# Main techniques to detect RNAs and proteins

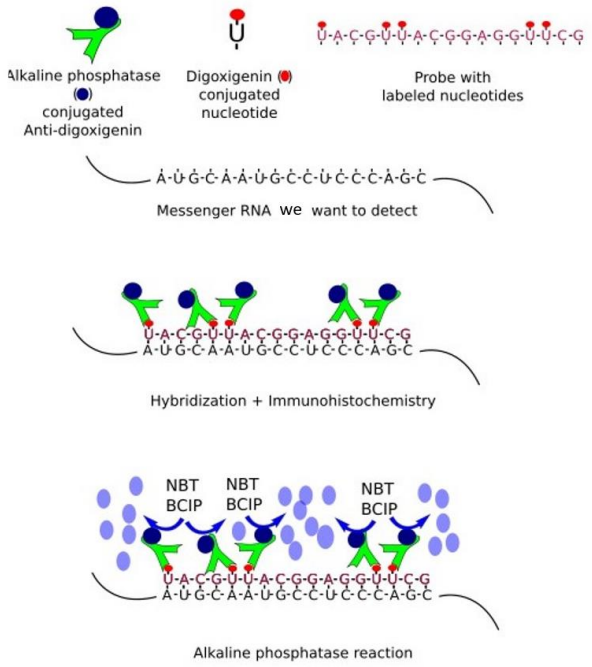
Patrick Pla with the help of Jérôme Artus, Caroline Borday and Sébastien Szuplewski

## *In situ* hybridization

To detect RNA (usually mRNA but we can also detect microRNA), the principle of complementarity of nucleic acid bases is used. We take advantage of the single-stranded nature of RNA and use an antisense RNA as a probe, a sequence complementary to part of the target RNA with modifications allowing its visualization. Length of the antisense RNA molecule can vary between 200 bp to 1000 bp.

First, the probe must be synthesized *in vitro*, by a transcription step from a plasmid containing the coding sequence of the gene of interest. This antisense RNA synthesis takes place in the presence of a UTP (uridine triphosphate) conjugated to digoxygenin, a compound produced by a particular group of plants (*Digitalis*). A labelled sense RNA (sequence identical to part of the target RNA) is also produced and is used as a negative control (it will not recognize itself) and will allow to evaluate the part of the labelling corresponding to background staining.

The embryo is fixed (typically with 4% paraformaldehyde) then permeabilized by detergents (Tween-20 or Triton X100) and proteinases (proteinase K), so that the probe and later antibodies can enter and leave the cells. Once inside the cells, hybridization occurs between the antisense RNA and the targeted RNA. To visualize the cells in which hybridization has occurred, we use an antibody that specifically recognizes digoxygenin. This antibody has been artificially conjugated to an enzyme, such as alkaline phosphatase. After incubation and repeated washes to remove any unbound antibodies, the embryo is incubated in a solution containing a substrate for the enzyme (traditionally an NTB/BCIP mixture for alkaline phosphatase) which yields a coloured product in blue-purple. By incubating the tissues with the substrates for longer time, lower quantities of mRNA can be revealed. This detection method is not finely quantitative, but it can be used to assess the relative level of expression of a gene of interest.

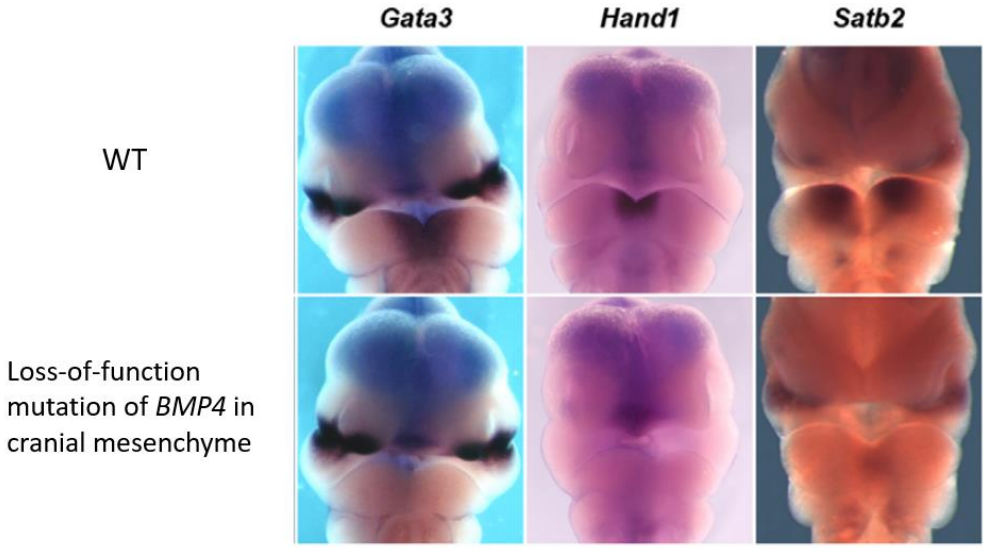



From https://mmegias.webs.uvigo.es/02-english/6-tecnicas/5-hibridacion.php

More modern methods use fluorescent-labelled antibody for the detection step or even directly fluorescent probes complementary to RNAs (RNA-FISH method).

Example of use:

The role of BMP4 on cranial mesenchymal development is investigated by studying the impact of its loss of function on expression of genes in this tissue.

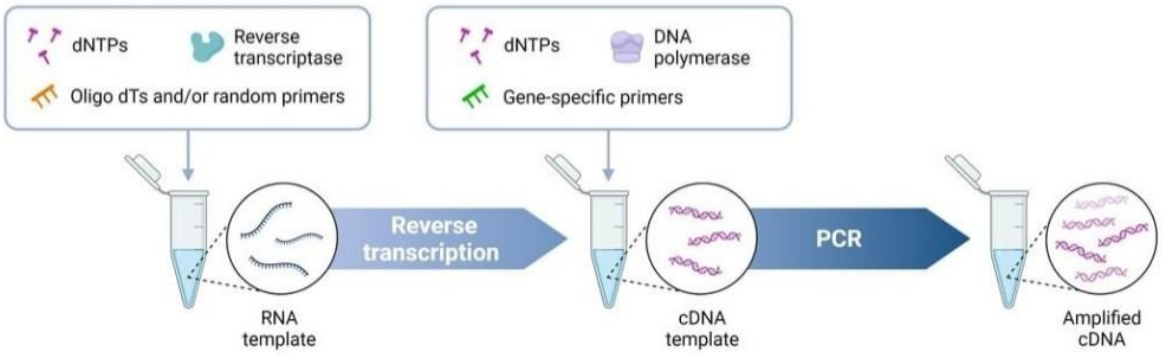



Frontal views of heads of mouse embryos (WT or mutated) treated in *in situ* hybridization with probes recognizing *Gata3*, *Hand1* or *Satb2* mRNAs. Note that the deficit of BMP4 causes only a partial and localized loss of expression of these genes. Adapted from Bonilla-Claudio *et al*. 2012

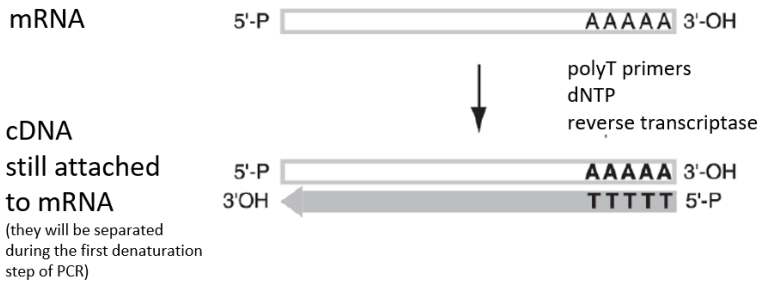
## RT-PCR and RT-qPCR

RNAs are extracted from tissue or cells and are reverse-transcribed into complementary DNA (cDNA) using a reverse transcriptase. Oligo dTs (a succession of T nucleotides) are used as primers to reverse-transcribe only mRNAs that have a polyA tail. This can create a bias towards an overrepresentation of the 3' parts of the genes. Alternatively, we can use primers with random sequences that will allow all RNAs to be reverse-transcribed, starting at random positions, but this will also include rRNAs that are very abundant but not always relevant for transcriptomic analysis.

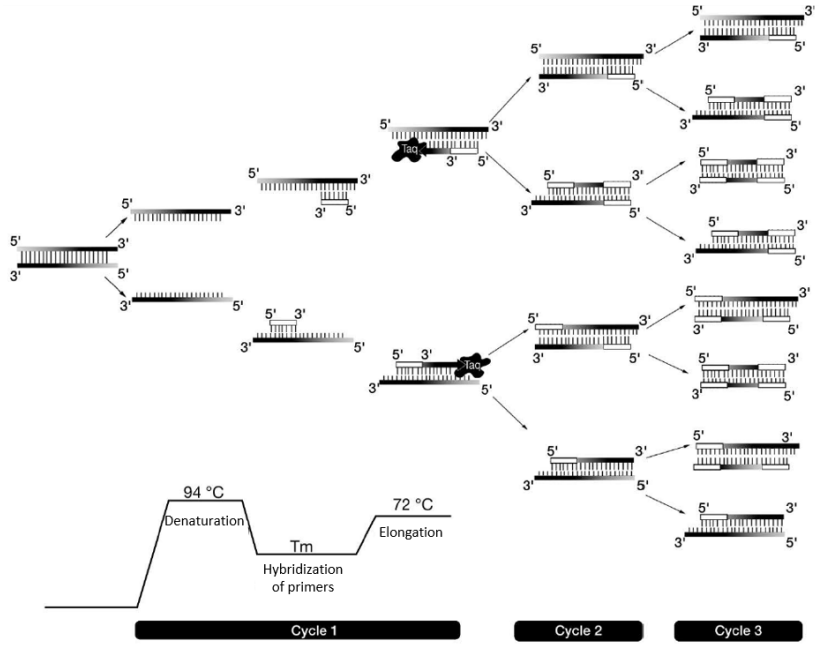
Then, specific fragments of the cDNA produced by reverse-transcription are amplified by PCR using pairs of specific oligonucleotide primers (around 20 nucleotides long) designed from the coding sequence of the genes of interest. These primers are necessary for the initiation of DNA synthesis by the DNA polymerase and act as guides to ensure that only the sequences of interest will be amplified and therefore visualized.




RT-PCR overview



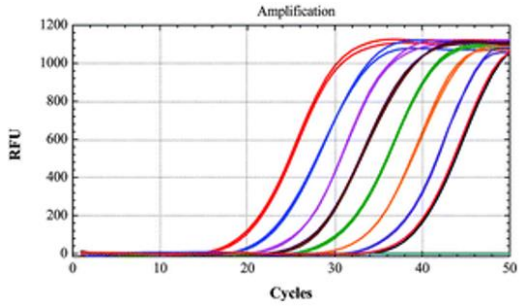

Reverse transcription step. Adapted from Denis Tagu, Stéphanie Jaubert-Possamai, Agnès Méreau, *Principes des techniques de biologie moléculaire et génomique, Éditions Quæ (2018)*




Three first cycles of a PCR. Adapted from Denis Tagu, Stéphanie Jaubert-Possamai, Agnès Méreau, *Principes des techniques de biologie moléculaire et génomique, Éditions Quæ (2018)*

The PCR can be classic: we observe the relative quantity of DNA obtained at the end of a given number of amplification cycles with a migration on an electrophoresis gel.

Alternatively, the PCR can be quantitative (qPCR or also real-time PCR): we observe the quantity of DNA after each cycle of amplification using a molecule which becomes fluorescent in the presence of double-stranded DNA (SYBRGreen). Other more complicated systems like Taqman can be used but the general principle is similar.



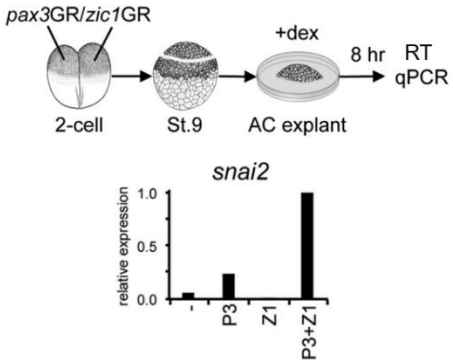

The different colors correspond to different samples with different quantity of mRNA. A threshold is defined (for example RFU=200) and the Cq is obtained for each of the samples, which is the number of PCR cycles necessary to obtain fluorescence that reaches this threshold. Theoretically if a sample A corresponds to cells having twice mRNA quantity comparatively to a sample B, the Cq of A must be smaller by 1 than the Cq of B (logarithms of 2 are used).

The qPCR method is quantitative because we can monitor the quantity of DNA that doubles at each PCR cycle for each pair of oligonucleotides used. This is not possible in classic PCR where the result is only observed after 30 to 40 cycles depending on the protocols.

In all cases, expressions of one or more housekeeping genes (genes whose expression is not expected to change between different conditions) have to be studied and used as a standardization reference (compensating for the small differences in RNA quantities or reverse-transcription efficiency between the samples).

Example of use:

The animal cap, which is the area around the animal pole of the *Xenopus* blastula, is destined to form the epidermis during normal development. However, these cells retain pluripotentiality and upon exposure to specific inducers, can differentiate into other tissues. Here, the potential of Pax3 and/or Zic1 to induce neural crest development is tested.




Increased quantity of *Snail2* (or *Snai2*) mRNA, a marker of neural crests, induced by the co-expression of Pax3 and Zic1 in the animal cap of *Xenopus* embryos. mRNAs encoding forms that can be induced by dexamethasone of the transcription factors Pax3 and Zic1 are injected (or not) into 2-cell stage *Xenopus* embryos. At the old blastula stage, the animal cap of these embryos is excised. Without intervention, the animal caps develop into epidermis. Dexamethasone is added and 8 h later, the RNAs are extracted and the expression of *Snai2* is studied by RT-qPCR. It is the combination of Pax3 and Zic1 which allows an optimal induction of the neural crest marker (Pax3 alone can induce *Snai2* expression but weakly). From: https://www.sciencedirect.com/science/article/abs/pii/S0006291X17325275

## RNAseq

*In situ* hybridization or RT-PCR/RTqPCR to detect mRNAs can be used to study a limited number of genes. New techniques, much more expensive, have allowed to detect the expression of several hundred or even several thousand genes simultaneously.

RNAseq takes advantage of the great progress in sequencing with the advent of NGS (Next Generation Sequencing or high-throughput sequencing) and can quantify most of the RNAs present in a tissue. Specifically, RNAs are extracted from the samples. They are fragmented and converted into complementary DNA (cDNA) with standard procedures using reverse transcriptase. Known adapter sequences are added to the ends. These adapters allow immobilization and PCR amplification. NGS can analyse these transcripts and quantify them after alignment of the obtained sequences on the genome. This is a global and unbiased approach (as you do not choose the genes whose expression you study).

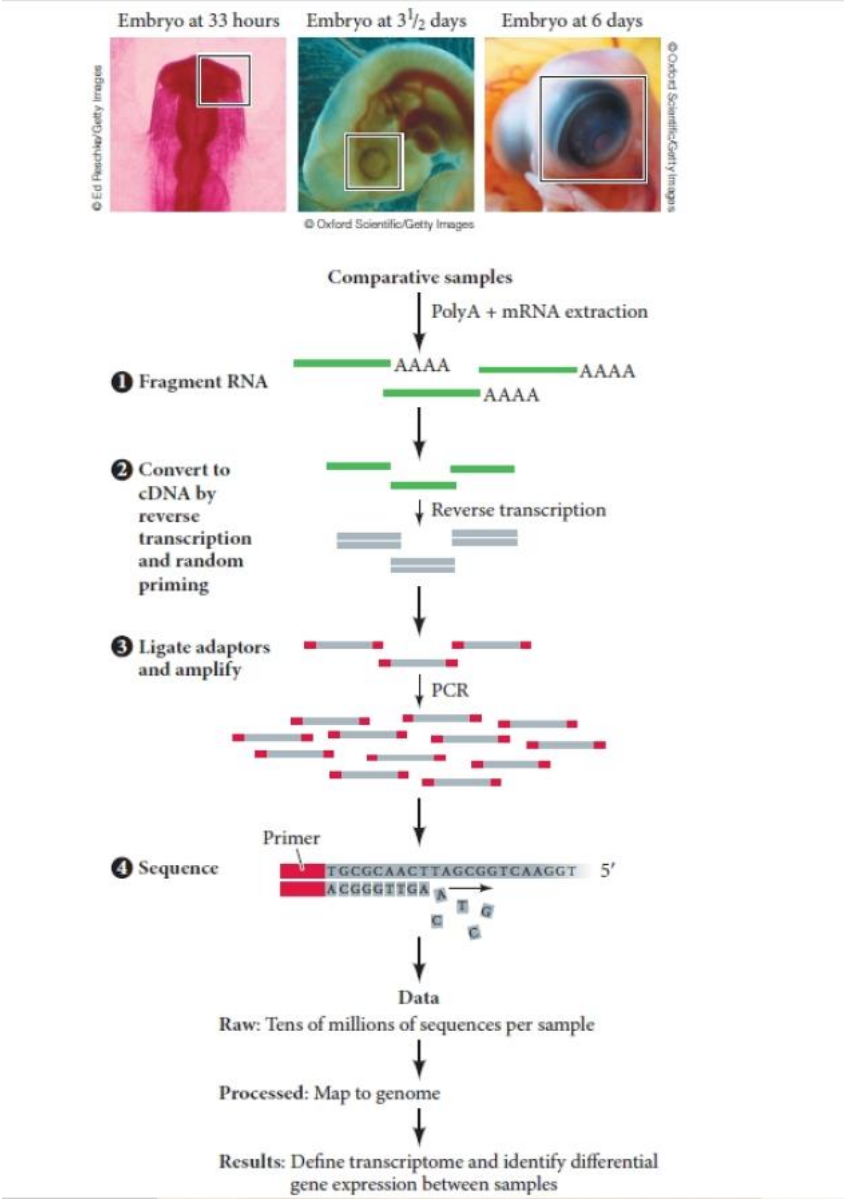


**FIGURE 3.33** Deep sequencing: RNA-Seq. (Top) Researchers begin with specific sorts of tissues, often comparing different conditions, such as embryos of different ages (chick embryos, as shown here), isolated tissues (such as the eye; boxed regions) or even single cells, and samples from different genotypes or experimental manipulations. (1) RNA is isolated to obtain only those genes that are actively expressed. (2) These transcripts are then fragmented into smaller stretches and used to create cDNA with reverse transcriptase. (3) Specialized adaptors are ligated to the cDNA ends to enable PCR amplification and immobilization for (4) subsequent sequencing. (After J. H. Malone and B. Oliver. 2011.

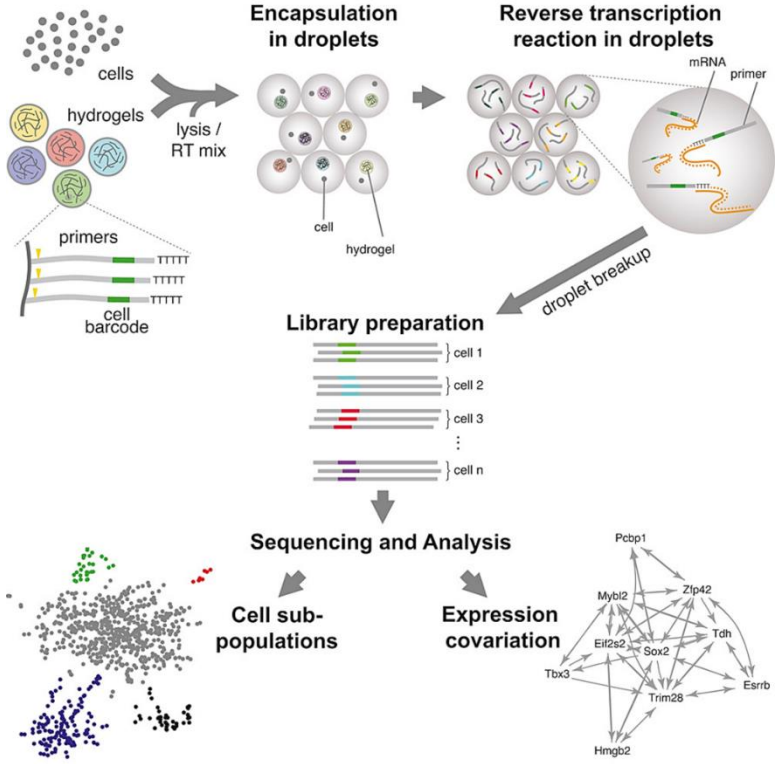
From Barresi, Gilbert, *Developmental Biology*, 12th edition, Sinauer Associates (2020)

RNA-seq is particularly powerful for comparing transcriptomes between samples differing only in certain experimental parameters (addition or not of a morphogen, for example). After statistical analysis, a list of sDEGs for Significantly Differentially Expressed Genes is obtained.

The development of fluorescence-activated cell sorting (FACS) and microdissection have allowed the precise isolation of tissues and cells from which RNAs are extracted, somewhat compensating for the lack of spatial resolution of this approach.

Recent advances in RNAseq sensitivity as well as microfluidics methods now make it possible to perform transcriptomic analysis of single cells (single cell RNAseq or scRNAseq). The term "isolated cells" would probably be more appropriate, but the term "single cells" has become commonplace.

See this video: https://www.cell.com/cms/10.1016/j.cell.2015.04.044/attachment/4318ad70-6cce-497a-8135-1a04319032bf/mmc7.mp4




Cells are encapsulated in droplets with lysis buffer, reverse transcription mix, and hydrogel microspheres bearing barcoded primers. After encapsulation, the primers are released. The cDNA in each droplet is labelled with a barcode during reverse transcription. The droplets are then broken up and material from all cells is linearly amplified before sequencing. All the sequences that will have the same barcode come from the same cell and therefore we can have access to the transcriptome of each cell. From: https://www.cell.com/cell/fulltext/S0092-8674(15)00500-0

Before the development of scRNAseq, we obtained only the "averaged" mRNA quantity from whole organs or tissue samples. scRNAseq allows the definition and characterization of individual cells at high resolution. This type of study has generally made it possible to unravel an unsuspected diversity of transcriptomic profiles within a given tissue that was thought to be homogeneous and to know more precisely the transition stages over time of a cell population. However, this technique has limitations because only the most abundant fraction of transcripts can be analyzed, leaving aside genes that are poorly transcribed or whose transcripts are quickly degraded but some of which can nevertheless have a significant impact. Also, spatial resolution is lost compared to *in situ* hybridization.

## Reporter genes

Studying the expression of a specific gene is sometimes complex whereas some genes produce proteins whose presence and activity is easily detectable. These genes are called reporter genes. The products of reporter genes must not be present naturally in the cells/tissues of interest, and they must not interfere with cell physiology to avoid disturbing the studied process. Reporter genes can encode:

- enzymes whose product is easily detectable: light for luciferase, blue/red staining for β-galactosidase (depending on Xgal/Redgal substrate, respectively)
- fluorescent proteins such as GFP and its derivatives.

Fluorescent proteins have the advantage to allow observations in living embryos, illuminated with the correct wavelength, whereas β-galactosidase staining usually requires fixation. Recently, luciferase activity can also be observed on living tissues and no longer only in protein extracts.

Coupled with transgenesis, reporter genes are powerful tools to study promoters and enhancers since reporter genes can be expressed in embryos under the control of these elements. The adequacy of the expression pattern of the reporter gene with the expression of the endogenous gene whose transcriptional regulatory elements are studied makes it possible to know whether the chosen fragments of promoters or enhancers are relevant or not.

Example of use:

NRSE, the neuron-restrictive silencer element, is a regulatory sequence that is present in several neuronal genes. Its role in restricting the expression of the neuron-specific *L1* gene is illustrated here.

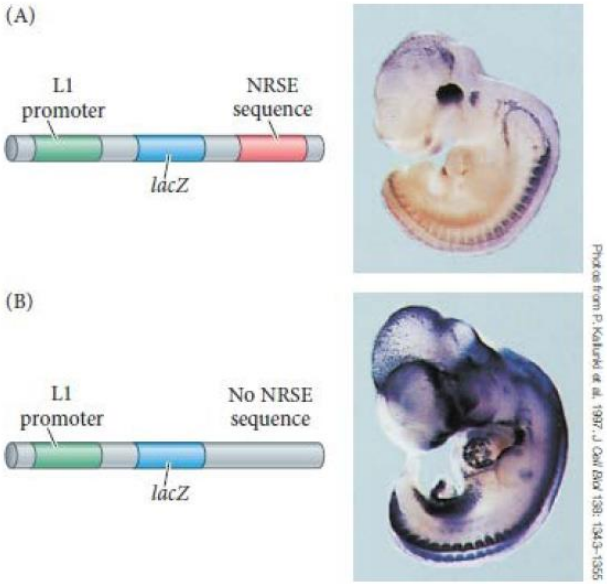



**FIGURE 3.9** A silencer represses gene transcription. (A) Mouse embryo containing a transgene composed of the L1 promoter, a portion of the neuron-specific *L1* gene, and a *lacZ* gene fused to the *L1* second exon, which contains the NRSE sequence. (B) Same-stage embryo with a similar transgene but lacking the NRSE sequence. Dark areas reveal the presence of β-galactosidase (the *lacZ* product).

From Barresi, Gilbert, *Developmental Biology*, 12th edition, Sinauer Associates (2020)

## Western blot

This is a method that separates proteins according to their size in a SDS-polyacrylamide gel submitted to electrophoresis (SDS-PAGE) and then transfers them to a membrane where specific proteins are identified using antibodies. The secondary/tertiary structures of the proteins and their charges are not relevant in western blot as they are incubated in DTT which disrupts disulfide bridges and in SDS which denatures and covers proteins with negative charges.

The antibodies used (for western-blot or other techniques using antibodies) can be of two types:

- a polyclonal antibody (in fact, a mixture of different antibodies) which are produced by an animal (rabbit, mouse, goat, etc.) after injection of the antigen of interest in its blood.
- A monoclonal antibody against a given antigen produced in hybridomas, that is to say B lymphocytes made "immortal" after fusion with cancerous myeloma cells.

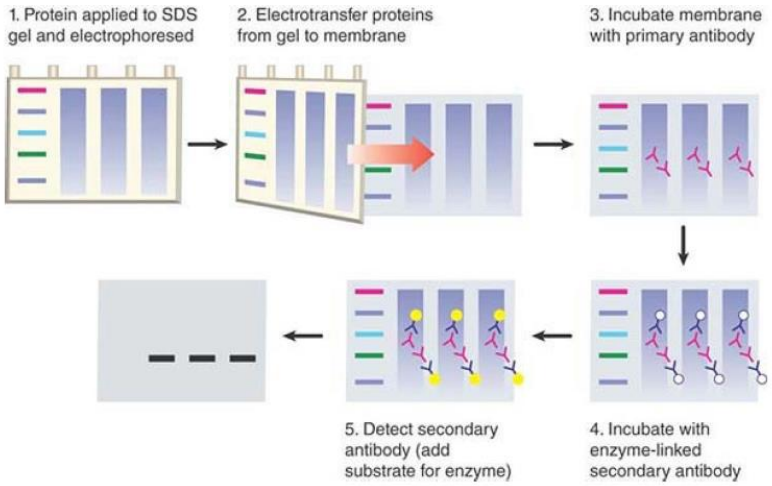


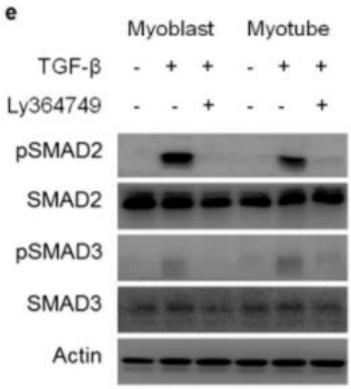

**FIGURE 2.23** In a western blot, proteins are separated by size on an SDS gel, transferred to a nitrocellulose membrane, and detected by using an antibody. The primary antibody detects the protein and the enzyme-linked secondary antibody detects the primary antibody. The secondary antibody is detected in this example via addition of a chemiluminescent substrate, which results in emission of light that can be detected on X-ray film.

From *Lewin's Gene XII*, Jones & Bartlett Learning (2018)

Western blot allows to compare the abundance of a particular protein in cells or tissues treated in different conditions. It can also compare the phosphorylation status of a particular protein in different conditions. In this case, the variations of the signal with an antibody targeting the phosphorylated protein should be compared with the variations of the signal with an antibody targeting all forms of this protein (phosphorylated or not), to distinguish between the variations of the phosphorylation and the variations of the general level of the protein. Other post-translational modifications can also be detected with this method.

Example of use:

Impacts of TGFβ signalling activation or inhibition in myoblasts or myotubes on phosphorylation, and then activity, of SMAD 2 and SMAD 3 are investigated.




Myoblasts or myotubes in culture are treated or not during one hour with the ligand TGFβ and in some cases Ly364947, an inhibitor of TGFβ receptors. Proteins of these cells were extracted and submitted to western blot with antibodies recognizing either a phosphorylated form of SMAD2 or SMAD3 (pSMAD2 or pSMAD3, respectively), all forms of SMAD2 or SMAD3 and actin. From https://www.mdpi.com/2073-4409/9/2/375/pdf

## Immunochemistry and immunofluorescence

Immunochemistry and immunofluorescence allow the localization of proteins of interest in tissues or cells, contrasting with western blot where spatial resolution is inexistent (or not precise if using dissected tissues or nuclear/cytoplasmic protein extracts as a starting material).

Immunochemistry is based on the revelation of an enzymatic activity associated to an antibody, whereas immunofluorescence is based on the emission of photons at a given wavelength from a fluorophore coupled to an antibody.

The use of one or two antibodies distinguishes direct or indirect immunodetection (shown below for immunofluorescence but it is the same for immunochemistry).

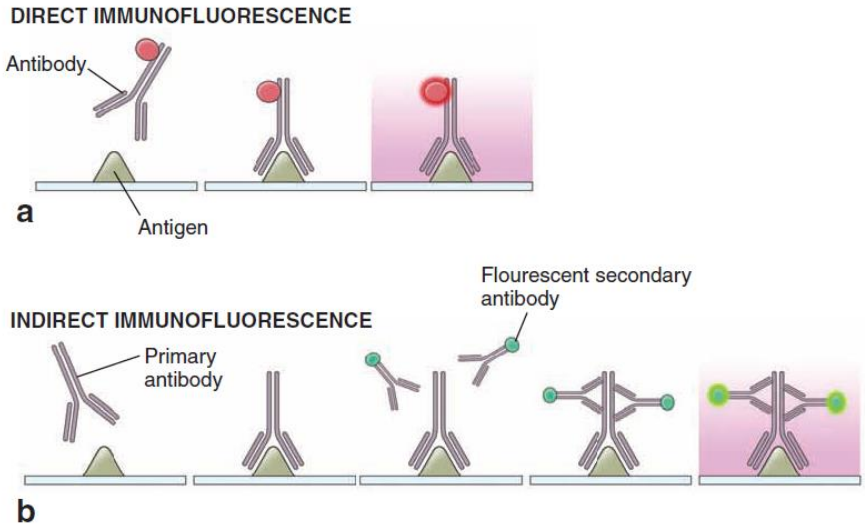


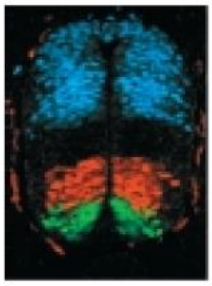

**FIGURE 1.5 ▲ Direct and indirect immunofluorescence. a.** In direct immunofluorescence, a fluorochrome-labeled primary antibody reacts with a specific antigen within the tissue sample. Labeled structures are then observed in the fluorescence microscope in which an excitation wavelength (usually ultraviolet light) triggers the emission of another wavelength. The length of this wavelength depends on the nature of the fluorochrome used for antibody labeling. **b.** The indirect method involves two processes. First, the specific primary antibodies react with the antigen of interest. Second, the secondary antibodies, which are fluorochrome labeled, react with the primary antibodies. The visualization of labeled structures within the tissue is the same in both methods and requires the fluorescence microscope.

From Wojciech Pawlina, *Histology : a text and an atlas 7th edition,* Wolters Kluwer ed. (2016)

Several indirect immunofluorescence detections can be made on the same sample provided that the origin of the antibodies (no common species for the production of the various primary antibodies) and the wavelengths of the fluorophores of the secondary antibodies are carefully chosen.

Example of use

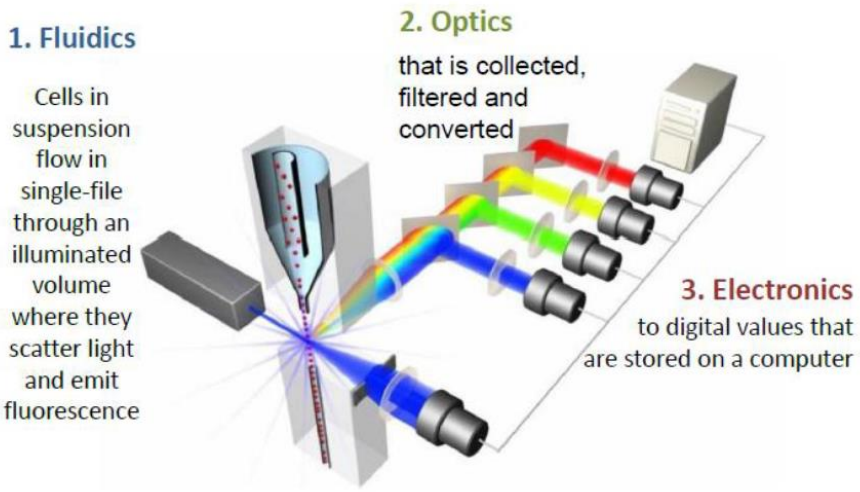
The dorso-ventral regionalisation of the neural tube is highlighted here using immunofluorescence to identify different cell types.



Transverse section of the neural tube of a chicken embryo treated in immunofluorescence with anti-Pax7 (revealed in blue), anti-Oligo2 (revealed in red) and anti-Nkx2.2 (revealed in green) antibodies. The dorsal region is on the top. From E. Dessaud et al. 2007. *Nature* 450: 717–720

## Flow cytometry and FACS (Fluorescent Activated Cell Sorting)

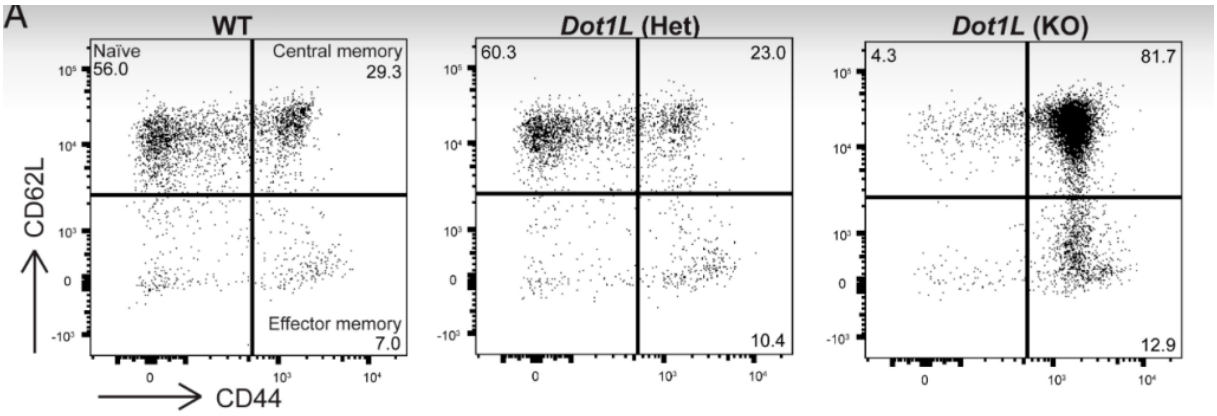
This technique consists in marking one or more cellular elements of interest with a fluorochrome (fluorochrome-coupled antibody or protein fusion with GFP or reporter gene based on the production of a fluorescent protein) then separating the cells to pass them one by one under a laser at the desired wavelength of excitation. The produced fluorescence of each cell is measured. If needed, cells can be sorted according to their fluorescence level, the flow being able to be automatically diverted towards different tubes according to the results obtained.




Principle of flow cytometry. From https://www.iitk.ac.in/che/pdf/resources/Flow-Cytometry-reading-material.pdf

Example of use:

The role of DOT1 in the establishment of different CD8 T lymphocyte subpopulations in the mouse spleen was studied by FACS.




Analysis of CD8+ T cell subgroups in the spleen by flow cytometry in WT mice, heterozygous mice for a loss-of-function of *Dot1L* gene (Het) and homozygous mice for this loss-of-function (KO). Each point represents the abundance pattern of CD62L (vertical axis) and CD44 (horizontal axis), which are markers of CD8+ T cells subgroups. Thresholds are empirically chosen to assess if a cell is positive or not for a particular marker (horizontal and vertical lines in the middle of the graph). The experiment reveals that *Dot1L-/-* (KO) mice show a significant loss of naïve CD8+ T cells (CD44-/CD62L+) and a massive gain in CD44+/CD62L+ cells, which are central memory T cells (MCTs). *Dot1L+/-* (Het) mice have a phenotype close to wild-type (WT) mice. Source: https://www.pnas.org/doi/10.1073/pnas.1920372117

| Technique | Which molecule is detected? | Which molecule for specificity? | Which methods of detection? | Spatial resolution? | Quantification? | Remarks |
|---|---|---|---|---|---|---|
| *In situ* hybridization | RNA (usually mRNA) | Antisense RNA | Colorimetric revelation or fluorescence | Yes | Relative for the same probe in various regions of a sample, or in various comparable samples during the same experiment | On fixed whole embryos or sections.<br>Note that the procedure often includes the use of antibodies (immunochemistry or immunofluorescence step) but **not** to detect proteins |
| RT-PCR | RNA (usually mRNA) | Primers for the DNA polymerase | Fluorescence on an electrophoretic gel loaded with BET or SyBR Green | No<br>(or partially if RNA is extracted from various tissues dissected from embryo) | Relative for the same primers between conditions<br>(with expression of housekeeping genes as internal controls) | Less used nowadays.<br>Sensitive.<br>The biological meaning of low expressions should be tested. |
| RT-qPCR | RNA (usually mRNA) | Primers for the DNA polymerase | Fluorescence in a qPCR machine | No<br>(or partially if RNA is extracted from various tissues dissected from embryo) | Relative for the same primers between conditions<br>(with expression of housekeeping genes as internal controls) and also quantitative (if a standard curve is done) | Very sensitive.<br>The biological meaning of low expressions should be tested. |
| RNAseq | RNA (usually mRNA) | No particular specificity as it is a global, unbiased approach | Fluorescence during sequencing | No<br>(or partially if RNA is extracted from various tissues dissected from embryo) | Quantitative | Can be done at a single cell level (scRNAseq) |
| Reporter gene | Protein | Reporter gene expression is controlled by specific regulatory sequences of the gene of interest | Light (luciferase) or colorimetric revelation (β-galactosidase) or fluorescence (GFP) | Yes | Relative for the same protein in various regions of a sample or in various comparable samples during the same experiment | Fluorescent proteins can be observed in live tissues/cells<br>whereas β-gal activity is generally assessed on fixed samples |

| | | | | | | |
|---|---|---|---|---|---|---|
| Western blot | Protein | Antibody (monoclonal or polyclonal) | Light or fluorescence | No (or partially if proteins are extracted from various tissues dissected from embryo) | Relative for the same antibody between conditions (with quantity of housekeeping proteins as internal controls) | |
| Immunochemistry/ immunofluorescence | Protein | Antibody (monoclonal or polyclonal) | Colorimetric revelation or fluorescence | Yes | Relative for the same antibody in various regions of a sample, or in various comparable samples during the same experiment | On fixed whole embryos or sections or cell cultures |
| Flow cytometry/FACS | Protein | Antibody or fluorescent protein (e.g. GFP) fused to a specific protein or - reporter gene encoding a fluorescent protein | Fluorescence | No (or partially with various dissected tissues from an embryo) | Relative for the same antibody in various cells of a sample | Can be done using live cells and in this case, it can be used to sort the cells in order to analyse them further or cultivate them separately. |