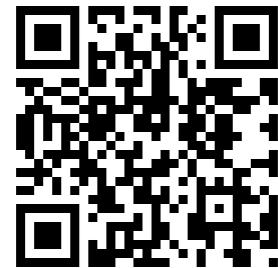


Prof. Dr. Boas Pucker

# **Identification of Biosynthesis Pathways**

# Availability of slides

- All materials are freely available (CC BY) - after the lectures:
  - eCampus: WBIO-A-08
  - GitHub: <https://github.com/bpucker/teaching>
- Questions: Feel free to ask at any time
- Feedback, comments, or questions: [pucker\[a\]uni-bonn.de](mailto:pucker[a]uni-bonn.de)



My figures and content can be re-used in accordance with CC BY 4.0, but this might not apply to all images/logos. Some figure were constructed using bioRender.com.

# Do you know metabolites produced by plants?



# Specialized plant metabolites



*Gossypium davidsonii*  
cellulose (cotton cloth)



*Daucus carota*  
anthocyanins (food color)

*Erythroxylum coca*  
cocaine (local anesthesia)

*Cinchona officinalis*  
quinine (malaria)



*Artemisia annua*  
artemisinin (malaria)



*Beta vulgaris*  
betalains (food color)



*Zea mays*  
 $\beta$ -carotene (provitamin A)

## Applications of plant metabolites

*Taxus brevifolia*  
paclitaxel (cancer)

*Atropa belladonna*  
scopolamine (motion sickness)

*Chrysanthemum cinerariaefolium*  
pyrethrin I (insect control)

*Digitalis purpurea*  
digoxin (congestive heart failure)

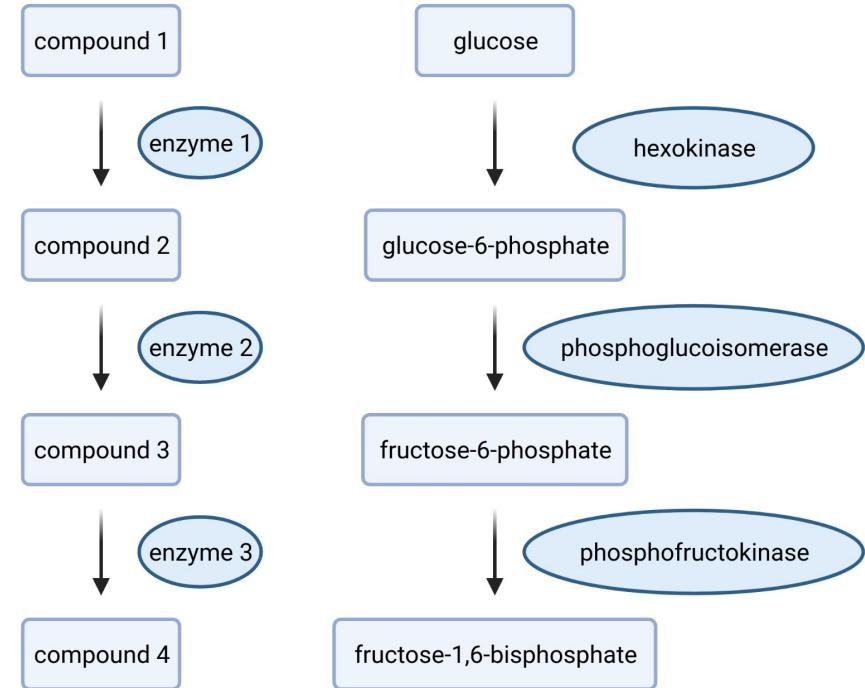
*Rauwolfia serpentina*  
ajmaline (cardiac arrhythmia)

*Crocosmia × crocosmiiflora*   
montbretin A (diabetes)

Partly inspired by 10.1126/science.aad2062  
Irmisch et al., 2020: 10.1104/pp.20.00522

# Biosynthesis pathways

- Biosynthesis pathways comprise steps catalyzed by different enzymes
- Chemical compounds are substrates/products of multiple steps
- How can we understand individual enzyme functions?

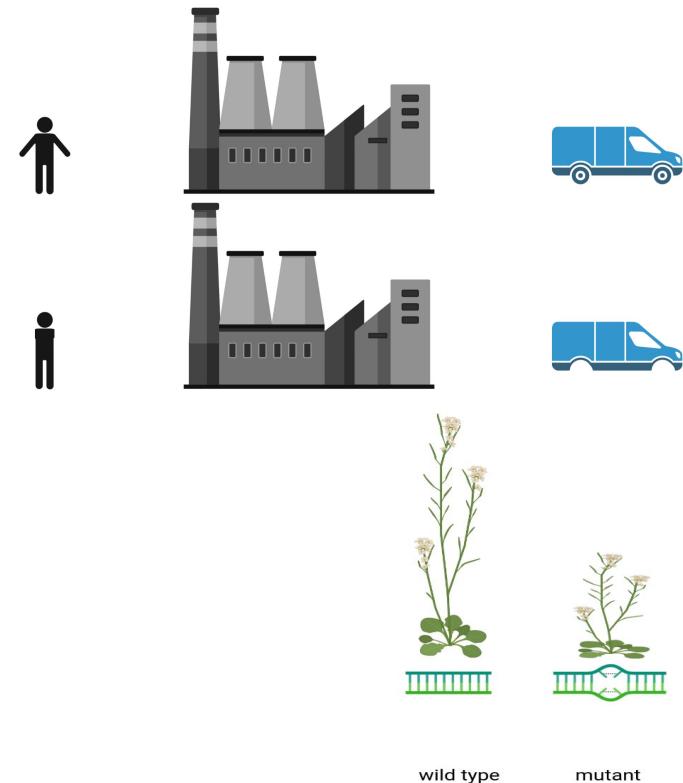


# How would you identify the function of a gene?



# How to reveal the function of a gene/enzyme?

- How would you find out what a car factory worker is doing?
- Bind/remove hands and arms and see how the product differs from the normal product
- Transfer to gene: knock-out a gene and see how the plant looks different compared to a wild type



# How can we do this in plants?

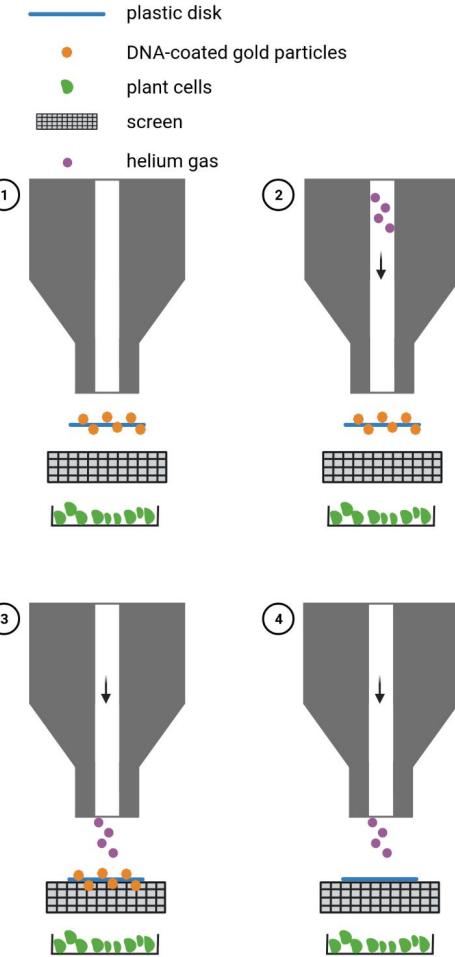
- How to knock-out a plant gene?
- Targeted knock-out through homologous recombination is impossible
- Random mutagenesis was only method for decades of plant research:
  - Transfer-DNA (T-DNA) insertions
  - transposons
  - chemical mutagenesis (Ethyl methanesulfonate, EMS)
  - radiation
- *Clustered Regularly Interspaced Short Palindromic Repeats & CRISPR-associated (CRISPR-Cas)*



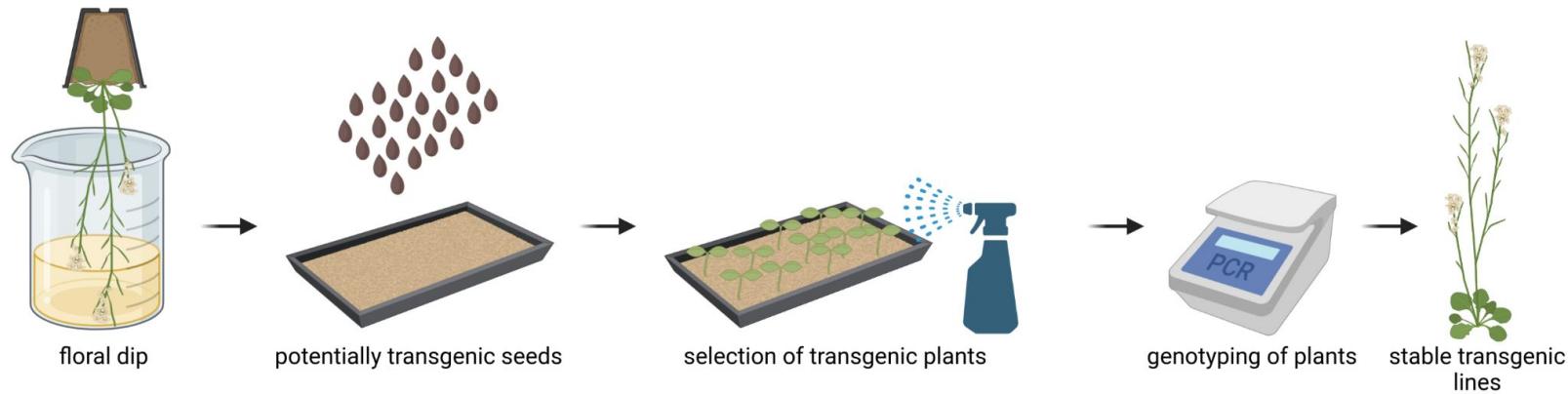
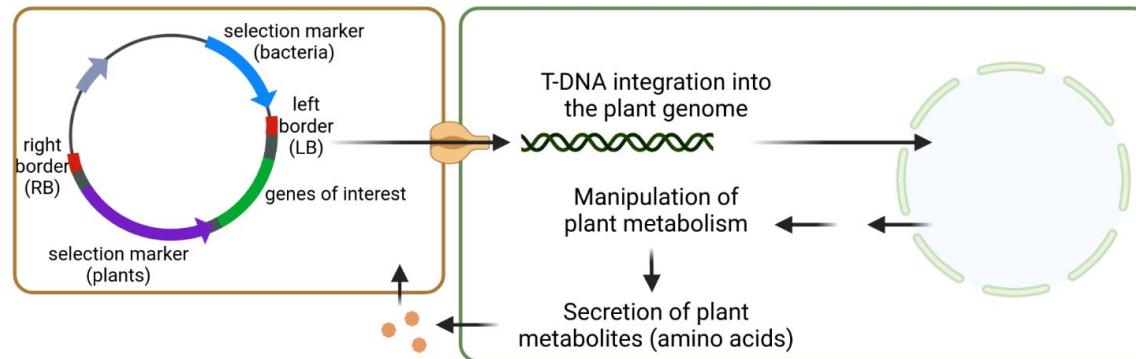
- Methods for stable transformation:
  - particle gun
  - Agrobacterium-mediated
- Integration of DNA in the genome is almost random (no effective homologous recombination)
- Transient transfection of protoplast can be done through  $\text{CaCl}_2$ -based methods or electroporation

# Particle gun

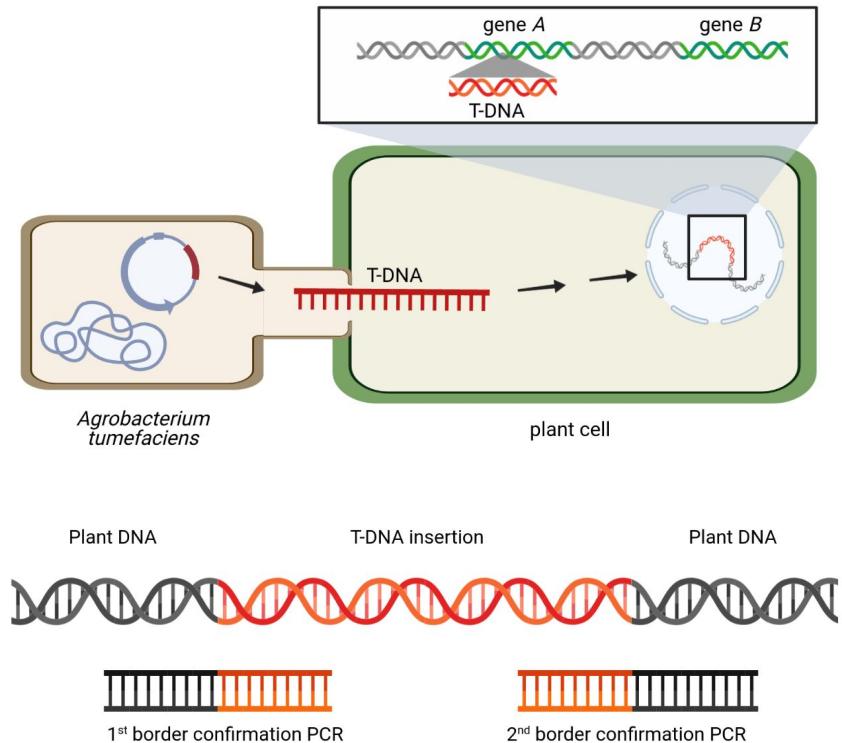
- Shooting DNA into plant cells
- Applicable to all plant species
- Optimization of conditions required



# Agrobacterium-mediated gene transfer (floral dip)

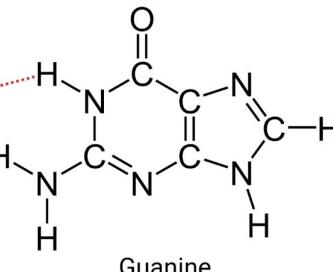
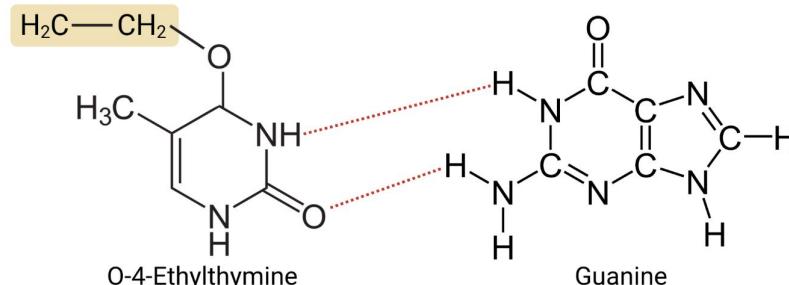
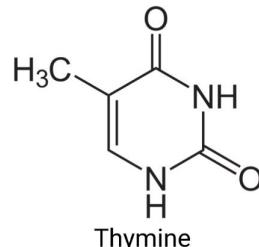
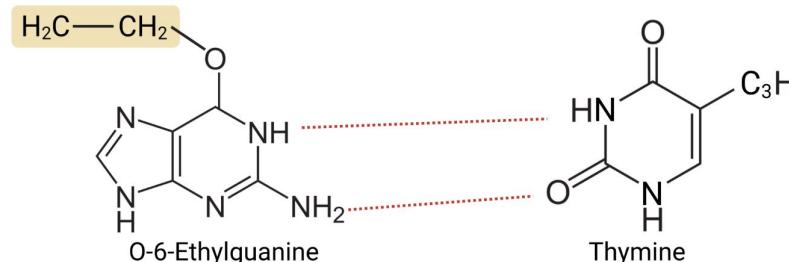
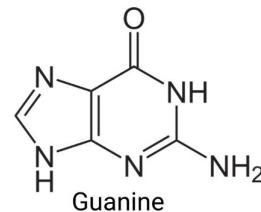
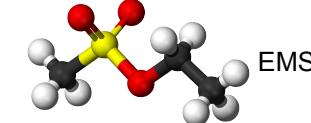


- Integration of DNA in the genome is almost random (no effective homologous recombination)
- T-DNA can disrupt genes
- T-DNA carries marker for selection of transgenic plants

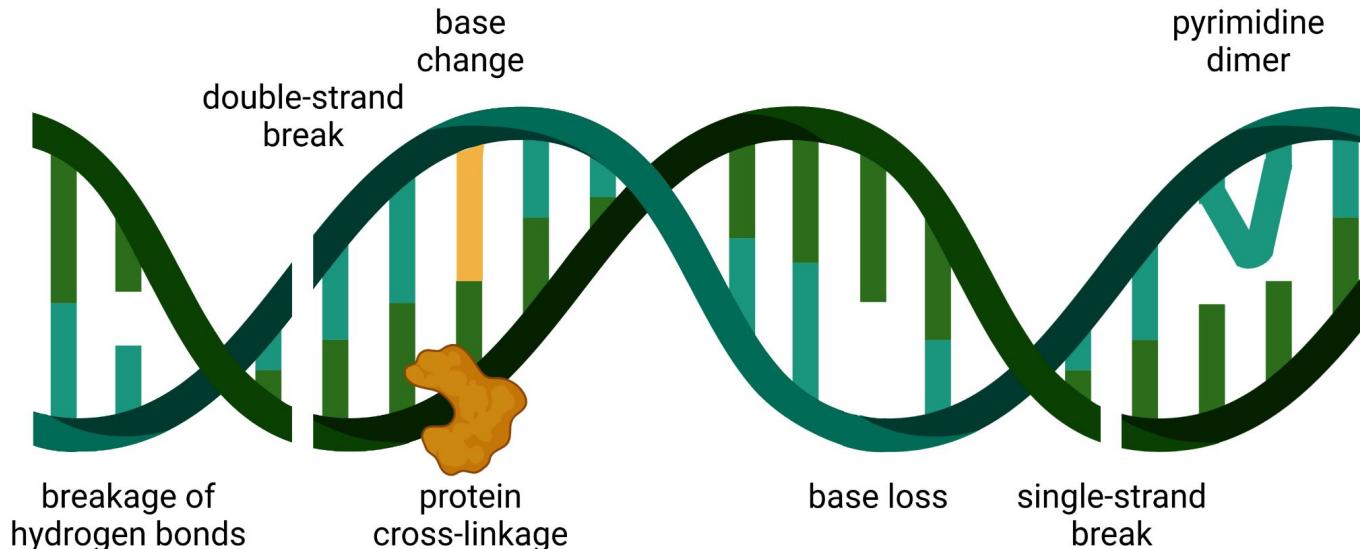


# Ethyl methanesulfonate (EMS)

- EMS causes mostly point mutations
- Seeds are incubated in EMS for mutagenesis

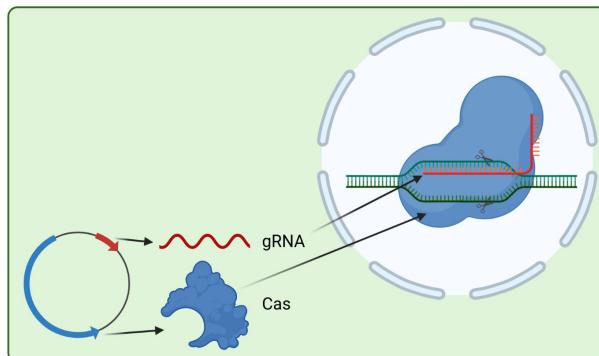


- Energy-rich radiation has a diverse range of effects on DNA
- Example: gamma radiation, ultra-violet (UV) radiation



# Opportunities using CRISPR/Cas9

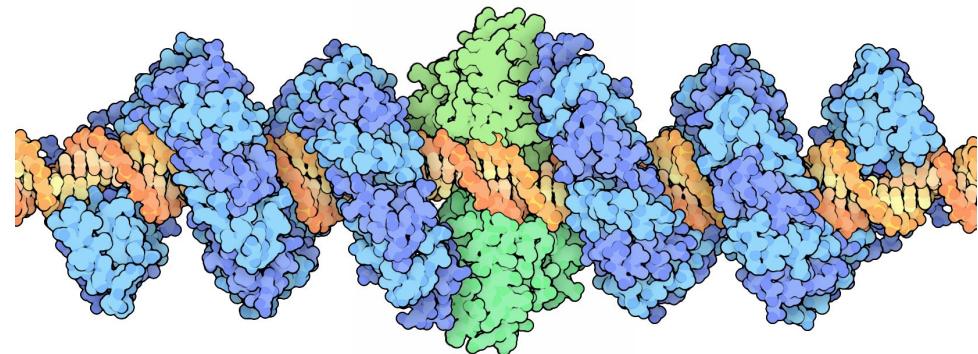
- CRISPR = Clustered Regularly Interspaced Short Palindromic Repeats
- Cas9 = CRISPR-associated
- Targeted modification in plant genomes
- Modifications:
  - Dead Cas protein (dCas, inactive) can be used to block promoter
  - Methyltransferases can be attached to Cas protein
  - ...



‘Emmanuelle Charpentier and Jennifer A. Doudna have discovered one of gene technology’s sharpest tools: the CRISPR/Cas9 genetic scissors. Using these, researchers can change the DNA of animals, plants and microorganisms with extremely high precision. This technology has had a revolutionary impact on the life sciences, is contributing to new cancer therapies and may make the dream of curing inherited diseases come true.’

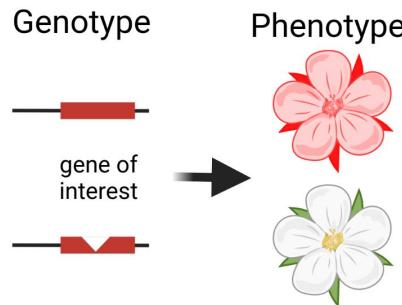
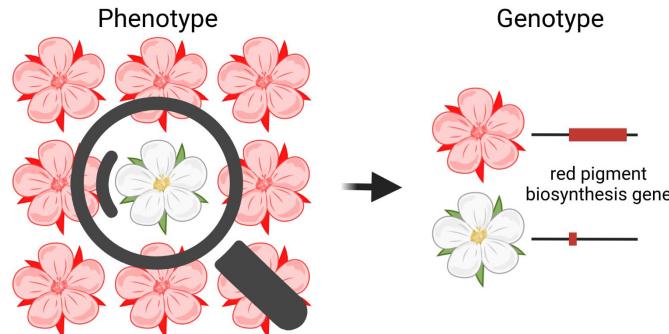
<https://www.nobelprize.org/prizes/chemistry/2020/press-release/>

- TALEN = Transcription Activator-Like Effector Nucleases
- TAL = DNA binding domain
  - highly repeated 33-34 amino acids
  - 12th + 13th amino acid determine specificity
- Nuclease to cut DNA at binding site



# Forward and reverse genetics

- Forward genetics = Interesting phenotype is observed and responsible gene is identified in the next step
  - Examples: transposon mutagenesis, ethyl methanesulfonate (EMS), radiation
- Reverse genetics = Finding the function of a known gene through targeted mutation of this gene
  - Examples: Transfer-DNA (T-DNA) collections, CRISPR-Cas



# Which properties of a model species do you know?



# Model systems - species

- Investigation of a research question requires a model species
- Important properties of model species:
  - Fully sequenced genome
  - Small size
  - Large number of seeds (offspring)
  - Accessible by genetic engineering
  - Short generation time
  - Large community
  - ...

*Arabidopsis  
thaliana*



[https://commons.wikimedia.org/wiki/File:Arabidopsis\\_thaliana.jpg](https://commons.wikimedia.org/wiki/File:Arabidopsis_thaliana.jpg) (CC-BY-SA)

# Model system - species (examples)

- *Arabidopsis thaliana* (thale cress): THE most important model system due to the genome sequence
- *Antirrhinum majus* (snapdragon): pioneer model plant; variation and genetics
- *Petunia x hybrida* (petunia): evolution and development
- *Zea mays* (maize): discovery of transposons by Barbara McClintock (Nobel prize)
- *Oryza sativa* (rice): model organism for research on crops

Parry et al., 2020: 10.1101/pld3.248

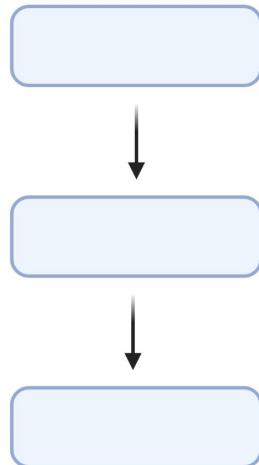
Schwarz-Sommer et al., 2003: 10.1038/nrg1127

Vandenbussche et al., 2016: 10.3389/fpls.2016.00072

Strable & Scanlon, 2009: 10.1101/pdb.em0132

Rensink & Buell, 2004: 10.1104/pp.104.040170

# Properties of a model biosynthesis pathway?



## Model system - pathway

- Easily detectable product
- Stable intermediates that can be quantified
- Gene knock-outs must not be lethal
- Conserved across many plant species
- Community working on the pathway
- Different branches
- ...

# Model system - pathway (examples)

- Flavonoid biosynthesis:
  - Anthocyanins: visible as pigments in the flowers of many plant species
  - Proanthocyanidins (PAs): visible as pigments of the seed coat that are responsible for the dark coloration
- Carotenoid biosynthesis
- Betalain biosynthesis (only in Caryophyllales)



Anthocyanin-pigmented flower



Proanthocyanidin-pigmented seeds



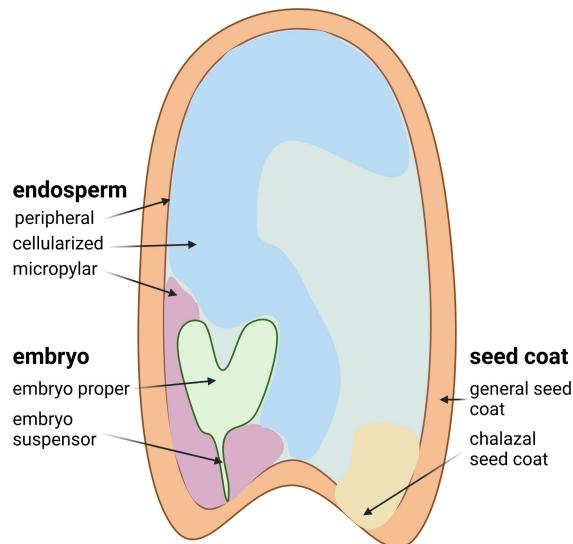
Carotenoid-pigmented carrots



Betalain-pigmented flowers

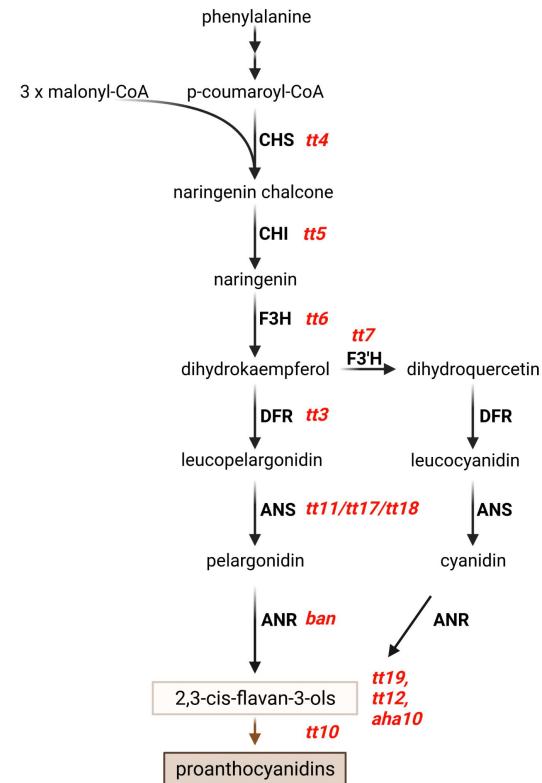
# Transparent testa

- Brown proanthocyanidins are located in the testa
- Transparent testa = lack of proanthocyanidins (no pigmentation)



# *Transparent testa* mutants

- Large collection of different *tt* mutants
- Numbering is based on order of discovery
- Checking if new *transparent testa* (*tt*) mutants are allelic with existing ones

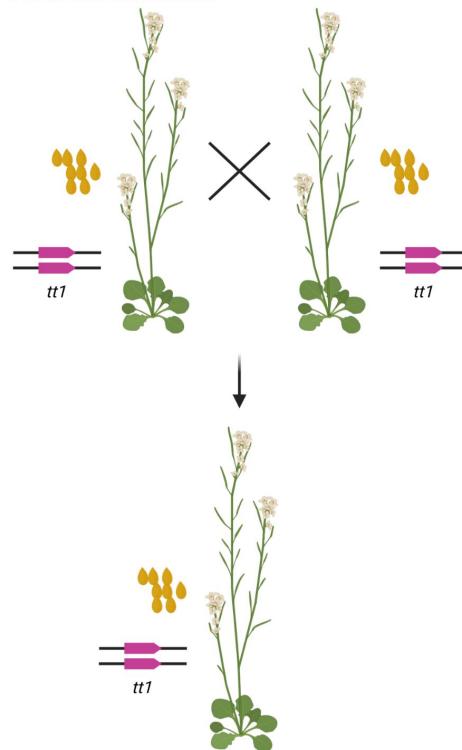


# How check if two mutations are in the same gene?

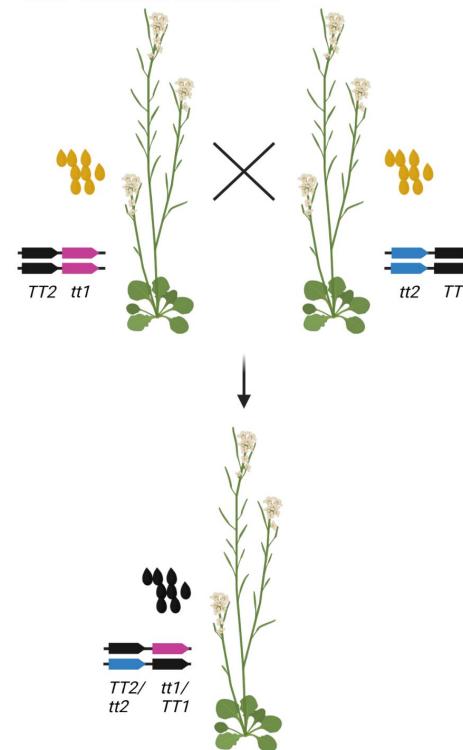


# Allelic mutants?

Allelic mutants:



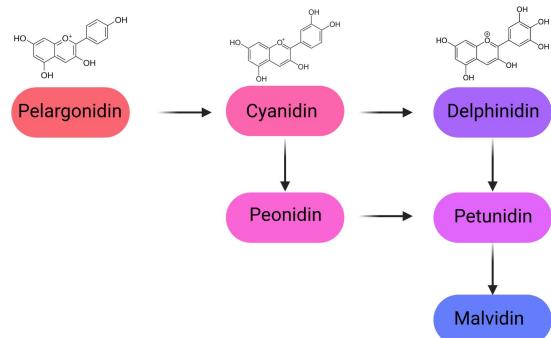
Non-allelic mutants:



$tt$  = transparent testa

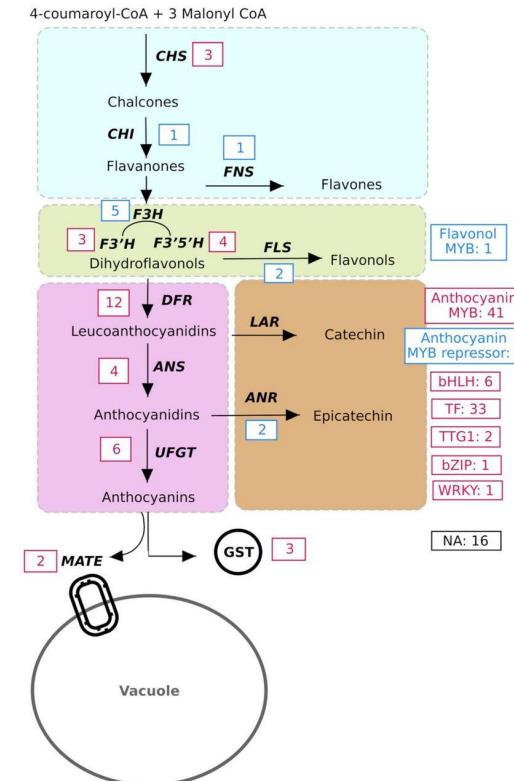
# Anthocyanin loss

- Anthocyanins are responsible for red to blue pigmentation of flowers
- Anthocyanins are derivatives from phenylalanine through the flavonoid biosynthesis
- *Digitalis purpurea* is an ornamental plant with biomedical potential
- Different varieties of *D. purpurea* show pigmentation differences



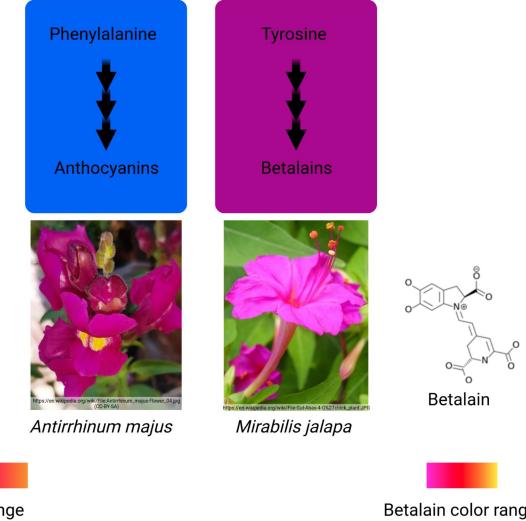
# Genetic factors blocking anthocyanin accumulation

- Transcription factors often responsible
- Genes at branch points often affected



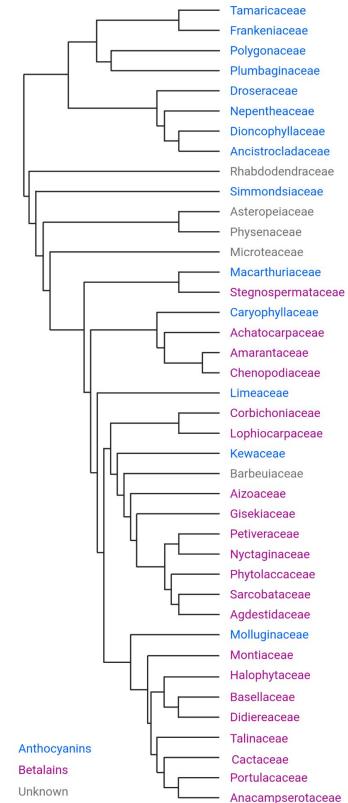
# Betalain pigmentation

- Betalains are tyrosine-derived pigments
- Production of betalains is restricted to one flowering plant order: Caryophyllales
- Similarities with the anthocyanin biosynthesis e.g. color of pigments
- Evolutionary benefit of betalains over anthocyanins remains unknown



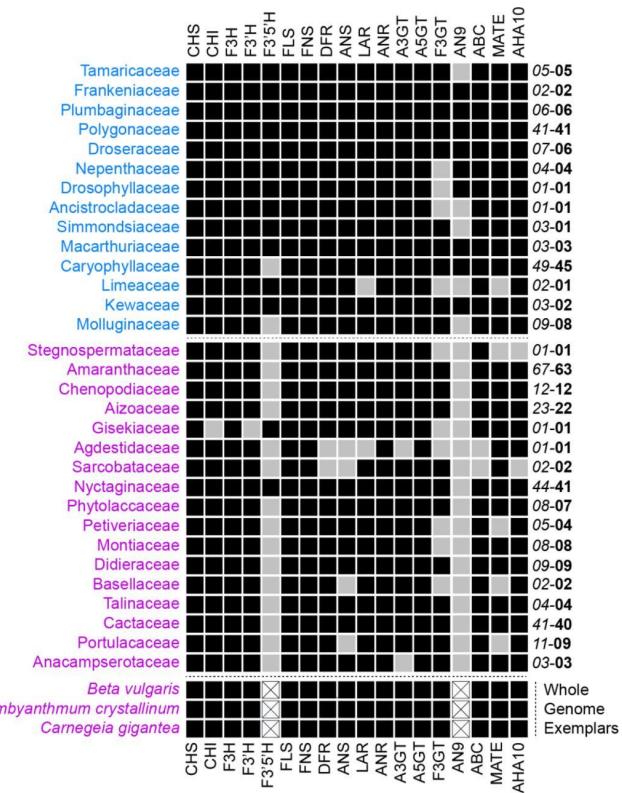
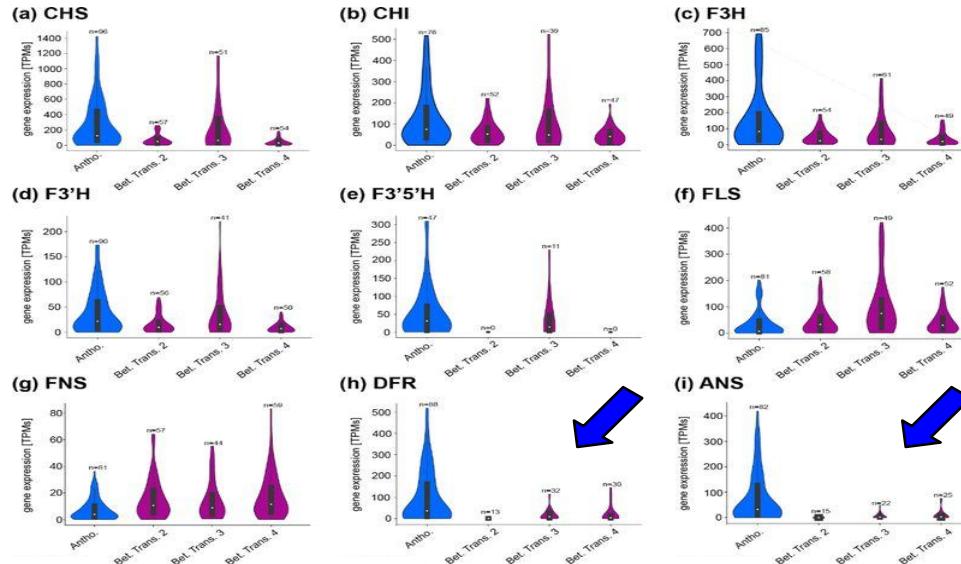
# Mutual exclusion of anthocyanins and betalains

- Anthocyanins & betalains have similar properties leading to partial redundancy
- Research on anthocyanins & betalains since 1970
- Both pigments have never been observed in the same species
- Mutual exclusion of both pigments is assumed:
  - Enzyme-encoding genes might be lowly expressed
  - Transcription factor(s) activating anthocyanin biosynthesis genes might have been co-opted for betalain biosynthesis
  - Anthocyanins transport in betalain-pigmented species seems blocked



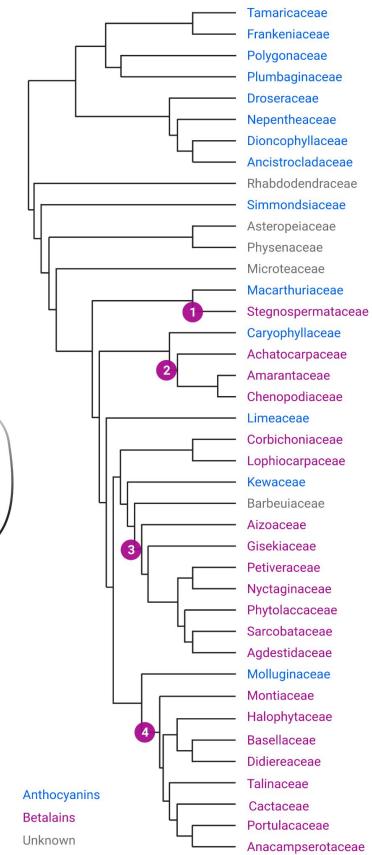
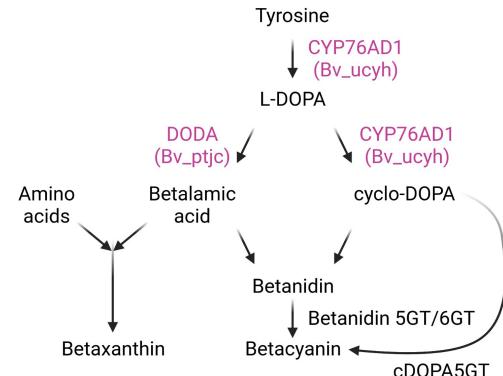
# Genetic factors explaining lack of anthocyanins in betalain-pigmented species

- AN9 ortholog loss (probably an enzyme)
- DFR and ANS lowly transcribed
- Transcription factor complex non-functional



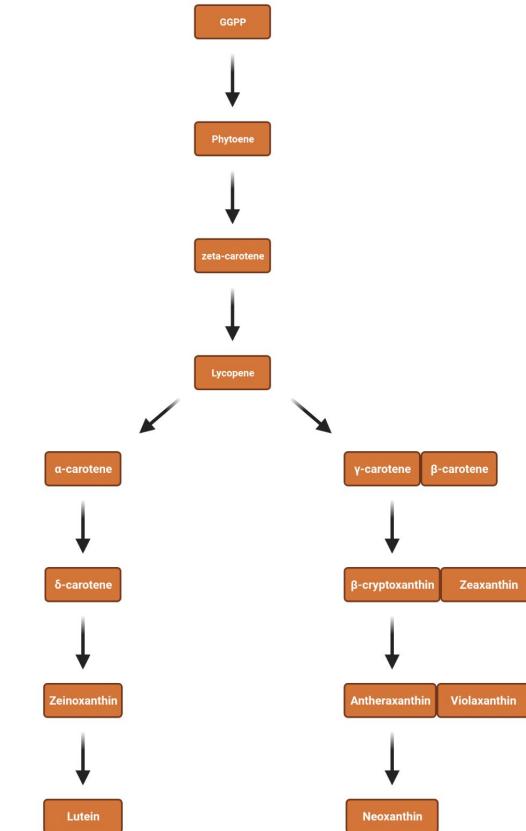
# Independent origins

- Betalain biosynthesis evolved four times independently
- Excellent model system for the investigation of pigment biosynthesis pathways
- Comparison of different betalain origins



# Model system - pathway: Carotenoid biosynthesis

- Carotenoids protect chlorophyll
- Almost all plants have chlorophyll and require this protection
- Are there plants without carotenoids?



# Why are carotenoids not a perfect model system?



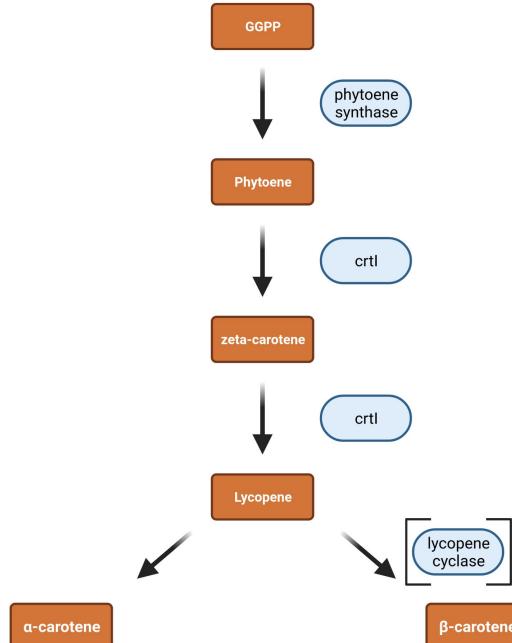
# Reconstruction of pathways in plants (golden rice)

UNIVERSITÄT BONN

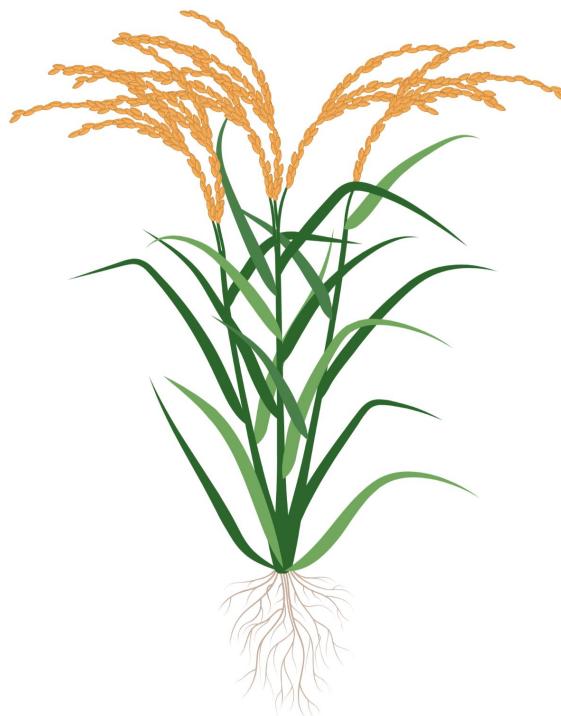
- Vitamin A deficiency causes numerous illnesses (in developing countries)
- Rice was engineered to increase the  $\beta$ -carotene content (provitamin A):
  - psy = phytoene synthase (*Narcissus pseudonarcissus*)
  - crtI = phytoene desaturase (*Erwinia uredovora*)
  - lcy = leucopene cyclase (already present)



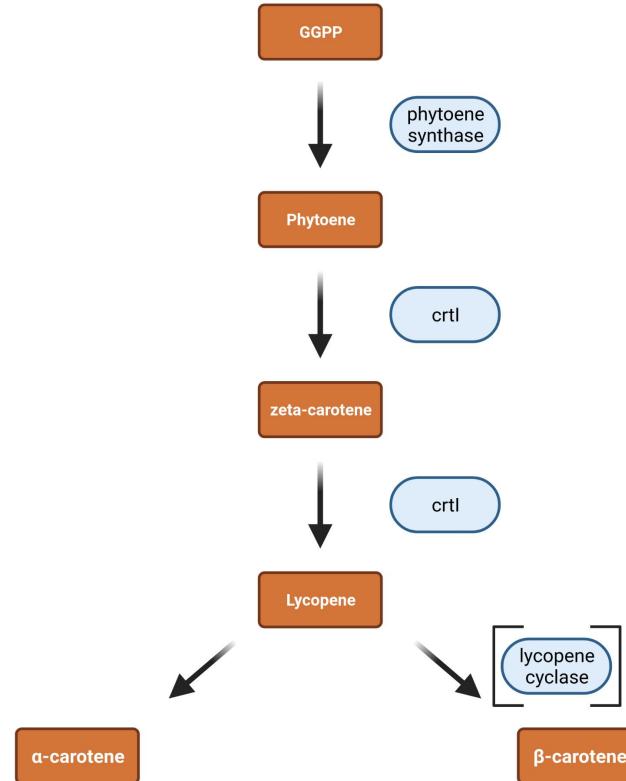
More details about vitamin A: <https://github.com/bpucker/teaching/blob/master/VitaminA.pdf>



# How to improve golden rice?



- General considerations:
  - Optimization of gene expression in tissues (endosperm)
  - Stability of system over time
  - Genomic integration locus should not influence other traits
- Golden Rice 2:
  - Selection of effective enzymes (maize psy is more efficient)
  - >20x increase in carotenoid production



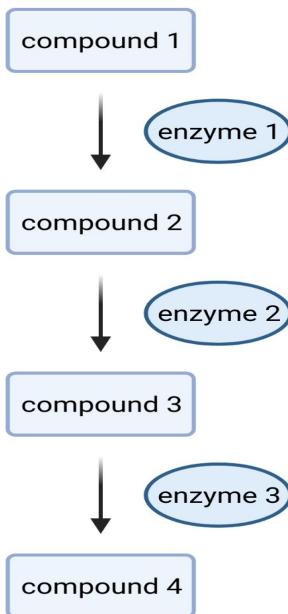
# Changing flower color of petunia

- Discovery of antisense transcripts by expressing *Chalcone Synthase (CHS)* heterologously in petunia
- Pigmentation was not enhanced, but reduced by RNA-mediated gene silencing (50x reduced *CHS* transcript)

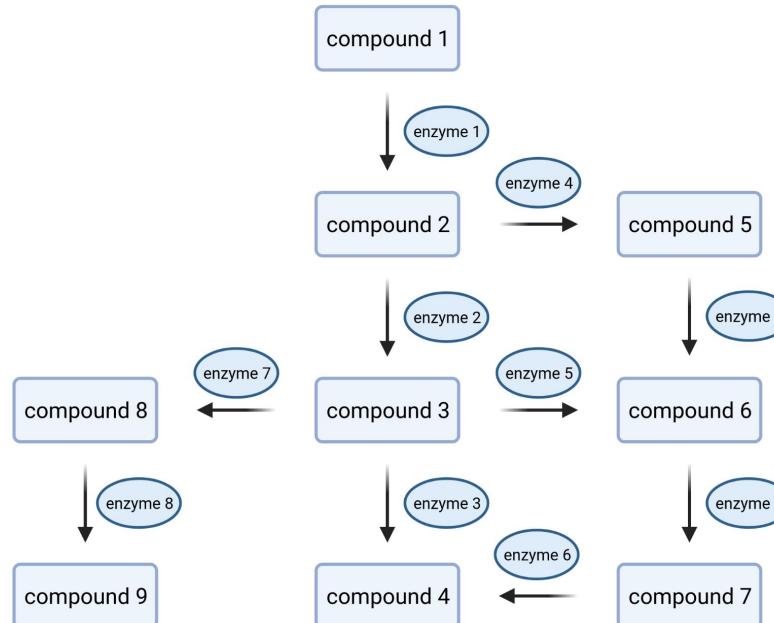


# Pathways are networks

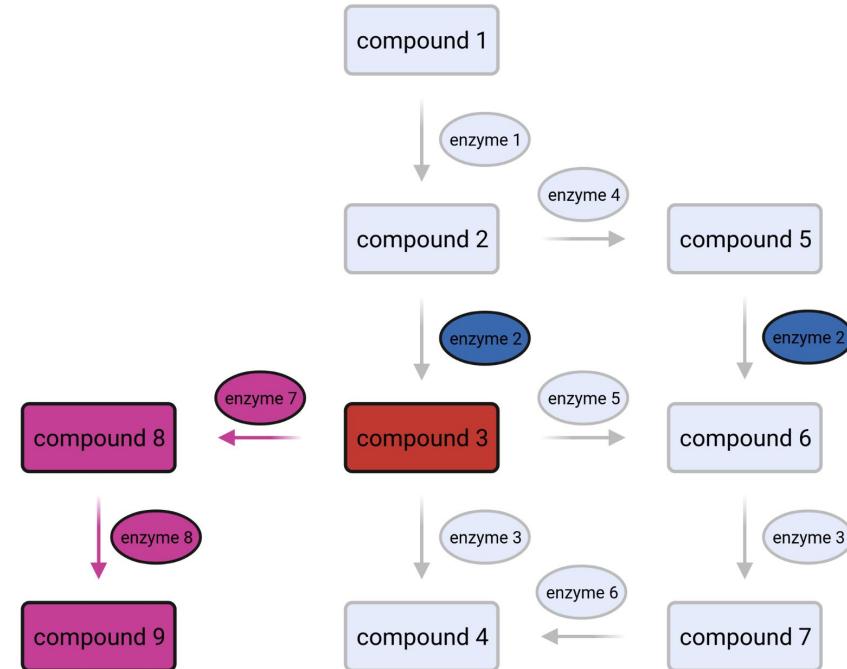
## Text book



## Reality

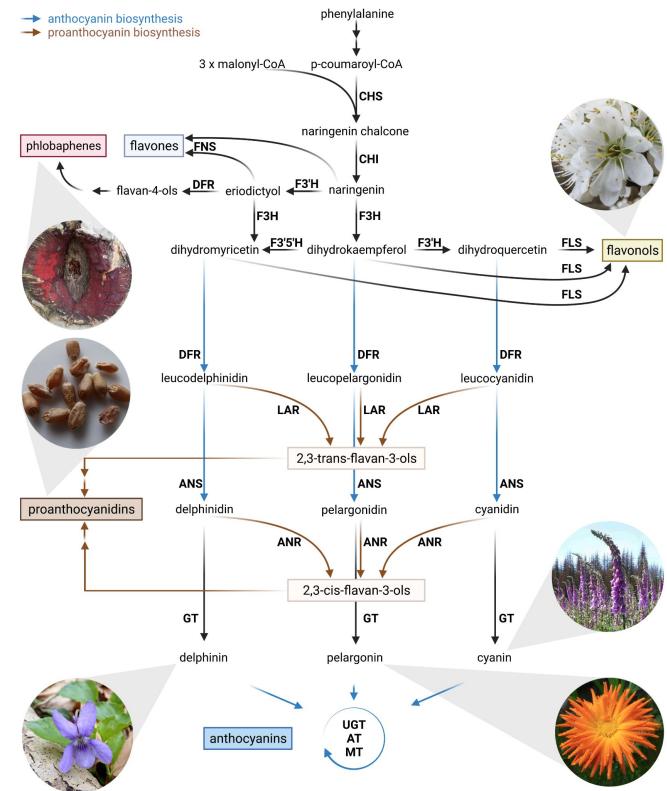


- Metabolic networks are characterized by hubs (central intermediates)
- Many enzymes show promiscuity i.e. catalyze different reactions
- Branches could be considered linear pathways
- Edge (reactions); nodes (compounds)



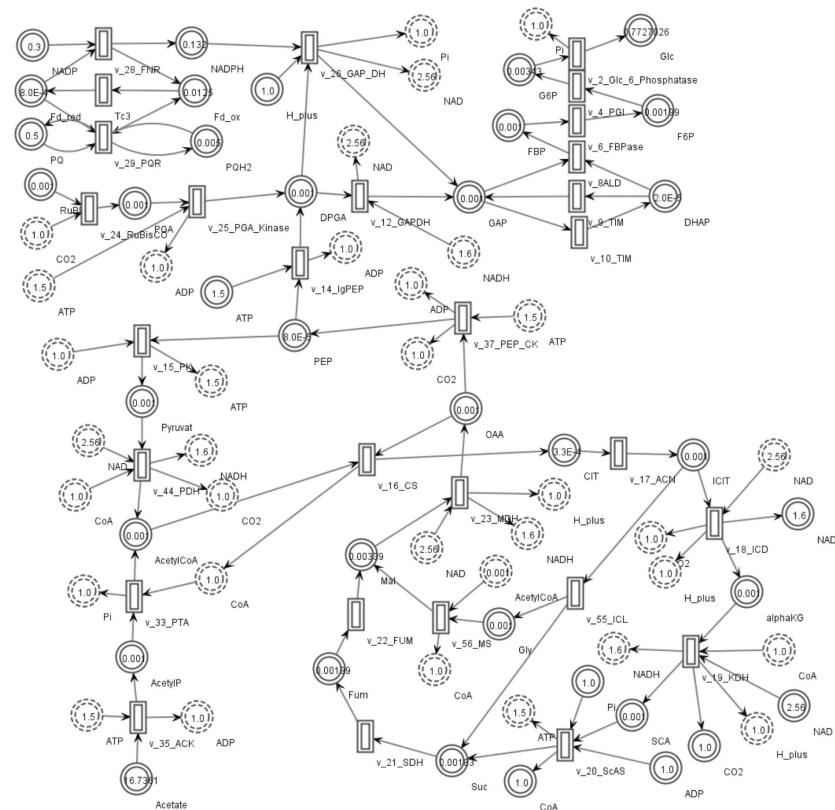
# Metabolic networks - example

- Flavonoid biosynthesis:
  - central compounds like naringenin
  - enzymes catalyze different reactions: flavonoid 3'-hydroxylase (F3'H), flavonol synthase (FLS), dihydroflavonol 4-reductase (DFR), ...
  - complex network with different branches: proanthocyanidins (brown), anthocyanins (blue)



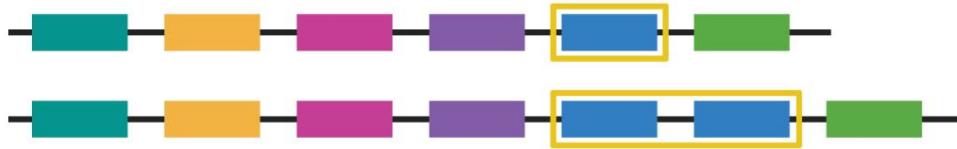
# Metabolic network - example 2

- Central carbon metabolism of *C. reinhardtii*
- Well studied system with quantitative information about different enzymatic reactions



# Importance of gene duplications (isoforms)

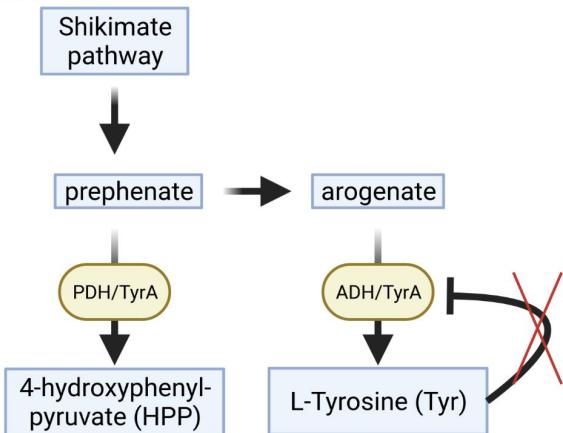
- Gene duplications form the basis for the evolution of novel genes
- Encoded enzymes can acquire novel functions
- Identification of gene duplications can reveal candidates



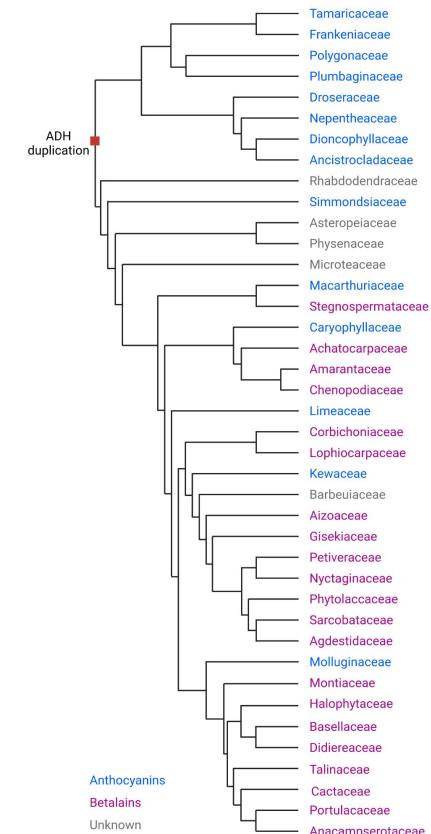
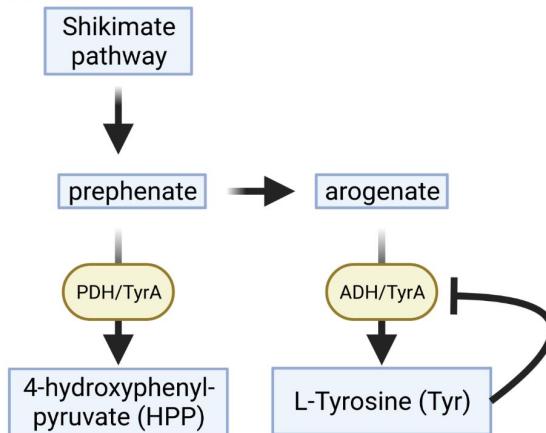
# Importance of gene duplications (isoforms)

- Arogenate dehydrogenase (ADH) duplication gave rise to feedback-resistant isoform
- ADHa might be required for betalain pigmentation

ADHa:

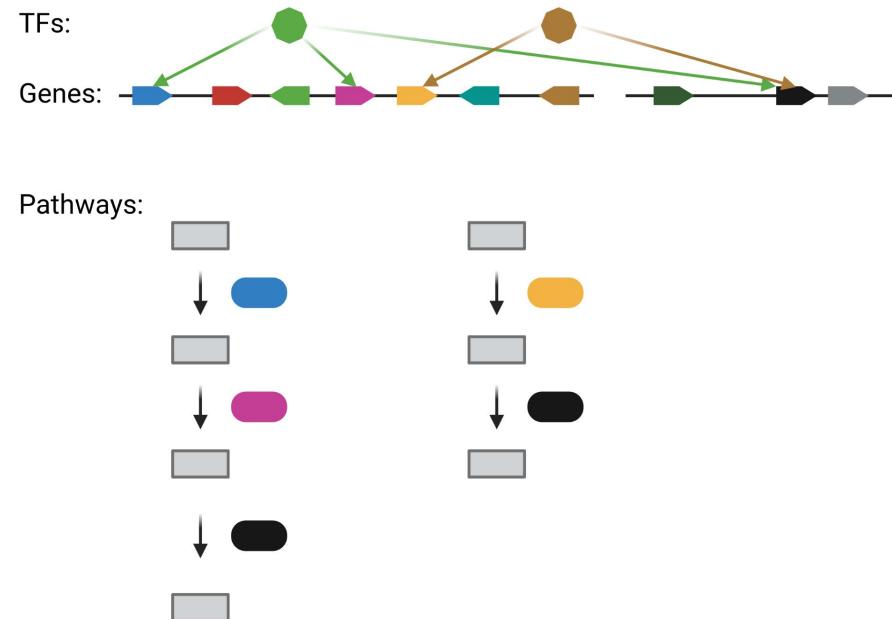


ADHb:



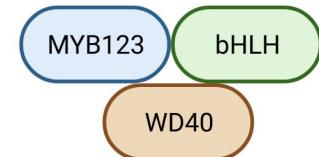
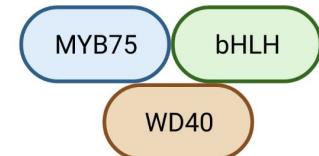
# Transcriptional control of biosynthesis pathways

- Biosynthesis pathways are controlled by transcriptional regulation
- Transcription factors (TFs) ensure activation of all required genes
- Genes can be involved in different pathways



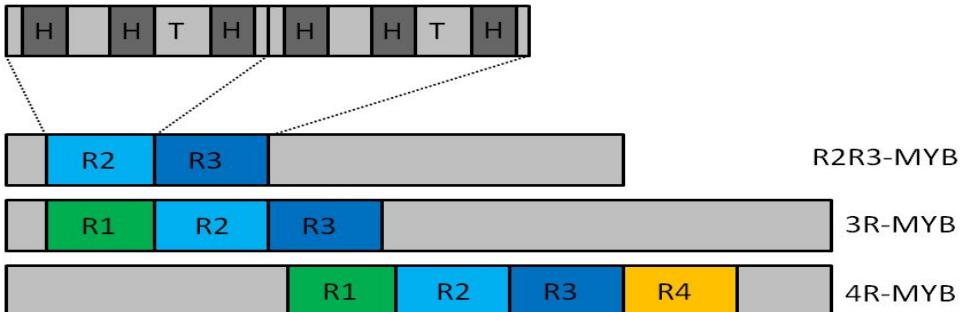
# Transcriptional control of the flavonoid biosynthesis

- Flavonoid biosynthesis is controlled by several transcription factor (TF) families
- MYBs are considered specific regulators of different flavonoid biosynthesis branches
- bHLHs are co-activators required for multiple branches
- WD40 (TTG1) is considered a scaffold protein for connection of MYB and bHLH
- MBW complex: MYB + bHLH + WD40
- WRKYs might be associated with the MBW complex

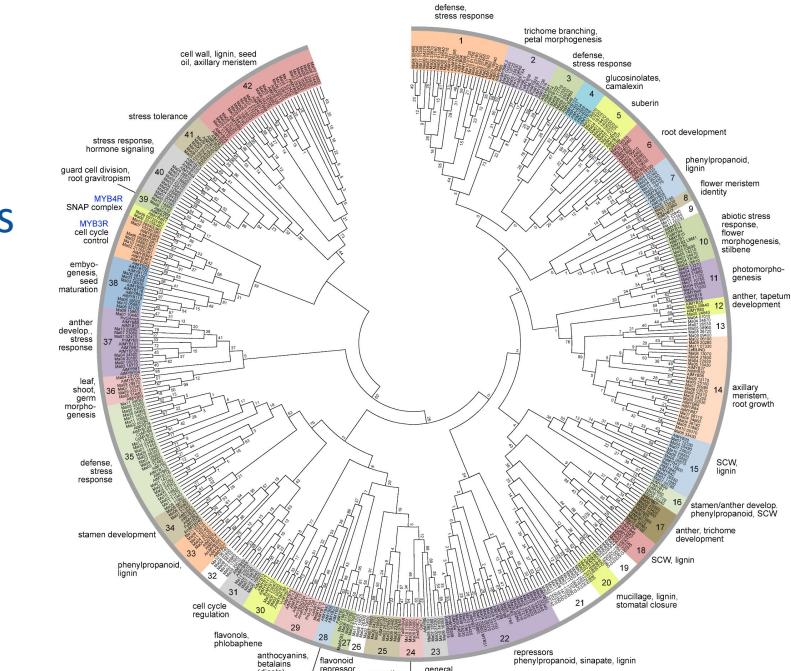


- MYB (myeloblastosis) is one of the largest TF families in plants
- Characterized by repeats (R)
- R2R3-MYBs are particularly important in plants
- Involved in numerous functions in plants

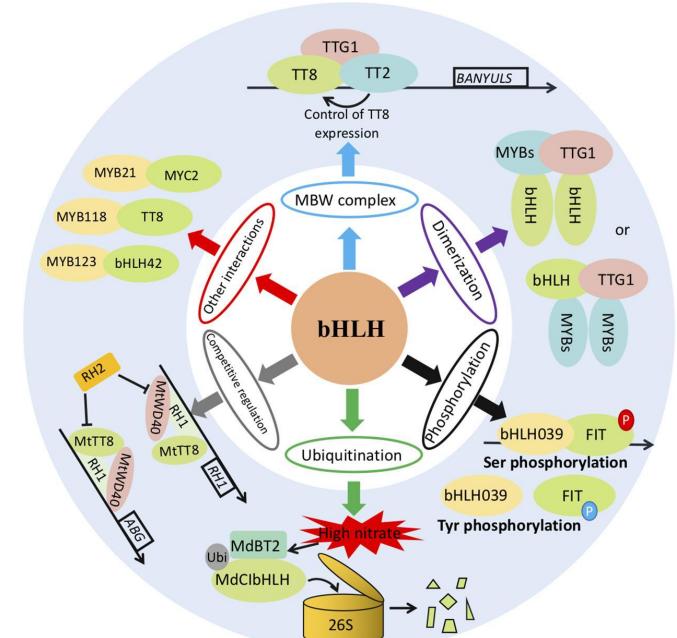
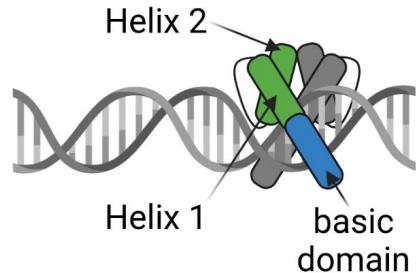
W-(X19)-W-(X19)-W-....-F/I-(X18)-W-(X18)-W-



Based on Dubos et al., 2010

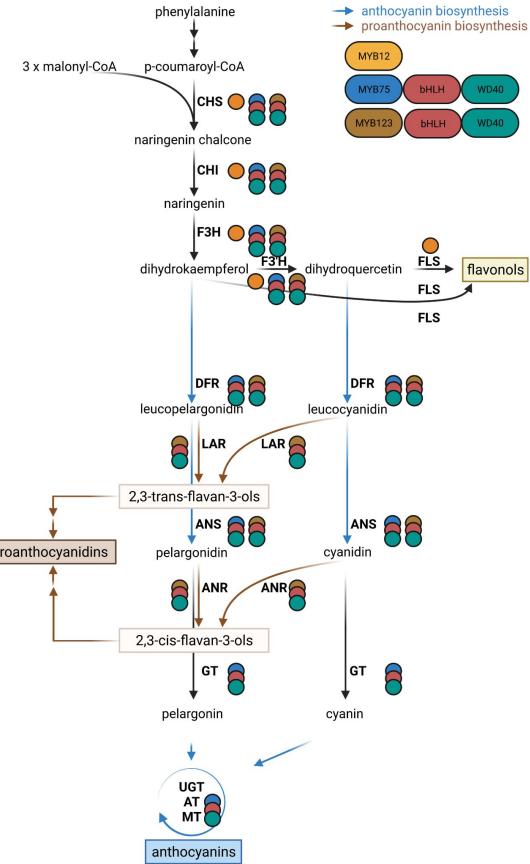


- bHLH (basic Helix-Loop-Helix) transcription factors form a large family in plants
- Transcriptional activation in cooperation with MYBs and independently
- bHLH transcription factors operate often in dimers



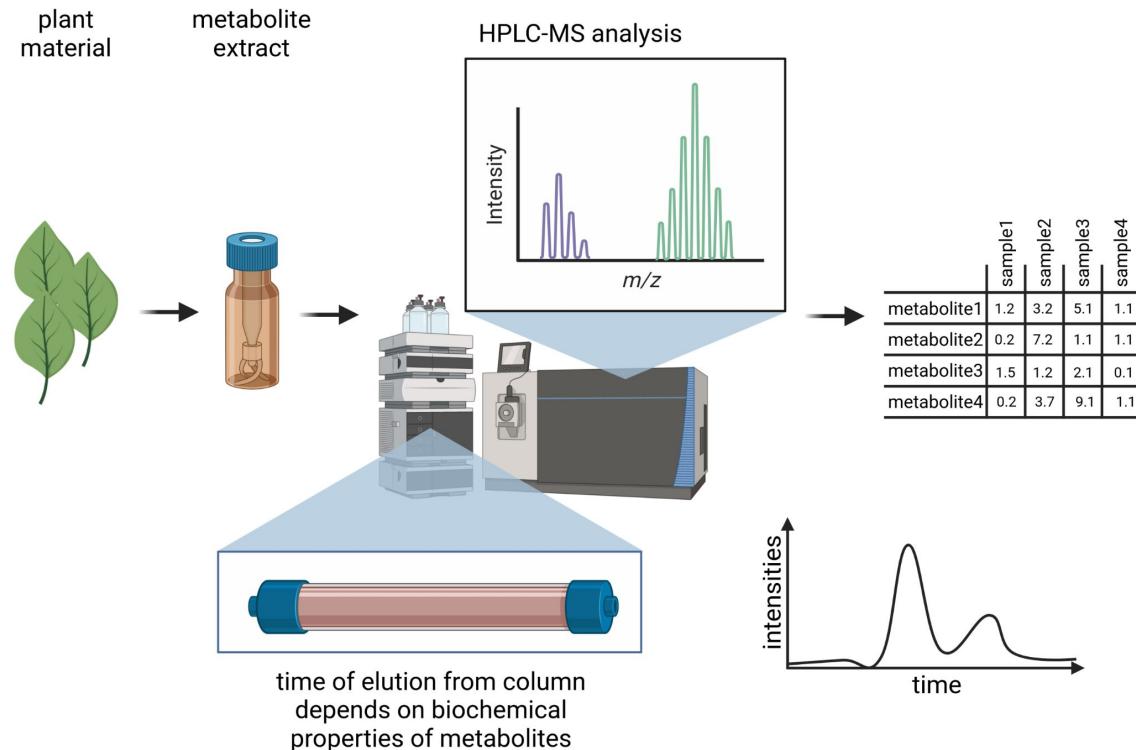
# Flavonoid biosynthesis mutants

- Example 1: *fls* or *myb12*
  - Flavonols (-)
  - Anthocyanins (+)
  - Proanthocyanidins (+)
- Example 2: *chs*, *chi*, or *f3h*
  - Flavonols (-)
  - Anthocyanins (-)
  - Proanthocyanidins (-)
- Example 3: *bhlh*, *wd40*, *dfr*
  - Flavonols (+)
  - Anthocyanins (-)
  - Proanthocyanidins (-)

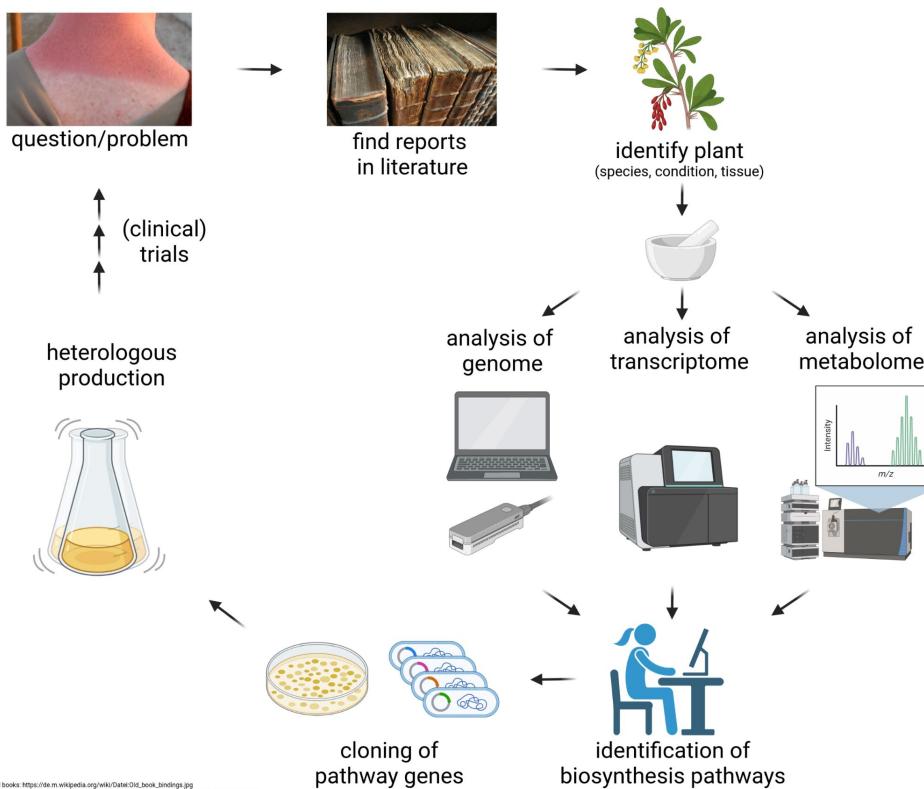


# Metabolite analytics

- Discovery of pathways requires quantification of metabolites
- HPLC-MS and GC-MS are standard methods for metabolite quantification
- Integration of metabolite data with other omics



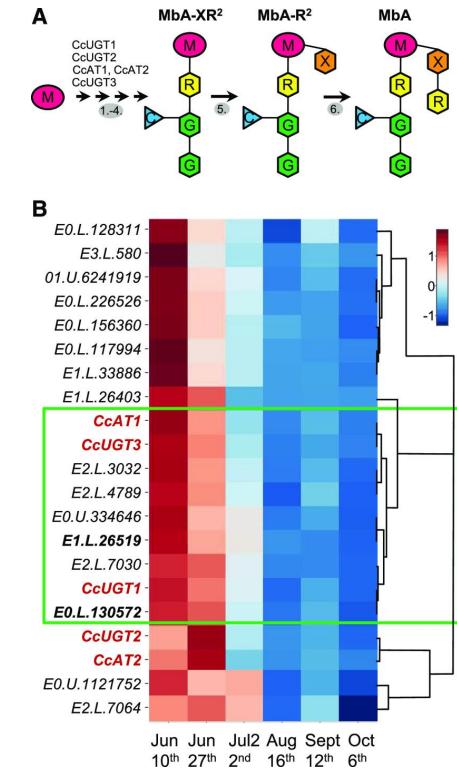
# Literature-based pathway search



old book: [https://de.wikipedia.org/wiki/Datei:Old\\_book\\_bindings.jpg](https://de.wikipedia.org/wiki/Datei:Old_book_bindings.jpg)  
sun burn: [https://en.wikipedia.org/wikis/Sunburn/media/File:Sunburn\\_Treatment\\_Practices.jpg](https://en.wikipedia.org/wikis/Sunburn/media/File:Sunburn_Treatment_Practices.jpg)

# Discovering a novel biosynthesis pathway: montbretin A

- Identification as inhibitor of  $\alpha$ -amylase in screen (enzyme assay)
- Identification of this myricetin-based metabolite in *Crocosmia x crocosmiiflora*
- Search for biosynthesis genes:
  - Co-expression analysis reveals candidate genes
  - Characterization of candidates through assays
  - Reconstitution of pathway through heterologous expression



# Summary

- Investigating gene functions through knock-outs (T-DNA, CRISPR-Cas9, EMS)
- Properties of model species and model pathways
- Flavonoid biosynthesis (anthocyanins, proanthocyanidins, flavonols)
- Mutual exclusion of anthocyanins and betalains
- Carotenoid biosynthesis and golden rice
- Metabolic networks and their evolution
- Transcriptional regulation (MYBs, bHLHs, ...)

# Time for questions!

# Questions

1. Which specialized metabolites are produced by plants?
2. How to reveal the function of a gene?
3. What are suitable mutagenesis methods in plants?
4. How to transform plants?
5. What is the difference between forward and reverse genetics?
6. Which characteristics of a model species are important?
7. Which characteristics of a model pathway are important?
8. Which biosynthesis pathways in plants are model systems?
9. Which transcription factors regulate the flavonoid biosynthesis?
10. What are the differences between anthocyanins and betalains?
11. Why is the carotenoid biosynthesis not an ideal model system?

1. Mutations in which genes could explain the loss of red flower pigmentation?
2. Loss of which genes could explain the absence of flavonols?

