Cellular and Molecular Biology I

BIO-205-11

Camille Goemans

V. Molecular and Cellular Biology in the lab

1. Model organisms
2. Cell cultures
3. Studying proteins
4. **Studying DNA**
    - DNA sequencing
    - DNA extraction
    - DNA amplification
    - **DNA cloning**

Eukaryotic cloning systems

- Eukaryotic genes may **not be expressed properly** in bacteria (introns, post-translational modifications, …)
- Researchers have then to use eukaryotic cells, like **yeasts**
- In some cases, yeasts do not have proteins required to modify a mammalian protein properly, in which case **insect cells** or **cultured mammalian cells** are used

- Recombinant DNA is introduced by **electroporation** or by **injection** (with microscopically thin needles)
- Once inside the cell, DNA is incorporated by **natural recombination**

How to introduce DNA in eukaryotic cells? = Transfection

1. Physical Methods

These methods use physical forces to deliver DNA into cells.

Method	Principle	Applications / Notes
Microinjection	Direct injection of DNA into the nucleus or cytoplasm using a fine glass needle under a microscope.	Used for animal embryos, oocytes, and transgenic animals; highly precise but labor-intensive.
Electroporation	Cells are exposed to brief electric pulses that create temporary pores in the plasma membrane, allowing DNA to enter.	Works well with many cell types, including mammalian, plant, and yeast cells; efficient but may cause some cell death.
Biolistic (Gene Gun) Method	DNA-coated microscopic particles (usually gold or tungsten) are shot into cells using high pressure.	Common in plant and fungal cells with tough cell walls; also used in some animal tissues.

How to introduce DNA in eukaryotic cells?

2. Chemical Methods

These rely on chemical carriers or compounds that facilitate DNA uptake.

Method	Principle	Applications / Notes
Calcium Phosphate Precipitation	DNA forms a precipitate with calcium phosphate, which is taken up by endocytosis.	One of the oldest and simplest methods; often used in mammalian cell lines.
Liposome-Mediated Transfection (Lipofection)	DNA is enclosed in lipid vesicles (liposomes) that fuse with the cell membrane to deliver genetic material.	High efficiency and low toxicity; widely used in human and animal cells.
Polymer-Based Transfection	Uses cationic polymers (like PEI or DEAE-dextran) to form complexes with negatively charged DNA, which are endocytosed.	Versatile and cost-effective; used in research and production.

How to introduce DNA in eukaryotic cells?

3. Biological (Viral) Methods

These exploit the natural ability of viruses to deliver genetic material into host cells.

Viral Vector	Type	Key Features / Uses
Retrovirus / Lentivirus	RNA viruses that integrate into the host genome.	Long-term, stable expression in dividing and non-dividing mammalian cells (e.g., in gene therapy).
Adenovirus	DNA virus; does not integrate into the genome.	Transient, high-level expression; useful for gene delivery in vivo.
Adeno-Associated Virus (AAV)	Small DNA virus with low immunogenicity.	Safe and efficient; used in human gene therapy.
Herpes Simplex Virus (HSV)	Large DNA virus; targets neurons.	Used in neuroscience and gene therapy for the nervous system.

How to introduce DNA in eukaryotic cells?

4. Genome Editing Technologies

Modern methods that allow precise, targeted modification of specific genes.

Technology	Mechanism	Applications / Notes
CRISPR-Cas9 System	Uses a guide RNA (gRNA) to target a specific DNA sequence, and Cas9 enzyme to cut DNA, enabling insertion, deletion, or correction.	Fast, accurate, inexpensive; used for gene knockout, knock-in, and therapy research.
RNA Interference (RNAi)	*Uses small interfering RNA (siRNA) or shRNA to silence gene expression post-transcriptionally.*	*Used to knock down genes (not permanently modify DNA).*

How to introduce DNA in eukaryotic cells?

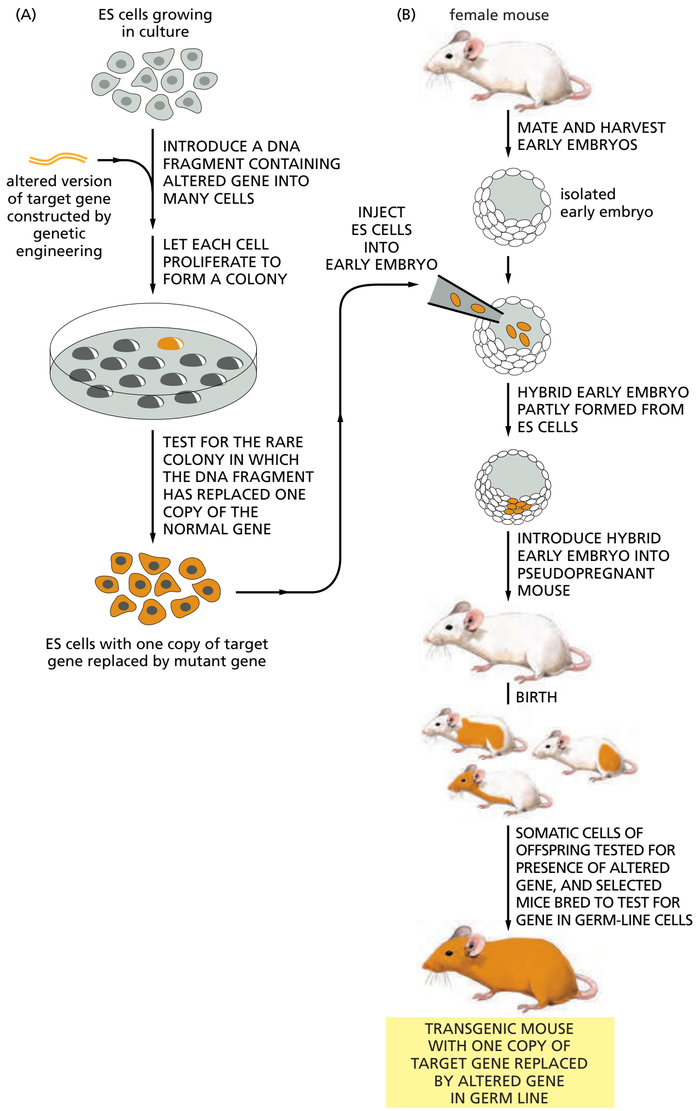
5. Plant-Specific Methods

Method	Description
Agrobacterium tumefaciens–Mediated Transformation	Uses a soil bacterium that naturally transfers part of its DNA (T-DNA) into plant cells — the basis for many genetically modified crops.
Protoplast Transformation	DNA introduced into plant cells without cell walls using electroporation or PEG treatment.

Engineering animals: transgenic organisms

- **Transgenic organisms** are organisms that have been genetically engineered
- **A transgene** is a foreign gene that has been added

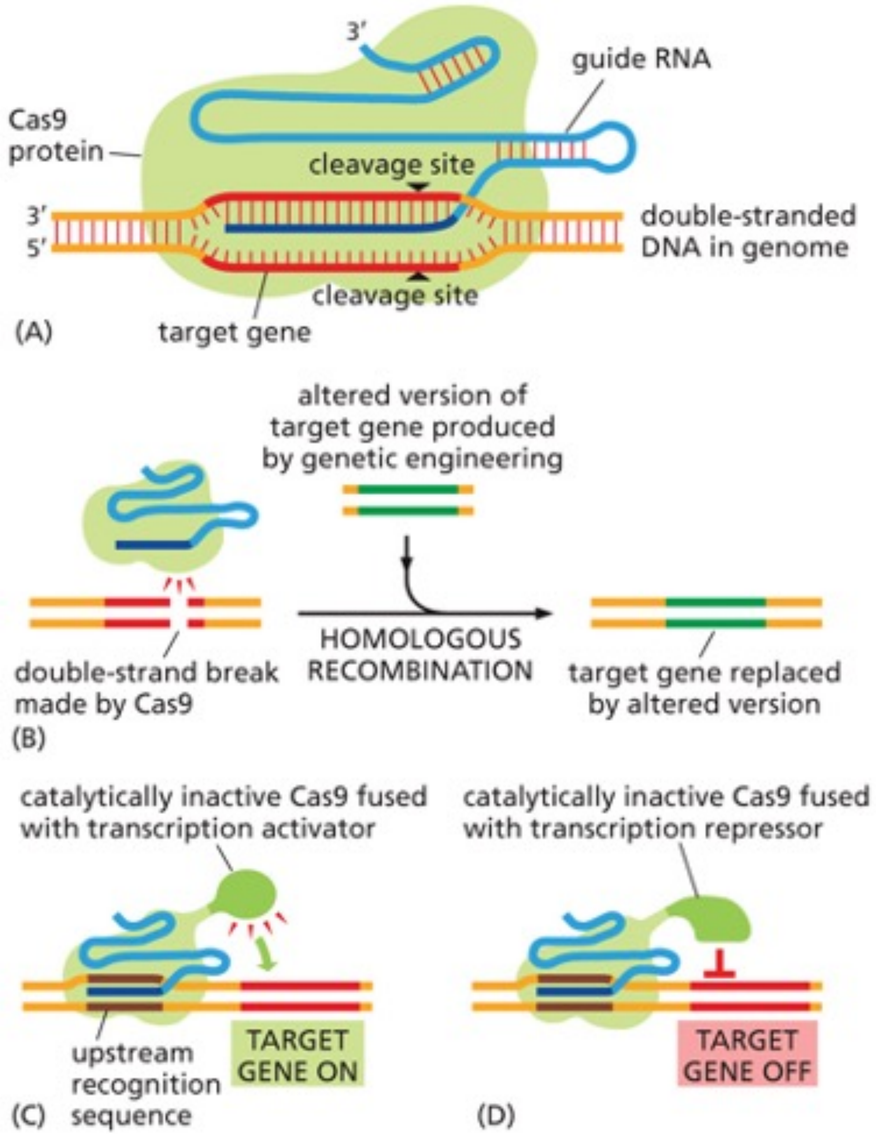
Often inserts at random;
Needs lots of screening



Also remember the Cre-LoxP system, CRISPR, RNAi, …

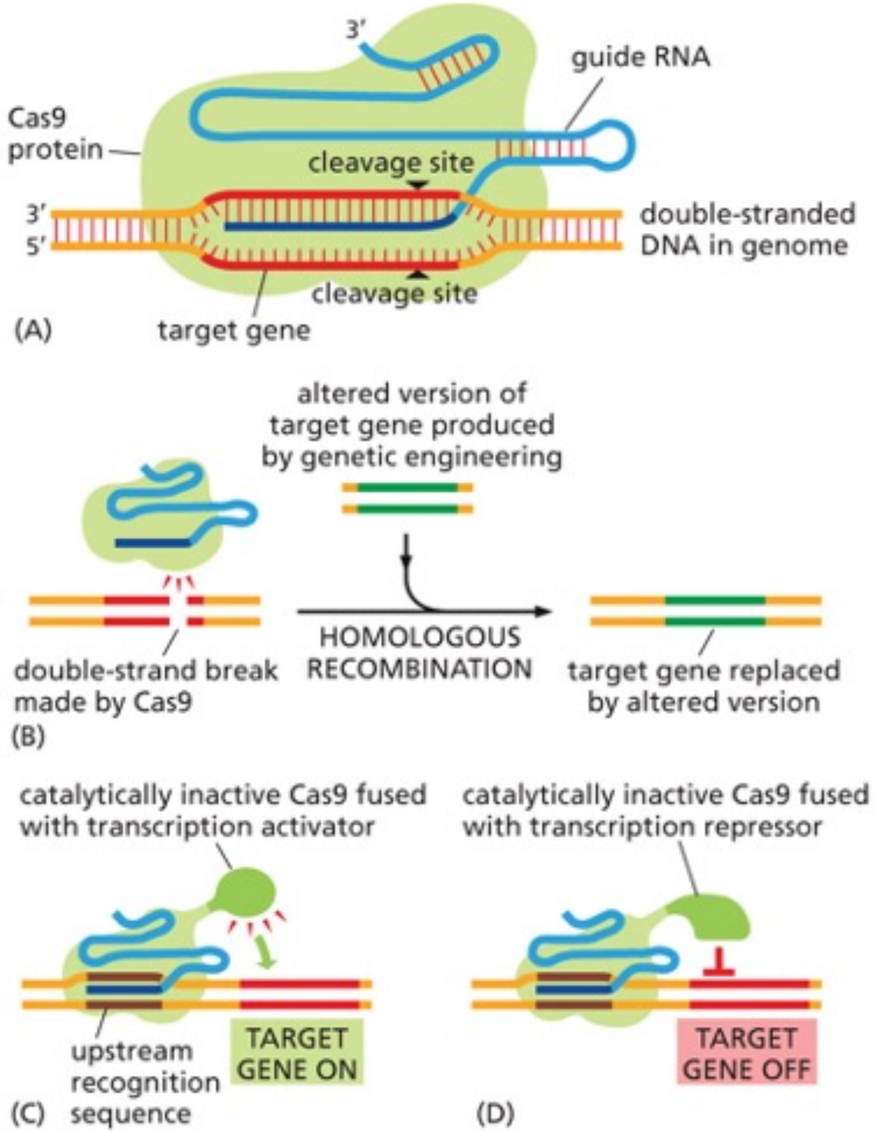
Engineering animals: the CRISPR system

- **Cas9 and guide RNA** are expressed in culture cells
- They associate and mediate **double-strand break** of the chosen DNA region
- This is repaired by non-homologous end joining leading to errors and often a **gene deletion phenotype**
- An **altered target gene** can be provided to be added by homologous recombination
- Can also be added to **turn genes on and off** by fusing Cas9 with transcription regulators



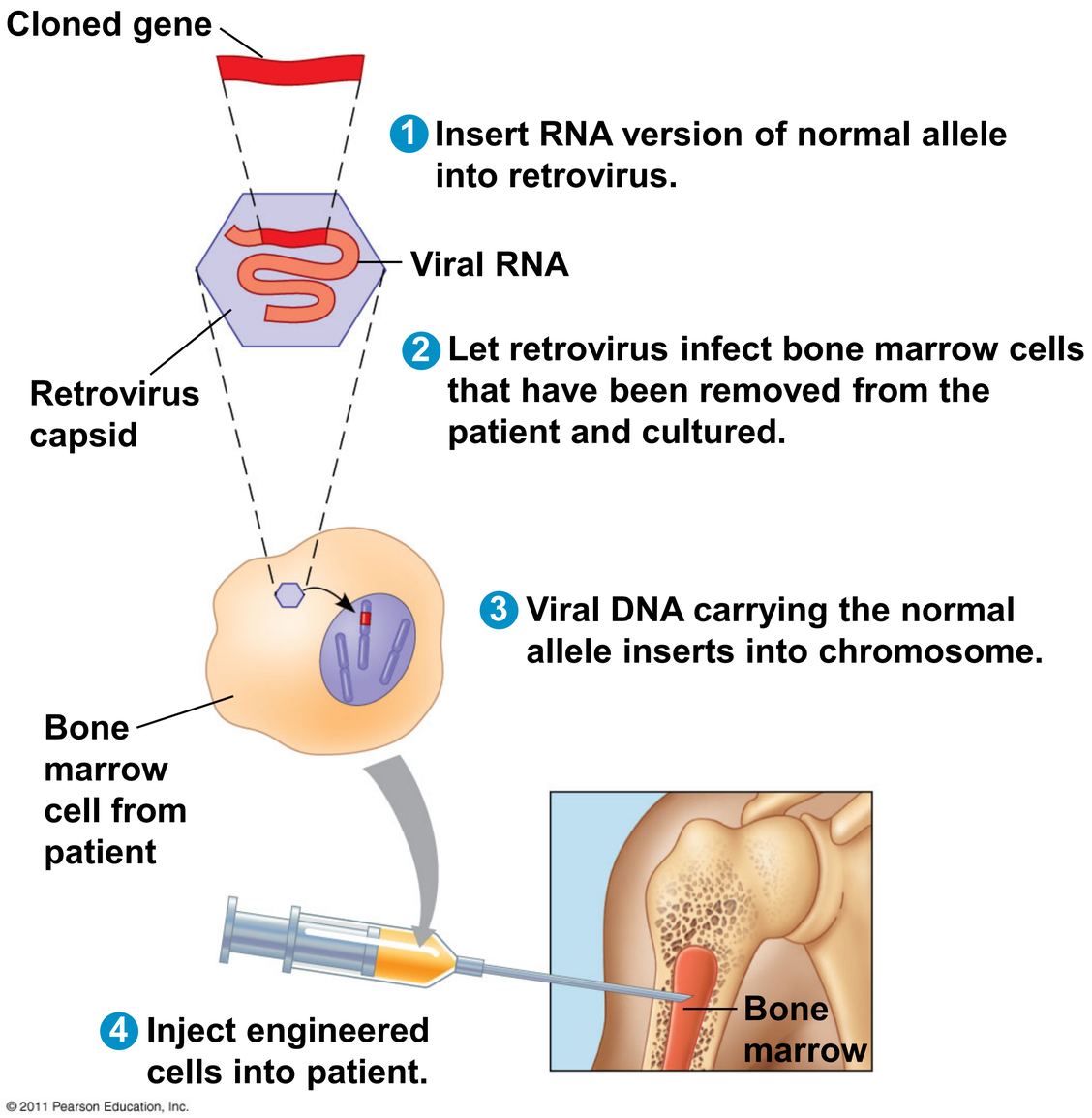
Engineering animals: the CRISPR system

- **easy** to design guide RNAs
- **numerous genes** can be controlled simultaneously (one Cas9, different guide RNAs)
- can virtually be adapted to **all organisms**



Practical applications of DNA-based technology

- Medical application
  - identify genes/mutations involved in **genetic diseases** which are target for **treatment or prevention**
  - **diagnostic** by PCR for disease-causing mutation
  - **gene therapy**, i.e. repairing a “bad” gene - technical and ethical questions



Practical applications of DNA-based technology

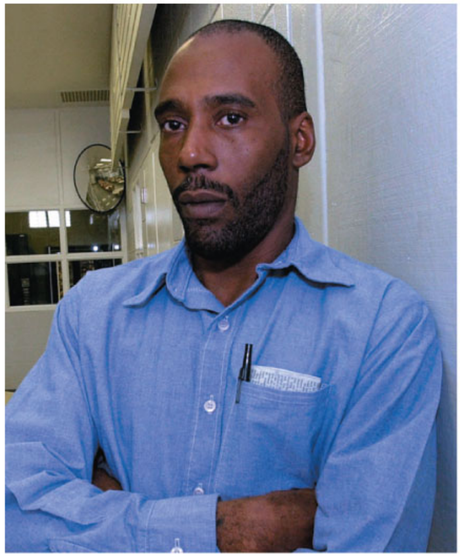
- Pharmacy
  - **transgenic animals** as **pharmaceutical factories**, producing large amounts of rare substances

  ➡ Goat produces anti-thrombin in their milk, which helps patients with genetic disease lacking this enzyme and that have blood clots in their vessels

Practical applications of DNA-based technology

- Forensic science
  - Each individual has a **unique DNA sequence** (=genetic profile) that can be obtained from body tissues or fluids
  - DNA testing can **identify individuals** with a high degree of certainty
  - As of 2013, more than **300 innocent people** have been released from prison as a result of old DNA analysis

(a) This photo shows Washington just before his release in 2001, after 17 years in prison.



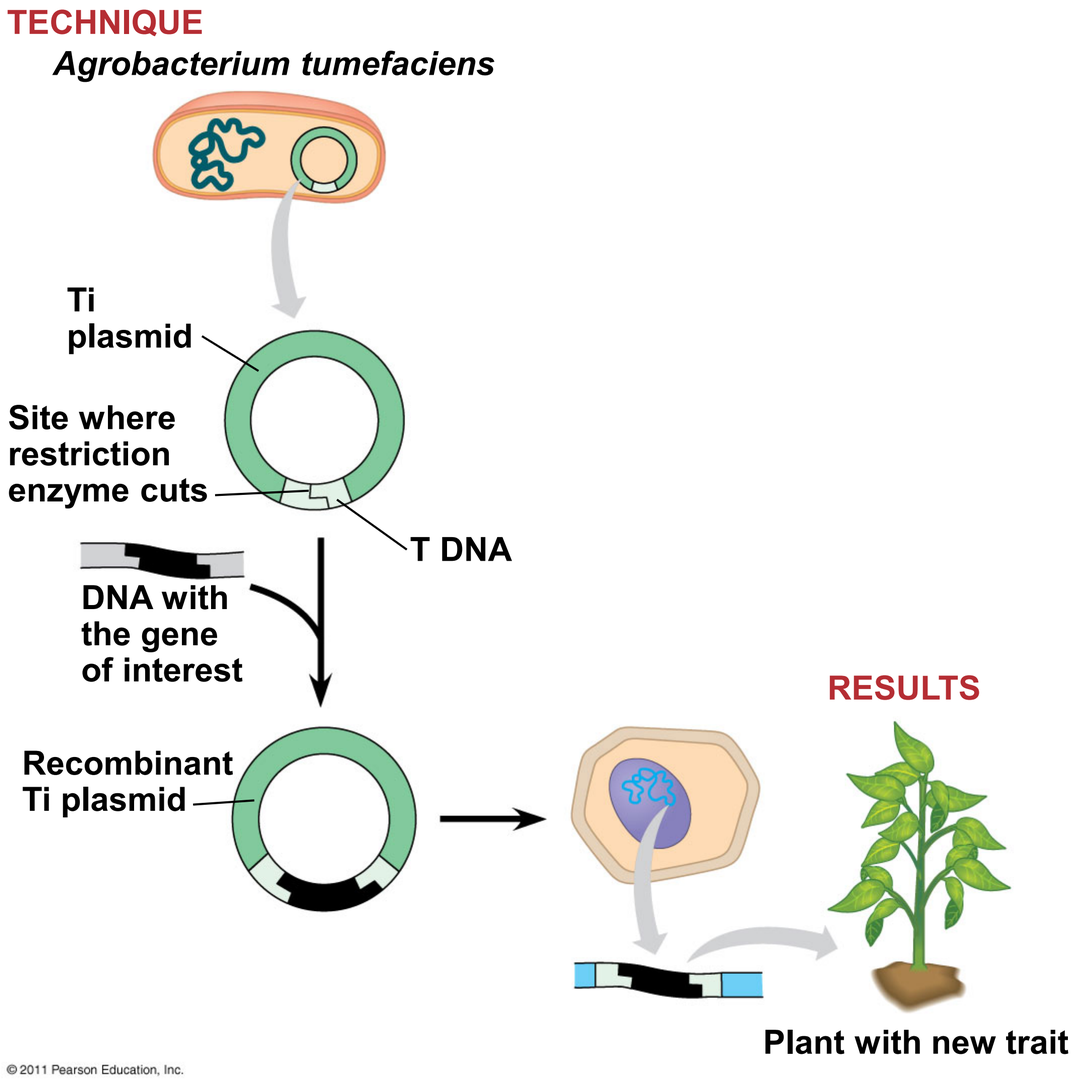
Source of sample	STR marker 1	STR marker 2	STR marker 3
Semen on victim	17,19	13,16	12,12
Earl Washington	16,18	14,15	11,12
Kenneth Tinsley	17,19	13,16	12,12

(b) These and other STR data exonerated Washington and led Tinsley to plead guilty to the murder.

Practical applications of DNA-based technology

- Environmental science
  - Genetic engineering to modify the **metabolism of micro-organisms** (degradation of toxic waste, mineral extraction, …)

- Agriculture
  - Improve **productivity** and **food quality**
  - Herbicide resistance, resistance to pests, resistance to salinity, improved nutritional value, …



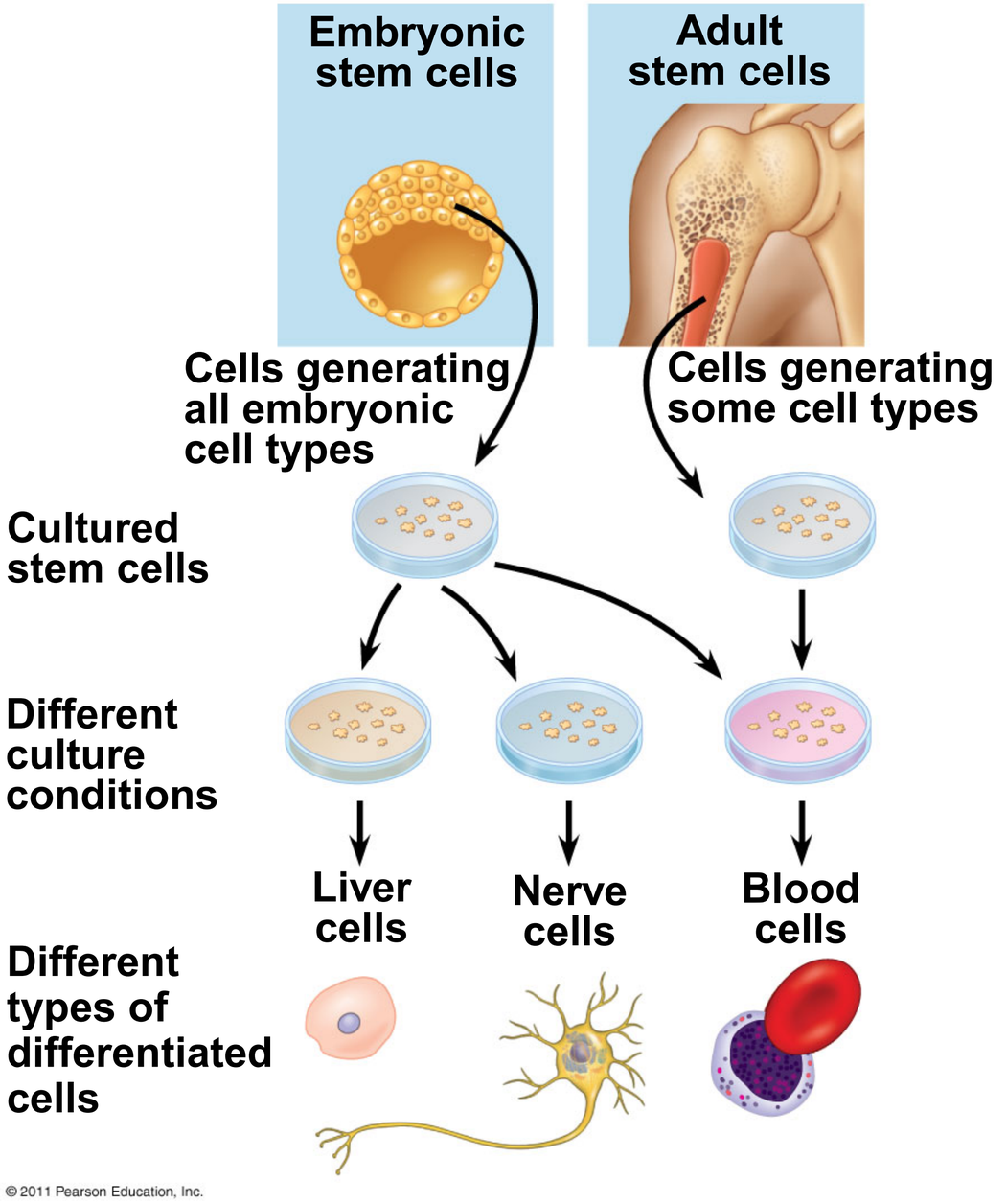
Potential benefits must be weighed against **potential hazards** of creating harmful products or procedures

V. Molecular and Cellular Biology in the lab

1. Model organisms
2. Cell cultures
3. Studying proteins
4. Studying DNA
5. **Stem cells**

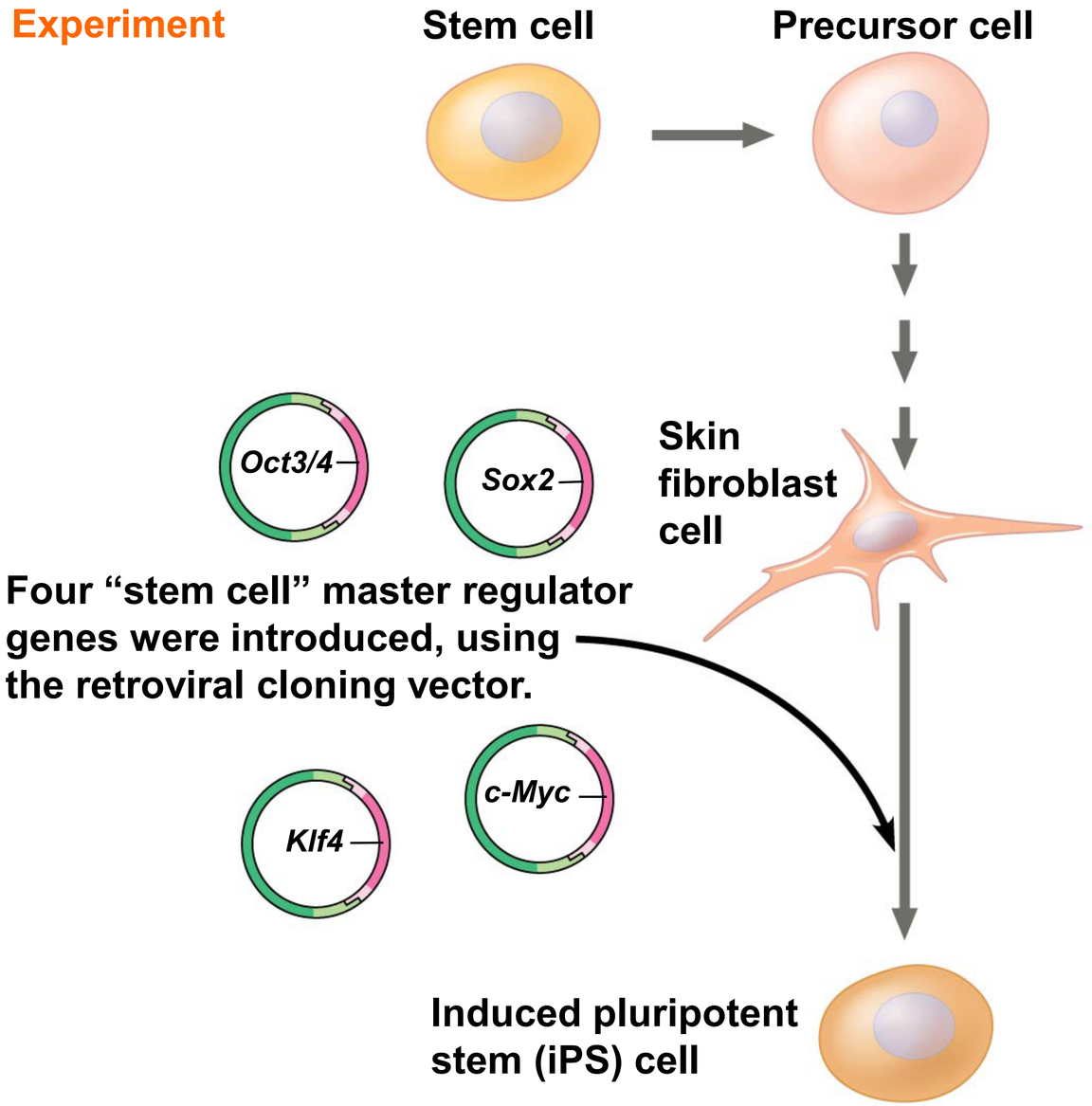
Stem cells (ES)

- **Stem cells** are unspecialized, reproduce themselves indefinitely and differentiate into specialized cells upon specific signals
- Stem cells from early embryos are called **embryonic stem cells (ES)** and can differentiate into **all cell types**
- In **adult** bodies, stem cells replace **non-reproducing specialized cells**
- ES are **pluripotent,** they can differentiate in many cell types
- The aim is to supply cell for the **repair of diseased organs**



Stem cells (iPS)

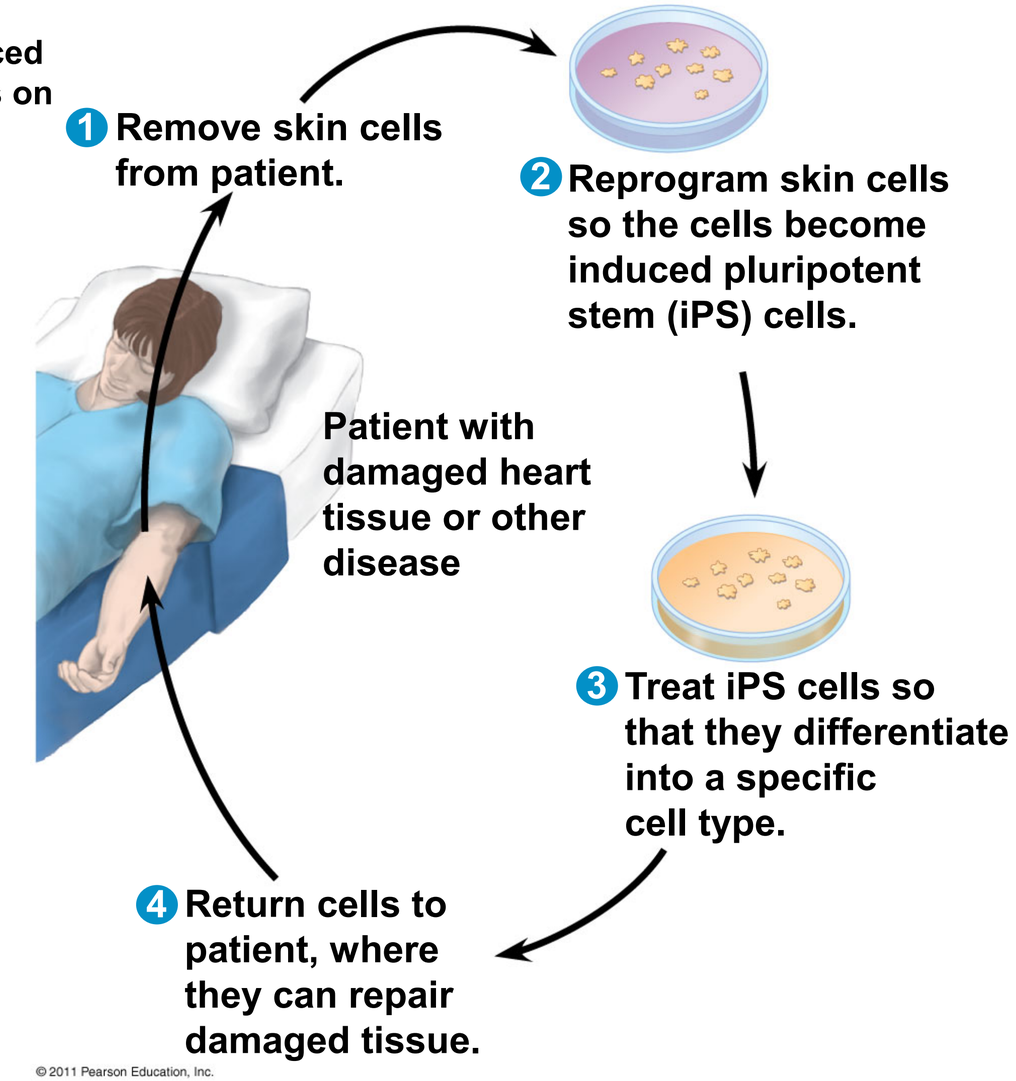
- **Induced pluripotent stem cells (iPS)** are differentiated cells that are reprogrammed to act as ES
- Retroviruses are used to induce extra copies of **4 master regulators** to produce iPS
- They can be used as **models** to study certain diseases



Stem cells (iPS)

- **Induced pluripotent stem cells (iPS)** are differentiated cells that are reprogrammed to act as ES
- Retroviruses are used to induce extra copies of **4 master regulators** to produce iPS
- They can be used as **models** to study certain diseases

Impact: The Impact of Induced Pluripotent Stem (iPS) Cells on Regenerative Medicine

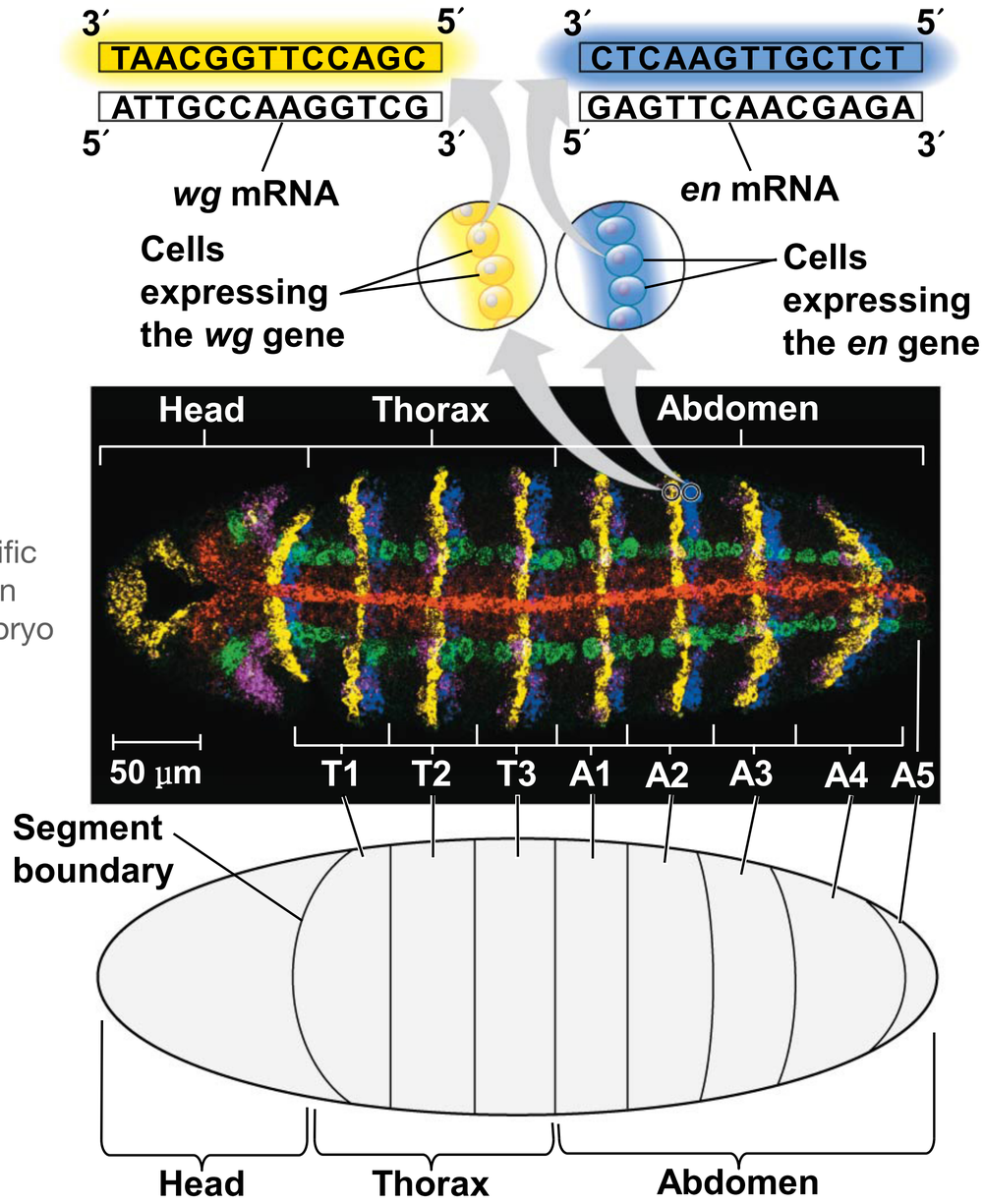


V. Molecular and Cellular Biology in the lab

1. Model organisms
2. Cell cultures
3. Studying proteins
4. Studying DNA
5. Stem cells
6. **Studying gene expression (mRNA)**

How to study gene expression ?

- by looking at the **mRNAs: *in situ* hybridization**
  - Where and when is the **RNA** produced?
  - Based on **nucleic acid hybridization**
  - Tissues are fixed and probes added
  - **No genetic engineering** required
  - Not good for **quantification**

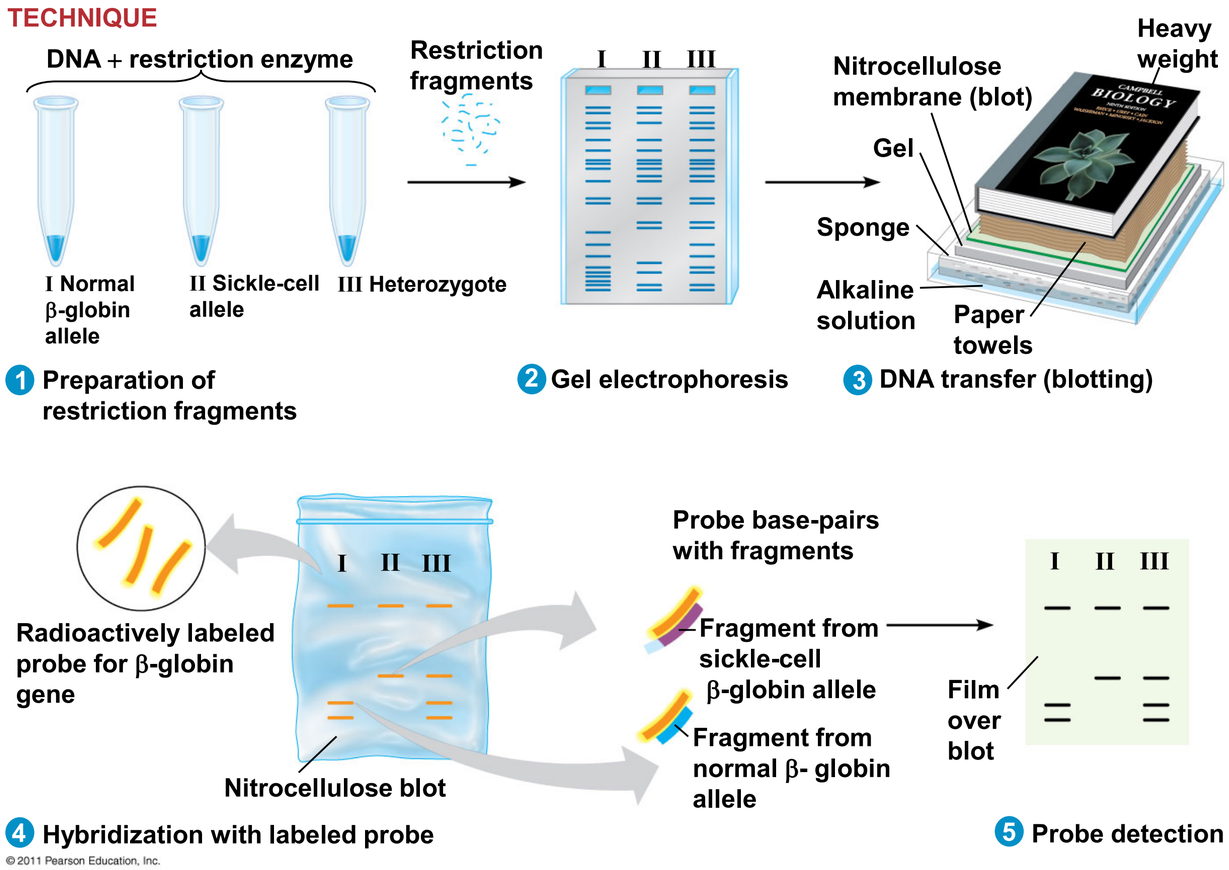


Expression of specific regulatory genes in Drosophila early embryo

Studying gene expression

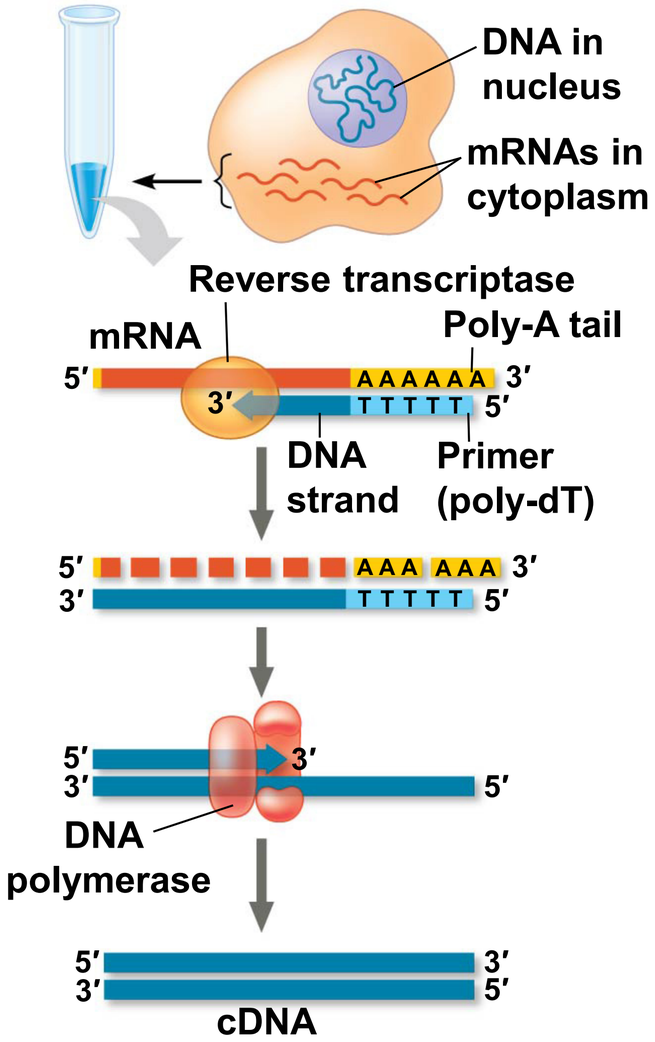
- using **Northern Blotting**: gel electrophoresis of mRNA followed by hybridization with a probe on a membrane

Here, southern blotting but same principle



Studying gene expression

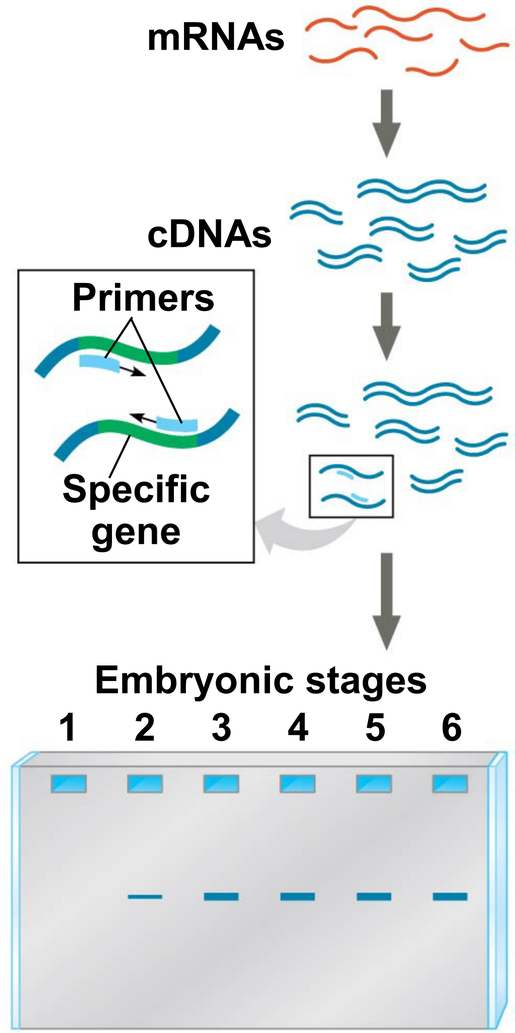
- using Reverse transcriptase-PCR or **RT-PCR**



Technique

1. **cDNA synthesis**
2. **PCR amplification**
3. **Gel electrophoresis**

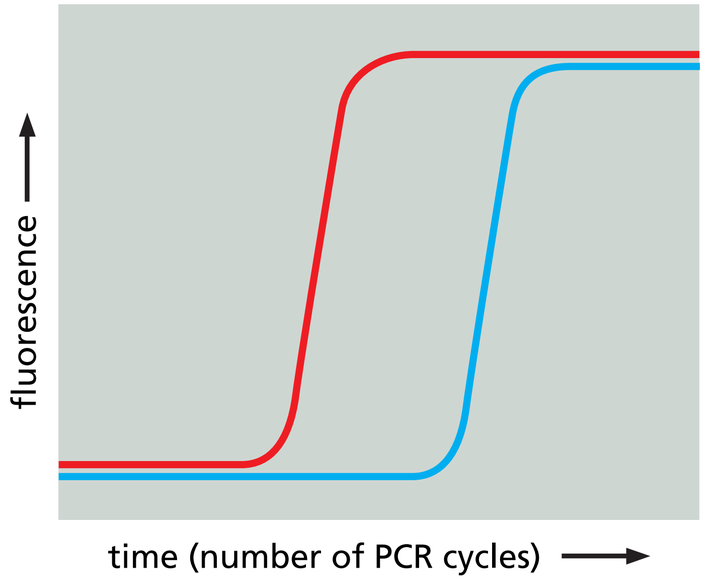
Results



Studying gene expression

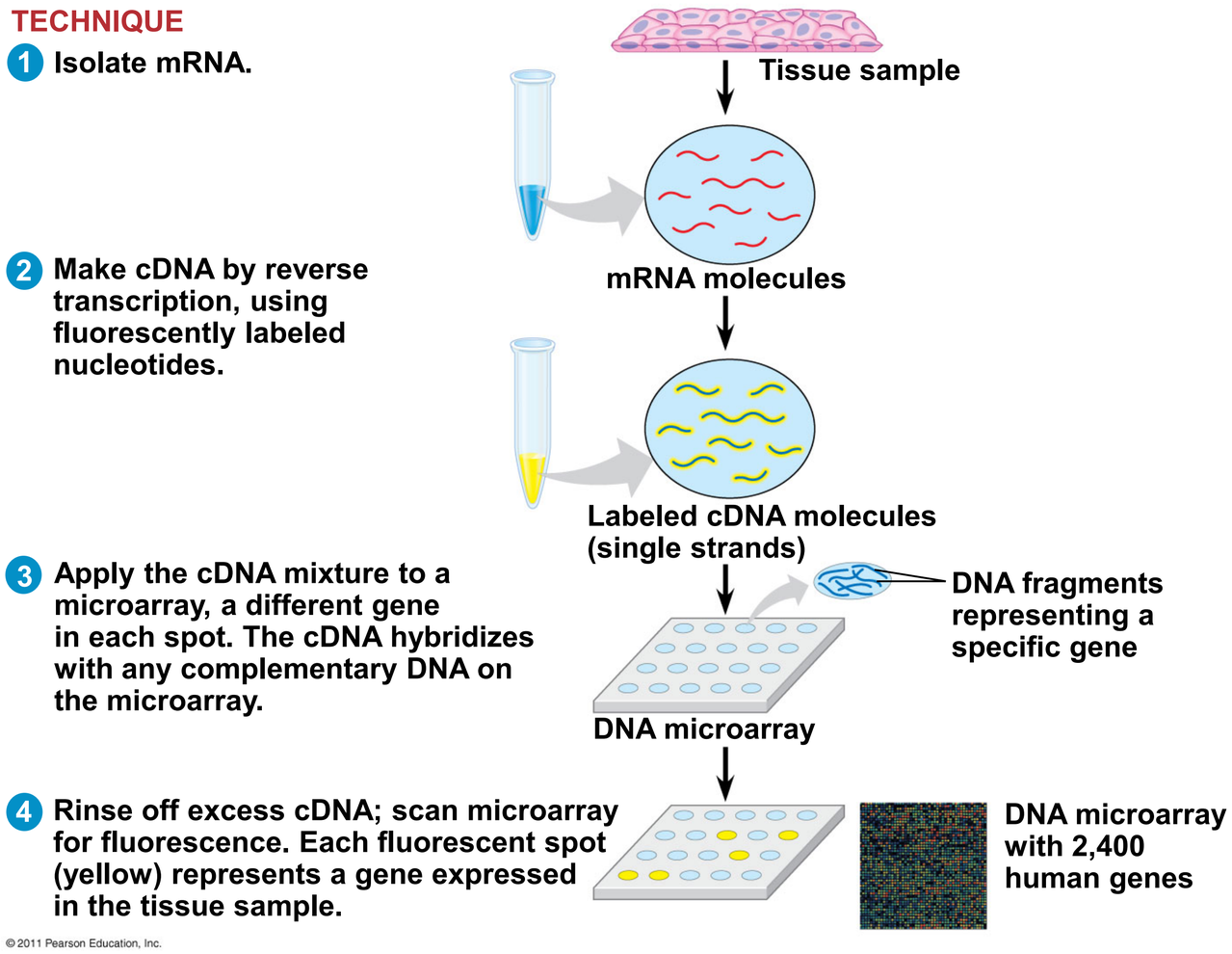
- or quantitative **RT-PCR (qRT-PCR)**

  - **Quantifying** RNAs using **quantitative reverse-transcription polymerase chain reaction**
  - Isolation of the **whole pool of RNAs** from a sample (no DNA!)
  - Addition of **primers specific to the RNA** of interest + reverse transcriptase+ DNA polymerase+ nucleotides
  - Addition of **chemical dyes** that are fluorescent when bound to dsDNA
  - Direct relationship between the **number of PCR cycles needed to detect the product** and **initial amount of RNA** in the sample



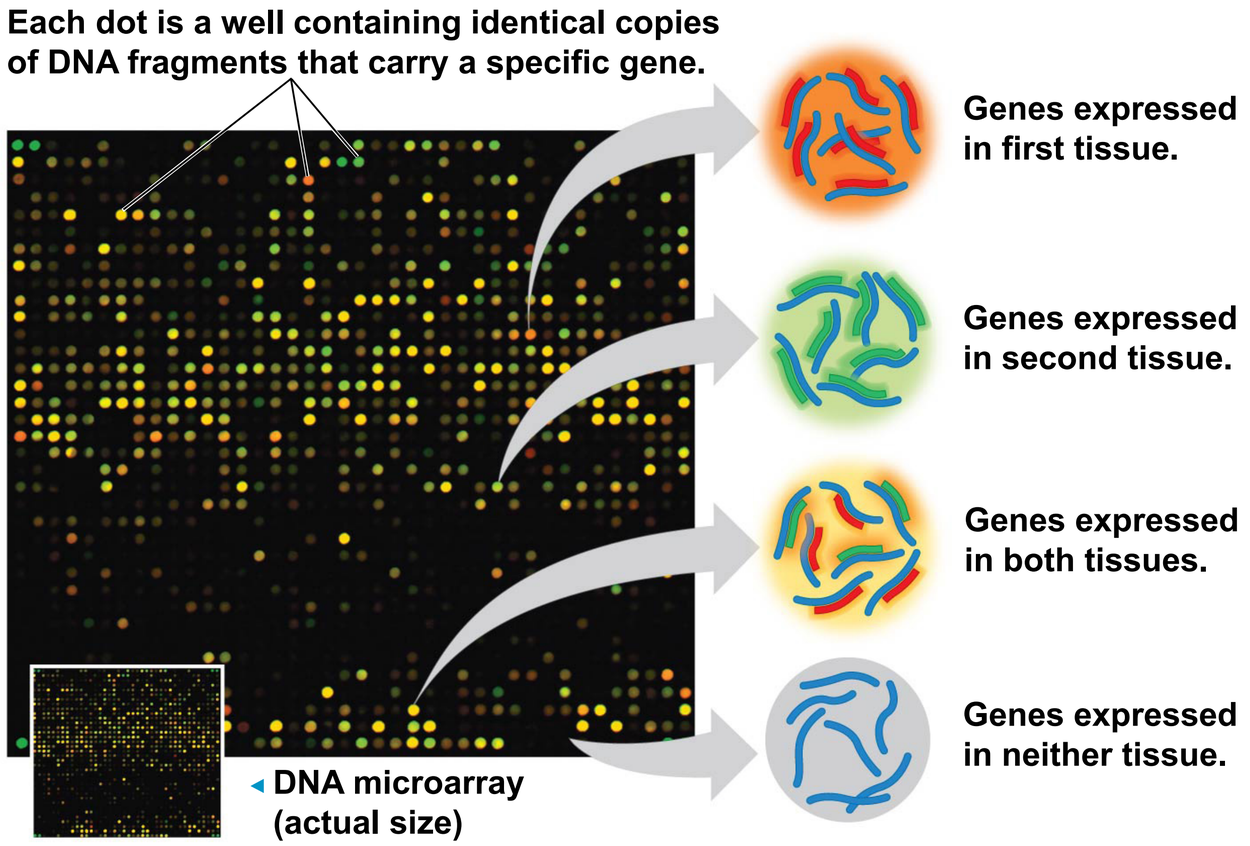
Studying gene expression

- using **microarrays**



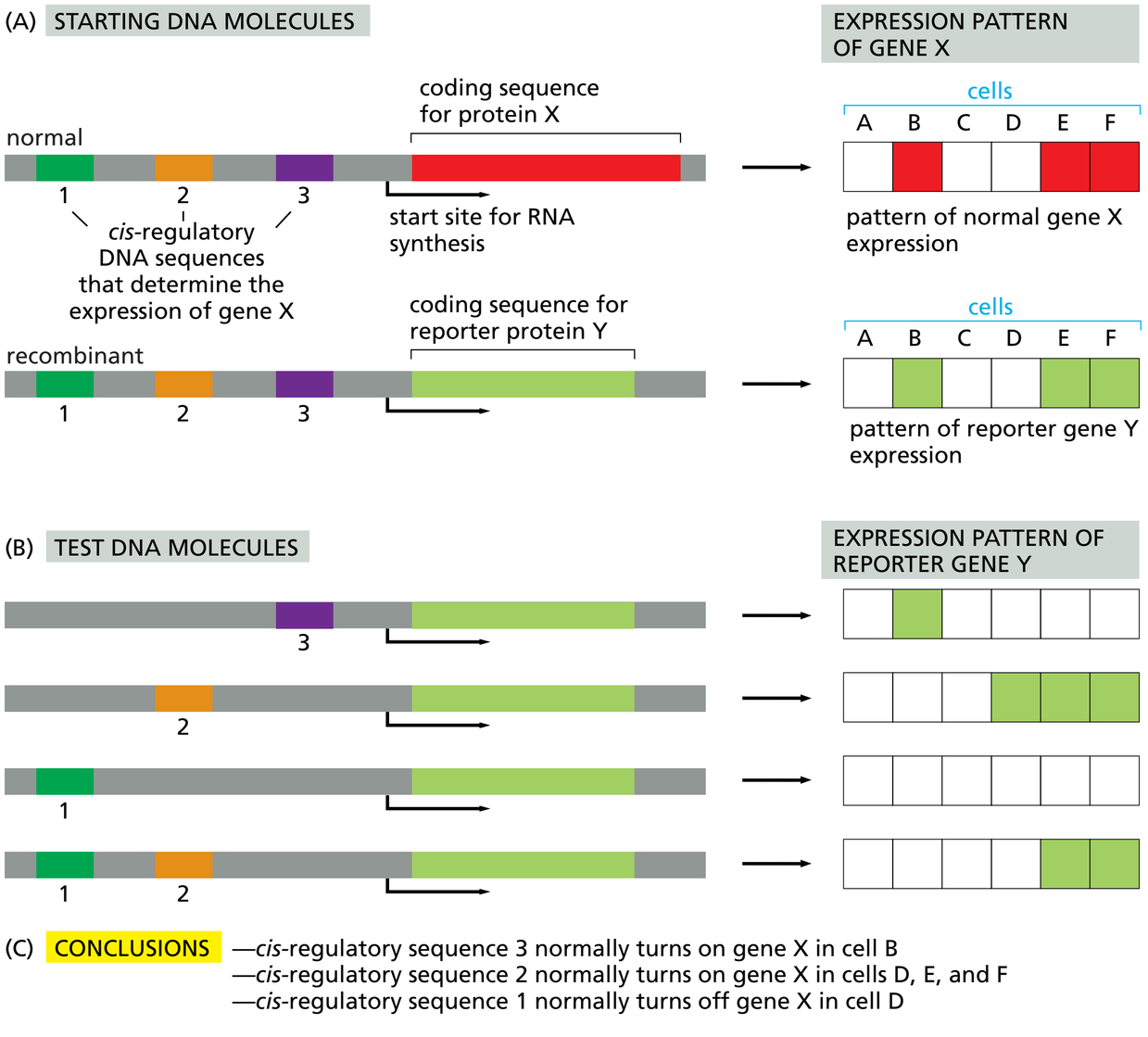
Studying gene expression

- using **microarrays**



Studying gene expression

- using **reporter genes**



Studying gene expression

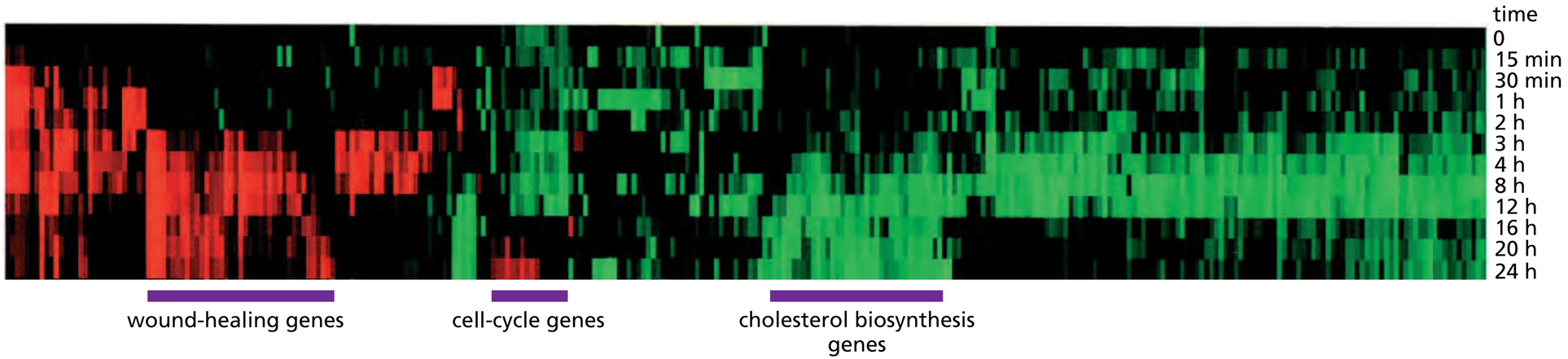
- using **RNA sequencing (RNA-seq)**
  - Measures **which genes are being transcribed** at a given time under given conditions
  - Uses **reverse transcriptase** that to copy all RNAs into cDNAs
  - cDNAs are then **fragmented** and **sequenced** by **next-generation sequencing**
  - Abundant **RNAs** have more **cDNA** copies and therefore more sequencing **reads**

Studying gene expression

- using **RNA sequencing (RNA-seq)**

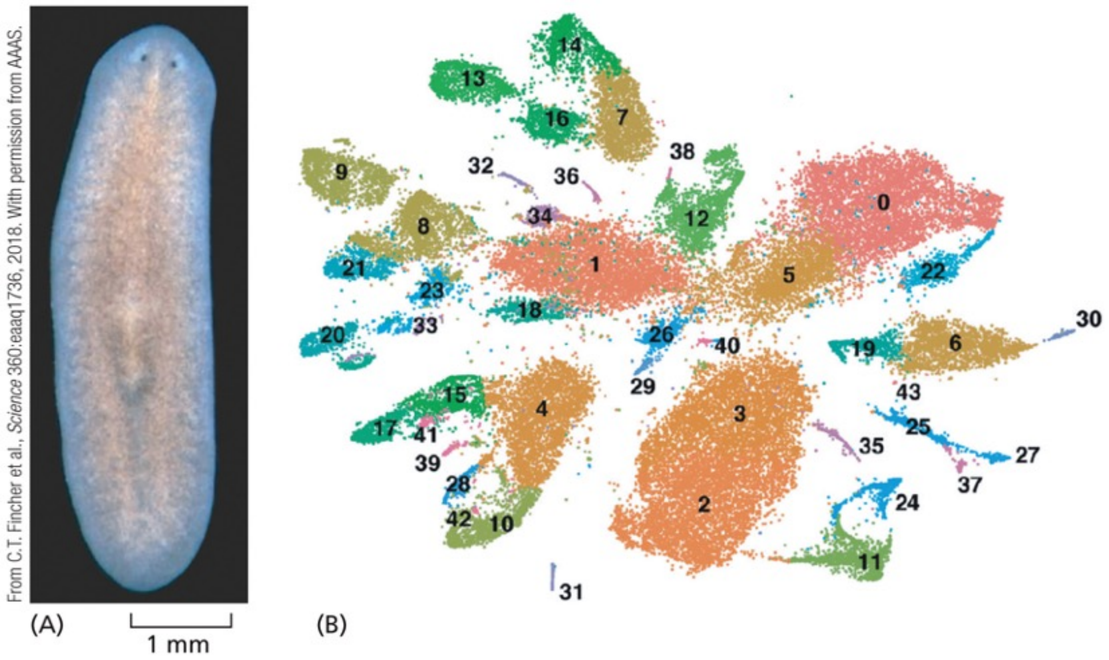
  ➡ **Identification** of RNAs present and **quantification**

  ➡ Using **cluster analysis** (computational approach), one can identify genes that are regulated together



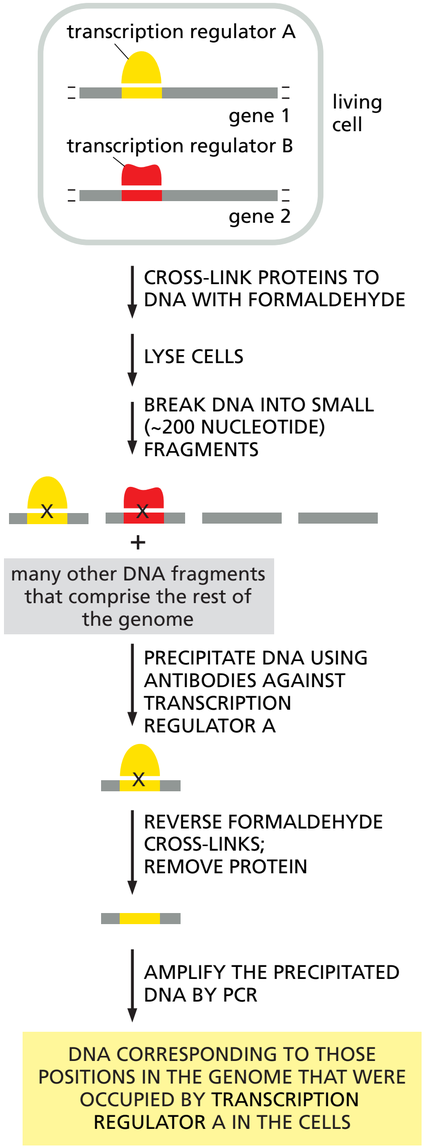
Studying gene expression

- using single cell **RNA sequencing (RNA-seq)**
  - Tissue dissociated into **single cells**
  - **Microfluidics** to separate single cells
  - Each cell is processed for **RNA-seq**
  - **Cluster analysis** algorithm that group cells with similar gene expression patterns



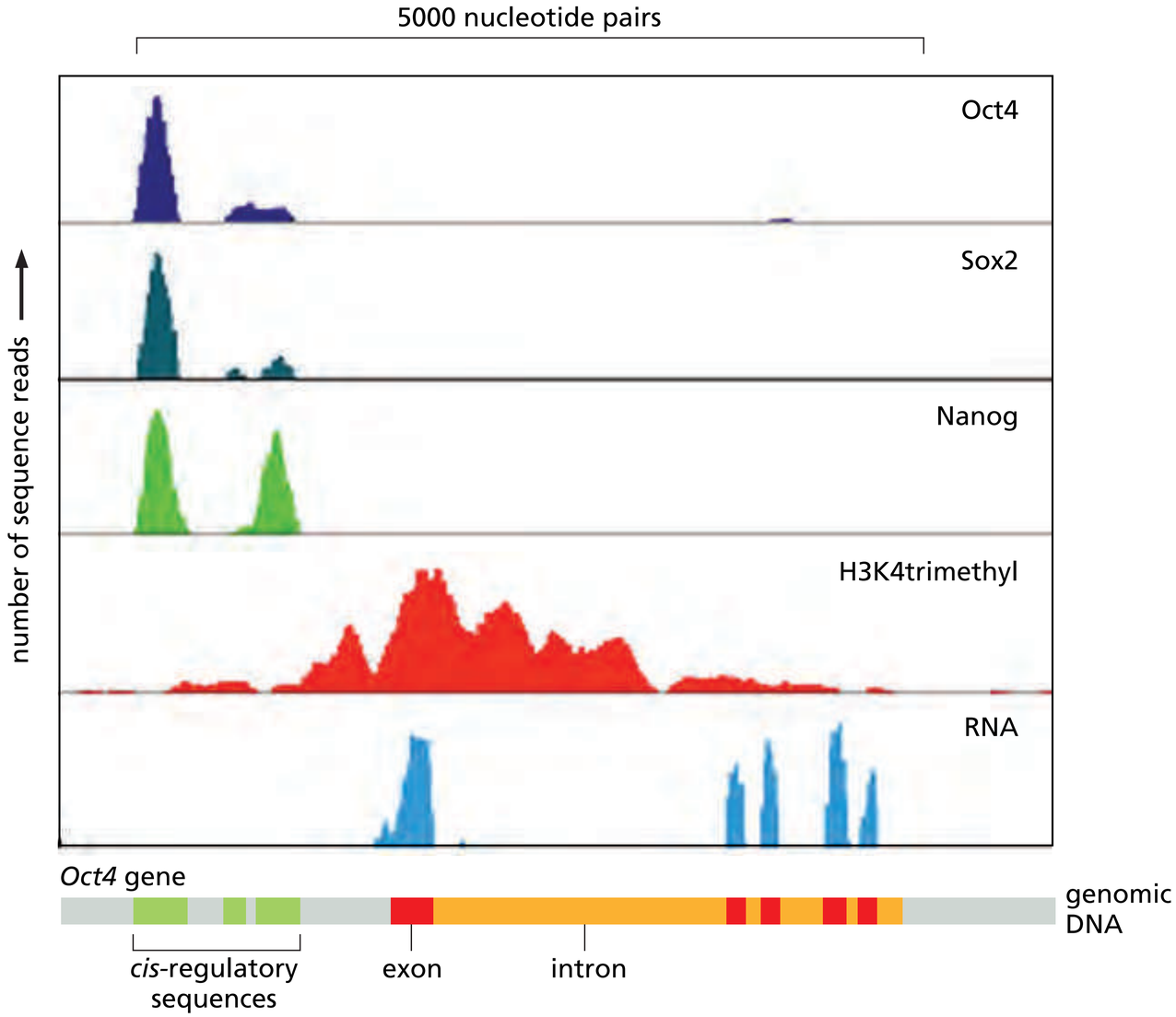
Studying gene expression

- Using **genome-wide chromatin immunoprecipitation** (ChiP) to identify sites on the **genome occupied** by transcription factors
  - **transcription regulators** are responsible for changing **transcription patterns**
  - Proteins are **cross-linked** to DNA
  - Cells are **open**
  - DNA is **fragmented**
  - **Antibodies** that recognise a specific transcription regulator are used to precipitate it with its bound DNA
  - DNA is **sequenced**
  - Can also be used to identify positions bound to specific **modified histones** (using antibodies that recognise those modifications)



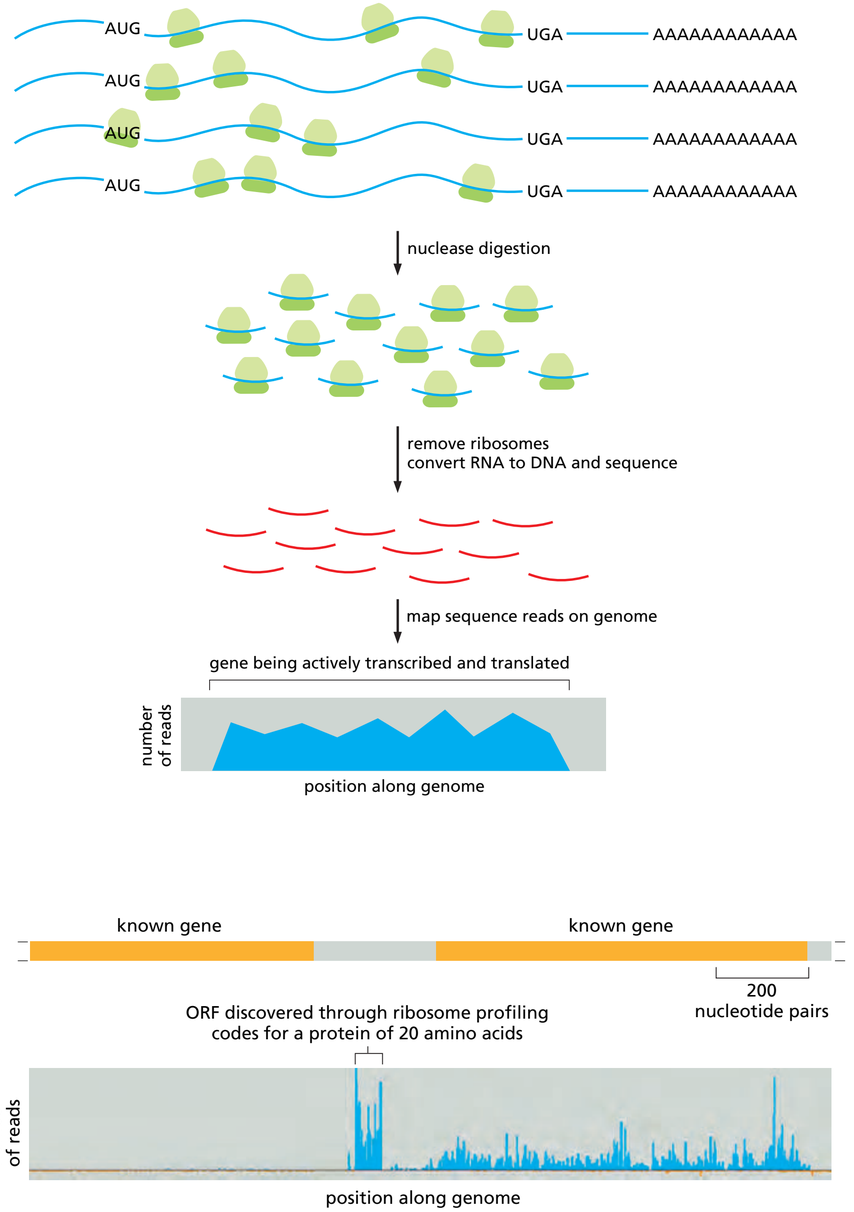
Studying gene expression

- Using **genome-wide chromatin immunoprecipitation** (ChiP) to identify sites on the **genome occupied** by transcription factors



Studying gene expression

- Using **ribosome profiling** to identify **RNAs being transcribed** at a given moment in the cell
  - Gives a map of the **instantaneous position** of ribosomes on each mRNA in the cells
  - Total **RNA** is exposed to a **ribonuclease**
  - RNA sequences **covered by ribosomes** are spared
  - Protected RNAs are converted to **DNA** and **sequenced**
  - Allowed the discovery of **new (small) ORFs**



V. Molecular and Cellular Biology in the lab

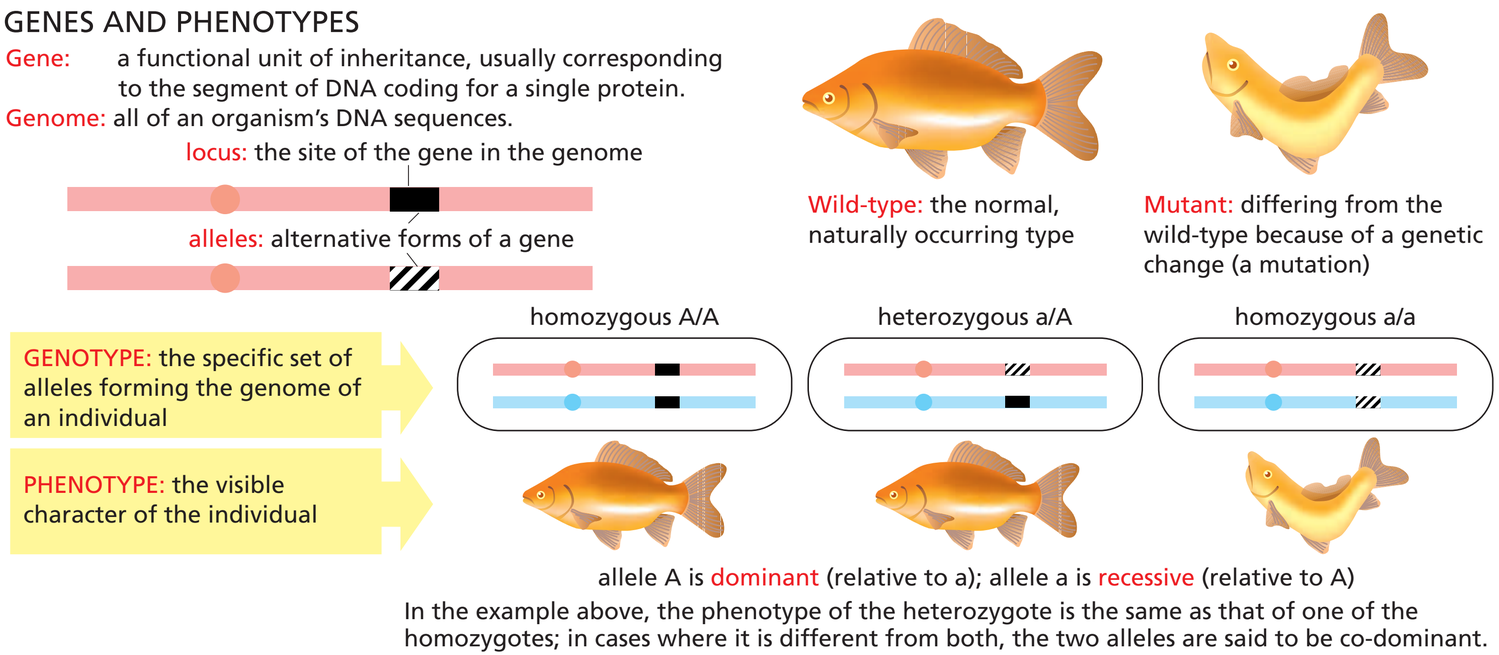
1. Model organisms
2. Cell cultures
3. Studying proteins
4. Studying DNA
5. Stem cells
6. Studying gene expression (mRNA)
7. **Uncovering gene function**

Determining gene function

- The classical approach is to **mutate or delete** a gene and look for associated **phenotype**
- **Mutations/deletions** are introduced by the **cloning methods** described before
- Gene expression can be silenced by **RNA interference**

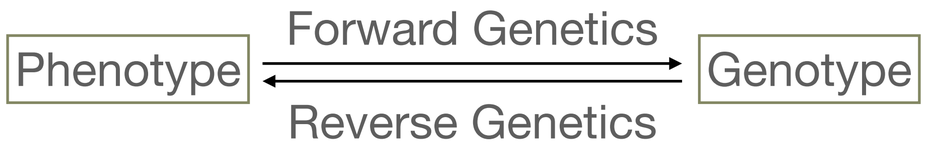
Determining gene function: deletion mutants

- We want to understand how genes (and the proteins they encode) **function**
- One of the most direct way is to **remove the gene (=deletion mutant)** and see what happens
- Basis of **genetics**



Determining gene function: deletion mutants

- We want to understand how genes (and the proteins they encode) **function**
- One of the most direct way is to **remove the gene (=deletion mutant)** and see what happens
- Basis of **genetics**



Determining gene function: creating mutations

- With **chemicals** or **radiations** that mutate DNA
- By **insertional mutagenesis** (with external DNA such as transposons)
- Using gene **cloning** (knock-out, knock-down, CRISPRi, …)

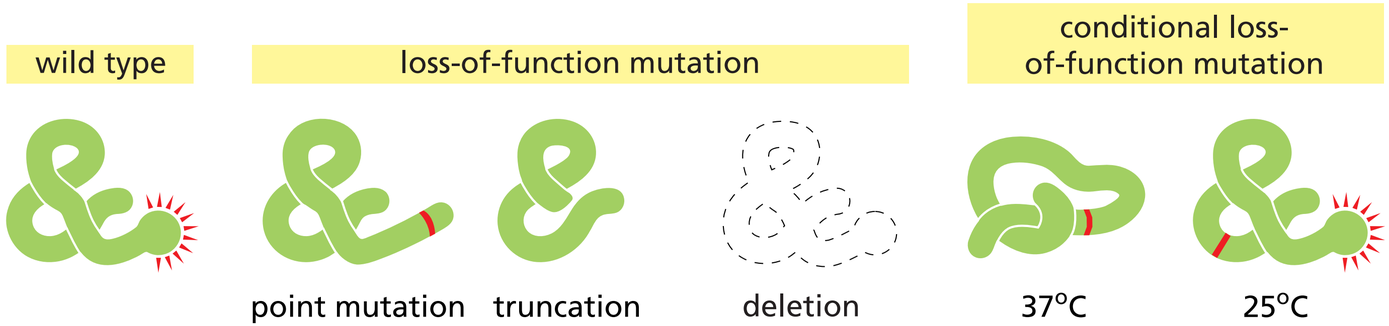
Identification of the mutated gene

- **Insertional mutagenesis:** by PCR
- **For random mutations:** whole genome sequencing

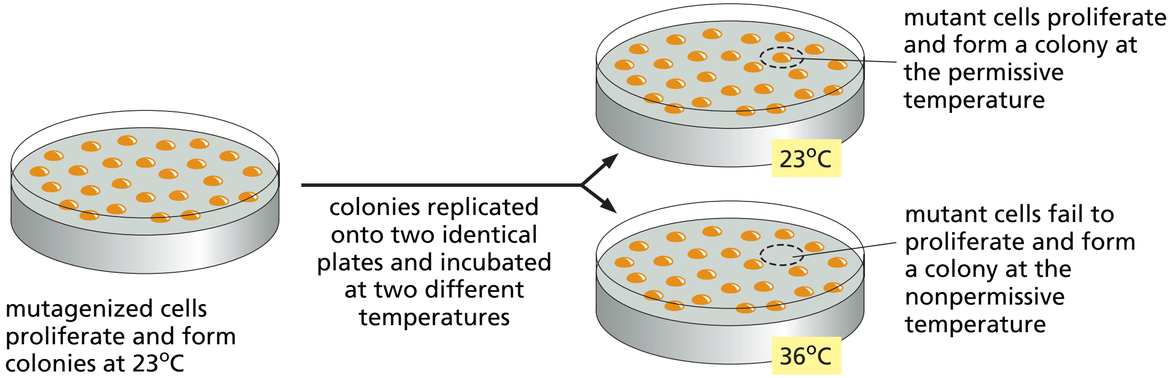
➡ re-introduction of the identified mutation in a clean WT background to prove causality

Gene mutations

- Can lead to gene **loss-of-function** = gene product does not work
- Can lead to gene **gain-of-function** = works too much or in a different way

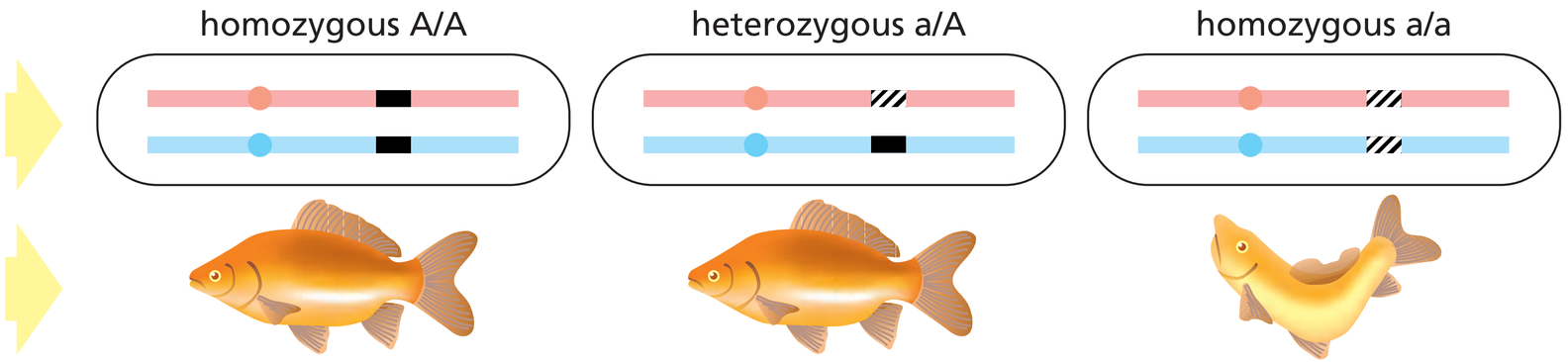


- **Conditional mutant** only shows a phenotype in a given condition



Gene mutations

- A mutation can be **dominant** or **recessive**
- **Dominant mutations** cause the mutant phenotype when only present in one copy
- **Recessive mutations** cannot cause the mutant phenotype when a wild-type copy is also present



allele A is dominant (relative to a); allele a is recessive (relative to A)

In the example above, the phenotype of the heterozygote is the same as that of one of the homozygotes; in cases where it is different from both, the two alleles are said to be co-dominant.

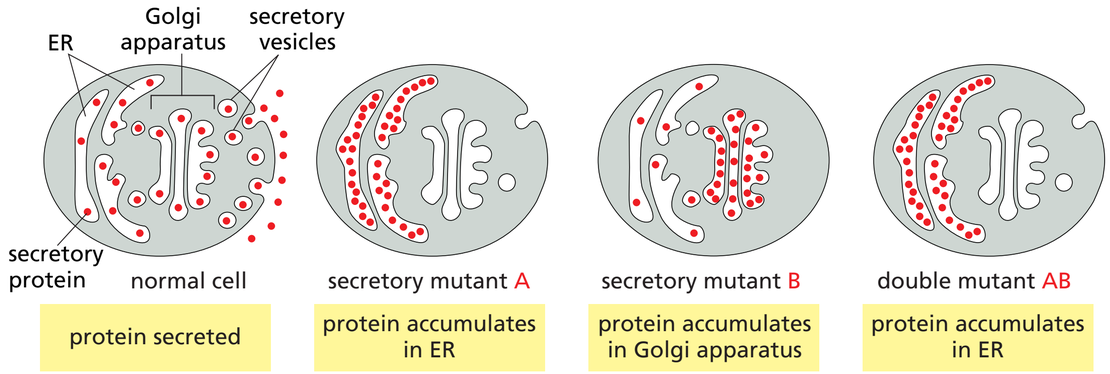
Epistasis analysis

- A **set of genes** participates in a given biological pathway - **in which order?**

Step A gene A	→	Step B gene B	gene A acts upstream of gene B; gene B acts downstream of gene A
Step A gene A*	→	Step B gene B	gene A null mutation — the process is **stopped** at step A even if gene B is functional
Step A gene A	→	Step B gene B*	gene B null mutation — the process is **stopped** at step B

- The term **epistasis** describes a certain relationship between genes, where one gene hides or masks the visible output, or phenotype, of another gene

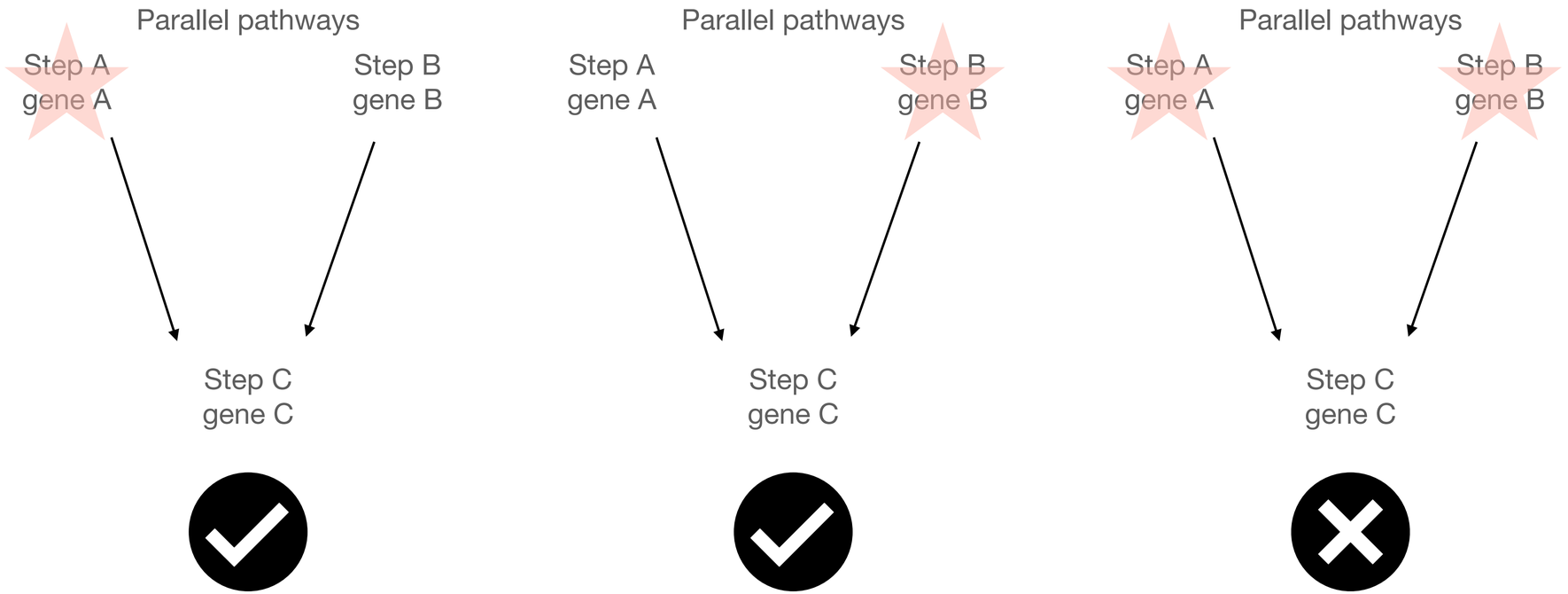
Do secreted proteins first go into the ER or the Golgi?



Epistasis analysis

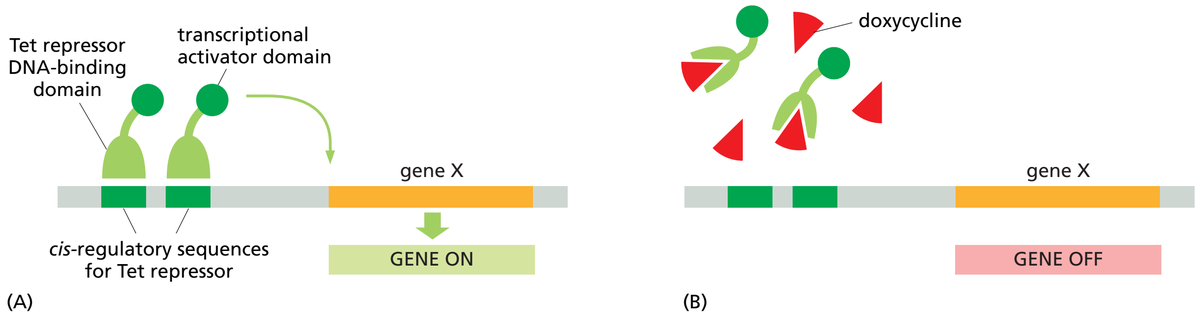
- **Synthetic phenotype**: the phenotype of a double mutant is more severe than each of the single mutants

- **Synthetic lethality**: the phenotype of a double mutant is death, whereas each of the single mutants survives



Starting from a known gene - reverse genetics

- Starting with a gene —> **making mutations** —> observing the phenotype
  - gene **deletions** (or gene knock-out) are only possible is the gene is not essential
  - **inducible** gene expression



  - **cell-type dependent** gene expression
  - **overexpression**
  - any other more **“subtle” mutation** (catalytic site, structure, fusing a marker,…)

Determining gene function: RNA interference

- Introduction of a **dsRNA molecule** with a sequence that matches the target sequence
- RNA **processing**
- RNA **binding** to complementary sequence (mRNA or non-coding RNA)
- Expression is **reduced**
- Frequently used in ***Drosophila*, mammalian cells or *C. elegans***

Determining gene function: RNA interference

- In *C. elegans*, the dsRNA can be injected directly in the intestine or it can be fed with *E. coli* producing the RNA

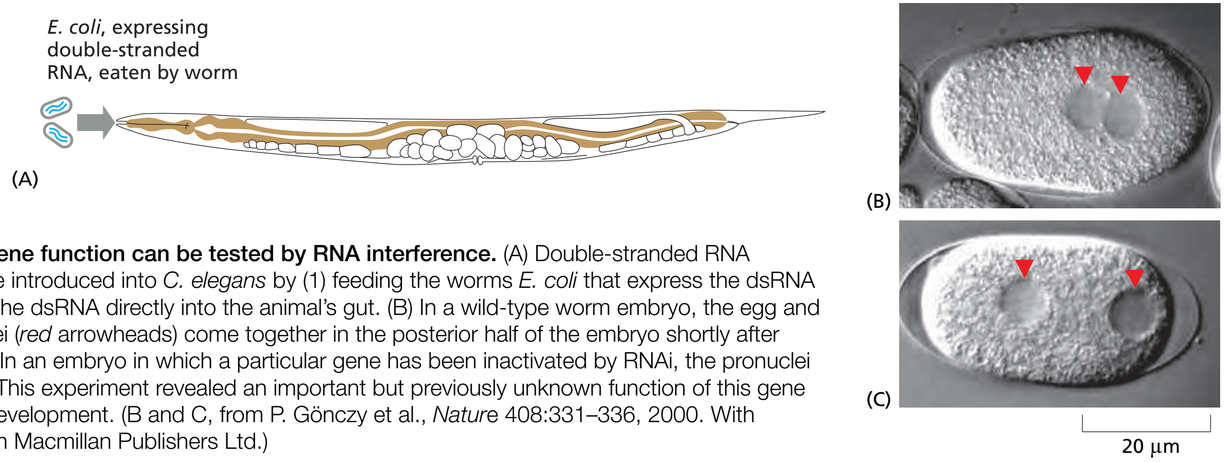
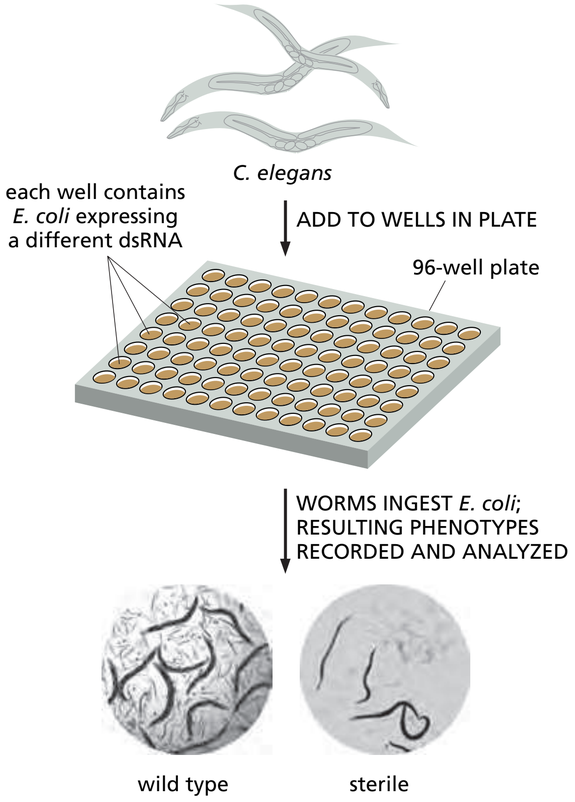


Figure 8–58 Gene function can be tested by RNA interference. (A) Double-stranded RNA (dsRNA) can be introduced into *C. elegans* by (1) feeding the worms *E. coli* that express the dsRNA or (2) injecting the dsRNA directly into the animal's gut. (B) In a wild-type worm embryo, the egg and sperm pronuclei (*red* arrowheads) come together in the posterior half of the embryo shortly after fertilization. (C) In an embryo in which a particular gene has been inactivated by RNAi, the pronuclei fail to migrate. This experiment revealed an important but previously unknown function of this gene in embryonic development. (B and C, from P. Gönczy et al., *Nature* 408:331–336, 2000. With permission from Macmillan Publishers Ltd.)



Building mutant libraries

Molecular Systems Biology (2006) doi:10.1038/msb4100050

www.molecularsystemsbiology.com
Article number: 2006.0008

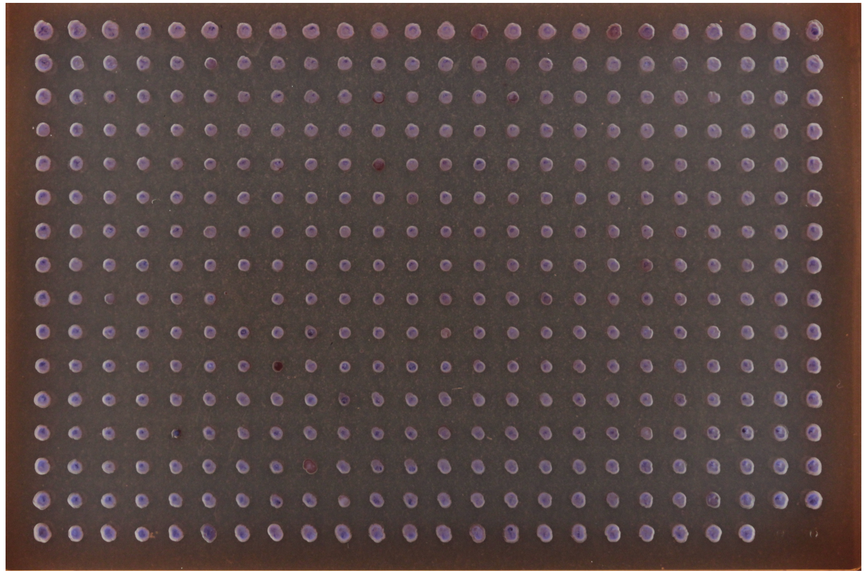
molecular systems biology

Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection

Tomoya Baba[1,2], Takeshi Ara[1], Miki Hasegawa[1,3], Yuki Takai[1,3], Yoshiko Okumura[1], Miki Baba[1], Kirill A Datsenko[4], Masaru Tomita[1], Barry L Wanner[4,*] and Hirotada Mori[1,2,*]

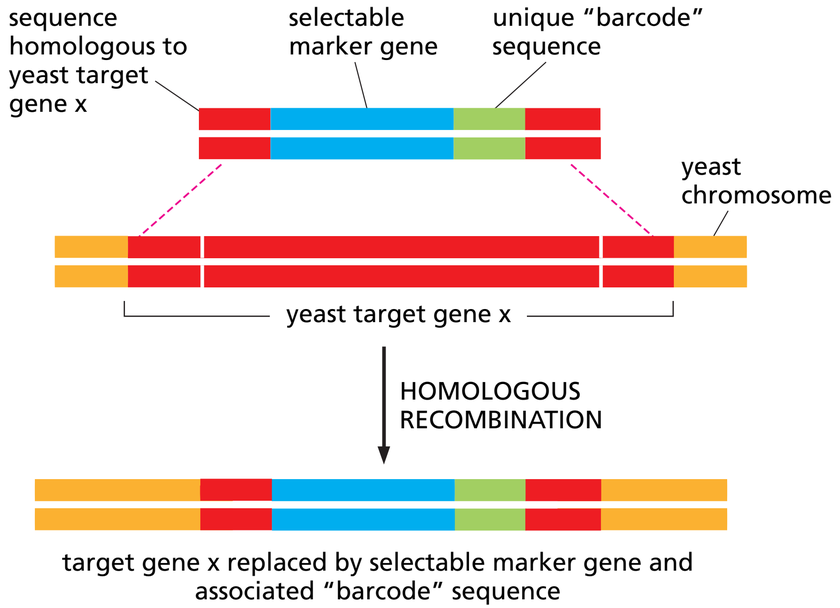
[1] Institute for Advanced Biosciences, Keio University, Tsuruoka City, Yamagata, Japan, [2] Graduate School of Biological Sciences, Nara Institute of Science and Technology, Ikoma, Nara, Japan, [3] CREST, JST (Japan Science and Technology), Kawaguchi, Saitama, Japan and [4] Department of Biological Sciences, Purdue University, West Lafayette, IN, USA
* Corresponding authors. BL Wanner, Department of Biological Sciences, Purdue University, West Lafayette, IN 47907-2054, USA. Tel.: +1 765 494 8034; Fax: +1 765 494 0876: E-mail: blwanner@purdue.edu or H Mori, Graduate School of Biological Sciences, Nara Institute of Science and Technology, 8916-5 Takayama, Ikoma, Nara 630-0101, Japan. Tel.: +81 743 72 5660; Fax: +81 743 72 5669; E-mail: hmori@gtc.naist.jp

Received 28.9.05; accepted 7.12.05



Building mutant libraries

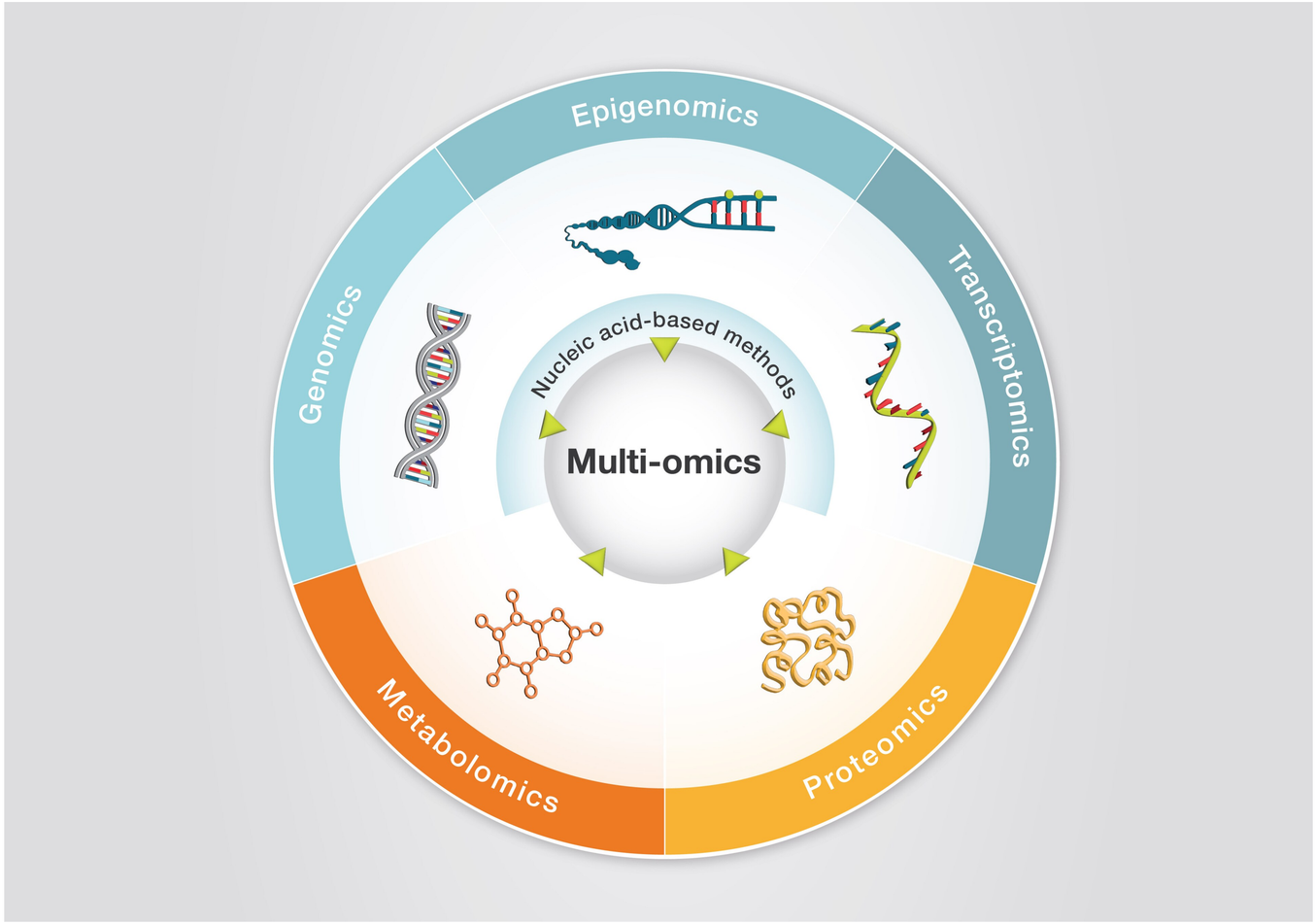
- Efforts to produce this in more **complex organisms**
- Invaluable resource to investigate **gene function on a genomic scale**
- Use of **DNA barcodes** to facilitate mutant identification



V. Molecular and Cellular Biology in the lab

1. Model organisms
2. Cell cultures
3. Studying proteins
4. Studying DNA
5. Stem cells
6. Studying gene expression (mRNA)
7. Uncovering gene function
8. **Omics methods**

(Multi-)omics approaches



(Multi-)omics approaches

“Omics” methods are techniques used to study all the molecules of a certain type in a living organism at once. Instead of looking at one gene or one protein, omics looks at the whole set to see the big picture.

- **Genomics** – studies all the DNA in an organism - based on sequencing
  Question it answers: What genes are there?
- **Transcriptomics** – studies all the RNA produced from DNA - based on sequencing
  Question it answers: Which genes are turned on or off?
- **Proteomics** – studies all the proteins made by the organism - based on mass spectrometry
  Question it answers: What proteins are present and what are they doing?
- **Metabolomics** – studies all the small molecules (metabolites) like sugars, fats, and amino acids - typically based on mass spectrometry
  Question it answers: What chemical reactions are happening in the cell?
- **Epigenomics** – studies chemical modifications on DNA and chromatine that affect how genes are used - based on different methods, including Chip-seq
  Question it answers: How is gene activity controlled without changing the DNA sequence?

Omics methods give a complete, system-wide view of how living things work, helping with disease research, drug development, agriculture, and understanding how organisms respond to the environment.


Have a nice holiday/exam prep !