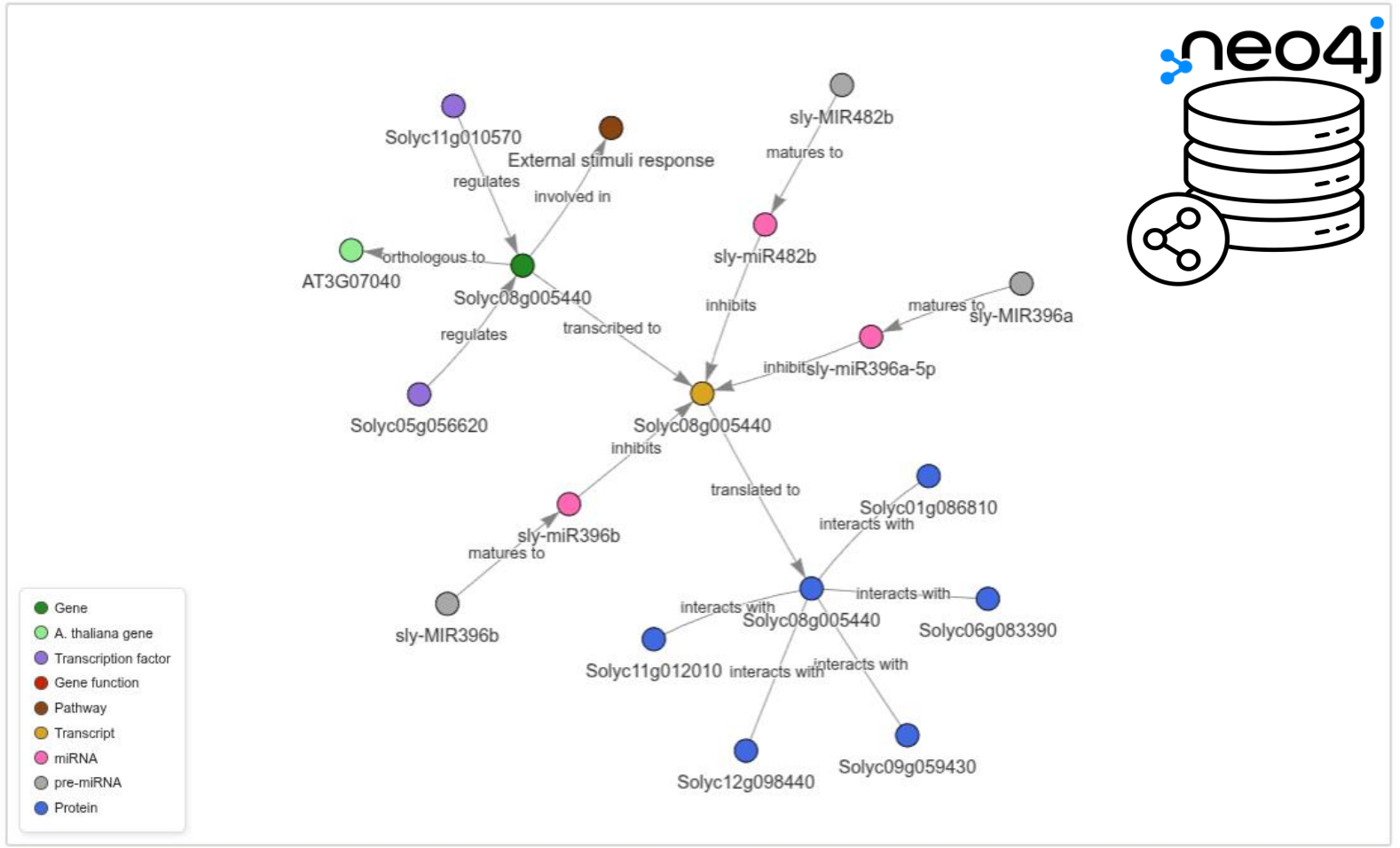
Select Protein Relationships:

Protein-protein i… ×

Output Table :

Source	Relationship	Target
Solyc06g083390	interacts with	Solyc08g005440
Solyc08g005440	translated to	Solyc08g005440
Solyc01g086810	interacts with	Solyc08g005440
Solyc08g005440	interacts with	Solyc09g059430
Solyc08g005440	interacts with	Solyc11g012010
Solyc08g005440	interacts with	Solyc12g098440
Solyc11g010570	regulates	Solyc08g005440

> Found 16 connections for 'Solyc08g005440' at the 'Gene, Transcript, Protein' level(s) – Relationships: Gene: TF regulation, Pathways, Arabidopsis orthologs | Transcript: miRNA repression (TarDB), miRNA repression (DPMIND), miRNA repression (PNRD), miRNA precursors | Protein: Protein-protein interaction

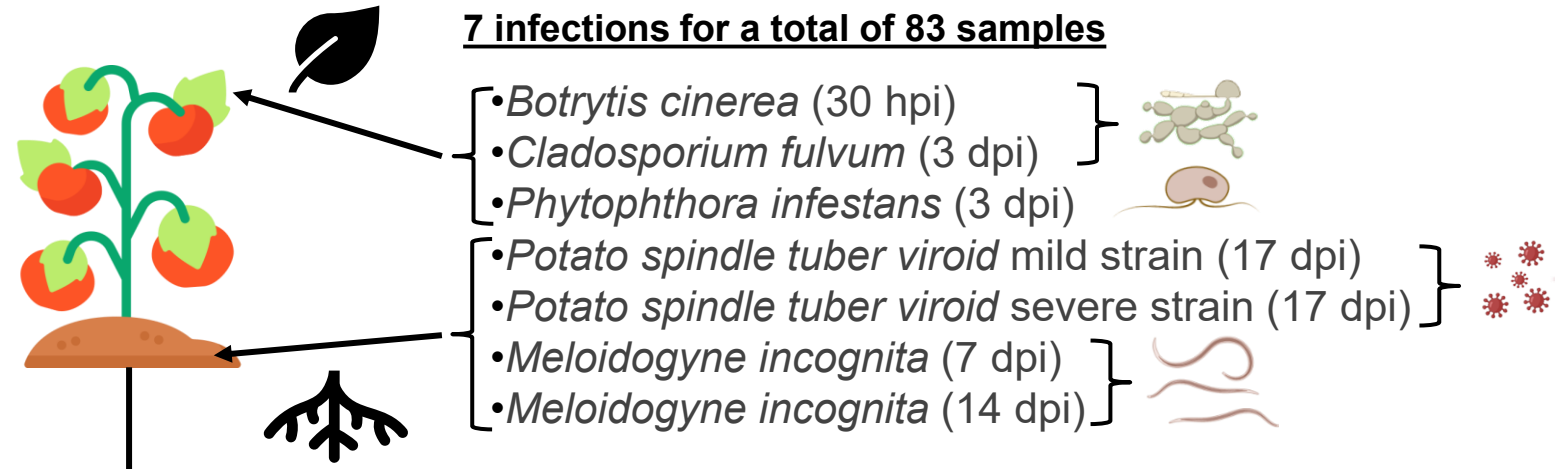


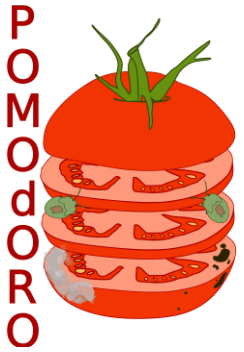
Click on nodes or edges to see their properties.

Display limitation: Only the first 100 edges are displayed.

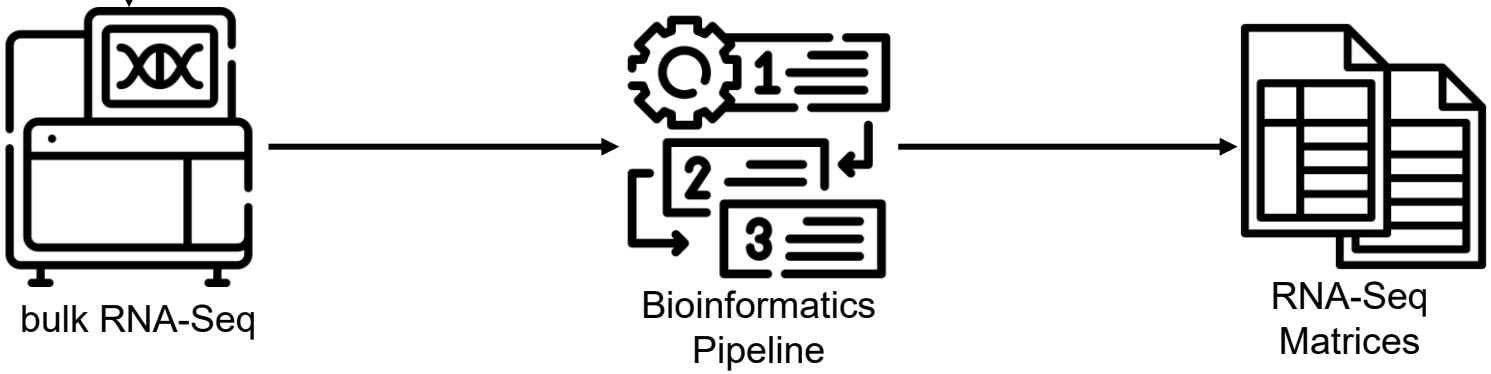
Multiple Stresses Pilot Data

Data





Database of RNA-seq data of tomato-biotic stress with metadata



Veronico, P. *et al. Front. Plant Sci.* **13**, 817185 (2022).

Courbier, S. *et al. Plant Physiol.* **187**, 1250–1266 (2021).

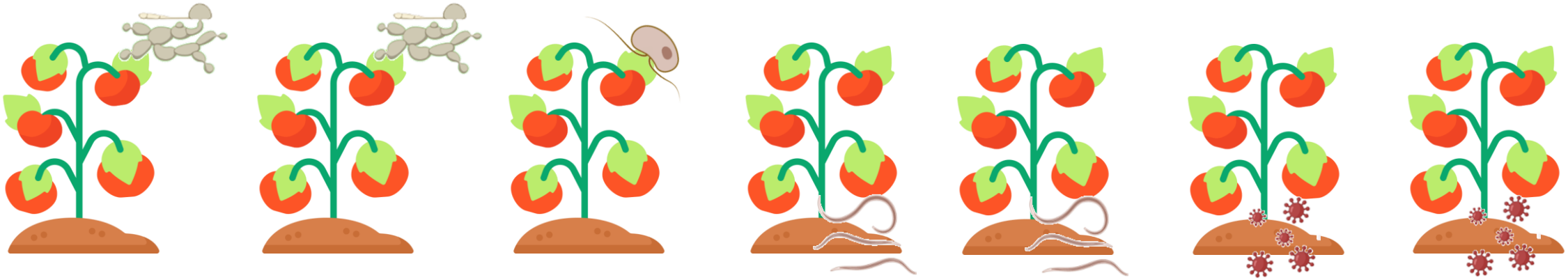
Góra-Sochacka, A., *et al. Viruses* **11**, (2019).

Liu, G. *et al. Front. Genet.* **14**, 1158631 (2023).

Canto-Pastor, A. *et al. Proc Natl Acad Sci USA* **116**, 2755–2760 (2019).

Traditional Approach

Gene Selection



Differential Expression Analysis

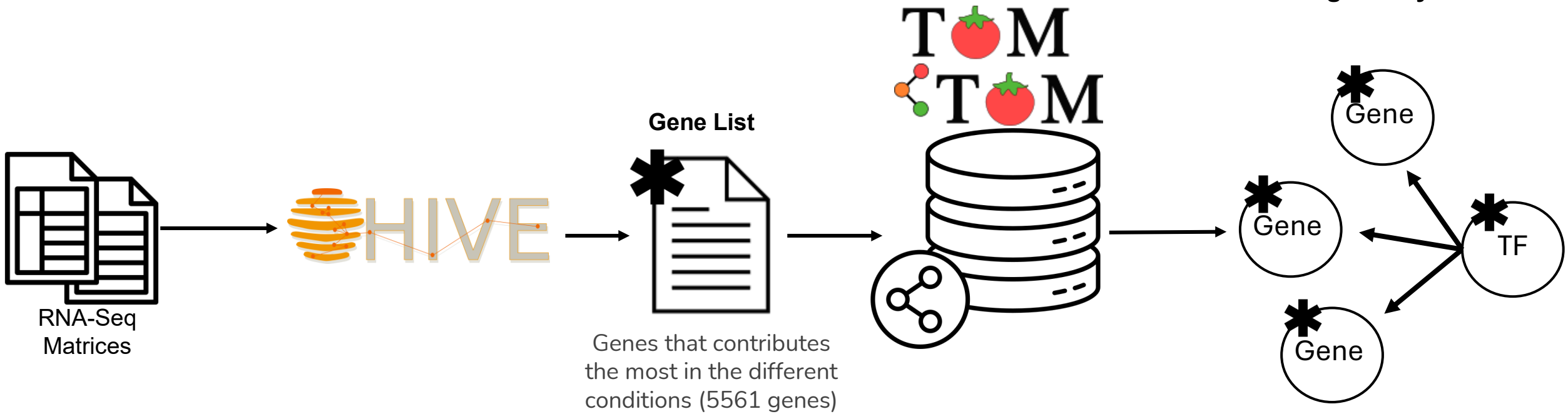
DEG List

50% of the *Sly* genome to analyse

Multi-transcriptomics analysis

Gene Selection

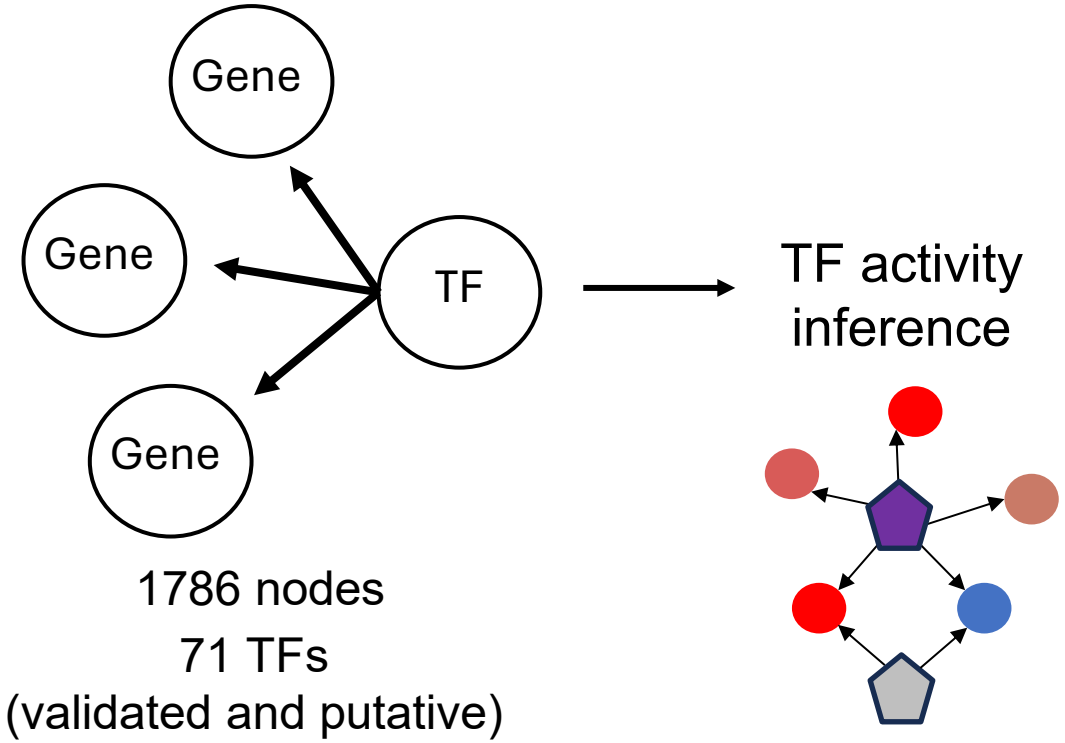
- **Gene Regulatory Network**



Calia *et al.* Modeling omics integration with HIVE identifies response signatures to multifactorial stress in plants. *Plant Physiology* **199**, (2025).

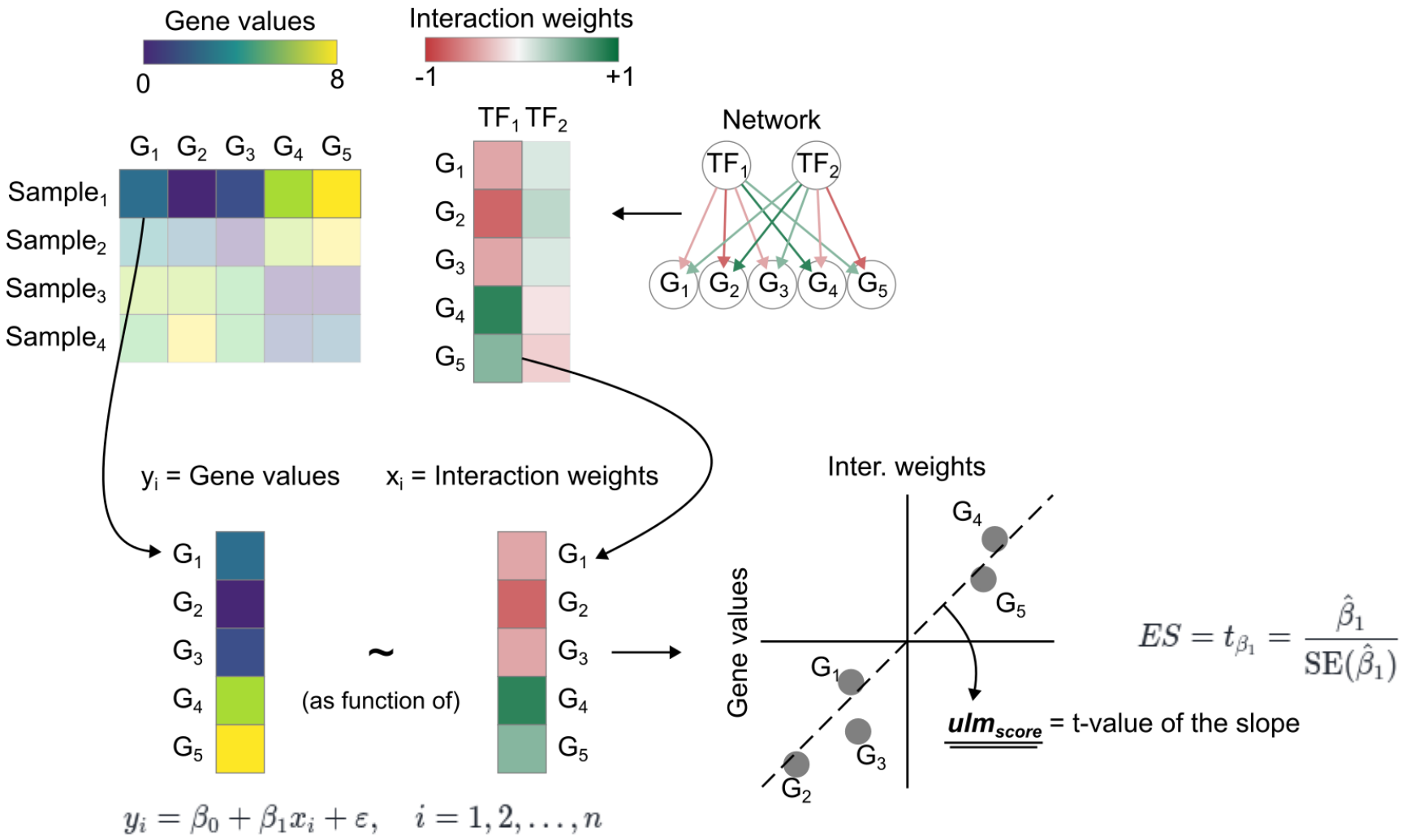
GENIAL Framework

- **Study the regulatory mechanisms involved in the genes of interest**



TF activity inference: Principles

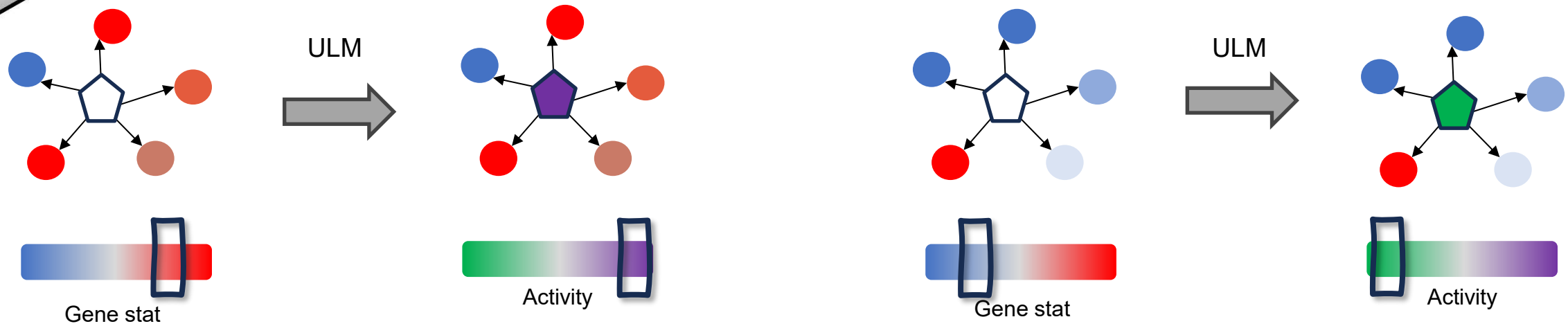
Univariate Linear Model (ULM)

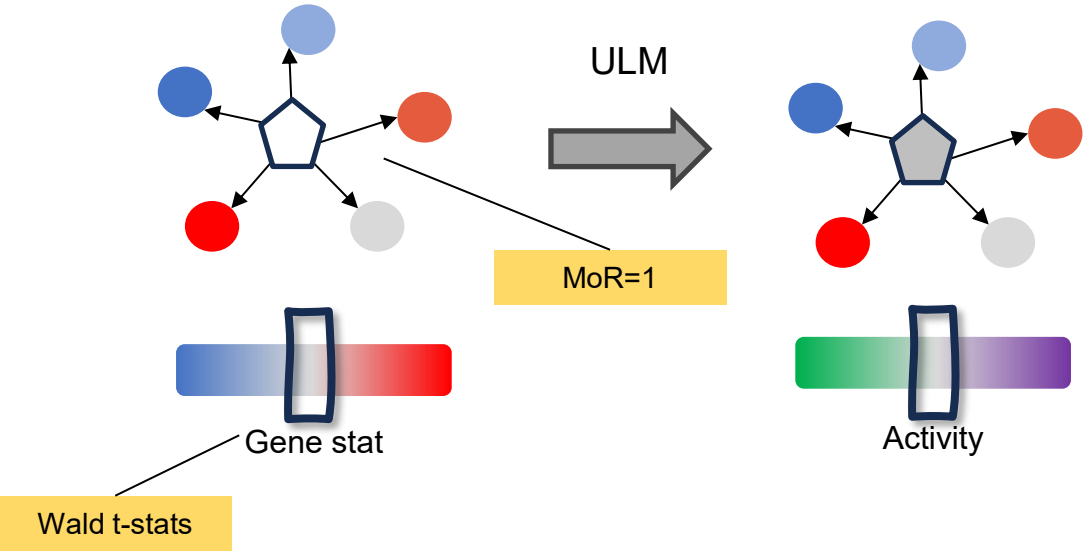


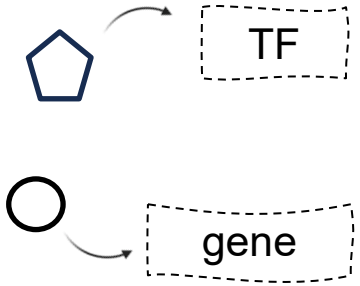
Badia-i-Mompel, P., et al. (2022). decoupleR: ensemble of computational methods to infer biological activities from omics data. Bioinformatics Advances 2, vbac016. https://doi.org/10.1093/bioadv/vbac016

TF activity inference: Principles

Activity Inference



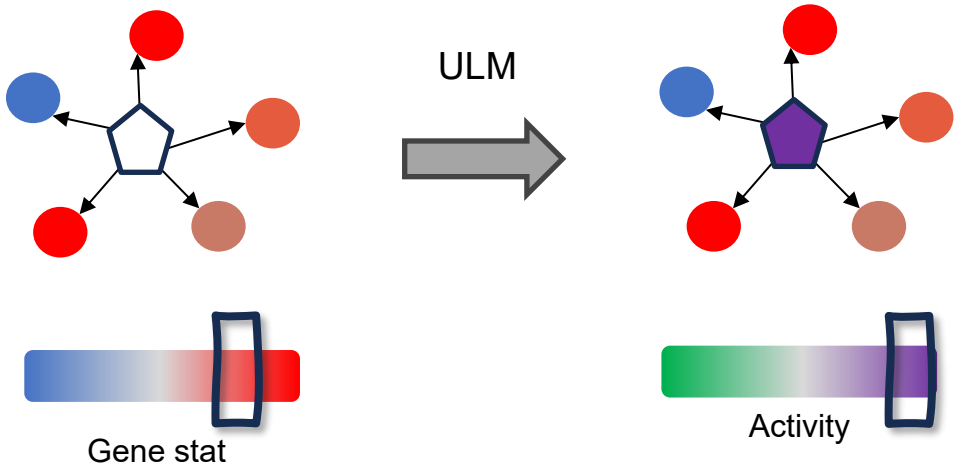




TFs with significant activity in response to specific or multiple pathogens attacks.

TF activity inference: Adaptation

BH correction was too stringent due to the relatively high number of targets/TF + noisy t-stats because unfiltered
Compared the obtained activity against the permutation null activity (1000-degree-preserving shuffle of the input network)

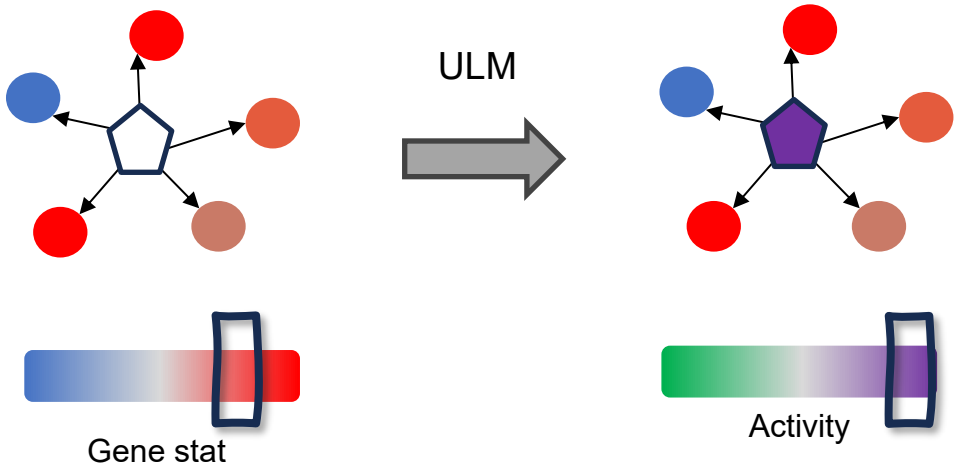


Real inferred activity = X
P-val <= 0.05 (>= 1 condition)

Could the activity be observed by chance?

TF activity inference: Adaptation

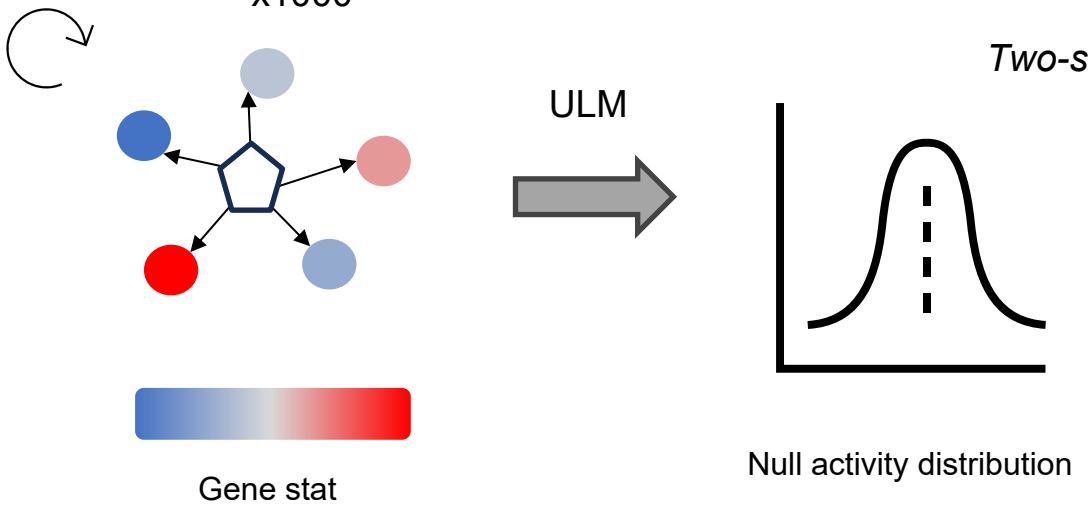
BH correction was too stringent due to the relatively high number of targets/TF + noisy t-stats because unfiltered
Compared the obtained activity against the permutation null activity (1000-degree-preserving shuffle of the input network)



Real inferred activity = X
P-val <= 0.05 (>= 1 condition)

Could the activity be observed by chance?

Degree-preserving permutation
x1000



Two-sided empirical p-val

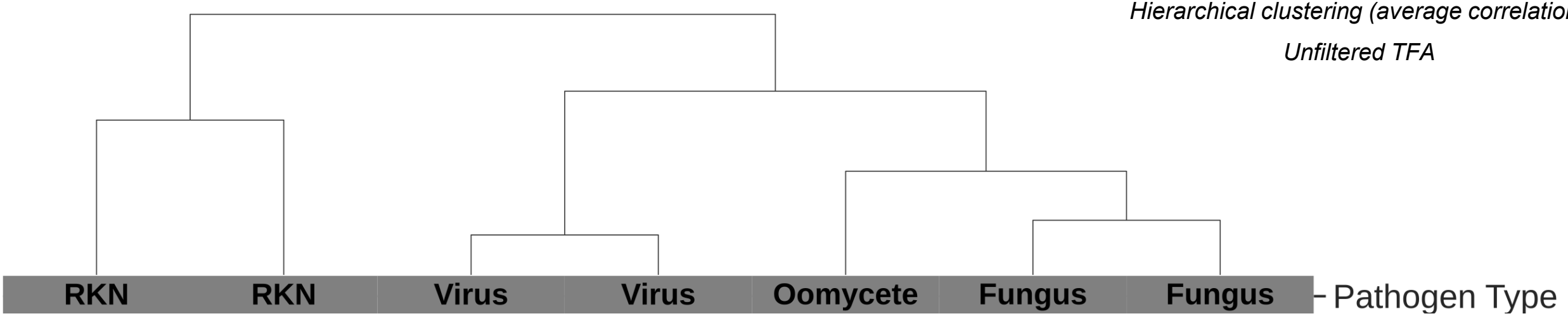
TF activity retained for later:
(p-val <= 0.05 & emp p-val <=0.05) in >= 1 condition

TF activity distinguish plant responses

Activity Inference

Hierarchical clustering (average correlation)

Unfiltered TFA



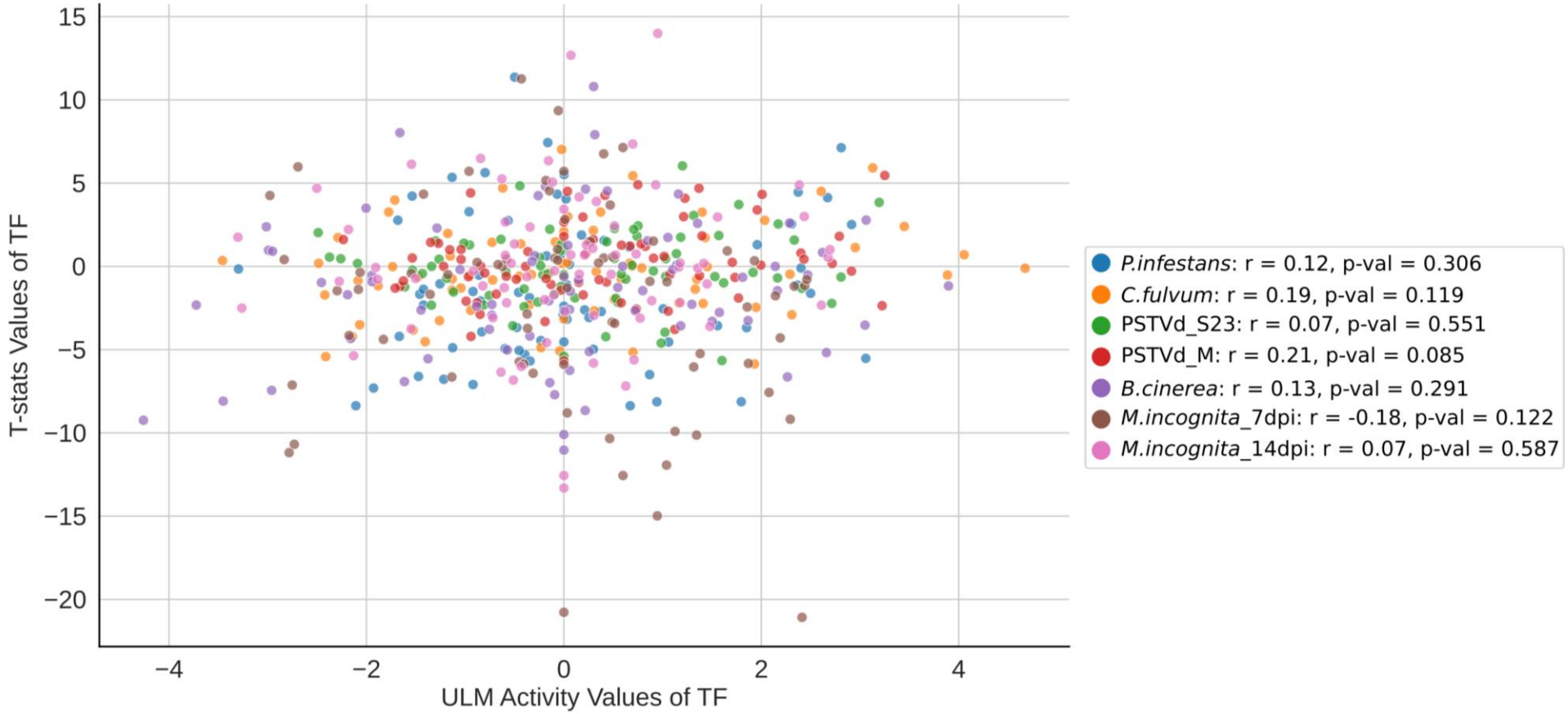
- **TF activity reflects the integrated data**
- **Not happening when looking at gene expression/t-stats**

TF activity VS TF t-stats

Activity Inference

- **Activity adds another information**

Pearson correlation
Unfiltered TFA

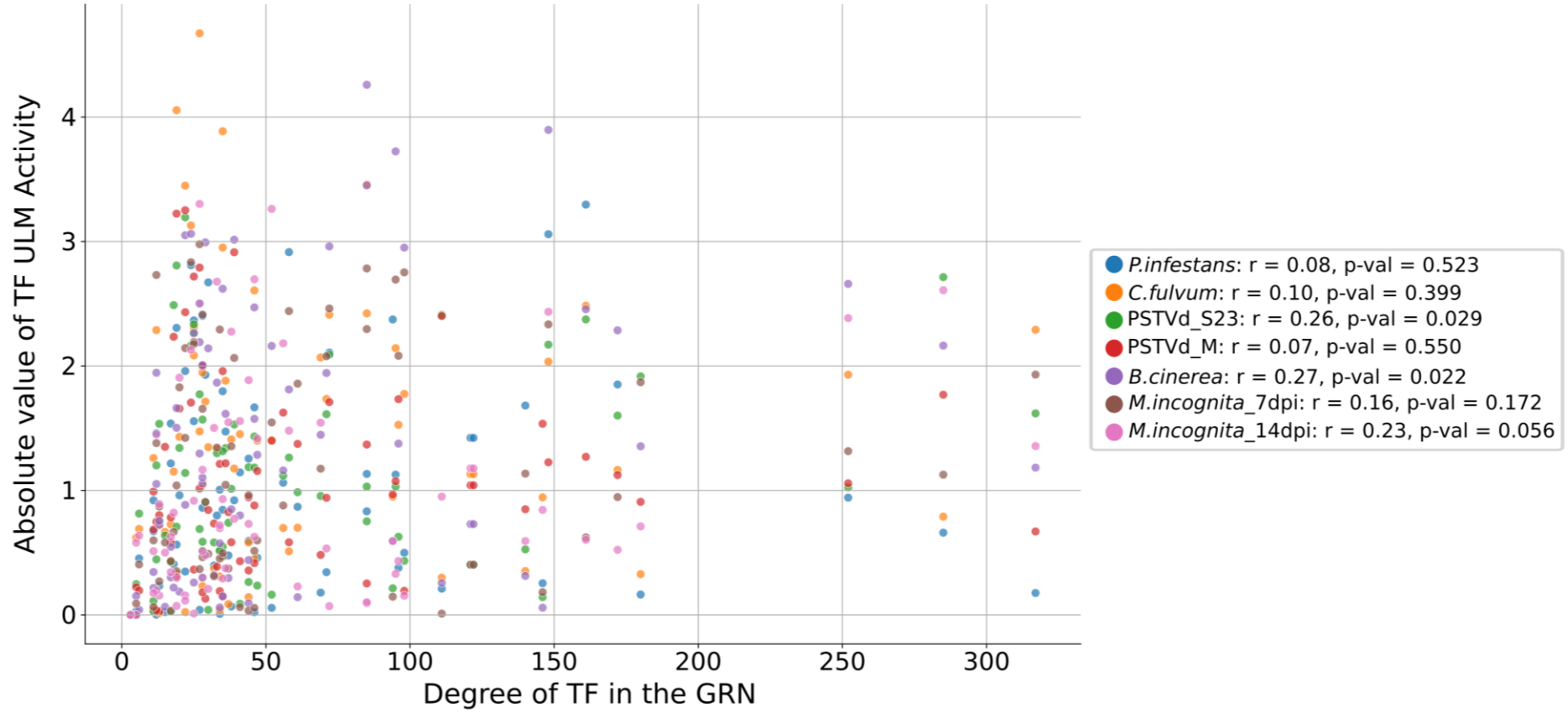


TF activity VS TF degree

Activity Inference

- **Activity adds another information**
- **Without having a size effect**

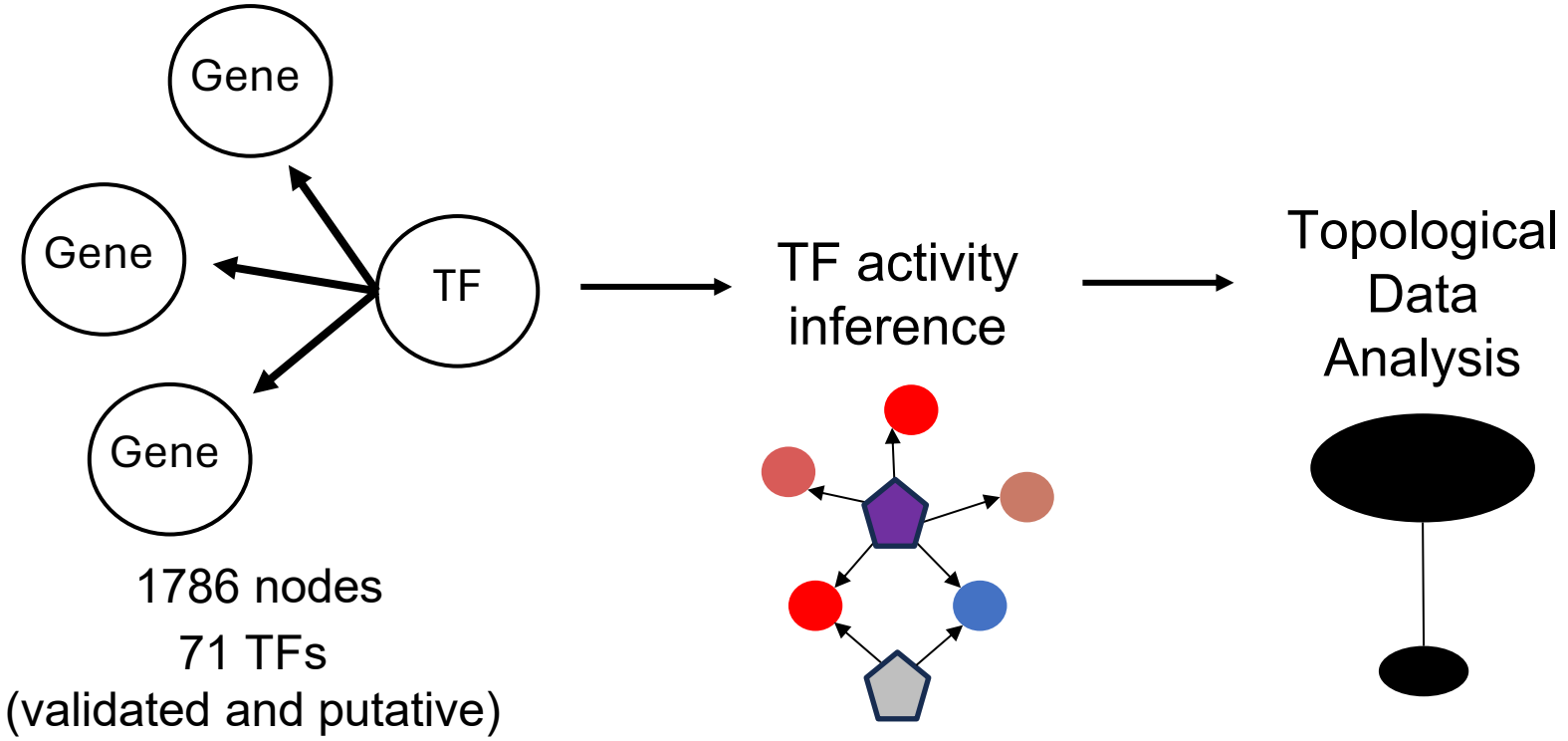
Pearson correlation
Unfiltered TFA



GENIAL Framework

GENIAL

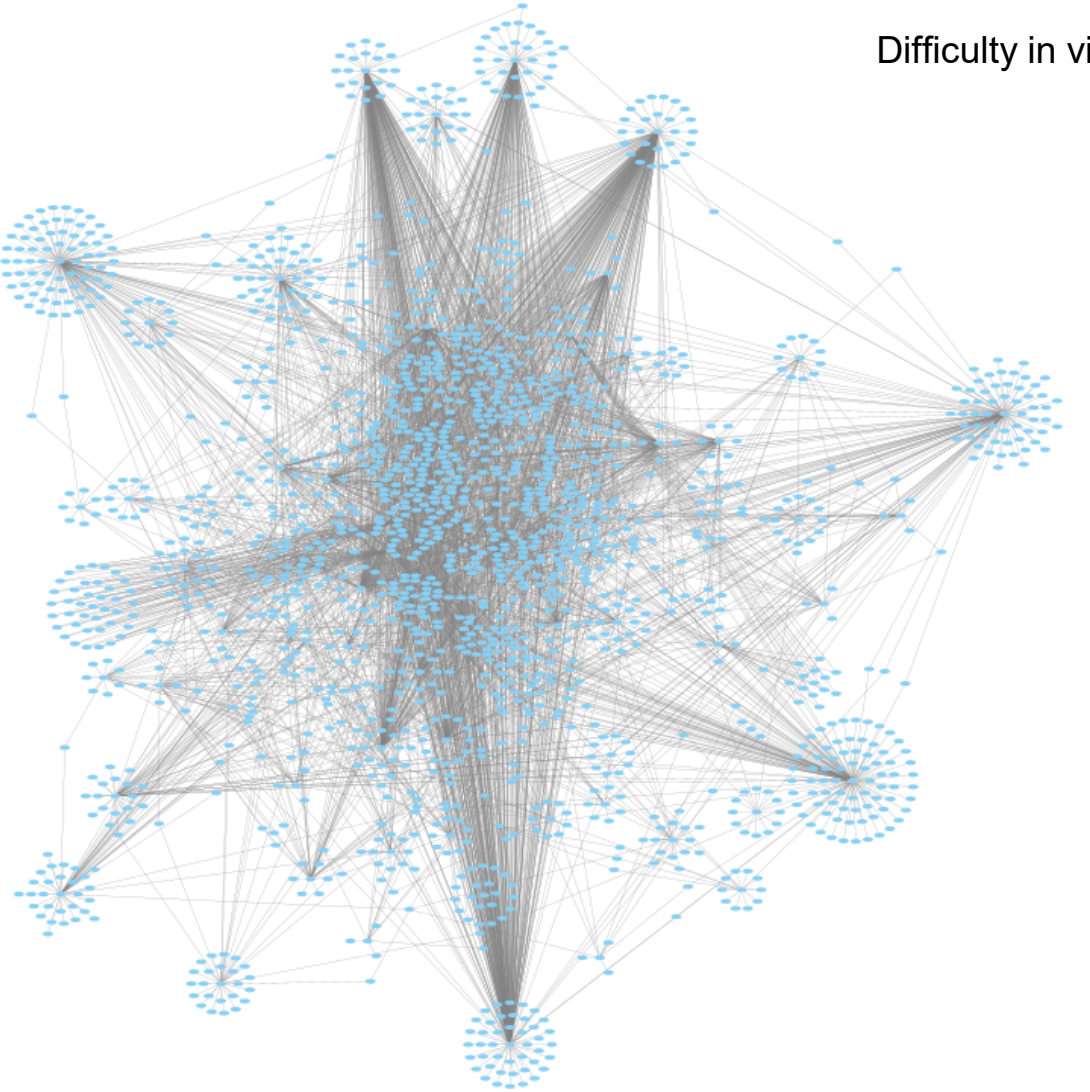
- **Study the regulatory mechanisms involved in the genes of interest**



Topological Data Analysis: Apply to GRN

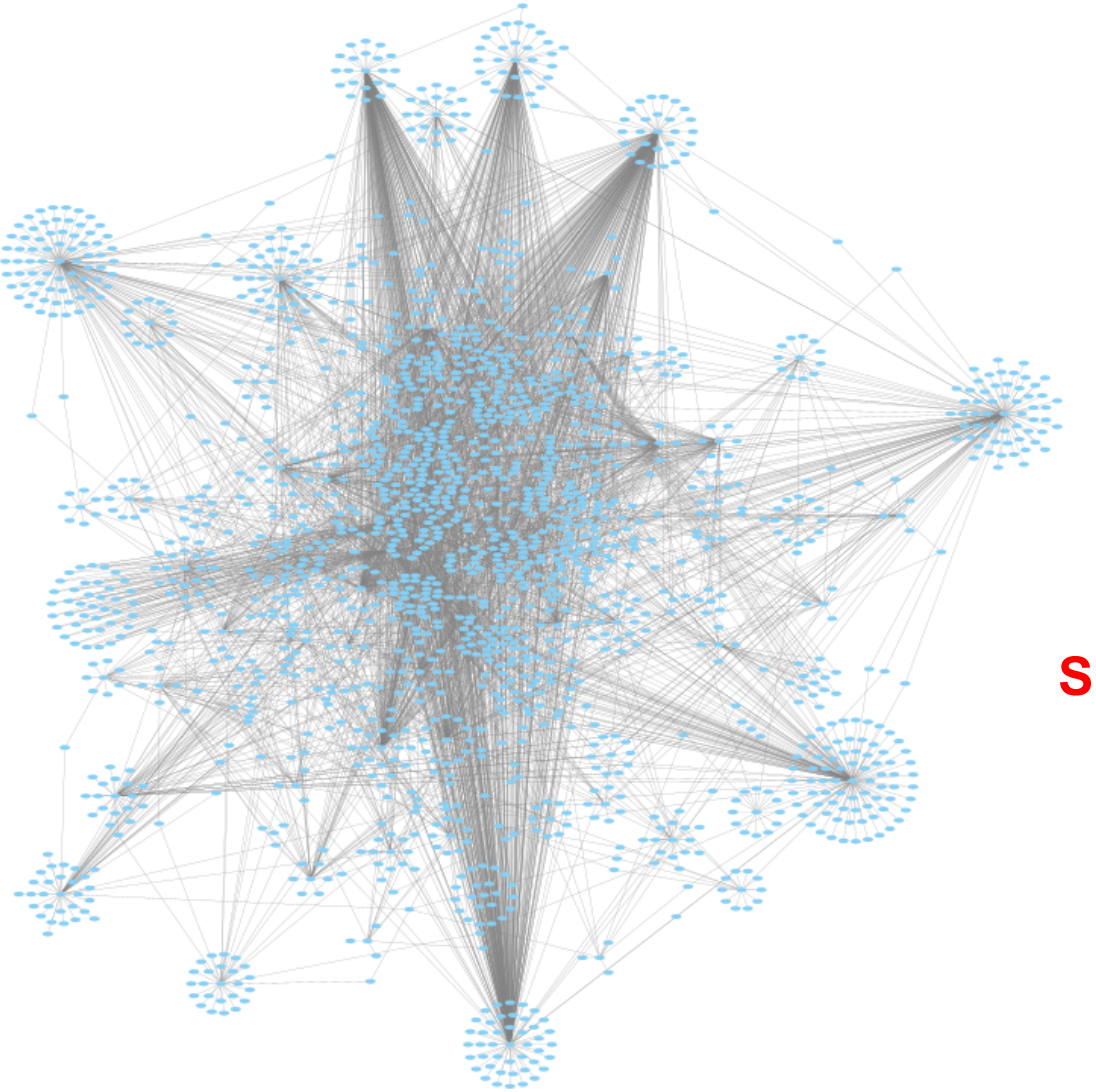
TDA

Difficulty in visualizing the TF activity in the GRN



Topological Data Analysis: Apply to GRN

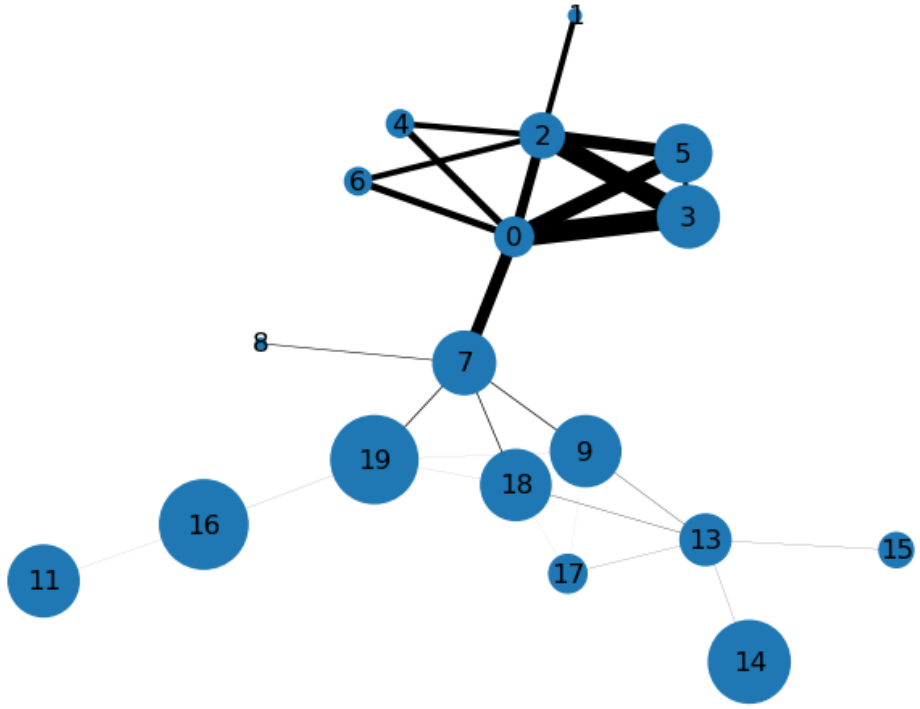
TDA



Mapper

TDA

Simplified View

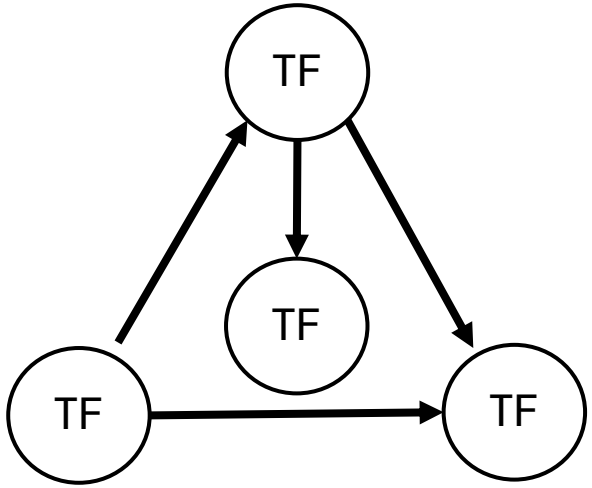


Topological Data Analysis: Apply to GRN

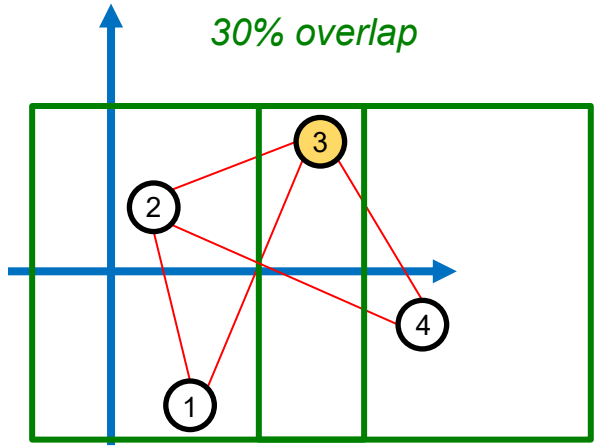
TDA

Filter = Eigenvectors of GRN Laplacian
Cover = 5x5, 30 % overlap
Clustering = GRN-based clustering

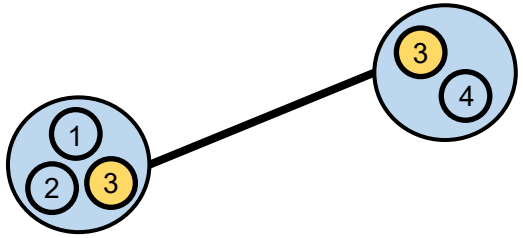
TF-TF subnetwork



Eigenvectors of the Laplacian



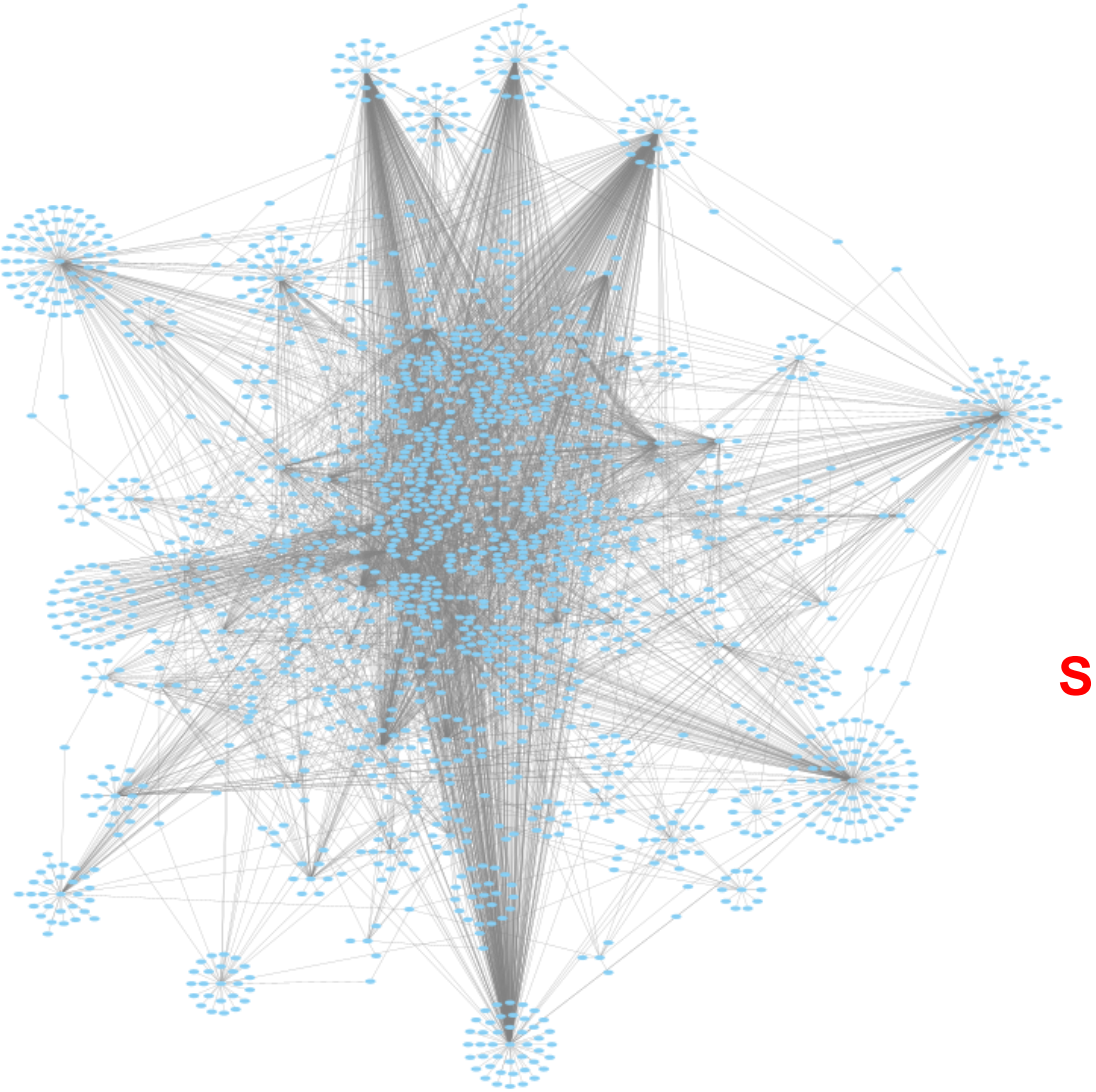
Cluster/connected if shared targets in GRN (in the same square)



Mapper graph

Topological Data Analysis: Apply to GRN

TDA

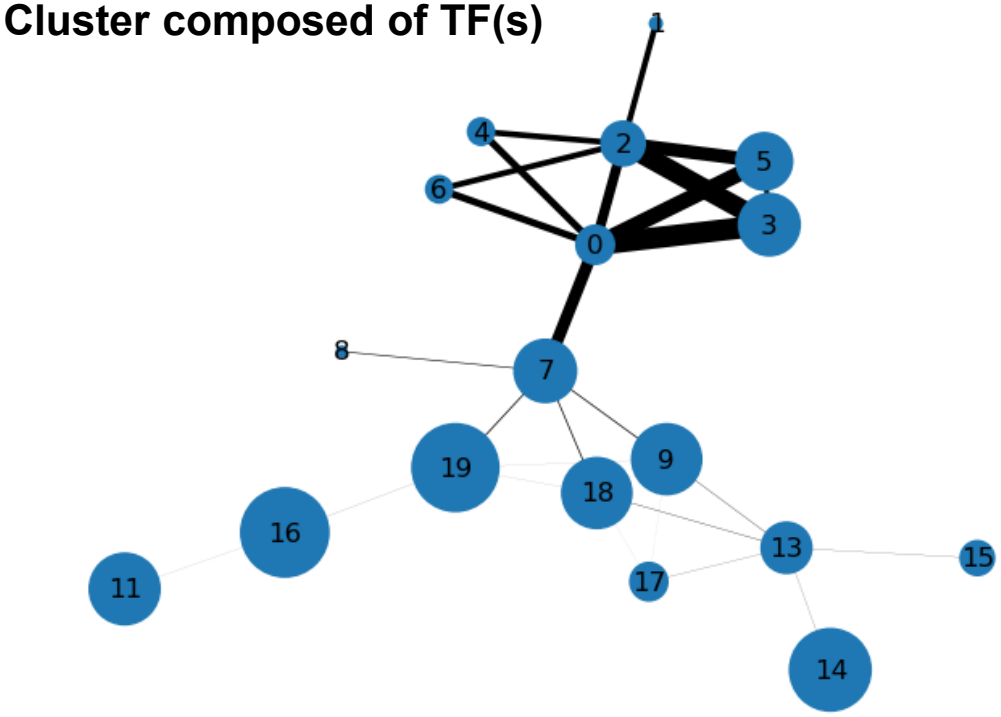


Mapper

TDA

Simplified View

Cluster composed of TF(s)



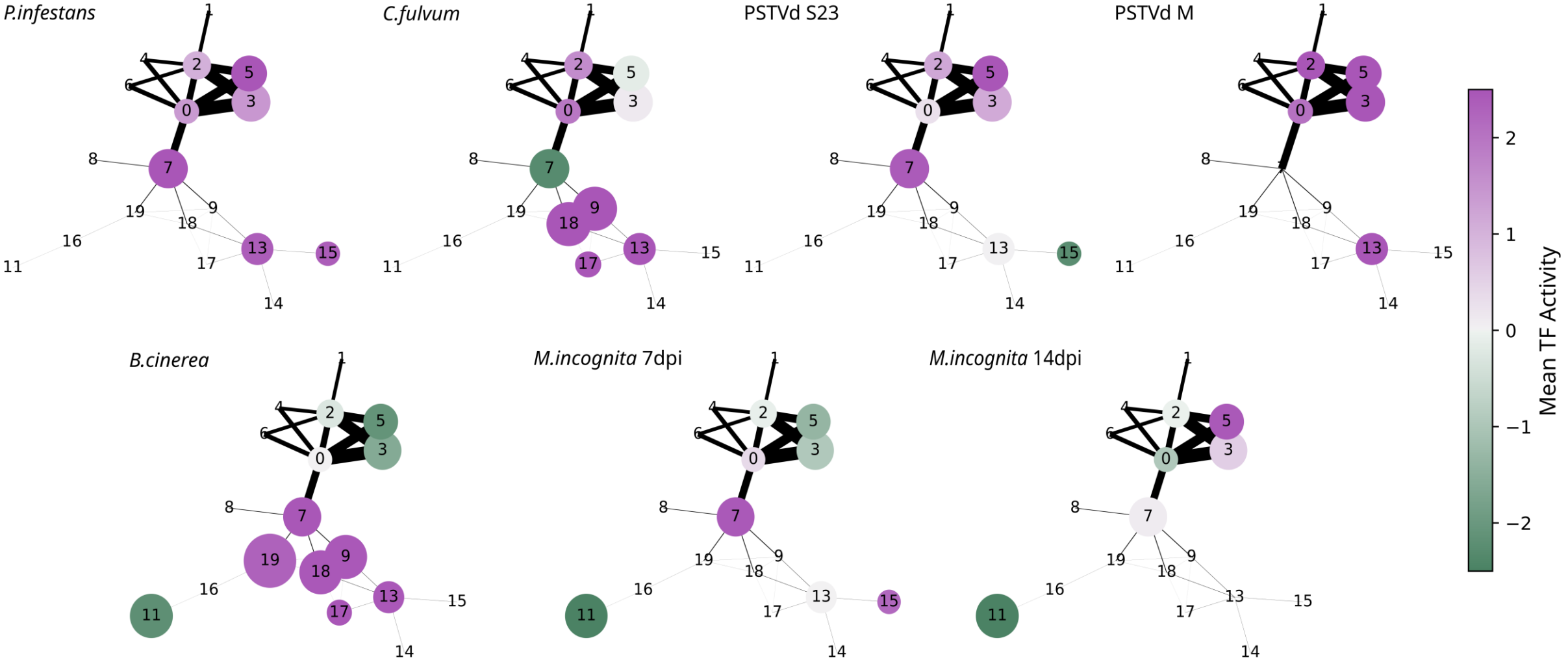
Edge size: number of shared targets between cluster

Node size: number of targets in a cluster

Clusters are identified by their number

Combining TDA and TF activity allow to summarize the response

TDA + TF Activity



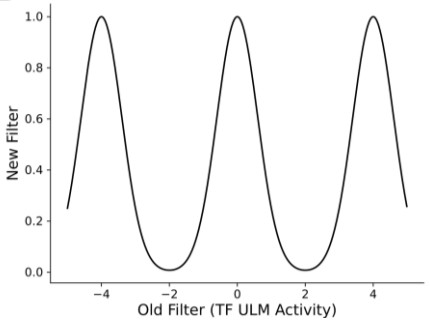
Combining TDA and TF activity allow to summarize the response

TDA + TF Activity

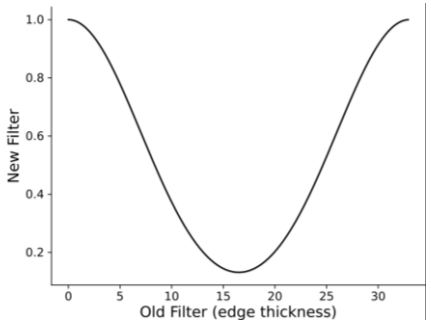
ToMATo: TDA graph clustering

Activity function

Once per condition



Connectiveness



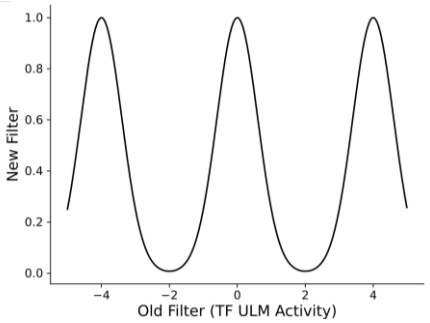
Combining TDA and TF activity allow to summarize the response

TDA + TF Activity

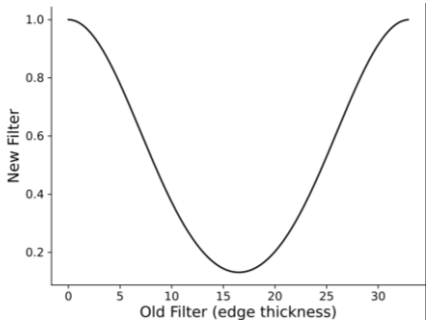
ToMATo: TDA graph clustering

Activity function

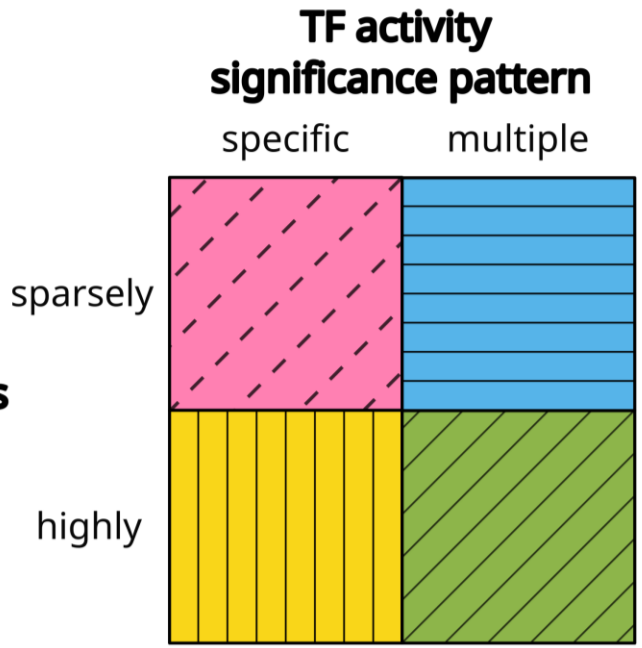
Once per condition

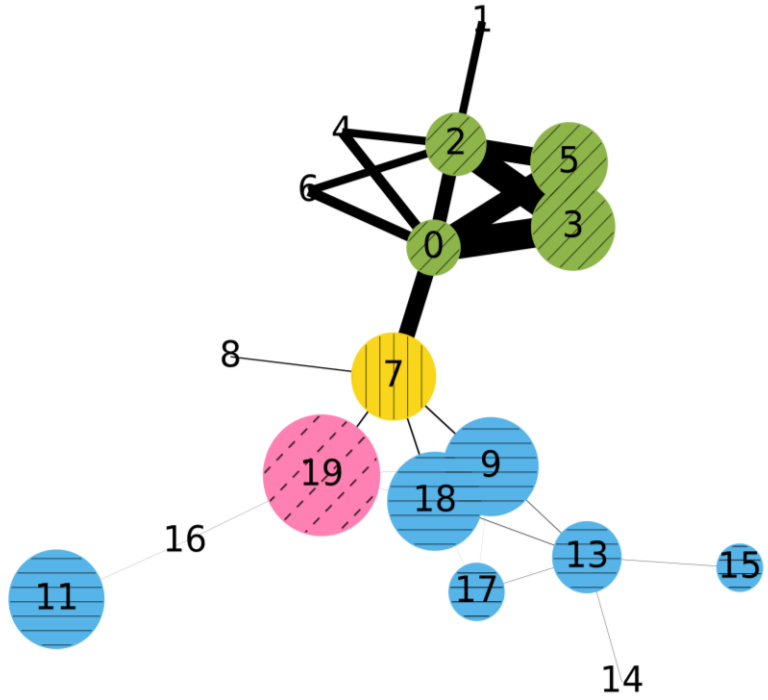


Connectiveness



- **Four configurations that summarize the response**



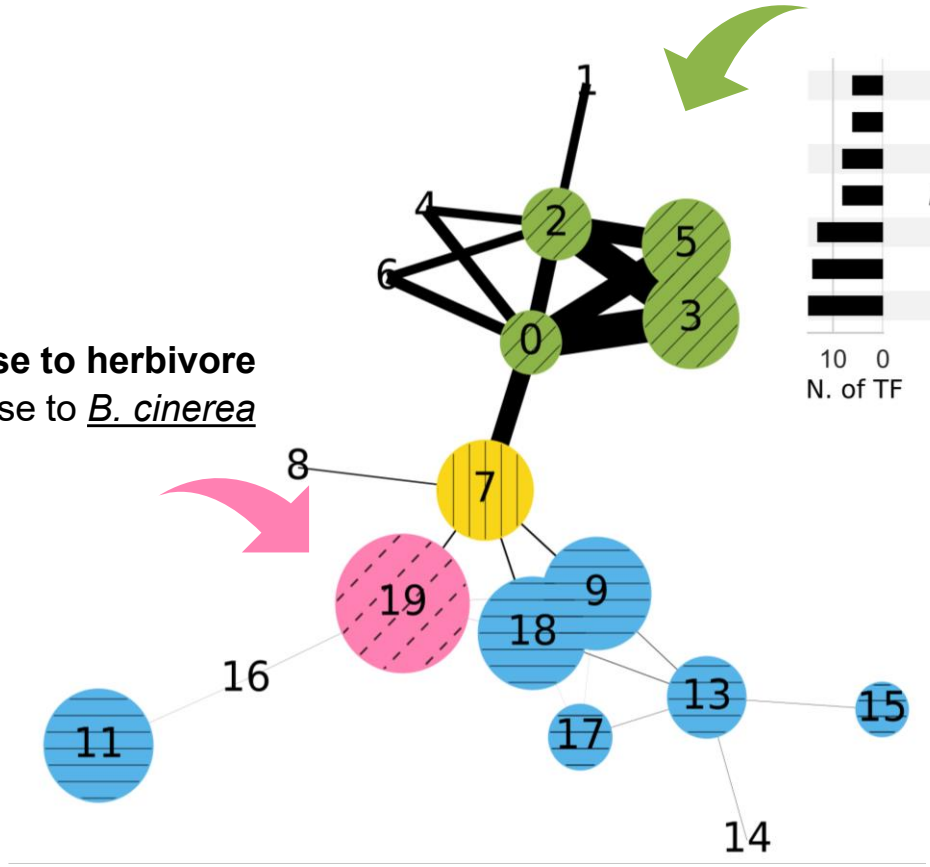


Combining TDA and TF activity allow to summarize the response

TDA + TF Activity

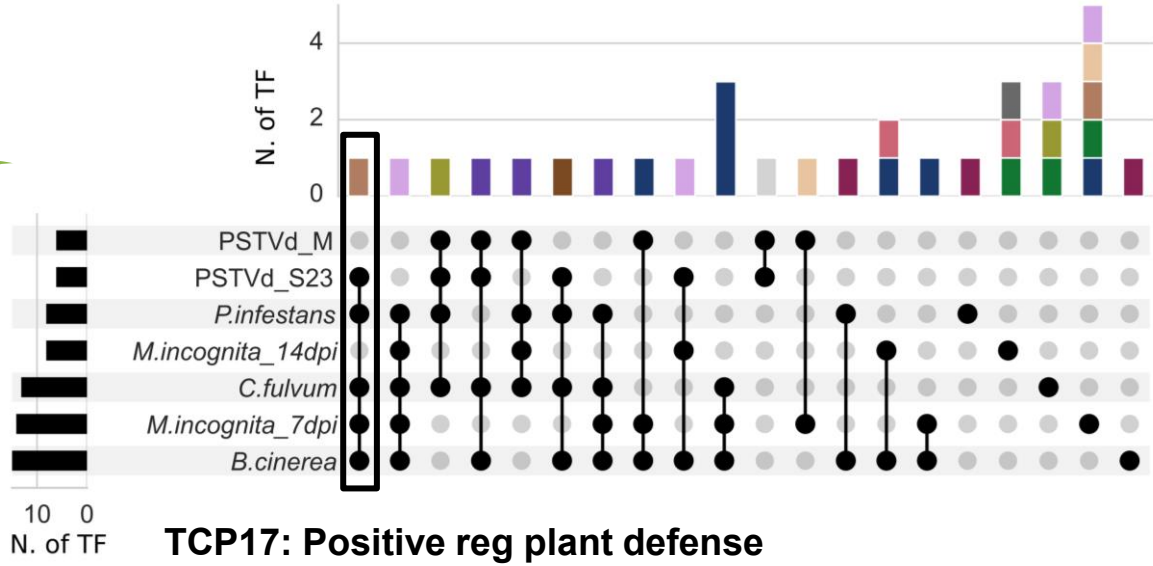
- **Find representative TFs in each configuration**

TF Family
- bZIP
- bHLH
- WRKY
- TCP
- NAC
- MYB_related
- MYB
- HSF
- GATA
- ERF
- Dof
- CPP
- E2F/DP
- ARF

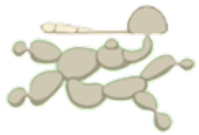


ERF16: response to herbivore
Found in response to *B. cinerea*





TCP17: Positive reg plant defense
Found in response to *B. cinerea*



Biological validation of representative TFs in response to *B. cinerea*

Validation

Virus-induced gene silencing (VIGS):
A transient gene knockdown technique using the plant antiviral RNA-silencing system to silence a target gene throughout PTGS.



We used VIGS to knockdown the previous ERF16 and TCP17 and observe the tomato response to *B. cinerea*



Biological validation of representative TFs in response to *B. cinerea*

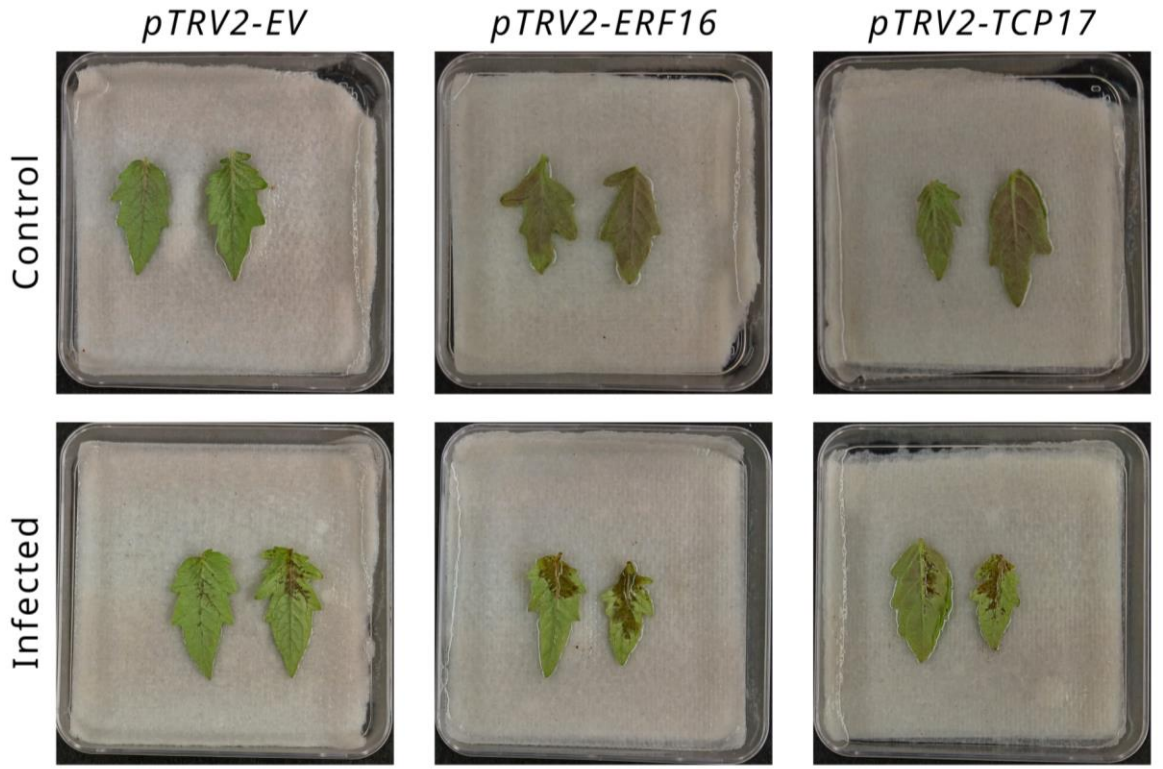
Validation

Virus-induced gene silencing (VIGS):
A transient gene knockdown technique using the plant antiviral RNA-silencing system to silence a target gene throughout PTGS.



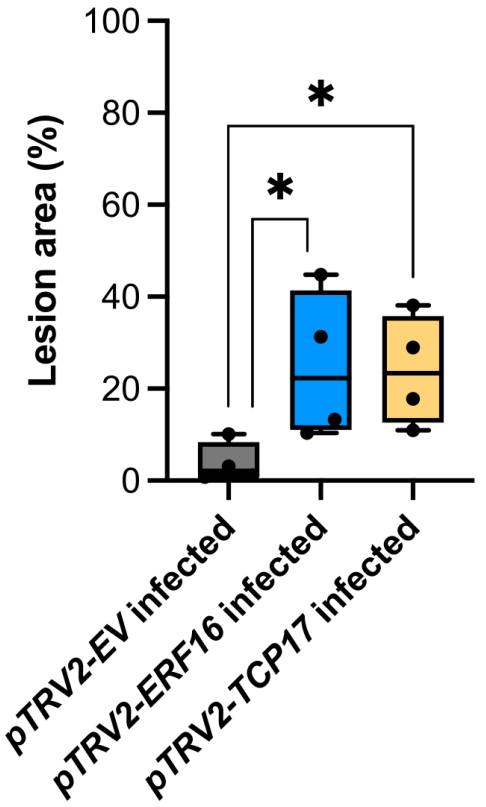
We used VIGS to knockdown the previous ERF16 and TCP17 and observe the tomato response to *B. cinerea*





Increase susceptibility of *B. cinerea* in perturbed tomato

Confirmed gene knockdown by qRT-PCR

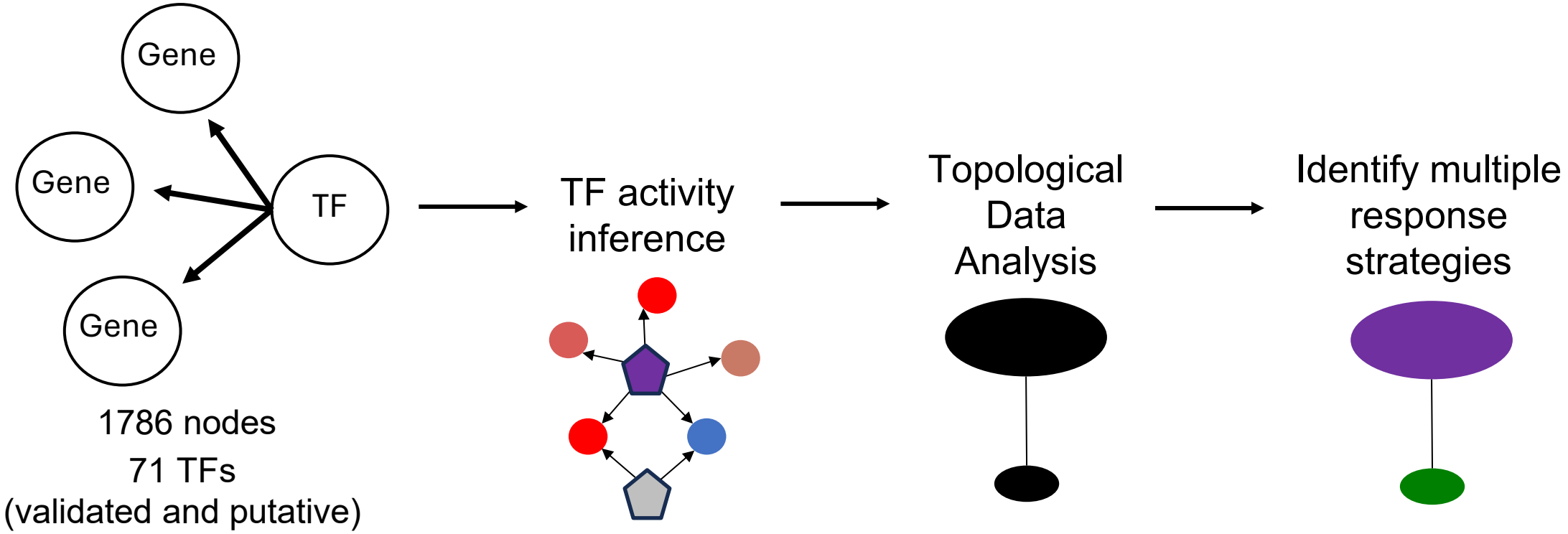


ERF16 and TCP17 are important in the tomato response to *B. cinerea*

GENIAL Framework

GENIAL

- **Study the regulatory mechanisms involved in the genes of interest**



Publication

PUBLISHED

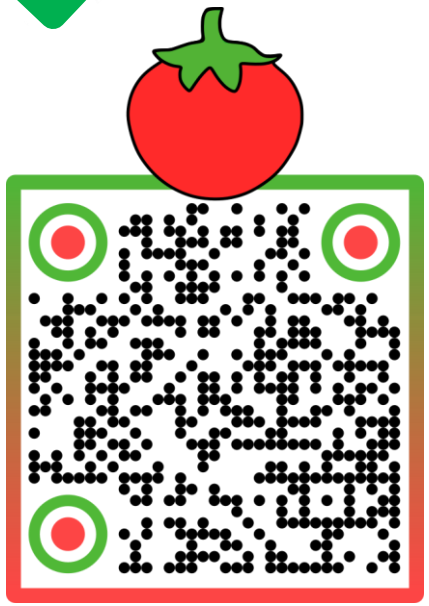
New Phytologist

Methods | Open Access

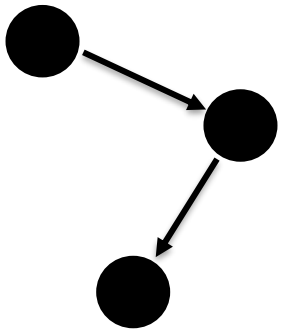
From knowledge graph to topological data analysis: a novel framework to analyze gene regulatory networks for tomato–multi-pathogen interactions

Maxime Multari, Mathieu Carrière, Xavier Amorós-Gabarrón, Krishna Kumar, Pratila Debnath, Alexina Damy, Ludo Andrianirina Mamisoa, Sebastian Lobentanzer ... See all authors

First published: 01 July 2026 | https://doi.org/10.1111/nph.71382 | VIEW METRICS



Conclusion



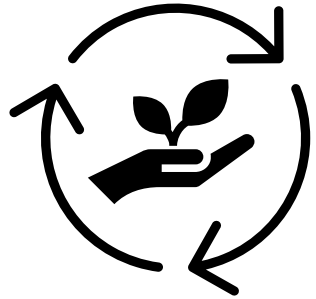
- TomTom: Tomato Knowledge Graph

TOM
TOM

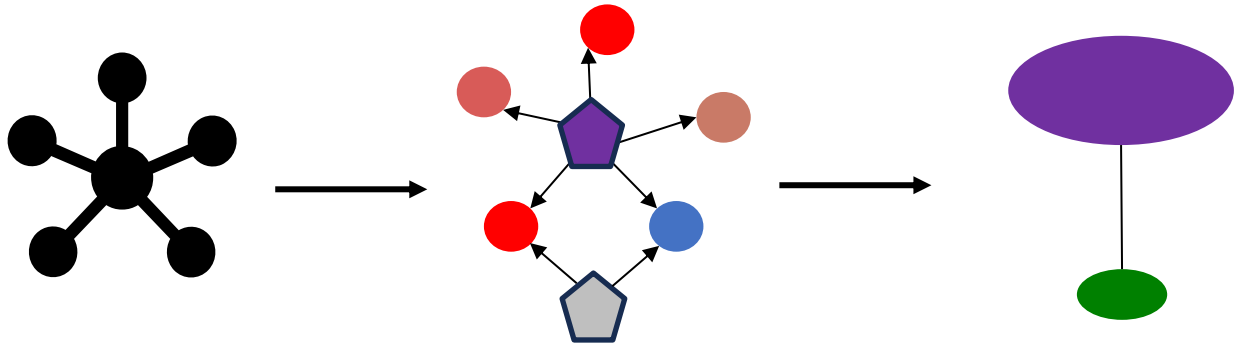
- Scalable to any questions involving Tomato



- GENIAL framework to identify key regulators in GRN
- Framework applicable to any species



GENIAL



Acknowledgments

Fundings:

UNIVERSITÉ CÔTE D'AZUR — Initiative d'Excellence (FRANCE 2030)

INRAE Département SPE

INRAE

Institut SOPHIA AGROBIOTECH

cnrs

METAPROGRAMME DIGIT-BIO

Stéphanie Jaubert

Silvia Bottini
Giulia Calia
Xavier Amorós
Alexina Damy

Inria
Mathieu Carrière
Ludo Andrianirina Mamisoa

IGZ
Justyna Olas
Krishna Kumar
Pratila Debnath

SAEZ-RODRIGUEZ Group

UniversitätsKlinikum Heidelberg

Aurélien Dugourd
Sebastian Lobentanzer
Julio Saez-Rodriguez

LIFE AND HEALTH SCIENCES
GRADUATE SCHOOL AND RESEARCH