CENTER OF PLANT SCIENCES

Sant'Anna
Scuola Universitaria Superiore Pisa

Introduction to Genomics 2024

Different Flavours in Genomics

-

Part B

Svenja Mager

svenja.mager@santannapisa.it

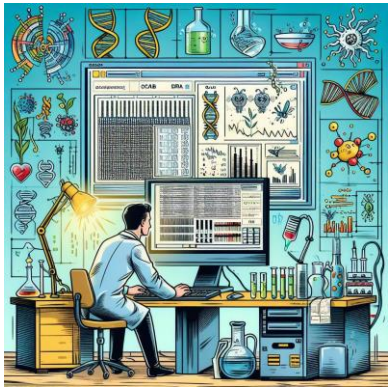
Transcriptomics Techniques Today

- Which are the main techniques used?
- How do Microarrays work?
- How does RNA Sequencing work?
- How to analyze transcriptomic data

Transcriptomics Technologies - Today

Where did we leave off last time?

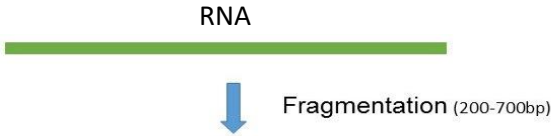
- Two dominant contemporary techniques: Microarrays and RNA-Sequencing (developed in the 1990's and 2000's, respectively)
- RNA-Seq overtook microarrays as dominant technique in 2015 due to several advantages
- Many different aspects can be changed in an RNA-Seq experiment and must be decided according to the needs for each project



Transcriptomics Technologies – RNA-Sequencing

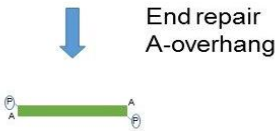
General steps (simplified)

Bulk RNA or mRNA, tissue or single-cell

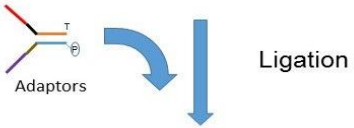


Fragment size: Illumina usually around 100bp, with long-read techniques up to 50kb possible

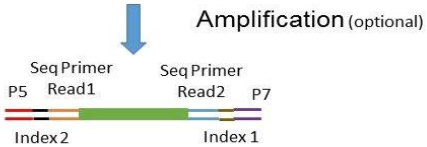
Single-end or paired-end, stranded or unstranded libraries possible



Ligation of adaptors to which primers can anneal (here: paired-end)

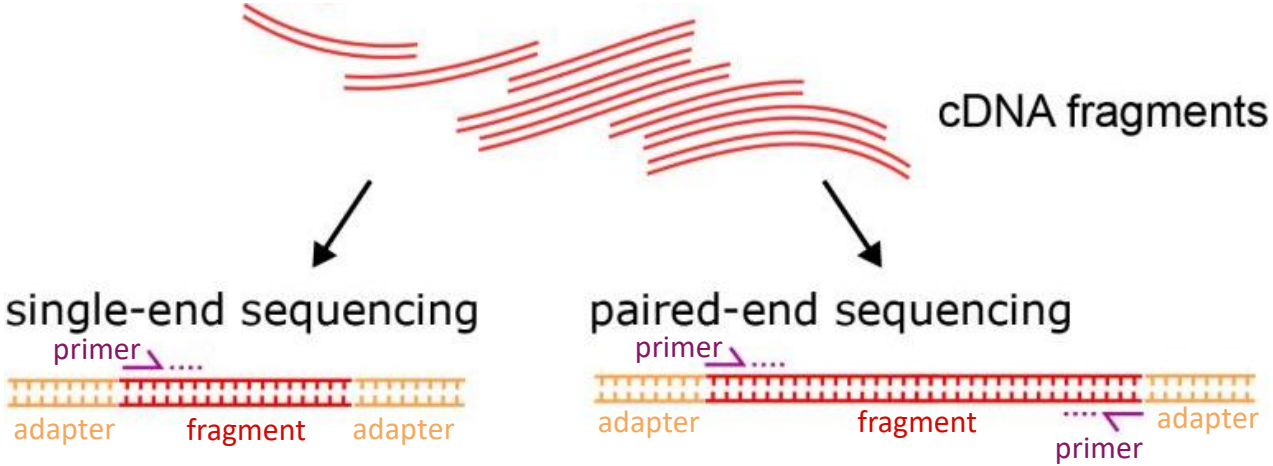


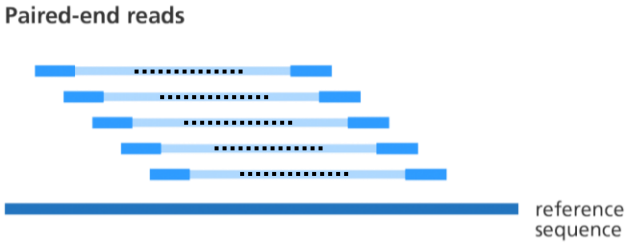
Amplify RNA material via PCR



Anneal primers and sequence reads

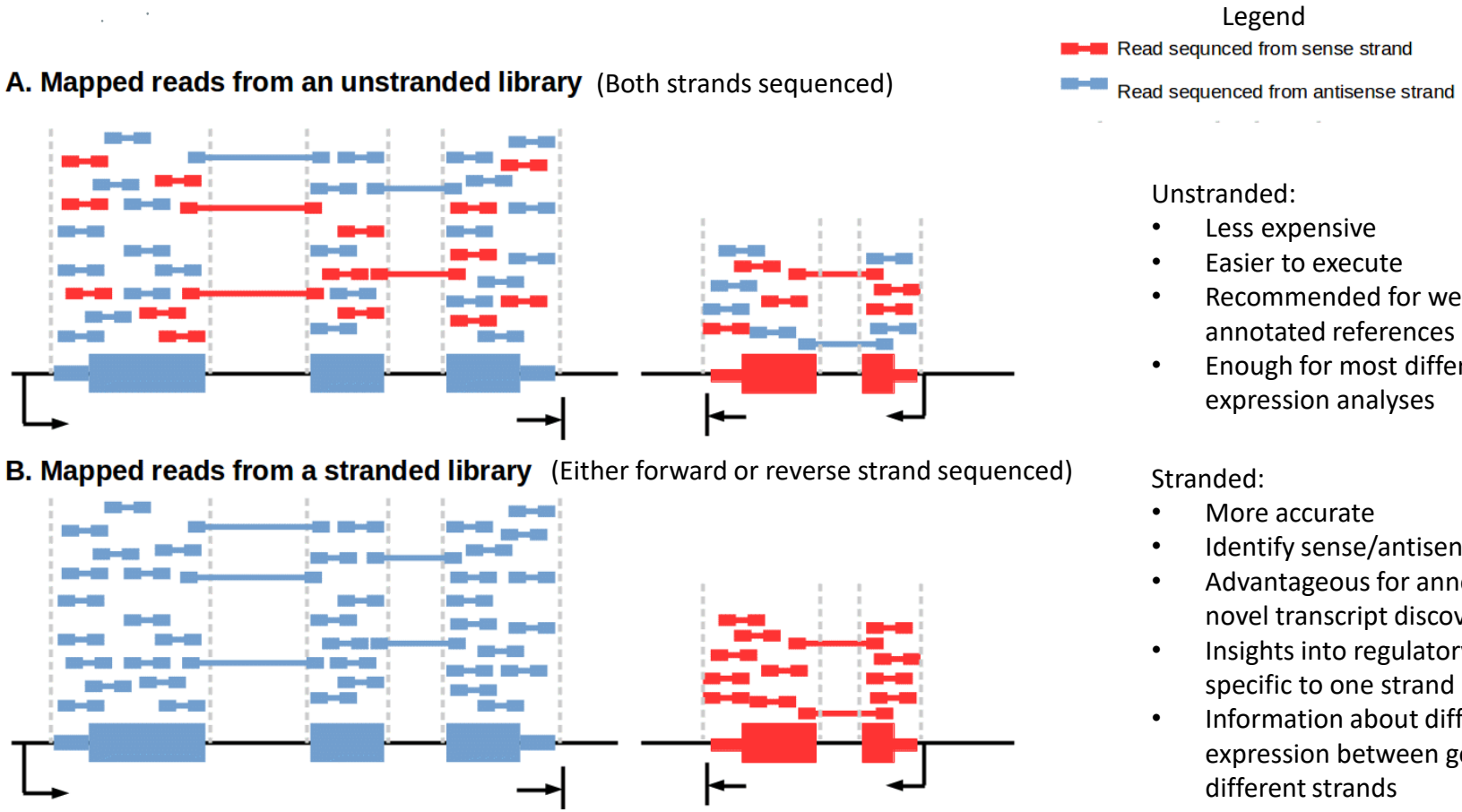
Transcriptomics Technologies - RNA-Sequencing





Transcriptomics Technologies - RNA-Sequencing

Unstranded vs Stranded Sequencing



Unstranded:

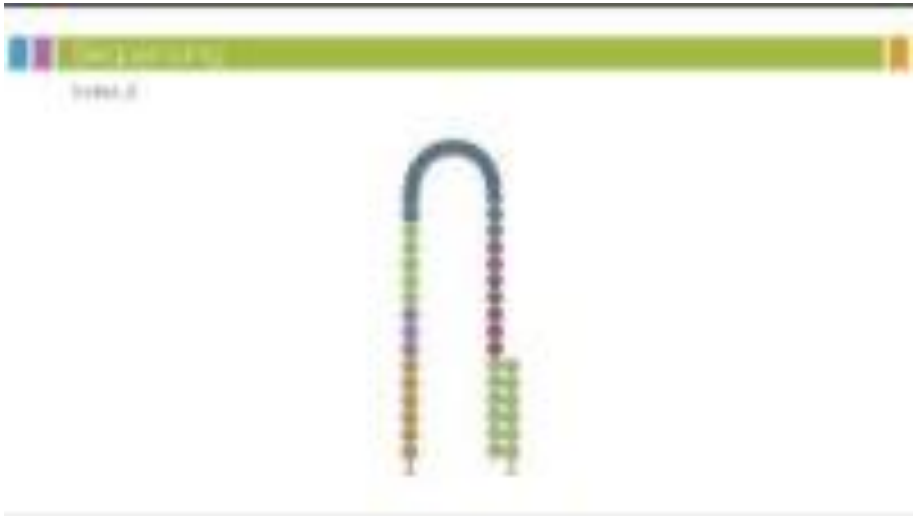
- Less expensive
- Easier to execute
- Recommended for well annotated references
- Enough for most differential expression analyses

Stranded:

- More accurate
- Identify sense/antisense transcripts
- Advantageous for annotation and novel transcript discovery
- Insights into regulatory mechanisms specific to one strand
- Information about differential expression between genes on different strands

Transcriptomics Technologies - RNA-Sequencing

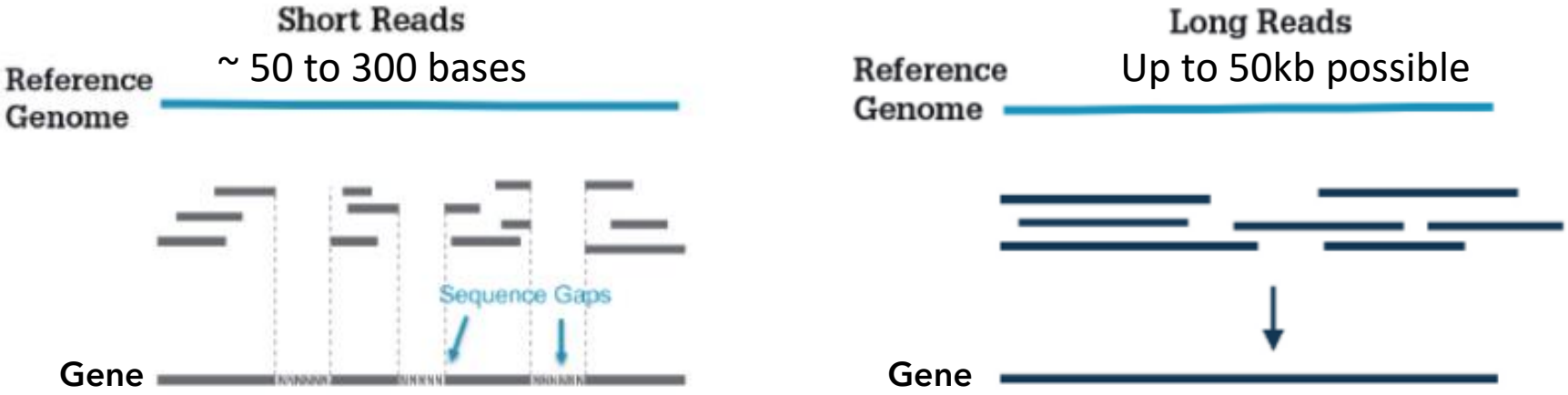
Illumina sequencing by synthesis



Transcriptomics Technologies - RNA-Sequencing

Short- vs Long-read Sequencing

Tremendous advancement for transcript detection **via long-read sequencing**
→ One read = one transcript (full-length cDNA) has become reality

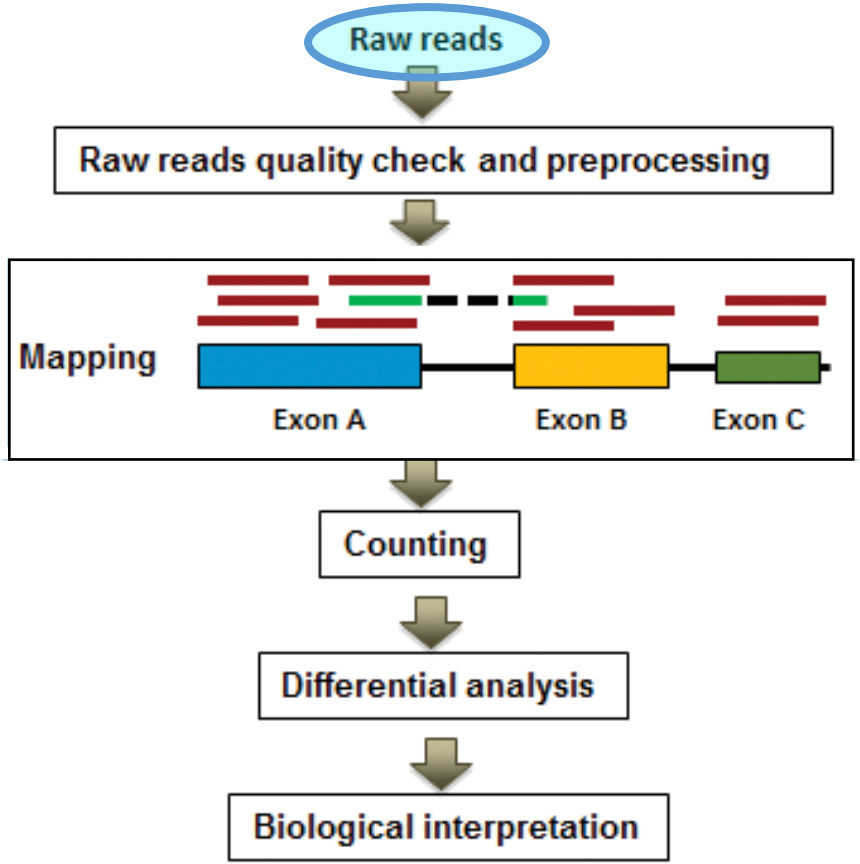


Transcriptomics Technologies - RNA-Sequencing

For comparison: long-read sequencing nanopore technique

Transcriptomics Technologies - RNA-Sequencing

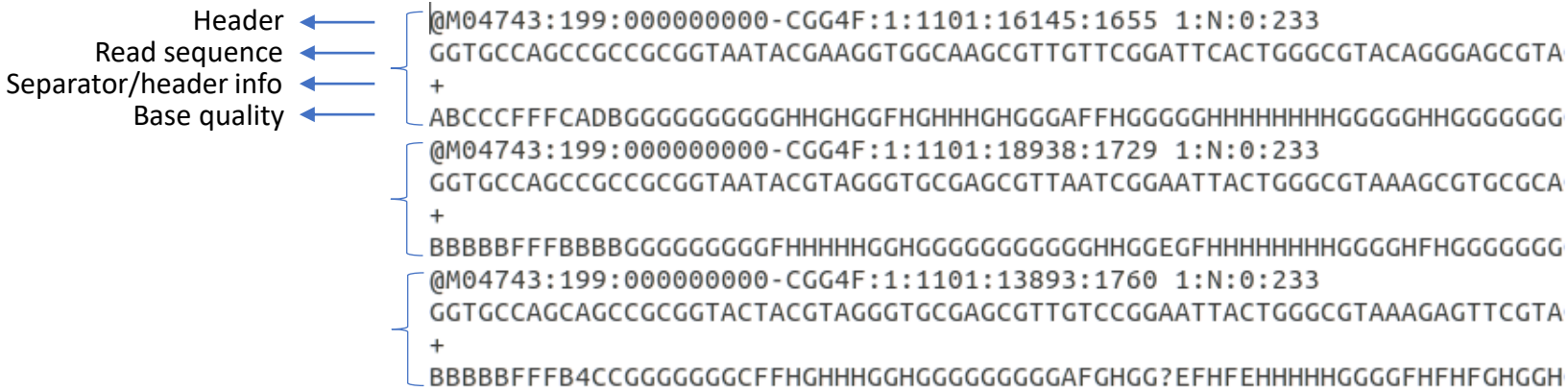
Data Analysis



Transcriptomics Technologies - RNA-Sequencing

Data Analysis

You receive data in fastq format:

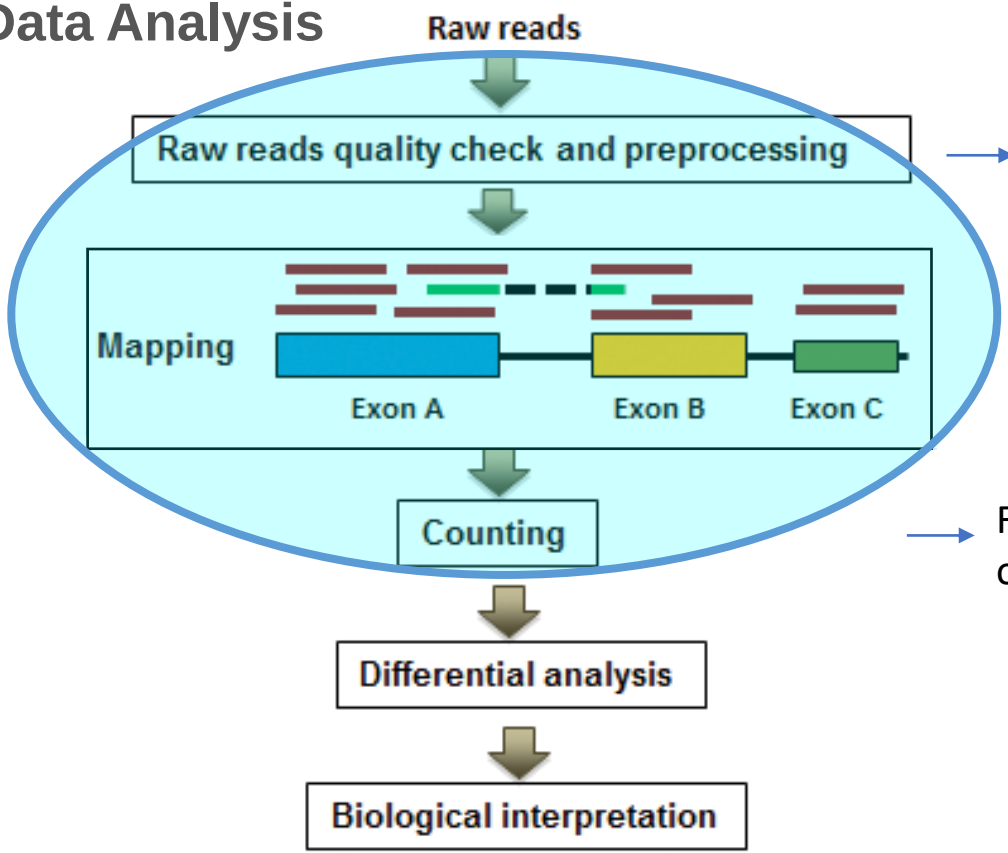


Header can contain information about

- the machine used
- flow cell id
- Lane
- Coordinates
- read direction (forward or reverse)
- …

Transcriptomics Technologies - RNA-Sequencing

Data Analysis



Filter your fastq reads (e.g. adapter sequences, short reads, low quality reads)

Filtered reads are mapped to the reference genome

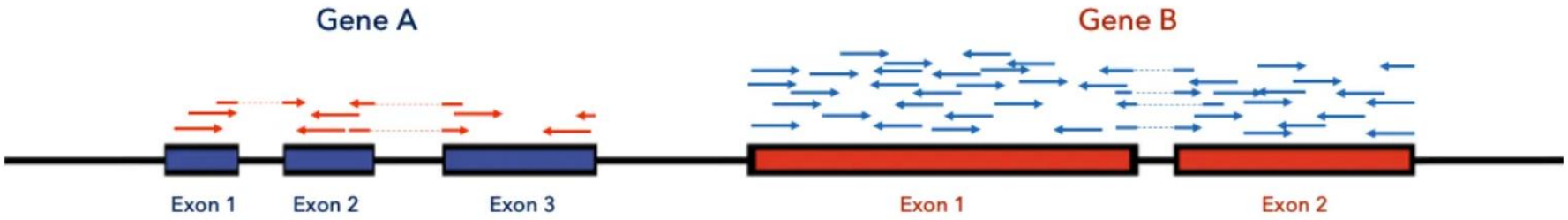
For each gene, the number of mapped reads is counted

Transcriptomics Technologies: RNA-Sequencing

Data Analysis

Read Counting

Assumption: the number of mapped reads for each gene is proportional to the expression of RNA



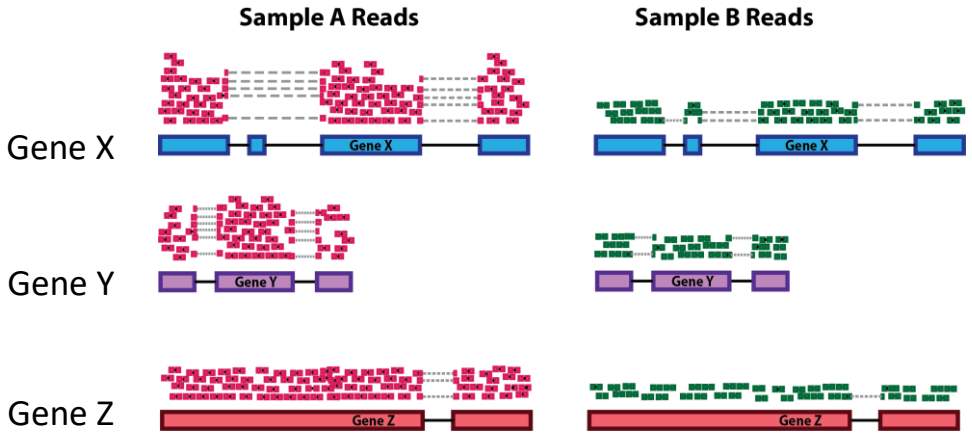
→ NOTE: counts have to be **normalized** according to the question to be answered

Transcriptomics Technologies: RNA-Sequencing

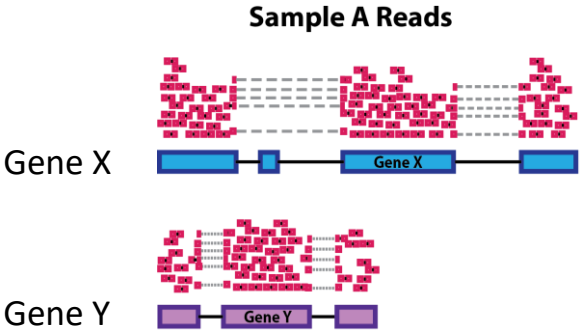
Data Analysis

Read Count Normalization

If you want to compare expression of a certain gene between two samples, you must normalize for sequencing depth



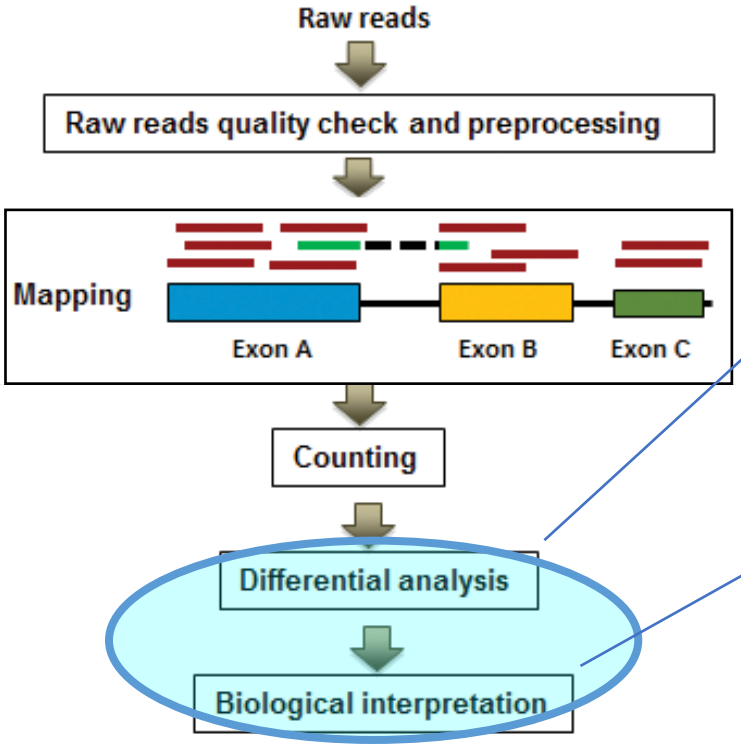
If you want to compare expression of two genes within the same sample you must normalize for gene length



Transcriptomics Technologies: RNA-Sequencing

Data Analysis

Differential expression analysis



Compare gene expression between two conditions, e.g. plants grown under normal conditions compared to heat stress conditions, or healthy cells vs diseased cells

Which genes are up-regulated, which are downregulated under certain conditions?
→ For example, make GO enrichment analysis to check for overrepresented gene categories

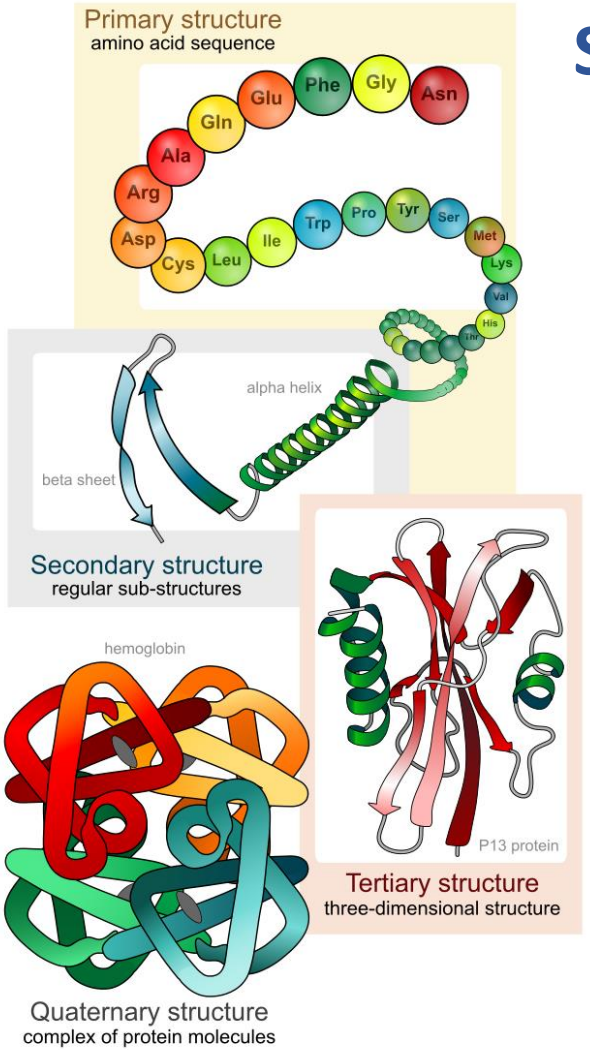
Can give insights into adaptability of plants or genes causing a diseases

Enables more precise breeding or finding cures for diseases

Structural Genomics

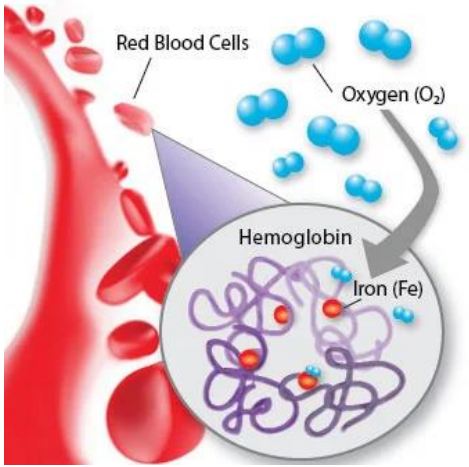
- What do Structural Genomics analyze?
- How?

Structural Genomics



Traditional protein structure determination efforts worked on single proteins and helped create protein structure databases

→ Most important method:
X-ray crystallography



Structural genomics aims at determining the different structures and their function of **all** proteins

Benefits from the databases and the large number of sequenced genomes, can predict and annotate protein structures computationally

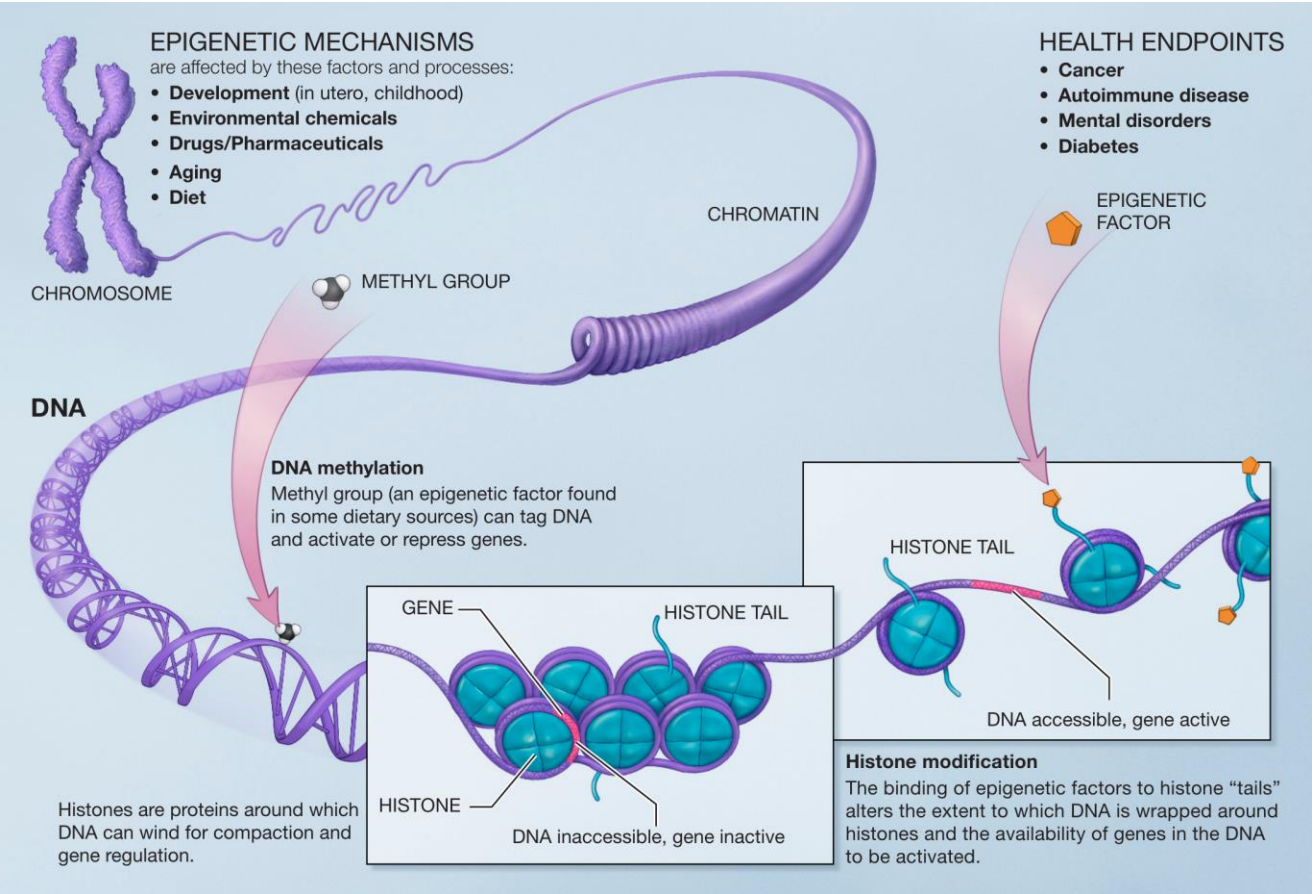
Epigenomics

- What are Epigenetic Mechanisms?
- Histone Modifications and how they are analyzed
- DNA methylation and how is it analyzed

Epigenomics

epi- (Greek ἐπι- "over, outside of, around")

Epigenetics is the study of how cells control gene activity without changes in the DNA sequence.

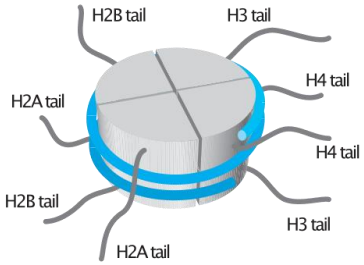


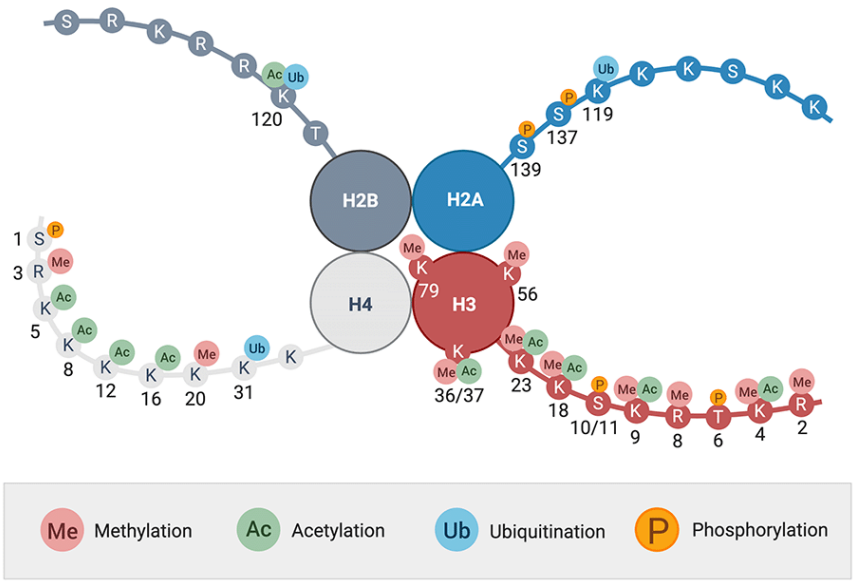
Many different functions!

In plants epigenetic mechanisms were shown to be involved in:

- Pathogen response
- Genome stability
- Protection from DNA damage
- Preserving nucleotide sequences
- Heterosis
- Imprinting
- Paramutation
- Regulation of transposable elements
- Gene expression regulation
- Recombination distribution and frequency

Epigenomics - Histone modifications





Histone residues can be modified in several ways

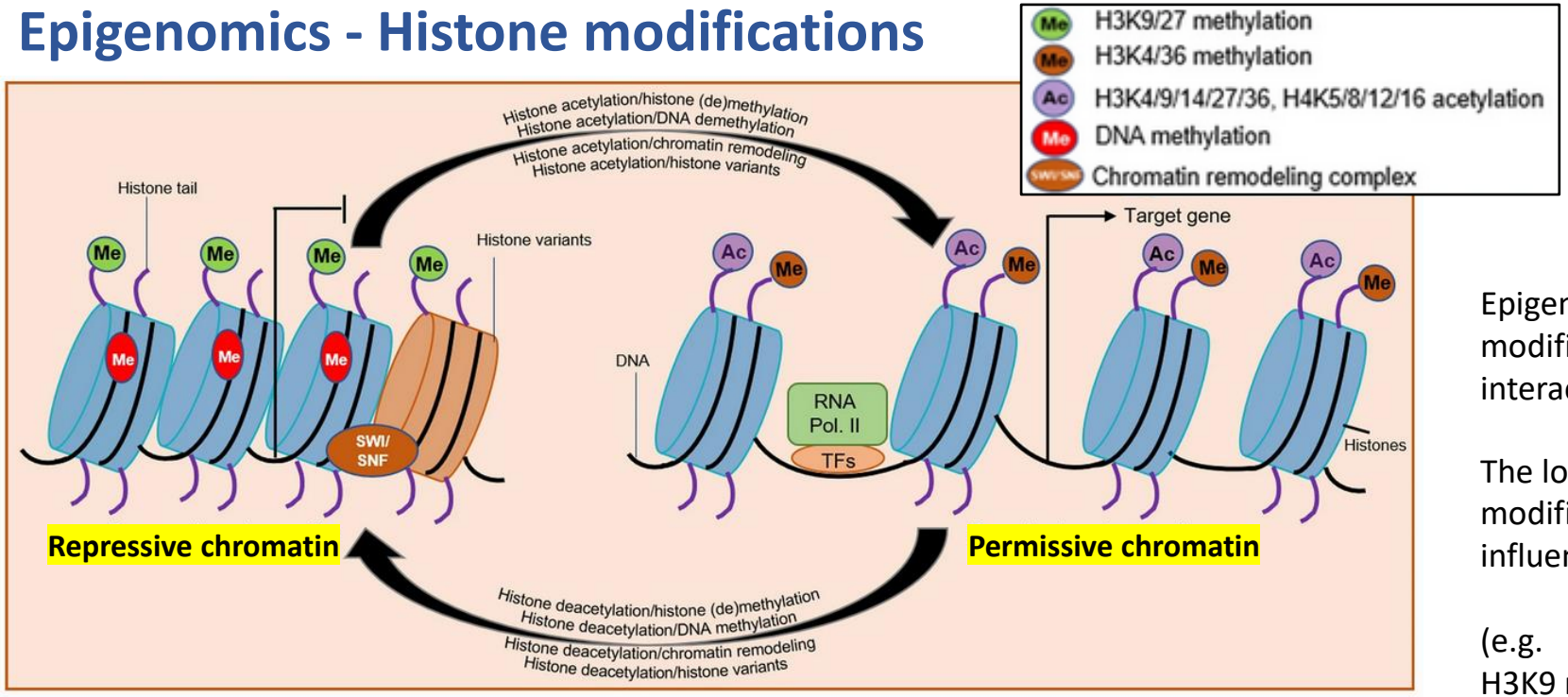
Most common:

- Acetylation
- Methylation
- Phosphorylation
- Ubiquitination

Further (less common):

- O-GlcNAcylation
- Sumoylation
- ADP-ribosylation
- Citrullination
- proline isomerization

Epigenomics - Histone modifications



Modulation of chromatin status by histone acetylation dynamics and other modifications. A combination of acetylation and methylation of different lysine residues in nucleosomal histones may cause expression (permissive chromatin) and repression (repressive chromatin) of genes. Besides, the combination of histone deacetylation and DNA methylation may also convert permissive chromatin to repressive chromatin which generally leads to gene repression. The regulation of gene expression is also affected by chromatin remodelling complex (such as SWI/SNF complex). Ac represents acetylation on lysine 4/9/14 /27/36 of H3, and lysine 5/8/12/16 of H4 histone (H3K4/9/14/27/36 and H3K5/8/12/16), Me in the green coloured circles represent methylation on lysine 9/27 of H3 histone (H3K9/27), Me in the brown-coloured circles represent methylation on lysine 4 and 36 of histone H3 (H3K4/36), Me in the red coloured circles represent DNA methylation (cytosine methylation), SWI/SNF: SWITCH/SUCROSE NONFERMENTING chromatin remodelling complex. *TFs* transcription factors, *RNA Pol. II* RNA polymerase II. This is a simplified figure describing only combinatorial effect of histone acetylation and histone/DNA methylation and chromatin remodelling

Epigenetic modifications interact

The location of modifications influence the effect

(e.g. H3K9 methylation = methylation on Lysine 9 on Histone subunit H3)

Kumar et al 2021

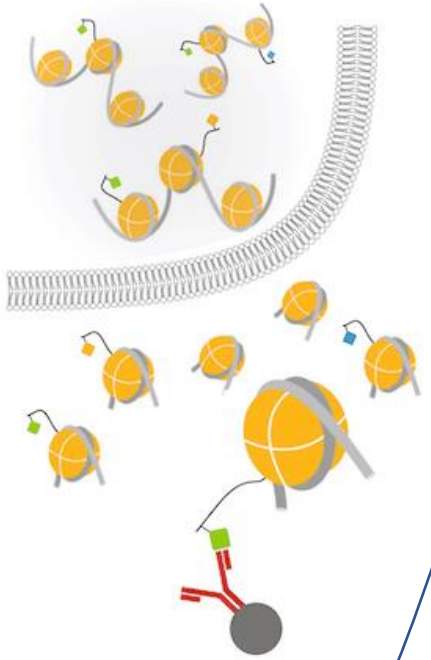
Epigenomics - Histone modifications

ChIP-Seq (chromatin immunoprecipitation sequencing)

Fix and lyse cells

Fragment chromatin

IP for target epigenetic factor



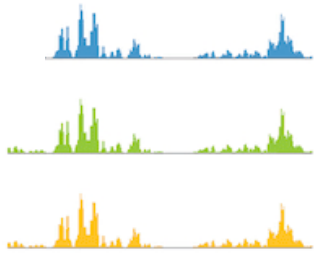
Antibodies are attached to magnetic beads and isolated using a magnet

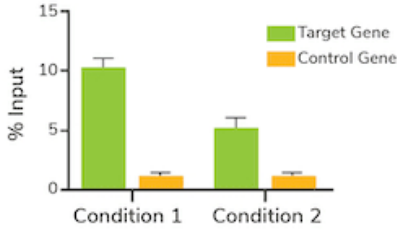
DNA Purification

Sequencing

qPCR

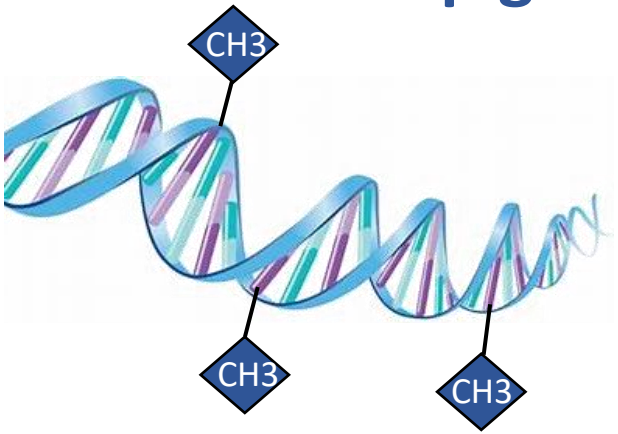
Analysis



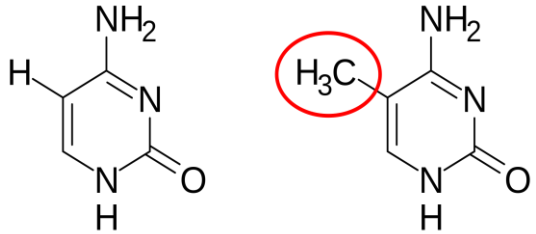


Kumar et al 2021

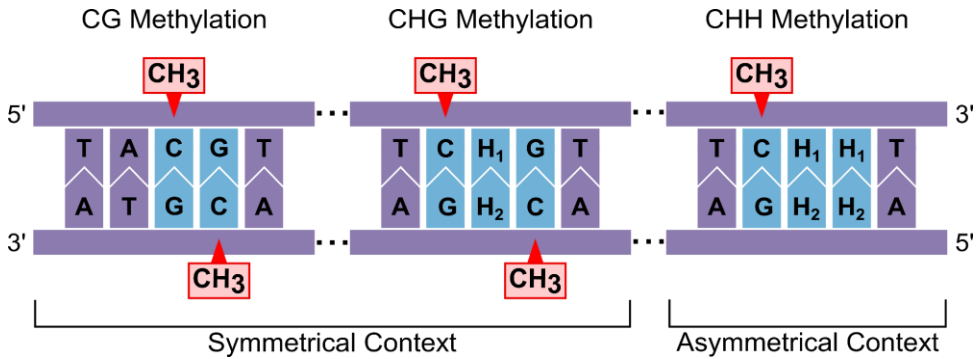
Epigenomics - DNA methylation



Methylgroup attached to Cytosine or Adenine (Adenine much less common)

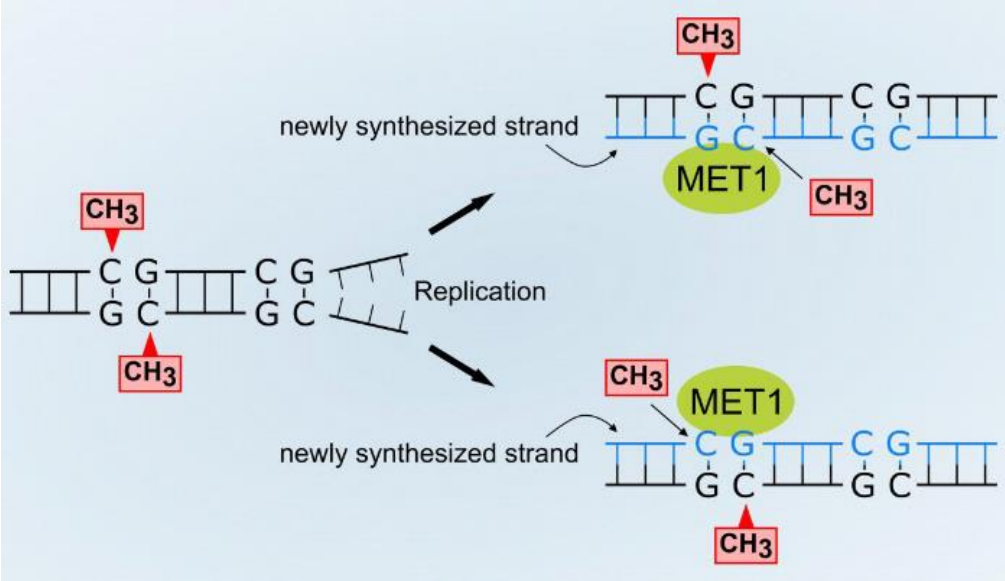


Three different “contexts” of Cytosine methylation

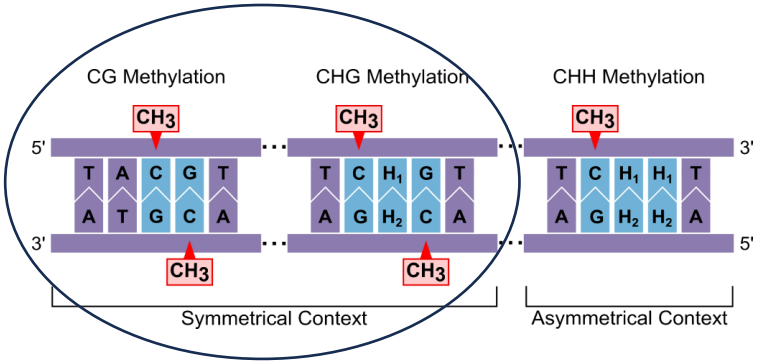


Epigenomics - Methylation can be maintained, *de novo* established and actively removed

Maintenance methylation

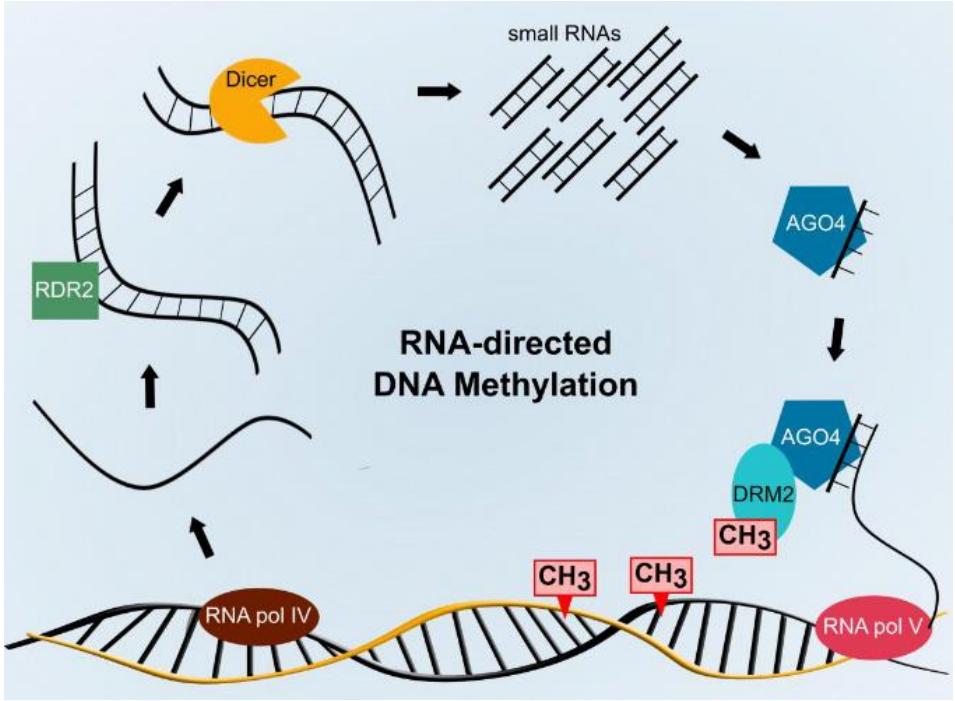


During replication, methylation information can be copied from the old to the new strand by methyltransferases *in the symmetric contexts*



Epigenomics - Methylation can be maintained, *de novo* established and actively removed

de novo methylation: RNA-directed DNA methylation



Occurs in all methylation contexts but most important in CHH due to the lack of maintenance methylation

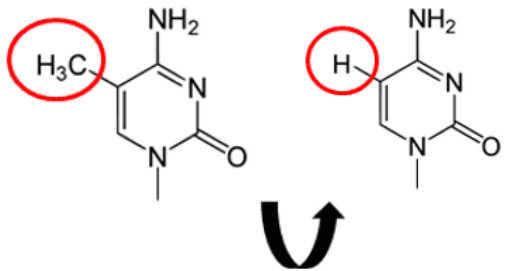
Non-coding RNAs and various proteins are in volved in RNA-directed (*de novo*) DNA methylation

Epigenomics - Methylation can be maintained, established *de novo* and actively removed

Demethylation

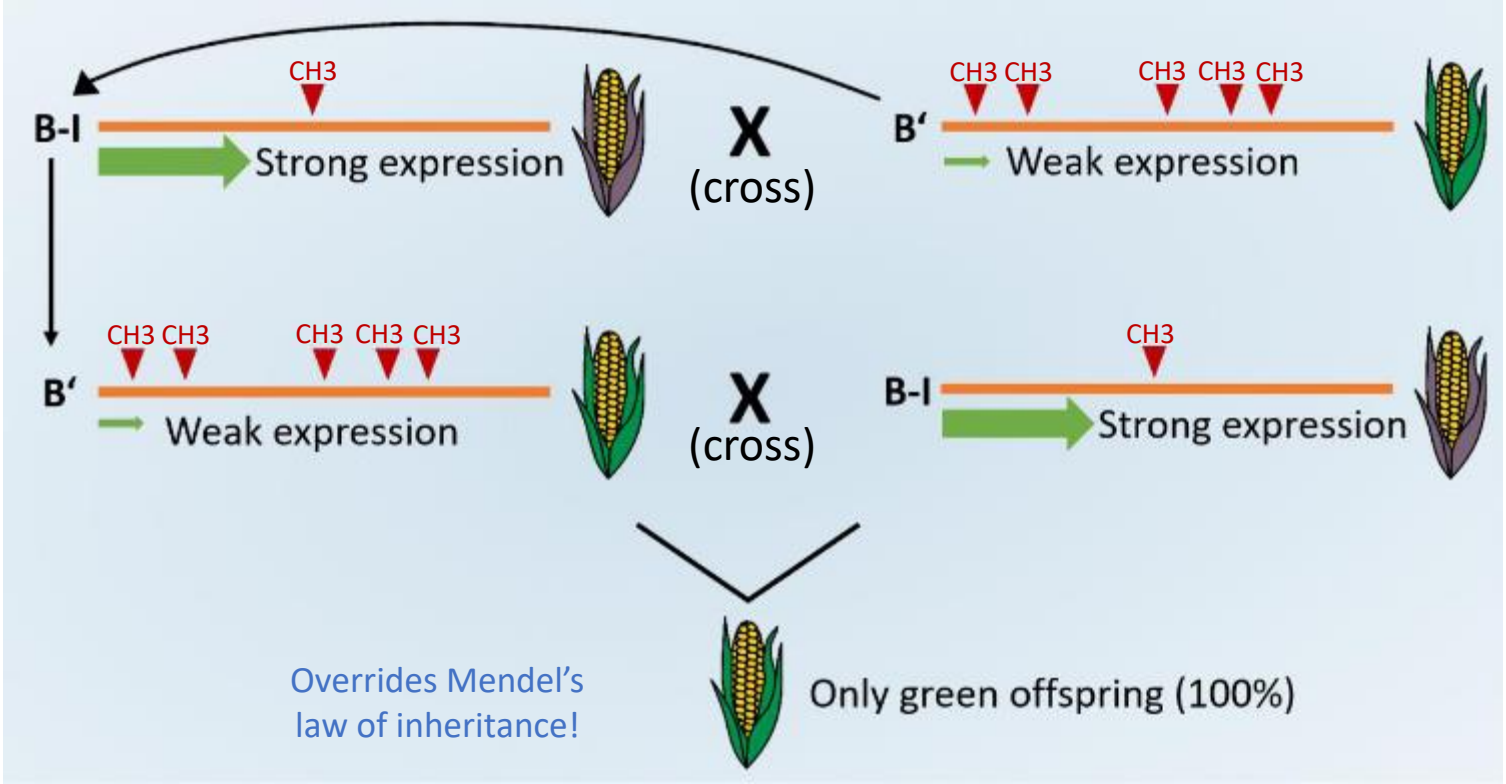
- Passively: lack of maintenance methylation during replication or DNA repair
- Actively: Base excision repair function (via DNA glycosylase domains in Repressor of silencing 1 (ROS1) and Demeter (DME) proteins)

→ The fact that methylation can be actively set and removed suggests a dynamic regulation of DNA methylation (influenced by environmental conditions)



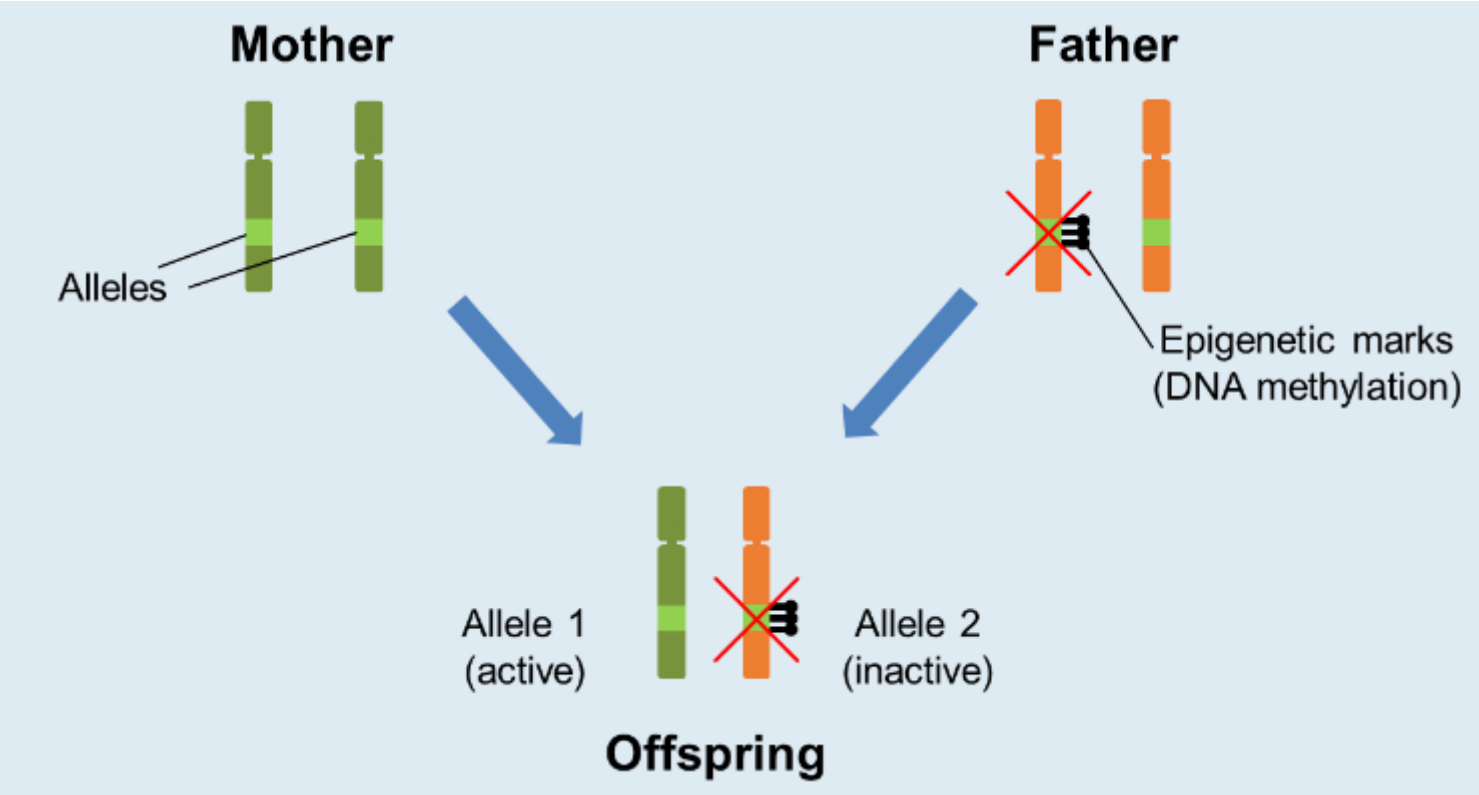
Epigenomics - Paramutation

Example: b1 locus in Maize
(involved in pigmentation pathway)



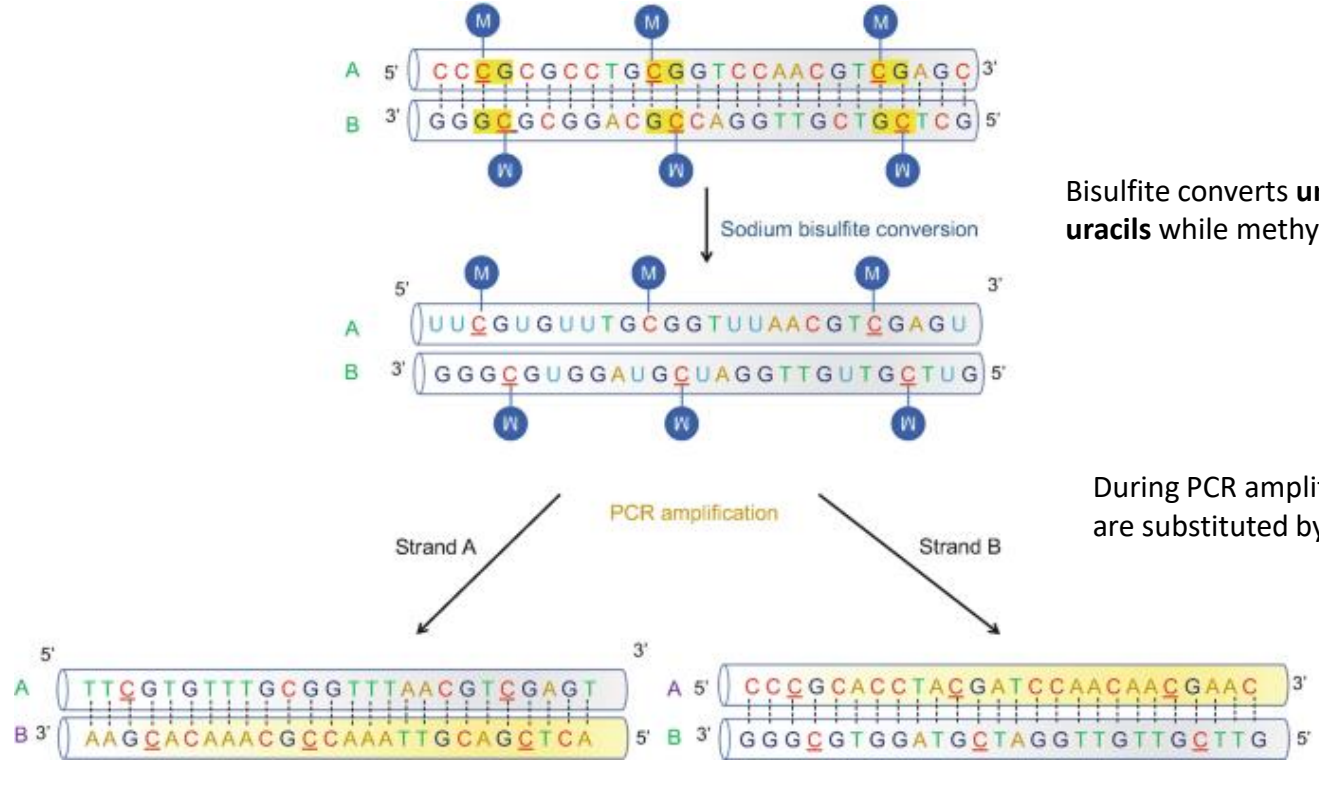
Alleles B-I and B' have identical DNA Sequence!

Epigenomics - Genomic Imprinting



Epigenomics - Methylome analysis

Whole genome bisulfite sequencing



Bisulfite converts **unmethylated** cytosines to **uracils** while methylated Cs are protected

During PCR amplification uracils are substituted by thymines

The converted samples are sequenced and then compared to the reference to deduce where methylation has preserved cytosines

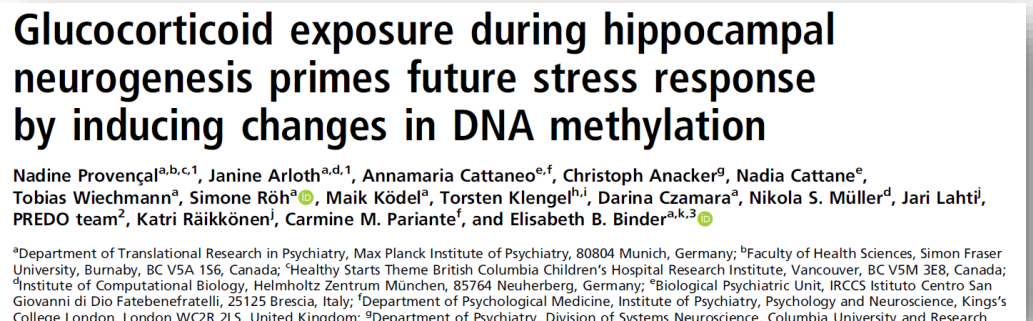
Epigenomics - Transgenerational Epigenetic Effects

Glucocorticoids (hormones) have multiple effects on fetal development (maturation of the lung, normal brain development etc)
And: Their secretion is also a known classic response to stress, providing energy for a “fight or flight” reaction

Increased pre-natal exposure to glucocorticoid in anxious mothers led to changes in the methylation pattern in the child

→ found enhanced transcriptional response to later glucocorticoid exposures

→ Suggest that high stress and anxiety level during pregnancy influence and increase the risk for psychiatric disorders later in the child’s life



Glucocorticoid exposure during hippocampal neurogenesis primes future stress response by inducing changes in DNA methylation

Nadine Provençal[a,b,c,1], Janine Arloth[a,d,1], Annamaria Cattaneo[e,f], Christoph Anacker[g], Nadia Cattane[e], Tobias Wiechmann[a], Simone Röh[a], Maik Ködel[a], Torsten Klengel[h,i], Darina Czamara[a], Nikola S. Müller[d], Jari Lahti[j], PREDO team[2], Katri Räikkönen[j], Carmine M. Pariante[f], and Elisabeth B. Binder[a,k,3]

[a]Department of Translational Research in Psychiatry, Max Planck Institute of Psychiatry, 80804 Munich, Germany; [b]Faculty of Health Sciences, Simon Fraser University, Burnaby, BC V5A 1S6, Canada; [c]Healthy Starts Theme British Columbia Children’s Hospital Research Institute, Vancouver, BC V5M 3E8, Canada; [d]Institute of Computational Biology, Helmholtz Zentrum München, 85764 Neuherberg, Germany; [e]Biological Psychiatric Unit, IRCCS Istituto Centro San Giovanni di Dio Fatebenefratelli, 25125 Brescia, Italy; [f]Department of Psychological Medicine, Institute of Psychiatry, Psychology and Neuroscience, Kings’s College London, London WC2R 2LS, United Kingdom; [g]Department of Psychiatry, Division of Systems Neuroscience, Columbia University and Research

Epigenomics - Transgenerational Epigenetic Effects

genes

MDPI

Review

Transgenerational Epigenetic Inheritance of Traumatic Experience in Mammals

Jana Švorcová

Department of Philosophy and History of Science, Faculty of Science, Charles University, 128 00 Prague, Czech Republic; jana.svorcova@natur.cuni.cz

Abstract: In recent years, we have seen an increasing amount of evidence pointing to the existence of a non-genetic heredity of the effects of events such as separation from parents, threat to life, or other traumatising experiences such as famine. This heredity is often mediated by epigenetic regulations of gene expression and may be transferred even across several generations. In this review, we


QUESTIONS?
Grazie ed Arrivederci!