CENTER OF PLANT SCIENCES

Sant'Anna
Scuola Universitaria Superiore Pisa

Introduction to Genomics 2024

Different Flavours in Genomics

-

Part A

Svenja Mager

svenja.mager@santannapisa.it

About this lecture

Besides the DNA sequence…

- ❑Functional Genomics
  - Gene transcription (Transcriptomics)
  - Gene translation (Proteomics)
  - Gene function/annotation
  - Regulation of gene expression
  - Protein–protein interactions
  - Metabolomics
- ❑Structural Genomics
  - Protein Structure
- ❑Epigenomics
  - DNA Methylation
  - Histone Modifications
- ❑Metagenomics
- ❑Genome Editing

Goal: Give you an impression about…

- the broadness of genomics
- the role it plays for life and its organisms
- the big scope of research question
- methods to confront these questions

About this lecture

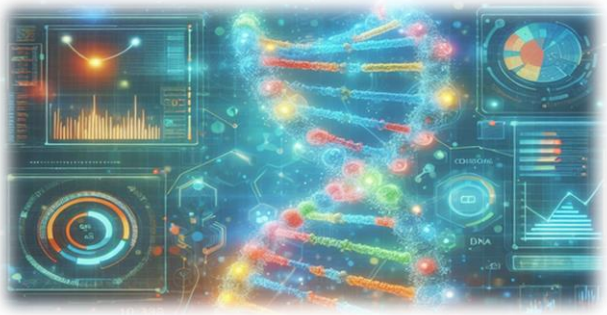
Main focus

Transcriptomics

- What is a transcriptome
- What functions does it have
- Analysis of gene expression and differentially expressed genes

Epigenomics

- What are epigenetic mechanisms
- What is their function
- Analysis of Chromatin status
- Analysis of DNA Methylation

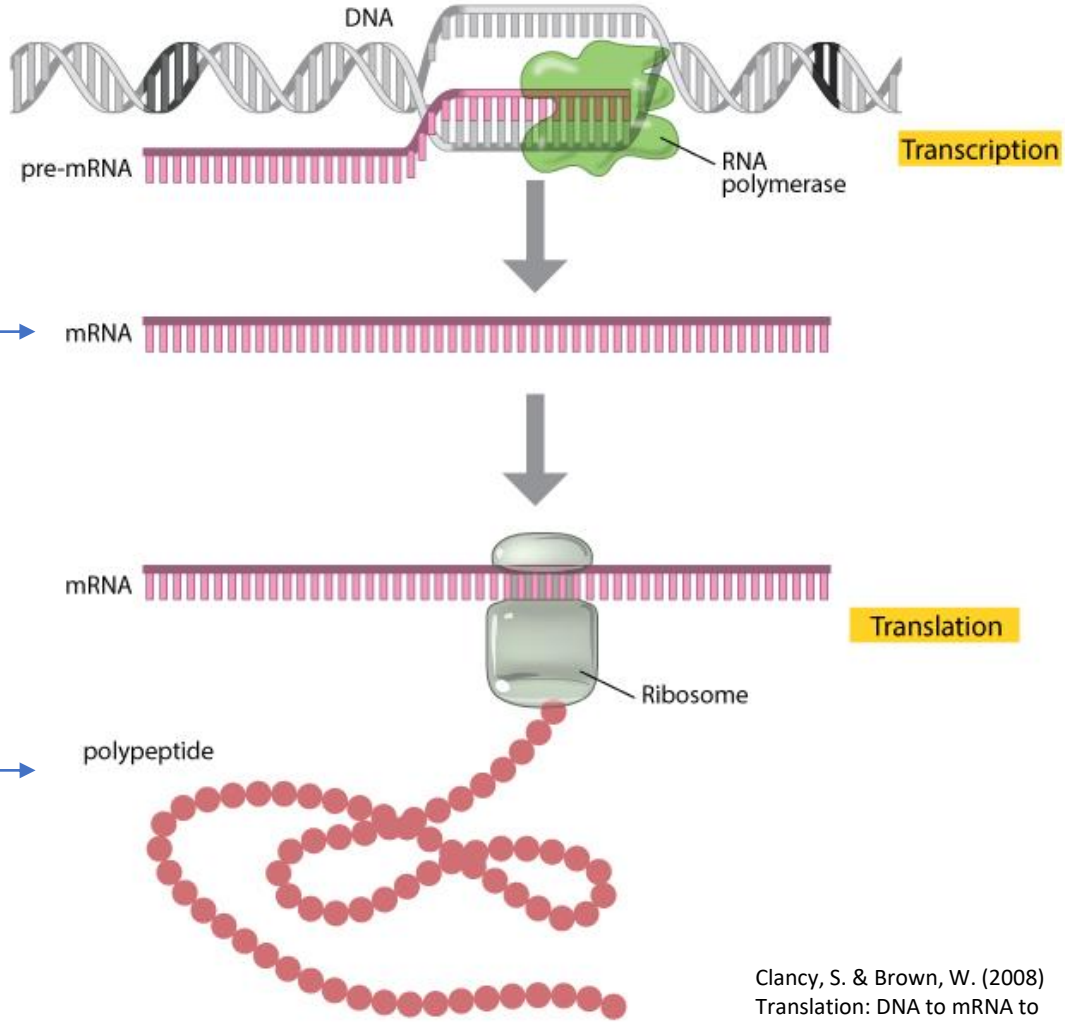


About this lecture

Reminder!

Transcriptome:
The complete set of transcribed genes at a certain time point in a cell/tissue/organism

Proteome:
The complete set of proteins present at a certain time point in a cell/tissue/organism



Clancy, S. & Brown, W. (2008) Translation: DNA to mRNA to Protein. Nature Education 1(1):101

Genomics Introduction

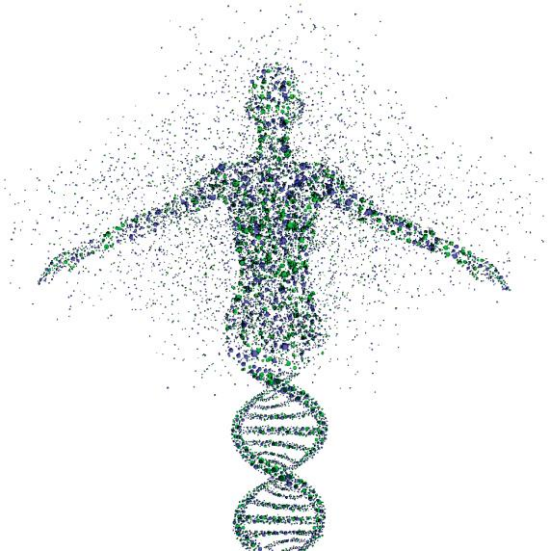
- Why are genomics important, what does it look at?
- From genotype to phenotype: how does DNA build an organism?
- Adaptability: How do organisms function in variable environments?

The genome is “static”

The genome is the same in all cells and tissues of an organism.

Exceptions include mutations and mitotic recombination.

→ if heritable, they are a characteristic of evolution

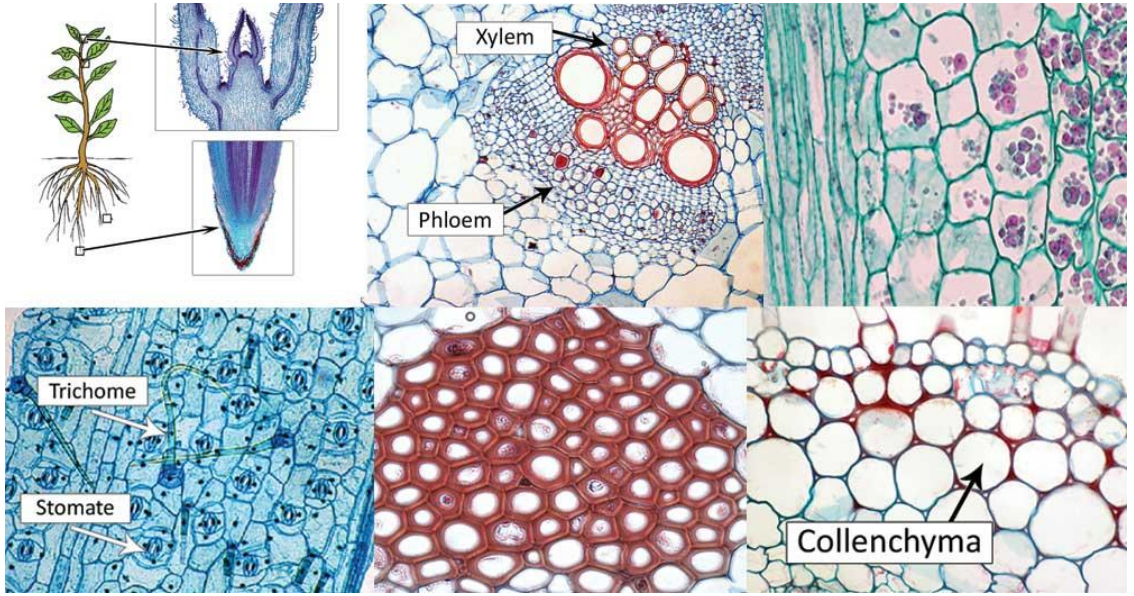


And transcriptome and proteome?

The transcriptome and proteome vary between cells and tissues!

Introduction - Transcriptome and Proteome vary

Example: Plants



Several cell types ….
- Parenchyma cells
- Collenchyma cells
- Sclerenchyma cells
- Xylem cells
- Phloem cells
- Meristematic cells
- Epidermal cells

… building tissues
- Dermal tissue
- Vascular tissue
- Ground tissue

… and organs
- Roots
- Stems
- Leaves

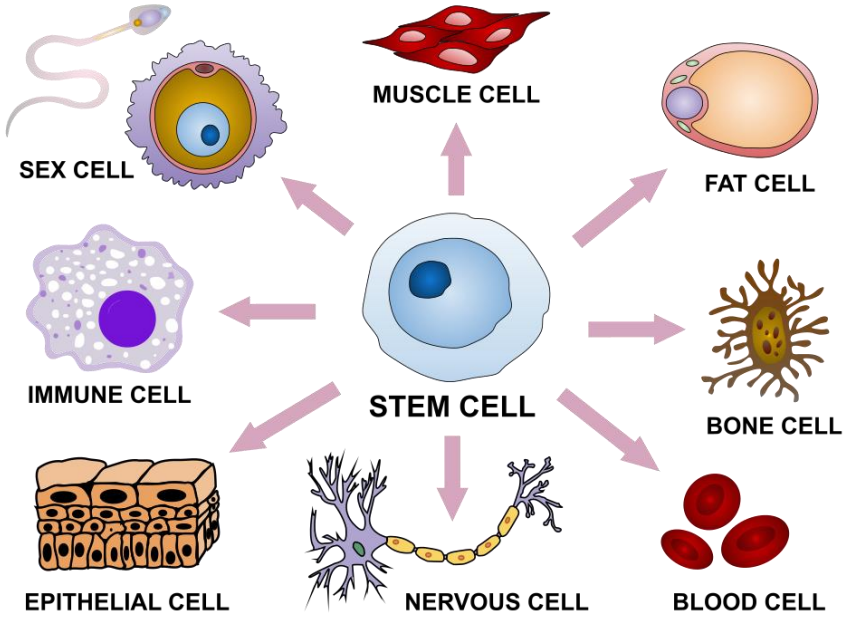
How do cells know what to be?
→ Cell differentiation

"Who am I?
What to do with my life?"

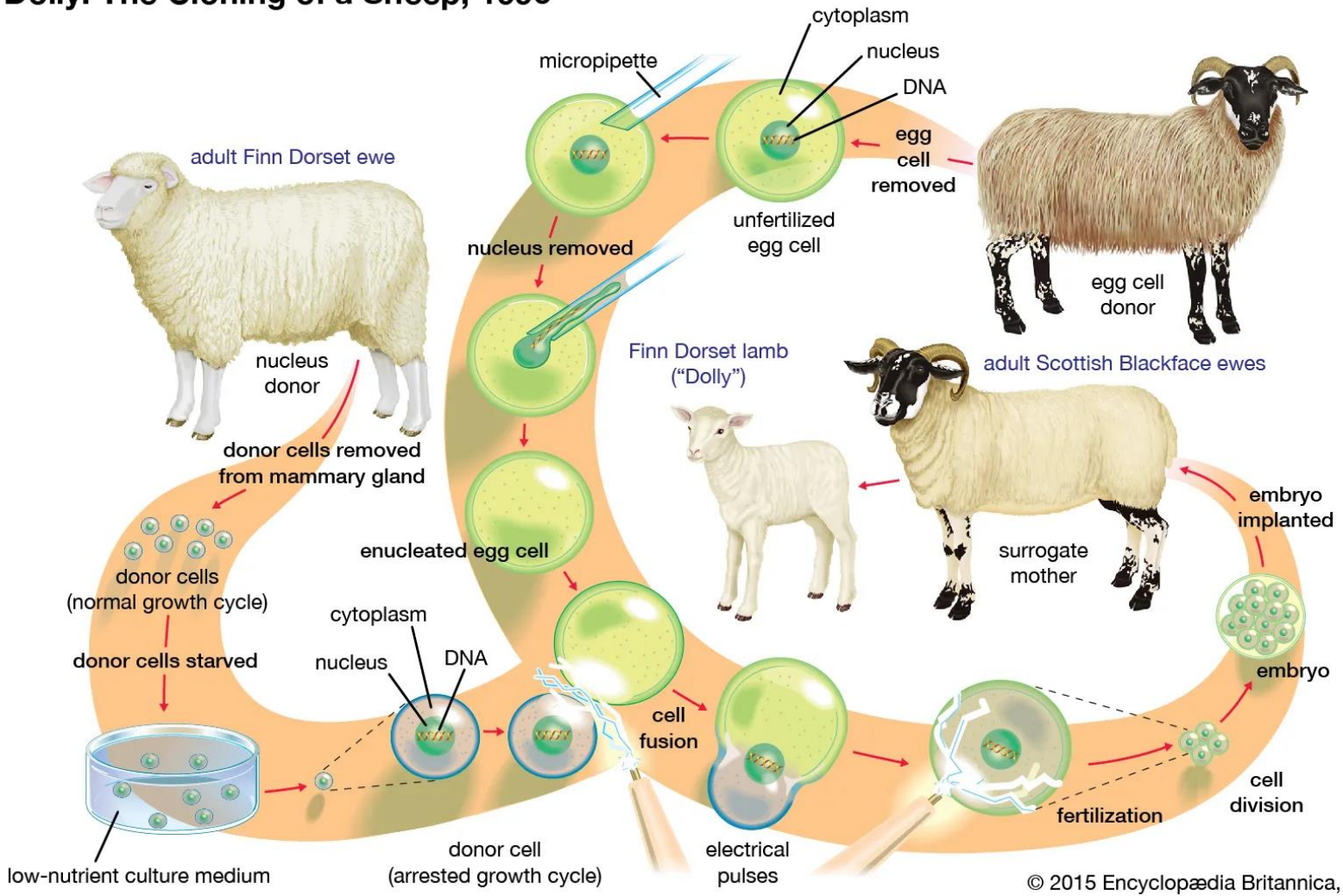
Differentiation changes a cell's size, shape, membrane potential, metabolic activity, and responsiveness to signals.

No change in DNA sequence involved.
Since each cell possesses the same genome, cell types must be regulated differently

Dolly was cloned from a mature cell



Dolly: The Cloning of a Sheep, 1996



How do cells know what to be?
→ Cell differentiation

Highly complex system of gene expression regulatory networks

Regulation by:

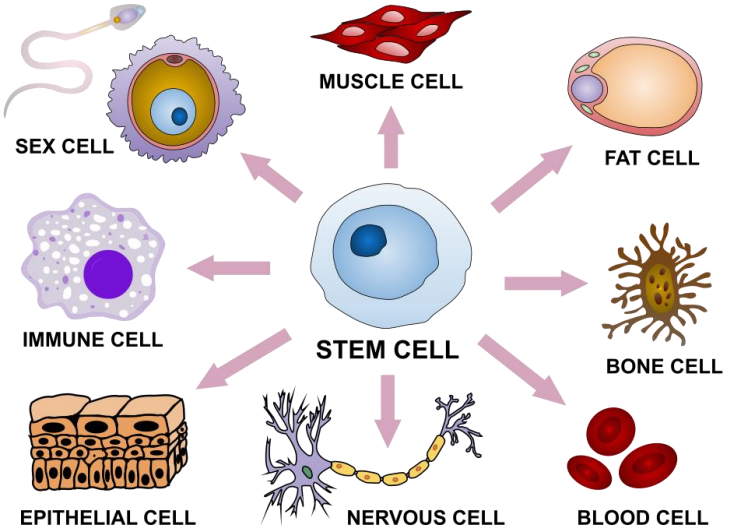
- cell-extrinsic factors

→ proteins secreted by other cells, temperature, oxygen etc.

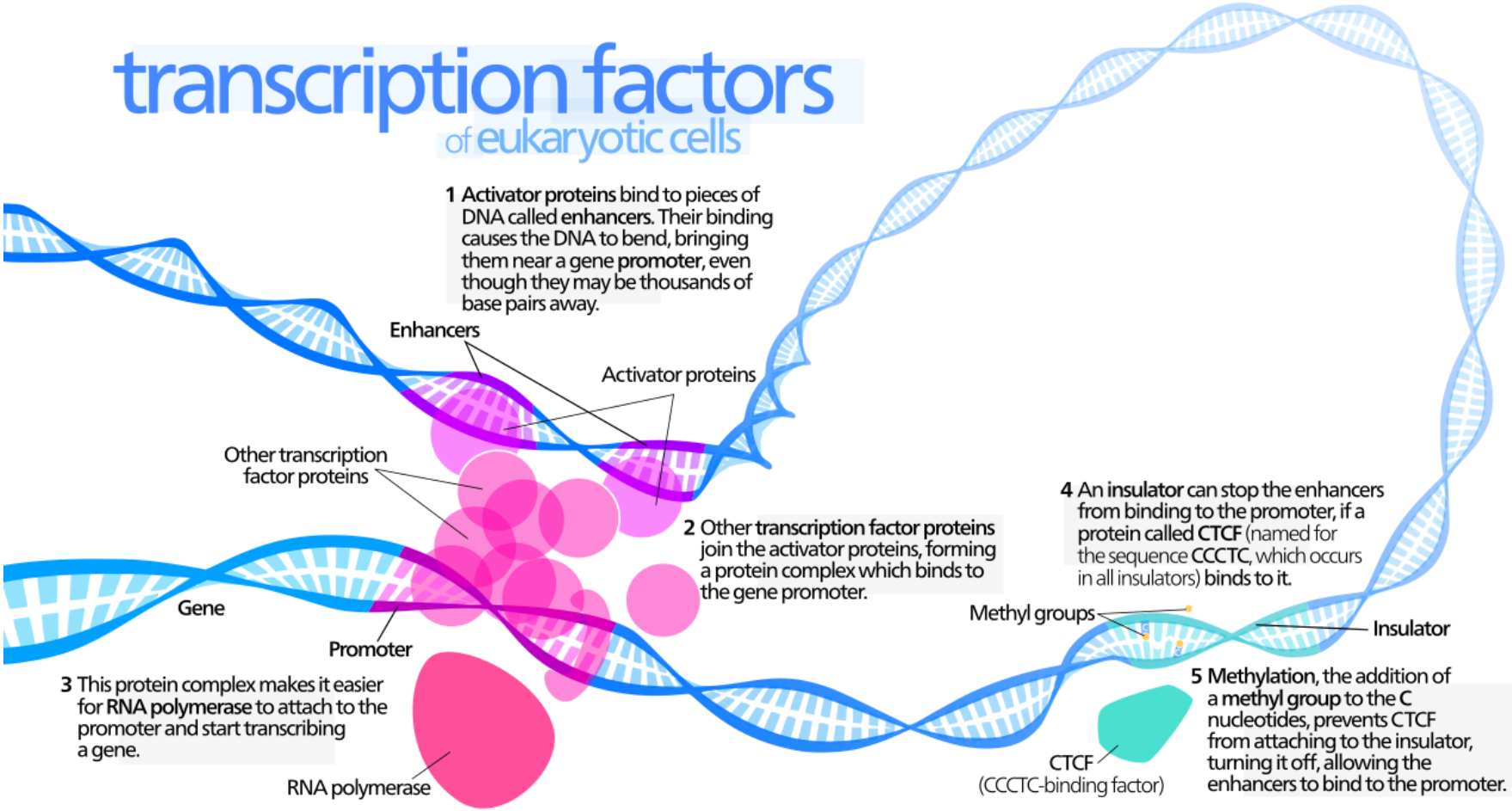
and

- cell-intrinsic factors

→ transcription factors, chemical modification of DNA and histones, chromatin remodeling (epigenetics)



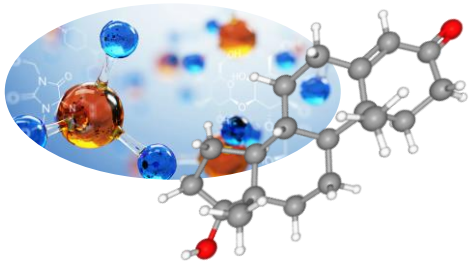
transcription factors
of eukaryotic cells



Other than cell differentiation?
→ Phenotypic Plasticity

Phenotypic Plasticity is the ability of a genotype to express different phenotypes in adaptation to varying environmental conditions

Most widespread (but not only) in basically immobile organisms that cannot move away from adverse conditions → plants



<u>Changes induced by phenotypic plasticity include</u>

- Morphological
- Physiological
- behavioral

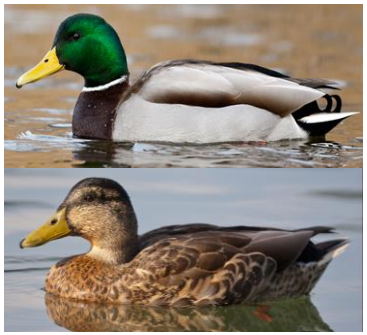
<u>Environmental stimulants can be:</u>

- Seasonal changes
- Chemicals, e.g. hormones
- Diet

Phenotypic Plasticity Examples

SEASONAL STIMULANTS

Arctic Fox
Camouflage in Summer and Winter



Male Mallard
during mating season and rest of the year

CHEMICAL CUES

Water fleas
produce spines and helmets when sensing predator-released chemicals

Freshwater Biology (2003) **48**, 1593–1602

Predator-induced phenotypic plasticity in the exotic cladoceran *Daphnia lumholtzi*

ANDREW R. DZIALOWSKI, JAY T. LENNON*, W.J. O'BRIEN† AND VAL H. SMITH
Department of Ecology and Evolutionary Biology, University of Kansas, Lawrence, KS, U.S.A.
**Present address: Department of Biological Sciences, Dartmouth College, 6044 Gilman Laboratory, Hanover, NH, U.S.A. 03755-3576*
†Present address: Department of Biology, University of North Carolina Greensboro, Greensboro, NC, U.S.A. 27402-6174

SUMMARY

1. The exotic cladoceran *Daphnia lumholtzi* has recently invaded freshwater systems throughout the United States. *Daphnia lumholtzi* possesses extravagant head spines that are

Phenotypic Plasticity Examples

DIET

Food determines a bee's fate!

Larva fed with:

- Royal Jelly (secretion from nurse bee glands) → 
- 'Beebread' (pollen and honey) → 

Phenotypic Plasticity Examples

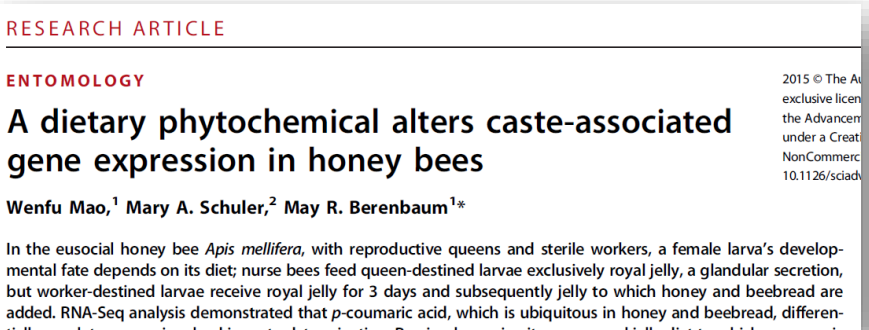
DIET

How?

"For years, people have wondered what components in royal jelly lead to queen development, but what might be more important is what isn't in royal jelly" (May Berenbaum)

Worker nutrition alters gene expression and modulates epigenetic changes!

- Organic compounds like p-Coumaric acid upregulate DNA methyltransferase production and downregulate peptides necessary for ovary development
- plant RNAs and miRNAs (important players in gene regulation and epigenetic modifications), delay development and decrease body and ovary size



RESEARCH ARTICLE

ENTOMOLOGY

A dietary phytochemical alters caste-associated gene expression in honey bees

Wenfu Mao,[1] Mary A. Schuler,[2] May R. Berenbaum[1]*

In the eusocial honey bee *Apis mellifera*, with reproductive queens and sterile workers, a female larva's developmental fate depends on its diet; nurse bees feed queen-destined larvae exclusively royal jelly, a glandular secretion, but worker-destined larvae receive royal jelly for 3 days and subsequently jelly to which honey and beebread are added. RNA-Seq analysis demonstrated that *p*-coumaric acid, which is ubiquitous in honey and beebread, differen-

10.1126/sciadv

PLOS | GENETICS

RESEARCH ARTICLE

Plant microRNAs in larval food regulate honeybee caste development

Kegan Zhu[1☯], Minghui Liu[1☯], Zheng Fu[1☯], Zhen Zhou[1], Yan Kong[1], Hongwei Liang[1], Zheguang Lin[2], Jun Luo[3], Huoqing Zheng[2], Ping Wan[3], Junfeng Zhang[1], Ke Zen[1], Jiong Chen[1,3]*, Fuliang Hu[2]*, Chen-Yu Zhang[1]*, Jie Ren[4]*, Xi Chen[1]*

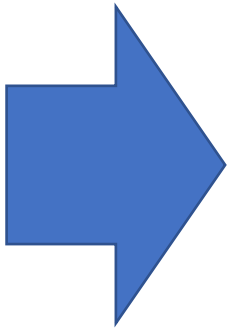
Transcriptomics Introduction

- What is analyzed in transcriptomics?
- Different types of RNA

Coming back to Genomics…

Cell differentiation and phenotypic plasticity are put into effect by **gene expression regulation**

→ Understanding the underlying mechanisms is a complex task and confronted with many different approaches



- Functional Genomics
  - **Gene transcription (Transcriptomics)**
  - Gene translation (Proteomics)
  - Gene function/annotation
  - Regulation of gene expression
  - Protein–protein interactions
  - Metabolomics
- Structural Genomics
  - Protein Structure
- **Epigenomics**
  - DNA Methylation
  - Histone Modifications

Transcriptomics vs Proteomics

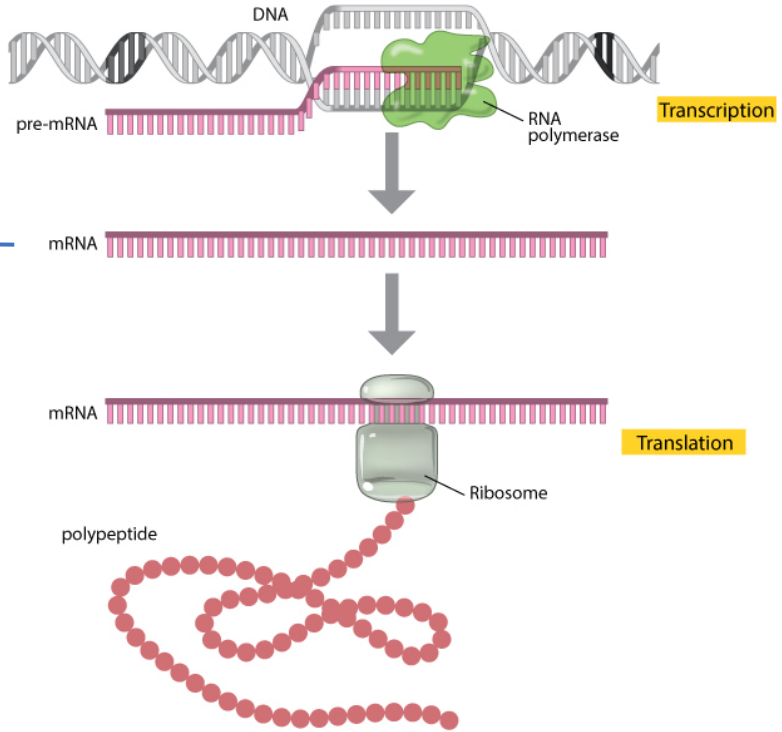
Why analyze both?

- Sequence and abundance of transcripts (mRNAs);
- coding and non-coding RNAs

Transciptomics

- Identification and expression level of proteins;
- Post-translational modifications

Proteomics



Transcript levels ≠ protein levels

- mRNA may be degraded rapidly or translated inefficiently
- many transcripts give rise to more than one protein
- Proteins maybe inactivated or activated later

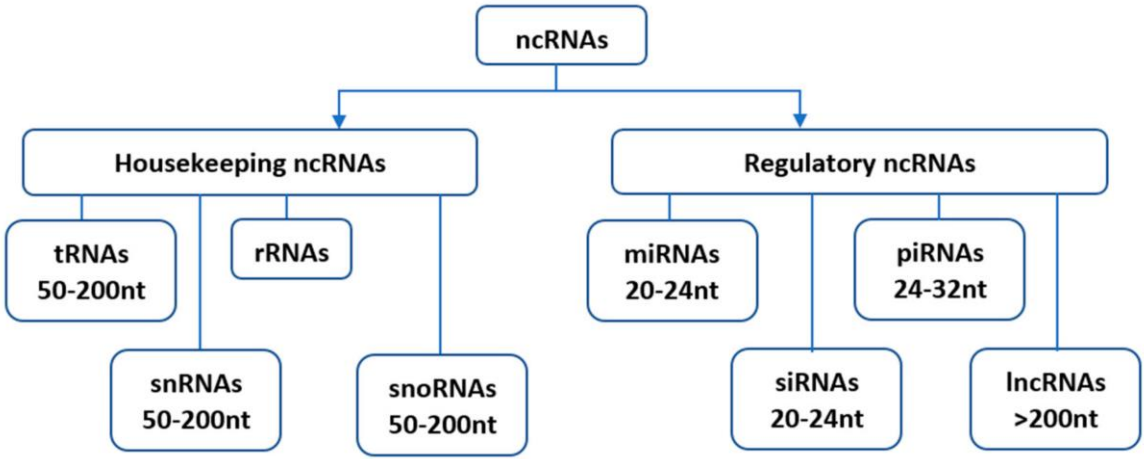
Transcriptomics - coding and non-coding RNA

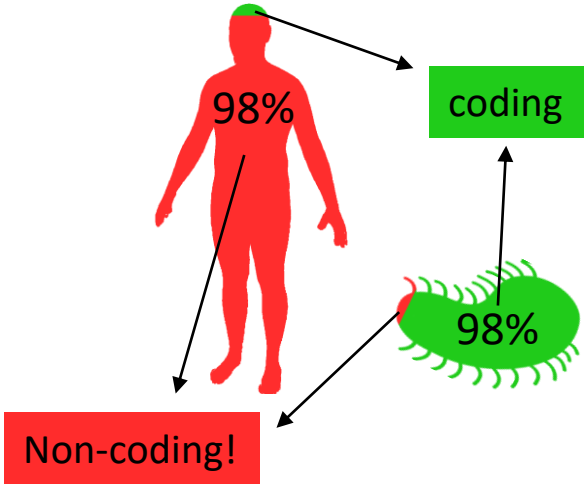
Often focus on **coding RNA** (translated to proteins), but:

→ only about 2-3% of eukaryote genomes is coding RNA!

The rest was considered “Junk” in the 60’s and 70’s

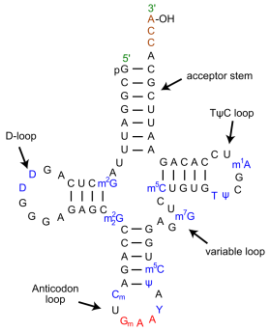
But, 80-90% of eukaryote genome transcribed at some point → **non-coding RNA**



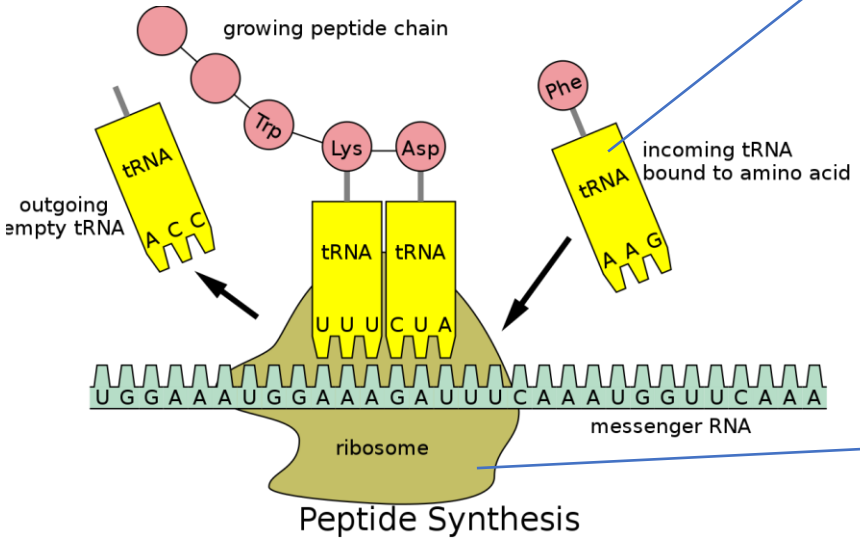


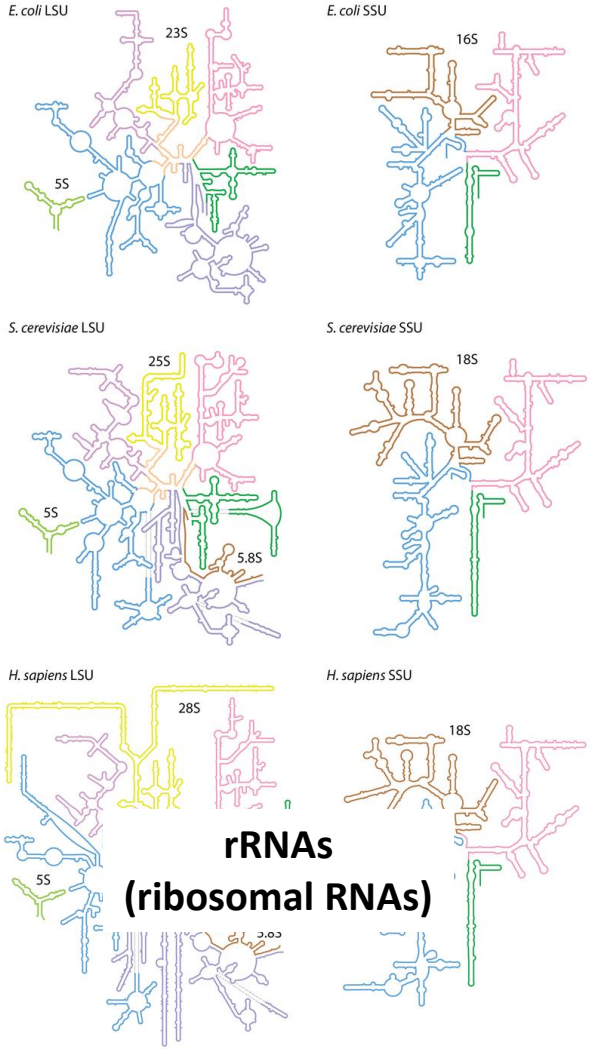
Waititu et al 2020

House-keeping non-coding RNAs



tRNAs (transfer RNAs)

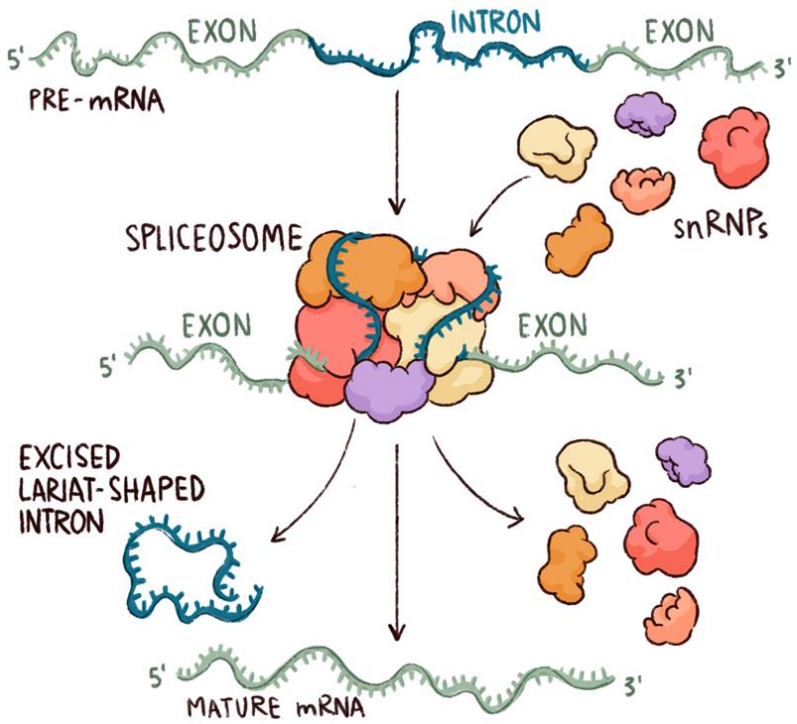




rRNAs (ribosomal RNAs)

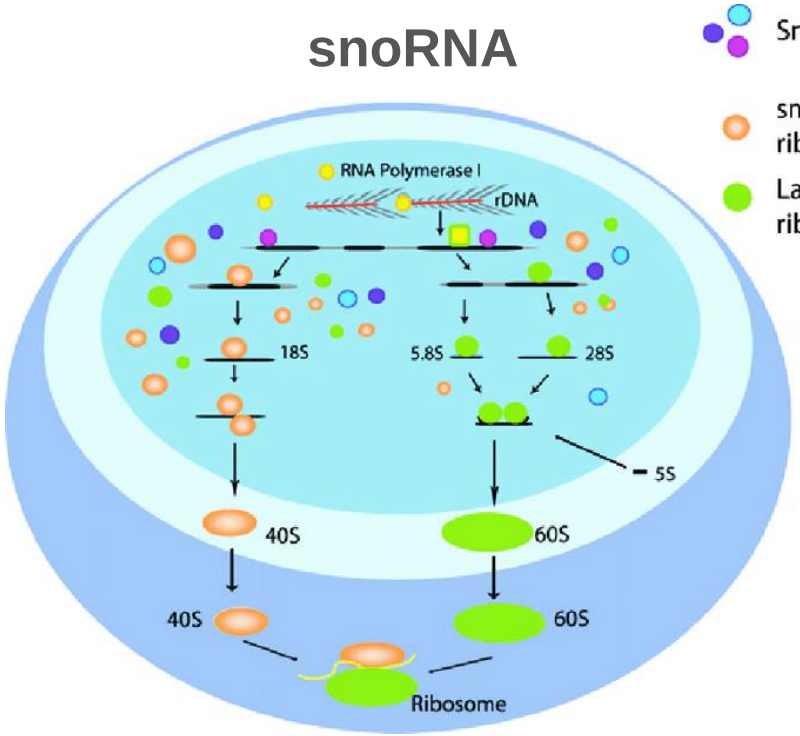
House-keeping non-coding RNAs

snRNA



Among the main functions of small nuclear RNA (**snRNA**) is Splicing

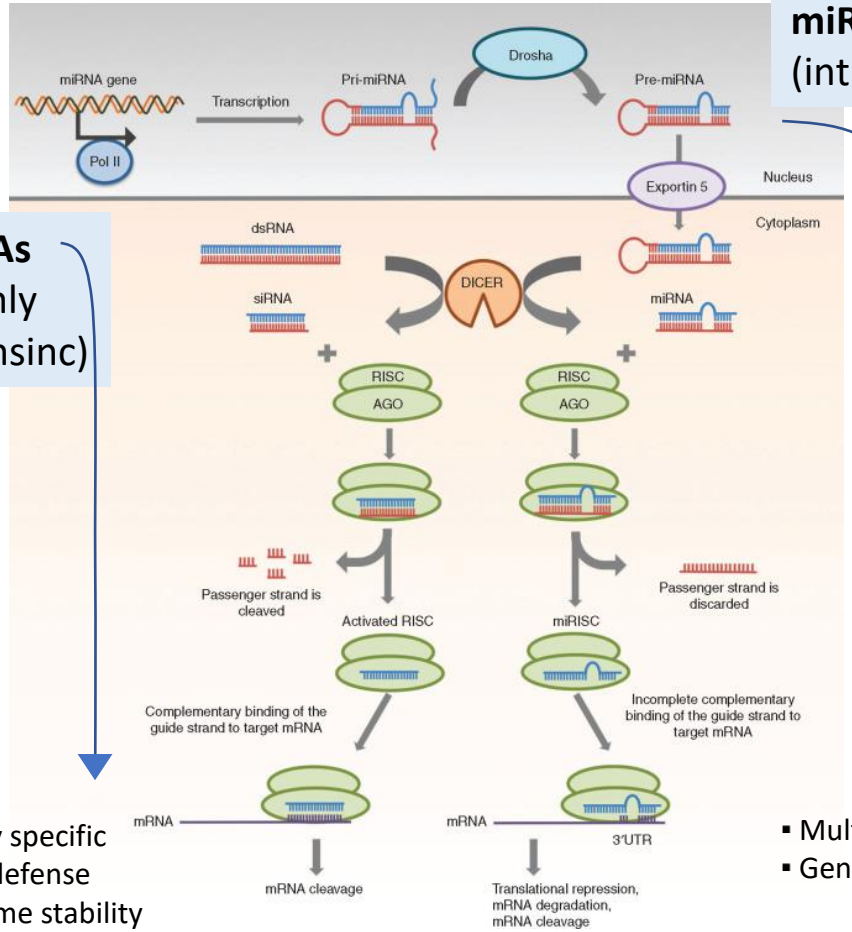
snoRNA



small nucleolar RNAs (**snoRNA**) stabilize the rRNA structure during ribosome biogenesis via 2'-O-methylation and pseudouridylation

Regulatory non-coding RNAs → miRNAs and siRNAs

miRNAs
(intrinsic)



siRNAs
(mainly extrinsinc)

- Highly specific
- Viral defense
- Genome stability

- Multiple targets
- Gene expression regulation

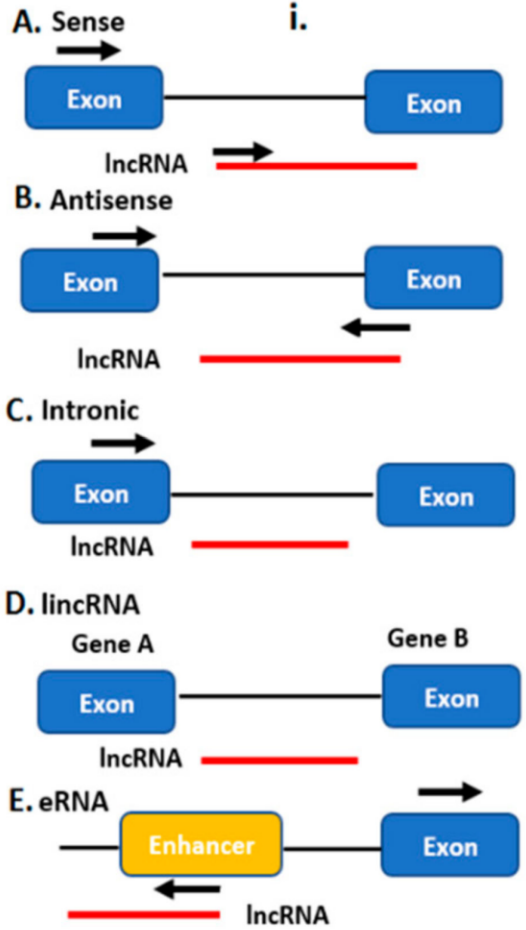
RNA Interference

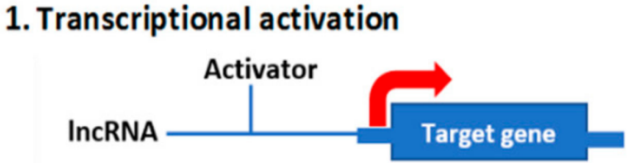
Micro RNAs (miRNAs) and small interfering RNAs (siRNAs) suppress gene expression by:

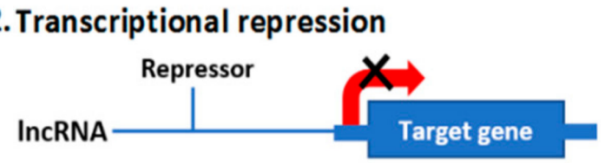
- Degradation of mRNA
- Inhibition of translation
- Heterochromatin formation (epigenetics)

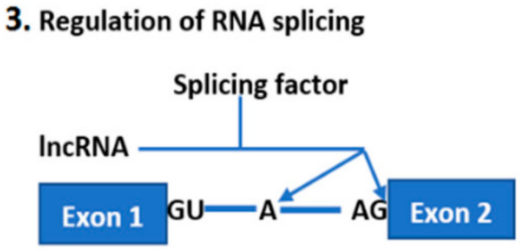
Lam et al 2015

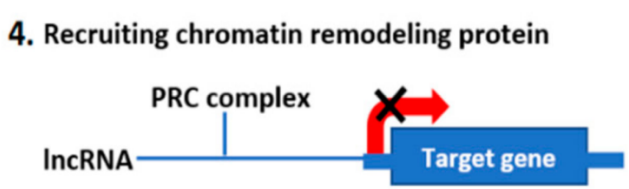
Regulatory non-coding RNAs → lnc-RNA (Long non-coding RNAs)

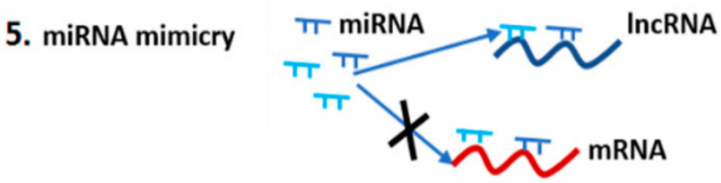






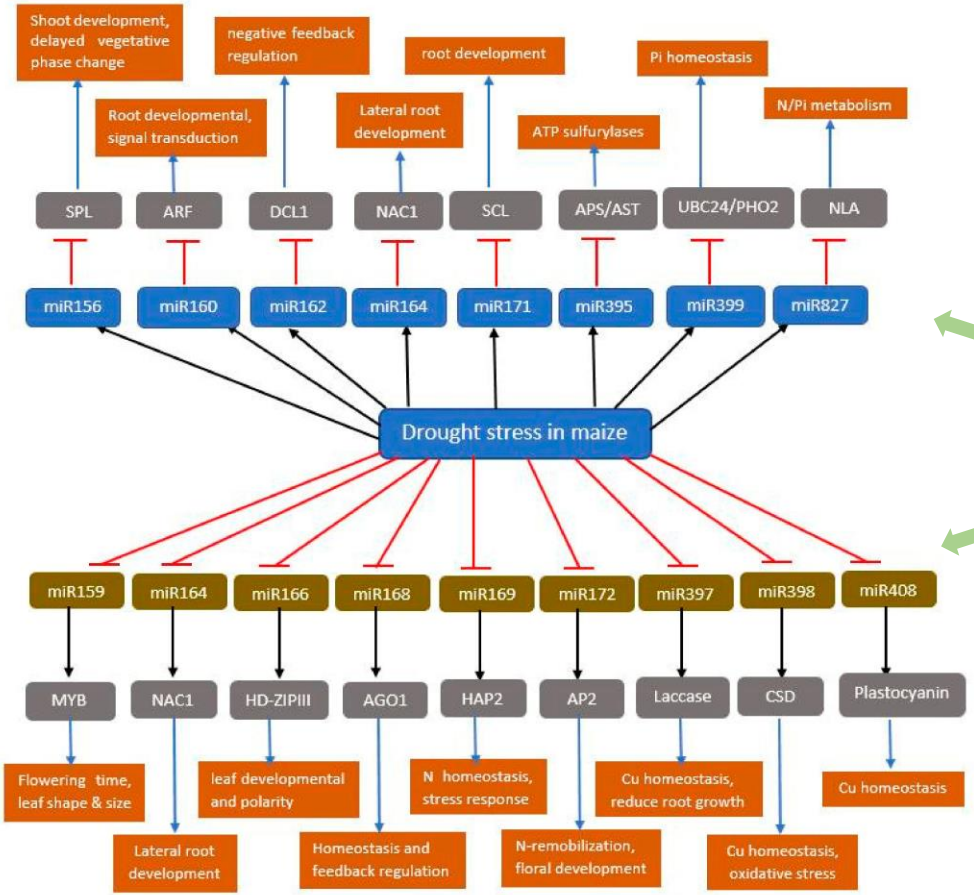






Waititu et al 2020

ncRNAs are part of a complex regulatory system



Example:
ncRNA in plant stress response

Up- and downregulation of miRNAs for drought stress response in maize

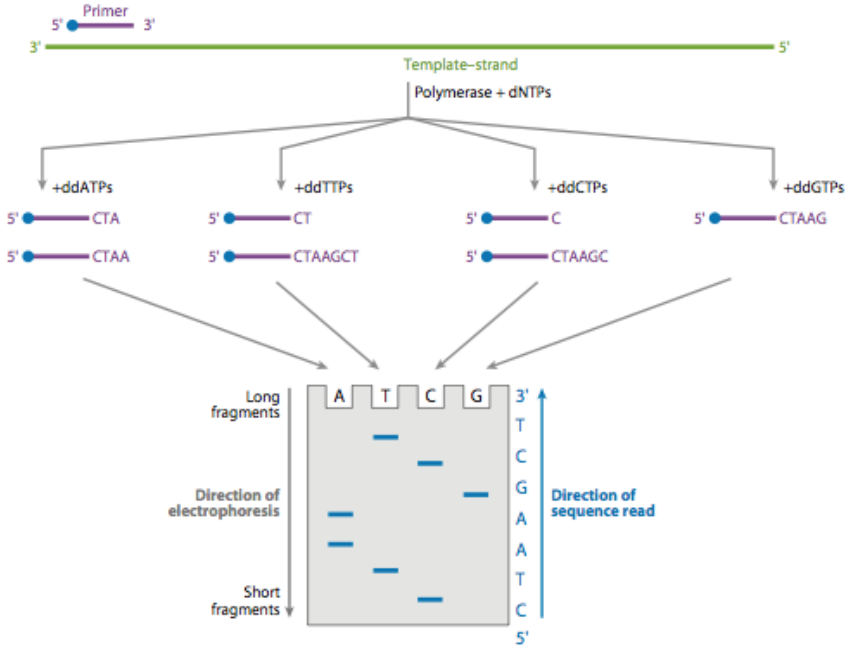
Transcriptomics Techniques in History

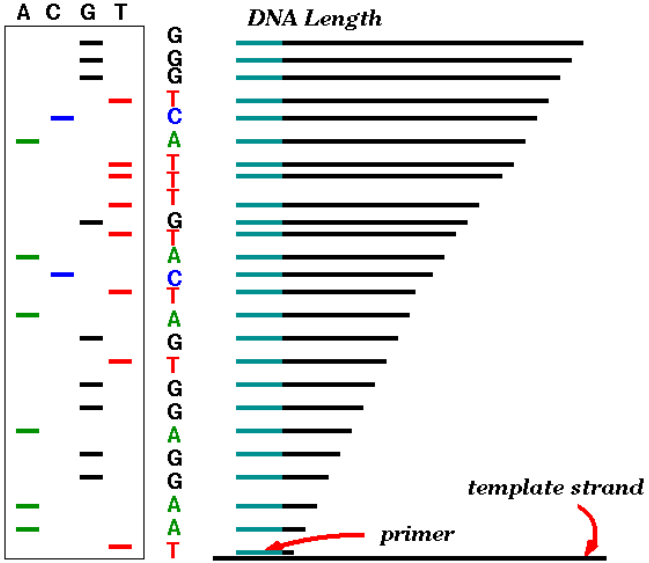
- When did transcriptomic analyses start?
- What were the first techniques?
- How did transcriptomics develop over time?

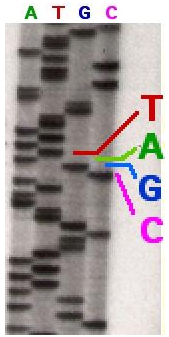
Transcriptomics Technologies - Beginnings

First attempts to capture and investigate transcripts date back to the 1970's and 1980's

- Sanger Sequencing, Expressed Sequence Tags
- Northern Blotting
- Reverse transcriptase quantitative PCR

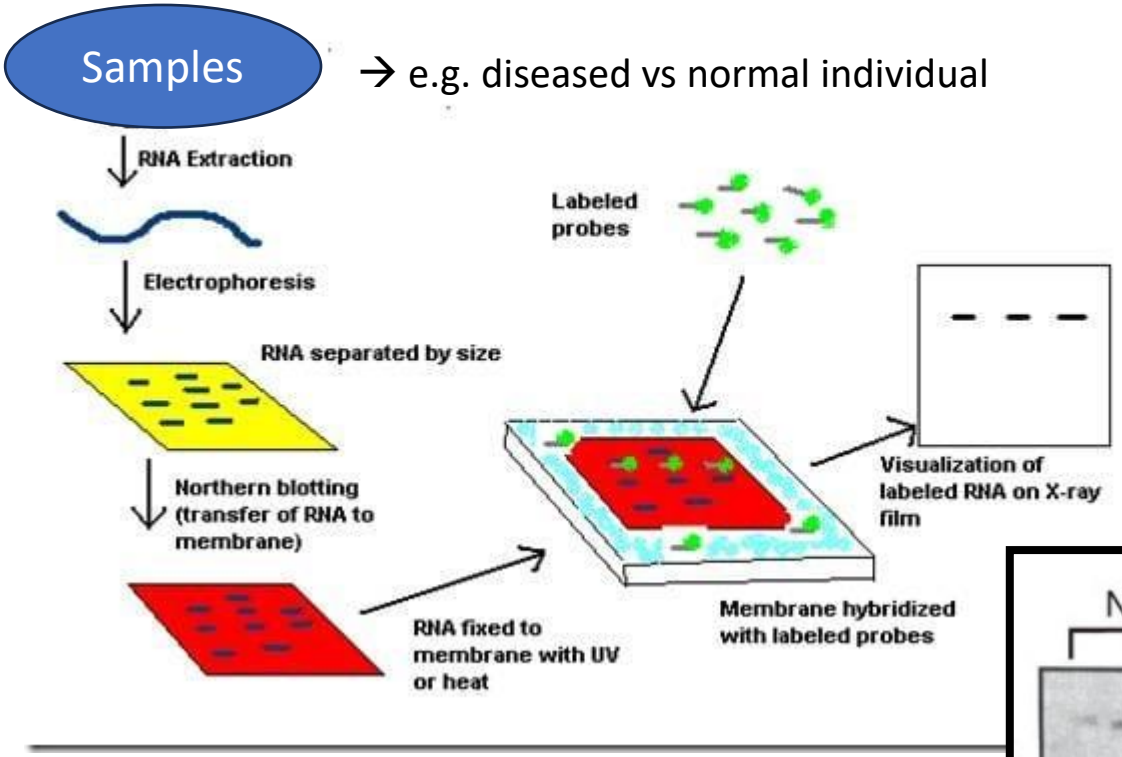




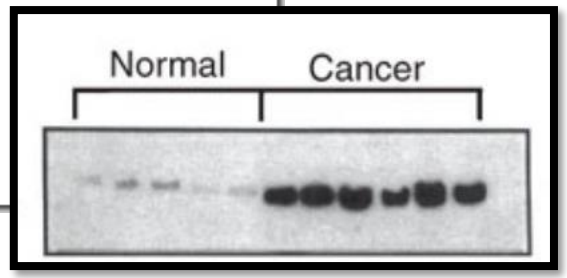


Transcriptomics Technologies - Beginnings

Northern Blotting

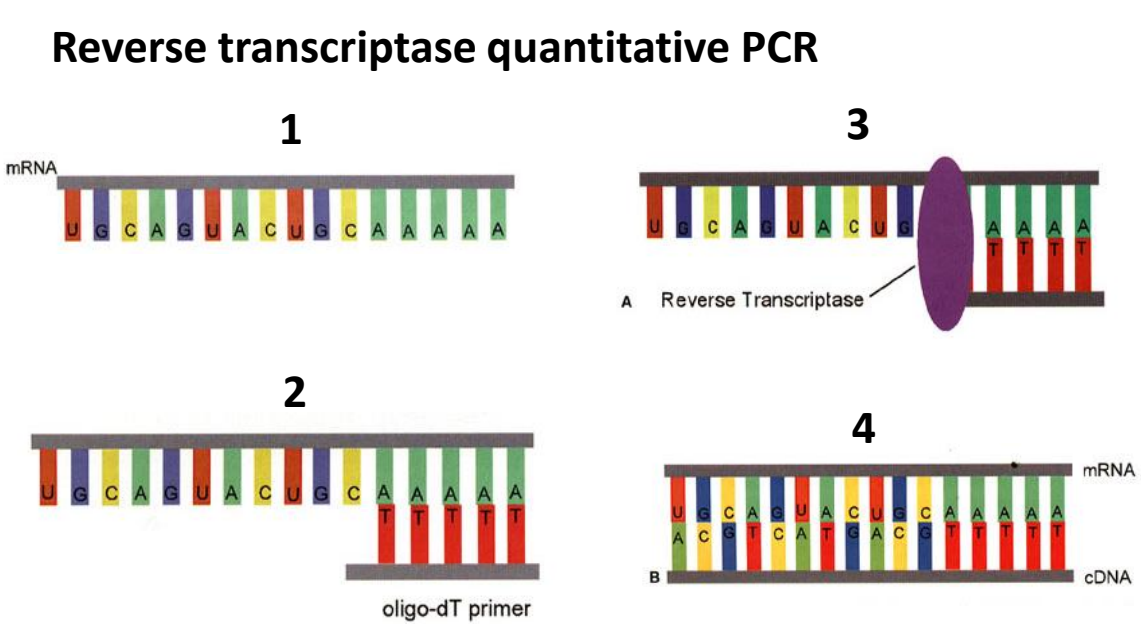


Northern Blotting allows to investigate one or a small number of genes



Transcriptomics Technologies - Beginnings

Reverse transcriptase quantitative PCR



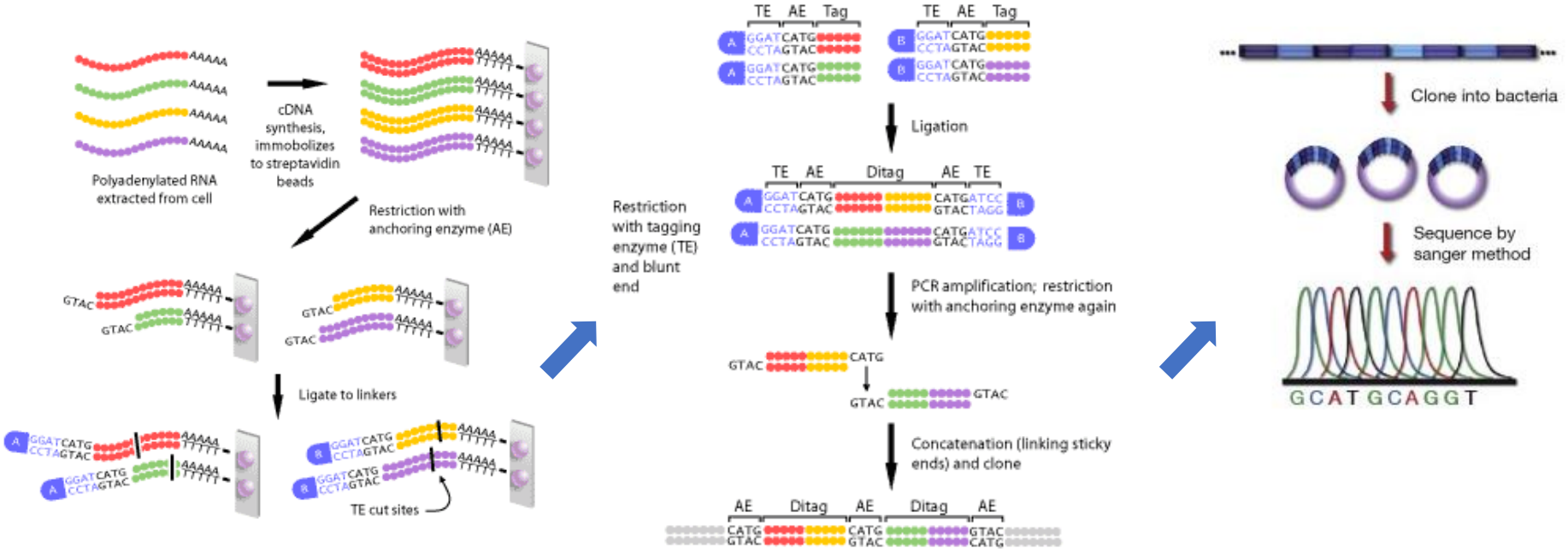
Sensitive, specific (with gene-specific primers) and rapid but only captures a tiny subsection of a transcriptome

→ visualization

Transcriptomics Technologies - Beginnings

“Transcriptome” as idea of the whole set of transcripts began in the 1990’s

One of the first methods was **Serial Analysis of Gene Expression (SAGE)** combined with Sanger Sequencing and later for a short time with high throughput sequencing



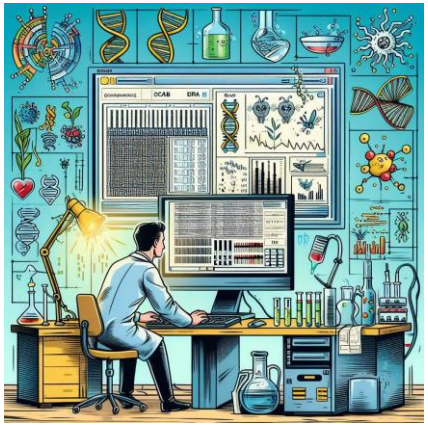
Rather complete transcriptome, but low accuracy and throughput, no splice information etc.

Transcriptomics Technologies - nowadays

Two dominant contemporary techniques:

Microarrays and RNA-Sequencing

Developed in the 1990’s and 2000’s, respectively



Transcriptomics Technologies - over time

Number of publications referring to different transcriptomics methods over the last three decades

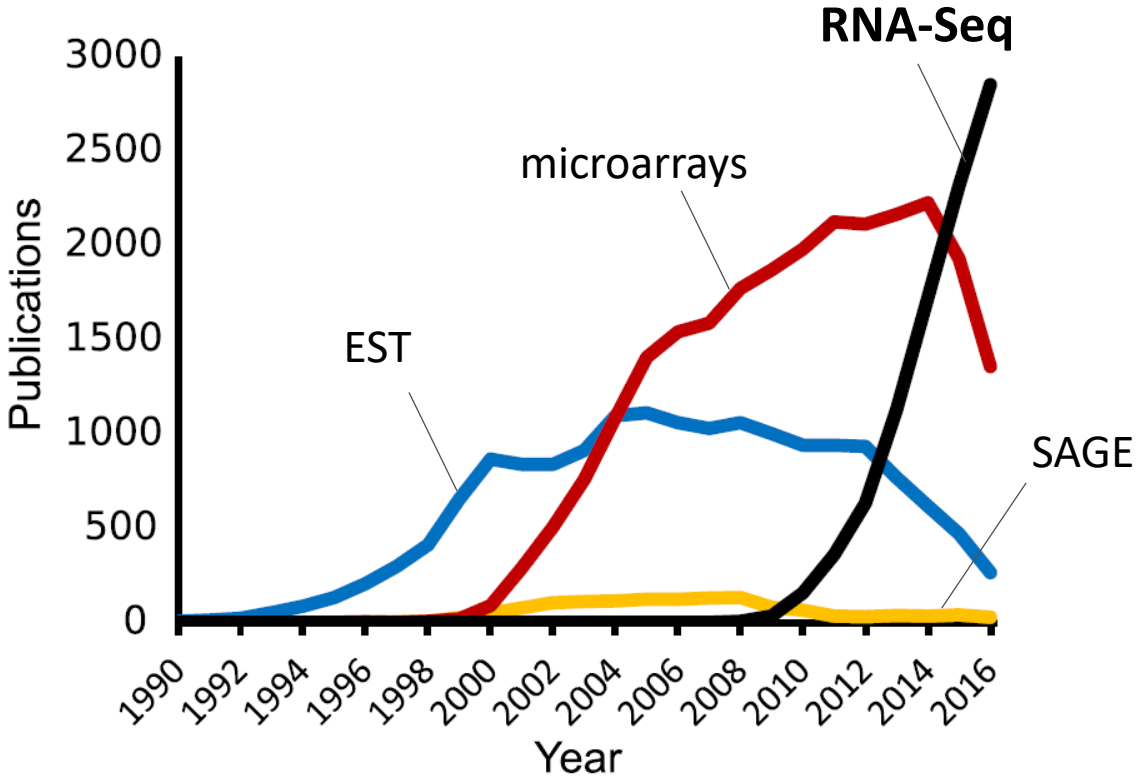
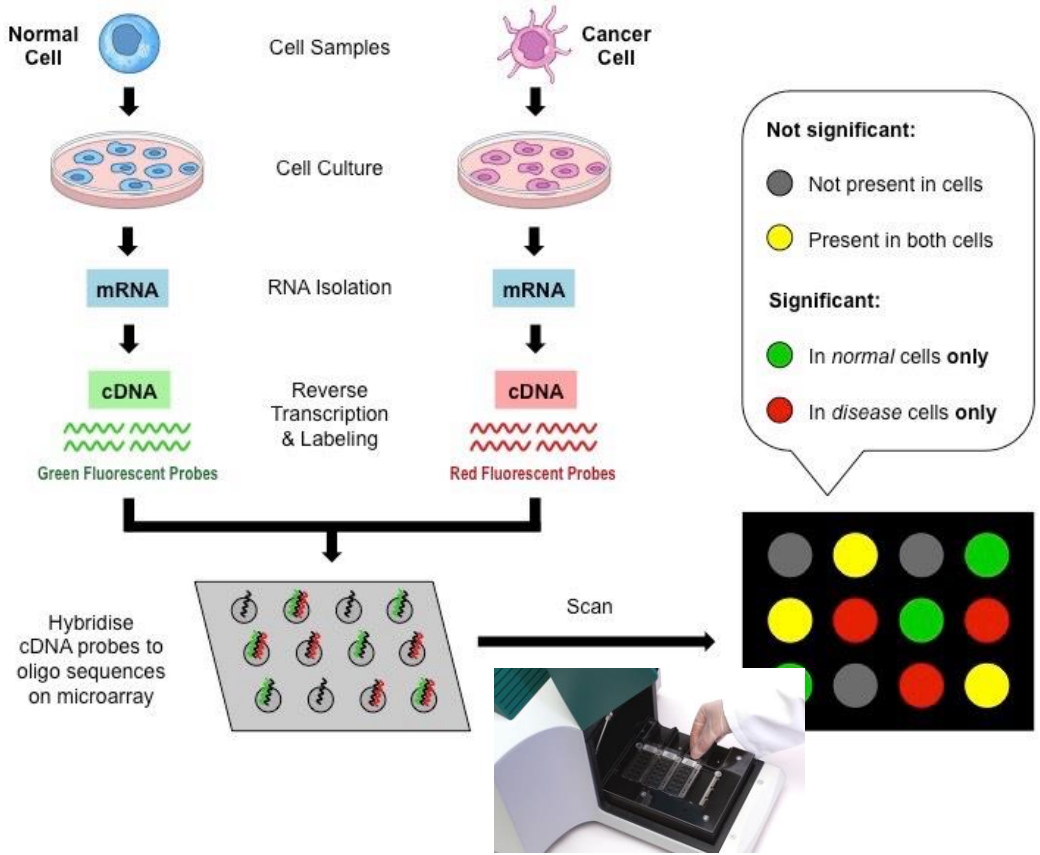


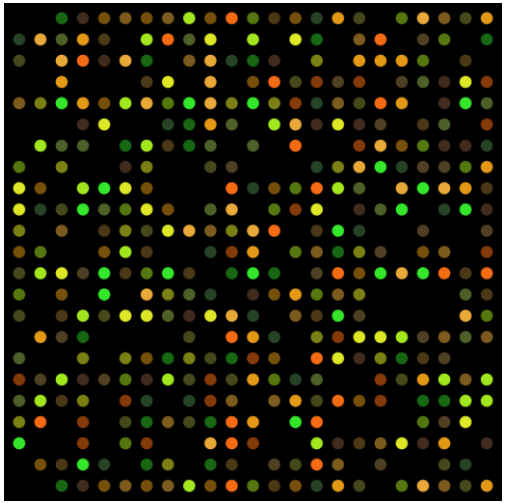
Fig 1. Transcriptomics method use over time. Published papers since 1990, referring to RNA sequencing (black), RNA microarray (red), expressed sequence tag (blue), and serial/cap analysis of gene expression (yellow)[12].

Transcriptomics Techniques Today

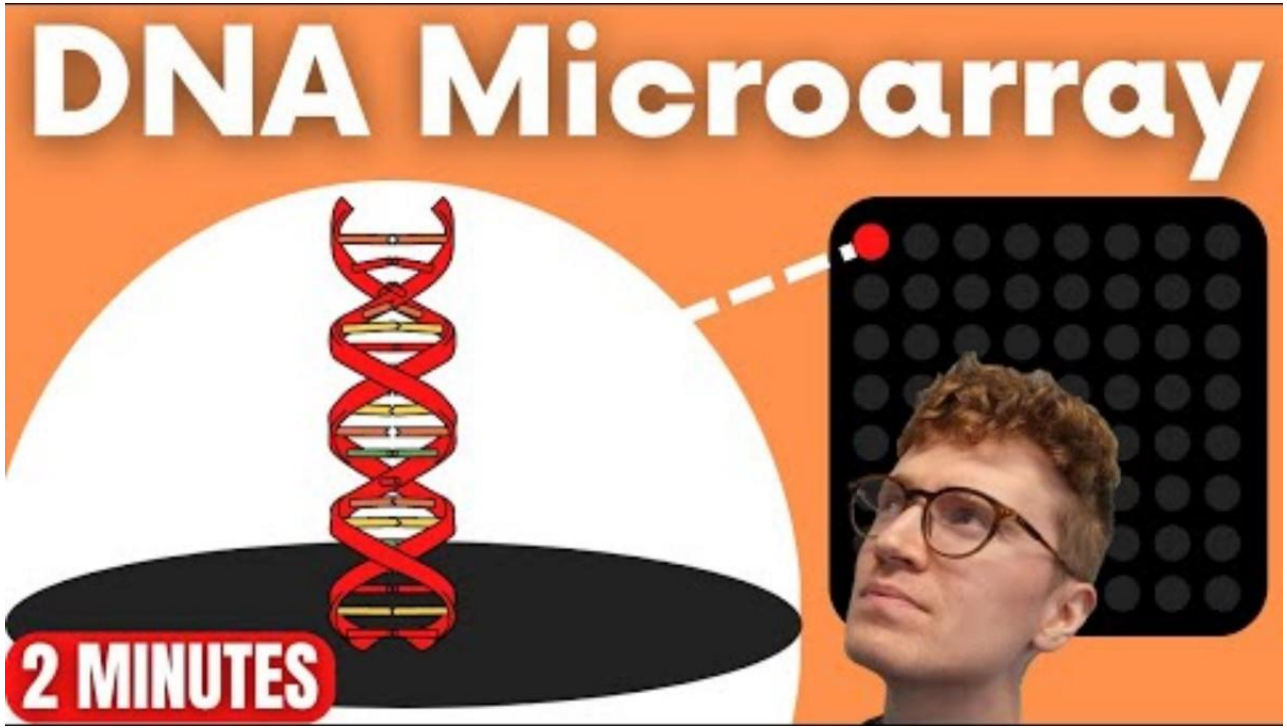
- Which are the main techniques used?
- How do Microarrays work?
- How does RNA Sequencing work?
- How to analyze transcriptomic data

Transcriptomics Technologies - Microarrays





Transcriptomics Technologies - Microarrays



High throughput and low cost ✓

BUT:

Reference needed! ✗

Transcriptomics Technologies - RNA-Sequencing

First descriptions in 2006 and 2008, overtook microarrays as dominant technique in 2015.

Advantages:

- Very high throughput
- Allows detection and quantification of transcripts
- Analyzing of non-coding RNAs
- Alignment to reference and *de novo* assembly possible
- Information about alternative splicing events

Several different RNA-Seq variations to decide for:

- the way of transcript enrichment
- the method of fragmentation
- length of fragments
- the method of amplification
- single or paired-end sequencing
- whether to preserve strand information and!


to be continued…
?
Questions
See you next time!