WAGENINGEN UR
For quality of life

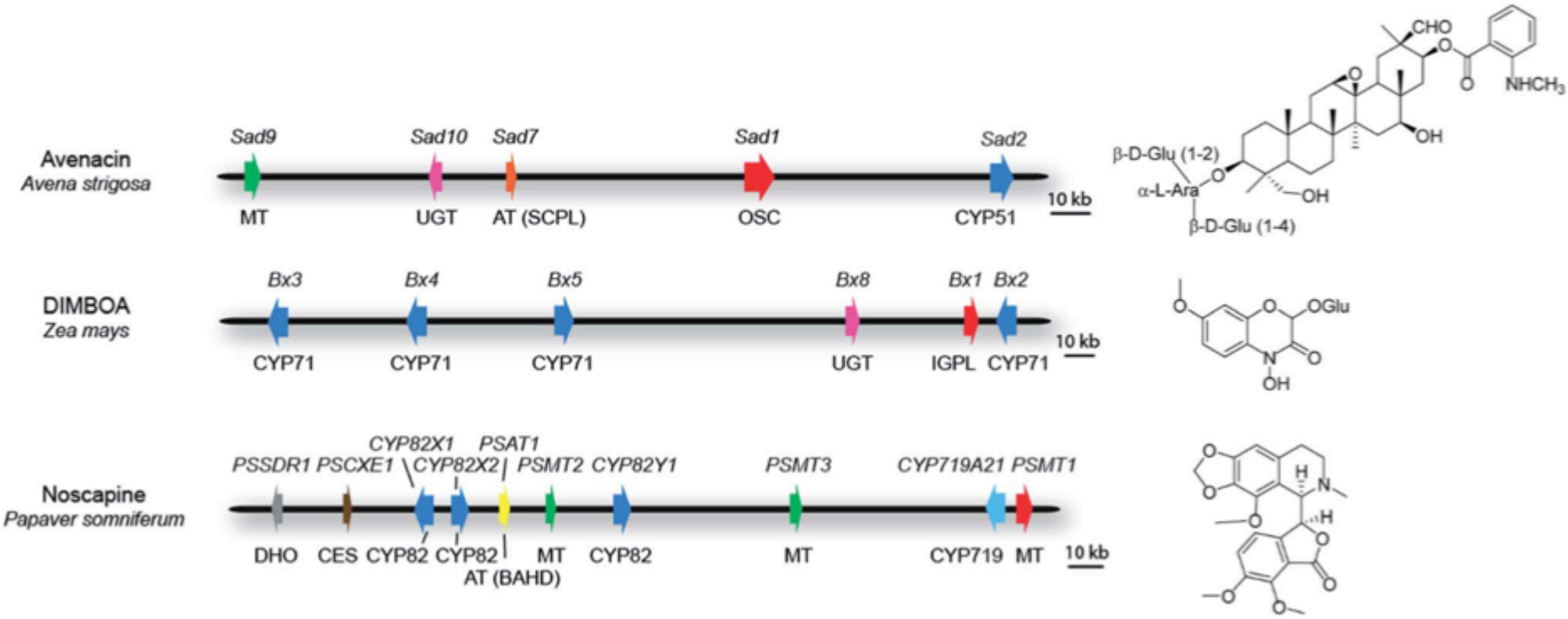
March 9th, 2021

# Computational Tools for Biosynthetic Pathway Discovery in Plants


Dr. Marnix H. Medema
Bioinformatics Group
Wageningen University, The Netherlands

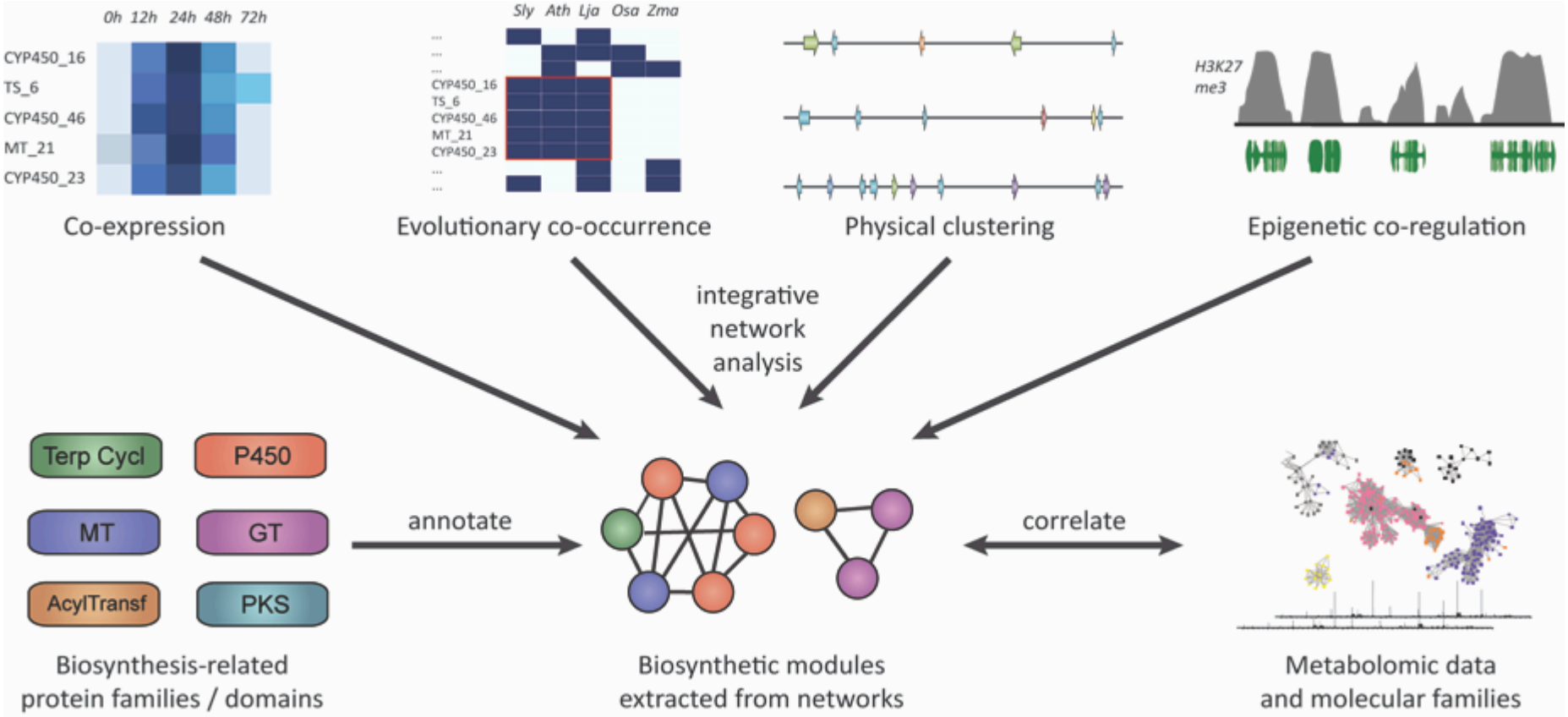
Computational Applications in Secondary Metabolite Discovery (CAiSMD)
Tuesday March 9TH, 2021


# NOT INFREQUENTLY, PLANT BIOSYNTHETIC PATHWAYS ARE ENCODED IN GENE CLUSTERS



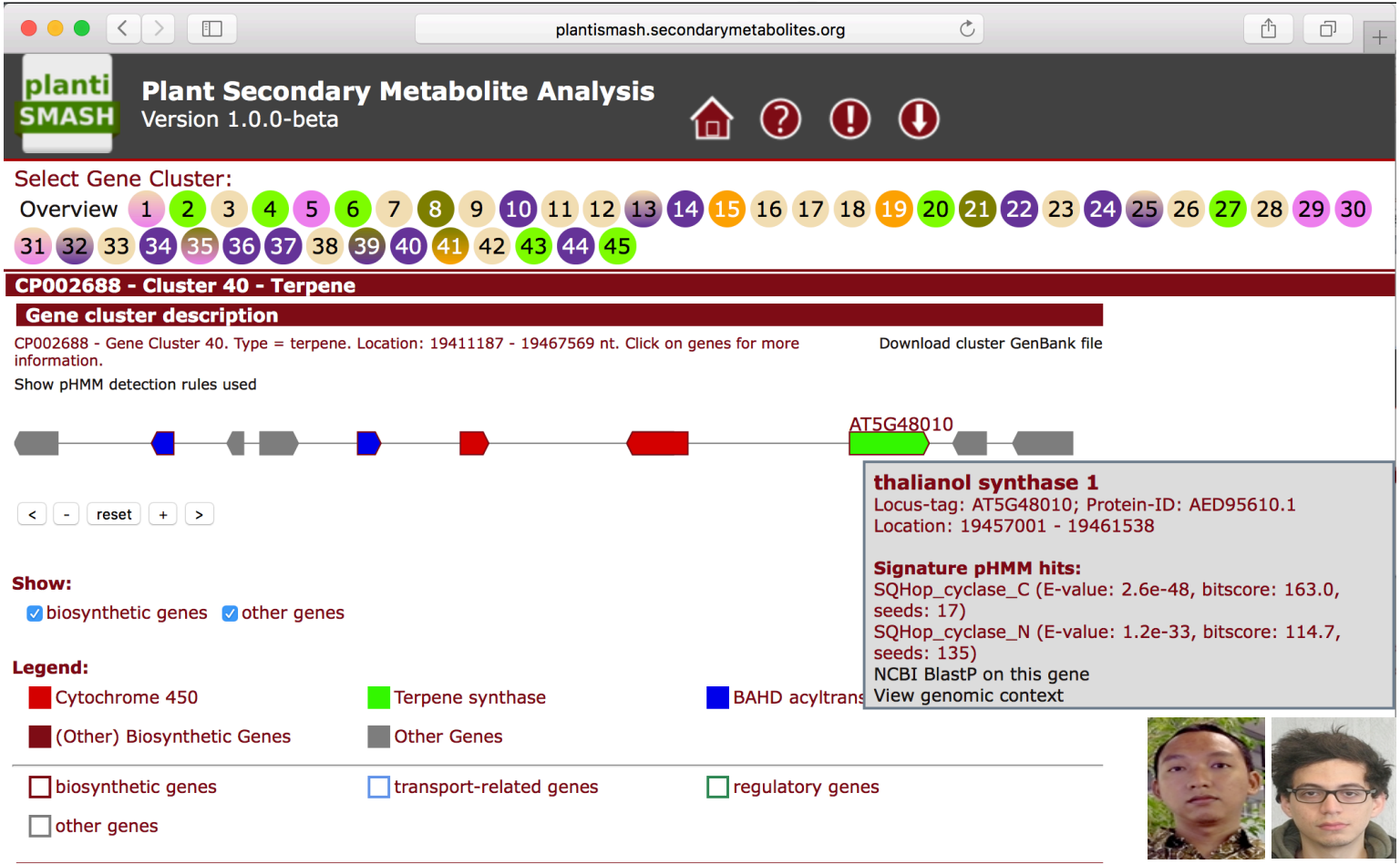

Medema & Osbourn, *Nat. Prod. Reports.* 33: 951-962 (2016)

# PLANT BIOSYNTHETIC PATHWAY DISCOVERY REQUIRES AN INTEGRATIVE APPROACH




Medema & Osbourn, *Nat. Prod. Reports.* 33: 951-962 (2016)

# PLANTISMASH: A SOFTWARE TOOL TO IDENTIFY BGCS IN PLANT GENOMES



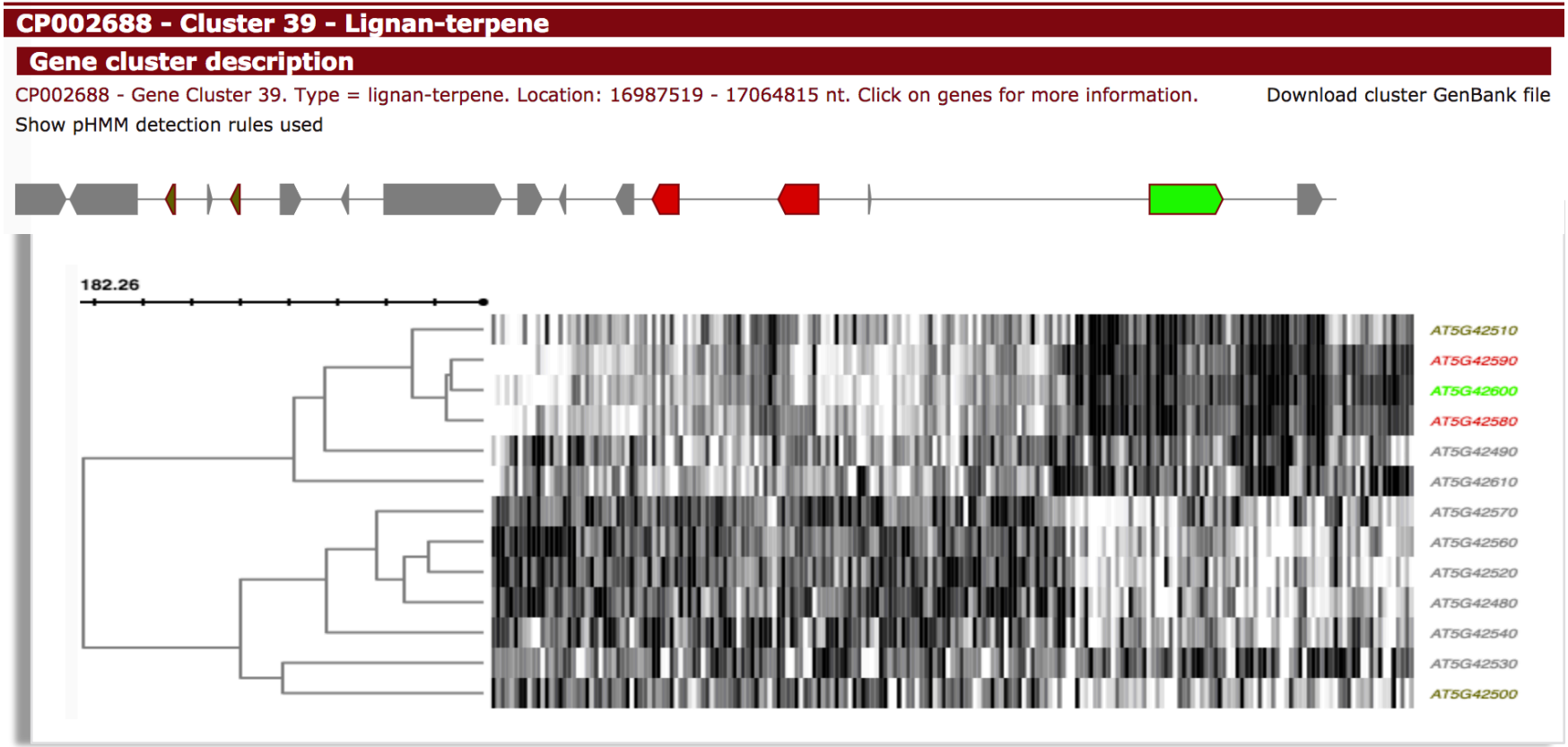





http://plantismash.secondarymetabolites.org

Kautsar, Suarez Duran et al. (2017) *Nucleic Acids Research* 45 (W1): W55–W63.

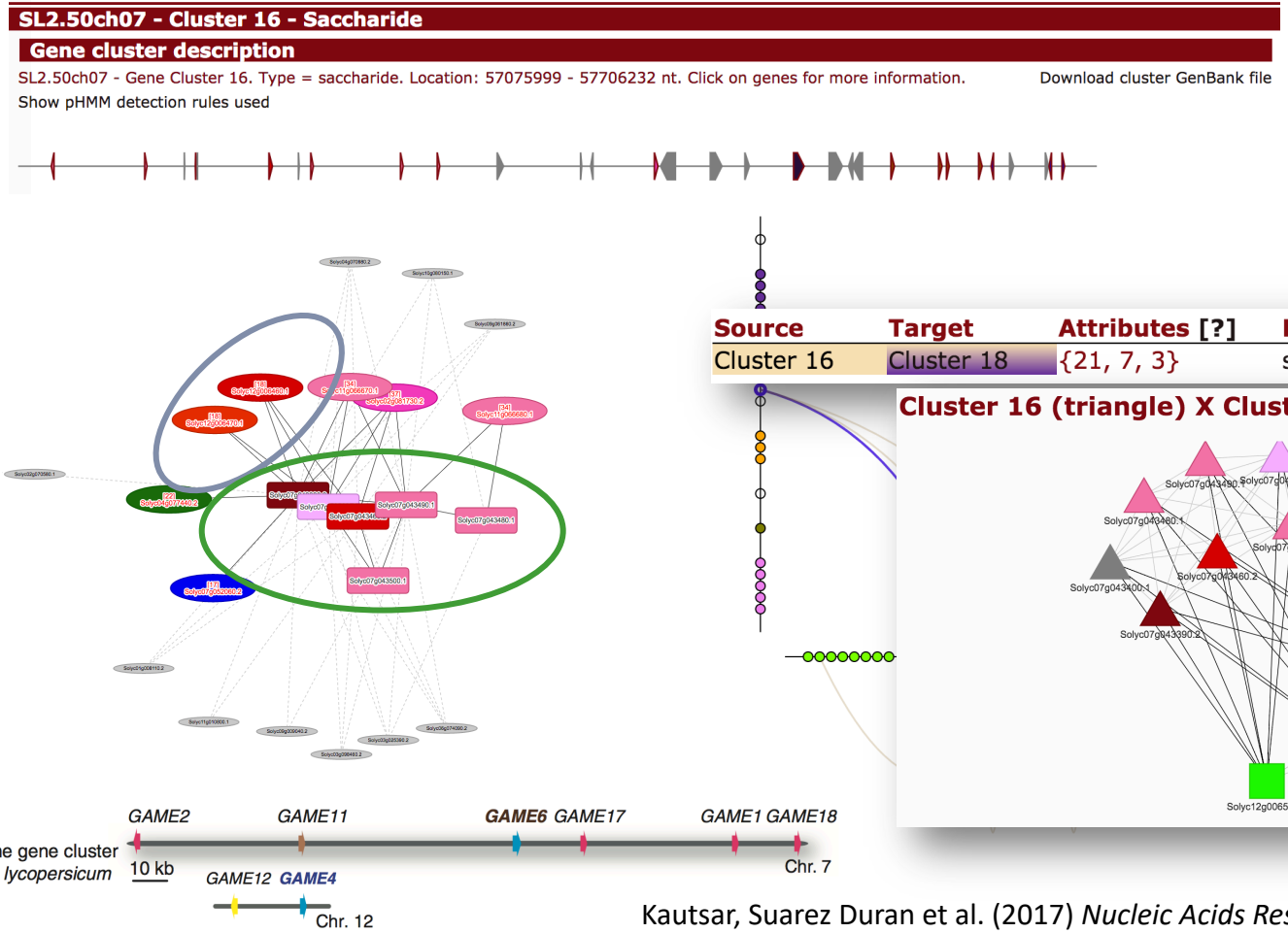
# EXPRESSION DATA CAN BE LOADED INTO PLANTISMASH TO IDENTIFY COEXPRESSION PATTERNS WITHIN CLUSTERS…




http://plantismash.secondarymetabolites.org

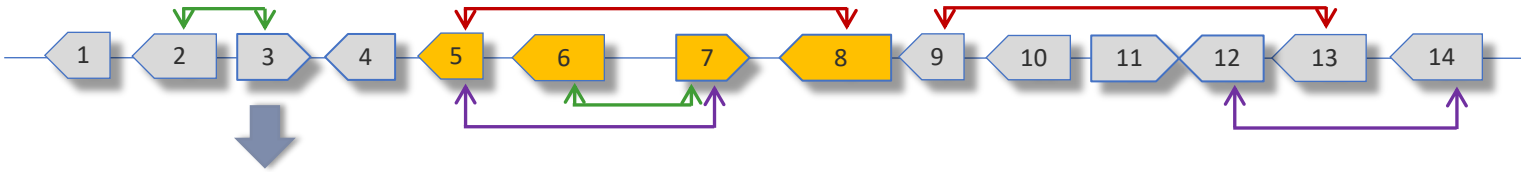
Kautsar, Suarez Duran et al. (2017) *Nucleic Acids Research* 45 (W1): W55–W63.

# … AND BETWEEN CLUSTERS




Kautsar, Suarez Duran et al. (2017) *Nucleic Acids Research* 45 (W1): W55–W63.

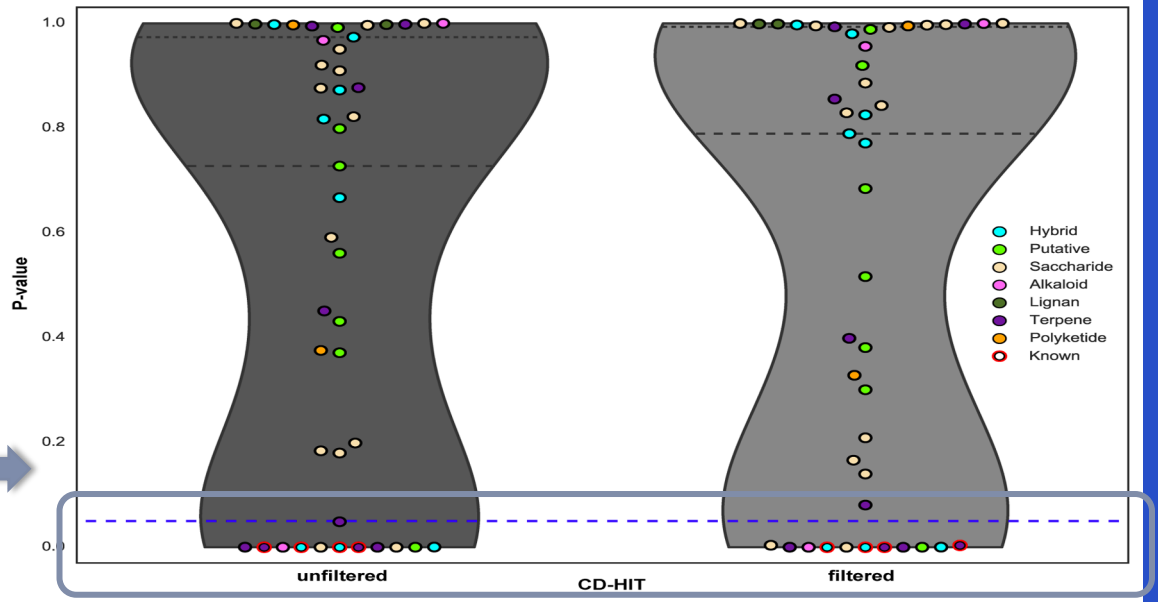
# STATISTICS INDICATE SIGNIFICANTLY COEXPRESSED GENE CLUSTERS: IN *ARABIDOPSIS*, ~25% OF ALL CANDIDATE BGCS ARE HIGHLY COEXPRESSED




$R_{(BGC)} = [R_{(5,8)}, R_{(5,7)}, R_{(6,8)}, R_{(5,6)}, R_{(6,7)}, R_{(7,8)}]$

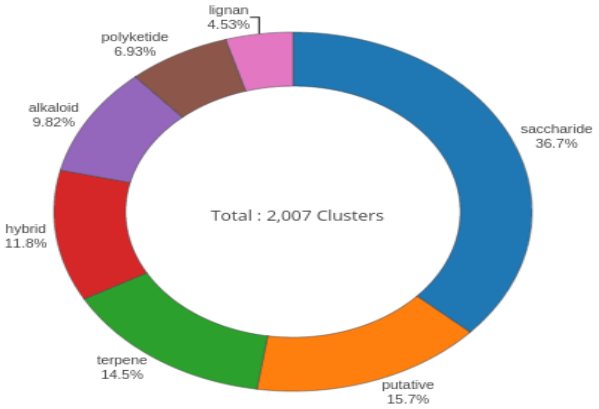
$R_{(Background)} = [R_{(1,4)}, R_{(9,12)}, R_{(10,13)}, R_{(11,14)}, R_{(1,3)}, R_{(2,4)}, R_{(9,11)}, R_{(10,12)}, R_{(11,13)}, R_{(12,14)}, R_{(1,2)}, R_{(2,3)}, R_{(3,4)}, R_{(9,10)}, R_{(10,11)}, R_{(11,12)}, R_{(12,13)}, R_{(13,14)}]$

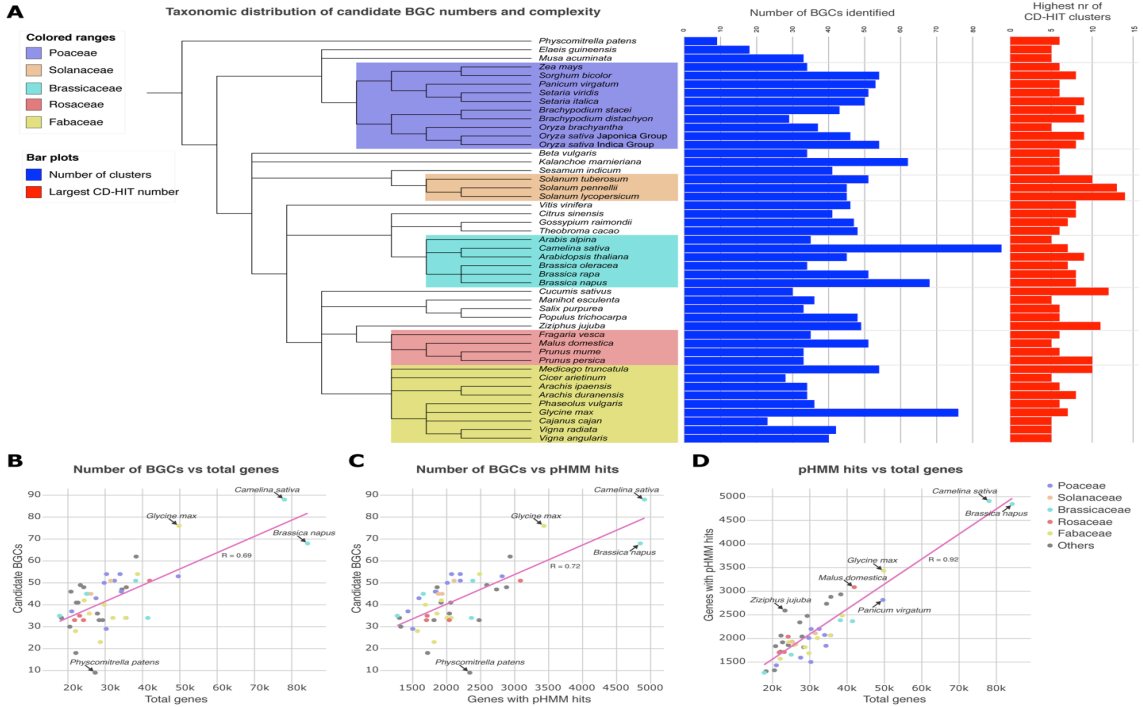
Wilcoxon rank sum (Mann-Whitney U) test




Kautsar, Suarez Duran et al. (2017) *Nucleic Acids Research* 45 (W1): W55–W63.

# GENE CLUSTER REPERTOIRES ARE DIVERSE THROUGHOUT THE PLANT KINGDOM




plantismash.secondarymetabolites.org/precalc

planti SMASH

Plant Secondary Metabolite Analysis

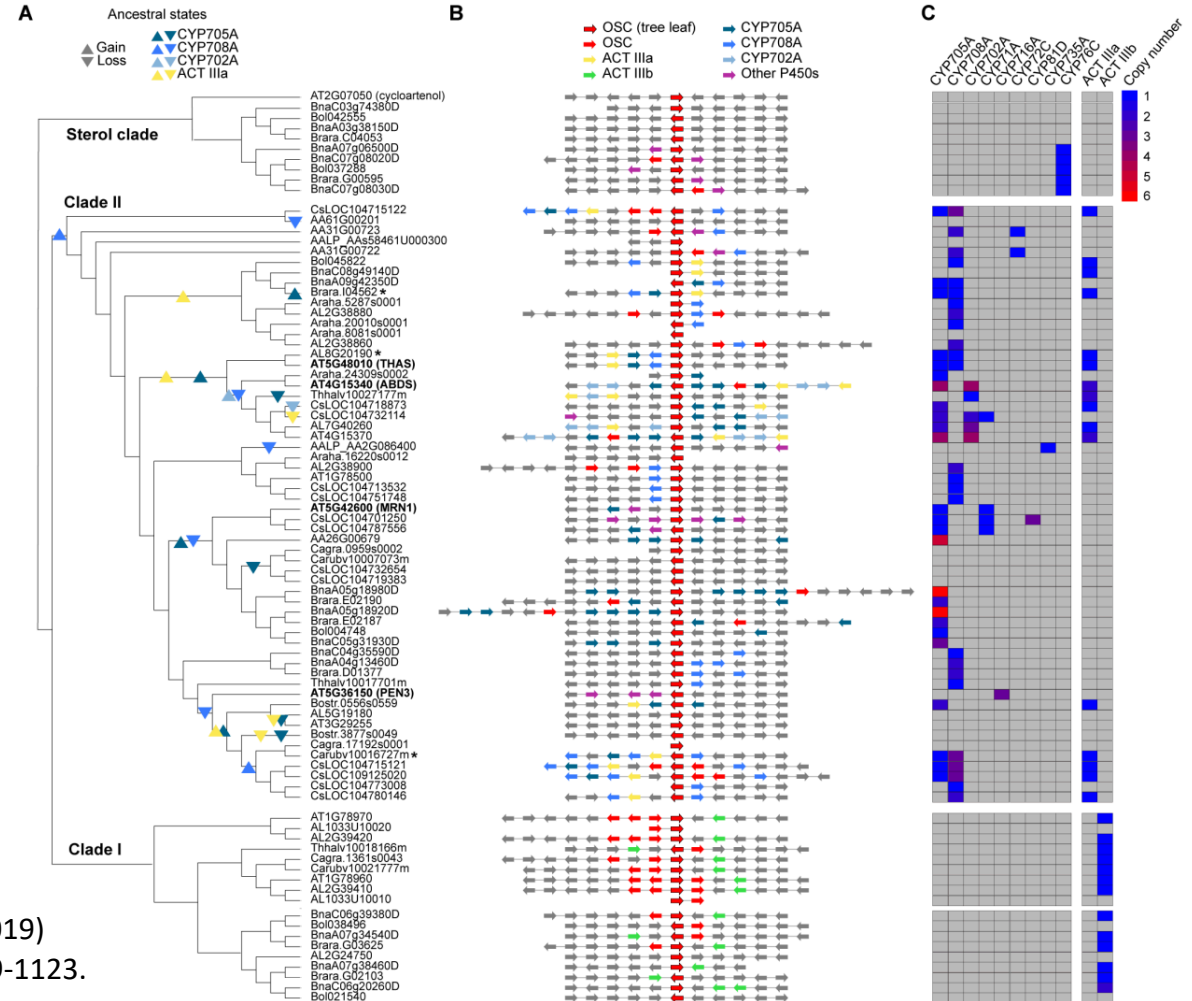
plantismash.secondarymetabolites.org

Kautsar, Suarez Duran et al. (2017) *Nucleic Acids Research* 45 (W1): W55–W63.

# PLANTISMASH IDENTIFIES MANY COMPLEX PLANT BIOSYNTHETIC GENE CLUSTERS


Kautsar, Suarez Duran et al. (2017) *Nucleic Acids Research* 45 (W1): W55–W63.

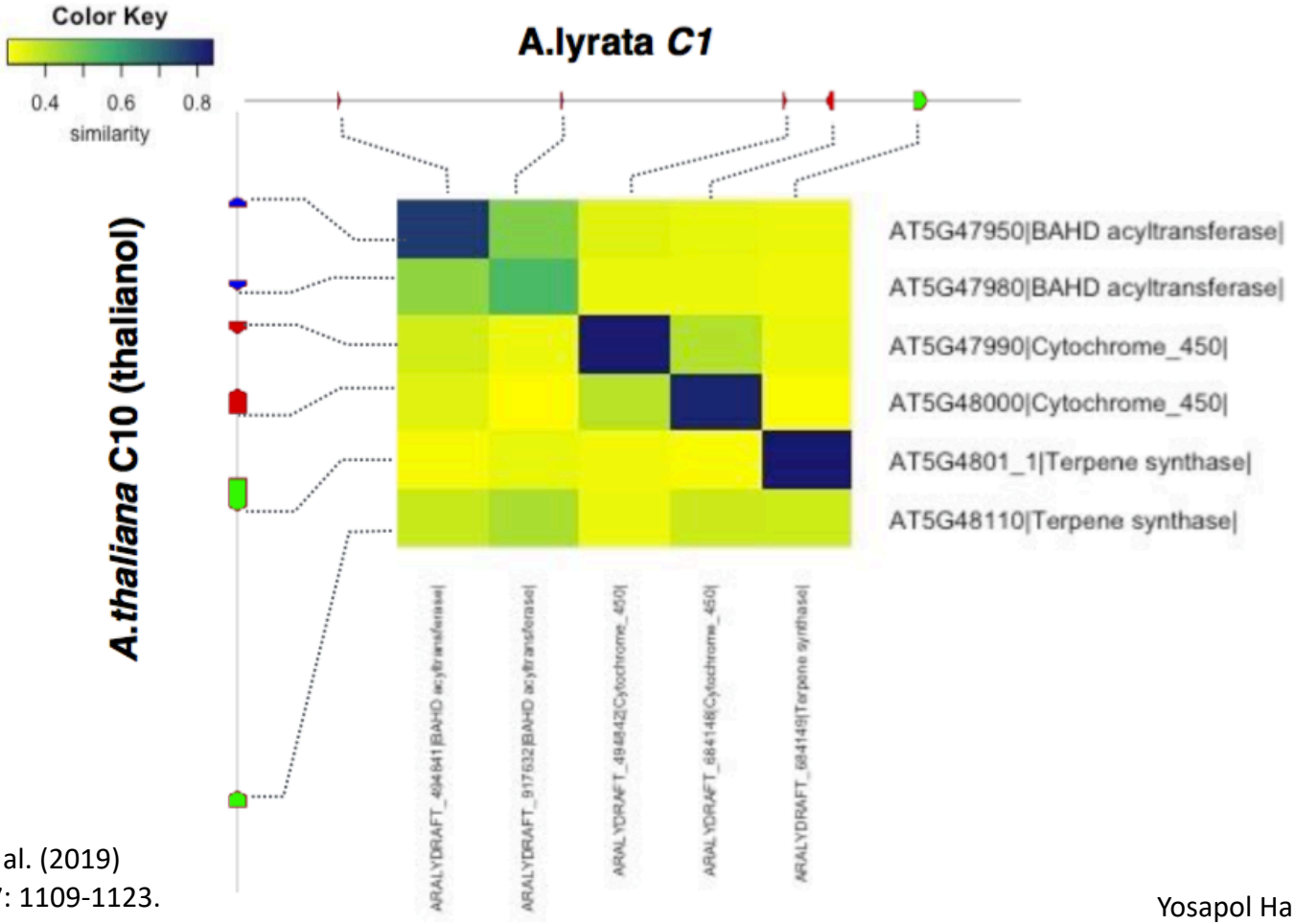
# PLANT TRITERPENE BIOSYNTHETIC LOCI ARE HIGHLY DYNAMIC




Liu, Suarez Duran et al. (2019)
*New Phytologist* 227: 1109-1123.

Hernando Suarez, Zhenhua Liu, Anne Osbourn

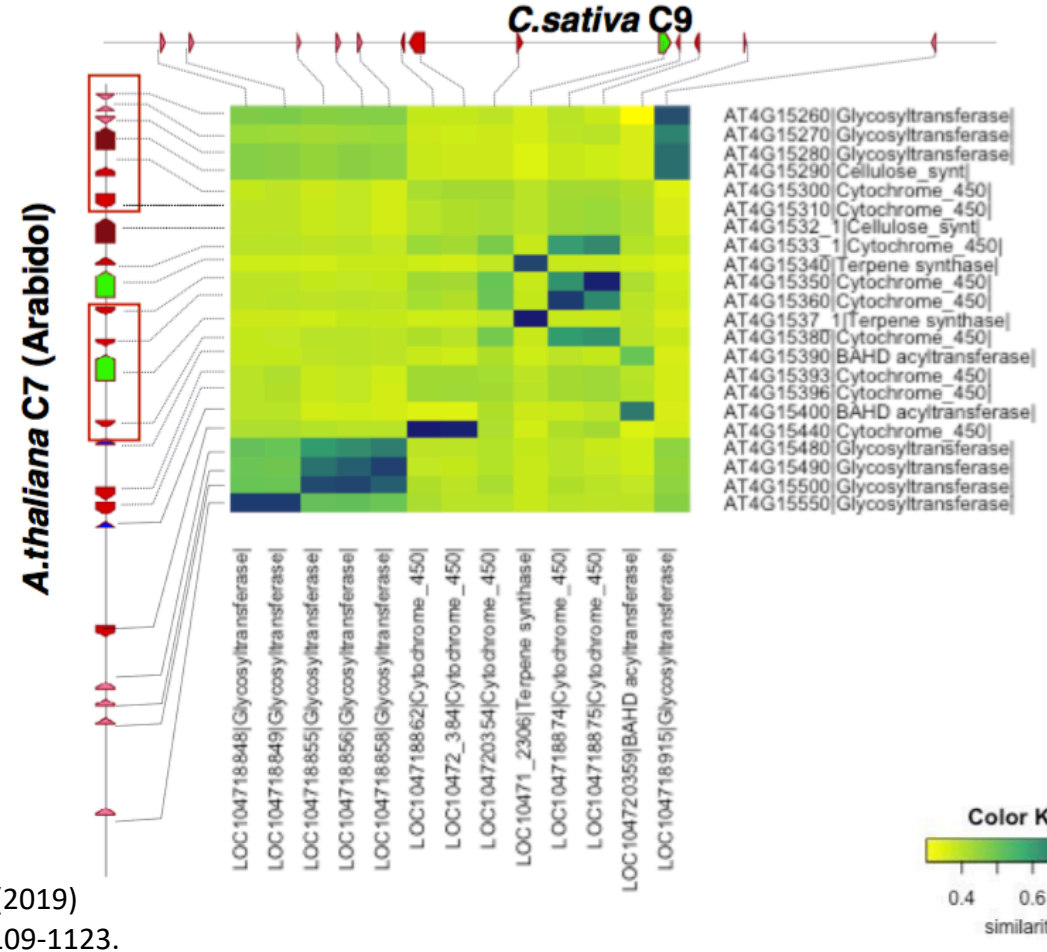
# WITHIN THE SAME PLANT GENUS, SOME GENE CLUSTERS ARE CONSERVED




Liu, Suarez Duran et al. (2019)
*New Phytologist* 227: 1109-1123.

Yosapol Harnvanichvech

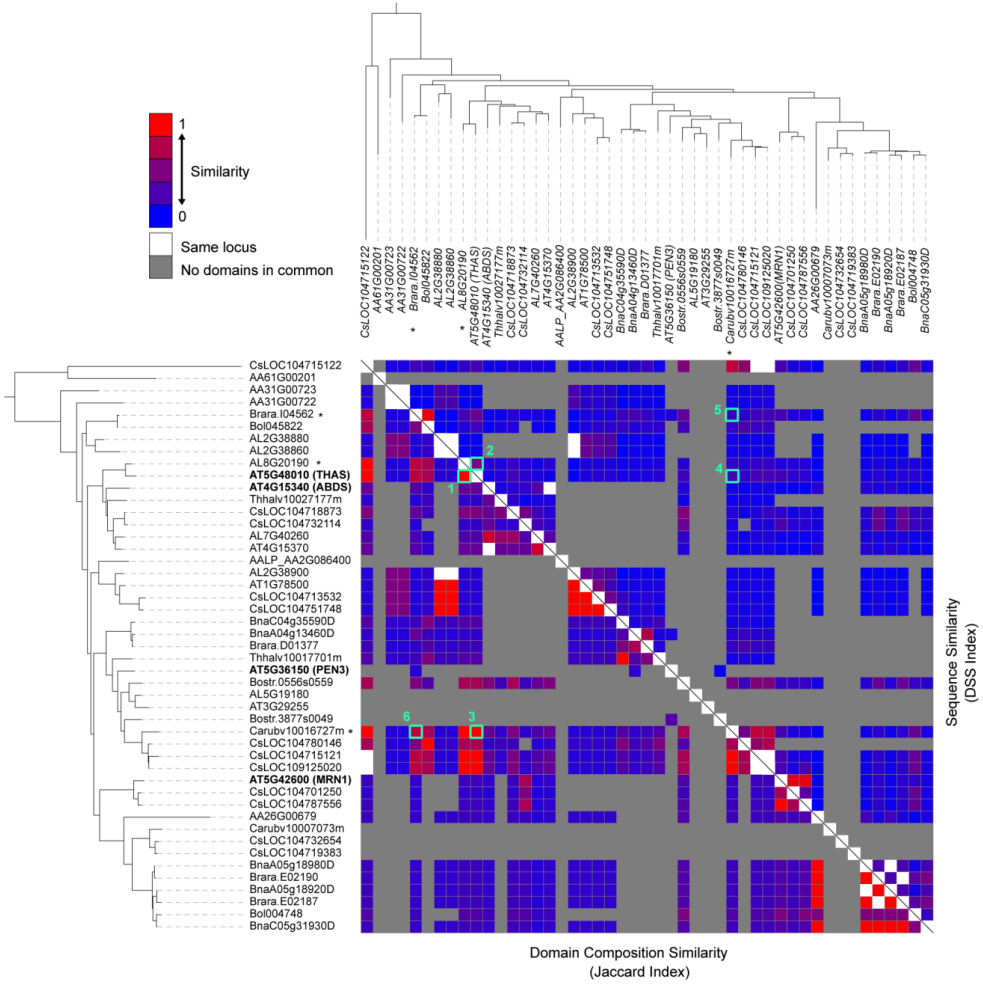
# ACROSS CLOSELY RELATED GENERA: LARGE DIFFERENCES!

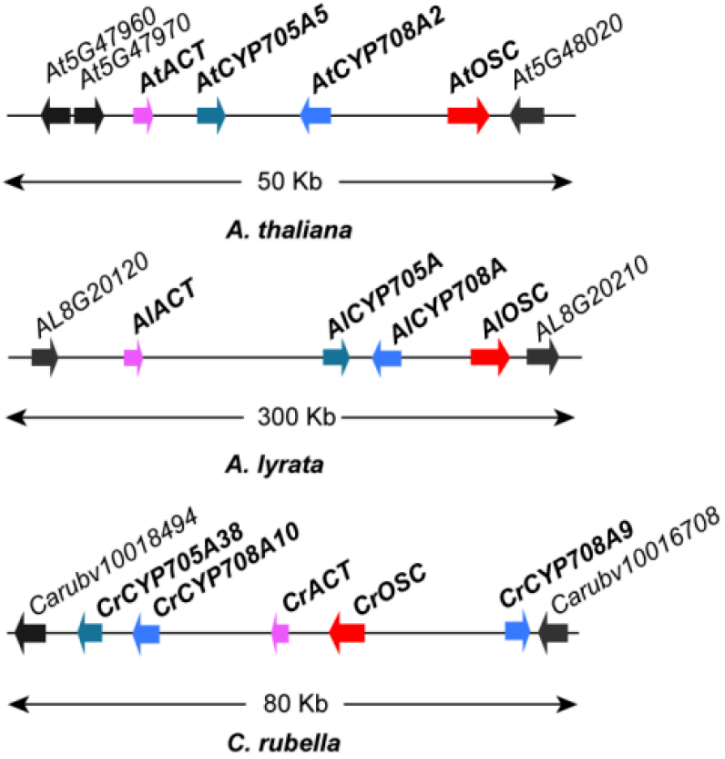



Liu, Suarez Duran et al. (2019)
*New Phytologist* 227: 1109-1123.

Yosapol Harnvanichvech

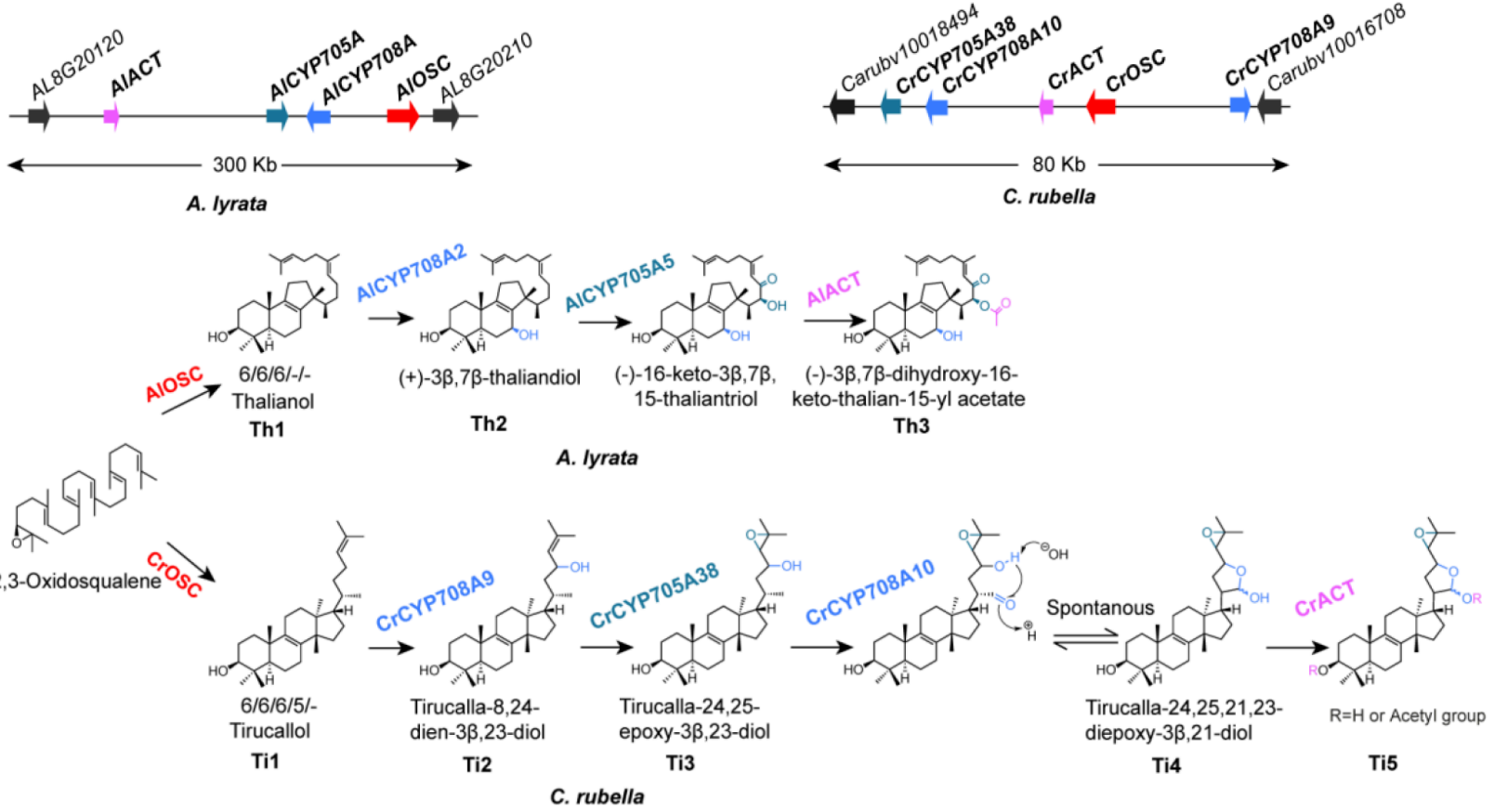
# PLANT TRITERPENE BIOSYNTHETIC LOCI ARE HIGHLY DYNAMIC



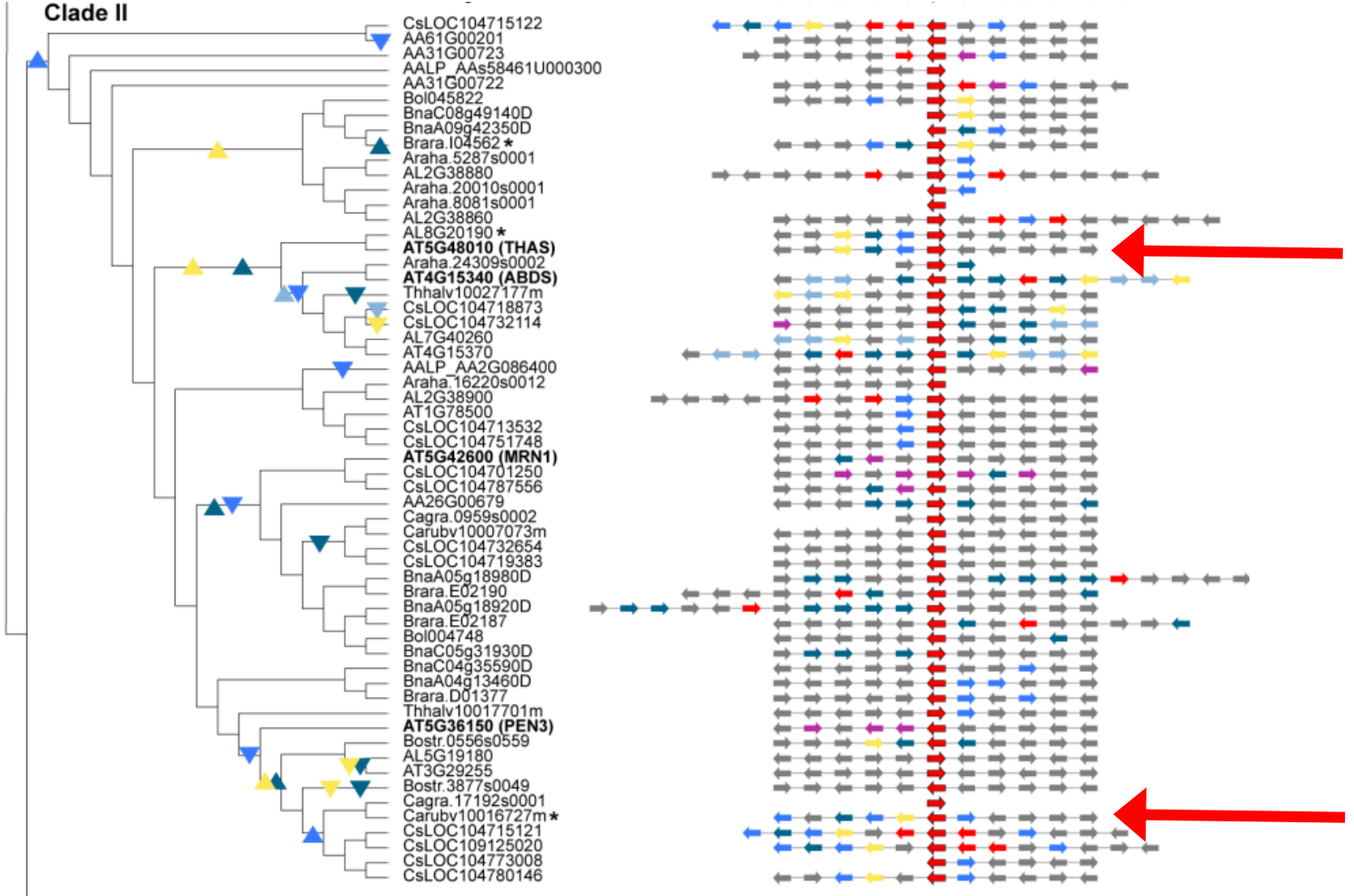

Liu, Suarez Duran et al. (2019)
*New Phytologist* 227: 1109-1123.

# GENE CLUSTERS WITH IDENTICAL ENZYME (SUB)FAMILY CONTENT MAY PRODUCE DIFFERENT CHEMISTRY



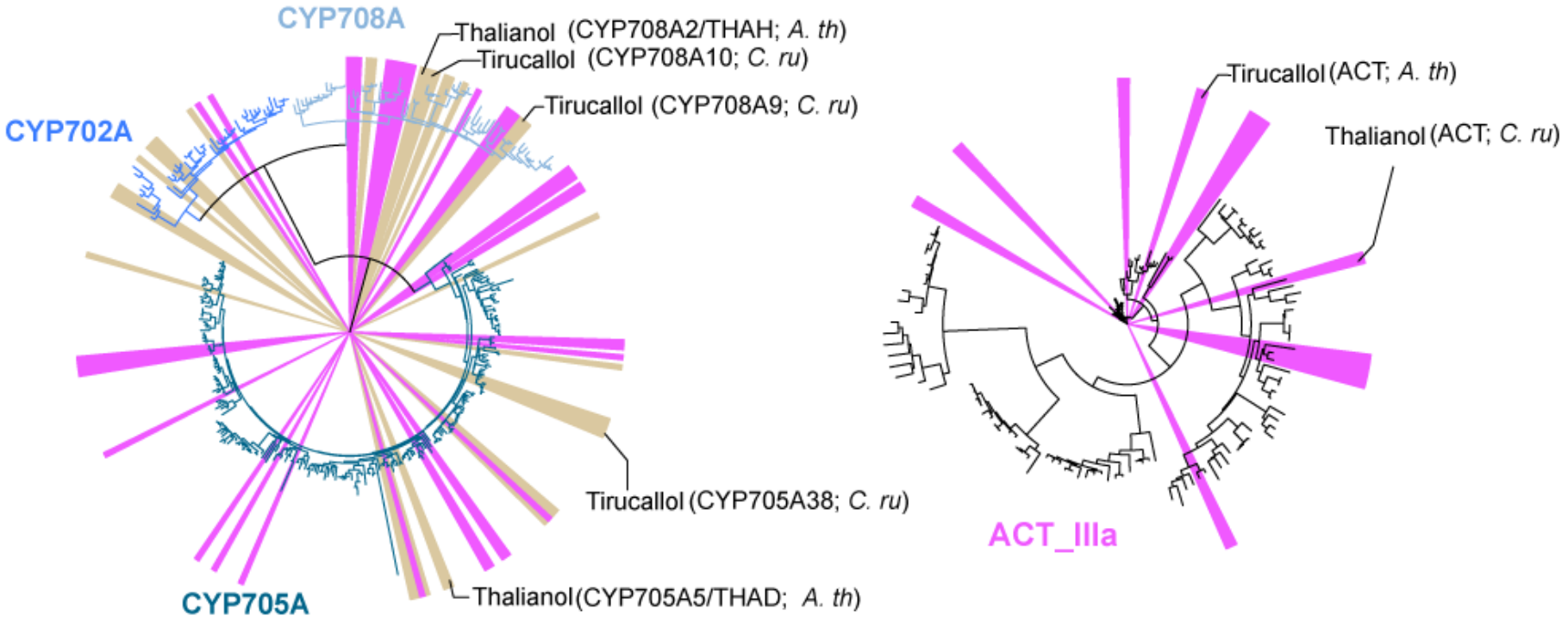

Zhenhua Liu, Anne Osbourn

# TRITERPENE SYNTHASE PHYLOGENY SUGGESTS POSSIBLE INDEPENDENT EVOLUTION



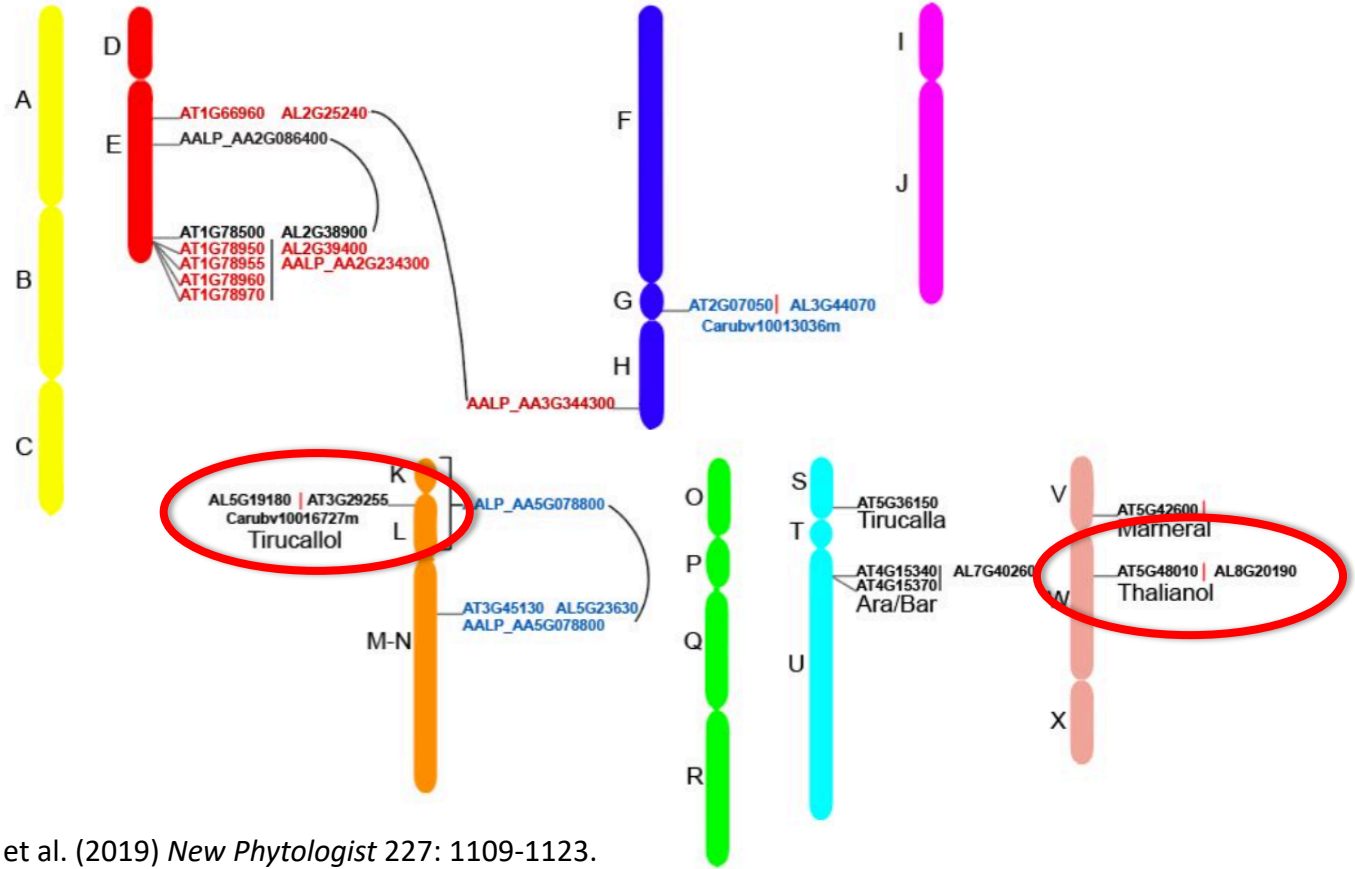

Liu, Suarez Duran et al. (2019) *New Phytologist* 227: 1109-1123.

# THE TIRUCALLOL AND THALIANOL BGCS CONTAIN P450S / ACYLTRANSFERASES FROM DIFFERENT CLADES OF THE SAME ENZYME SUBFAMILY



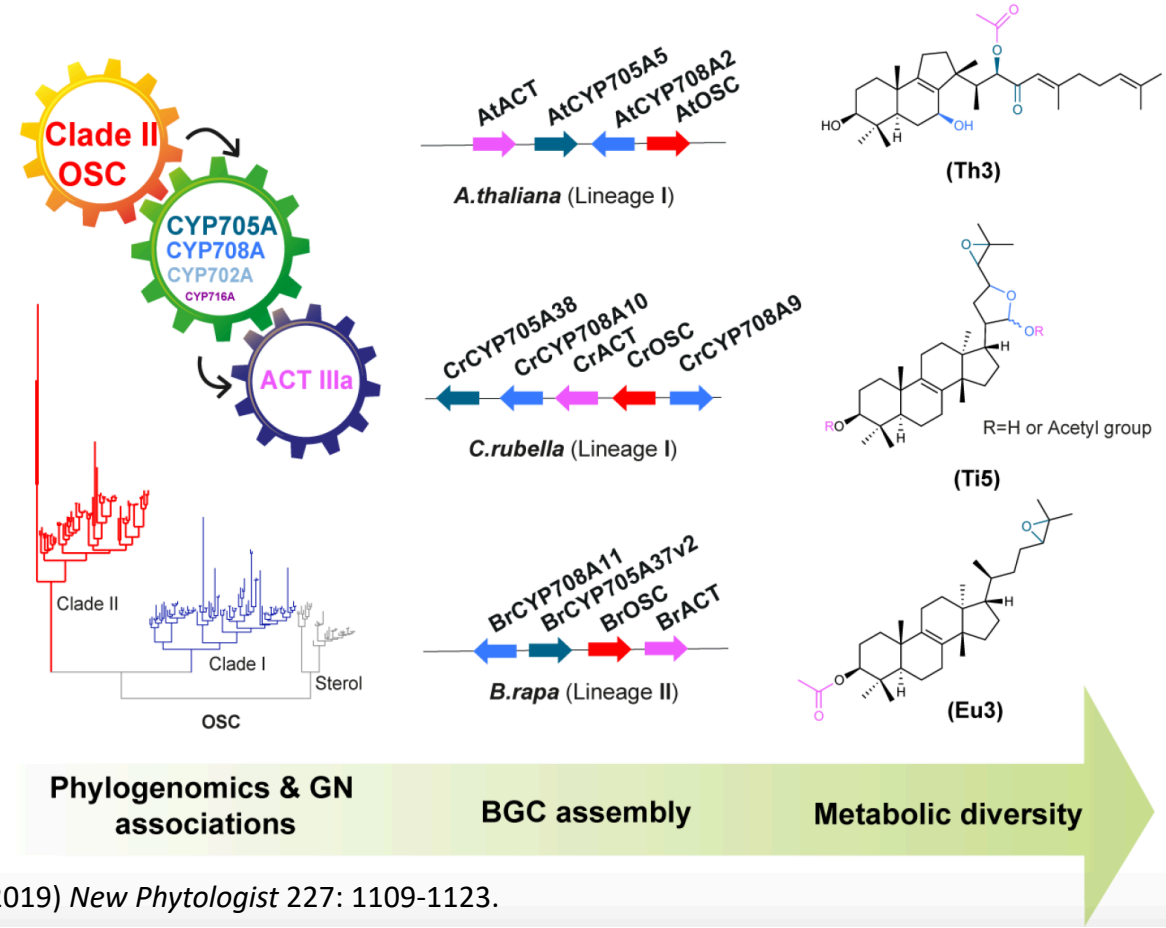

Liu, Suarez Duran et al. (2019) *New Phytologist* 227: 1109-1123.

# TRITERPENE SYNTHASE GENES MAP BACK TO DISTINCT CHROMOSOMAL LOCATIONS ON THE ANCESTRAL CRUCIPHER KARYOTYPE



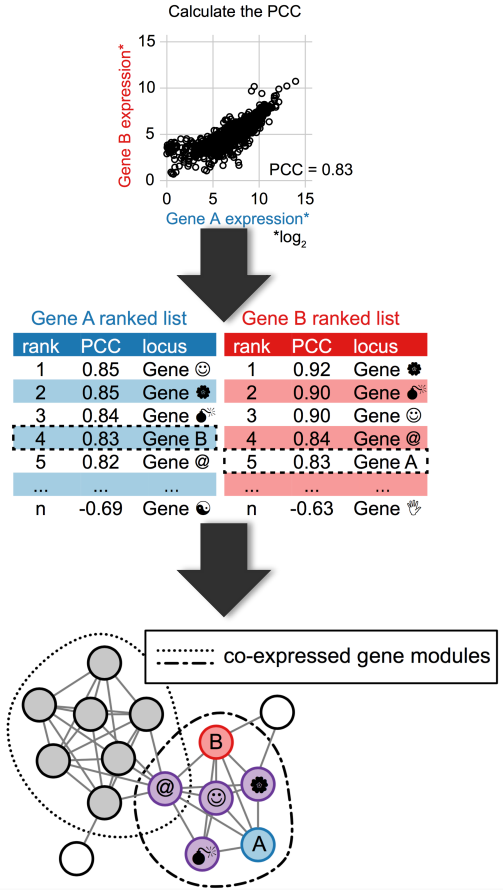

Liu, Suarez Duran et al. (2019) *New Phytologist* 227: 1109-1123.

# INDEPENDENT, DYNAMIC COMBINATORIAL EVOLUTION OF TRITERPENE DIVERSITY IN BRASSICACEAE: A MODEL FOR SYNTHETIC BIOLOGY

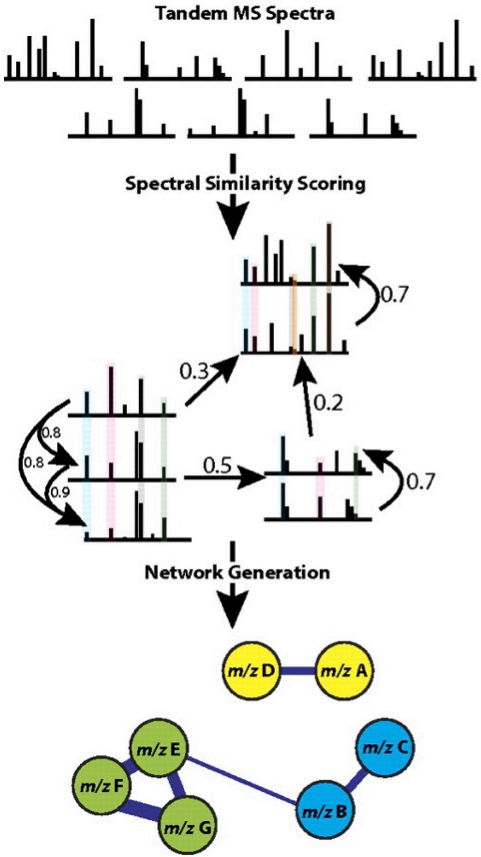



Liu, Suarez Duran et al. (2019) *New Phytologist* 227: 1109-1123.

# RECENT ADVANCES IN TRANSCRIPTOME/METABOLOME ANALYSIS PROVIDE FURTHER OPPORTUNITIES FOR UNSUPERVISED DISCOVERY

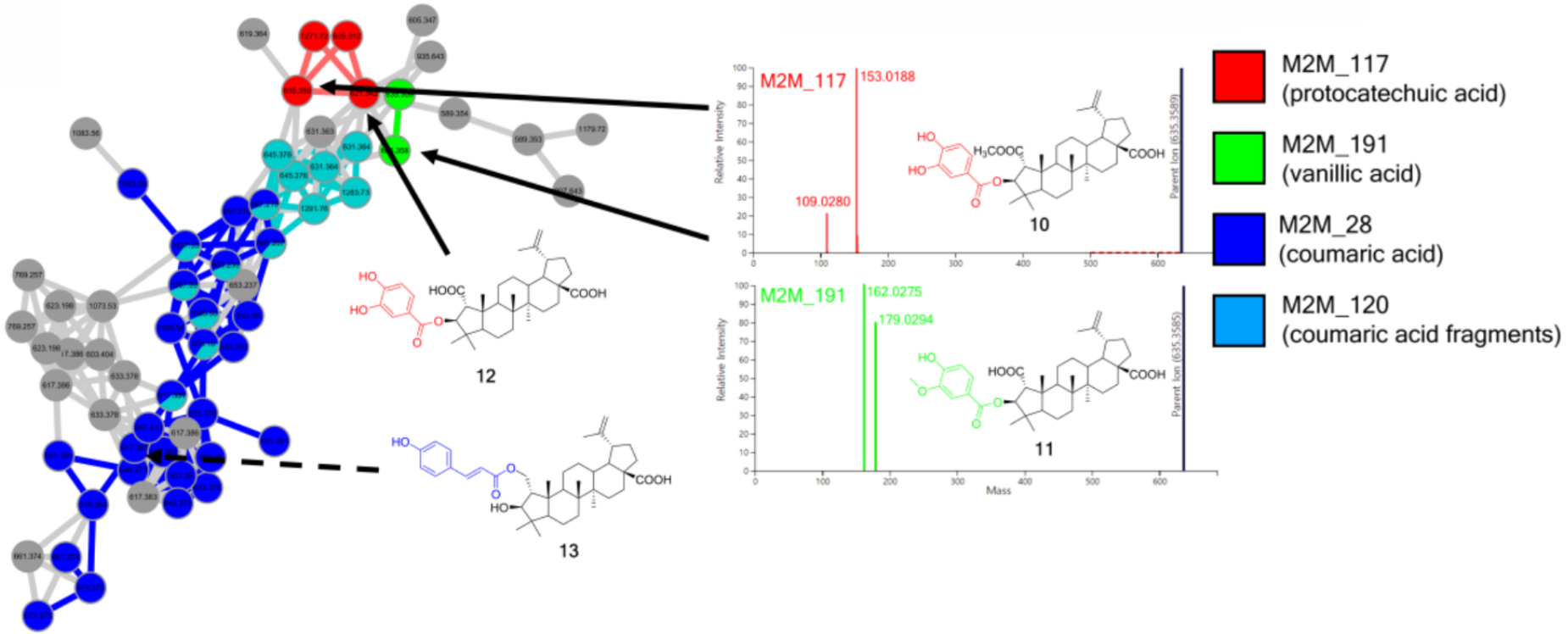



Wisecaver et al. (2017) *The Plant Cell.* 29 (5) 944-959.



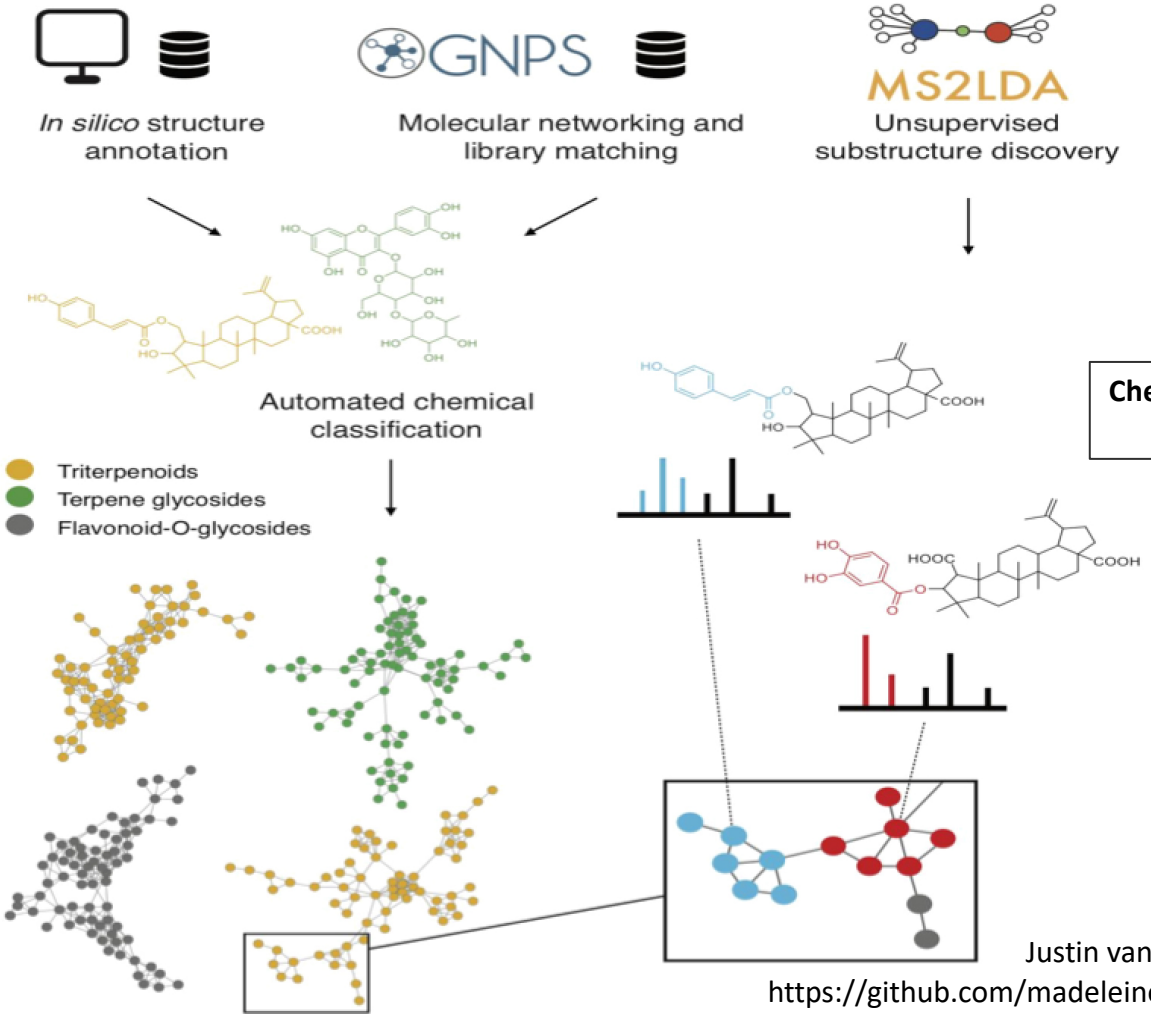

Watrous et al. (2012) *PNAS.* 109: E1743–E1752.

# MOLECULAR NETWORKS CAN BE ANNOTATED USING MOTIF FINDING AND OTHER ALGORITHMS FOR SUBSTRUCTURE PREDICTION




Kang, Ernst, van der Hooft (2018) *The Plant Journal*, doi:10.1111/tpj.14292

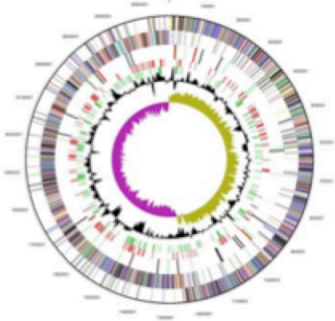
# THE MOLNETENHANCER WORKFLOW GENERATES ANNOTATED MOLECULAR NETWORKS




**Check out Hands-on Session 1 ( Justin van der Hooft)!**

Justin van der Hooft, Madeleine Ernst
https://github.com/madeleineernst/pyMolNetEnhancer

# THE PAIRED OMICS DATA PLATFORM CONNECTS GENOMICS WITH METABOLOMICS



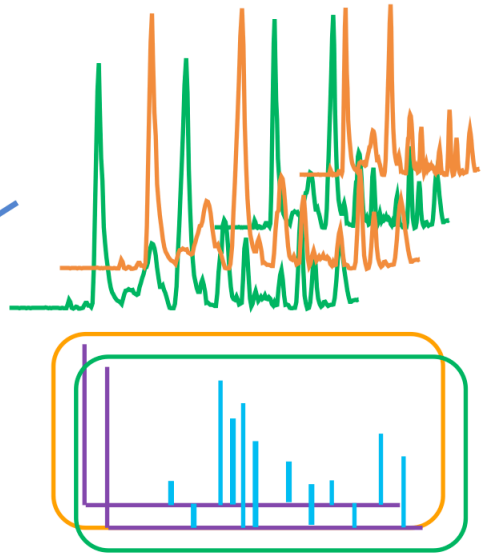
Dr. Justin van der Hooft

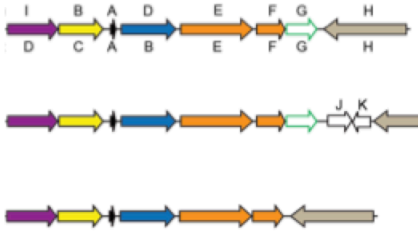
Dr. Michelle Schorn

Dr. Stefan Verhoeven

**Paired Data Platform**

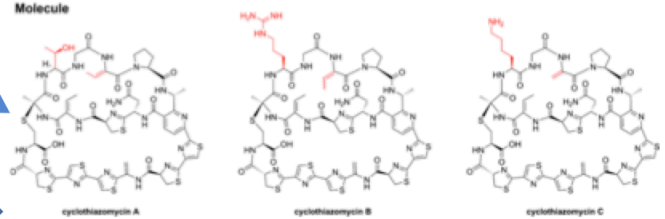
| Genome | MS/MS |
|---|---|
| Name | MASSive / |
| NCBI Acc. # | GNPS link |
| Taxonomy ID | Experimental |
| antiSMASH | Parameters |






Gene clusters
Gene cluster families

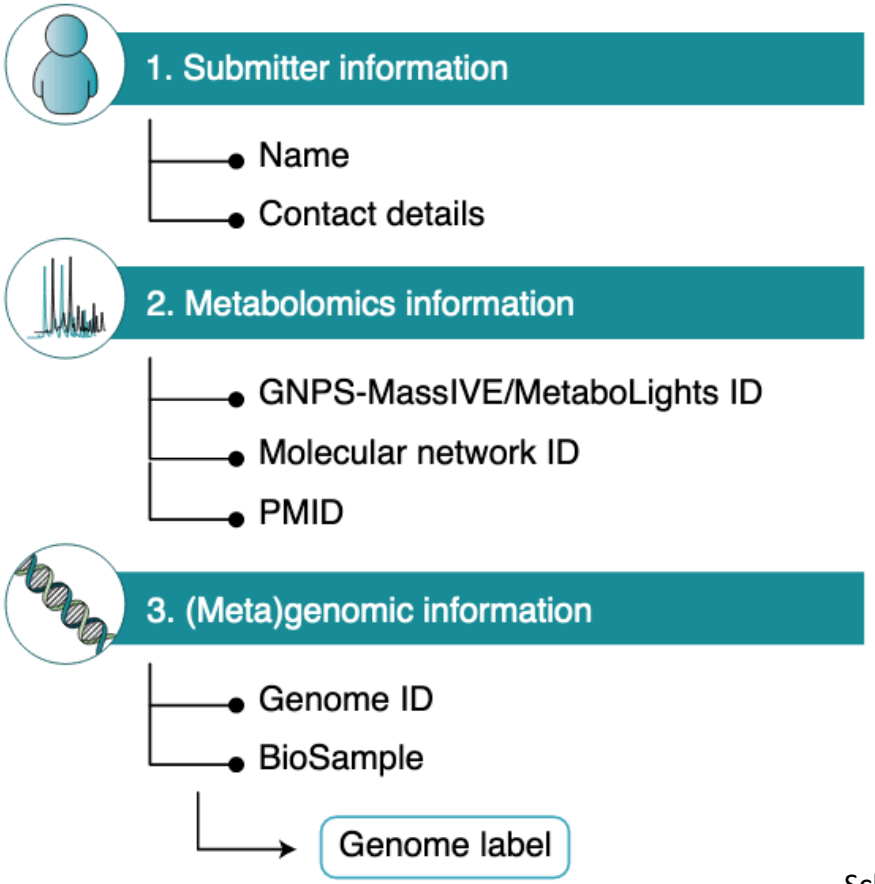
**LINK!**




Molecules
Molecular families

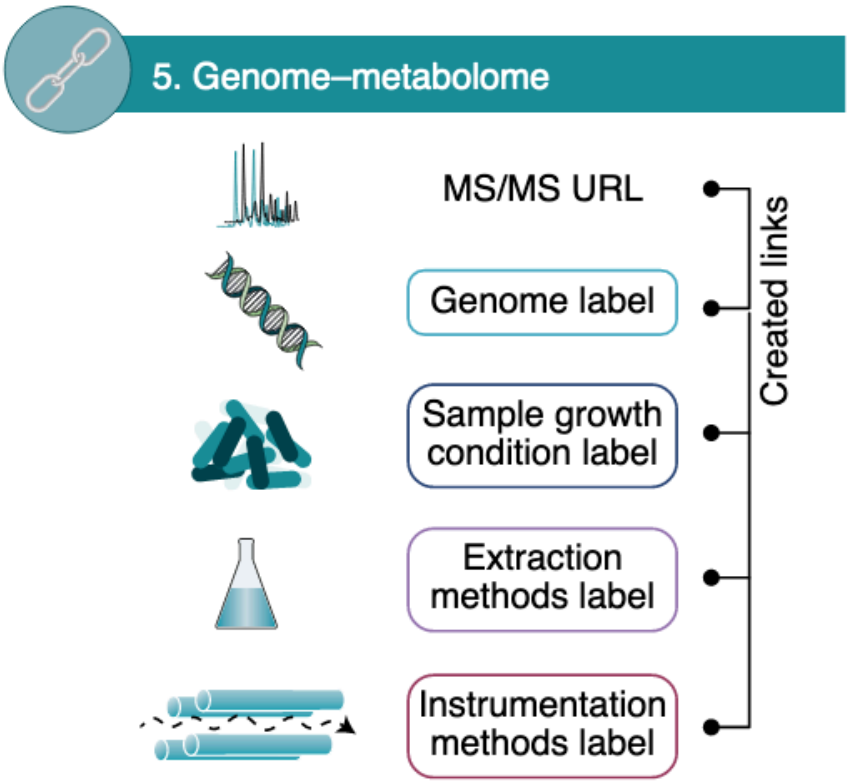
Schorn et al. (2021) *Nature Chemical Biology,* doi:10.1038/s41589-020-00724-z

# THE PAIRED OMICS DATA PLATFORM CONNECTS GENOMICS WITH METABOLOMICS




Schorn et al. (2021) *Nature Chemical Biology,* doi:10.1038/s41589-020-00724-z

# THE PAIRED OMICS DATA PLATFORM CONNECTS GENOMICS WITH METABOLOMICS




Schorn et al. (2021) *Nature Chemical Biology,* doi:10.1038/s41589-020-00724-z

# THE MEDEMA RESEARCH GROUP @ WAGENINGEN UR

*Principal Investigator:*
- Marnix Medema

*Postdocs & PhD students:*
- **Hernando Suarez Duran** – *until Dec 2019*
- **Satria Kautsar** – *until Dec 2020*
- **Michelle Schorn** – *until Feb 2019*
- Vittorio Tracanna
- Victòria Pascal Andreu
- Barbara Terlouw
- Lotte Pronk
- Hannah Augustijn
- Zachary Reitz
- Mohammad Alanjary
- Huali Xie

*WU Bioinformatics colleague PI:*
- **Justin van der Hooft**

*MSc students:*
Jeanine Boot, Bram van Wersch, Liana van Grieken, Matthias van den Belt, Thijs Finnegan


Part of the Wageningen UR Bioinformatics Department, headed by Prof. Dick de Ridder

Funding:
NWO Netherlands Organisation for Scientific Research
ZonMw
WAGENINGEN UR For quality of life
erc European Research Council Established by the European Commission
GNPN
The Graduate School EXPERIMENT PLANT SCIENCES

# MANY THANKS TO COLLABORATORS

John Innes Centre
Unlocking Nature's Diversity

- Anne Osbourn
- Zhenhua Liu
- Charlie Owen

UC San Diego

- Pieter Dorrestein
- Kyo Bin Kang

DTU Technical University of Denmark

- Tilmann Weber
- Kai Blin
- Simon Shaw

netherlands eScience center

- Lars Ridder
- Florian Hüber
- Stefan van der Hoeven