

# Phylogenomic Mining of the Mints Reveals Multiple Mechanisms Contributing to the Evolution of Chemical Diversity in Lamiaceae

Mint Evolutionary Genomics Consortium\*

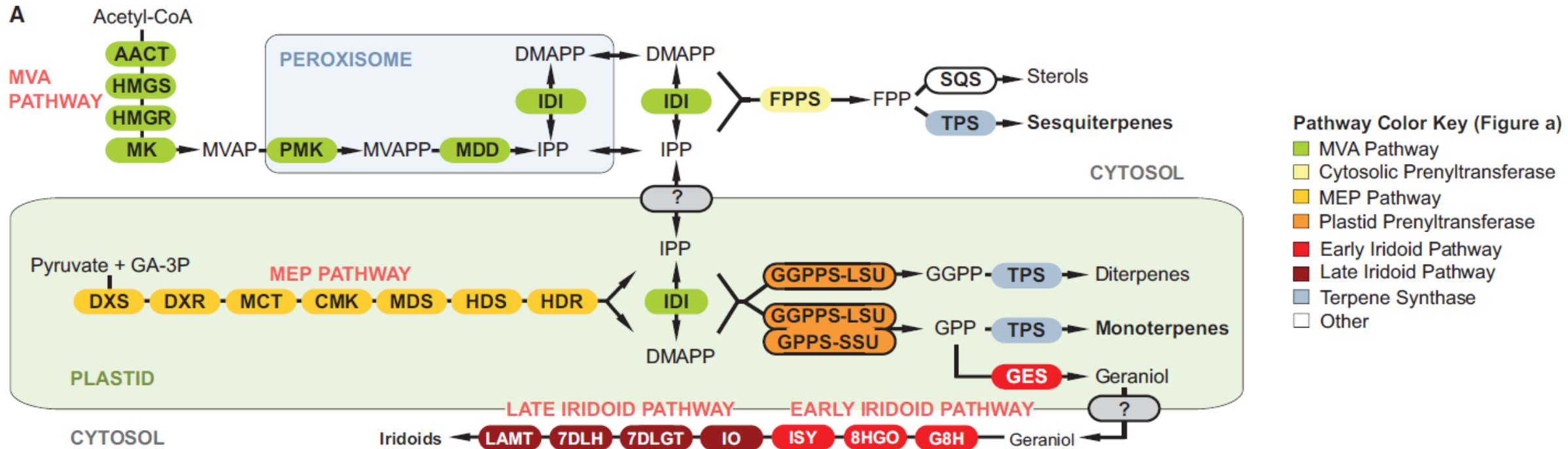
\*Correspondence: [buell@msu.edu](mailto:buell@msu.edu)

<https://doi.org/10.1016/j.molp.2018.06.002>

Published August 2018

# Background

- Chemical diversity is a key driver in evolutionary diversification
- Specialized metabolites are made from a small group of precursors
  - Ex: terpenes are precursors for monoterpenes, sesquiterpenes, and iridoids
- Lamiaceae is a family rich in terpenoids
  - For plants – anti-microbial, anti-herbivory, and pollinator attraction
  - For us – health, food flavoring, and agricultural purposes
    - Menthol – peppermint
    - Nepetalactone – catnip
    - Agnuside – progesterone normalization
    - Caryphyllene – food aromatic (pepper)



## Key points:

- Monoterpenes and iridoids are both derived from GPP
- GPP is produced by the MEP pathway
- Sesquiterpenes are produced by the MVA pathway
- MVA and MEP pathways compete for IPP

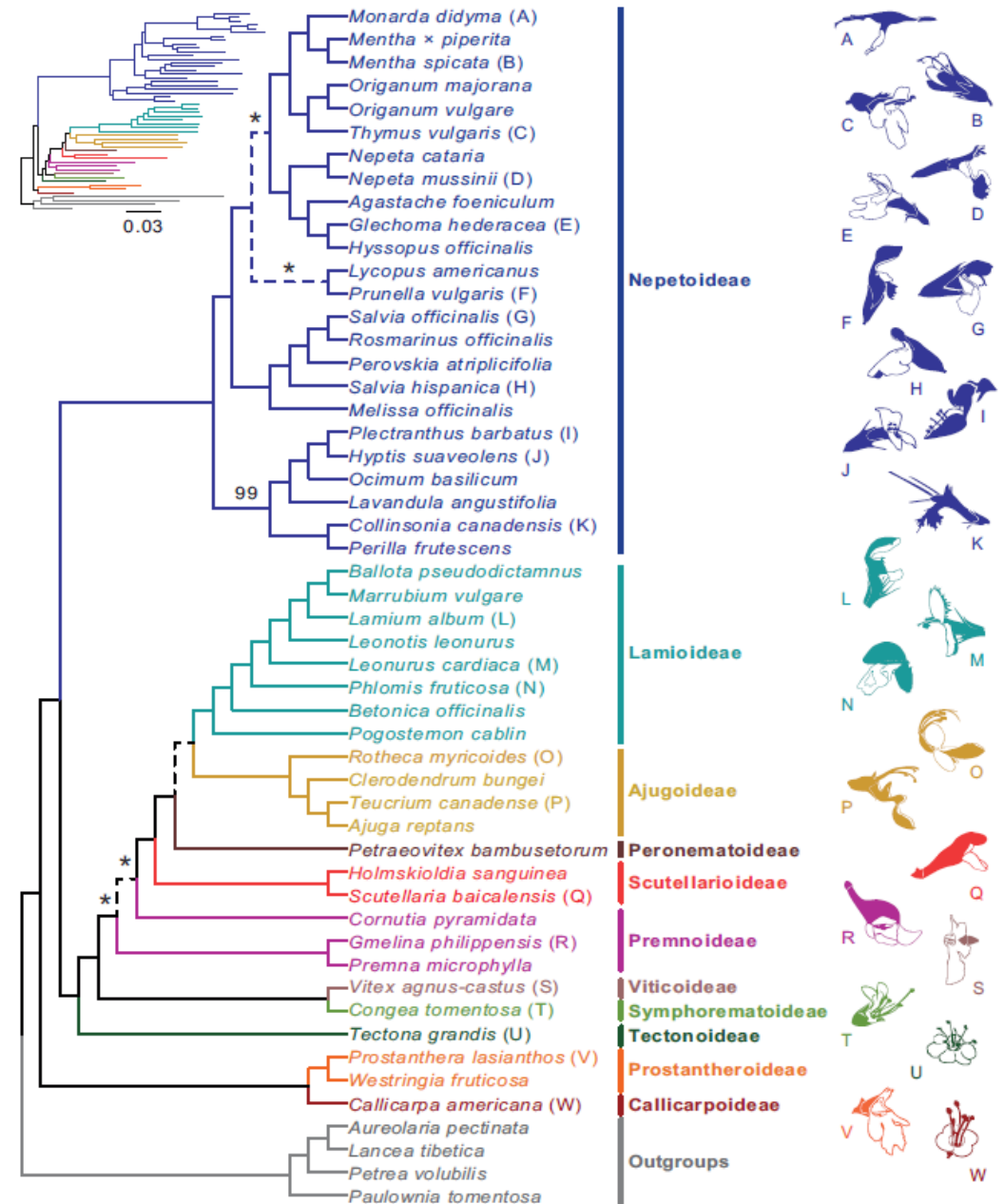
# Background

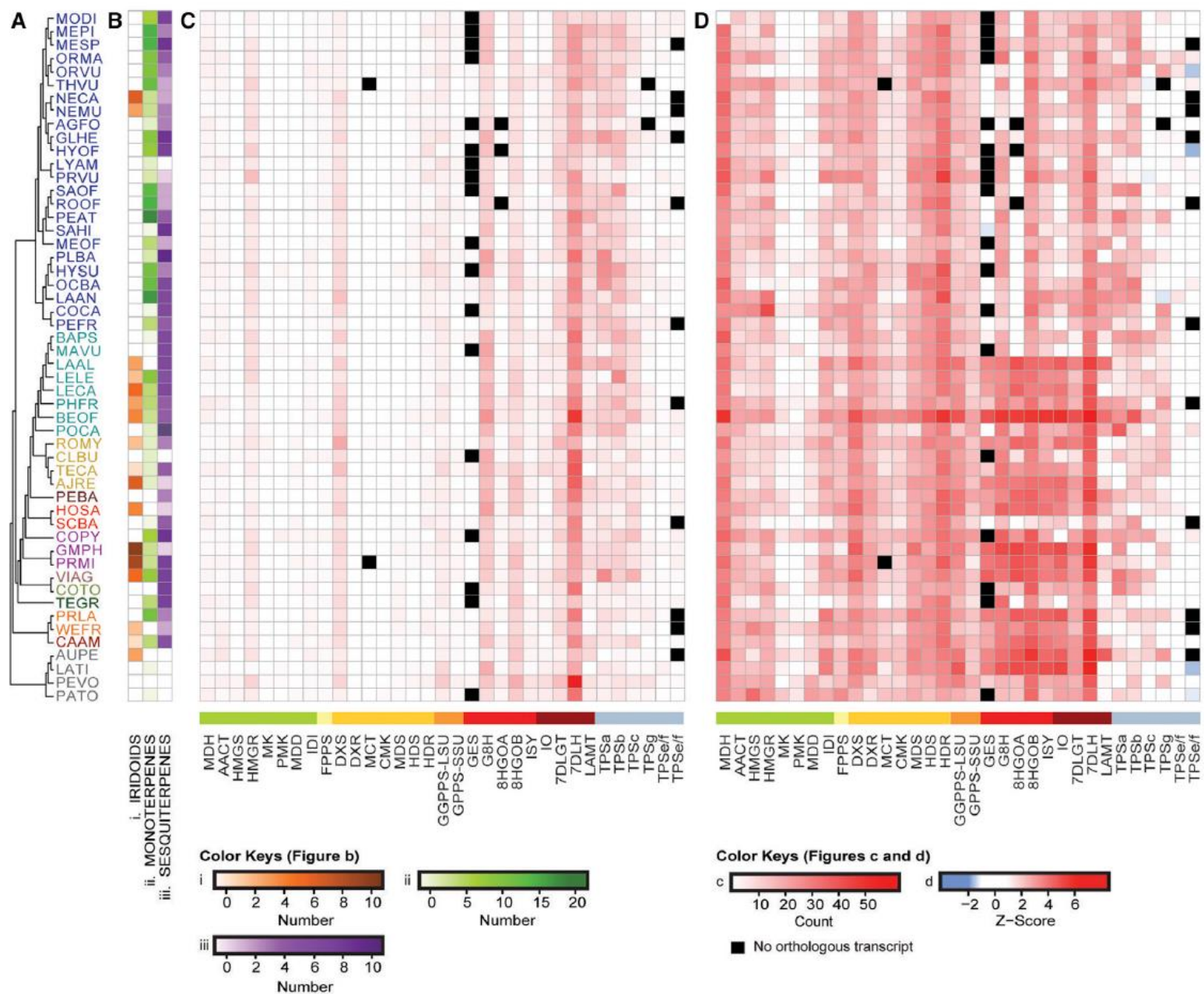
- Extensive flavonoid diversity exists in Lamiaceae
  - GM-MS and LC-MS identified 64 iridoids, 44 monoterpenes, and 39 sesquiterpenes across all species investigated
- However, evolutionary mechanisms (i.e. selective pressures) driving diversity are unknown

**Goal** – analyze transcriptomes and metabolomes of 48 phylogenetically diverse species of Lamiaceae to reveal evolutionary mechanics underlying chemodiversity

# Transcriptome-based phylogeny

- RNA extracted from young leaves of 48 species and 4 outgroups
- Transcriptome assembled for each species
- 520 single-copy nuclear genes identified from angiosperm database to be present in transcriptome of all 52 species
  - Maximum-likelihood tree constructed based on sequences of these 520 genes

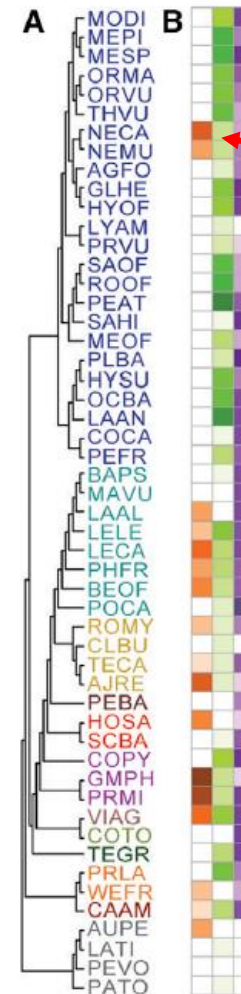






# Metabolite profiling

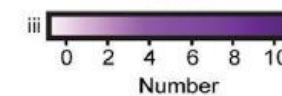
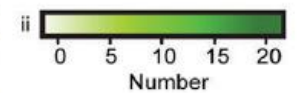
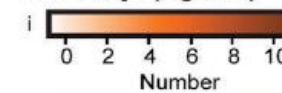
- Untargeted metabolite profiling of young leaf tissue
- GC-MS used to identify monoterpenes, sesquiterpenes, and 2 volatile iridoids
- LC-MS used to identify iridoids
- Data is “binary” – 2 possibilities for each metabolite
  - Detected
  - Not detected



Iridoid production is lost in Nepetoidae, but restored in *Nepota* family

i. IRIDIDS  
ii. MONOTERPENES  
iii. SESQUITERPENES

Color Keys (Figure b)

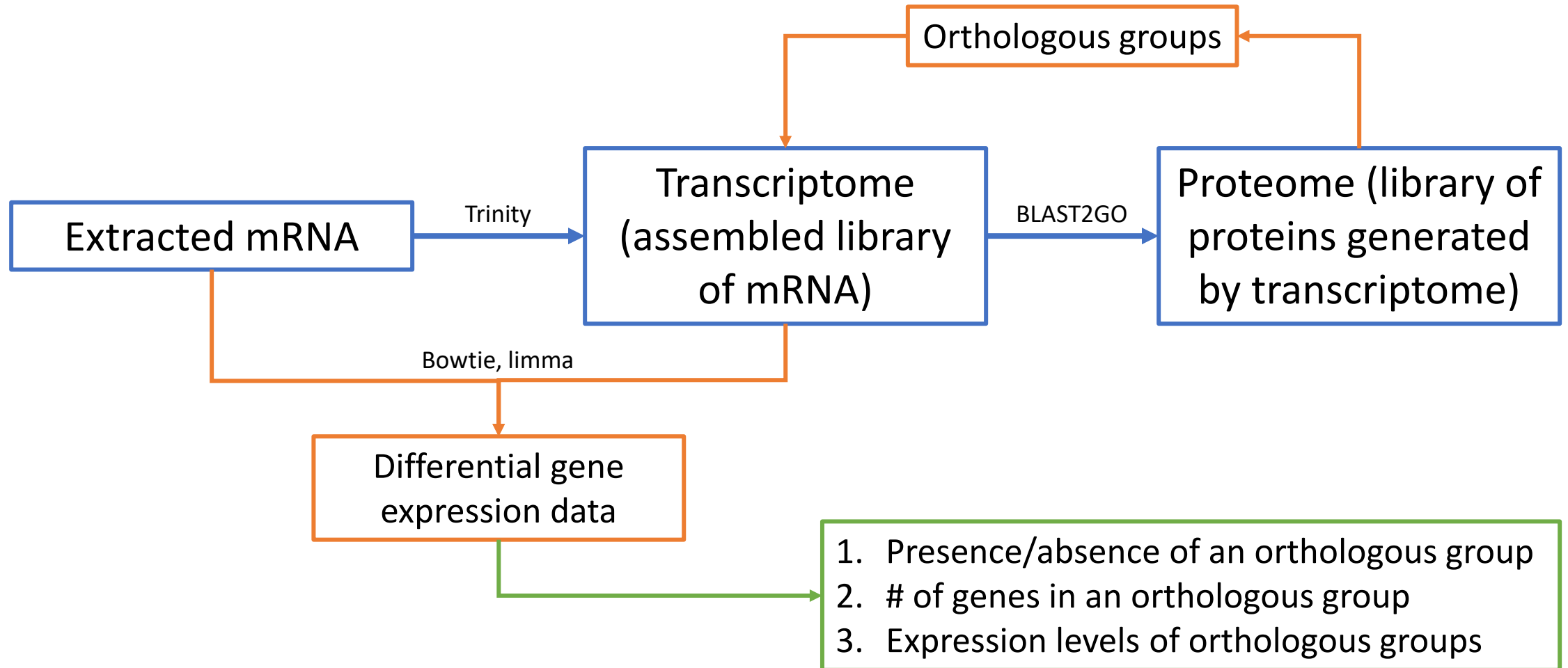


# Orthologous grouping of proteins

- Previously constructed transcriptomes used to predict proteomes for each species
- Proteins predicted from proteomes clustered into orthologous groups
  - Orthologous – from common ancestor (i.e. a family/class of proteins with similar function)
- Analysis of genes/protein groups involved in TCA cycle used to validate transcriptional representation
  - Detected a similar number of genes with little variation in copy number
  - Detected genes had similar levels of expression (Z-scores)
  - Conclusion – read depth is sufficient for analysis

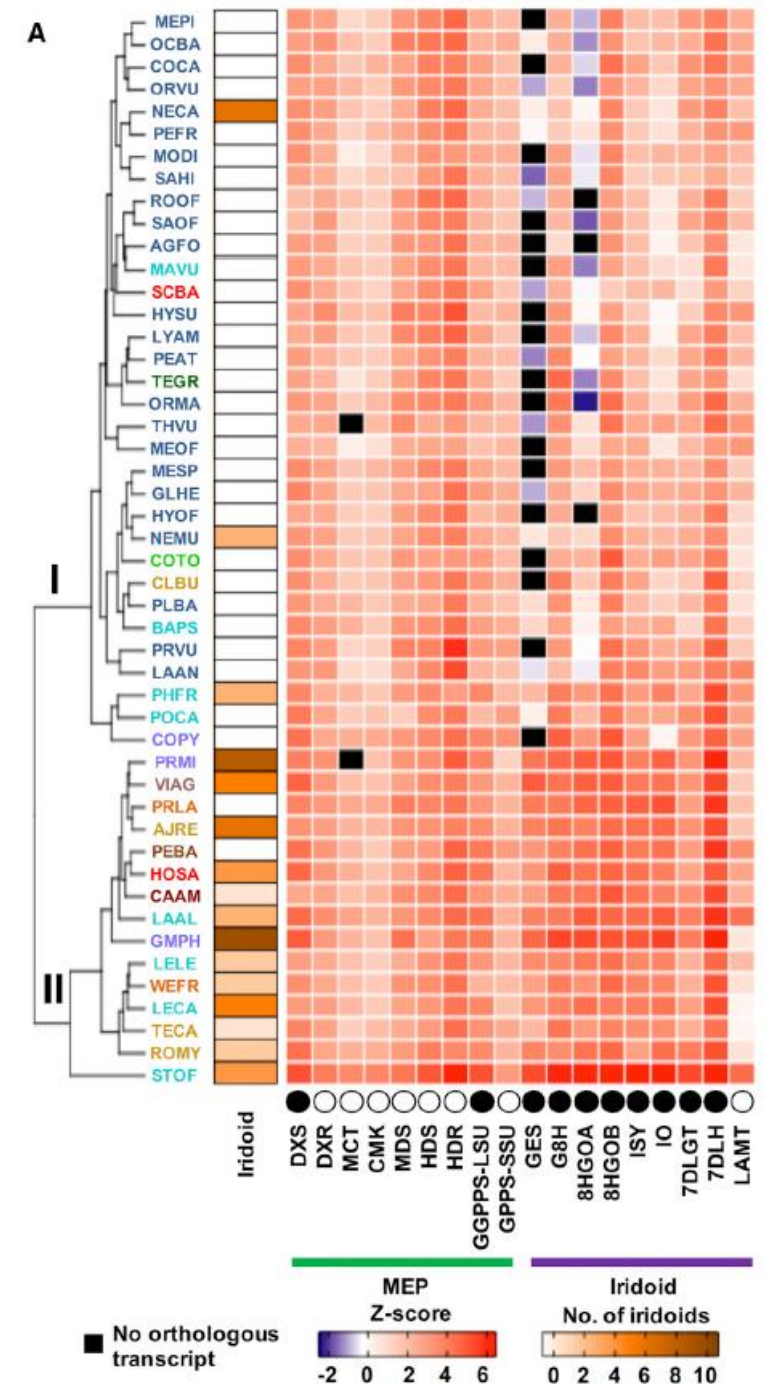


# Overview of phylogenomic method



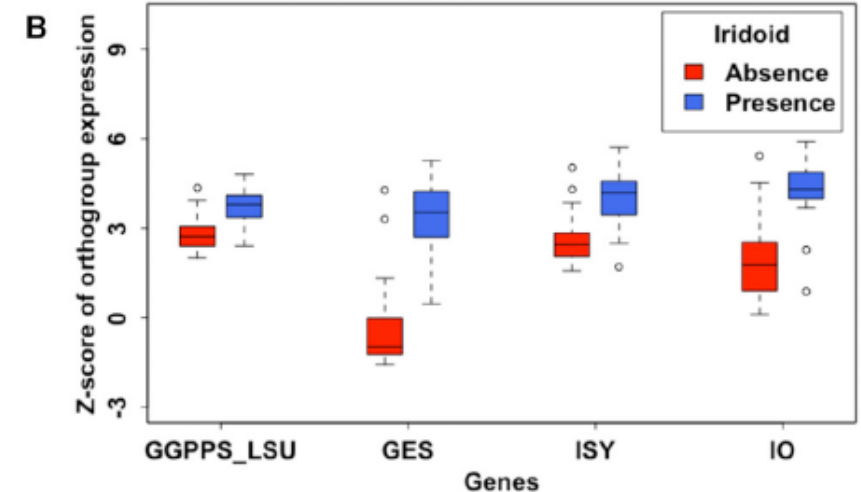
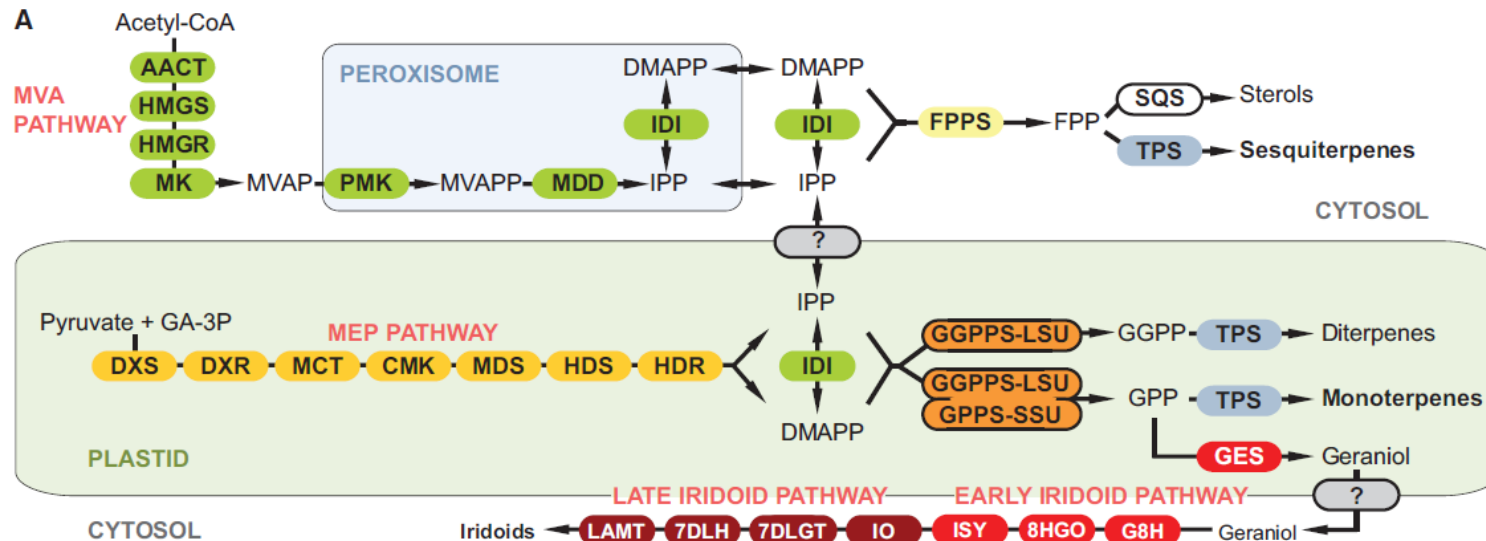
# Iridoid pathway

- Species clustered/ordered by iridoid pathway gene orthogroup expression levels
- 2 distinct clusters:
  - Group I – non-iridoid producing
  - Group II – iridoid producing
- 3 iridoid producing species (NECA, NEMU, PHFR) clustered in Group I
  - Use of genes from different orthogroups to produce iridoids
  - Independent evolution of iridoid pathway



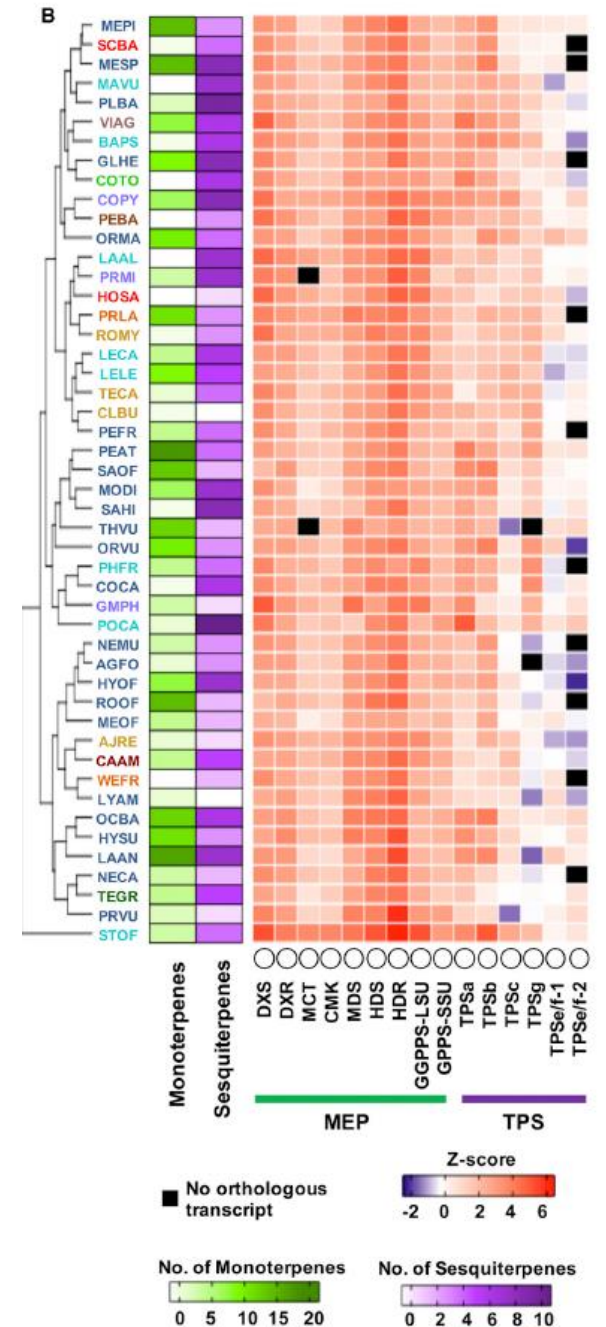
# Iridoid pathway

- 4 orthogroups had very significant differences ( $P < 0.001$ ) in expression between Groups I and II
- Many species in Group I (no iridoids) had no detectable GES expression – key gatekeeping step



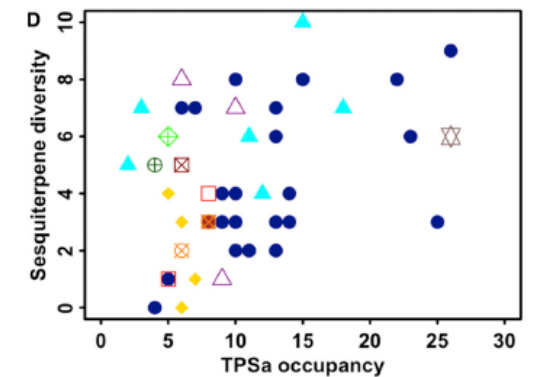
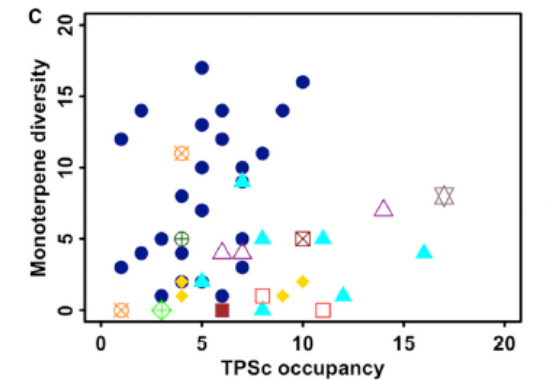
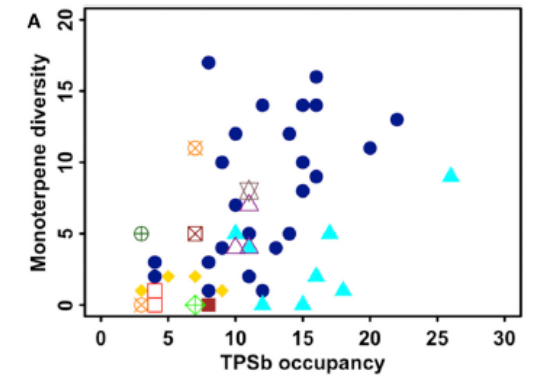
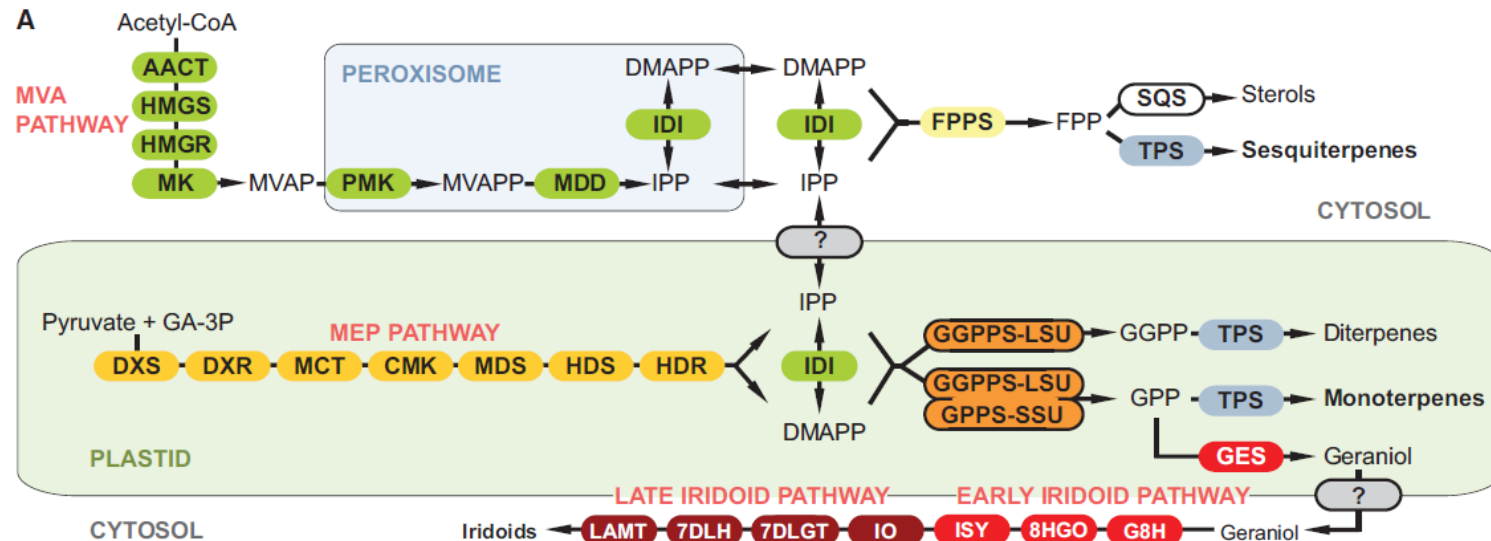
# Mono- and sesquiterpene pathways

- Species clustered/ordered by mono- and sesquiterpene pathway gene orthogroup expression levels
- Overall weak relationship between orthogroup expression levels and mono- and sesquiterpene diversity
  - Primary driver is not gene expression, but rather gene family expansion (i.e. number of genes in each orthogroup)



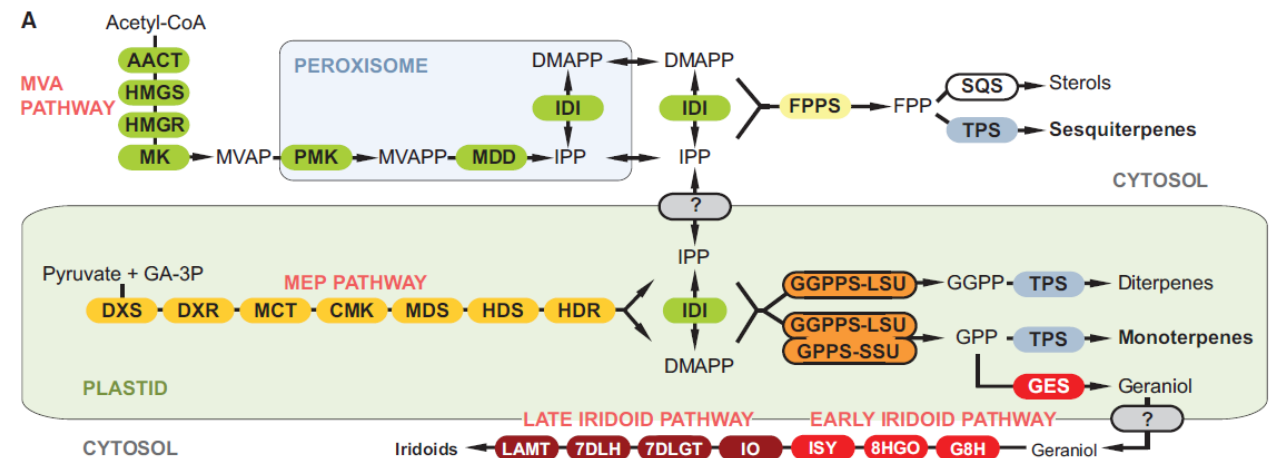
# Mono- and sesquiterpene pathways

- Number of genes in TPS orthogroups controls mono- and sesquiterpene diversity
  - TPSb and TPSc – monoterpenes
  - TPSa – sesquiterpenes



# Mono- and sesquiterpene pathways

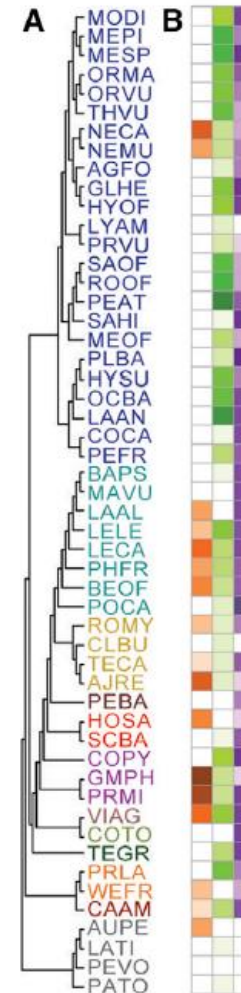
- Negative relationship between GGPPS-LSU, MDS, and HDS orthogroup occupancy and sesquiterpene diversity
- Positive relationship between MDD orthogroup occupancy and sesquiterpene diversity
- Conclusion - Sesquiterpene precursors are provided primarily by MVA pathway, MEP



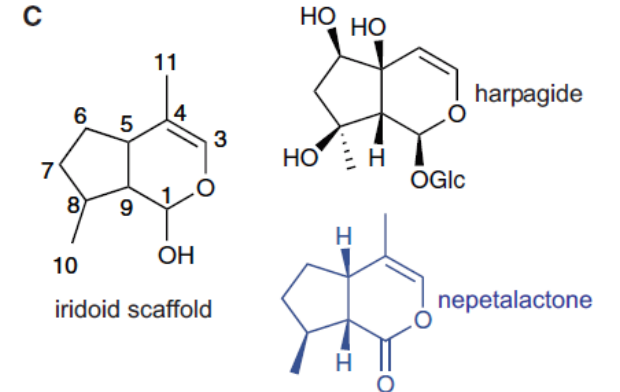


# Discussion

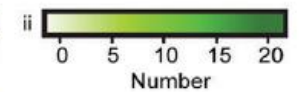
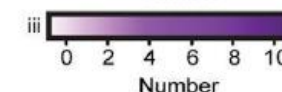
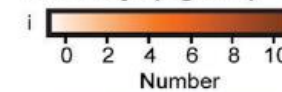
- In Nepetoideae, mono- and sesquiterpene volatiles serve as key defense compounds, in place of iridoid glycosides used by other species
  - Structural analysis shows that primary form of iridoids in family *Nepeta* is nepetalactone – much more volatile than traditional iridoids used by non-Nepetoideae species
  - Nepetalactones resemble insect pheromones – play role in plant-insect signaling



i. IRIDIDS  
ii. MONOTERPENES  
iii. SESQUITERPENES



Color Keys (Figure b)





# Discussion

- In general, production of iridoids and monoterpenes is inversely correlated
  - GES controls flux into iridoid or monoterpene pathway
- Iridoid production is associated with enhanced gene expression
- Mono- and sesquiterpene production is associated with increased orthogroup size

