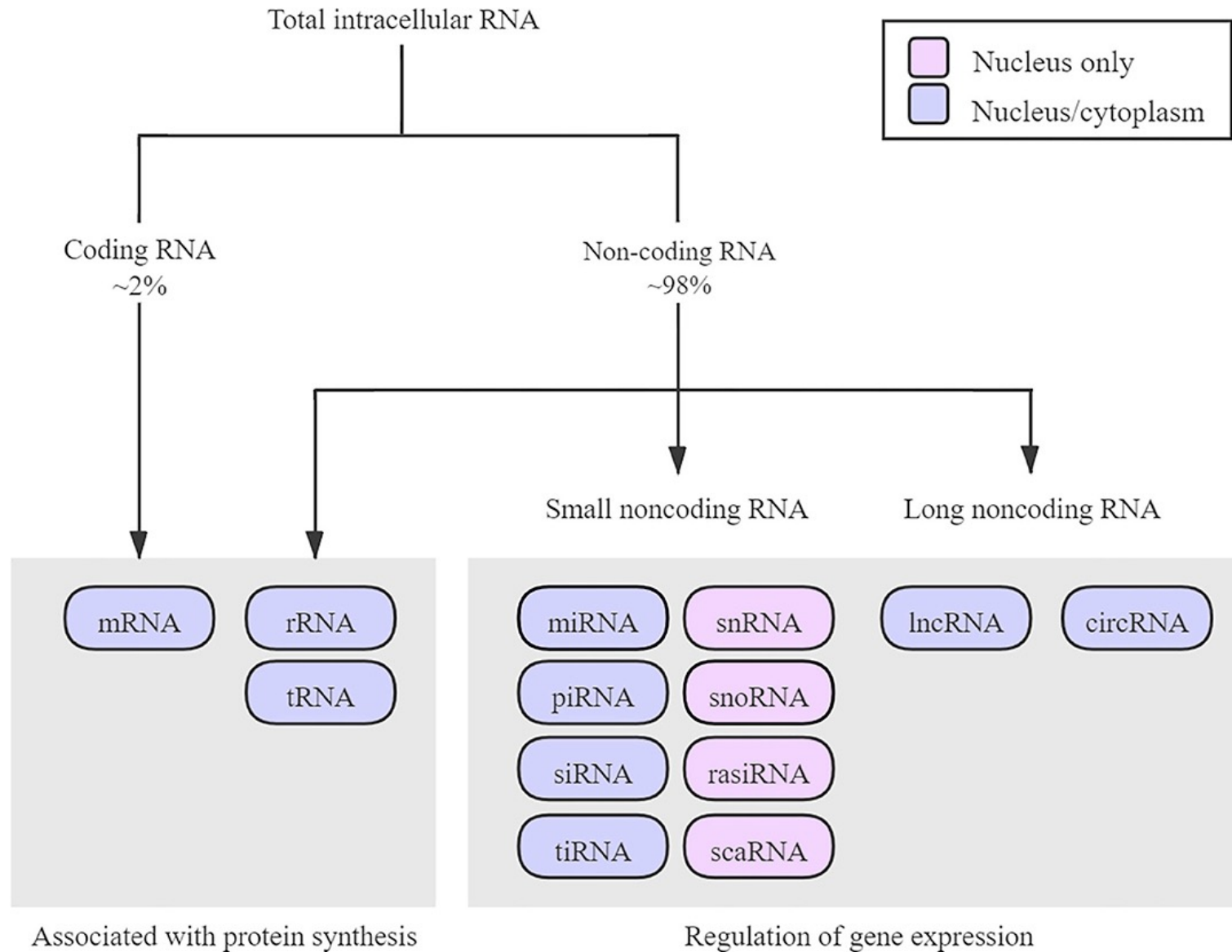


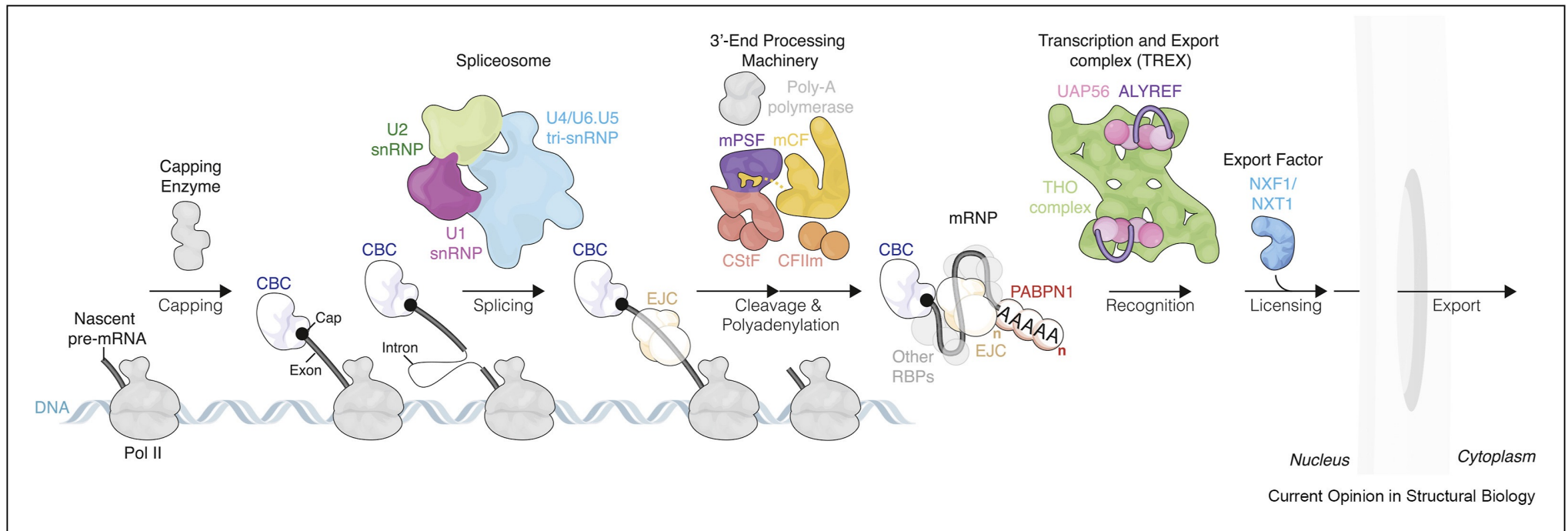
HOW TO STUDY GENE EXPRESSION

RNA classes

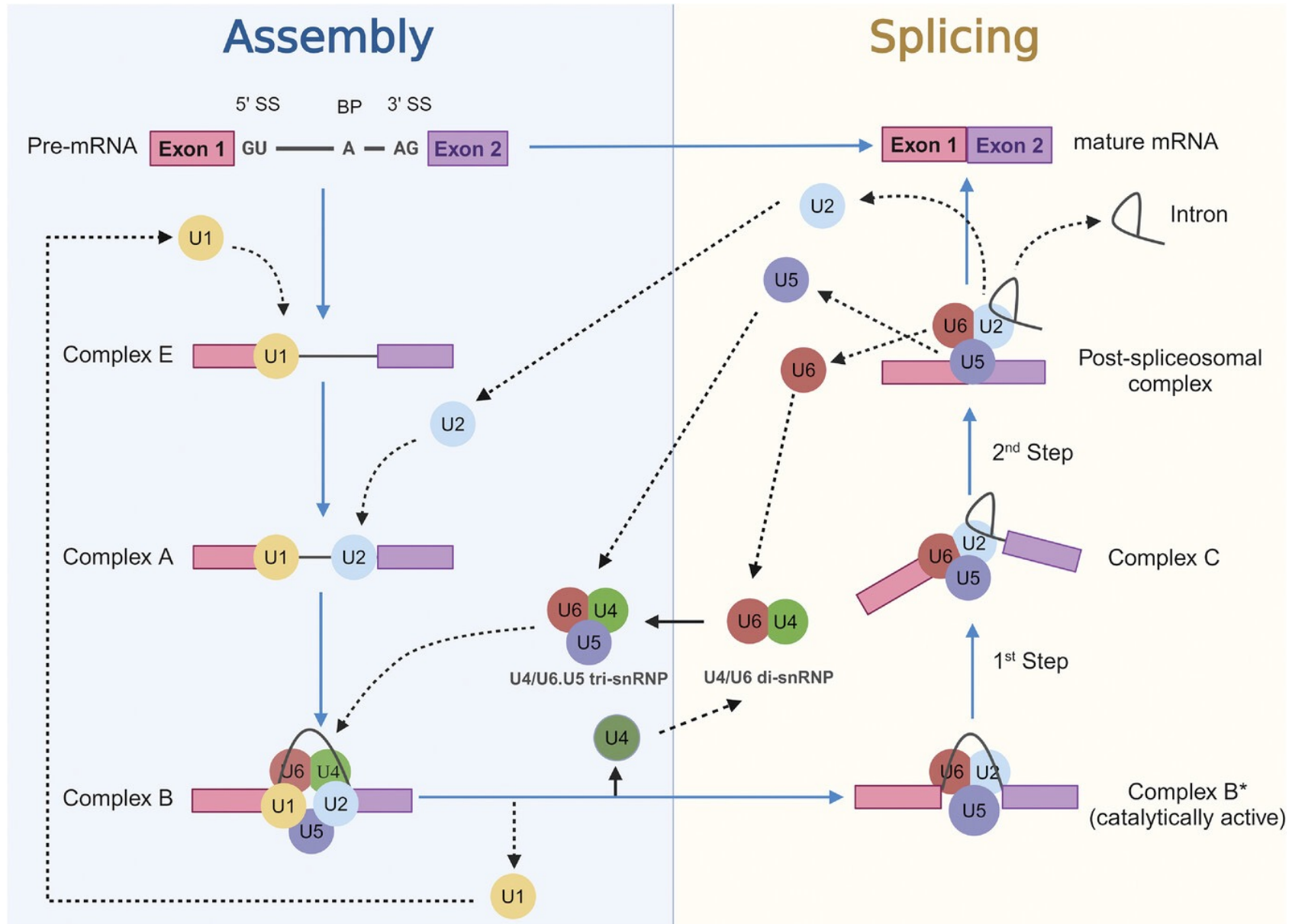


mRNA

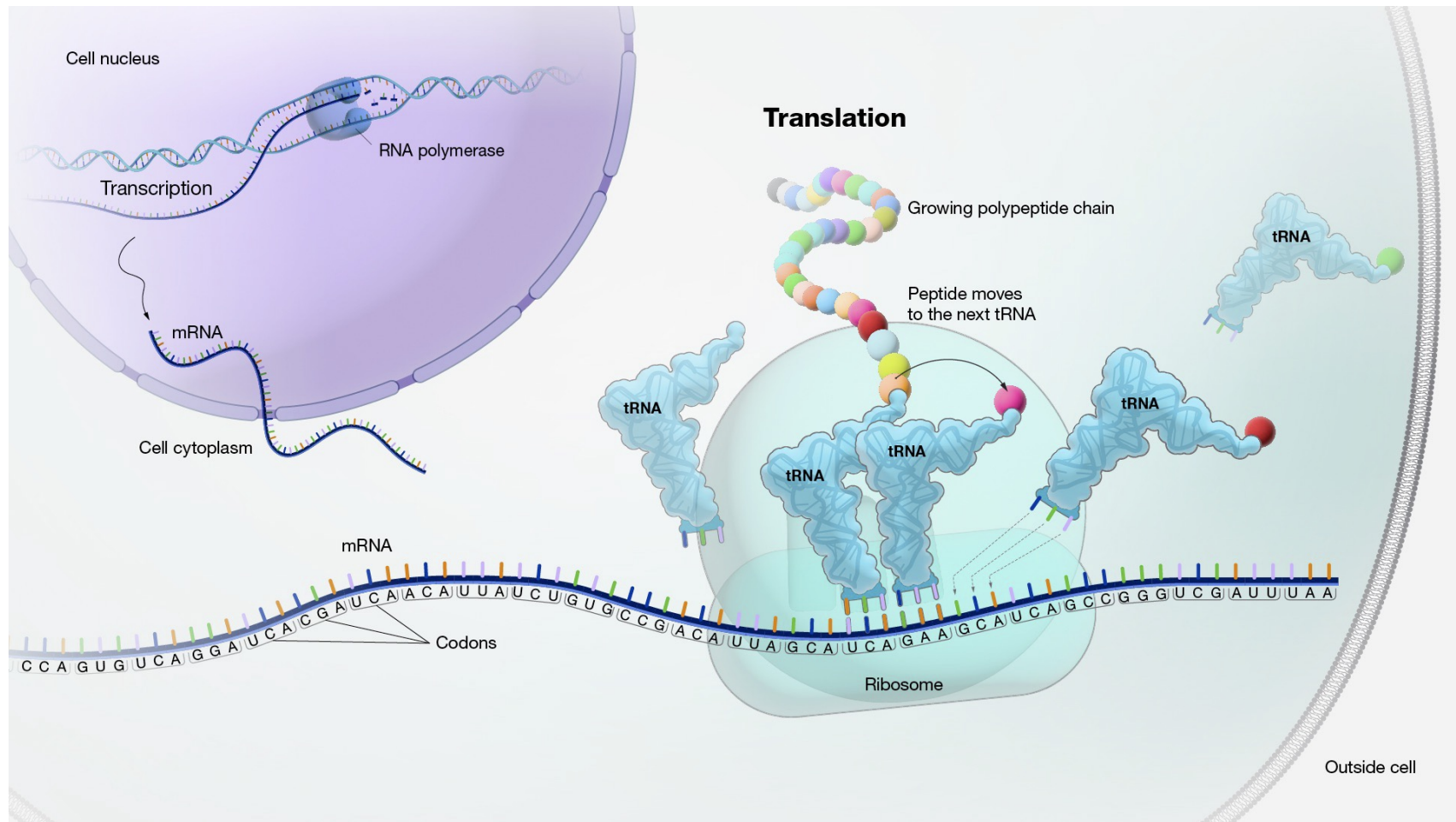
1.5-2% of total RNA



Alternative splicing

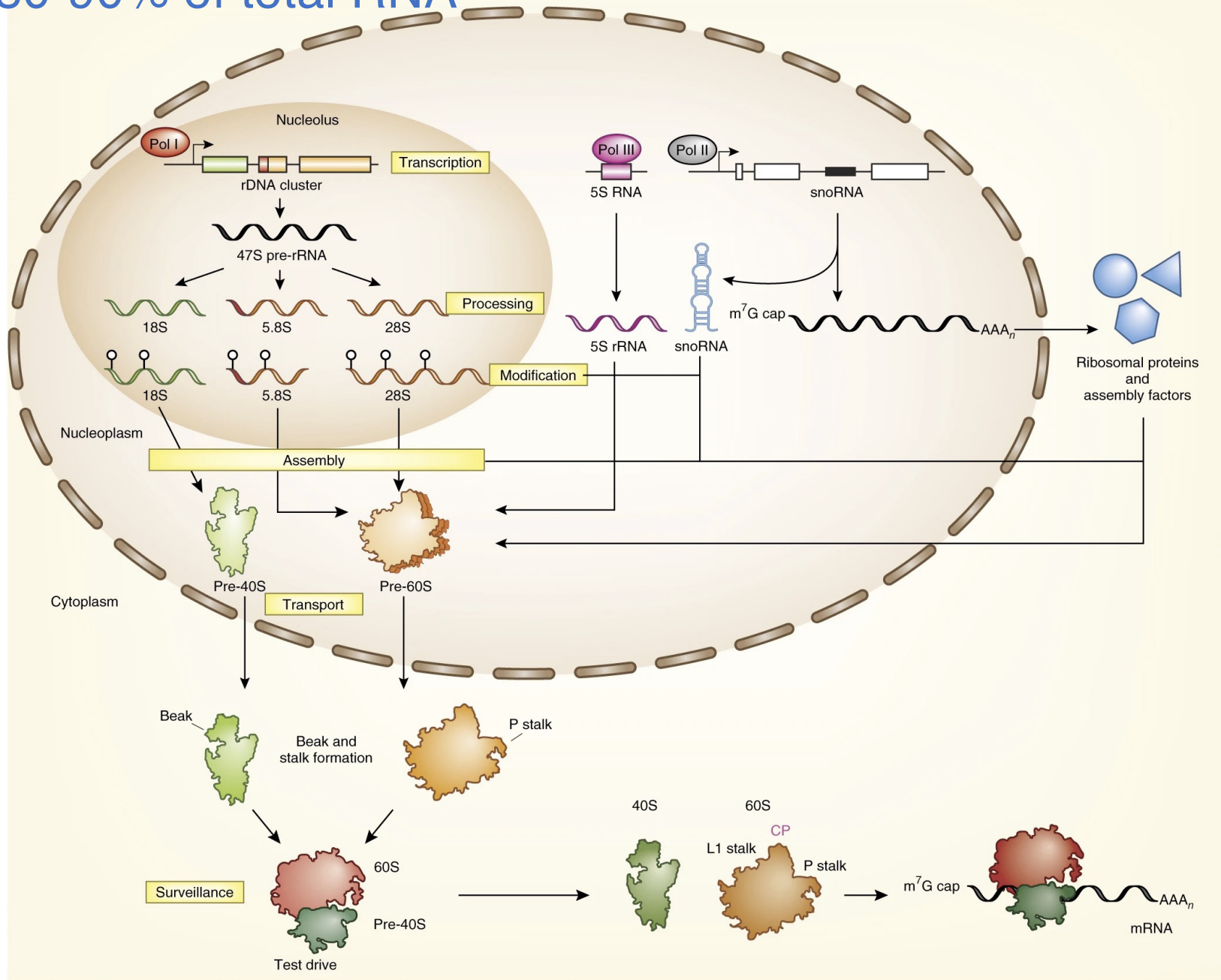


Translation

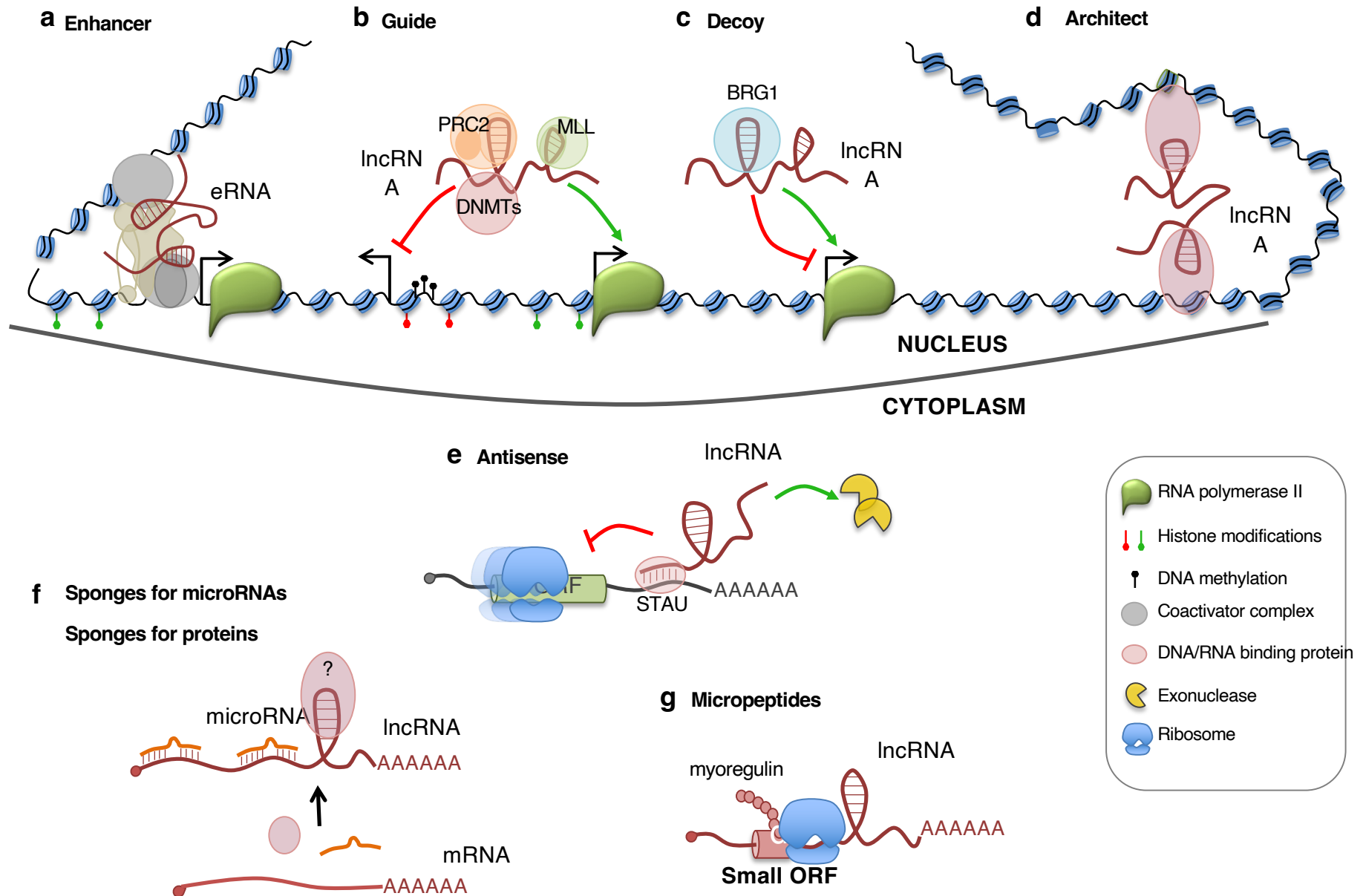


rRNA

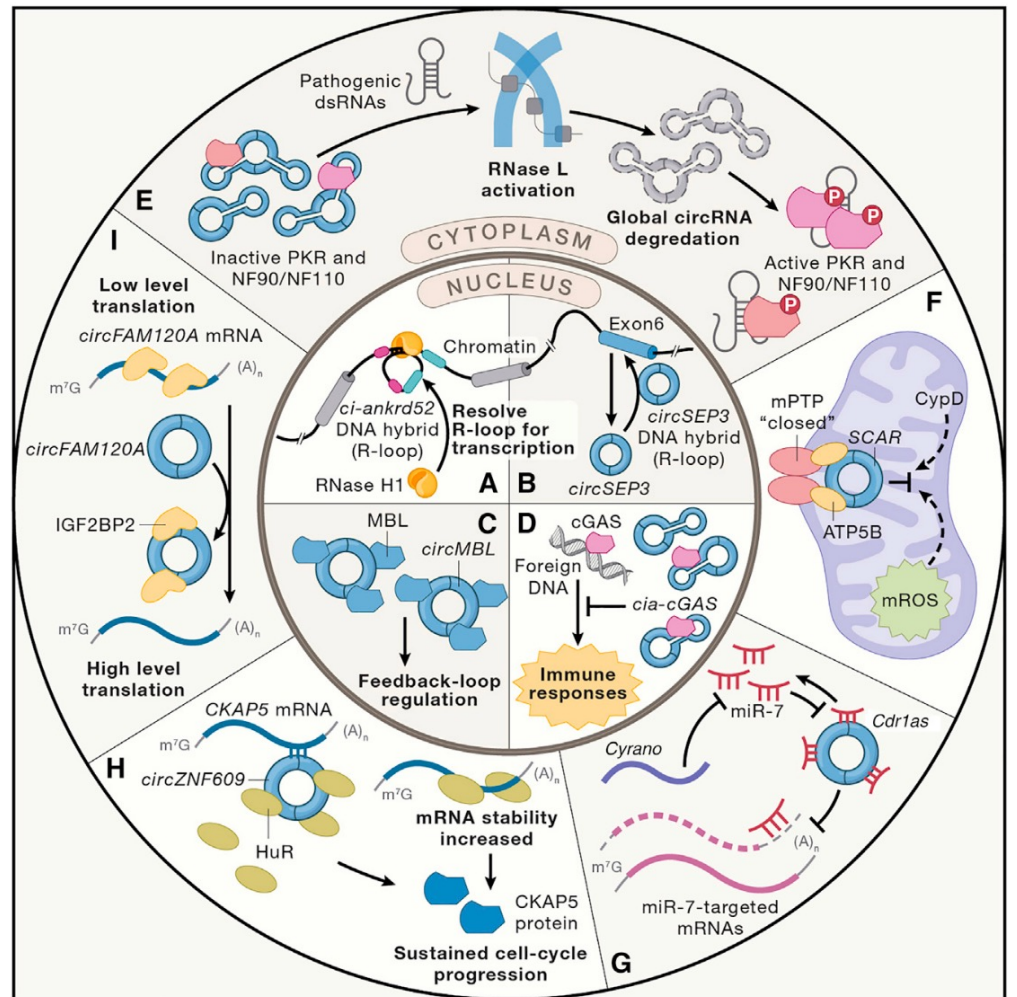
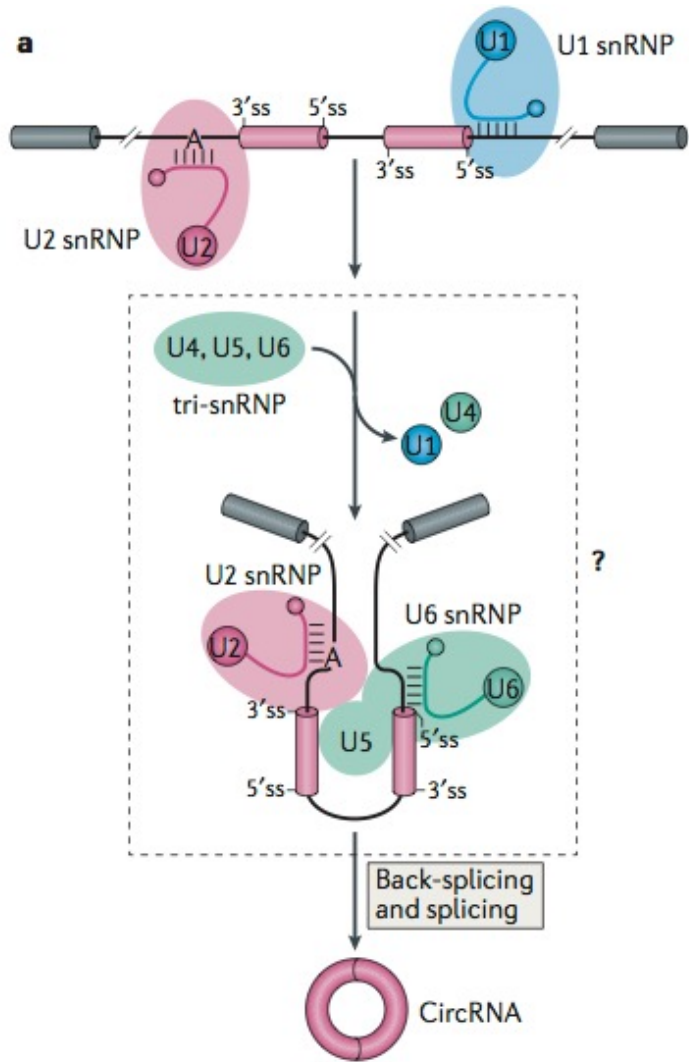
80-90% of total RNA



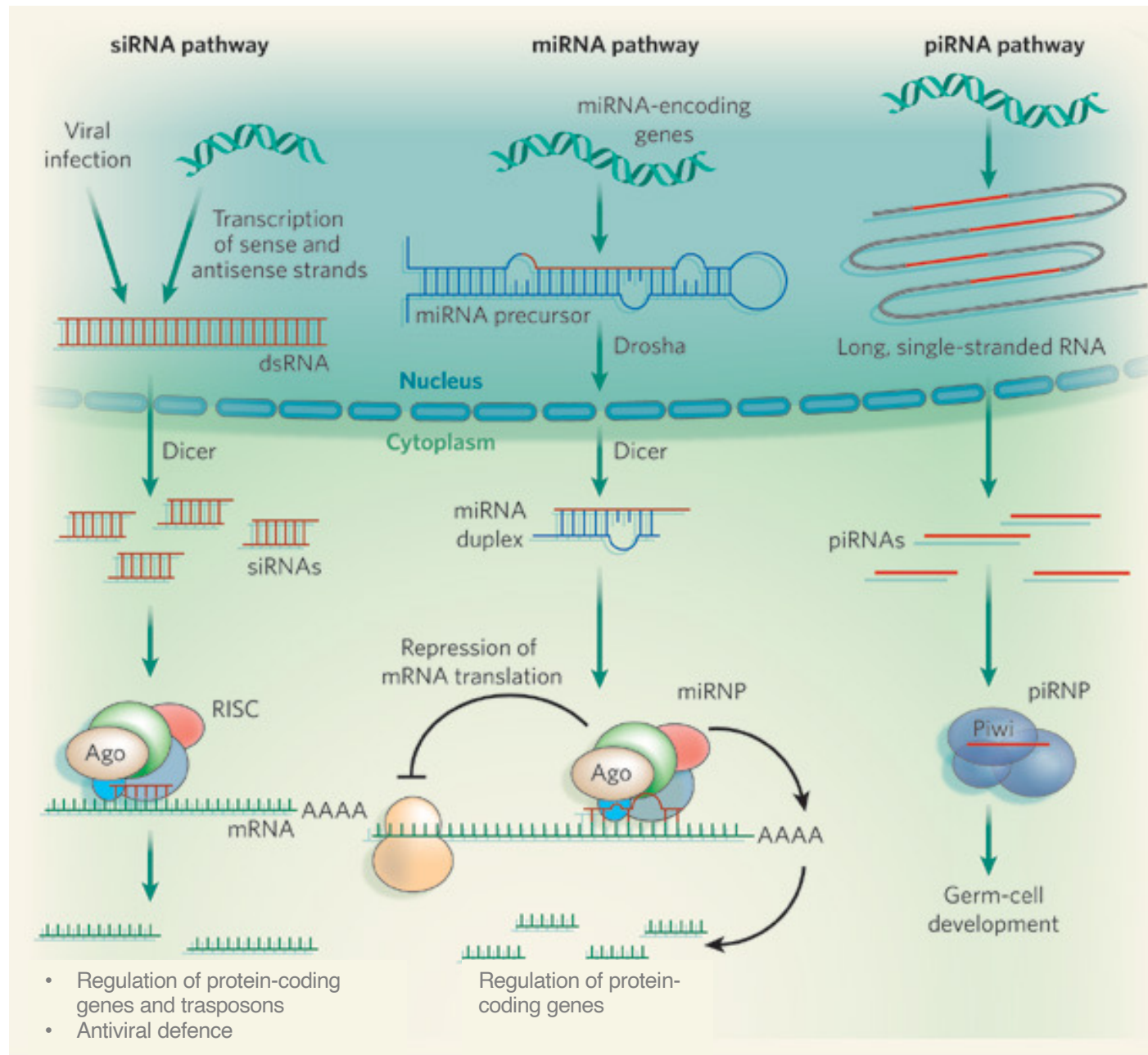
lncRNA



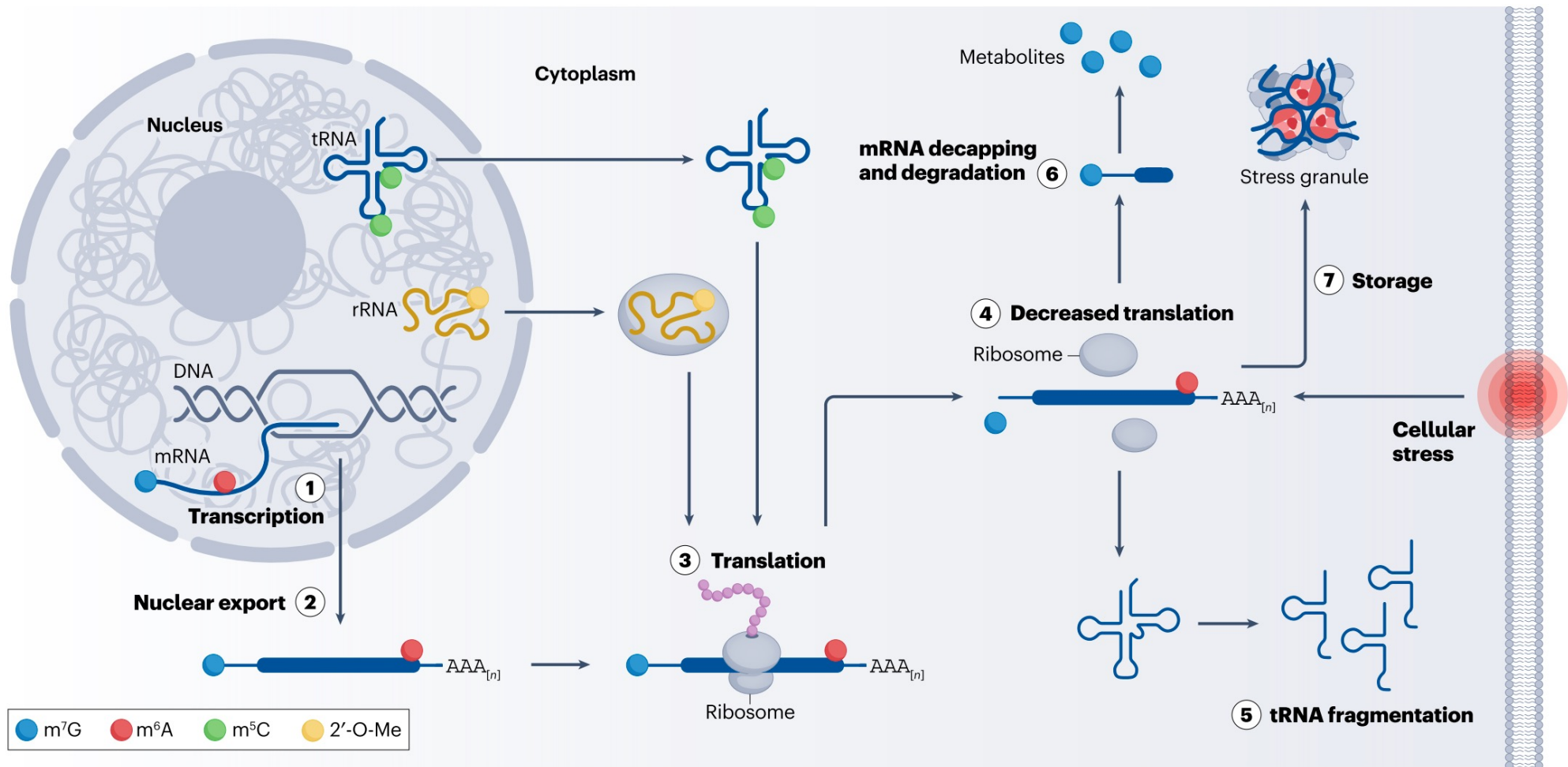
circRNA



siRNA, microRNA e piwiRNA



Chemical modifications of RNA can impact gene expression at different levels.



HOW TO STUDY GENE EXPRESSION AND GENE EXPRESSION REGULATION

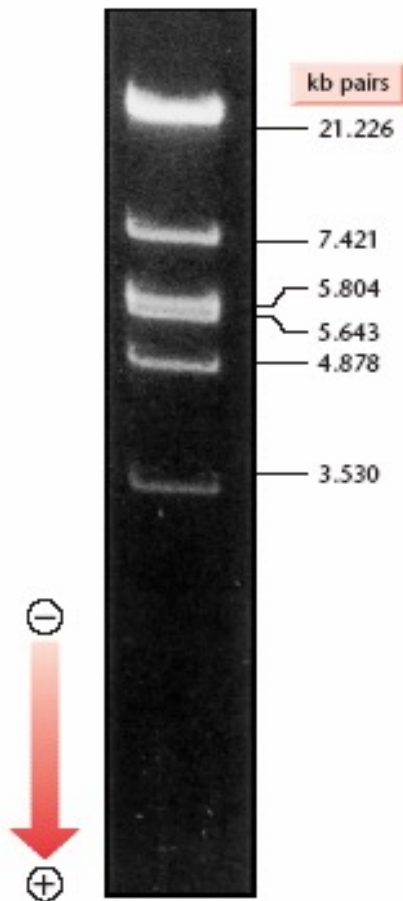
- RNA analysis
- Protein analysis
- Analysis of transcriptional and translational regulation
- Functional analyses (e.g., localization, interactions)
- Analysis of chemical modifications

Methodologies for the study of RNA

- **RNA Electrophoresis** – quality and processing
- **Northern blot** – processing and gene expression
- **Electrophoretic Mobility Shift Assay (EMSA)** –
RNA/protein interaction
- **Real-time PCR** – gene expression
- **Immunoprecipitation or pull-down** – RNA/protein
interaction

Electrophoresis of nucleic acids

Electrophoresis is a biochemical technique which allows to separate molecules with charge based on their different molecular weight. The electrophoretic separations is carried out through a gel.

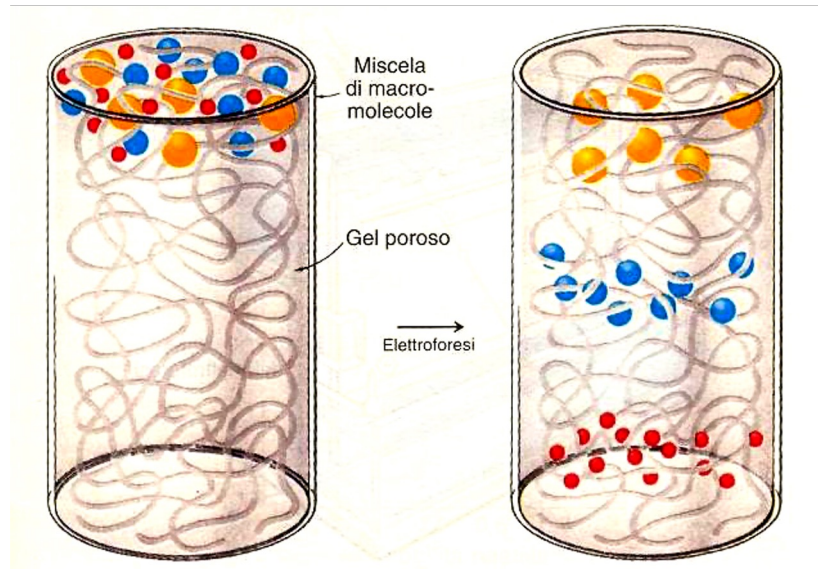


The gel can be made by:

✓ **agarose** (separation range : 0.5-20 kb)

✓ **polyacrylamide** (separation range 10-500 bp)

Electrophoresis of nucleic acids



Sieve effect



For linear fragments of DNA and/or RNA the migration distance is inversely proportional to the size of the molecule (that corresponds to its length in bases)

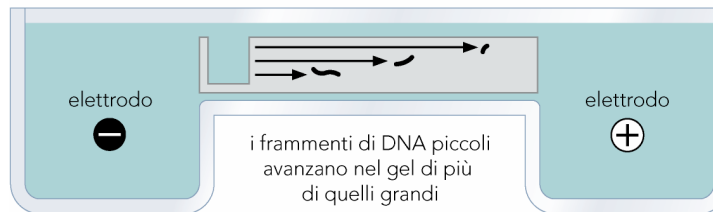
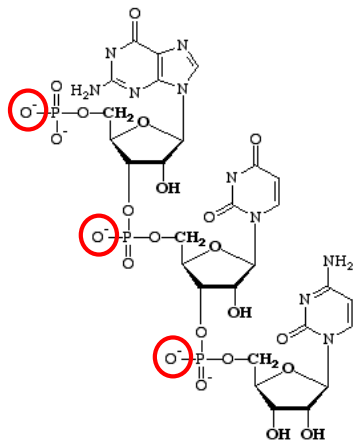
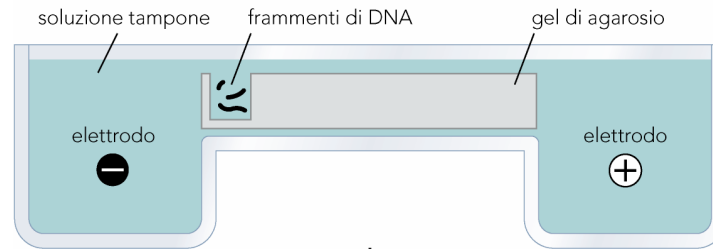
Electrophoresis of nucleic acids

Electrophoresis

✓ **horizontal**

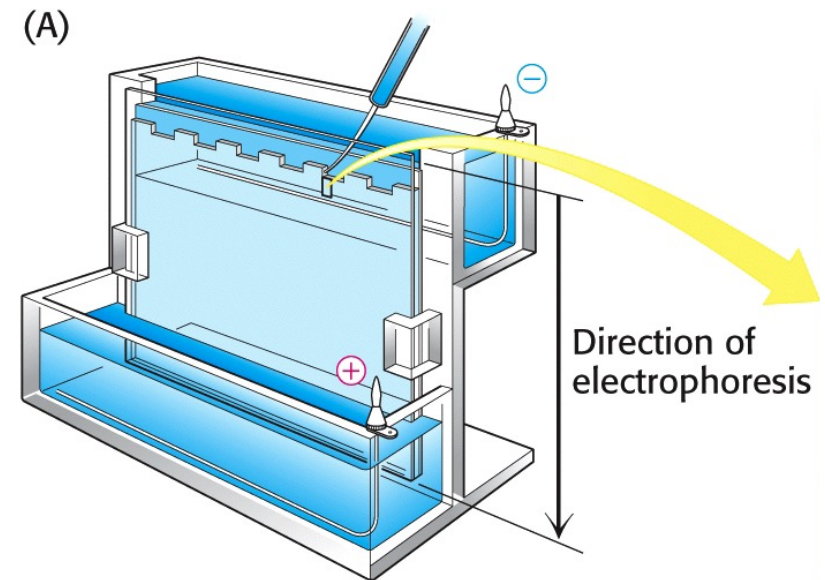
agarose

camera di elettroforesi



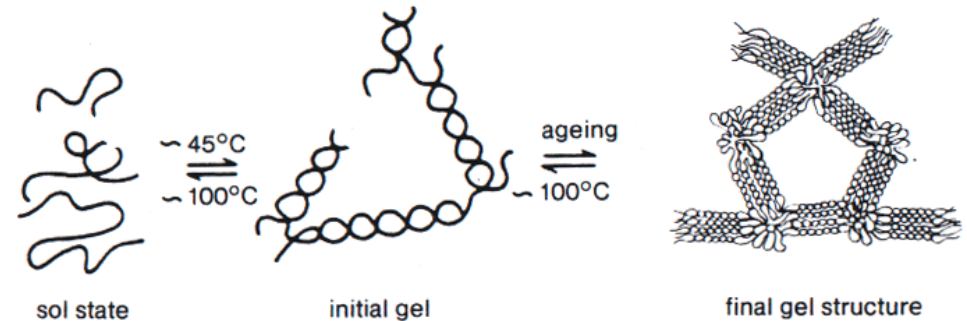
✓ **vertical**

acrylamide

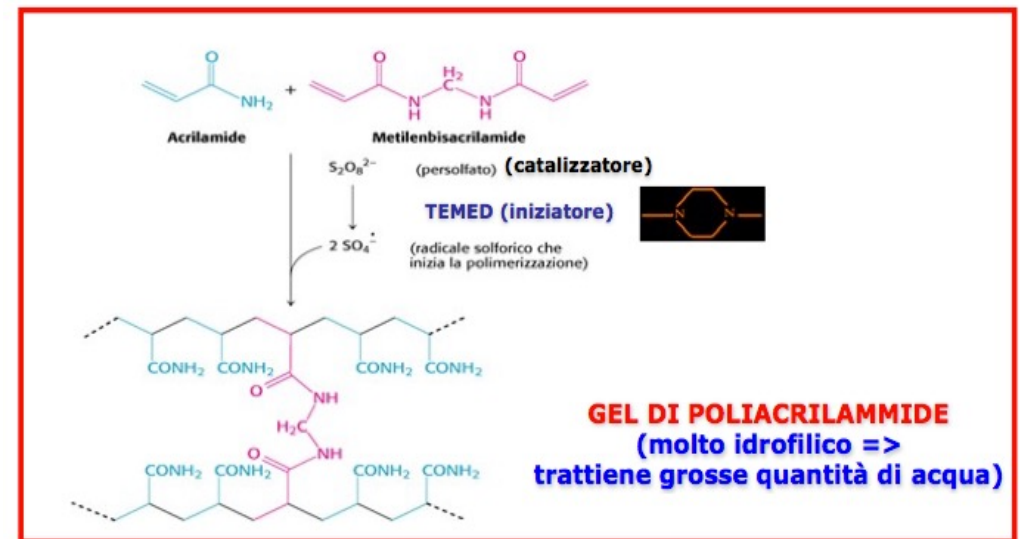


Electrophoresis of nucleic acids

The agarose is a polysaccharide composed of D-galactose and 3,6-anhydro-L-galactose residues



The polyacrylamide gel is formed by the copolymerization of acrylamide and of an agent which forms crosslinks (usually N, N'-methylene bisacrilamide) to form a three-dimensional lattice

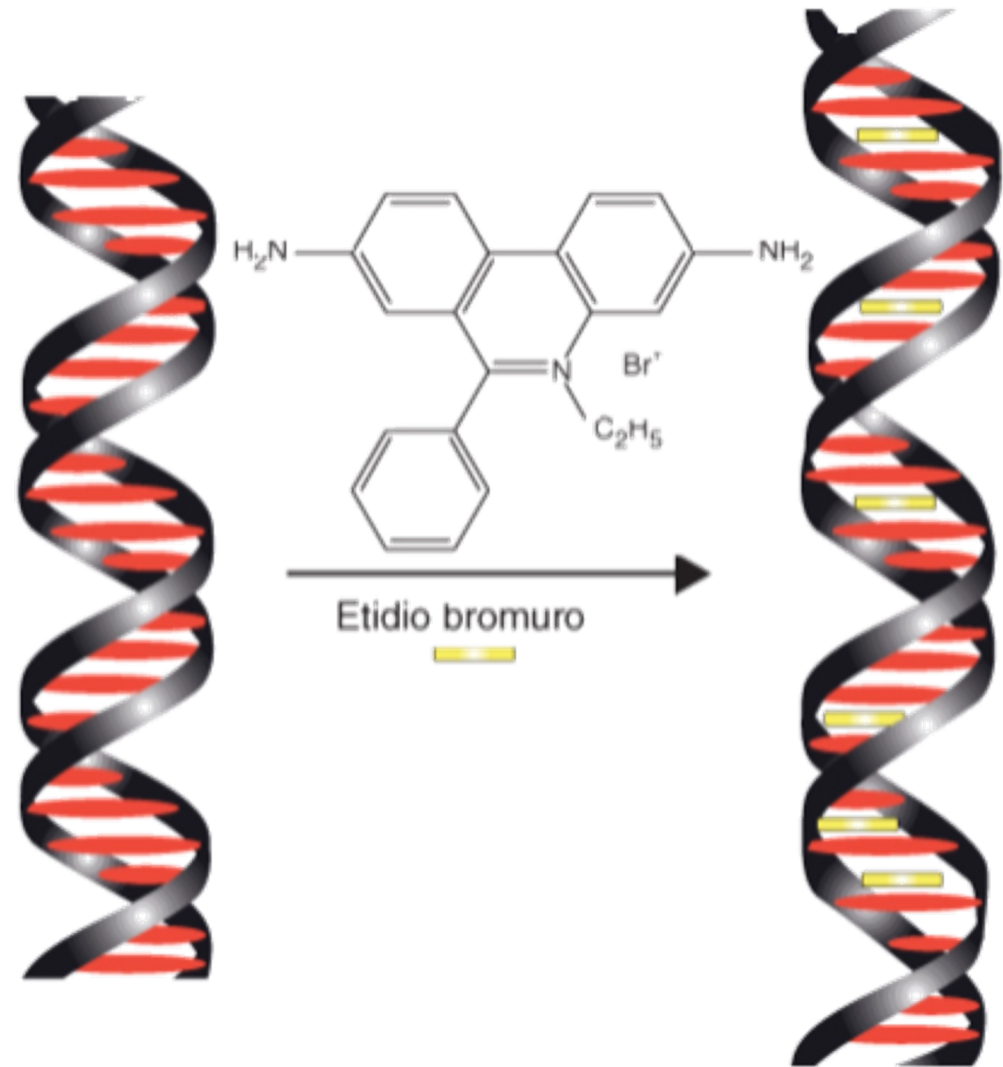


Electrophoresis of nucleic acids

ethidium bromide (EtBr):

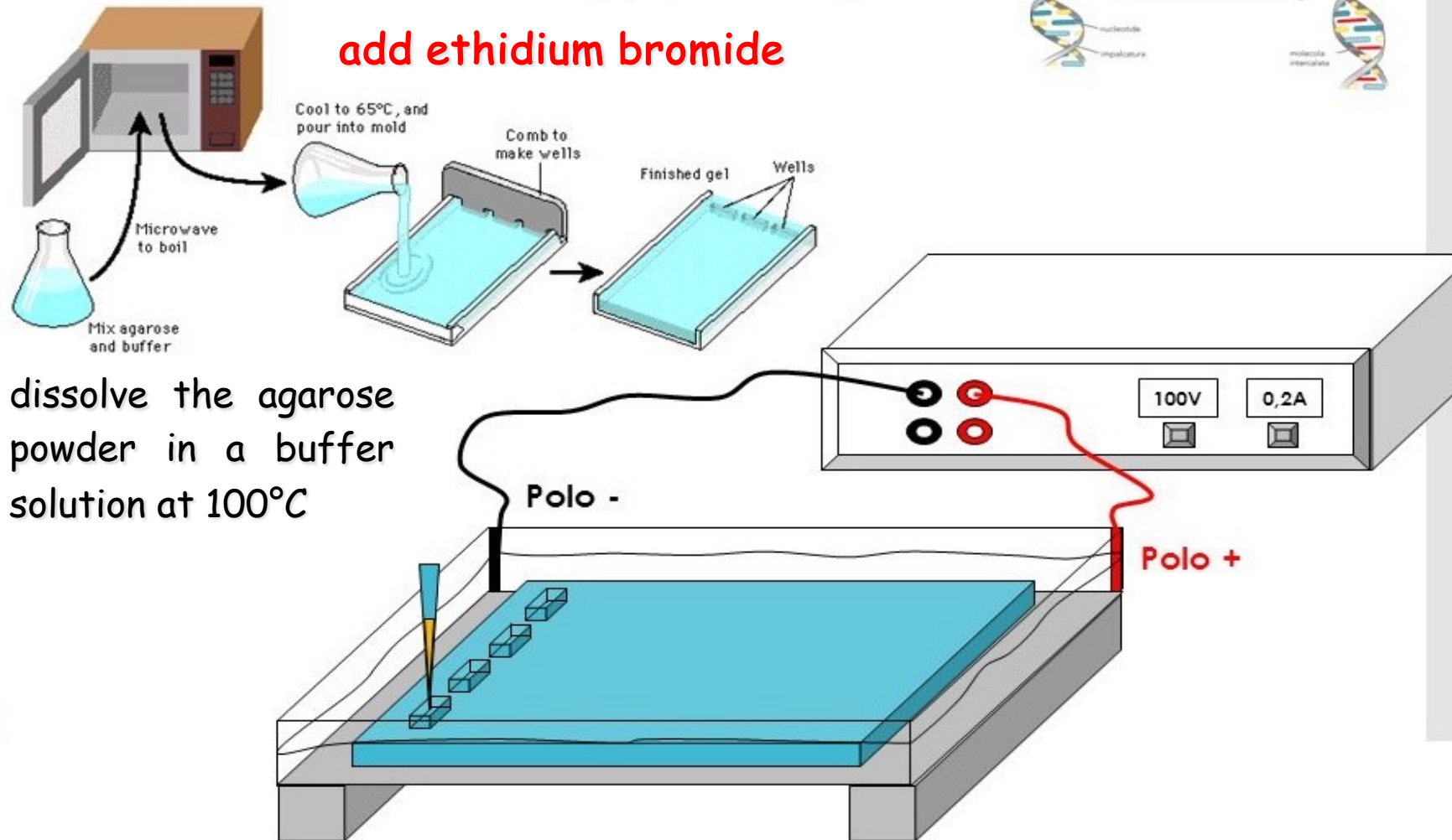
Intercalating agent, fluorescent dye that absorbs UV light at 300 nm resulting yellow-orange

Useful both for visualise and quantify the sample: the intensity of the fluorescence is, in fact, proportional to the amount of the sample.



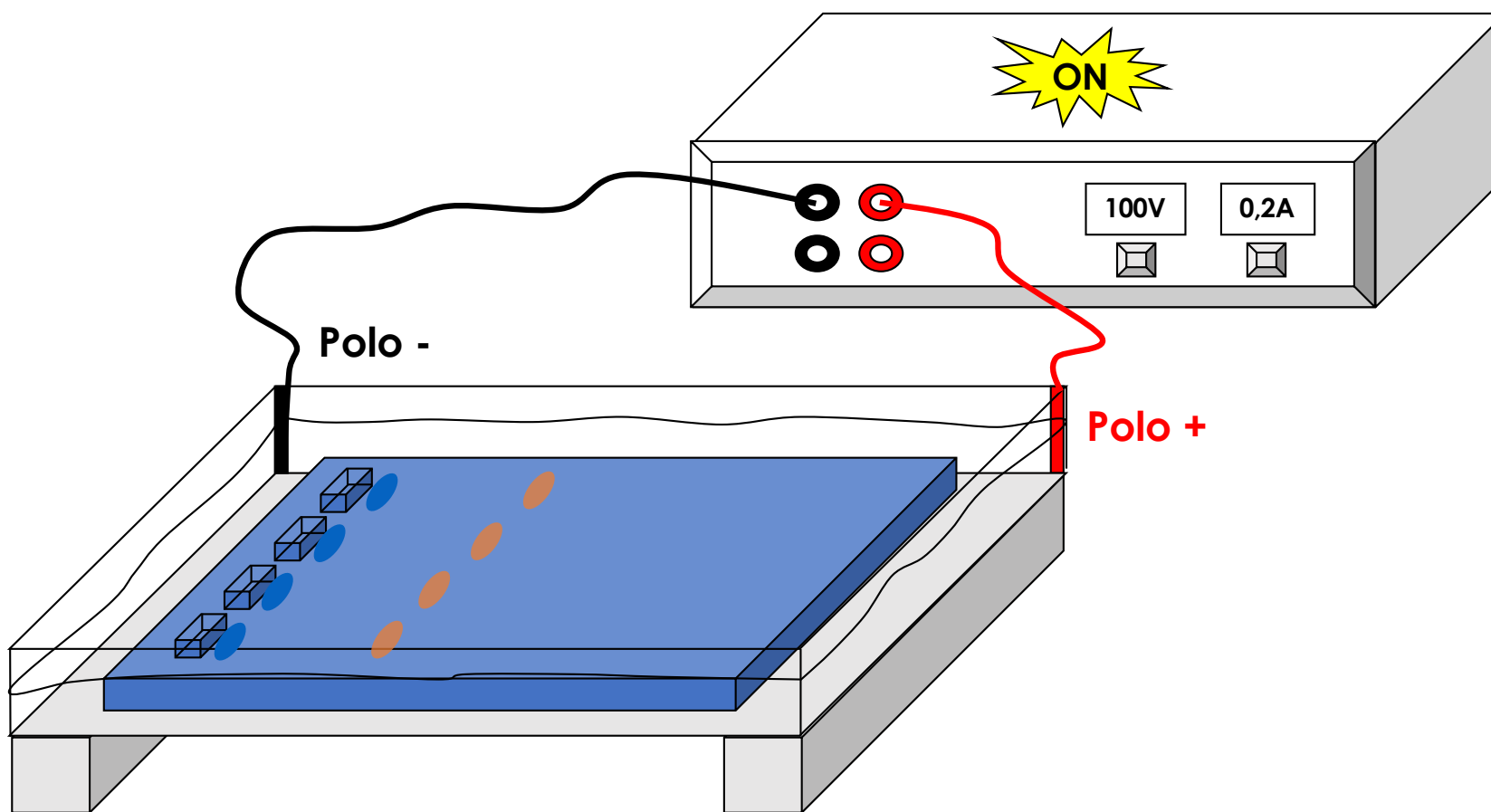
How to prepare agarose gel:

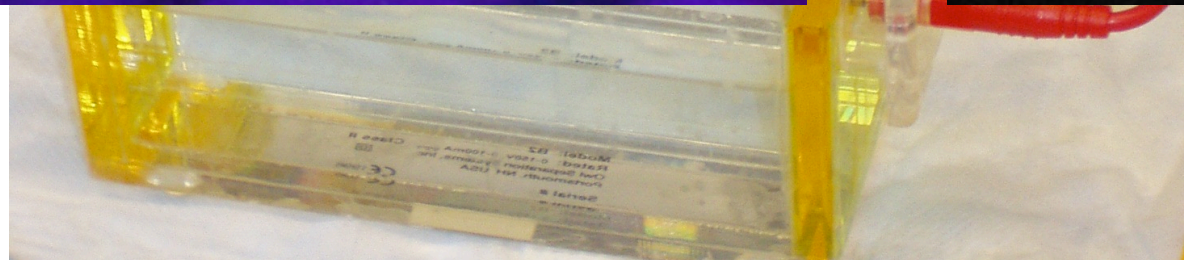
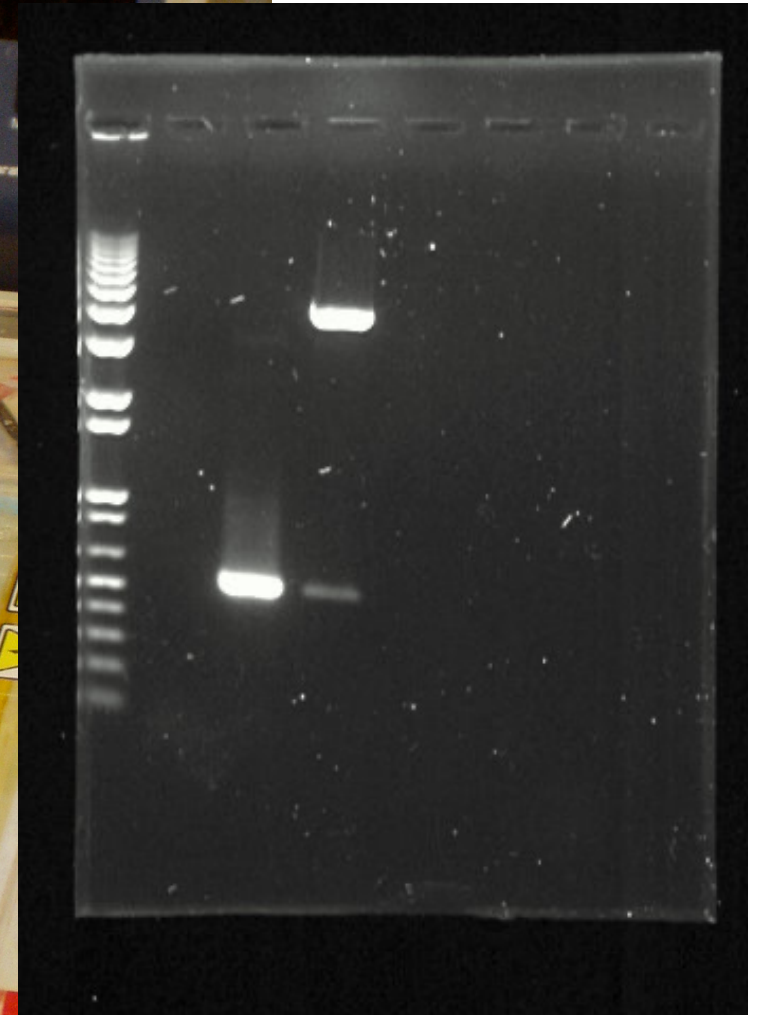
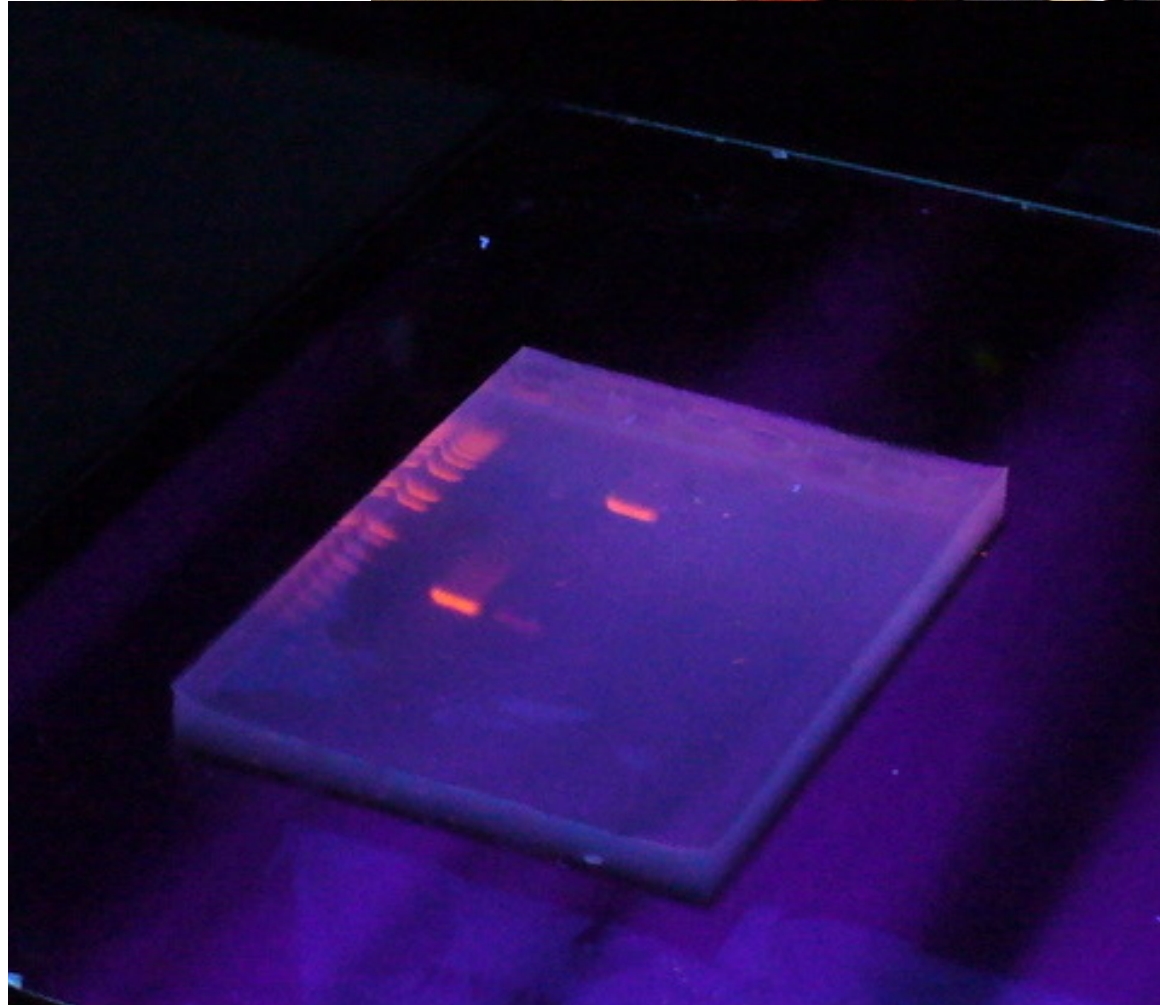
- 1% agarose gel in 1X TAE electroforesis buffer
- Run Time: 40 min.
- Voltage: 80 V
- Samples: 1/20 final Volume (5 μ l) + Loading buffer



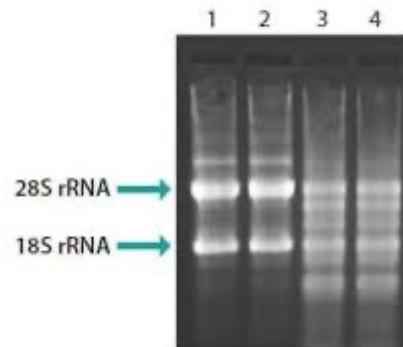
<https://www.youtube.com/watch?v=wXiiTW3pflM>

<https://www.youtube.com/watch?v=U2-5ukpKg> Q

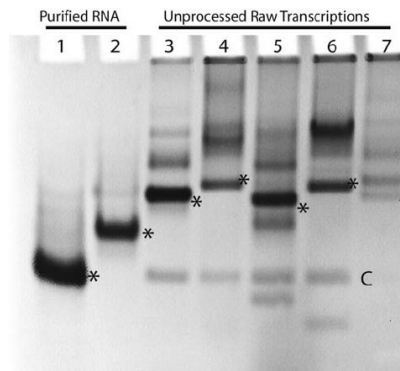




Quality check of RNA

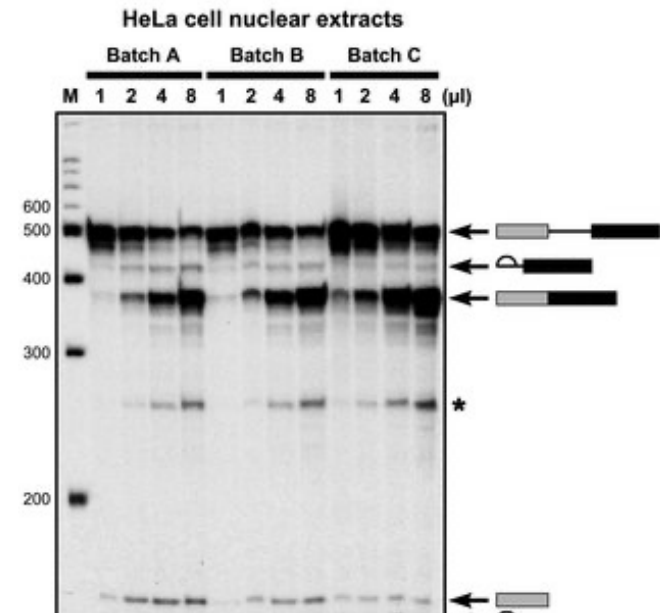


Total RNA
(agarose gel)



In vitro transcribed RNA
(acrylamide gel)

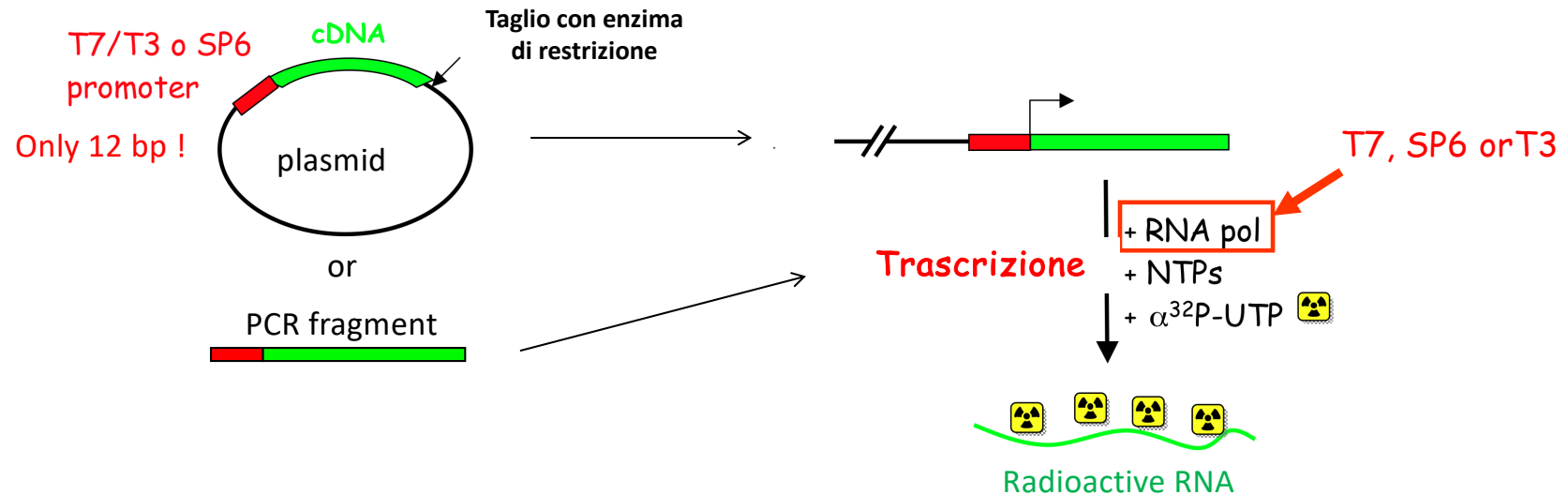
RNA processing



In vitro transcribed RNA
+
protein extract
(acrylamide gel)

How to generate RNA transcripts in vitro

in vitro trascription



Most of the enzymes used in molecular biology come from bacteriophages (e.g., T3, T7, SP6) because they have a low molecular weight and are monomeric

Transcription termination is achieved by making the RNA polymerase 'fall off' the template, which must therefore be linear.

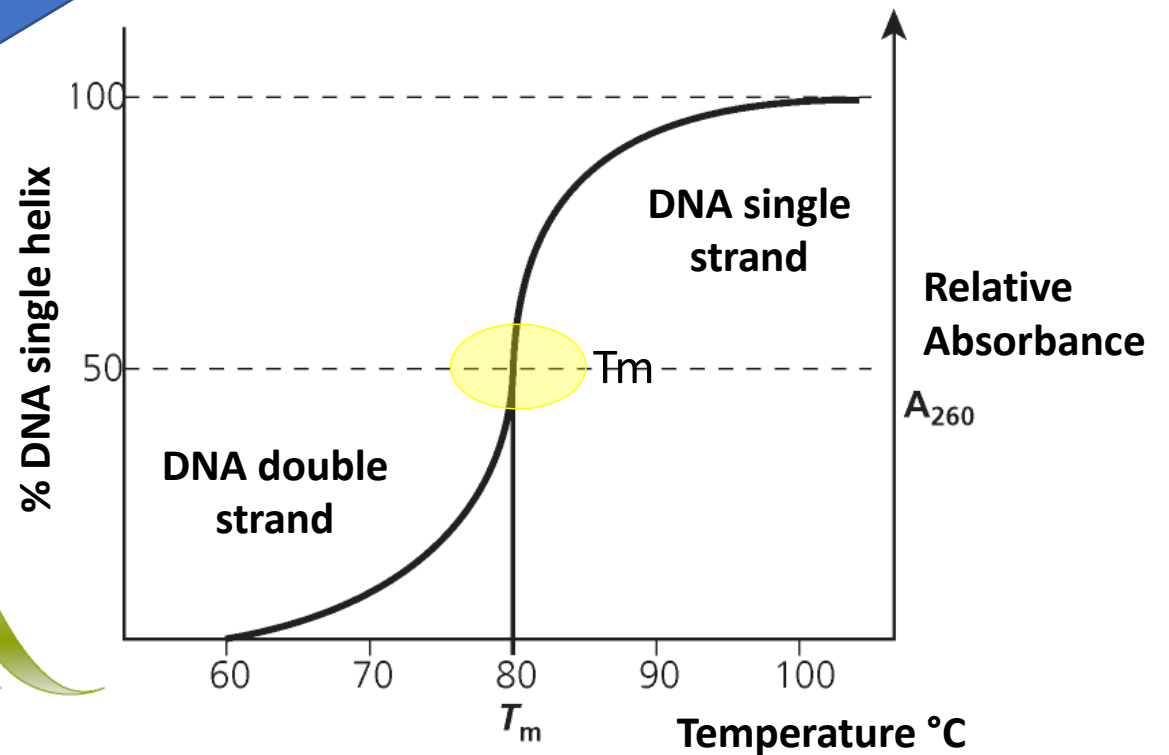
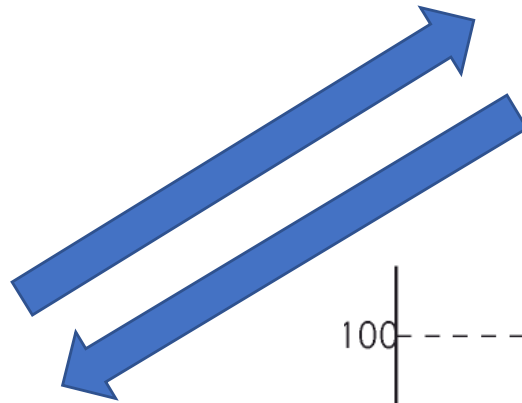
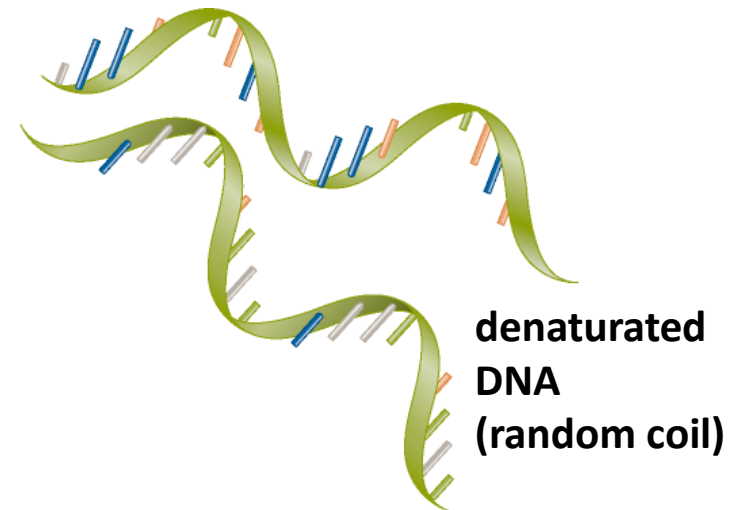
Chemical synthesis:

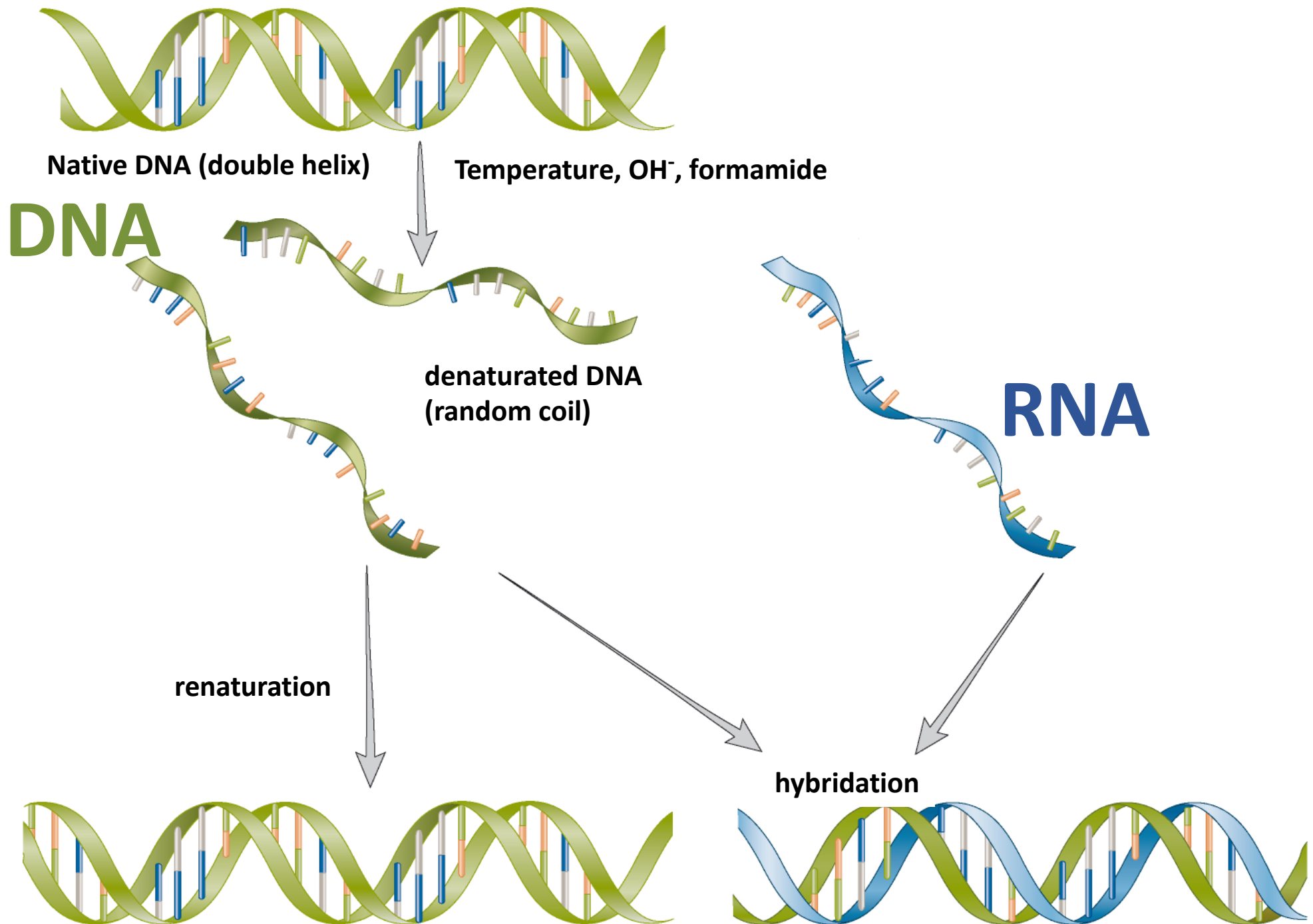
Oligonucleotide synthesizers can produce RNA molecules up to about 200 nt

Non radioactive tracers

- ✓ fluorochromes (direct labelling)
- ✓ digoxigenin or Biotin (undirect labelling)

**Preliminary: Nucleic acid can
be denaturated and
renaturated**

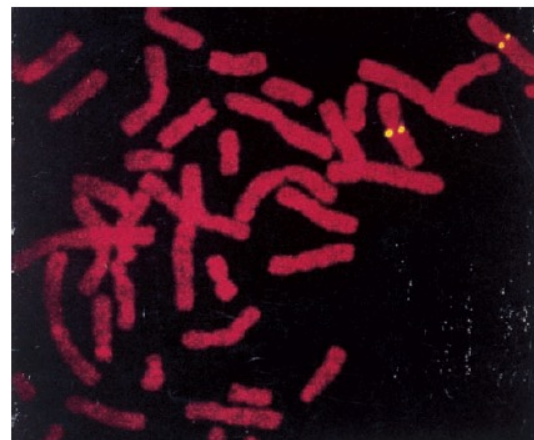
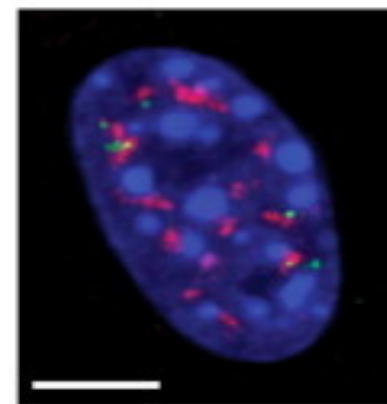
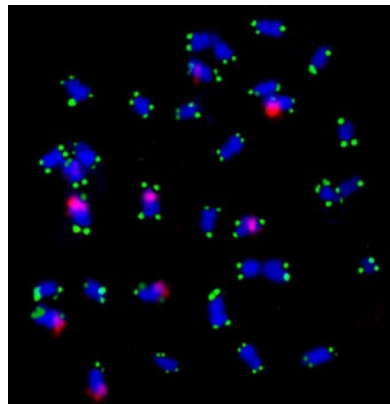
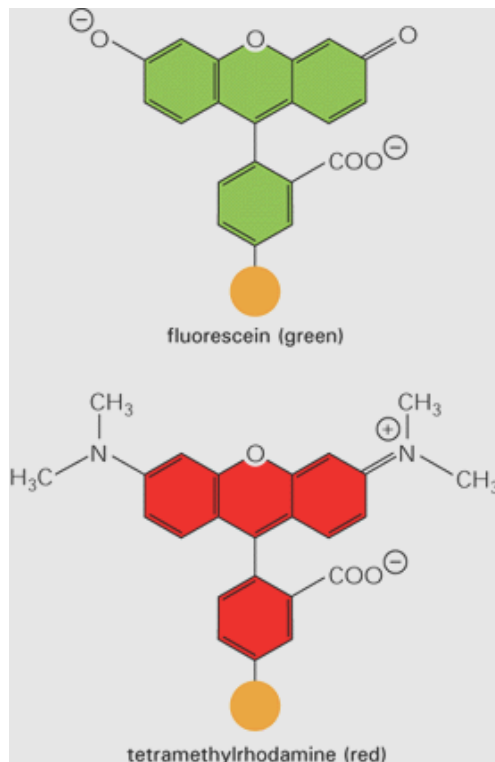




Non radioactive tracers

Fluorescence

It works with nucleotides linked to fluorochromes (Fluorescein, rhodamine) that can be detected a fluorescence microscope or other fluorescence detectors



Non radioactive tracers

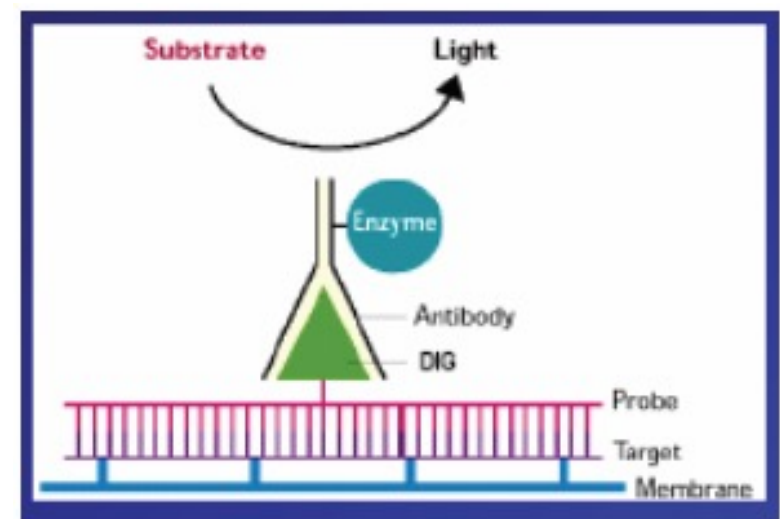
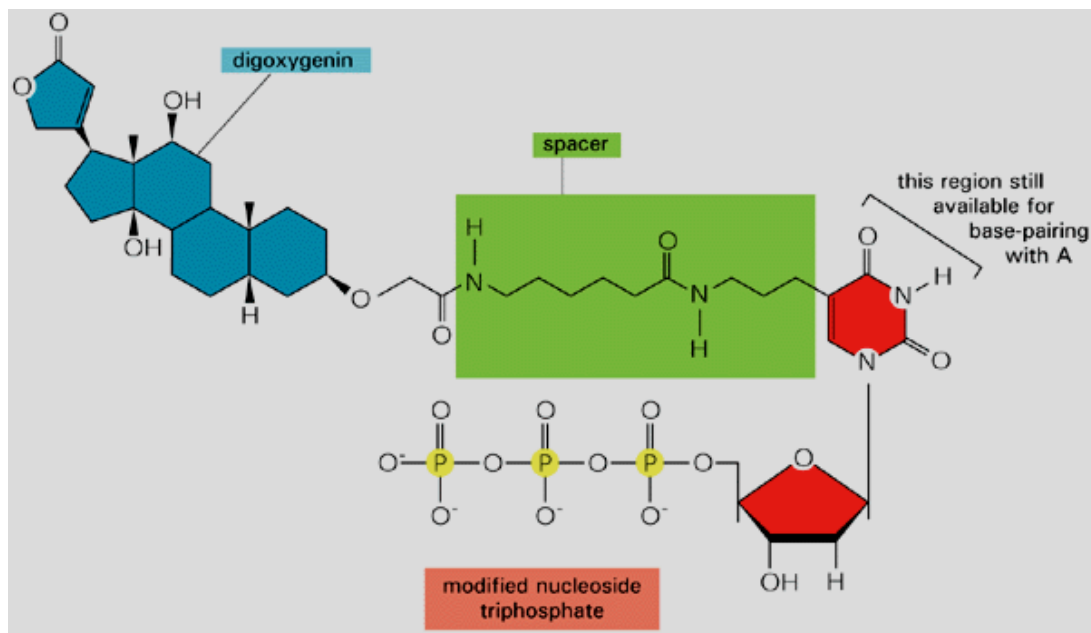
Chemiluminescence

Nucleotides digoxigenin marked:

Digoxigenin is a cardiotonic steroid isolated from the *Digitalis purpurea*.

The probe is bound and recognized by an immuno-enzymatic system that uses an antibody against the digoxigenin (anti-DIG). The antibody is ligated to an alkaline phosphatase. The presence of such a complex is revealed by adding a chromogenic or luminogenic substrate that only “emits” when is processed by the alkaline phosphatase.

The high specificity of the antibody ensures the high specificity of the detection.

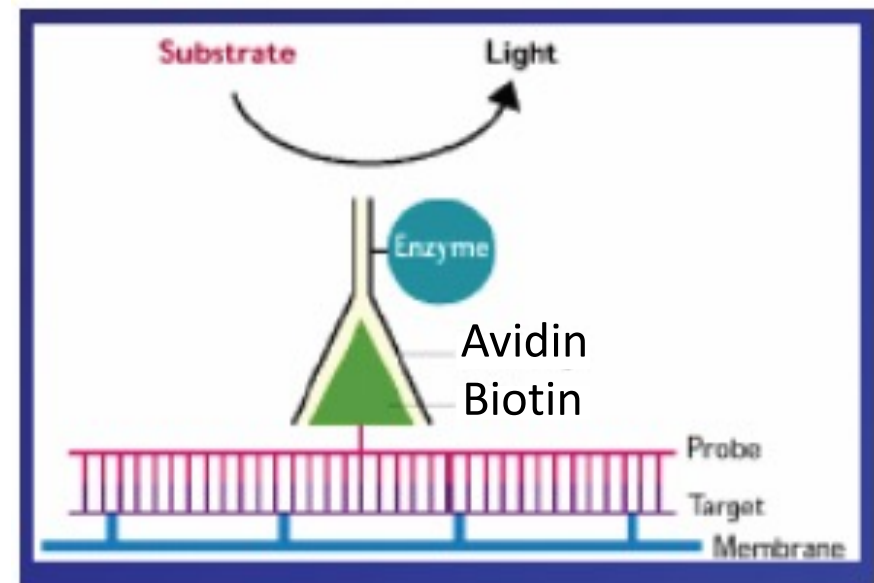
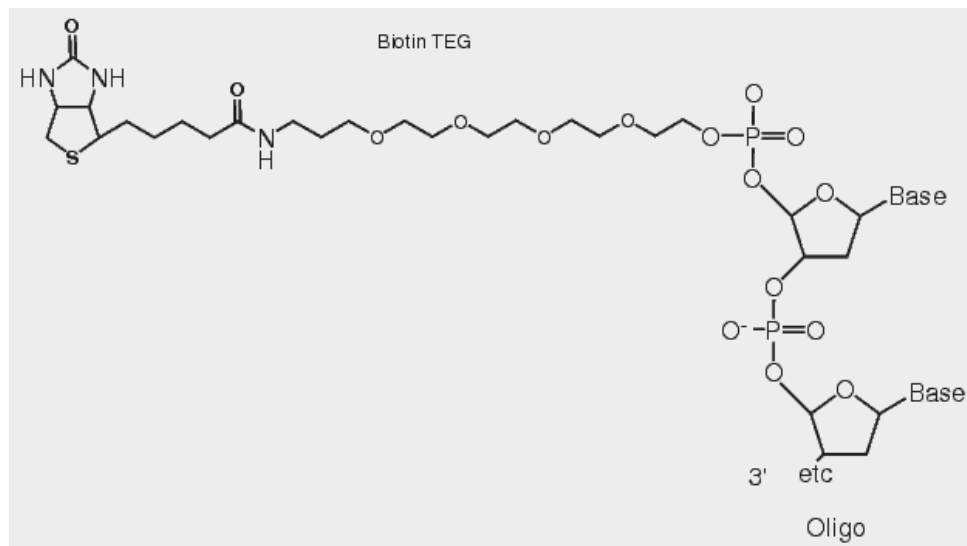


Non radioactive tracers

Biotinilated nucleotides

Biotinilated nucleitides can be incorporated in the DNA both terminally and internally.

The Biotin is specifically bound by avidin or streptavidin. Therefore it is possible to use avidin or spreptavidin marked with fluorophores or linked to enzymes (phosphatases or peroxidases). Providing chromogenic or fluorogenic substrates for these enzymes you allow the production of a measurable signal.

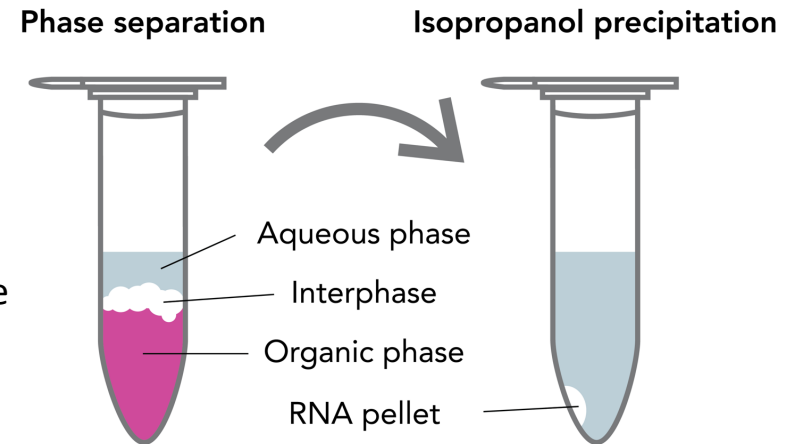


RNA extraction

The most commonly used reagents for RNA extraction from biological samples are based on guanidinium isothiocyanate and phenol (e.g., TRIzol). The low pH of the reagent allows for the separation of RNA from DNA and proteins. The guanidinium salt acts as a chaotropic agent to denature proteins, and phenol is an organic compound also used to extract nucleic acids and proteins.

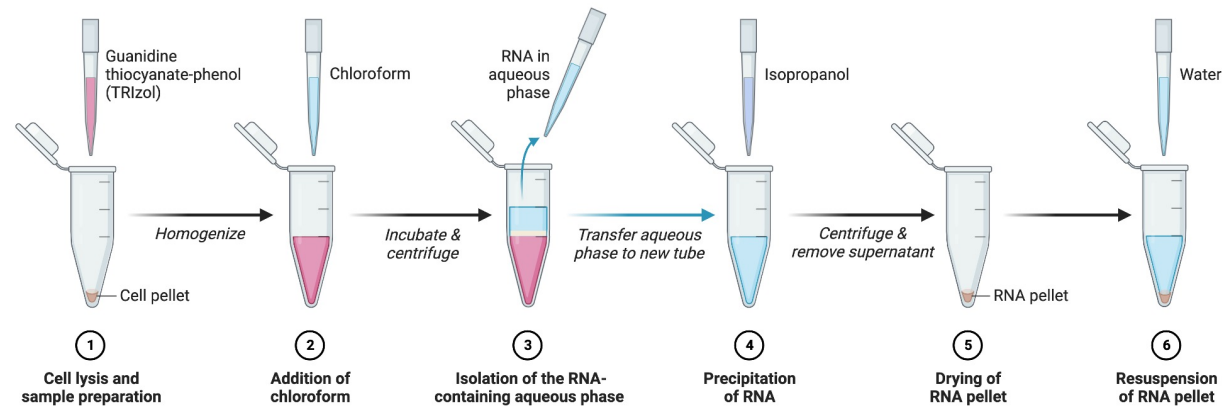
After sample solubilization and homogenization, RNA, DNA, and proteins are differentially extracted by adding a phase separation reagent (chloroform). The solution separates RNA from DNA and proteins into different layers. A clear, upper aqueous phase primarily contains RNA, while the intermediate interphase and the lower organic phase contain DNA, proteins, and lipids.

Subsequently, the RNA in the upper aqueous phase is collected via alcoholic precipitation. The pellet is washed in an ethanol solution, air-dried, and resuspended in deionized water.



RNA extraction

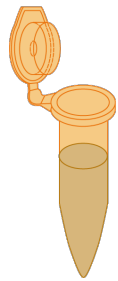
Classic method (ethanol precipitation)



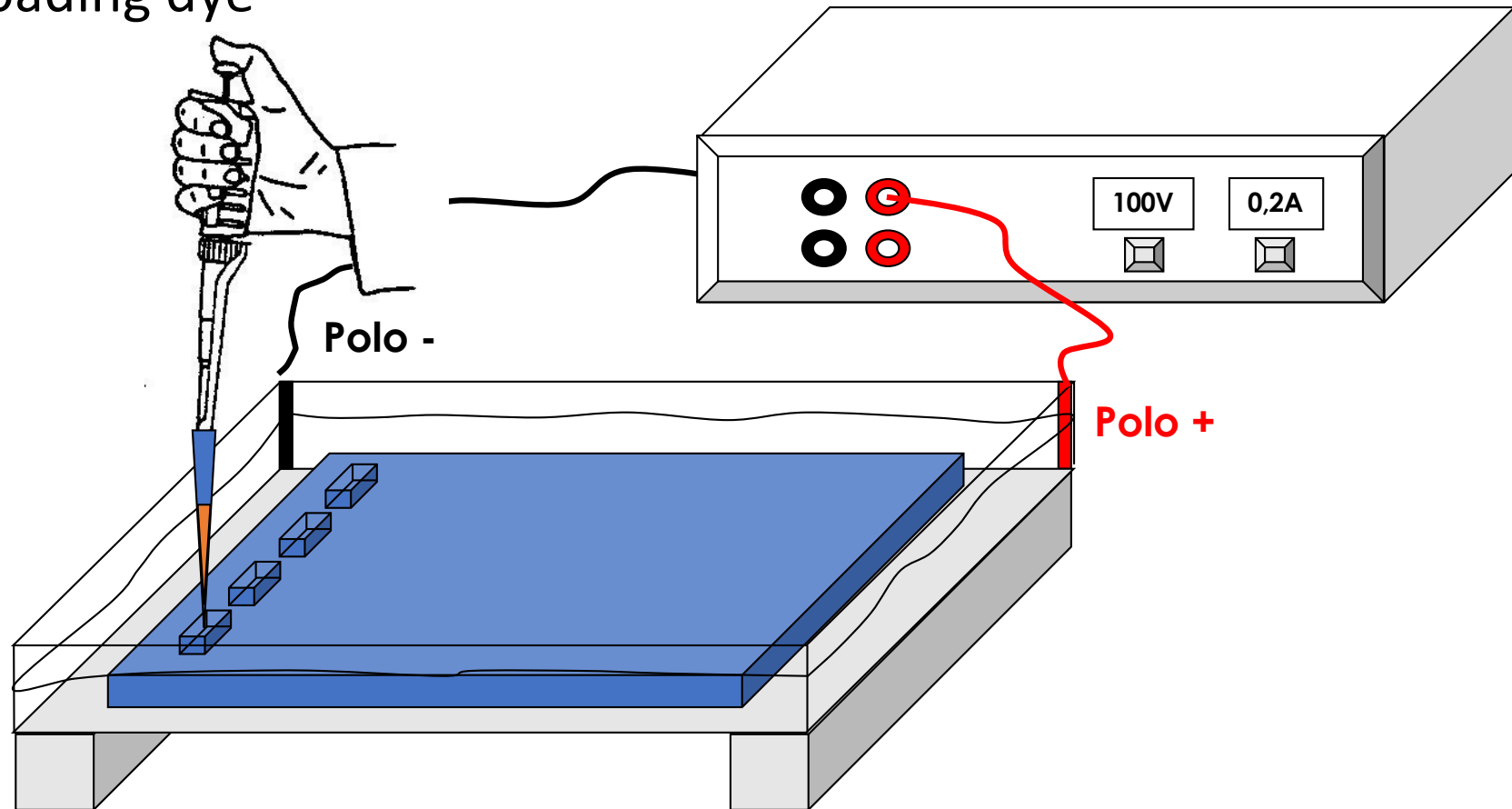
Rapid method (with column)





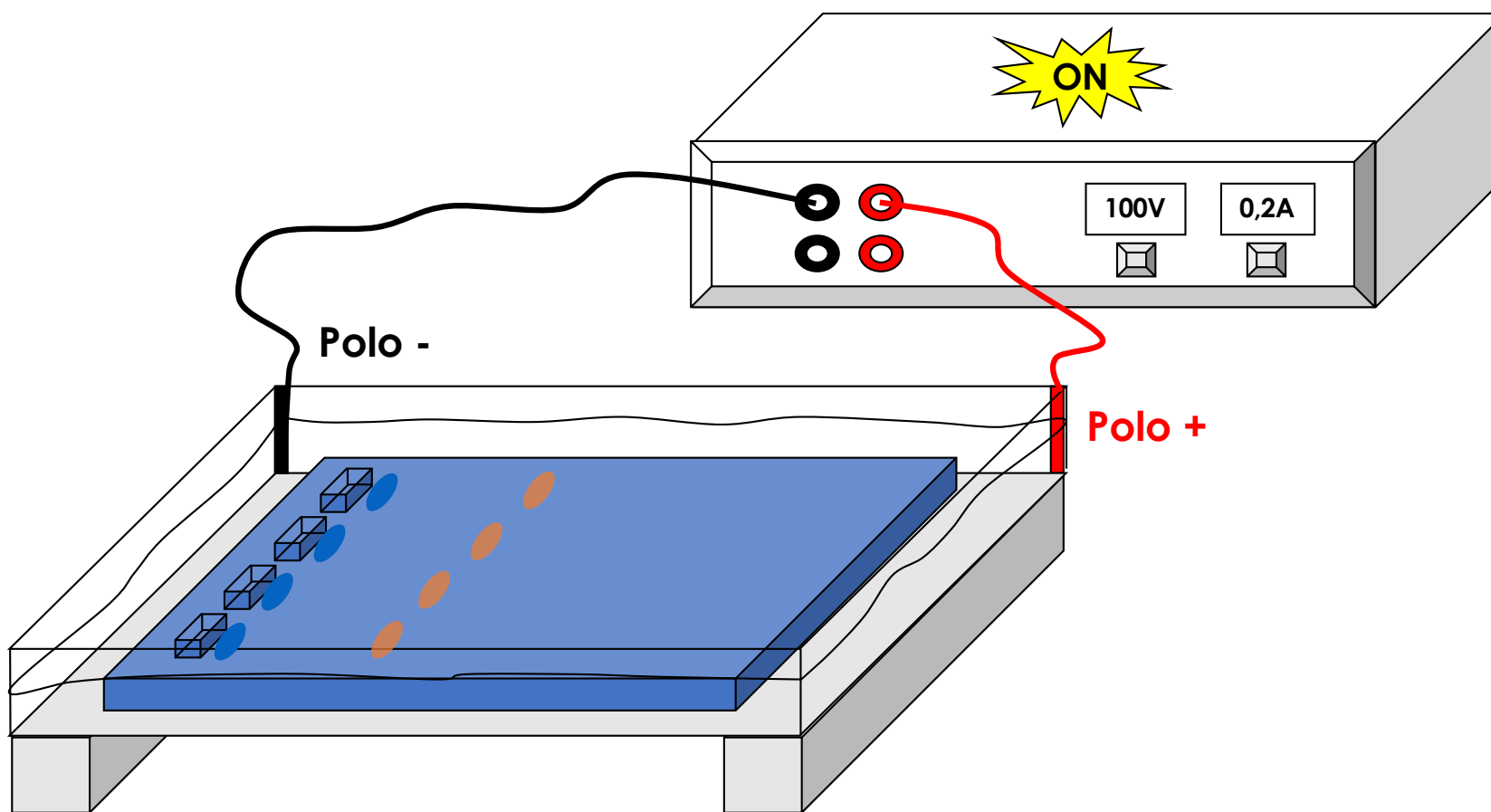


+ Loading dye

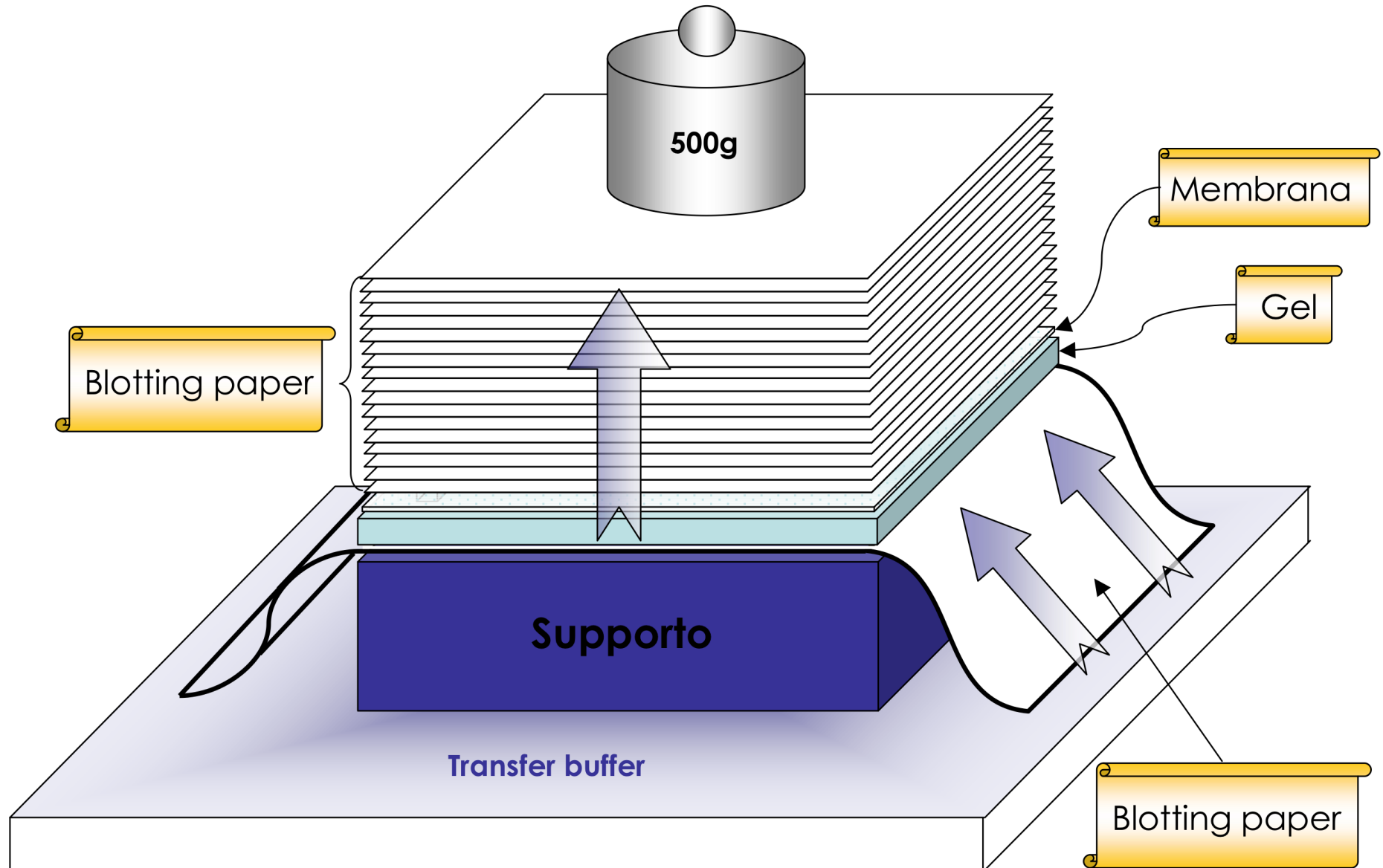


Loading dye helps the sample loading in the well

It contains **glycerol**, **bromophenol Blu** and **xilencianol Blu** that migrate in the gel at different speeds.

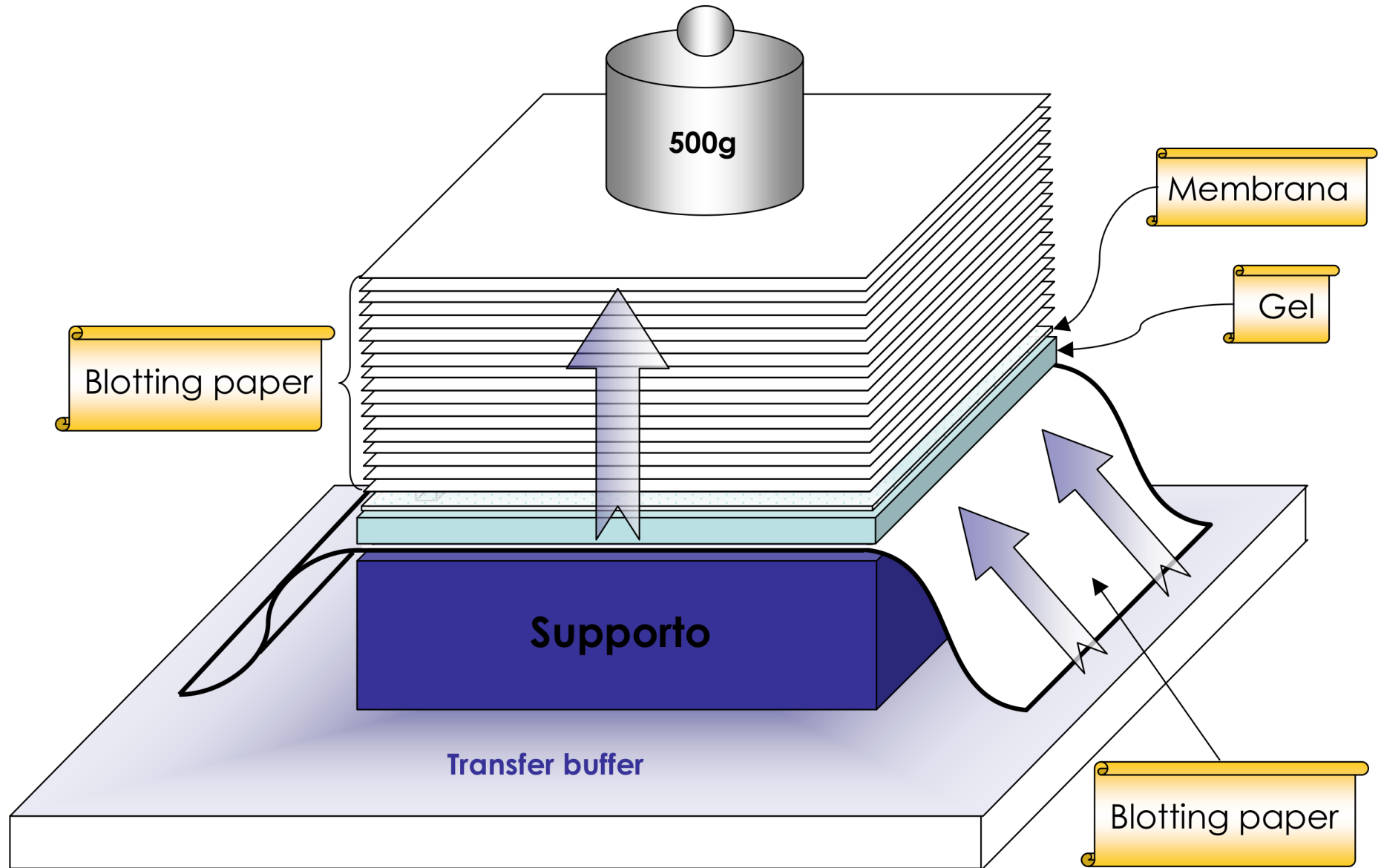


BLOT: transfer the nuclei acids from the gel to the membrane by diffusion (for agarose gel)



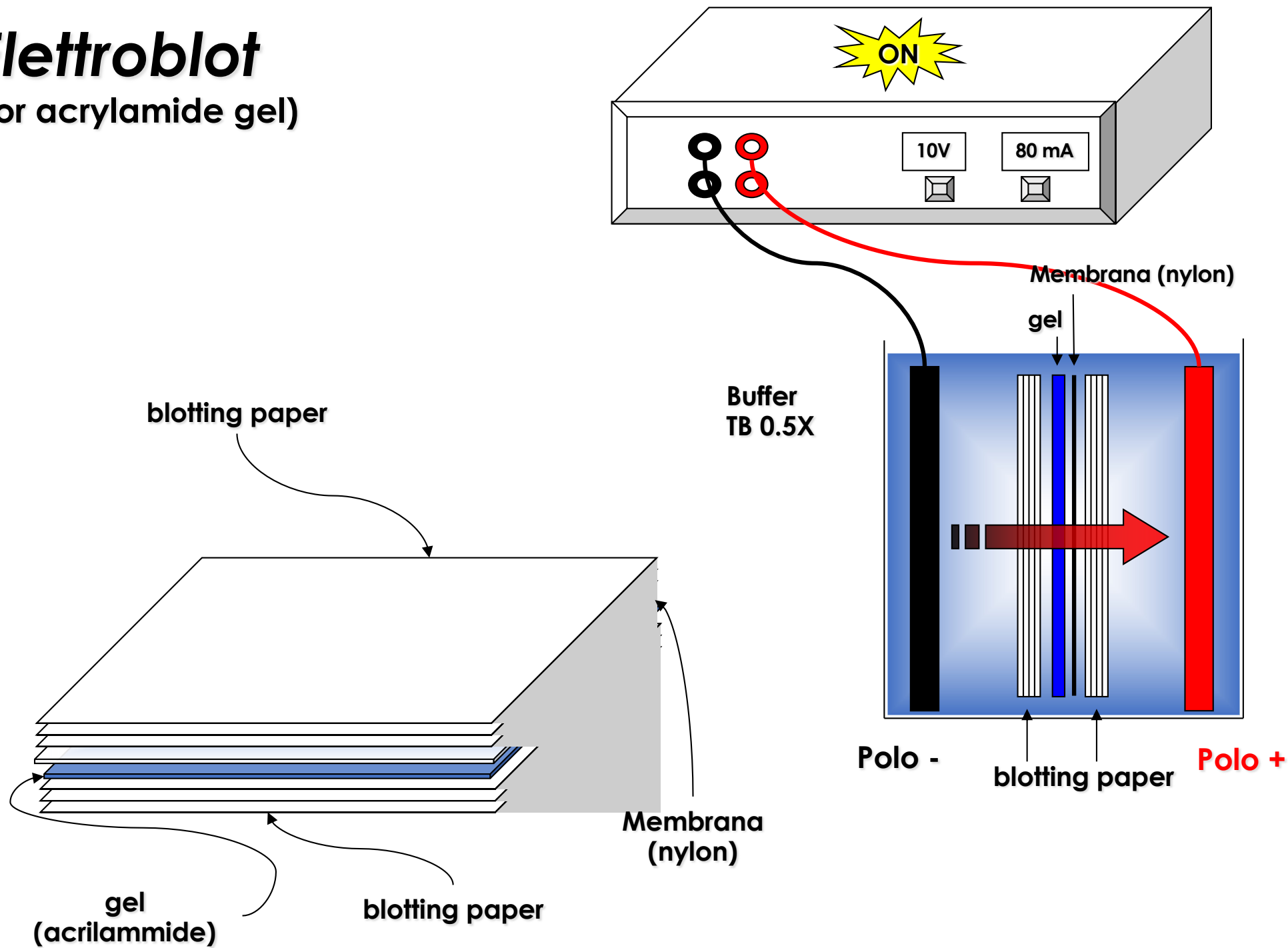
The nucleic acids will be fixed to the membrane by:

- UV irradiation for nylon membrane
- boiling for 2hrs at 80°C for cellulose nitrate membrane

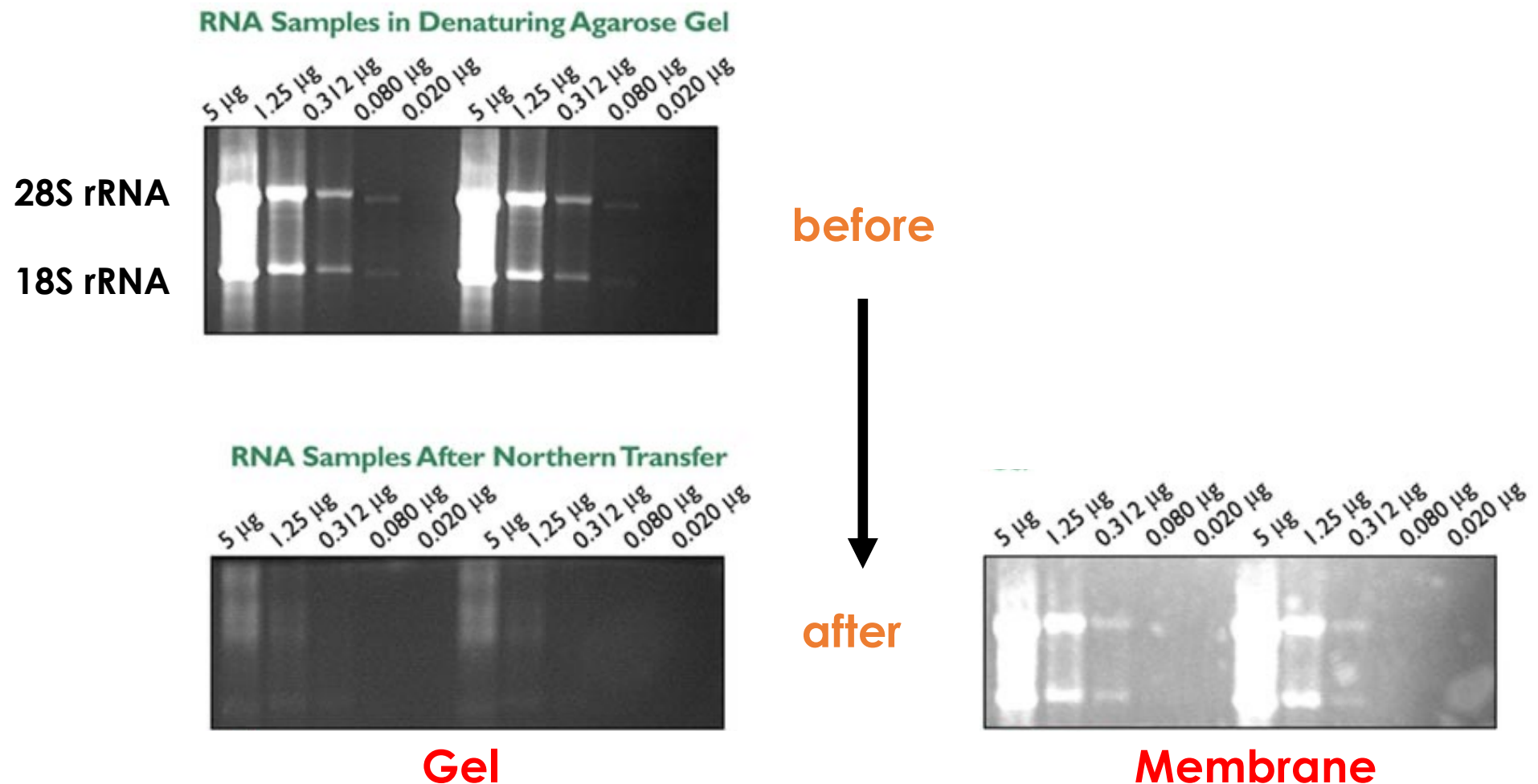


Elettroblot

(for acrylamide gel)

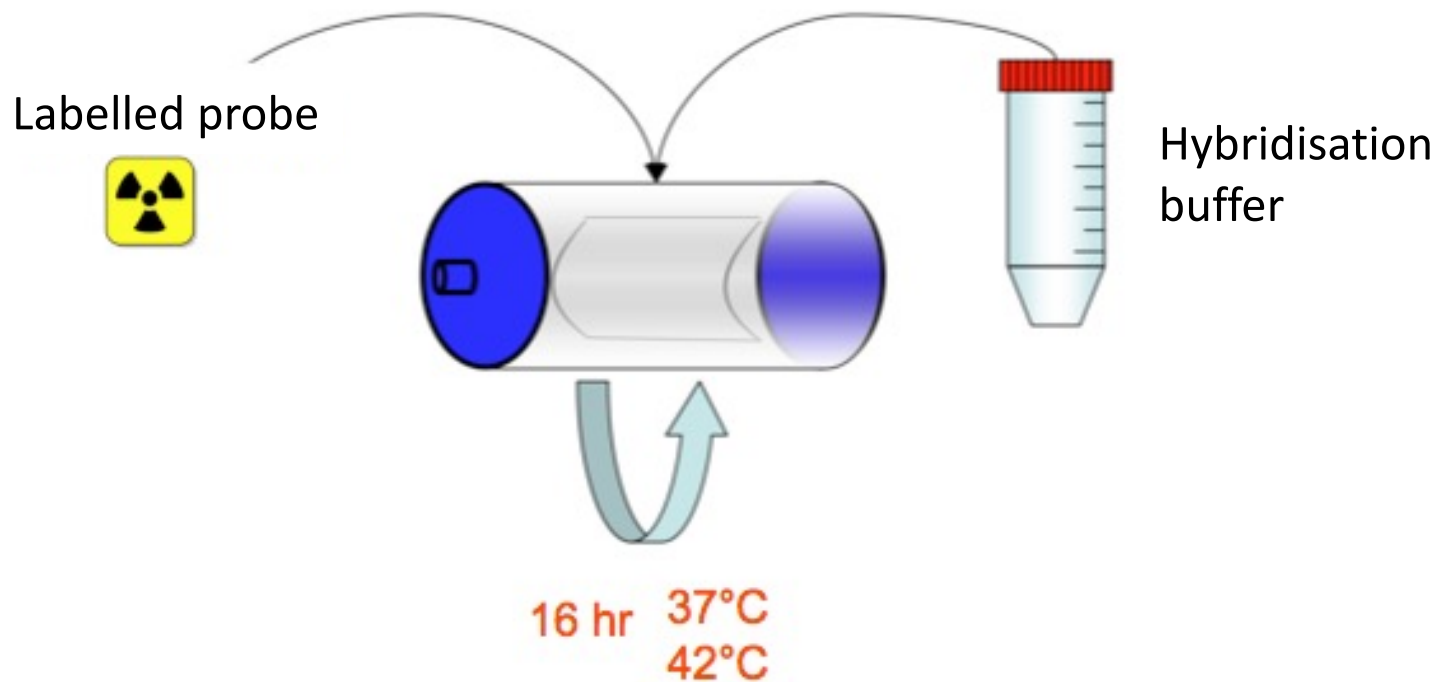


Ethidium bromide staining helps to check the proper RNA transfer on the membrane



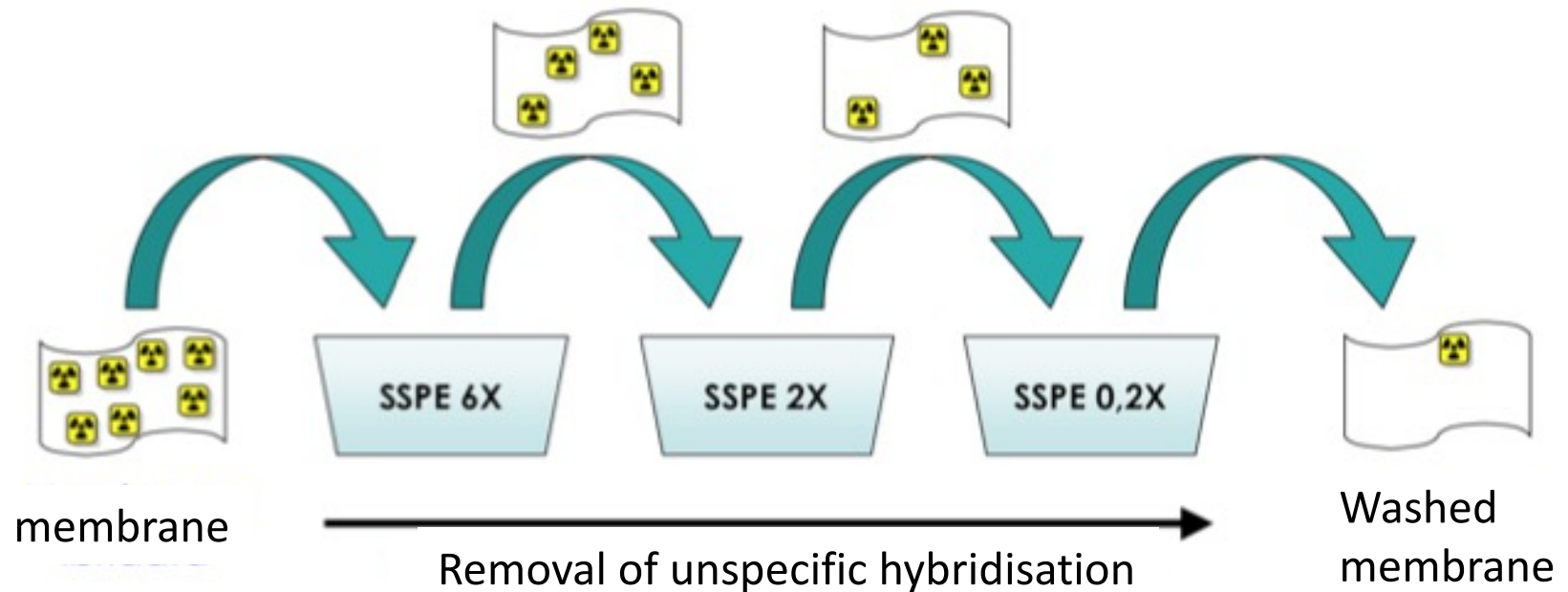
Hybridisation

A specific labelled probe is added to the membrane in the hybridisation tube for 16 hrs at 37-42 °C in the hybridisation buffer

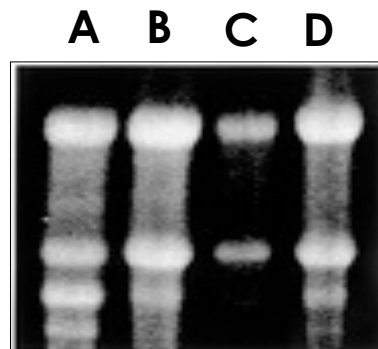
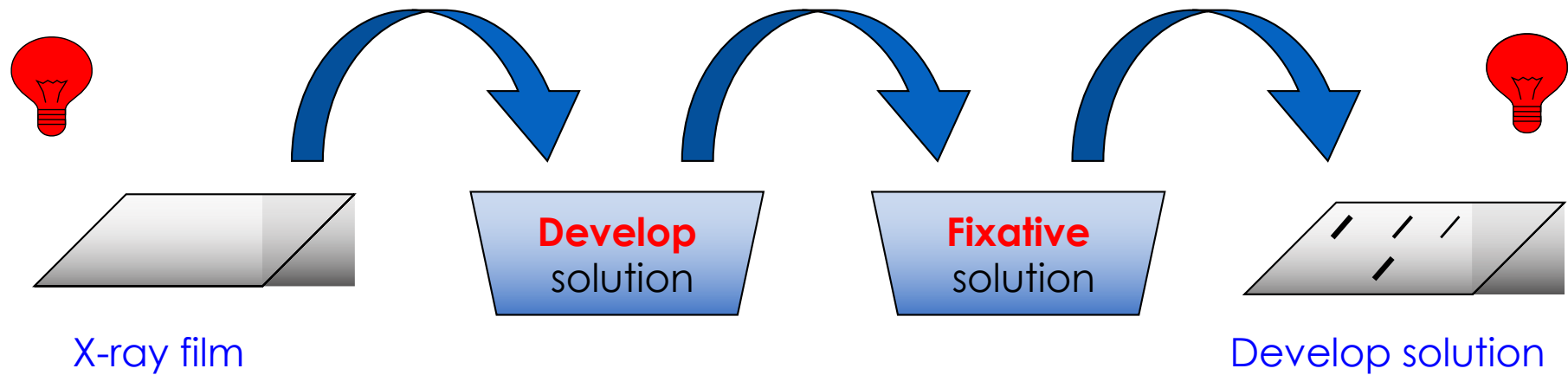


Washes

Several washes are needed in order to wash out the unspecific binding to the membrane. The unspecific hybridisation is not facilitated by high temperature and low salt. In this way will be removed the molecules weakly bound to the membrane



Detection of the signal



Gel



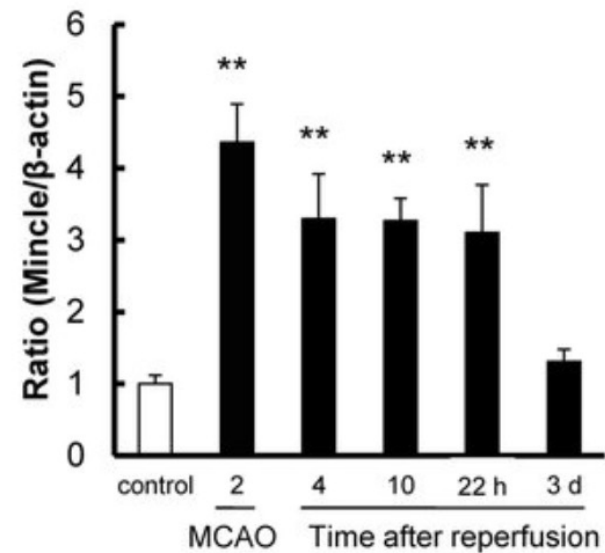
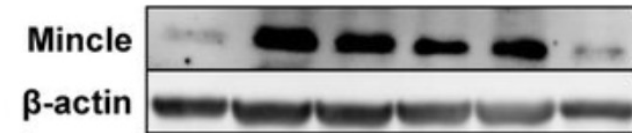
X-ray

← Target RNA

APPLICATIONS

It provides several pieces of information:

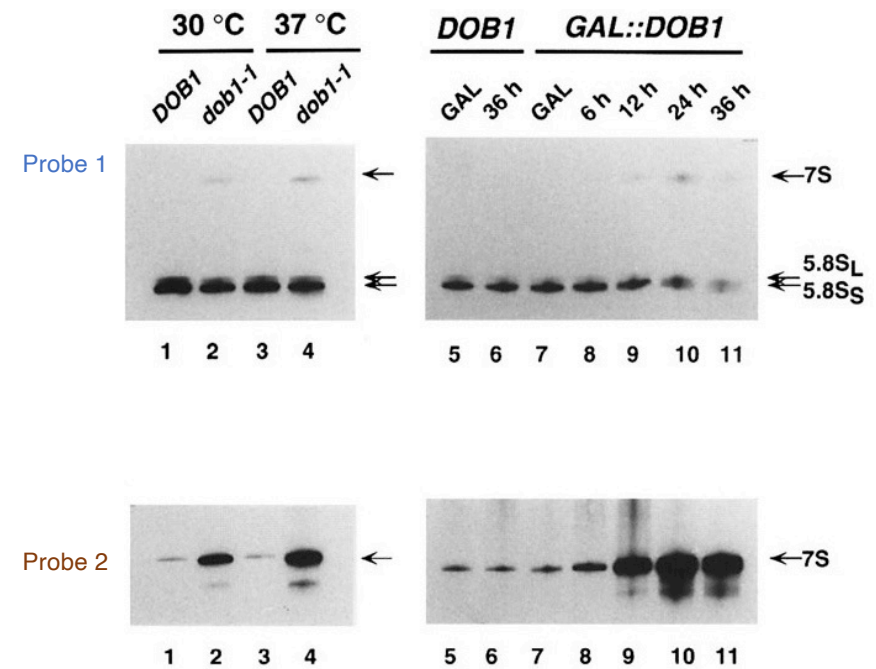
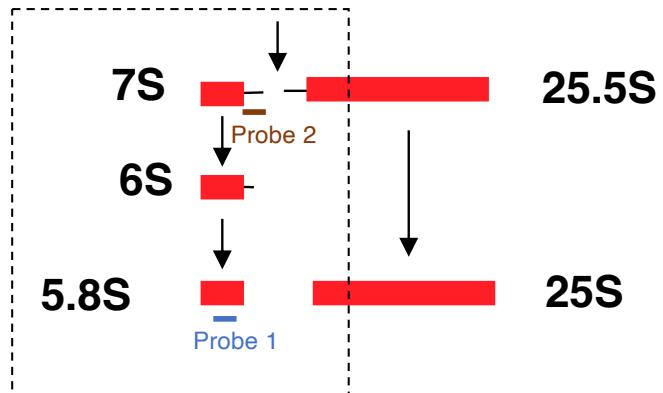
- 1- Expression levels (relative to the control gene)
- 2- Length of the RNA produced



REPLACED BY RT-PCR

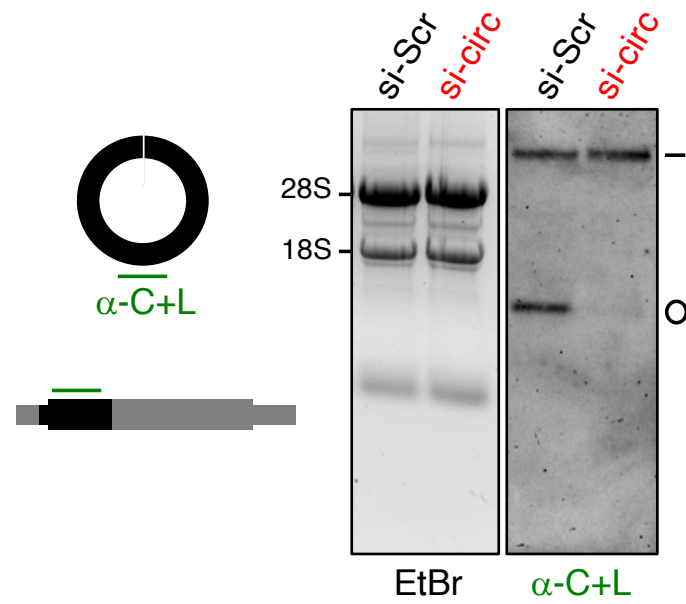
APPLICATIONS

rRNA processing



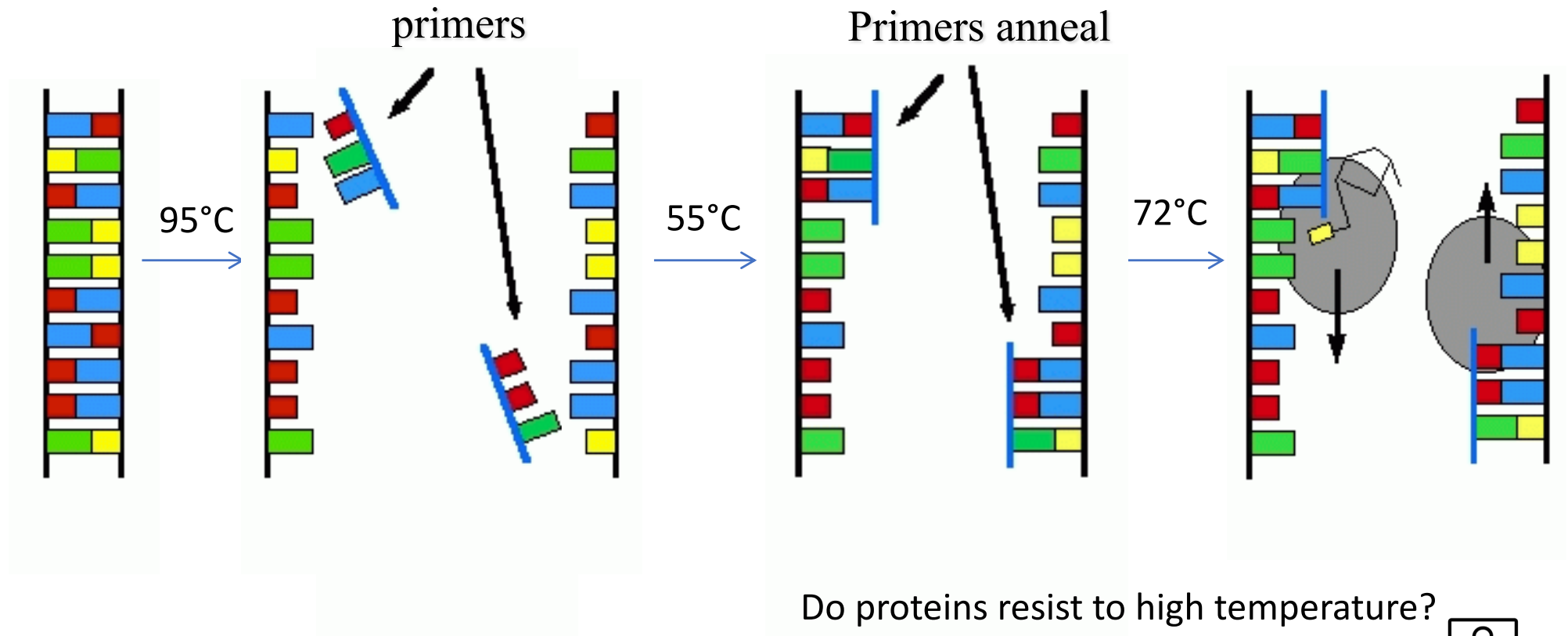
APPLICATIONS

circRNA

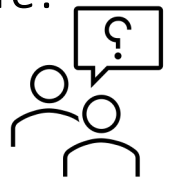


The reaction

PCR is based on base complementarity and the DNA ability to renature after being denatured



Do proteins resist to high temperature?



The Thermocycler



2.000-3.000 Euro

Metal block programmed to adjust the sample to specific temperatures for a defined time and n of cycles

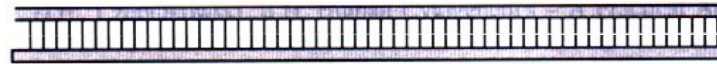
Baking time



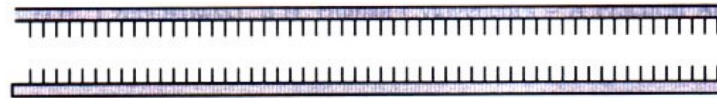
| | |
|------------------------|----------------------------------|
| template | y (ng) |
| primer F (10mM) | 1mM (10 μ l) |
| primer R (10mM) | 1mM (10 μ l) |
| dNTP 2.5mM | 0,25 mM (10 μ l) |
| Polymerase | x |
| Buffer 10X | 1X (10 μ l) |
| MgCl ₂ 50mM | \cong 2mM (2 μ l) |
| H ₂ O | up to final volume (100 μ l) |



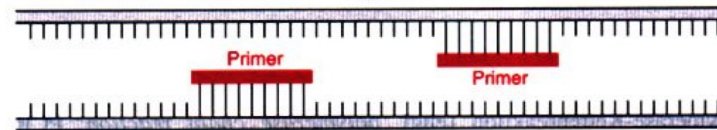
I cycle...



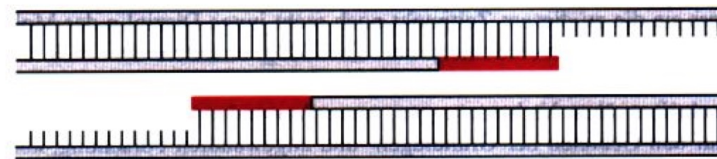
dsDNA



DENATURATION

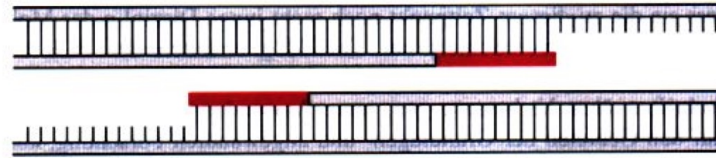


ANNEALING

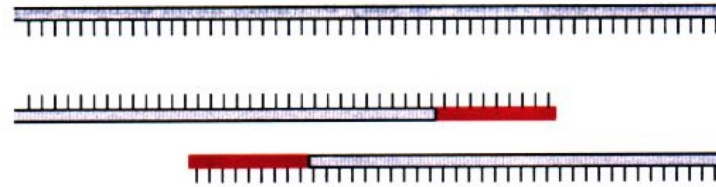


SYNTHESIS

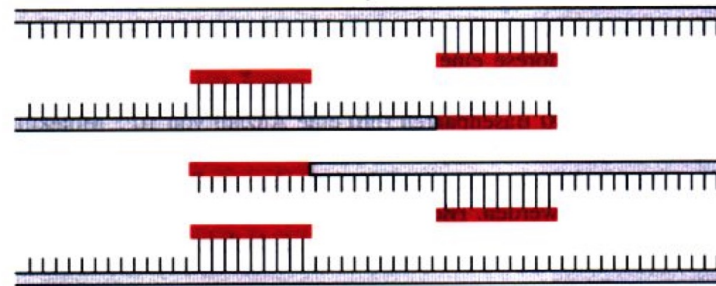
..II cycle



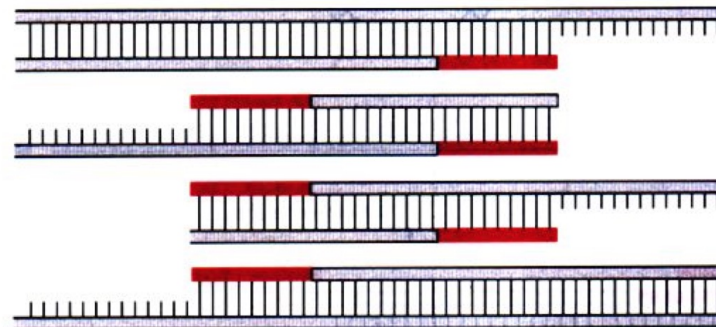
dsDNA



DENATURATION

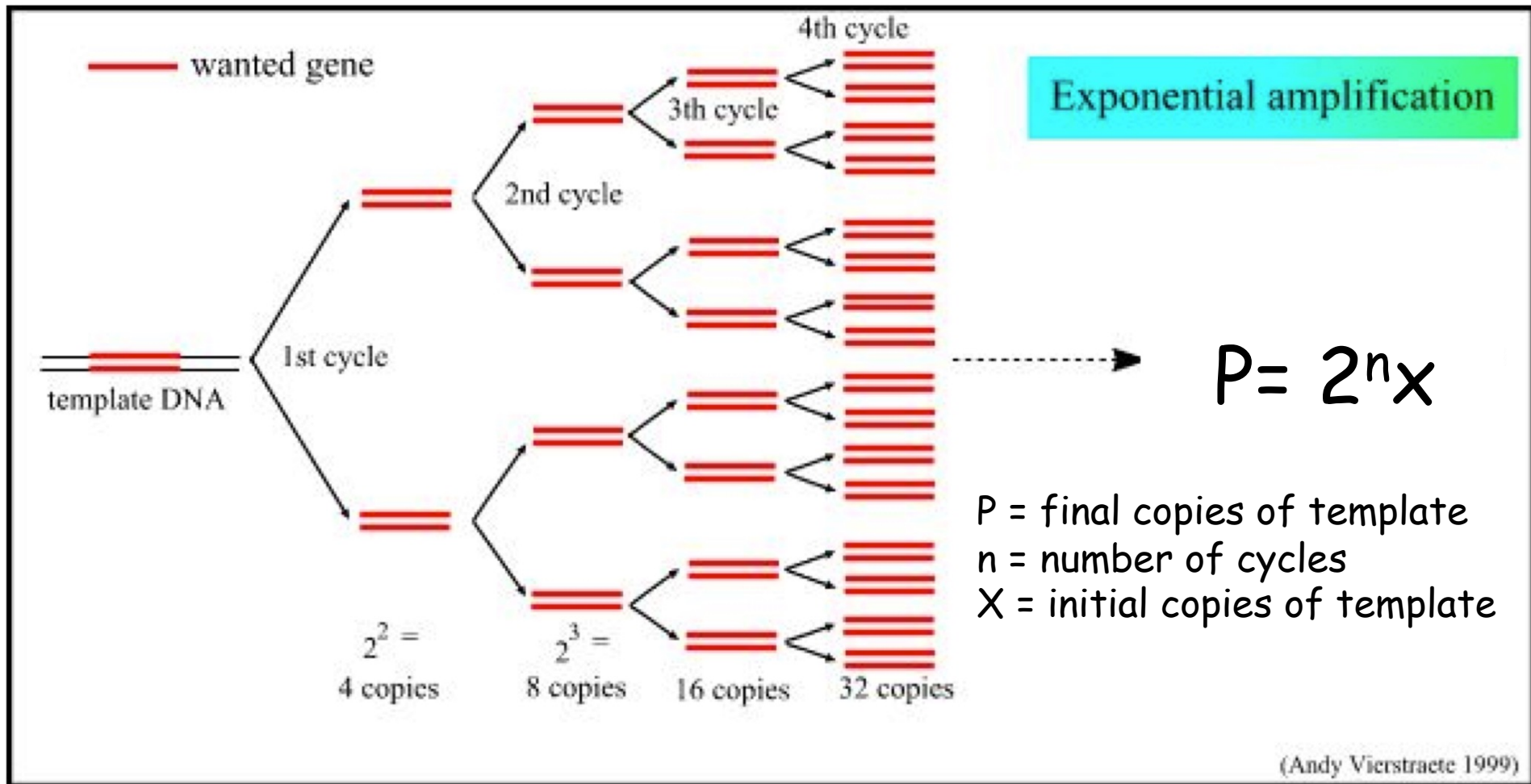


ANNEALING



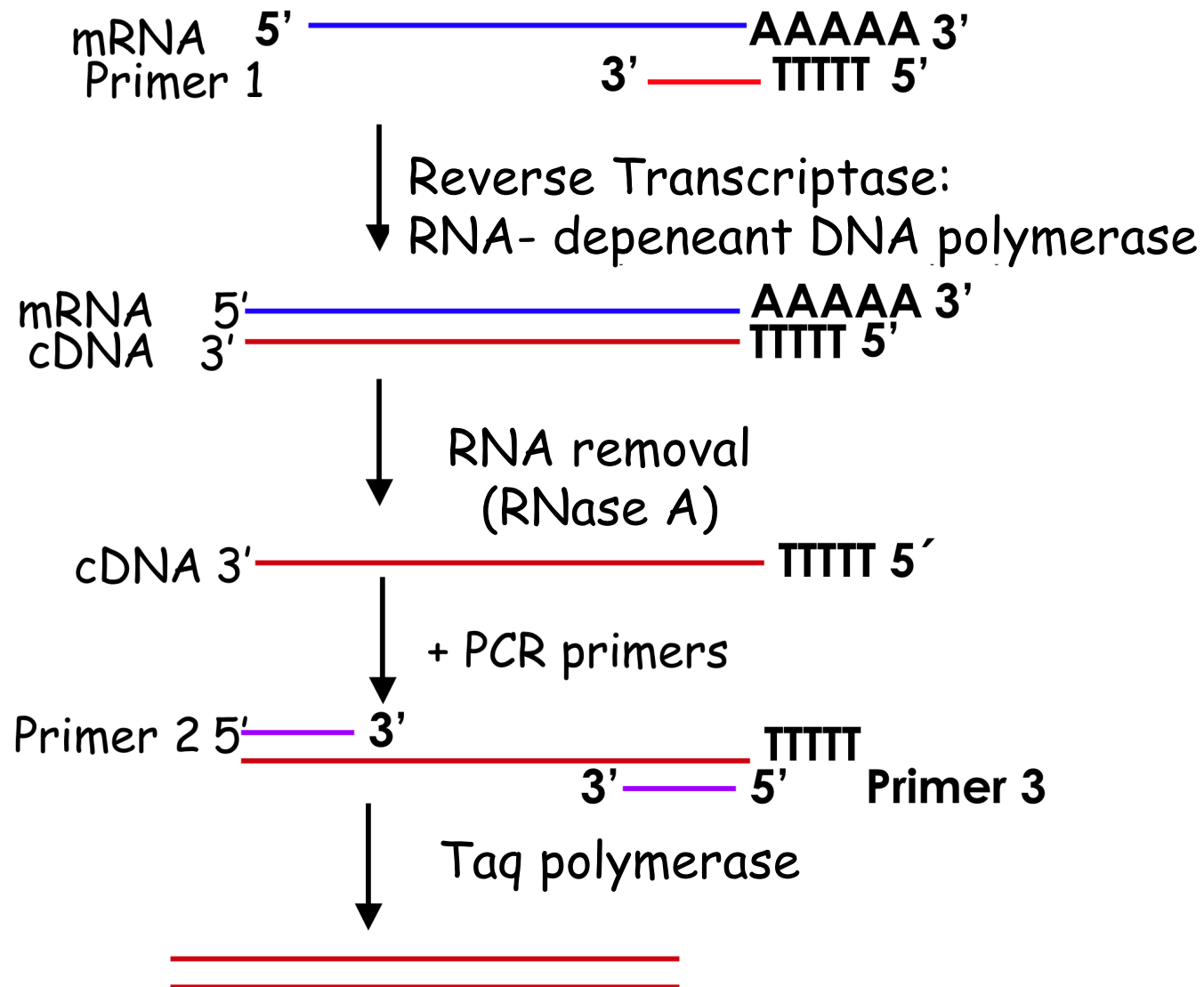
SYNTHESIS

....n cycles



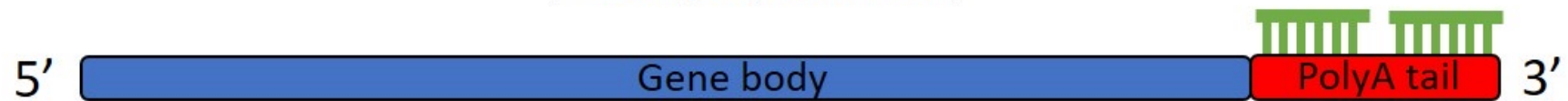
The RT-PCR

RT-PCR: Reverse Transcription PCR



Many RT primers

Oligo(dT) primers



Random hexamers



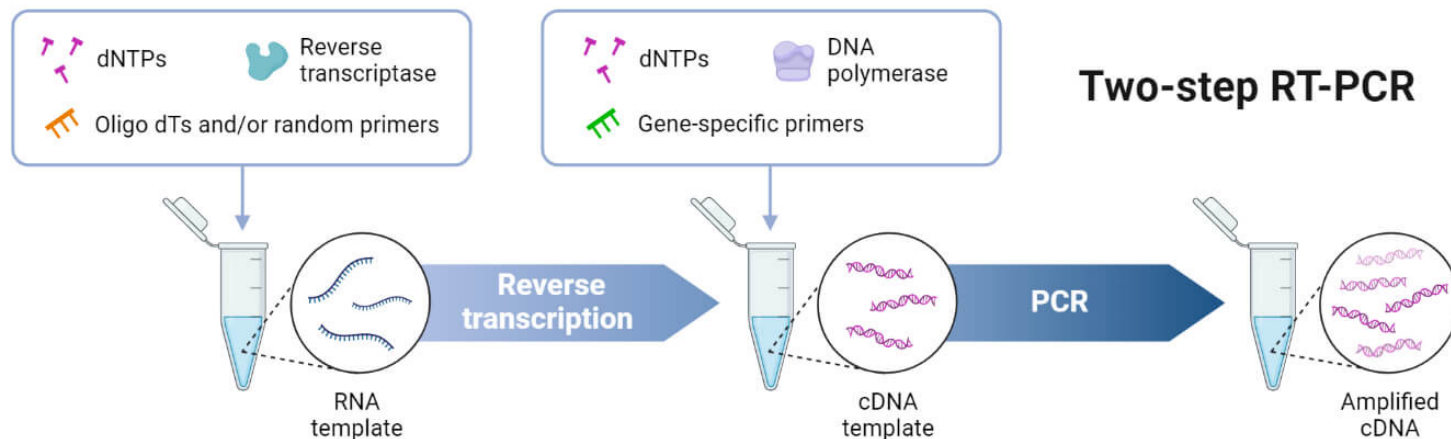
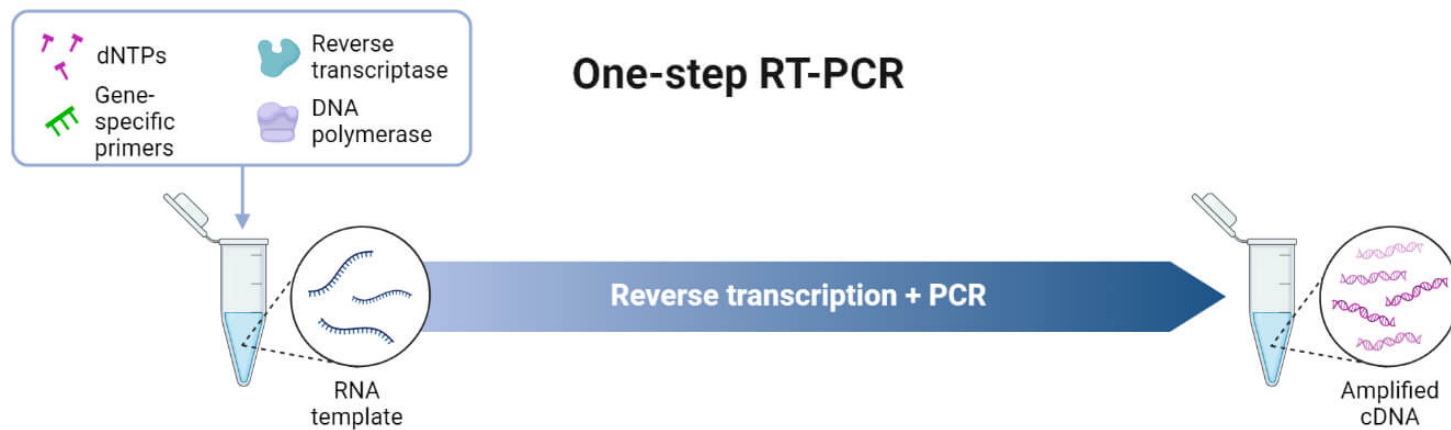
Region-specific primers



One-step vs Two-step RT-PCR

RT-PCR proceeds with two steps:

1. cDNA synthesis (to make cDNA from RNA)
2. PCR (to amplify cDNA)



Semiquantitative RT-PCR

Semiquantitative PCR is a gene expression quantification technique based on classic PCR. This technique exploits the principle that if you perform the same PCR reaction on two samples (i.e., using the same primers and the same number of cycles), you can quantify the intensity of the resulting band and determine the relative amount of initial cDNA present in your two samples.

This technique is used to get an idea (albeit a rough one) of the expression levels of a gene under different biological conditions (e.g., healthy cell vs. tumor cell, cell at an early differentiation stage vs. differentiated cell).

PROS

- Speed and ease of use
- High sensitivity (amplification is possible starting from low quantities of DNA, even if it's of low quality)

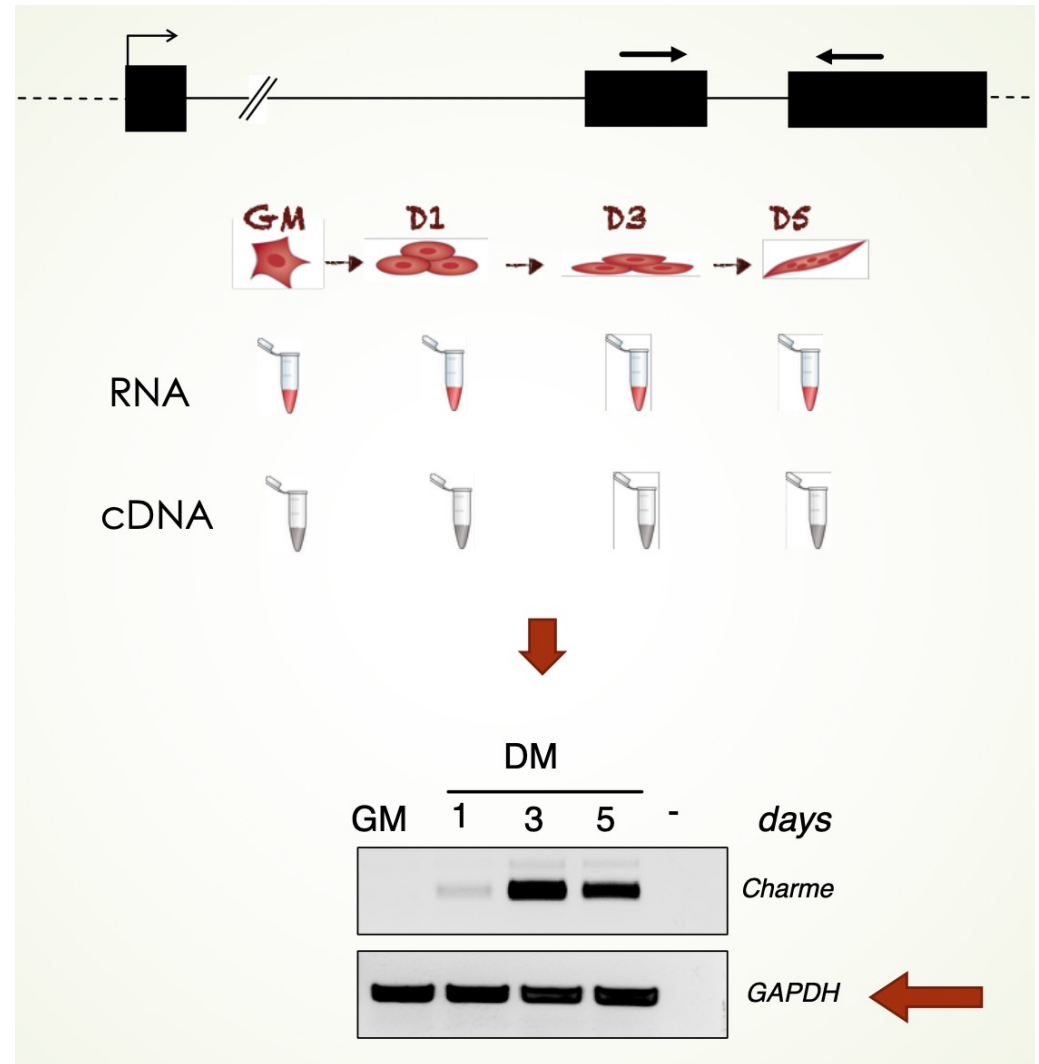
CONS

- Difficulty in obtaining a **correct quantification** of the final product

Semiquantitative RT-PCR

Example: I want to quantify the expression of certain genes that I suspect are important during muscle differentiation.

- perform an in vitro differentiation.
- extract the RNA from my cells.
- reverse-transcribe the RNA into cDNA (random hexamers or oligo-dT).
- perform PCR with specific oligos for my gene (Target) and an endogenous control (e.g., GAPDH).
- run an agarose gel electrophoresis.
- quantify the resulting bands.



PCR: Theoretical Yield

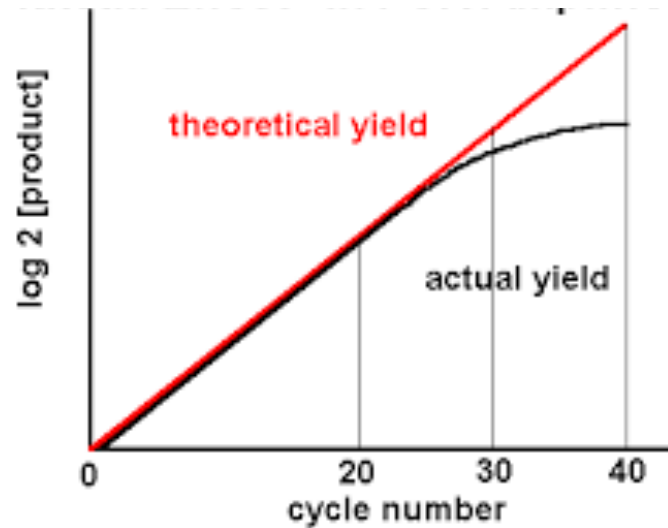
$$P = (2)^n T$$

The PCR product (P) increases exponentially with the number of cycles (n).

Theoretically, P is directly proportional to the number of starting DNA copies (T).

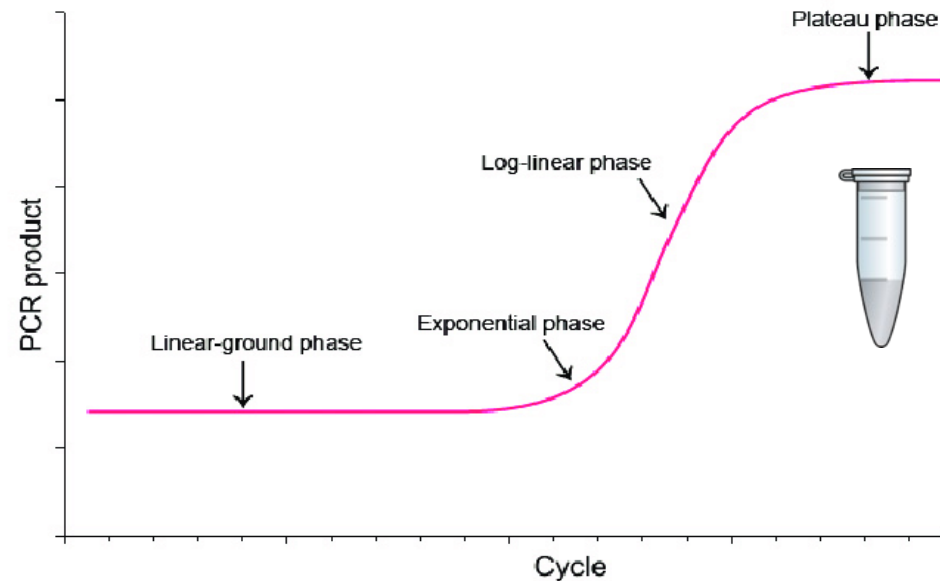
The plateau effect

PCR doesn't go forever

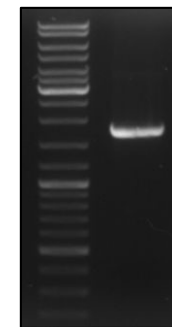


"PLATEAU" effect

- Reagents are going over
- Taq polymerase activity decreases
- Reannealing of the filaments



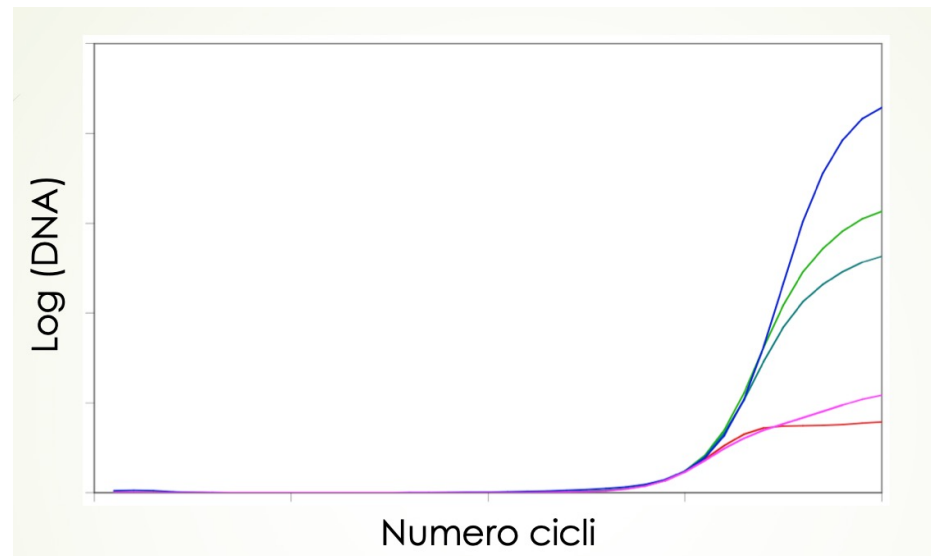
End-point analysis



How can we obtain data from the exponential phase?
How could we **quantify DNA** in the reaction in *real-time*?

The plateau effect

Effective Yield



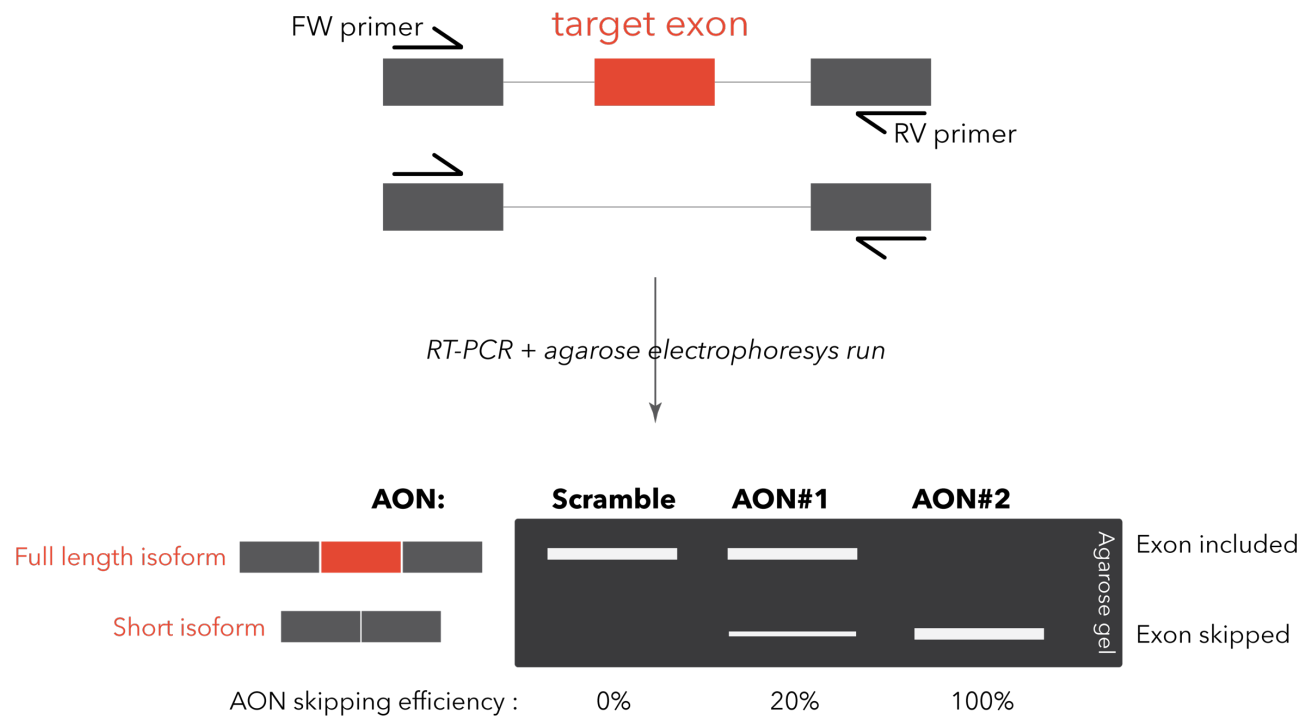
Even if the initial quantity of template is the same, the plateau is reached at different times and in different cycles.

At the end of the reaction, the quantity of amplified product obtained is NOT always proportional to that of the starting DNA.

PCR is therefore a primarily QUALITATIVE, NOT QUANTITATIVE methodology.

Semiquantitative RT-PCR

It is used for the study of alternative splicing because it allows the visualization of different isoforms with a single amplification



For a QUANTITATIVE estimate of the amplified product, it is necessary to use the data obtained during the exponential phase of the amplification process, where the **PCR product is proportional to the initial template**. Therefore, every amplified sample needs to be monitored at every replication cycle

How can we obtain data from the exponential phase?

How could we **quantify DNA** in the reaction in *real-time*?

qRT-PCR: The PCR that “quantifies”

REAL TIME PCR or qRT-PCR allows to quantify the synthesis of PCR product at every stage of amplification *in real-time*

Thanks to **fluorescent reporter dyes that bind to DNA**

Quantitative measurement: **the fluorescence signal is proportional to the number of target DNA copies**



qRT-PCR mix

- Template
- DNA Polymerase
- dNTPs
- **Reporter dye**
- Passive reference (ROX)
- Specific Primers



25.000-80.000 Euro

Reporter dyes

Reporter DYE:

NON – SPECIFIC

Fluorescent intercalant

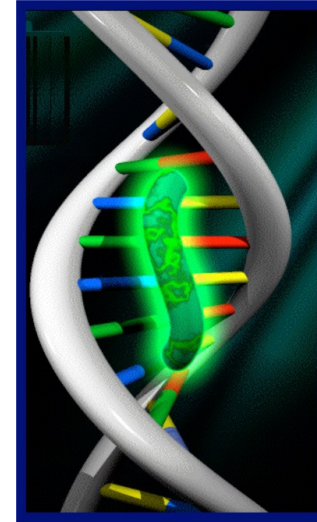
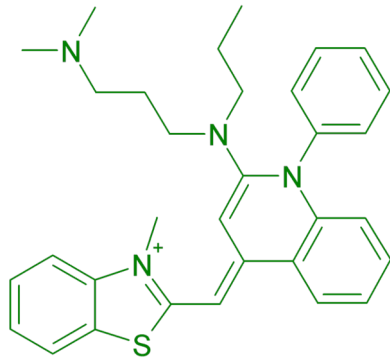
(**SYBER GREEN**)

SPECIFIC

Molecular probes marked with fluorescent molecules

(**TAQ-MAN, Molecular beacon, etc.**)

Syber green



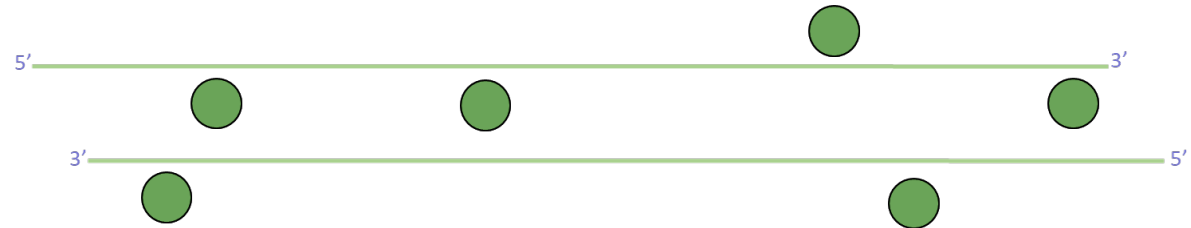
Syber green is not fluorescent in solution, it is when it's inside the minor groove of DNA

In dsDNA, Syber green absorbs blue light and emits green light

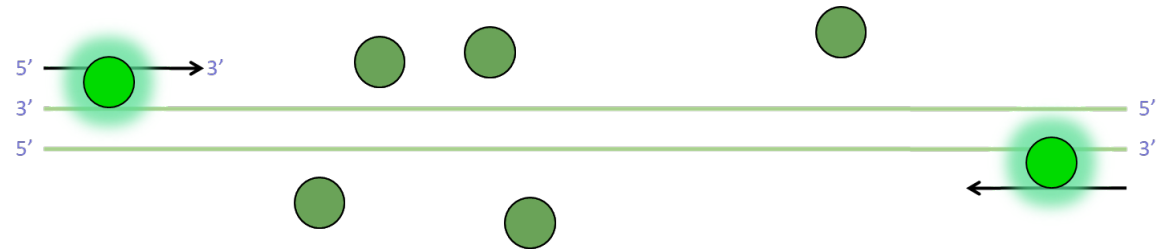


Syber green

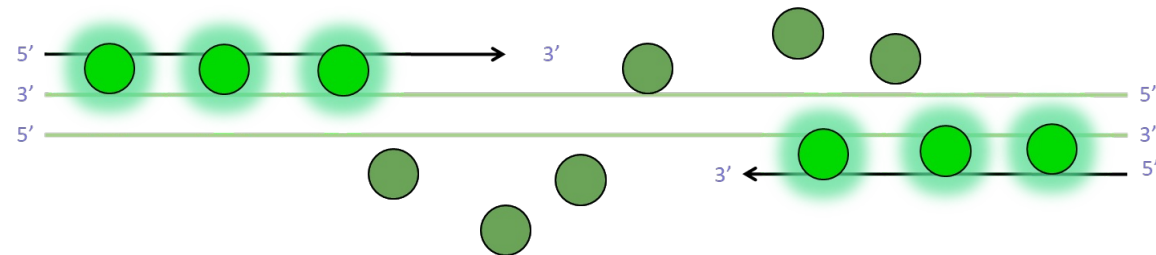
1) Denaturation



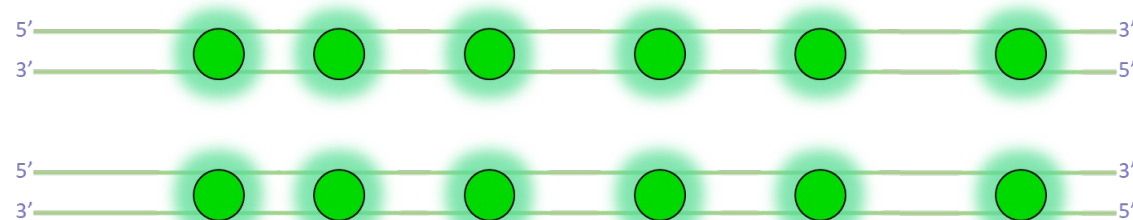
2) Annealing



3) Extension



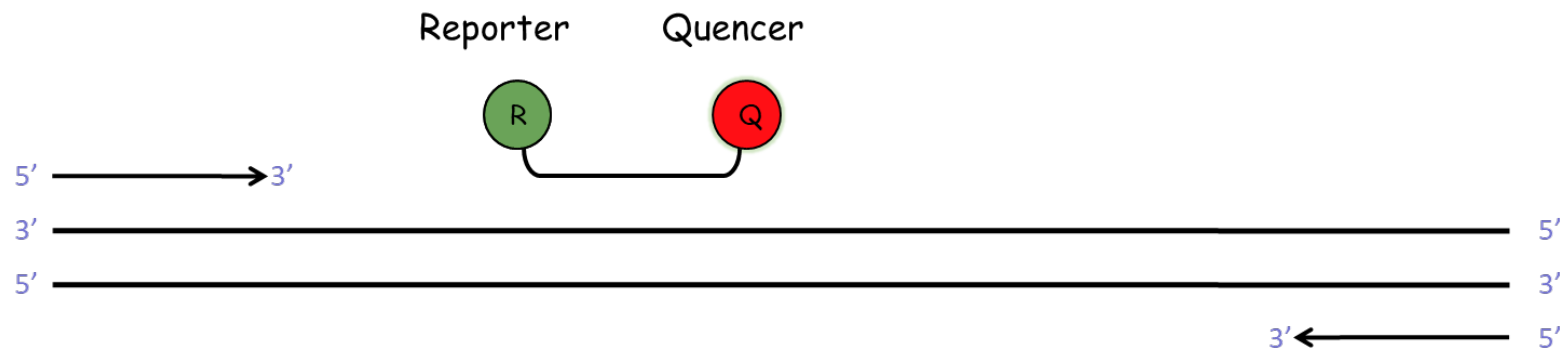
4) End of first cycle



Taq-man probe

The TaqMan probe is an oligonucleotide that, as the primers used during the PCR, is complementary to target sequence.

It is designed to be complementary to a region inside target sequence and contains a «Reporter» fluorophore at the 5' and a «Quencher» molecule at the 3'.

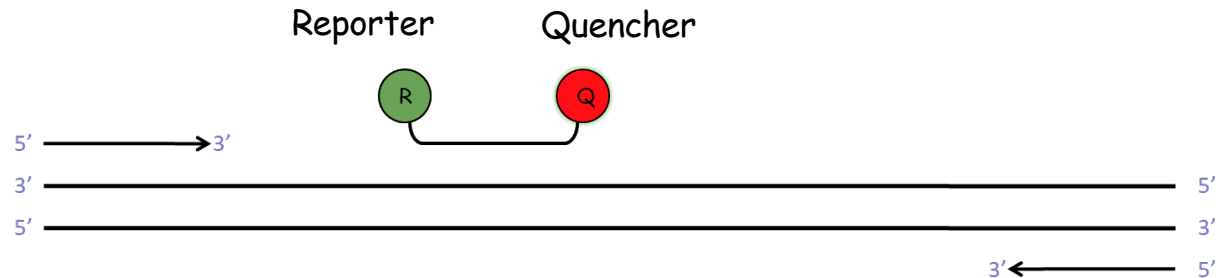


Taq-man probe

The TaqMan probe is an oligonucleotide complementary to a region in the target sequence. It contains a «Reporter» fluorophore at the 5' and a «Quencher» molecule at the 3'.

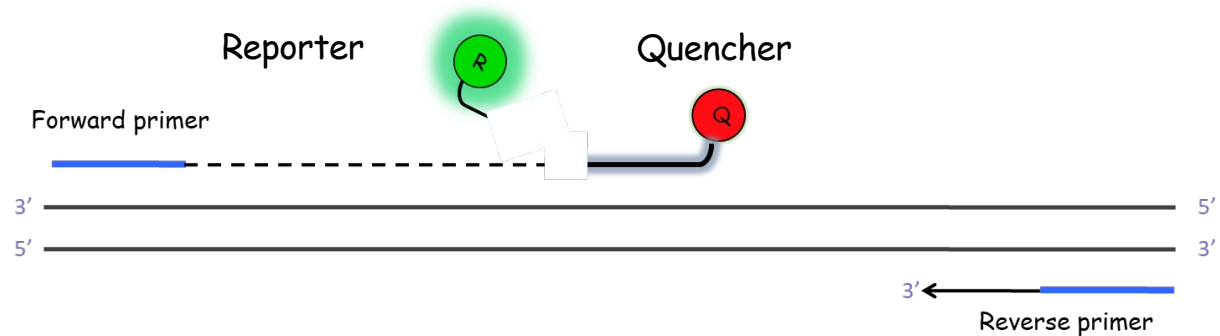
1) Denaturation / Annealing

Photons emitted by R are absorbed by the Q = no fluorescence

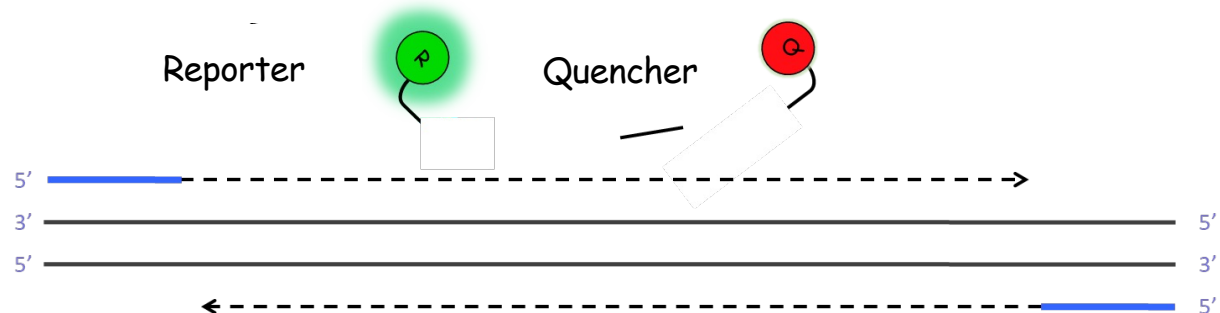


2) Extension

DNA polymerase 5'→3'
exonuclease activity cuts off the R

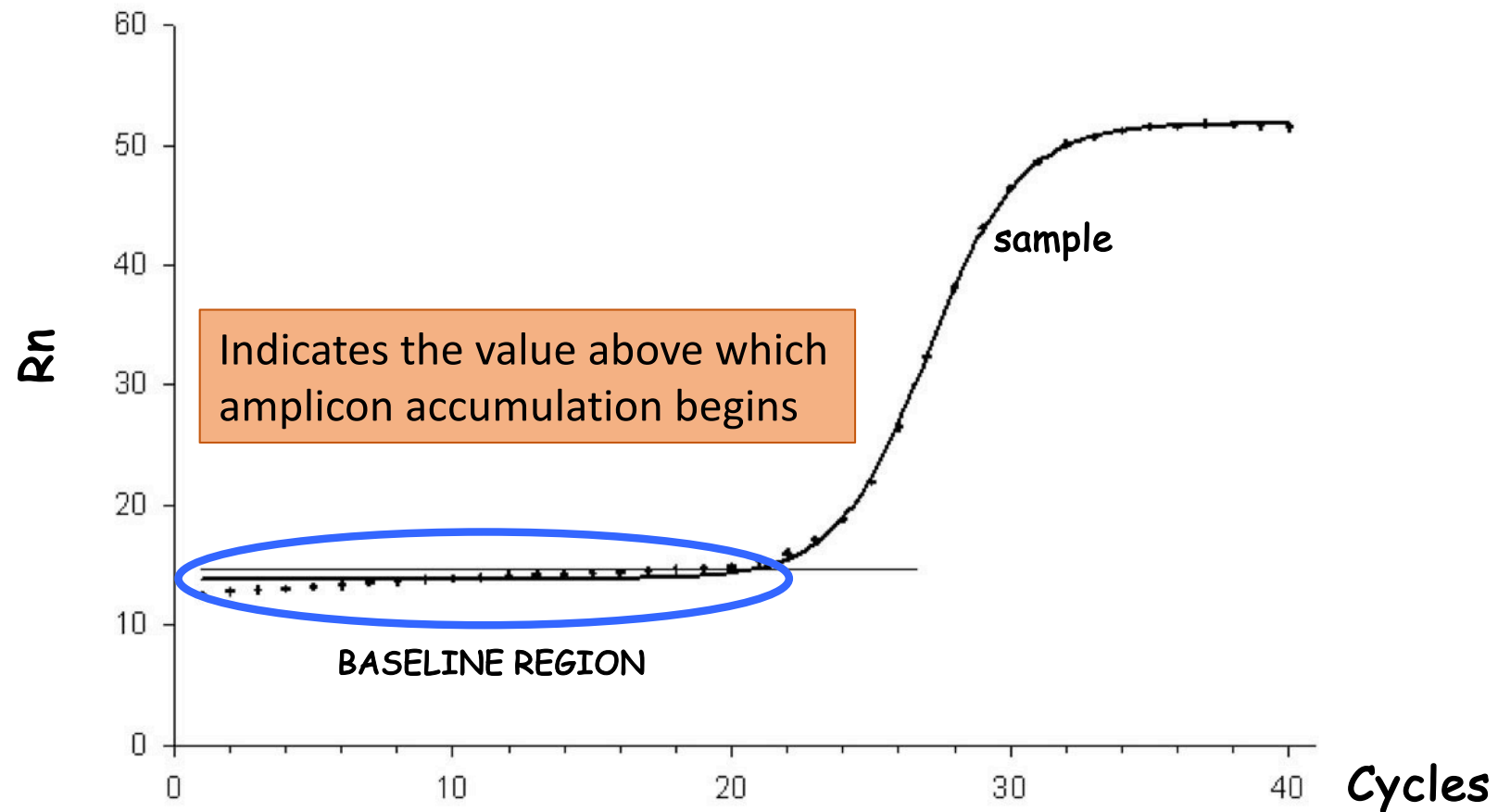


Once far from the Q,
the R can fluoresce



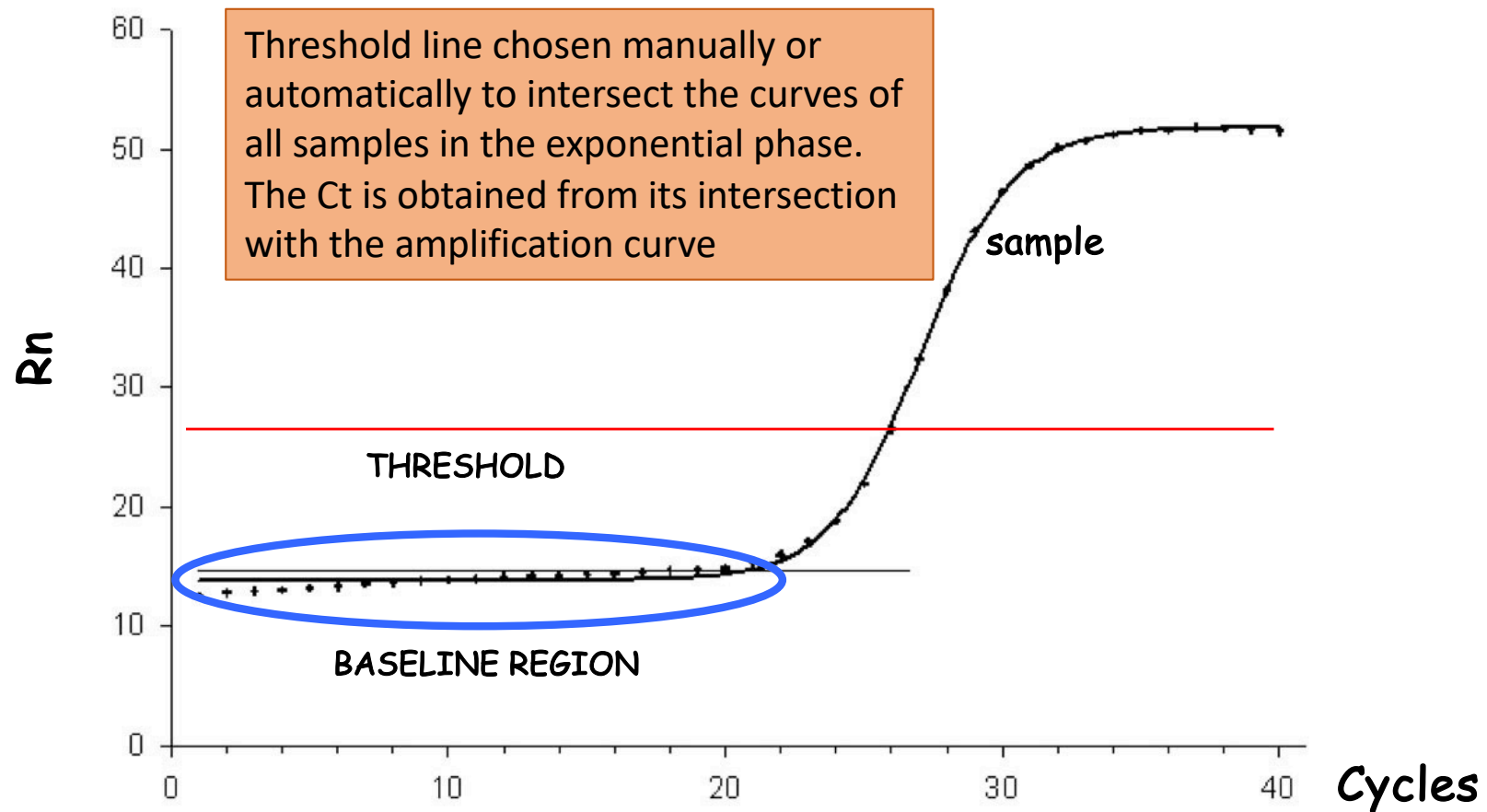
Amplification plot

The fluorescence (R_n) is recorded at every cycle



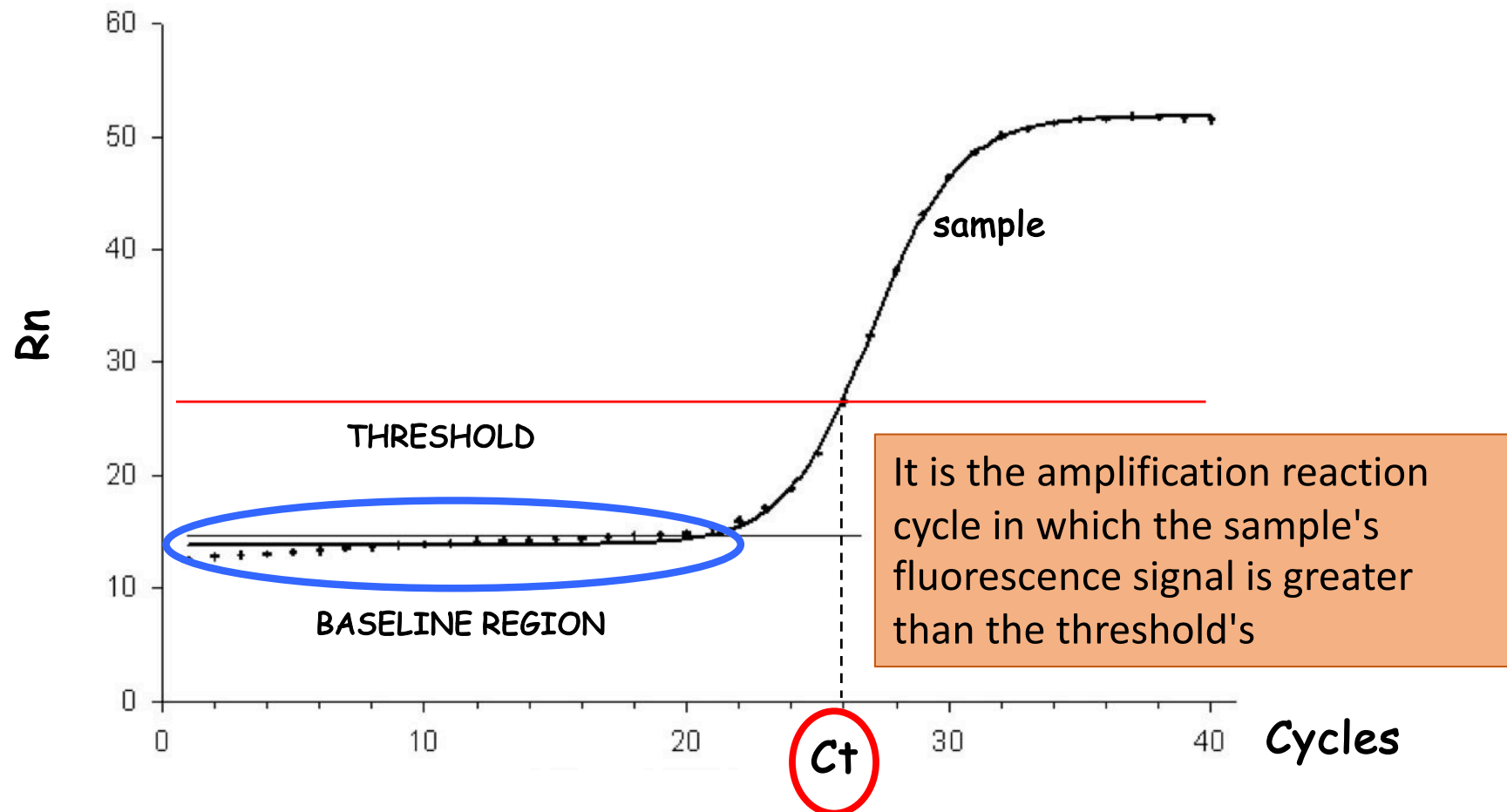
Amplification plot

The fluorescence (R_n) is recorded at every cycle

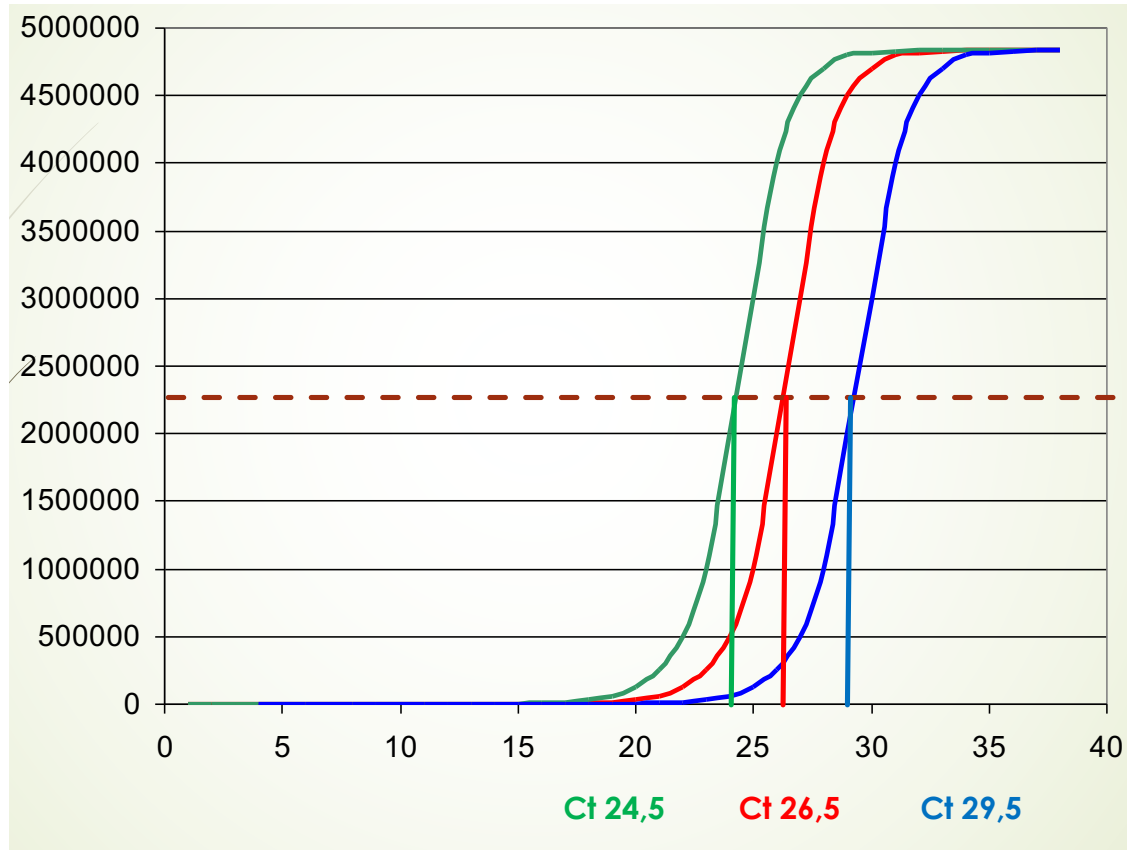


Amplification plot

The fluorescence (R_n) is recorded at every cycle



Amplification plot



Target quantity in each sample:

A: 1 000 copies

B: 8 000 copies

C: 32 0000 copies

THRESHOLD

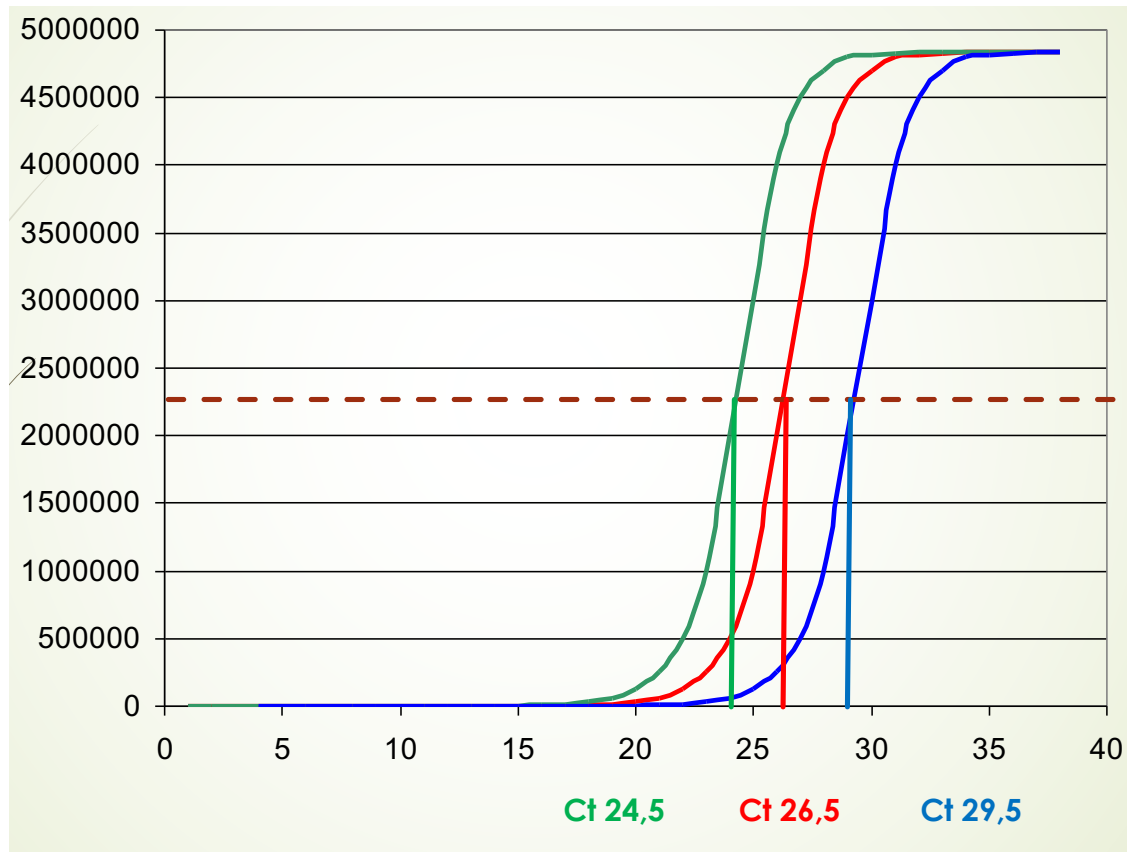
Cts are the measure we need!

← More target in the sample Less target in the sample →

For a given target, each target in each sample has its own Ct

Ct values are inversely proportional to the initial amount of the target in the sample

Amplification plot



Target quantity in each sample:

A: 1 000 copies

B: 8 000 copies

C: 32 0000 copies

THRESHOLD

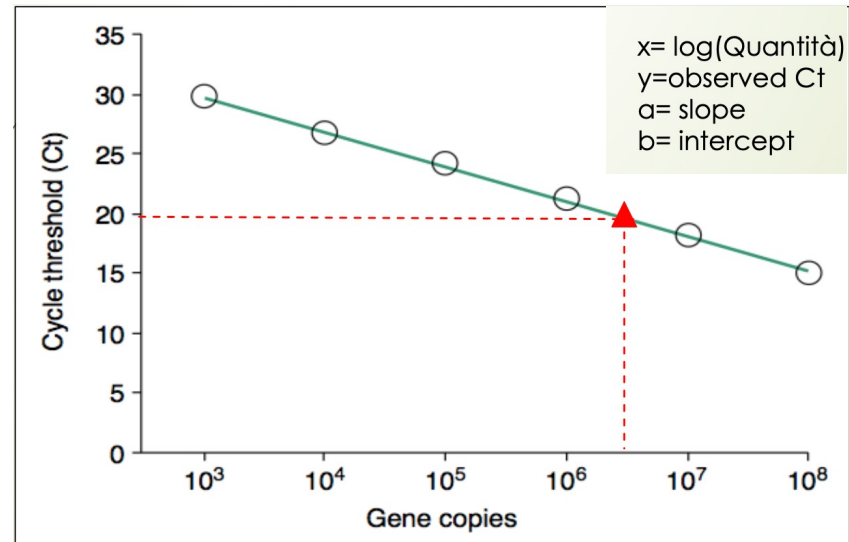
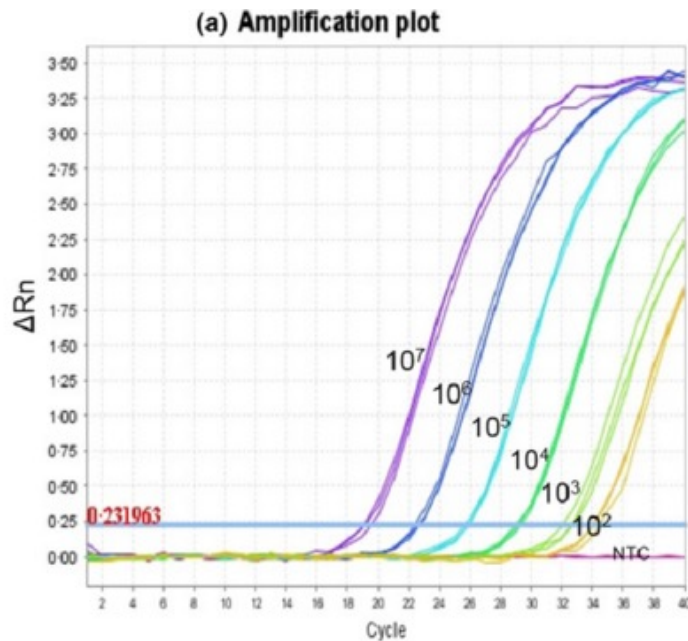
Cts are the measure we need!

Several factors, besides the initial concentration of the target, influence the absolute value of the Ct, such as:

- **Efficiency of the PCR reaction** (normally dependent on the primers used and the Master Mix)
- **pH and salt concentration of the Master Mix**

Absolute Quantification

Serial dilutions of a sample with a known concentration (standard) are used for the realization of a standard curve, described by the following equation: $y=ax+b$.



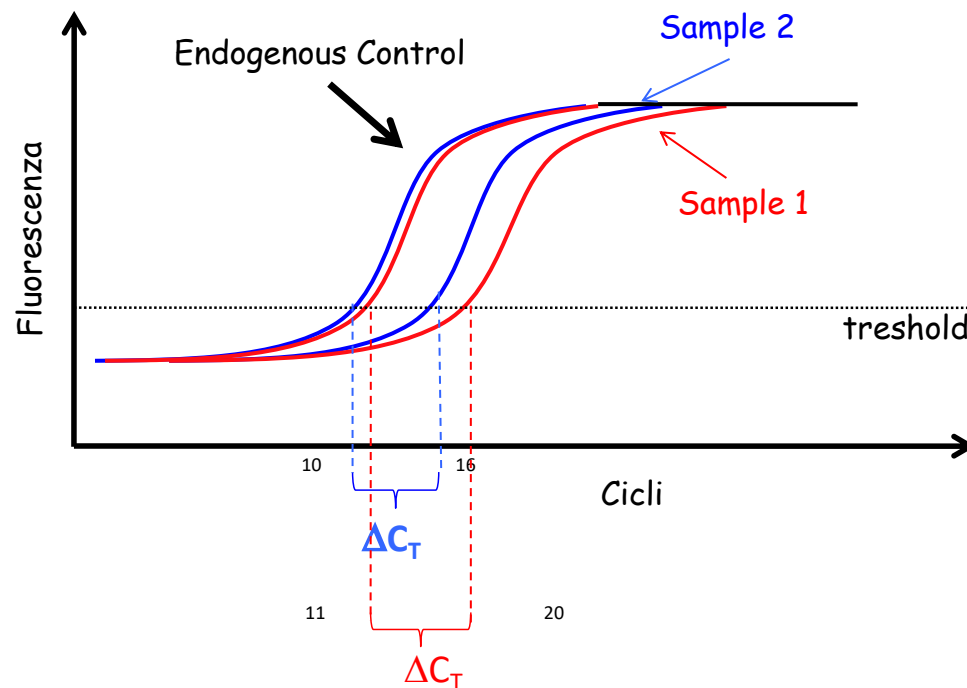
The standard curve is necessary for:

- The evaluation of the **EFFICIENCY** of the **REAL-TIME PCR reaction** is associated with the **R² value** or **coefficient of determination**, which represents a proportion between the variability of the data and the correctness of the statistical model used. **R²≥0.99 indicates an excellent correlation** between the obtained data and the standard curve.

- The determination of the unknown concentration sample through interpolation with the standard curve.

Relative quantification

Performed by comparing the Ct values to determine the change in expression of a TARGET gene relative to another gene chosen as the CALIBRATOR.



The values thus obtained will be normalized against an endogenous control, which is a gene constitutively expressed in all analyzed samples, necessary to normalize the data with respect to the amount of DNA loaded and variations in reaction efficiency

Relative quantification

Normalization:

$$\Delta Ct_{\text{sample1}} = Ct_{\text{sample1}} - Ct_{\text{reference gene}}$$

$$\Delta Ct_{\text{sample2}} = Ct_{\text{sample2}} - Ct_{\text{reference gene}}$$

Is there difference between my samples?

$$\Delta Ct_{\text{sample2}} - \Delta Ct_{\text{sample1}} = \Delta\Delta Ct$$

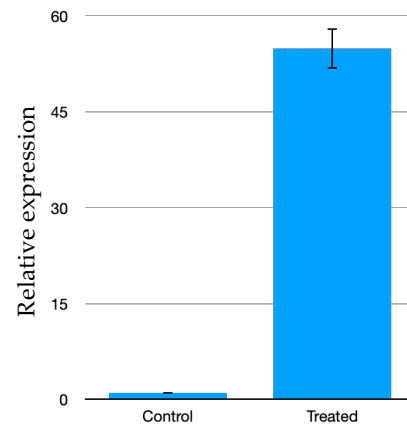
$$\text{fold change} = 2^{-\Delta\Delta Ct \text{ sample}}$$

Relative quantification

Biological replicates 3
technical replicates 2

| Houskeeping gene | | | | Gene of interest | | | | | | | |
|------------------|----------|----------|--------|------------------|----------|----------|--------|--|---|------------------------|---------|
| Sample | Ct rep.1 | Ct rep.2 | Avg.Ct | Sample | Ct rep.1 | Ct rep.2 | Avg.Ct | $\Delta Ct = \text{Avg Ct gene of interest} - \text{Avg Ct houskeeping}$ | $\Delta\Delta Ct = \Delta Ct - \text{Avg } \Delta Ct \text{ control}$ | $2^{-\Delta\Delta Ct}$ | |
| Control 1 | 17.19 | 17.16 | 17.18 | Control 1 | 30.57 | 30.53 | 30.55 | 13.38 | -0.18 | 1.1 | Control |
| Control 2 | 16.96 | 16.95 | 16.96 | Control 2 | 30.73 | 30.37 | 30.55 | 13.60 | 0.04 | 1.0 | 1.0 |
| Control 3 | 17.07 | 17.15 | 17.11 | Control 3 | 30.76 | 30.82 | 30.79 | 13.68 | 0.13 | 0.9 | |
| Treated 1 | 18.04 | 17.95 | 18.00 | Treated 1 | 26.11 | 25.54 | 25.83 | 7.83 | -5.72 | 52.7 | Treated |
| Treated 2 | 17.99 | 17.91 | 17.95 | Treated 2 | 25.7 | 25.56 | 25.63 | 7.68 | -5.87 | 58.5 | 54.9 |
| Treated 3 | 17.9 | 17.86 | 17.88 | Treated 3 | 25.64 | 25.74 | 25.69 | 7.81 | -5.74 | 53.4 | |
| | | | | | | | | | | | |
| | | | | | | | | Avg. ΔCt control | | | |
| | | | | | | | | 13.55 | | | |

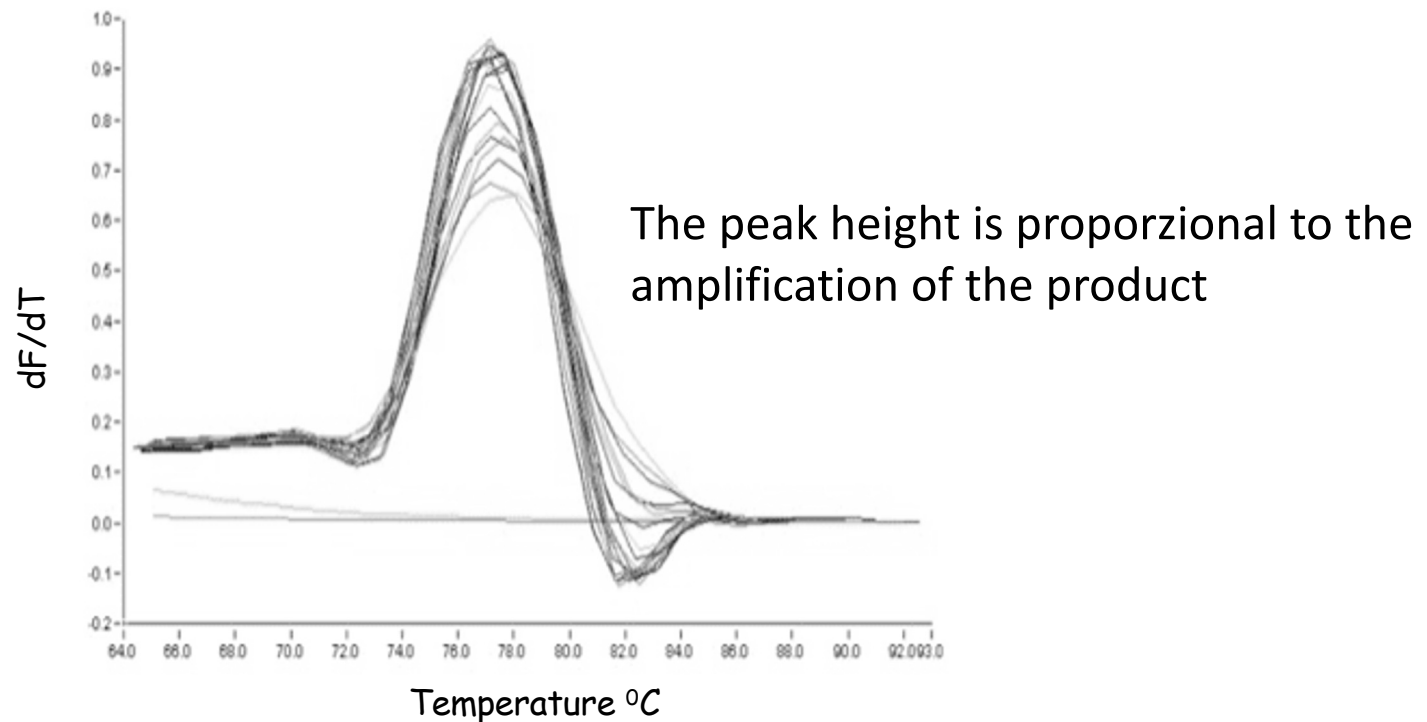
Gene of interest
(relative expression with SD)



The average (ΔCt Avg) of the control is also subtracted from the ΔCt 's of the individual controls to obtain an SD (Standard Deviation) on the value of 1

The Melting curve

At the end of the amplification, the samples are heated, and the change in fluorescence energy is monitored to generate a dissociation curve or 'MELTING CURVE'."

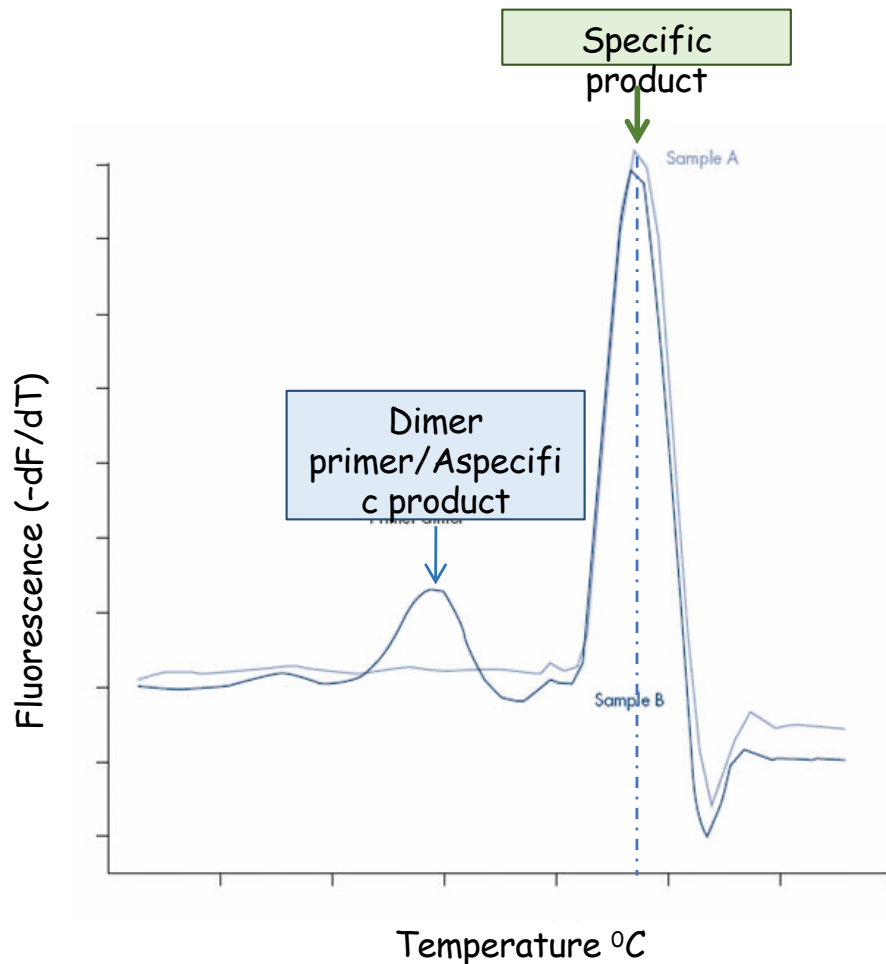


The analysis of the 'MELTING CURVE' consists of evaluating the dissociation characteristics of double-stranded DNA during the temperature increase.

The energy required to break the H-bonds between the nitrogenous bases of the two strands of the DNA molecule is dependent on their length, GC content, and complementarity

The Melting curve

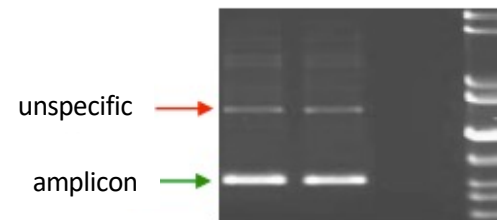
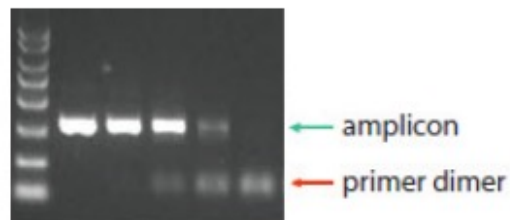
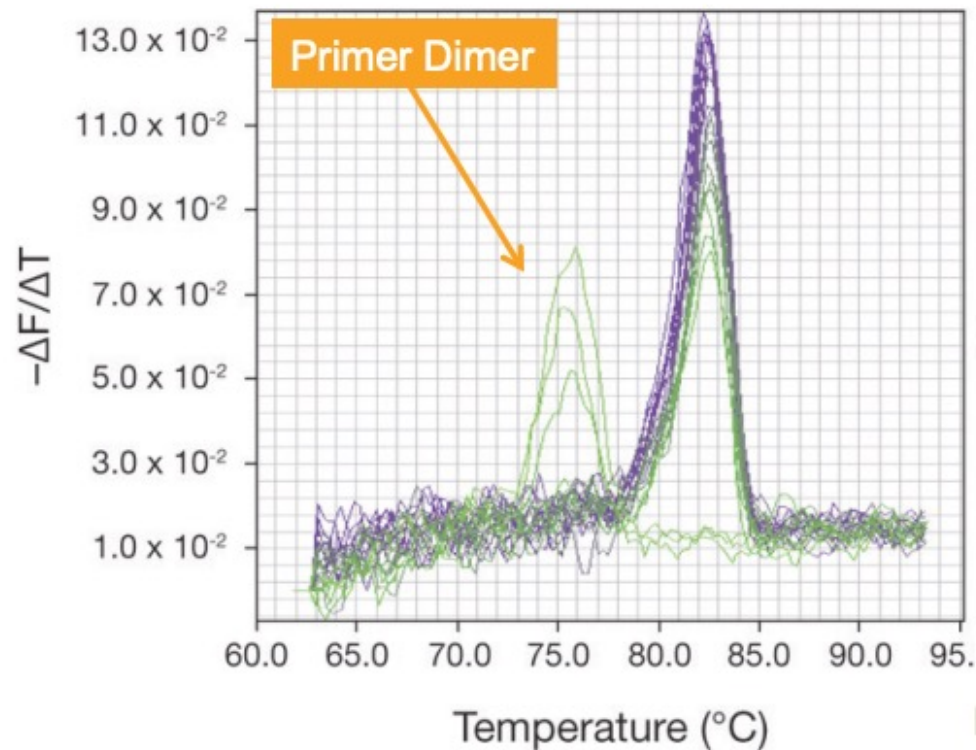
From the analysis of the dissociation curve, it is clearly possible to highlight the presence of dimers, DNA or PCR products resulting from primer misannealing/contamination.



SYBR Green will in fact be associated with all double-stranded DNA molecules, and dissociation peaks will be observed that are different from the one related to the specific product.

Taqman method is not able to detect primer dimers

The Melting curve analysis



REAL-TIME PCR (qRT-PCR)

SYBR-GREEN

PROS

- Simple methodology; the same primers as classic PCR can be used
- Less expensive compared to Taqman probes

CONS

- Does not discriminate between specific and non-specific amplifications
- It is often necessary to test the primers to avoid the formation of non-specific products

Taqman

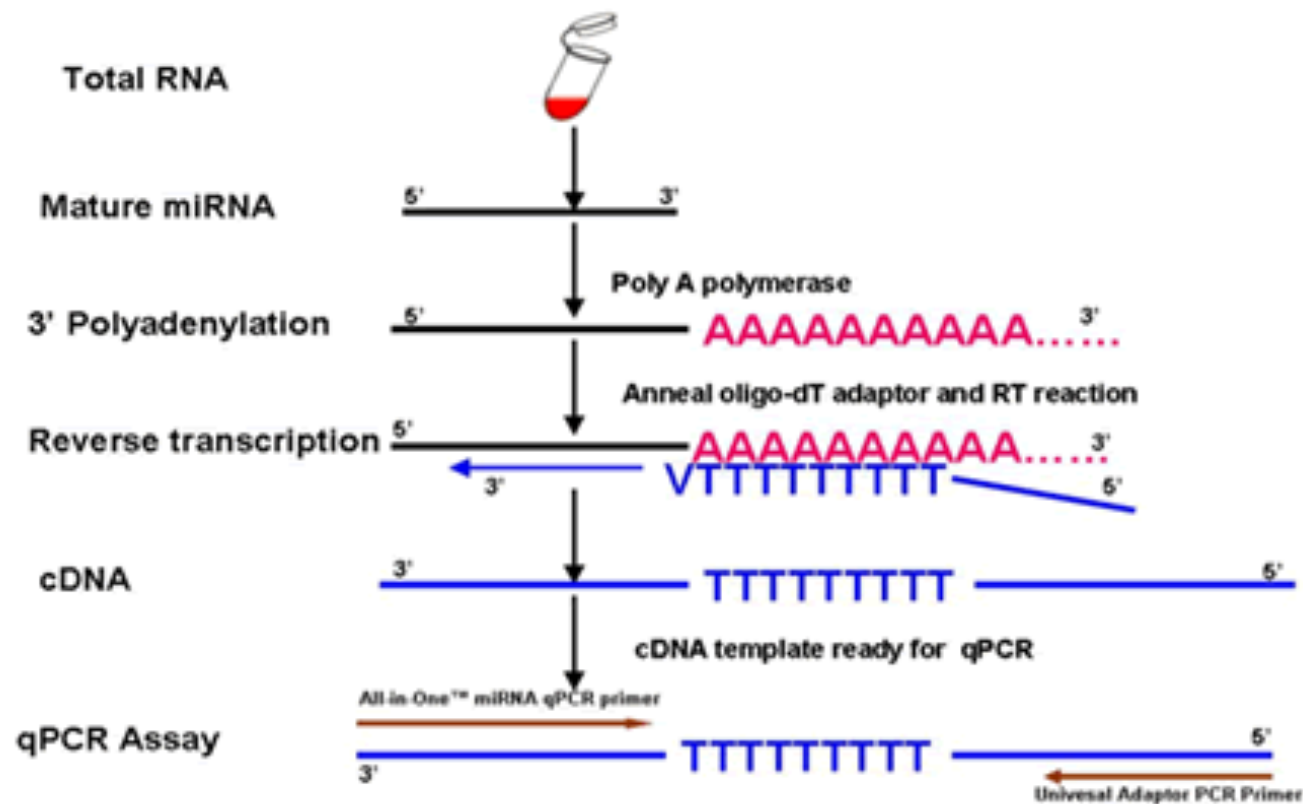
PROS

- High specificity

CONS

- expensive

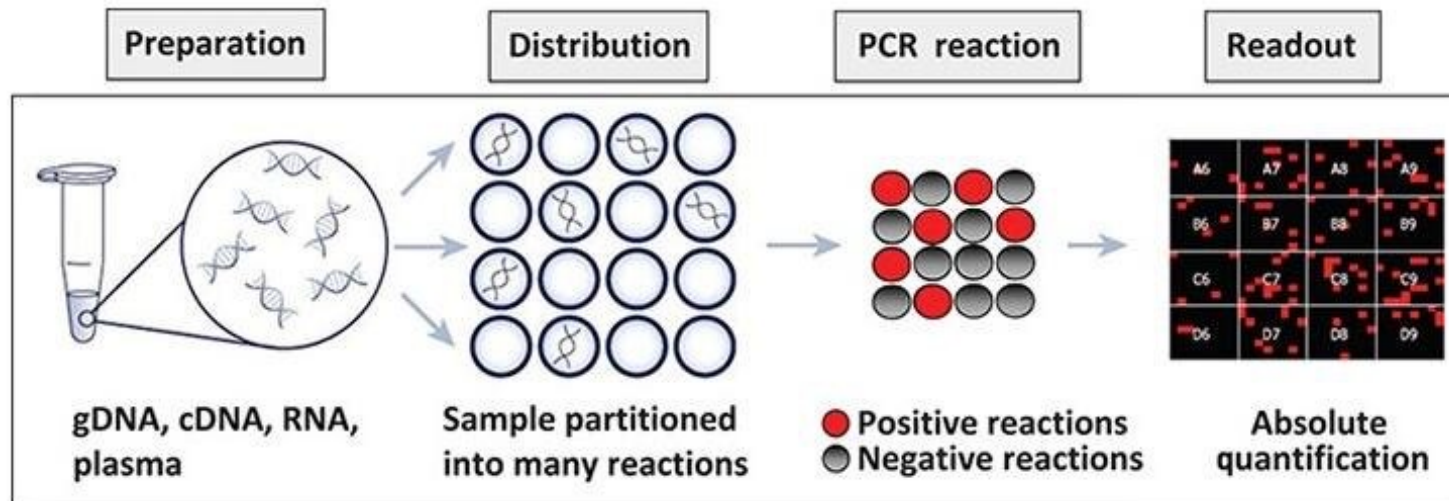
microRNA qRT-PCR Detection



Digital PCR (dPCR)

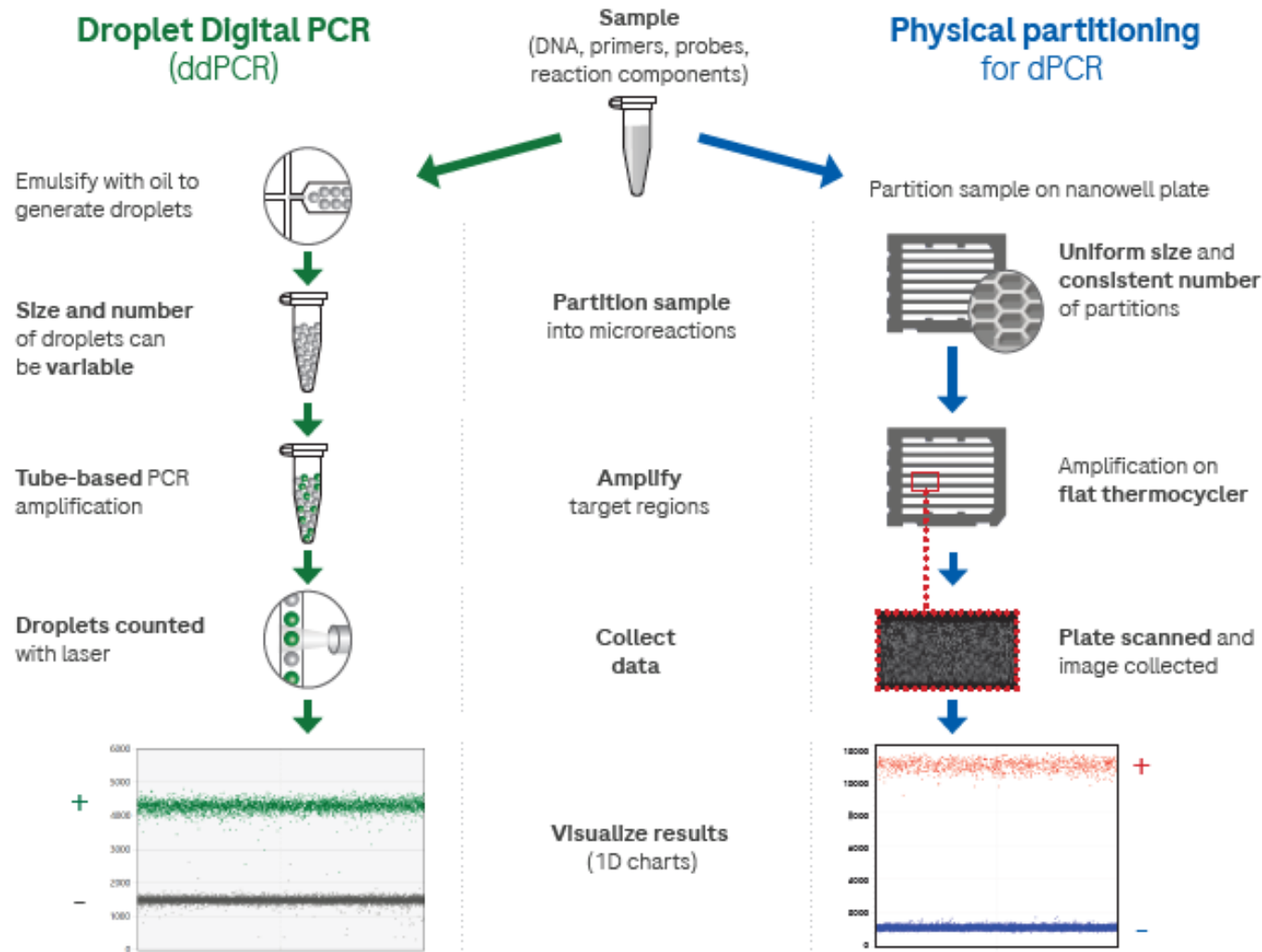
The "**digital PCR**" (dPCR) allows for highly sensitive absolute quantification of **nucleic acids** without the need for standard curves. Its fundamental principle is to partition a reaction mixture containing target nucleic acid molecules into tens of thousands of independent **microreaction units** (such as **microdroplets** or **microwells**), with individual **PCR amplification** taking place within each unit.

Poisson statistical analysis of the number of positive and negative units yields an **absolute quantification** of the target sequence.



Digital PCR (dPCR)

Two types of dPCR



Digital PCR (dPCR)

Contrary to real-time qPCR, dPCR doesn't rely on every amplification cycle to determine the relative quantity of the target molecule; instead, it uses **Poisson statistics** to determine the **absolute quantity** of the target after **end-point amplification**. Since the target molecule is randomly distributed across all available partitions, the **Poisson distribution** estimates the average number of molecules per partition (zero, one, or more) and calculates the copies of the target molecule per positive partition. The **Poisson statistical analysis** of the number of positive and negative reactions provides a precise and **absolute quantification** of the target sequence."

λ is the expectation value for these events and will tell what the most likely average value of copies/partitions is

$$\lambda = -\ln \left(\frac{\text{number of valid partitions} - \text{number of positive partitions}}{\text{number of valid partitions}} \right)$$

Computing the estimated concentration of target genes in copies/ μ l

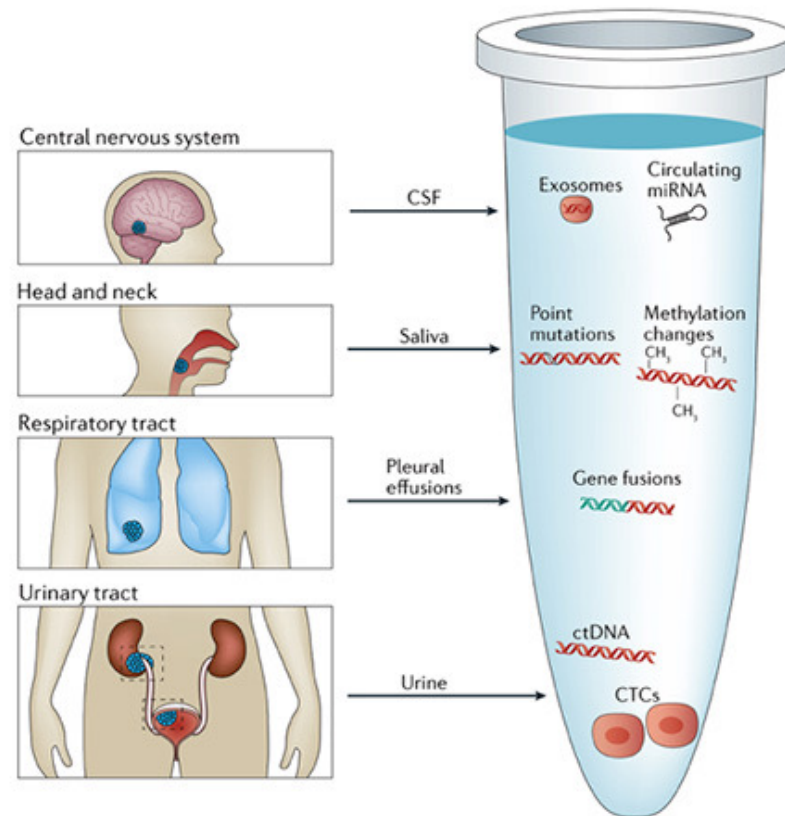
Given:
Valid partitions: 8000
Positive partitions: 4000

$$\lambda = -\ln \left(\frac{\text{number of valid partitions} - \text{number of positive partitions}}{\text{number of valid partitions}} \right)$$
$$\lambda = -\ln \left(\frac{8000 - 4000}{8000} \right) = 0.693 \quad \text{Calculated target copies per partition}$$

Given:
Partition volume is $V = 0.23 \text{ nl}$

$$\lambda_{\text{volume}} = \frac{\lambda}{V [\mu\text{l}]}$$
$$\lambda_{\text{volume}} = \frac{0.693}{0.23} \times 1000 = 3013 \text{ copies}/\mu\text{l} \quad \text{Calculated target copies per } \mu\text{l}$$

Body fluids as a source of molecular information derived from the tumor



Body fluids as a source of molecular information derived from the tumor

Circulating tumor DNA (ctDNA), which is fragmented, cell-free tumor DNA found in blood plasma, is an invaluable source of cancer-specific mutations and represents an excellent resource for developing minimally invasive diagnostic tests. ddPCR-based monitoring is used for the early diagnosis of metastases and for therapy modification.

Mutation diagnosis has been shown to precede the clinical diagnosis of metastasis in 86% of breast cancer patients.

