

Prof. Dr. Boas Pucker
PBPM-BP-03

Availability of slides

- All materials are freely available (CC BY) - after the lectures:
 - eCampus: PBPM0 - Plant Biochemistry, Physiology and Molecular Biology (LEC)
 - GitHub: <https://github.com/bpucker/teaching/PBPM>
- Questions: Feel free to ask at any time
- Feedback, comments, or questions: [pucker\[a\]uni-bonn.de](mailto:pucker[a]uni-bonn.de)

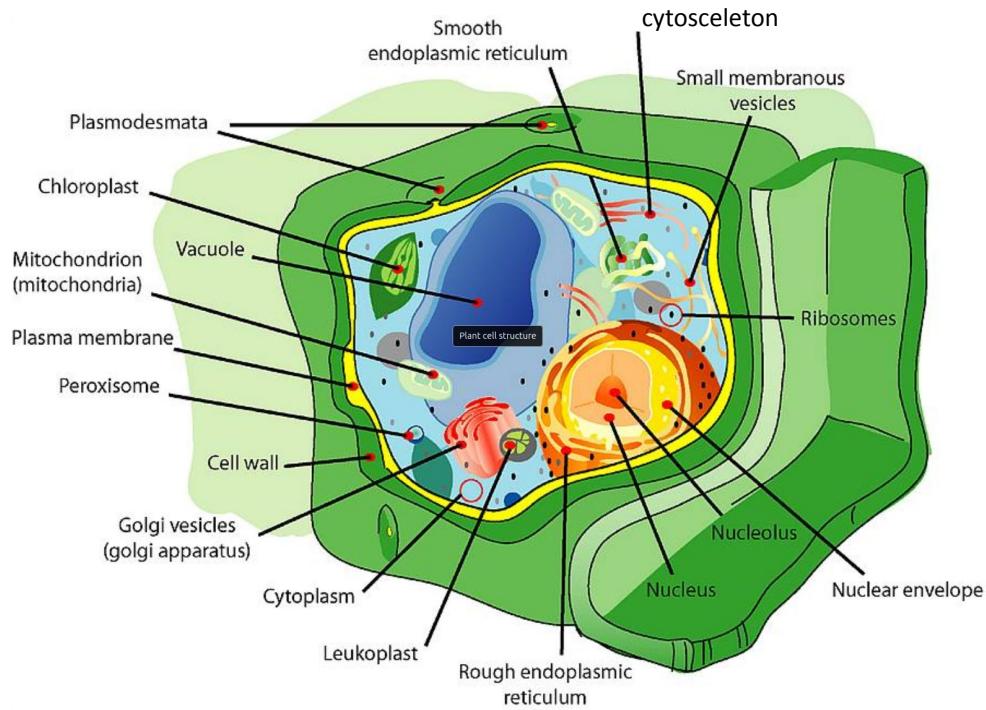


My figures and content can be re-used in accordance with CC BY 4.0, but this might not apply to all images/logos. Some figure were constructed using bioRender.com.

- History of DNA
- Sequencing methods: Sanger, Illumina, PacBio, ONT
- Genome sequence assembly
- Structural annotation
- Functional annotation

Components of a plant cell

- Cell wall
- Cell membrane
- Cytoplasm
- Nucleus
- Mitochondria
- Chloroplast
- Endoplasmatic reticulum (ER)
- Ribosomes
- Golgi apparatus
- Vacuoles

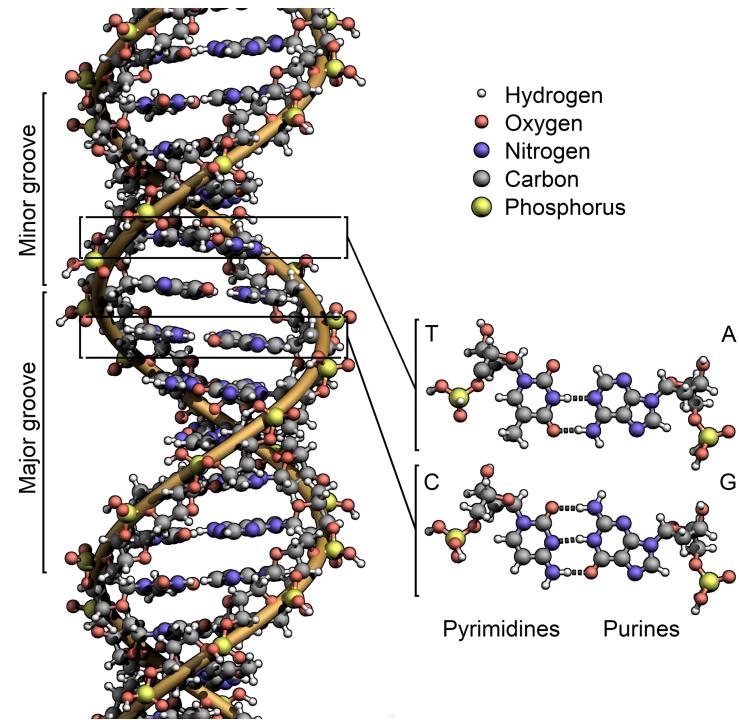


Molecular components of a plant cell

- Carbohydrates: sugar, cellulose, starch
- Inorganic ions: Na^+ , K^+ , Ca^{2+} , Mg^{2+} , Cl^- , PO_4^{3-} , Fe^{2+}/β^+ , Zn^{2+}
- Vitamins and coenzymes: vitamin C, vitamin E
- Lipids: membrane components; energy storage
- Proteins: enzymes, cytoskeleton,
- Nucleic acids: DNA, mRNA, rRNA, tRNA
- Specialized metabolites: flavonoids, betalains, terpenes, ...

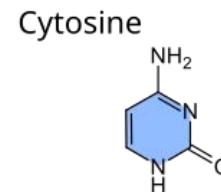
Nucleic acids

- Chain of nucleotides: sugar and phosphate backbone + bases
- Double strand of antiparallel complementary single strands
- Pairing of A-T and C-G
- Orientation: 5' > 3'
- Biological extension only at 3'-OH



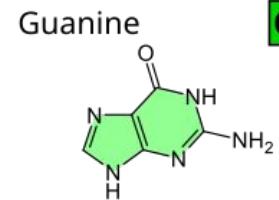
DNA building blocks

Cytosine (C) → dCTP (deoxycytidine triphosphate)



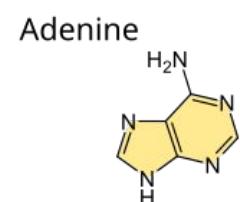
C

Guanine (G) → dGTP (deoxyguanosine triphosphate)



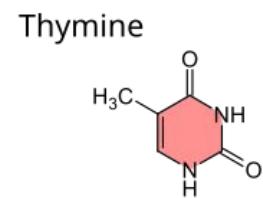
G

Adenine (A) → dATP (deoxyadenosine triphosphate)



A

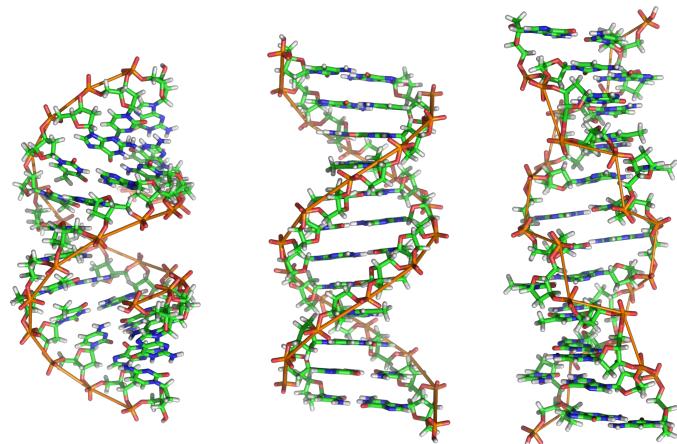
Thymine (T) → dTTP (deoxythymidine triphosphate)



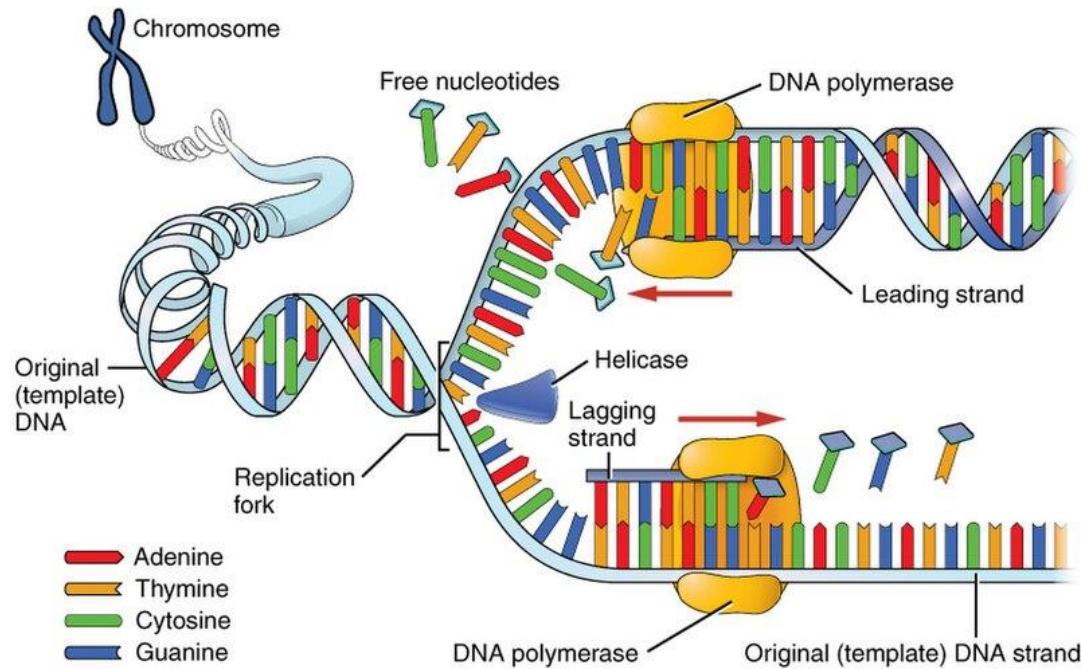
T

Different DNA structures

- **A-DNA:** Right-handed, shorter and wider; forms under dehydrated/low-humidity conditions; ~11 bp per turn; base pairs tilted and displaced from helix axis; appears in DNA–RNA hybrids
- **B-DNA:** Right-handed, most common physiological form; occurs under normal cellular hydration; ~10 bp per turn; bases nearly perpendicular, centered on helix axis
- **Z-DNA:** Left-handed helix; forms in high GC sequences or under high salt/supercoiling; Zig-zag sugar-phosphate backbone; ~12 bp per turn; linked to gene regulation

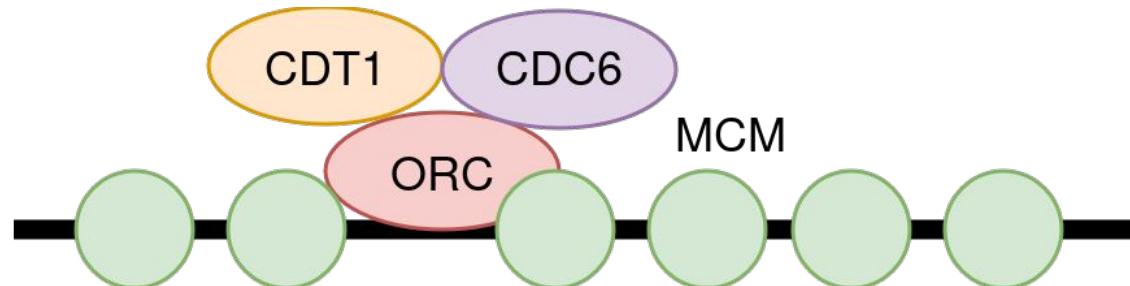


- DNA structure proposed a mechanism for duplication
- Helicase separates DNA strands to enable access
- Synthesis always from 5' to 3'
- Continuous synthesis vs. lagging strand synthesis



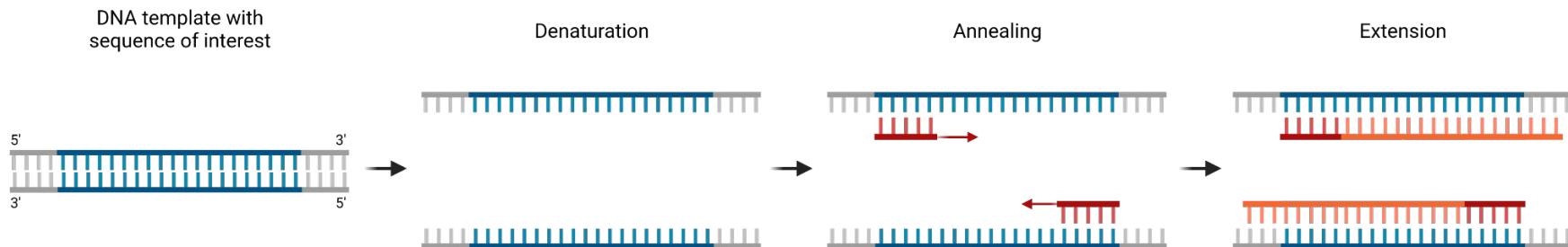
DNA replication origins

- Replication starts at AT-rich regions
- Epigenetic modification can determine replication origins
- Plant genomes are huge thus replication needs to start in multiple places

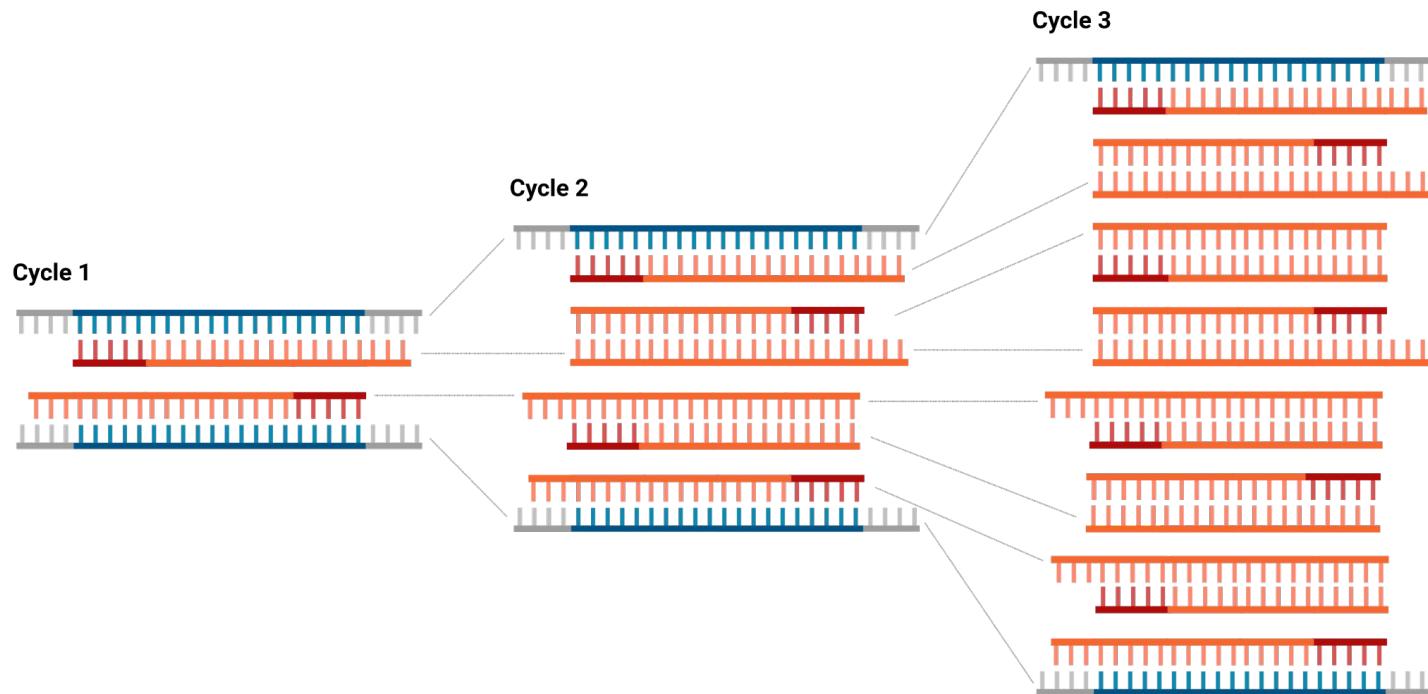




- Exponential amplification of defined DNA fragment
- PCR amplifies DNA fragment enclosed by primers
- Modification pattern of DNA is lost in amplification

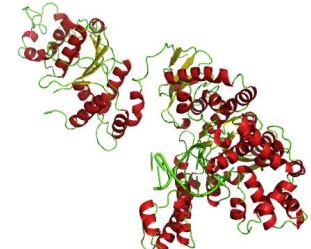
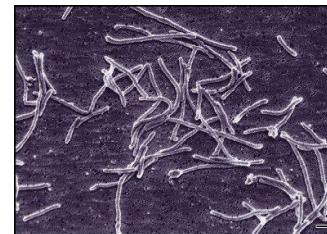


Exponential amplification



PCR: thermostable DNA polymerase

- Thermotolerant bacterial species isolated from hot springs in Yellowstone national park
- Conditions: 50-80°C
- *Thermophilus aquaticus* is source of a thermostable polymerase: Taq



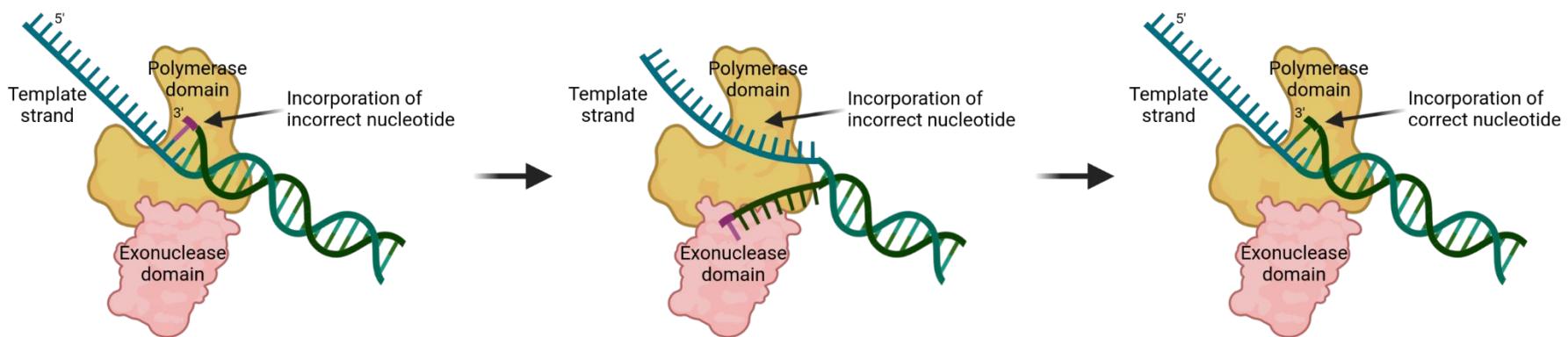
<https://commons.wikimedia.org/wiki/File:Taq.png> CC BY-SA 3.0

PCR: polymerase properties

- Thermostability: stable at high temperatures for a long time
- Extension rate: fast polymerases speed up experiments
- Fidelity: low error rate desired for specific applications
- Processivity: generation of large amplicons required for certain applications

PCR: polymerase proofreading

- Taq polymerase lacks proofreading activity (error rate: 10^{-6})
- Fusion polymerase offers proofreading activity (error rate: 10^{-9})



PCR: components

- Polymerase: enzyme responsible for synthesis of complementary DNA strand
- Template DNA: basis for synthesis of new DNA strand
- Primers: determine which region of the template DNA is amplified
- dNTPs: building blocks of new DNA strand
- Buffer (Mg^{2+} source): provides correct condition for the polymerase
- Water: most biological/biotechnological reactions take place in water

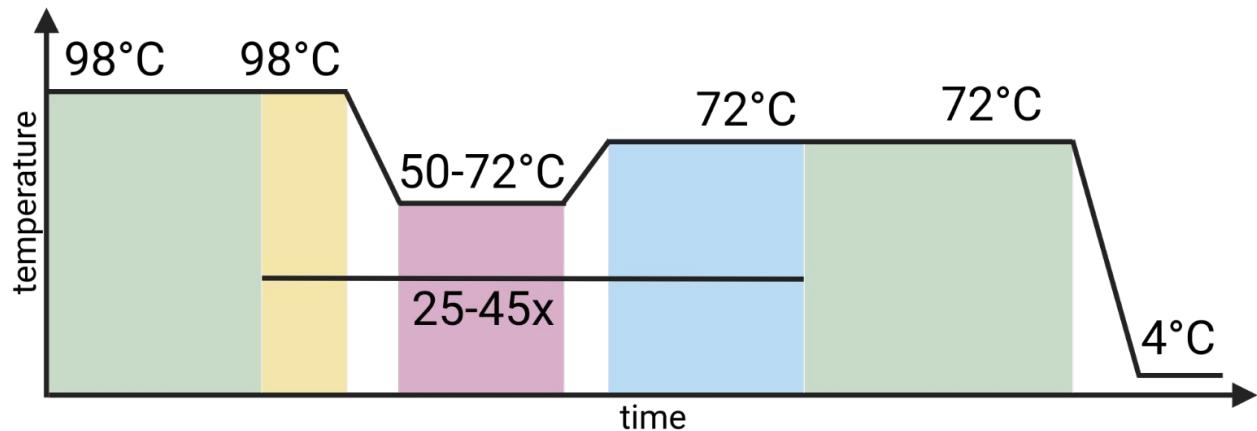
PCR: Thermocycler

- Quick temperature change poses technical challenge
- Heating the lid prevents condensation of evaporated solution
- Cooling function preserves PCR product after end of run
- Numerous programs can be saved for various applications
- Device also suitable for other incubation steps



PCR: temperature profile

- Denaturation:
separation of DNA
strands
- Annealing: primers
bind to DNA
- Elongation:
Polymerase
amplifies DNA

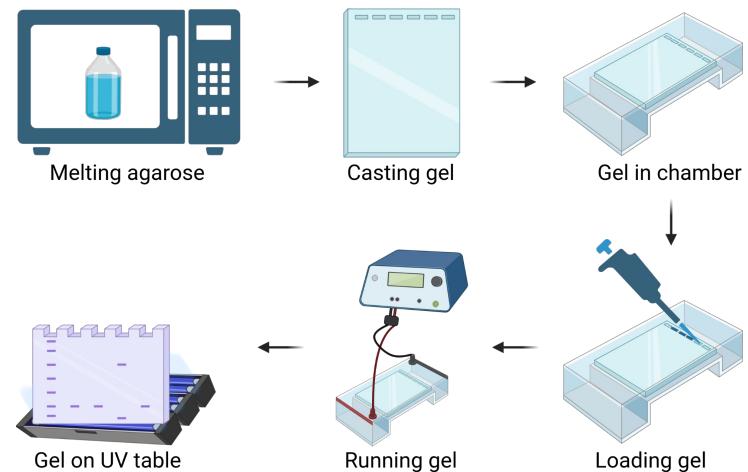
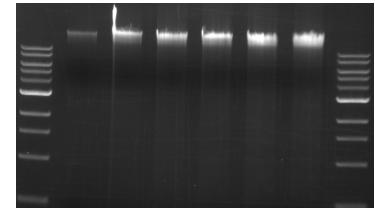


PCR: primer design

- Primer length determines specificity (about 20nt)
- GC clamp: 2xG/C at the 3' end
- Avoid self complementarity
- Avoid long homopolymers
- Primers must bind specifically to desired position on template

PCR: agarose gel electrophoresis

- Classical type of a PCR
- Running 25-45 cycles
- PCR is only exponential during the initial phases, but runs into saturation
- Result of an endpoint PCR is usually assessed on an agarose gel

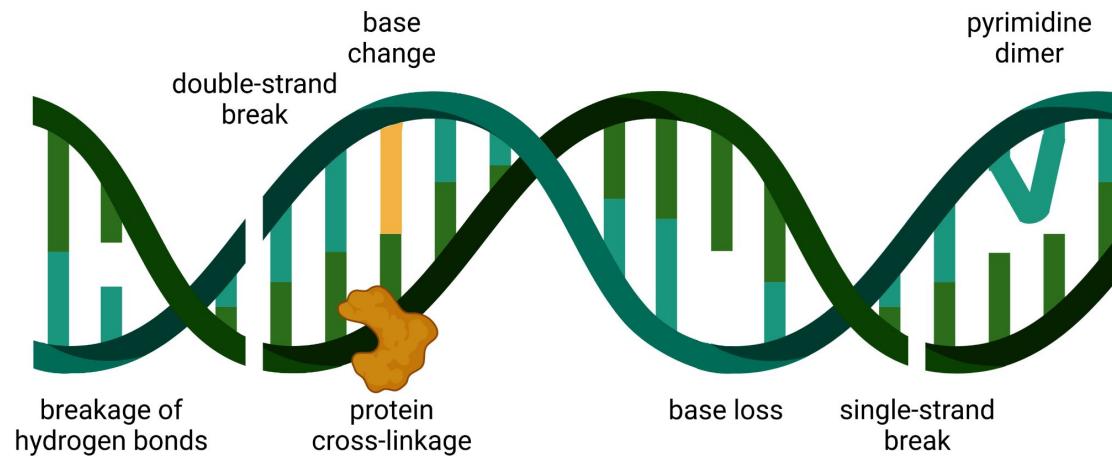


PCR: applications

- Genotyping: analyze the genetic basis of plants
- Cloning: amplify DNA pieces for genetic engineering
- Diagnostics: check for the presence of certain DNA in a sample
- qPCR: quantification of DNA or RNA (after reverse transcription)

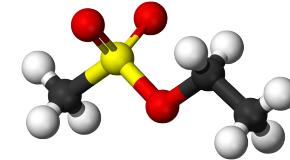
DNA mutations through radiation

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

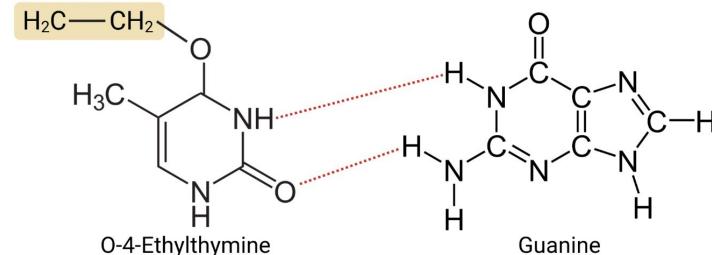
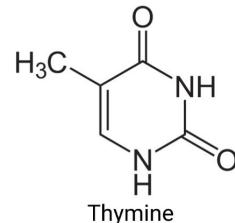
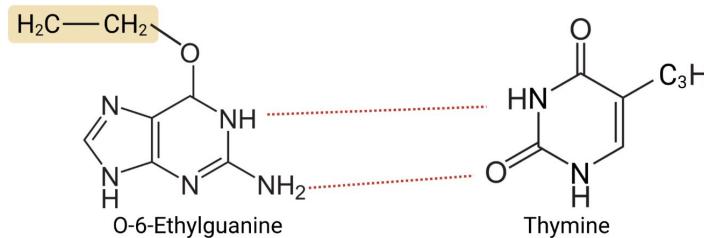
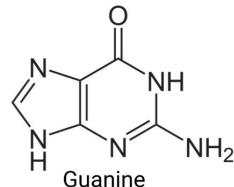


DNA mutations through EMS

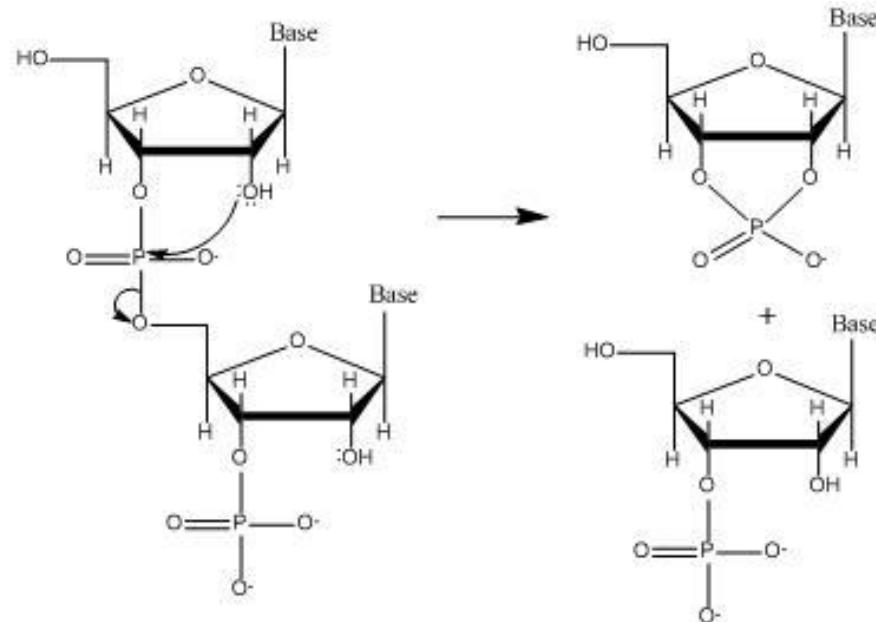
- EMS = Ethyl methanesulfonate
- EMS causes mostly point mutations
- Seeds are incubated in EMS for mutagenesis



EMS



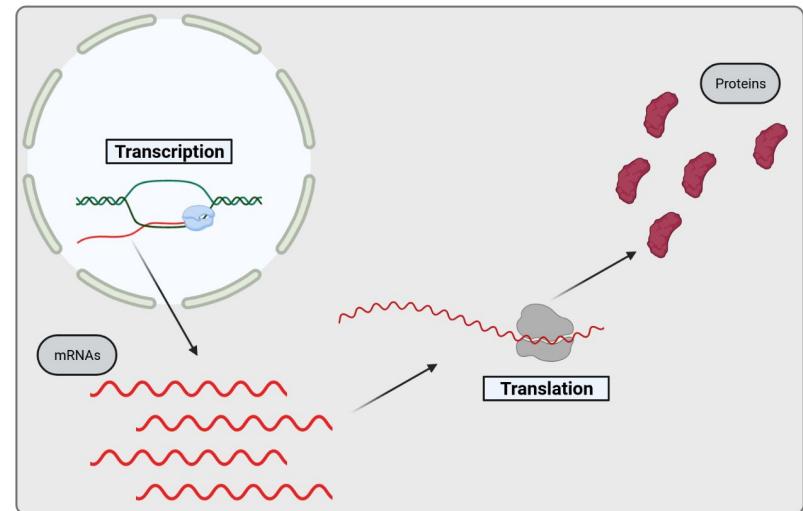
- RNA is less stable than DNA
- Additional hydroxy group at ribose allows more reactions
- Thymine replaced by Uracil



Protein biosynthesis (transcription)

Protein biosynthesis (overview)

- Gene expression = formation of gene product (i.e. a protein)
- Transcription of DNA by RNA polymerase and translation of mRNAs by polymerase
- Transcript abundance is often used as proxy (=gene expression)



- Transcription = generation of RNA based on DNA template
- RNA polymerase catalyzes the reactions
- RNA synthesis from 5' to 3'
- Initial transcript contains exons and introns
- Splicing is required to remove introns from transcript
- Addition of polyA tail (template independent)

Features of plants vs. bacteria

- Multiple RNA polymerases in eukaryotes
- mRNA: 5'-cap & poly-A tail in eukaryotes
- Spatial separation of transcription and translation in eukaryotes
- Exon-intron structure only in eukaryotes

- 5' splice site (donor site): Begins with GU (in most introns)
 - Branch point: An adenine (A) located near the 3' end of the intron
 - Polypyrimidine tract: A region rich in U and C before the 3' splice site
 - 3' splice site (acceptor site): Ends with AG
-
- Spliceosome assembly
 - First transesterification (lariat formation)
 - Second transesterification (exon joining)
 - Intron degradation

Alternative splicing leads to diversity

- Exon skipping
- Intron retention (common in plants)
- Alternative 5' or 3' splice sites

Gene structure (DNA)

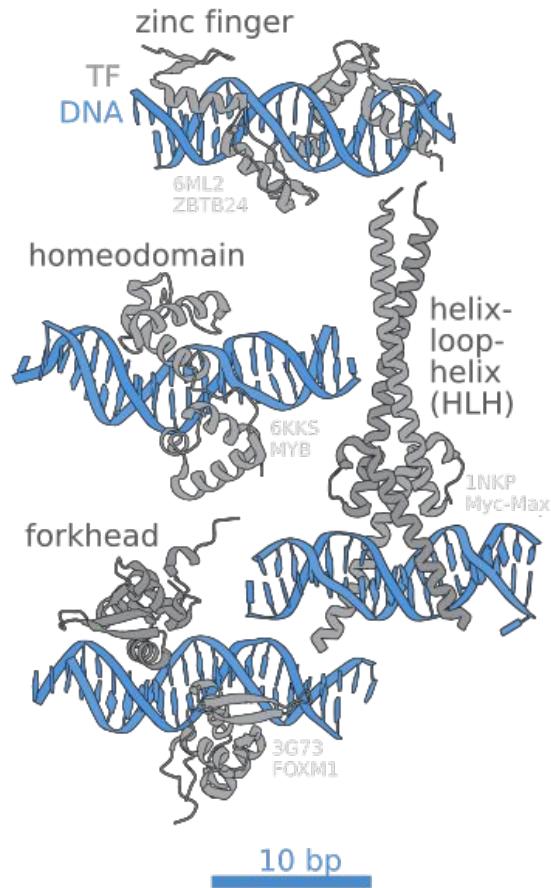


Environmental conditions influence splicing

- Modification (like phosphorylation) of splicing factors
- Alteration of the composition and activity of splicing machinery
- Modified splicing machinery affects splicing patterns
- Alteration of epigenetic marks affect speed of RNA polymerase II on DNA
- Elongation rate can favor different splicing outcomes (intron retention or exon skipping)
- Plants may use alternative splicing to sense temperature changes
- Example: proteins in ABA metabolism

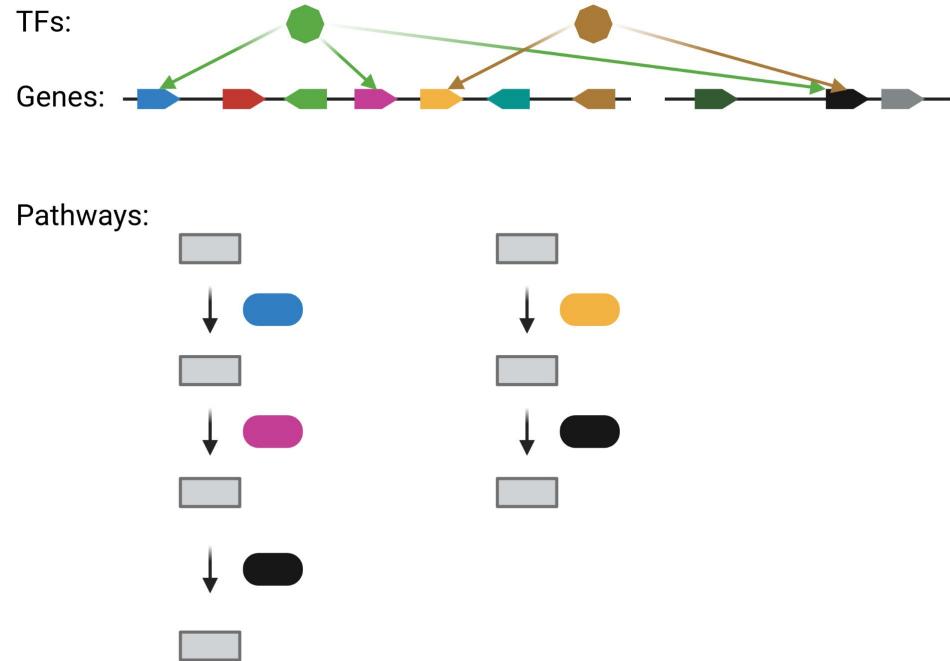
Transcription factors

- DNA-binding proteins facilitate transcription
- Interaction with RNA polymerase
- cis-regulatory elements (CRE) are TF binding sites



Transcriptional regulation 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



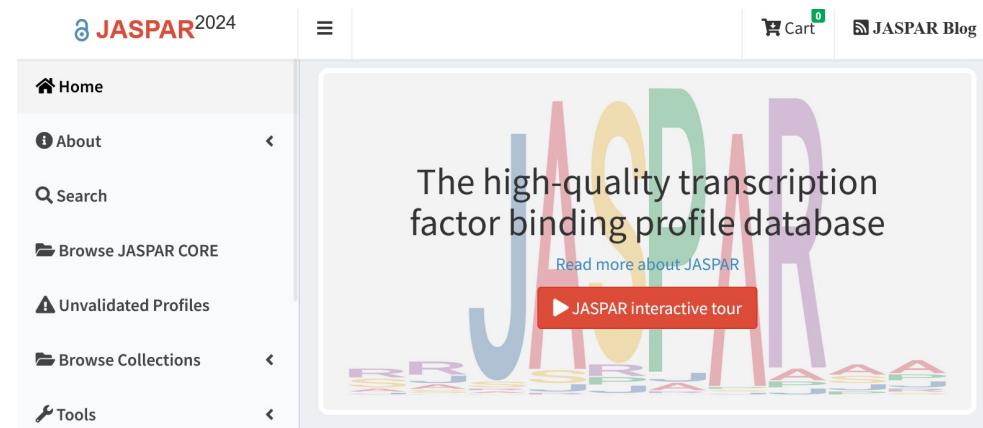
TF binding sites: cis-regulatory elements

- Cis-regulatory elements (CREs) are DNA motifs in the promoter region
- Composition of CREs determines TF binding
- Not all CREs are always accessible (methylation, repressors)

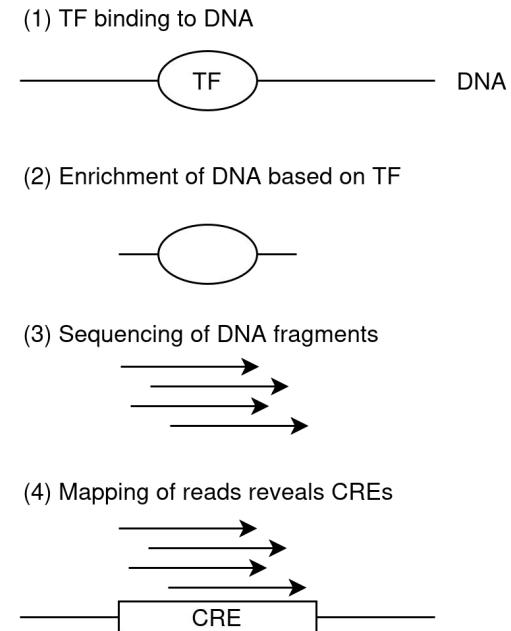


CRE identification

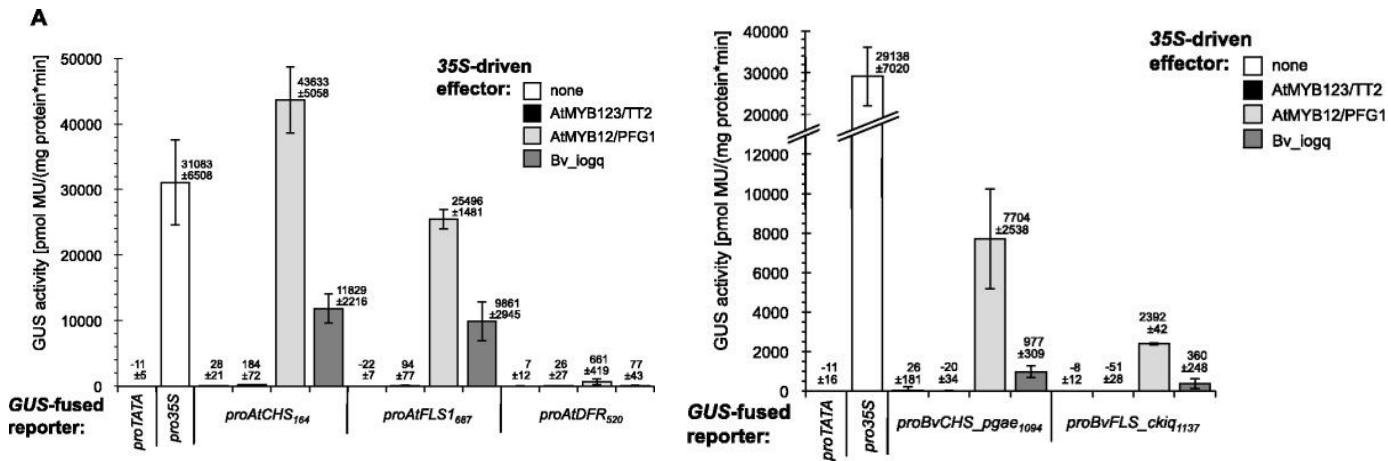
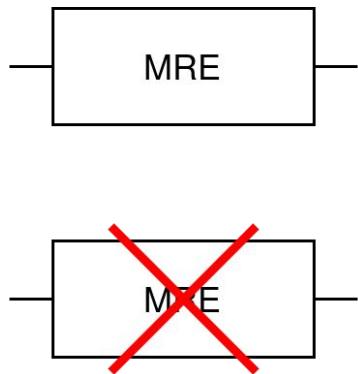
- Databases like PlantCARE and JASPAR provide information and allow identification in sequences
- JASPAR is largest collection of CRE and TF information
- Characterization of CRE via:
 - DAP-seq
 - Mutagenesis experiments
 - Y1H



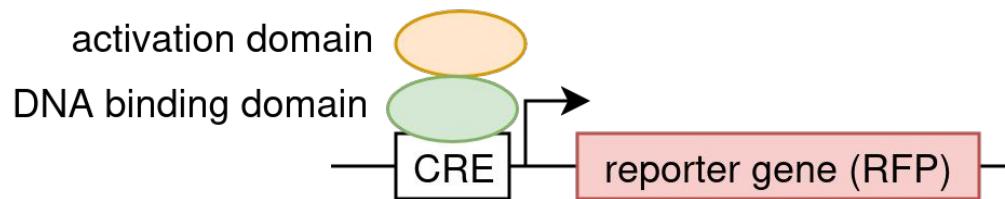
- Transcription factor of interest is heterologously expressed
- Purified protein is mixed with DNA of species of interest
- Fragmentation of DNA and enrichment based on protein
- Release of bound DNA and sequencing
- Read mapping to reference genome sequence reveals CREs



Testing mutated CREs

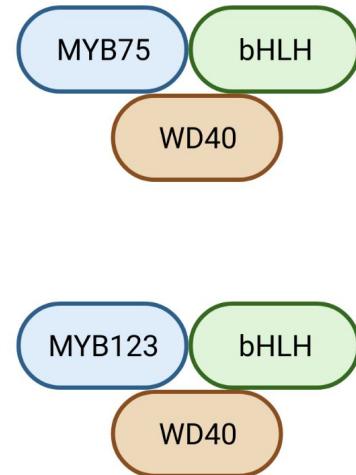


- Putative transcription factor is fused with an activation domain
- Putative CRE is placed upstream of a reporter gene
- If TF binds to CRE, activation of reporter gene is detectable
- Reporter gene is a selection/screening marker

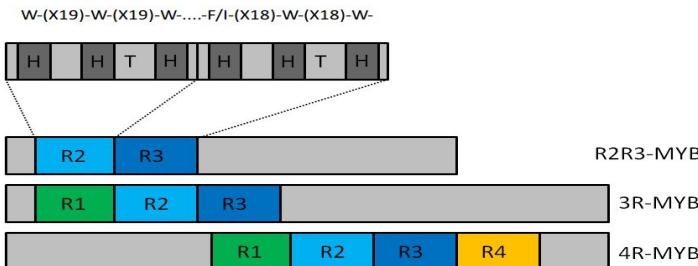


Example: MBW complex

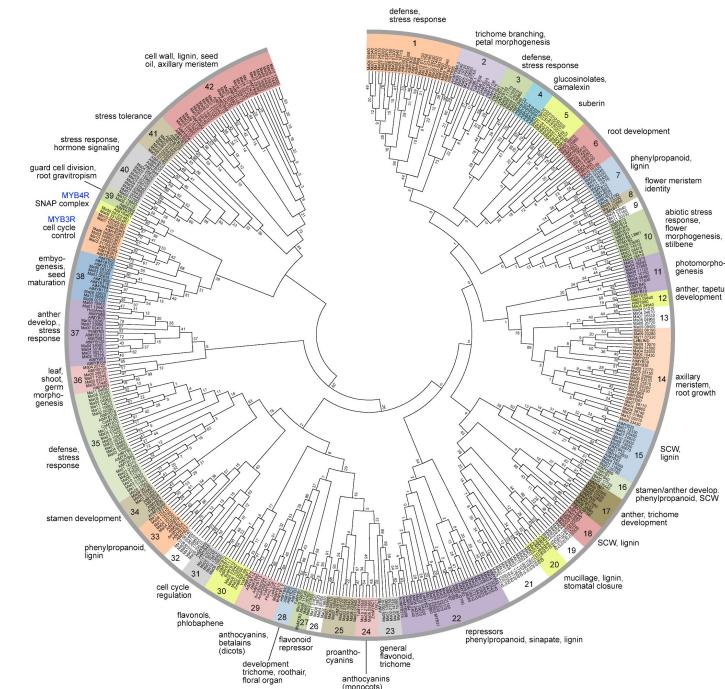
- Flavonoid biosynthesis is controlled by several transcription factor (TF) families
- MYBs are considered specific regulators of different flavonoid biosynthesis branches
- bHLHs are co-activators required for multiple branches
- WD40 (TTG1) is considered a scaffold protein for the connection of MYB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



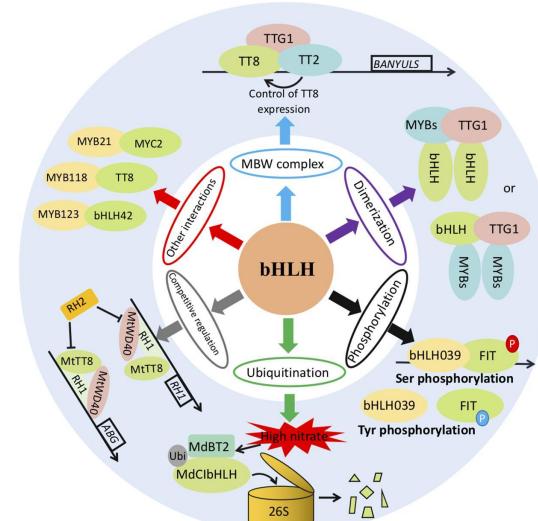
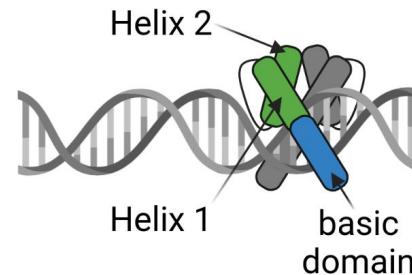
- MYB (myeloblastosis) is one of the largest TF families in plants
- Characterized by repeats (R)
- R2R3-MYBs are particularly important in plants
- Involved in numerous functions in plants



Based on Dubos et al., 2010

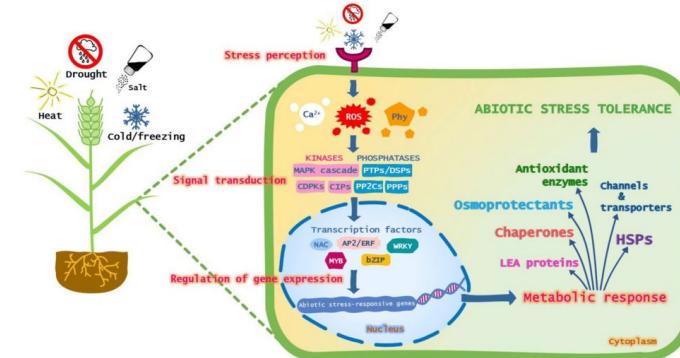
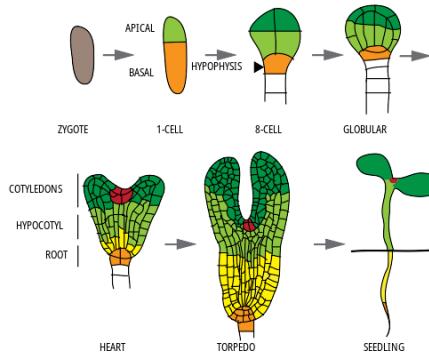


- bHLH (basic Helix-Loop-Helix) transcription factors form a large family in plants
- Transcriptional activation in cooperation with MYBs and independently
- bHLH transcription factors operate often in dimers



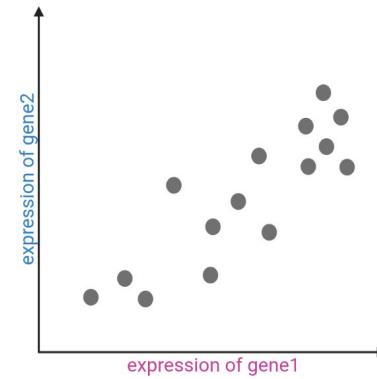
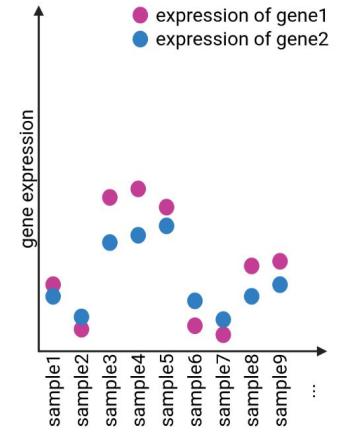
TFs: control of development vs. stress response

- WRKYs are important for stress response
- MADS box are important for development
- Plants are in trade-off to continue development vs. responding to stress



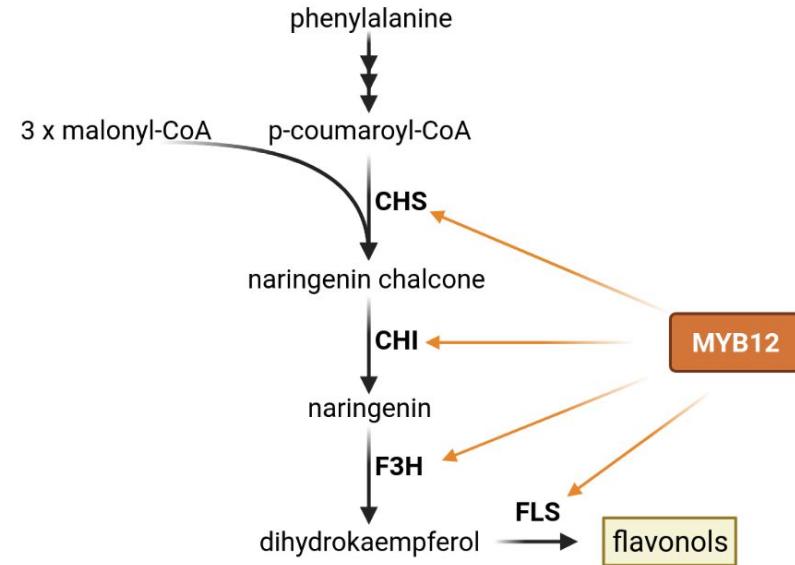
Concept of coexpression

- Genes can show similar expression values across numerous samples
- Reality usually results in similar, but not identical patterns
- Different samples could be different plant parts of plants cultivated under different conditions



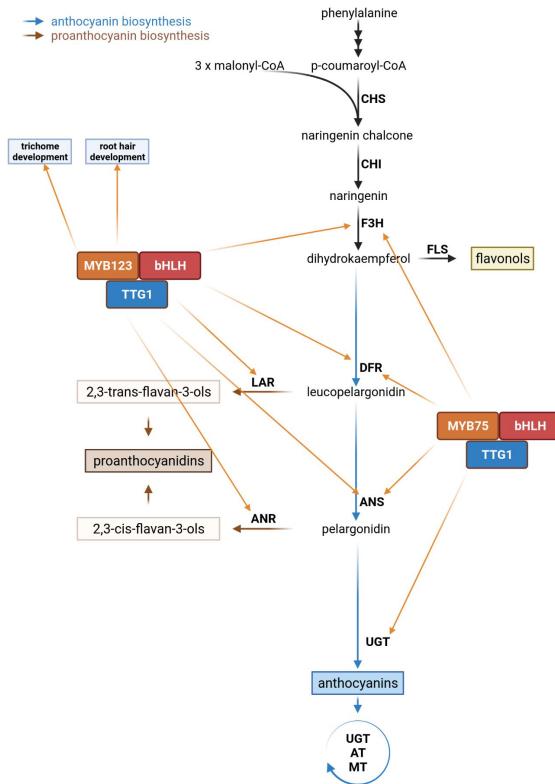
Molecular basis of coexpression

- Shared transcription factor can explain similar expression patterns
- Example: MYB12 controls the flavonol biosynthesis through activation of *CHS*, *CHI*, *F3H*, and *FLS*
- Expectation: *CHS*, *CHI*, *F3H*, and *FLS* should show a similar expression pattern



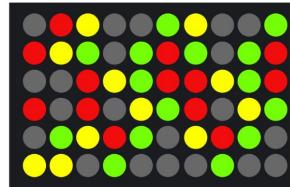
Nothing is perfect

- Genes can be regulated by multiple TFs (e.g. *DFR* by MYB123 and MYB75)
- TFs can control different processes (e.g. proanthocyanidins, trichome development, root hair development)
- Co-expression of TFs and structural genes in pathways is not perfect

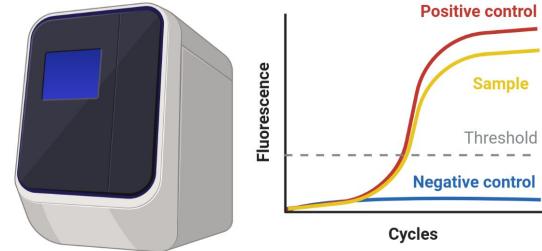


Types of expression data

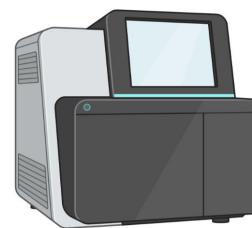
- Microarray



- RT-qPCR

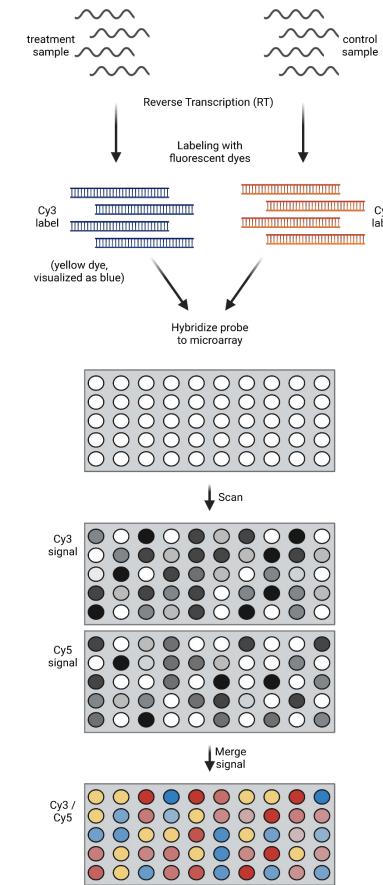


- RNA-Seq

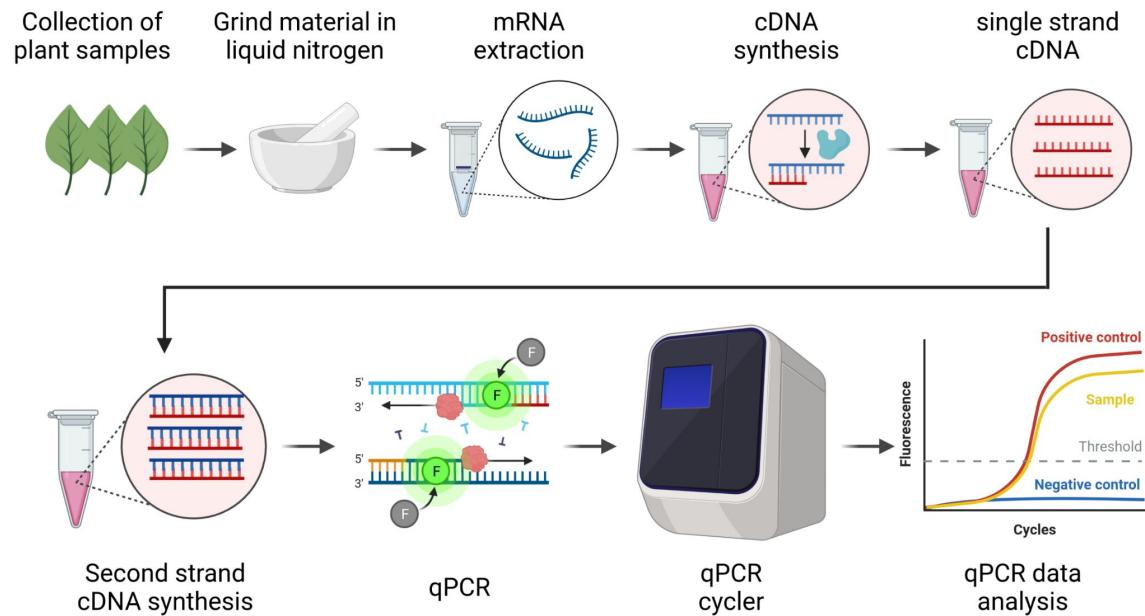


Microarray

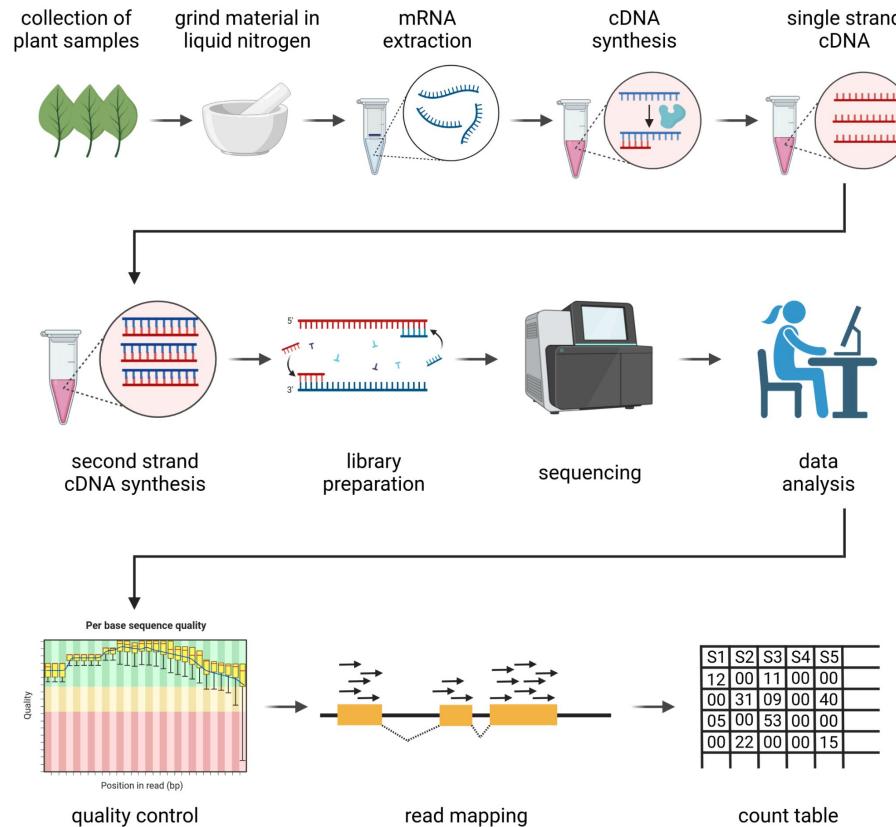
- Transcript abundances are compared
- Cy3 and Cy5 are fluorescent labels
- Fluorescence intensity indicates transcript abundances
- Dynamic range is small due to saturation of signal
- Only genes represented on the microarray can be studied
- High investment costs for microarray generation



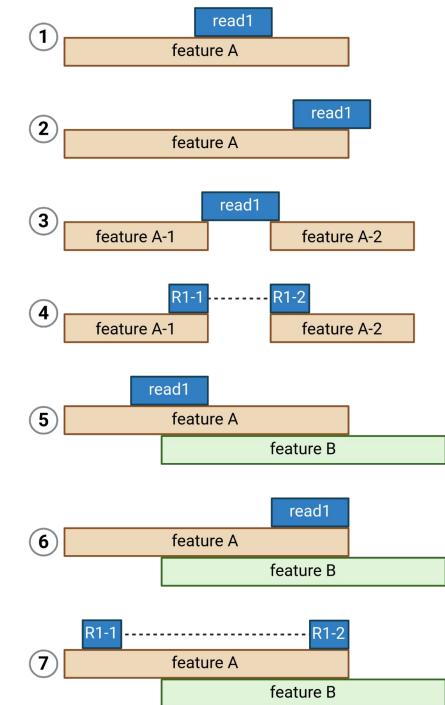
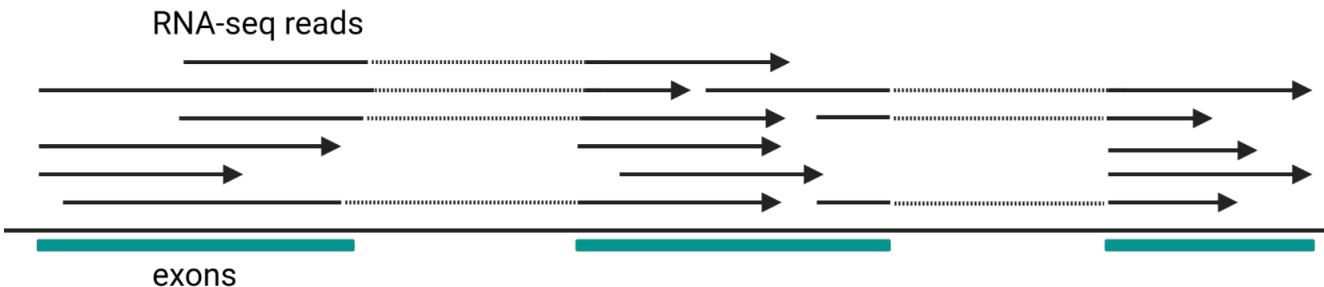
- Quantification of cDNA based on incorporation of fluorescent dyes



RNA-Seq (overview)

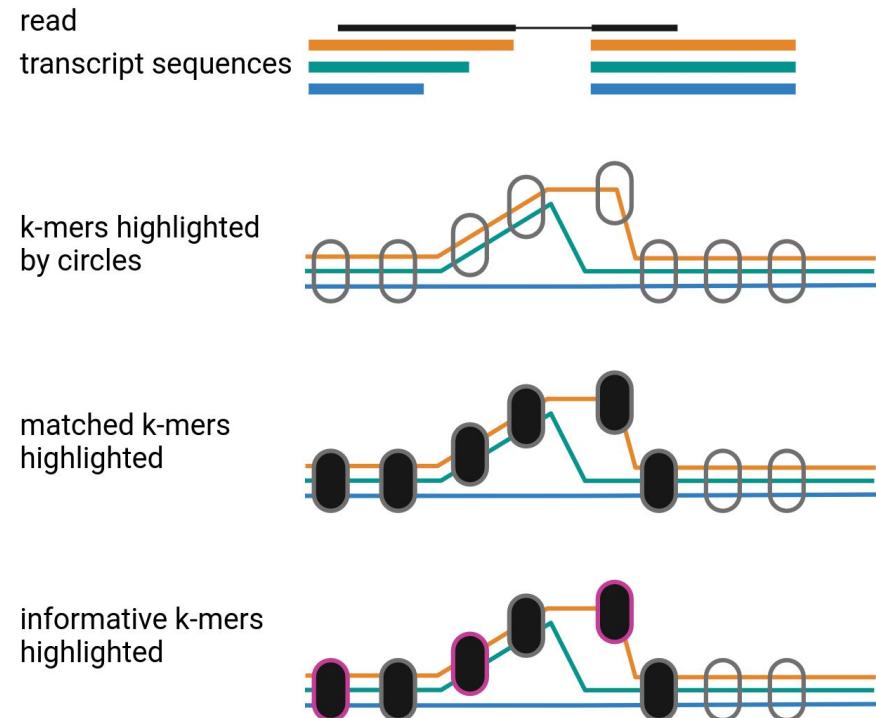


RNA-seq: read mapping



RNA-seq: mapping-free analysis

- Much faster than proper read mappings with STAR/HISAT2
- Context of k-mers is considered and not only individual k-mers



- Components of a plant cell
- Nucleic acids
- DNA replication / PCR
- Transcription & splicing
- Transcription factors and cis-regulatory elements

Time for questions!

Questions

1. Please list 5 important components of a plant cell!
2. Please explain 5 important properties of the DNA!
3. How does a PCR work?
4. How can you mutagenize DNA?
5. What are the differences between replication of leading and lagging strand?
6. What is transcription?
7. Why are transcription factors important in plants?
8. Which steps determine the amount of a protein in a plant cell?
9. Which methods can be used to investigate the proteome of a plant cell?