



**POLYTECH<sup>®</sup>**  
MONTPELLIER

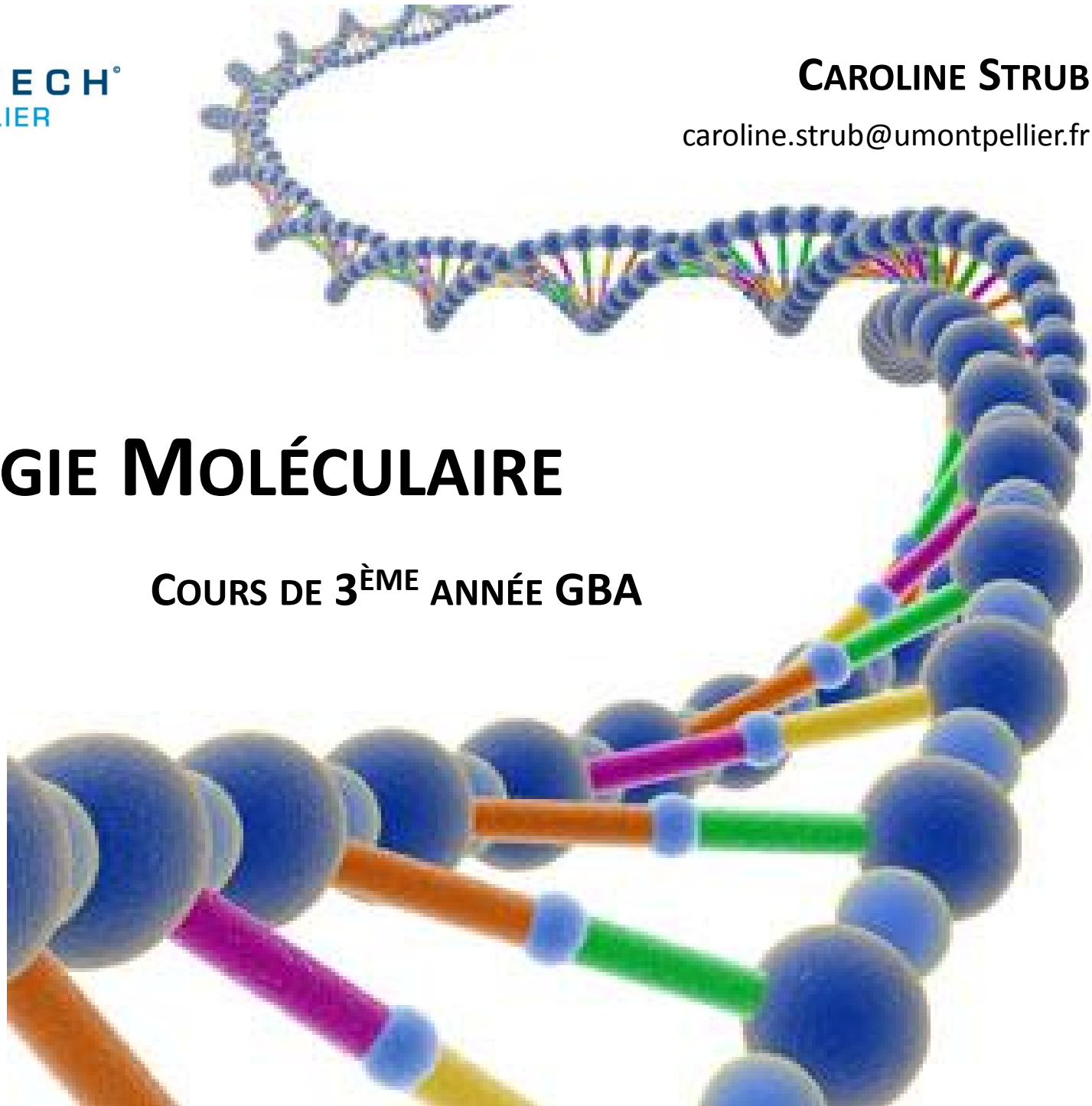
**CAROLINE STRUB**

caroline.strub@umontpellier.fr

# **BIOLOGIE MOLÉCULAIRE**

**COURS DE 3<sup>ÈME</sup> ANNÉE GBA**

**2016 - 2017**



“ 6 x 1,5h de cours

“ 1,5h de TD

“ Etude d'un article scientifique par groupe de 4

= > évaluation : Poster (Format A4) + oral (15-20 minutes ) en Janvier 2017

= > plus d'info pendant les TP

# SOMMAIRE

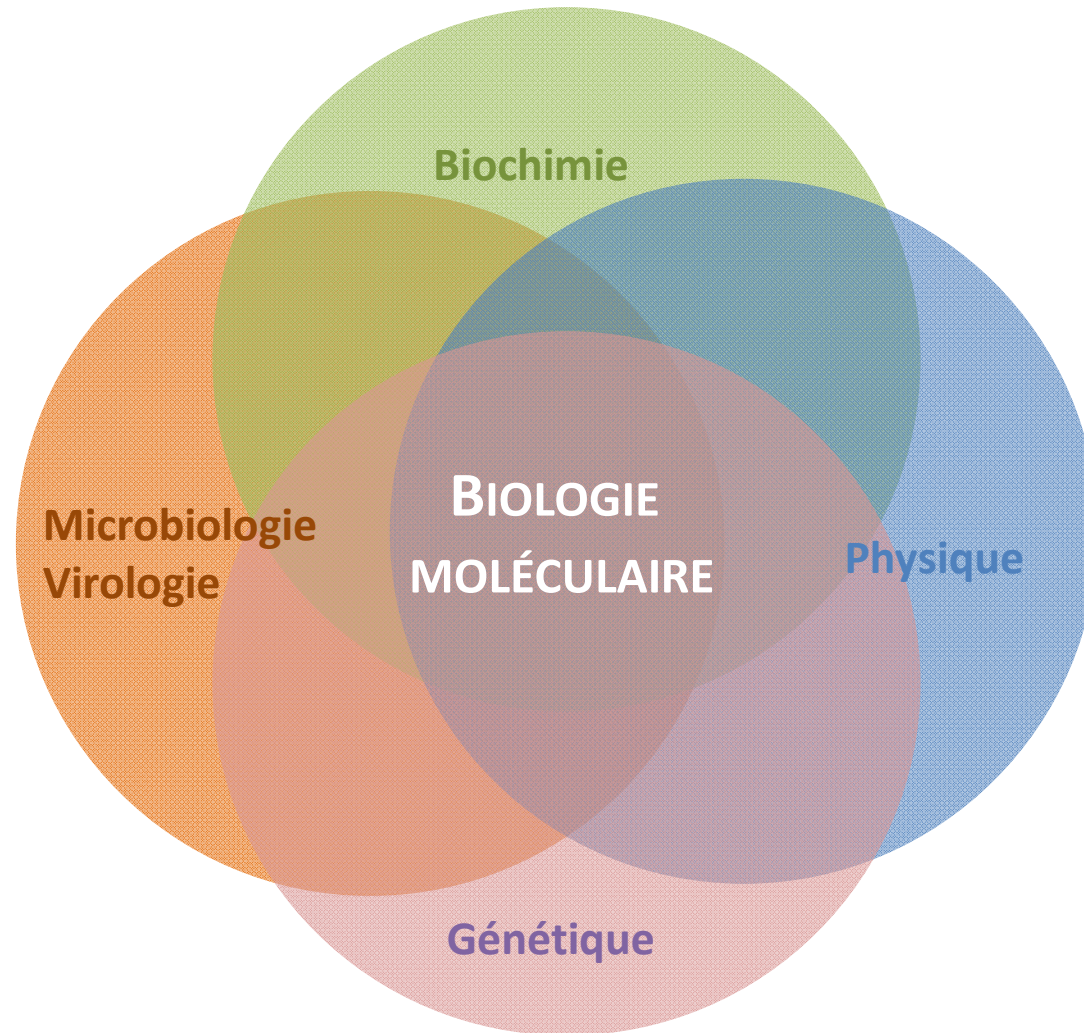
## Biotechnologies

Techniques de génie génétique (ER, Clonage, PCR, amplifications isothermales...)

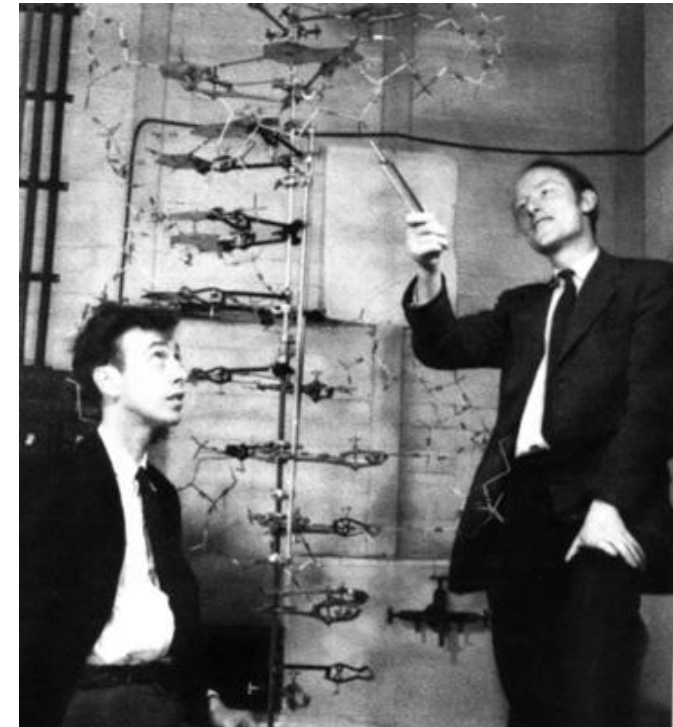
Techniques de séquençage, méthodes de séquençage nouvelle génération (NGS).

Introduction à la biologie synthétique.

# DÉFINITION ET HISTORIQUE



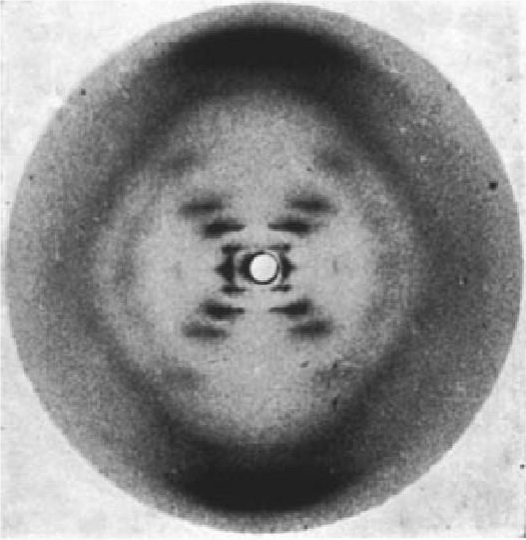
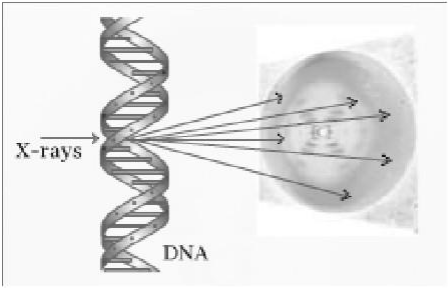
1953 : Structure de l'ADN  
(Crick et Watson)



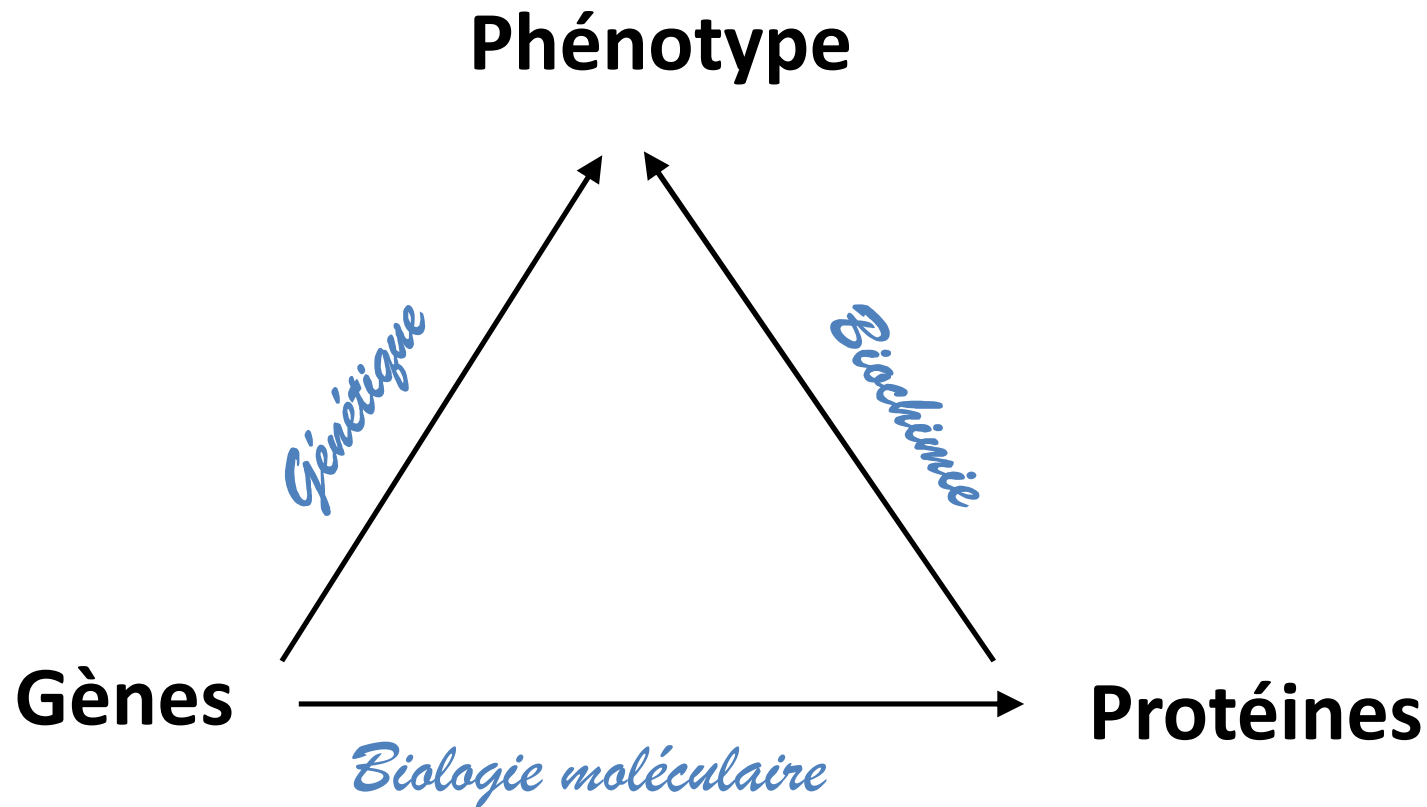
Prix Nobel 1962 (Watson, Crick et Wilkins)  
Structure moléculaire de l'ADN



Rosalind Franklin,  
The dark lady of DNA



Cristallographie au rayon X sur  
« fibre de thymus de veau »

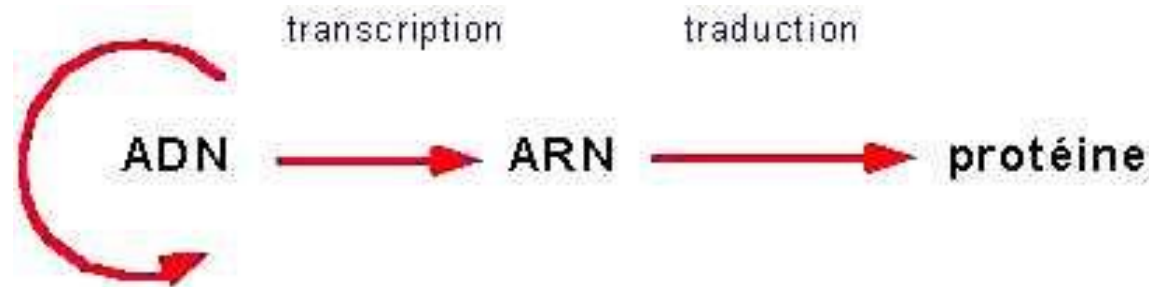


❑ Etude des mécanismes de fonctionnement de la cellule à l'échelle moléculaire

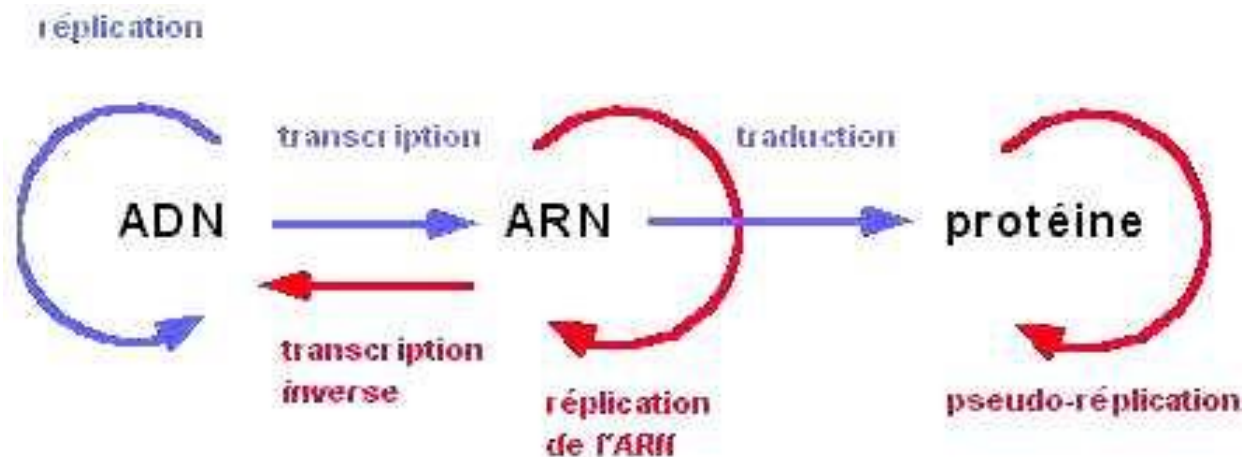


❑ Ensemble des techniques s'appuyant sur la manipulation d'acides nucléiques (ADN, ARN), appelées aussi techniques de génie génétique, biotechnologies.

- ❑ Le dogme central de la biologie moléculaire (W. Crick 1956)



- ❑ Le dogme central de la biologie moléculaire (années 2000)



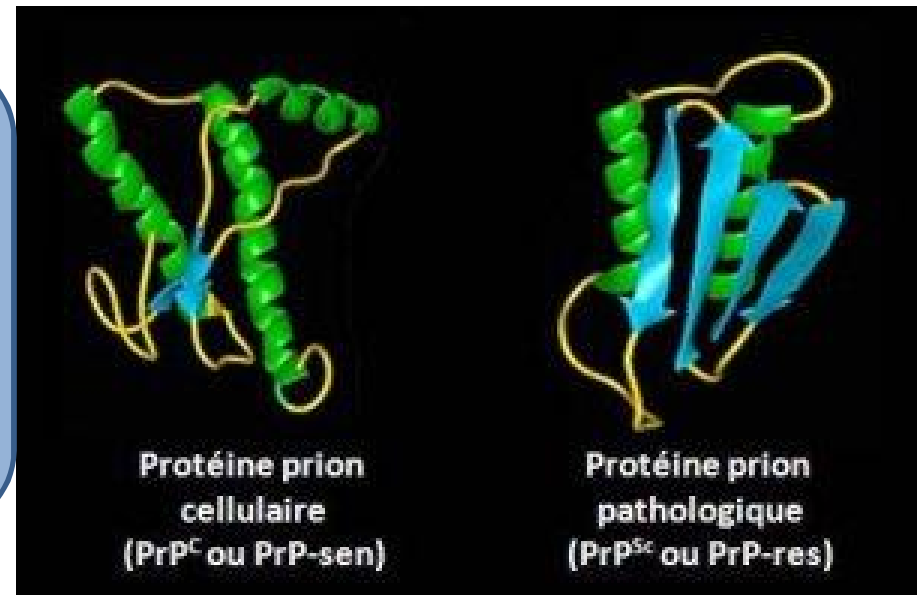
Les différences par rapport au dogme des années 60 :

Les rétrovirus (HIV), les virus à ARN et les prions

- ❑ Les années 2010 : Isogène, promotérome, épigénome...

- ❑ **Protéine infectieuse / PRoteinaceous Infectious particle**  
Découverte en 1982 par S. Prusiner (Prix Nobel 1997)
- ❑ Agent pathogène non conventionnel => Dégénérescence du système nerveux centra  
Pas d'acide nucléique, pas de réponse immunitaire

Le prion dérive d'une protéine endogène :  
structure I<sup>R</sup> identique  
structures II<sup>R</sup>, III<sup>R</sup> et IV<sup>R</sup> différentes.



- ❑ Propagation du prion

Conversion de la protéine « sauvage »  
en protéine « prion »  
Polymérisation de type amyloïde

=

Processus **auto-catalytique**  
(cas sporadique ou contamination)

## Les prions

6

❑ Chez les mammifères : maladies neurodégénératives :

**Encéphalopathies spongiformes transmissibles** (Aspect spongieux du cerveau)

❑ Les EST se définissent par l'accumulation dans le système nerveux central, de particules infectieuses de nature protéique (prions) :

Chez l'homme : le kuru (Tribu Papou 1957) Prix Nobel D. Gajdusek 1976, La Maladie de Creutzfeldt-Jakob (hormone de croissance...)

Chez d'autres mammifères : La scrapie chez le mouton, **l'ESB chez les bovins**

❖ L'Encéphalopathie Spongiforme Bovine

1996 : La crise sanitaire de **la maladie de la vache folle**

(Royaume Uni)

190000 animaux diagnostiqués



Bétail nourrit aux farines animales

- ❑ Apparition d'une nouvelle forme d'EST en GB à la fin des années 90 proche de la maladie de Creutzfeldt-Jakob sur jeunes adultes

➔ Transmission inter-espèce de l'ESB ?

214 victimes humaines (jusqu'à 2009)

Temps d'incubation de l'EST ?

- ❑ BILAN de cette crise sanitaire majeure :

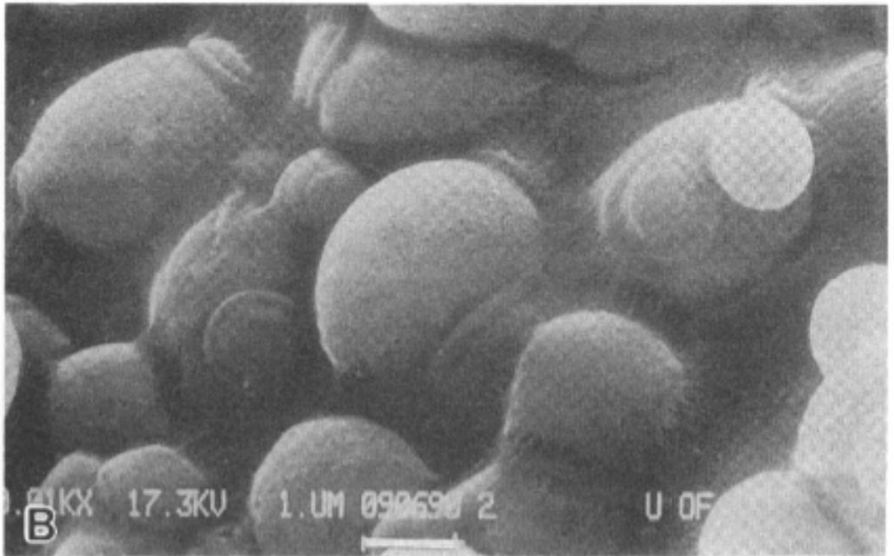
- “ Effondrement de la consommation de la viande de bœuf
- “ Interdiction des farines animales dans l'alimentation du bétail
- “ Développement du principe de précaution

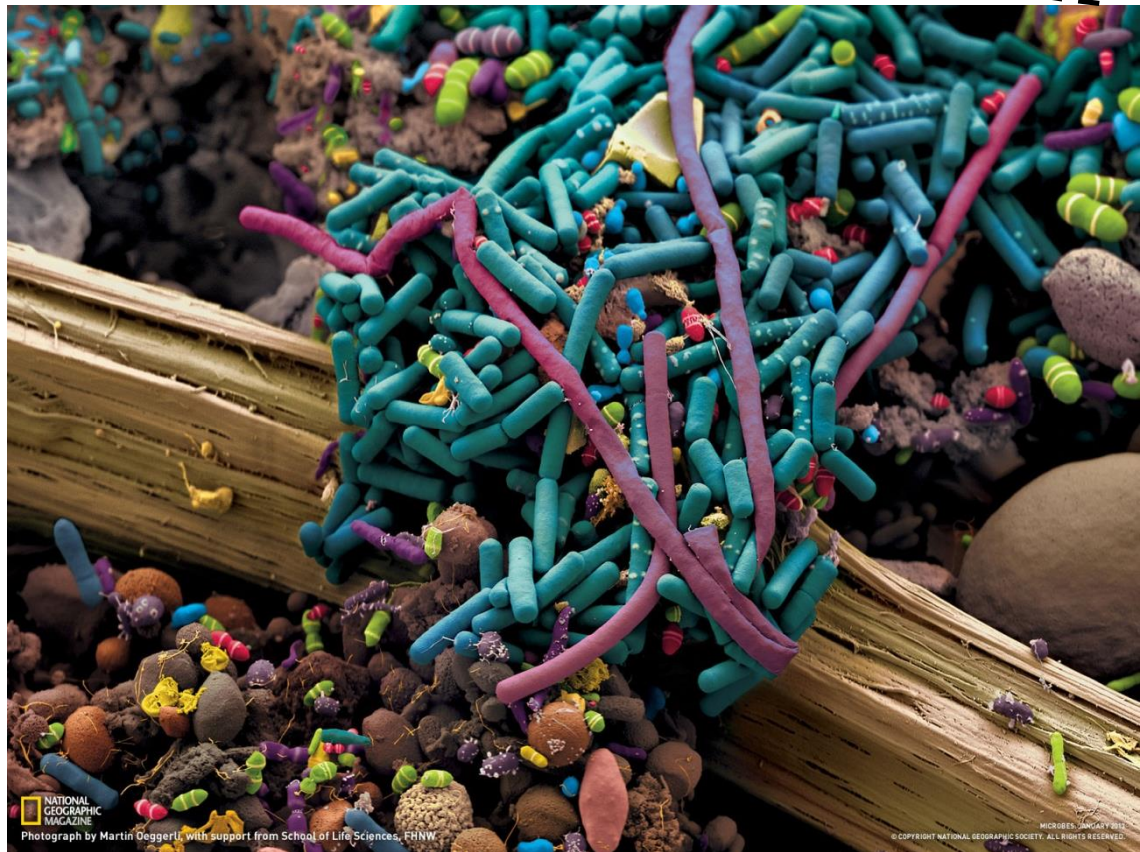
















# SOMMAIRE

## Biotechnologies

Techniques de génie génétique (ER, Clonage, PCR, amplifications isothermales...)

Techniques de séquençage, méthodes de séquençage nouvelle génération (NGS).



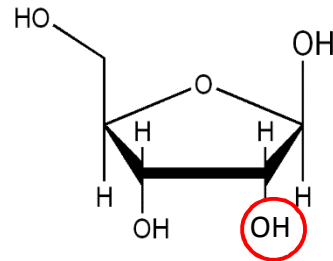
# I - MATÉRIEL GÉNÉTIQUE

## 1.1 Structure moléculaire des acides nucléiques

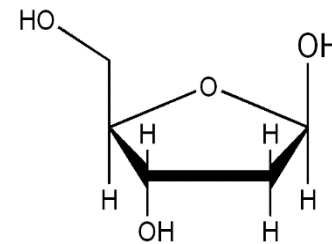
ARN = acide ribonucléique

ADN = acide désoxyribonucléique

PENTOSES



Ribose

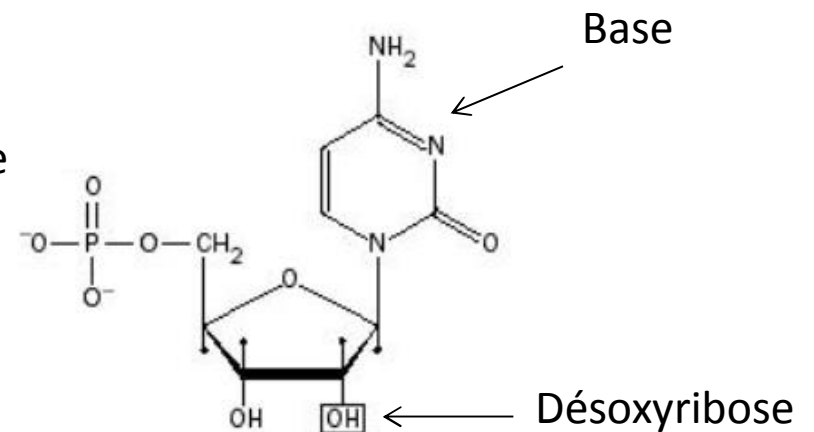


Désoxyribose

Les acides nucléiques sont constitués de chaînes linéaires de **nucléotides** reliés par des ponts phosphodiesters

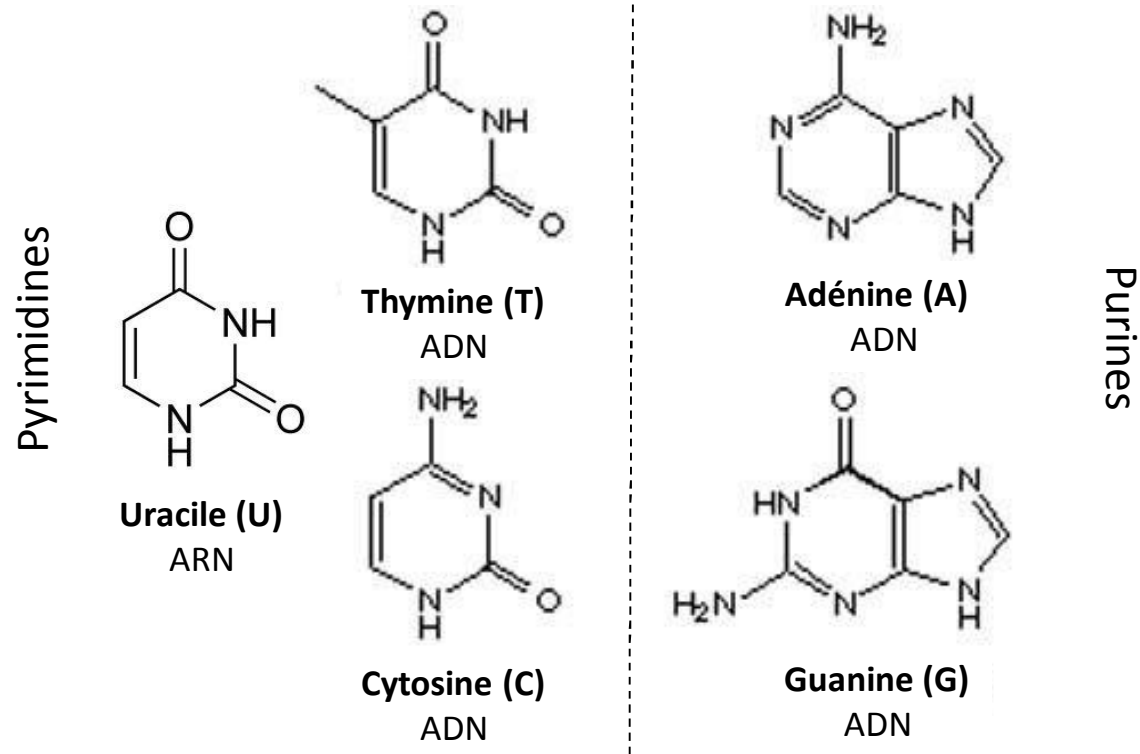
Structure d'un **NUCLÉOTIDE**

⇒ Base + Pentose + Groupement phosphate



Nucléosides = Base + Pentose

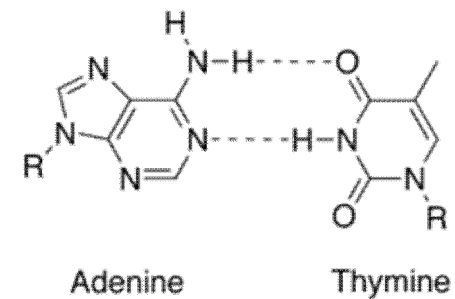
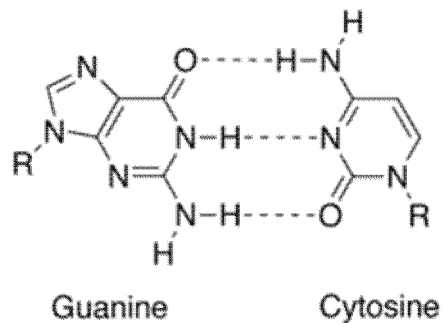
## BASES AZOTÉES



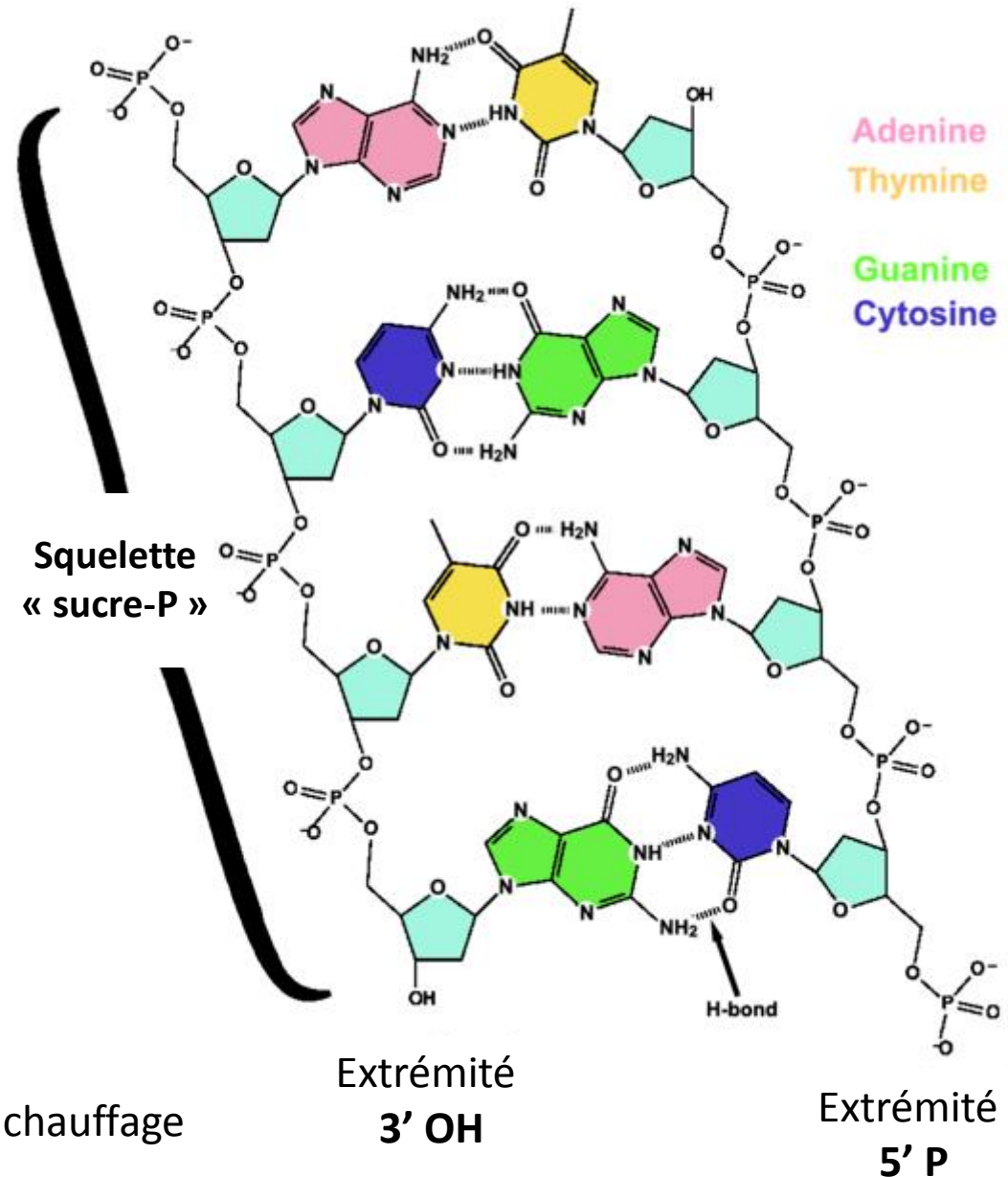
L'appariement des bases azotées s'effectue par liaisons hydrogènes

**CG** (ADN et ARN) : 3 liaisons H (E = -63 kJ)

**AT** (ADN) / **AU** (ARN) : 2 liaisons H (E = -21 kJ)

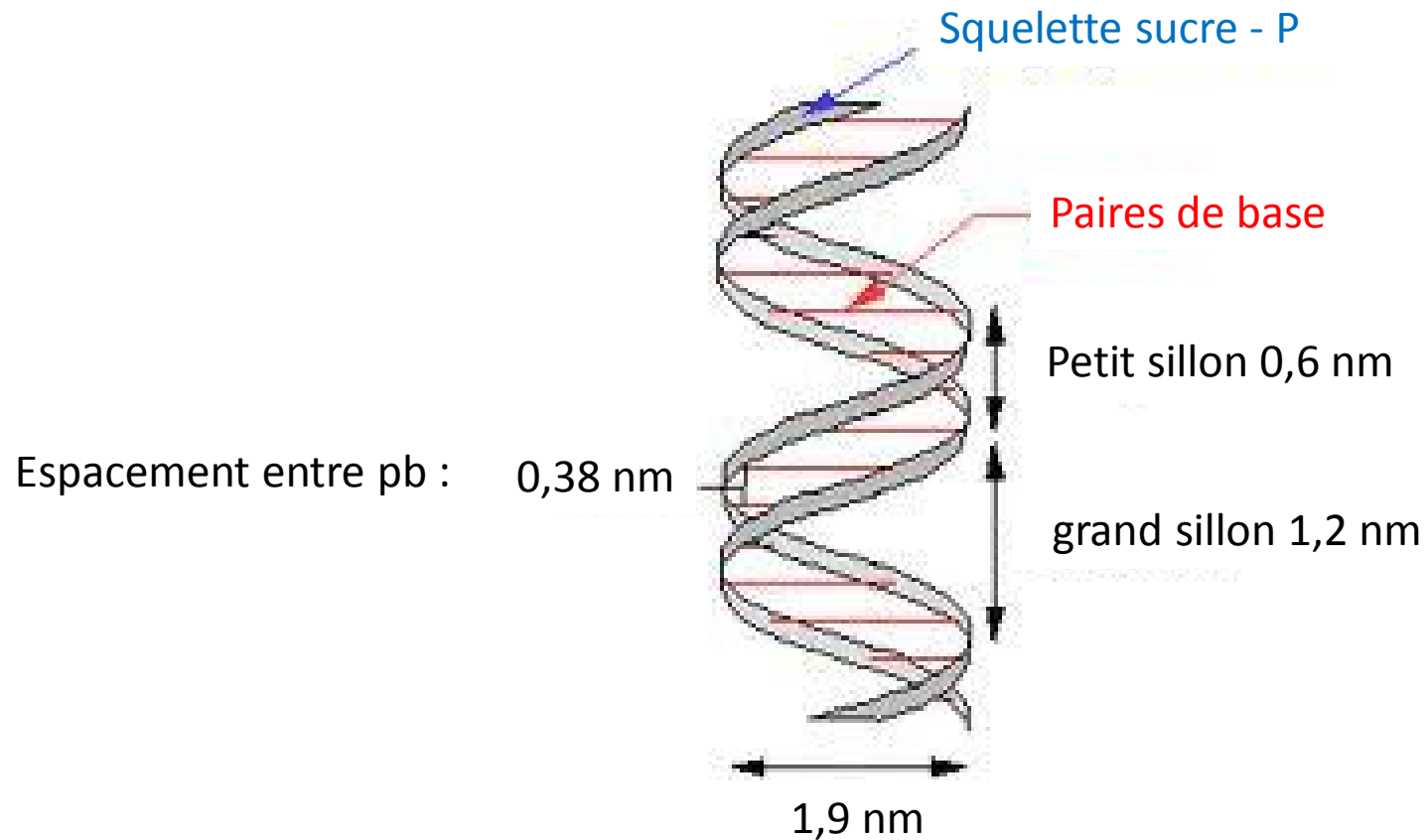


- L'ADN, support de l'information génétique est constitué de deux brins (DS « double strands ») complémentaires et antiparallèles (= bicaténaire)
- Acide faible
- Sens d'enroulement droit
- Super enroulements
- Etat intermédiaire : relâché
- Dénaturation de l'ADN double brin par chauffage



- La double hélice d'ADN possède un petit et un grand sillon (fixation de molécules « intercalantes » de l'ADN)

Pb / tour : 10,4  
Rotation / pb  $34,6^\circ$





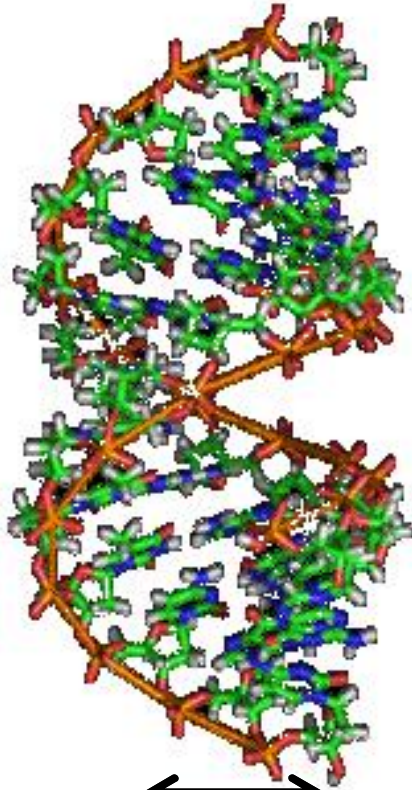
## STRUCTURE ALTERNATIVES DE L'ADN DS

- L'ADN est une molécule polymorphe : importance pour la régulation de l'expression génique

tour d'hélice :      2,53 nm  
                              11 pb / tour

3,54 nm  
10,4 pb / tour

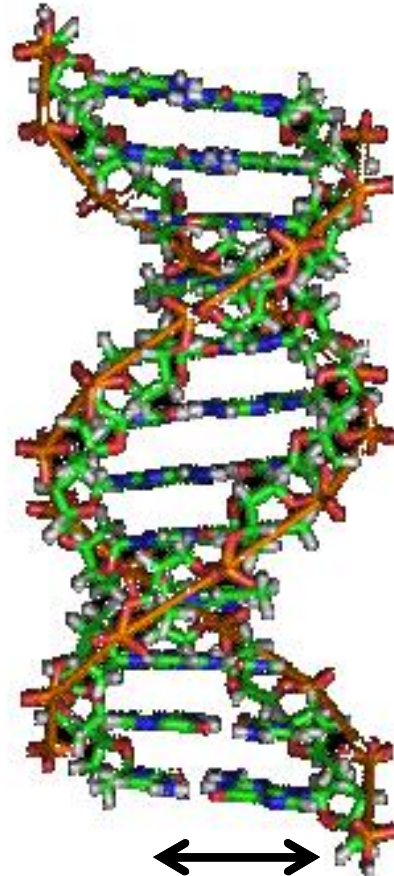
4,56 nm  
12 pb / tour



↔  
2,55 nm



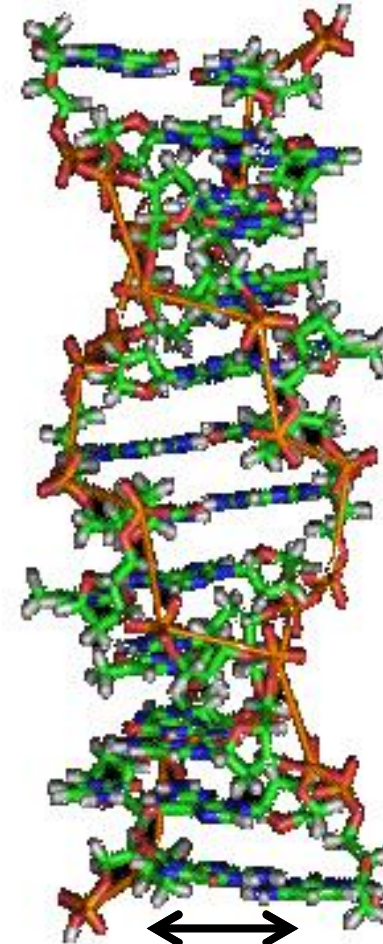
Dimère ADN/ARN



↔  
2,37 nm



La plus commune



↔  
1,84 nm



Séquences GC



## 1.2 Matériel génétique

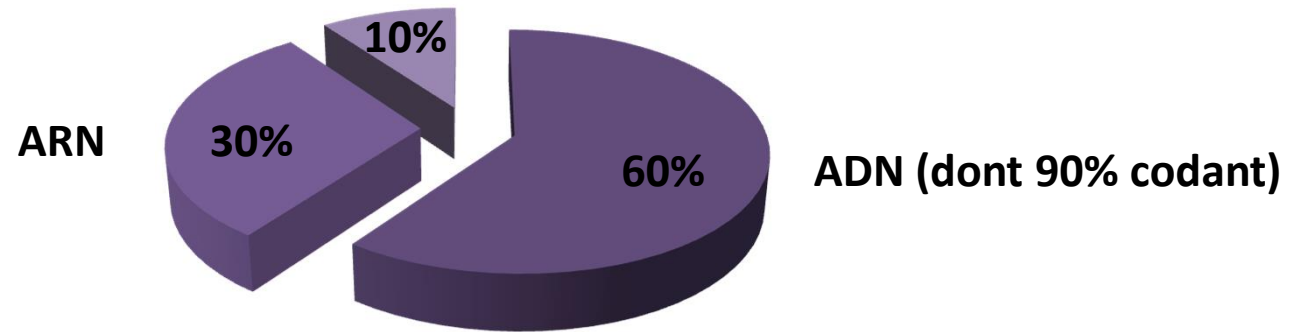
### 1.2.1 Génophore de procaryotes

□ Chromosome circulaire pelotonné (dble hélice enroulée en super hélice)

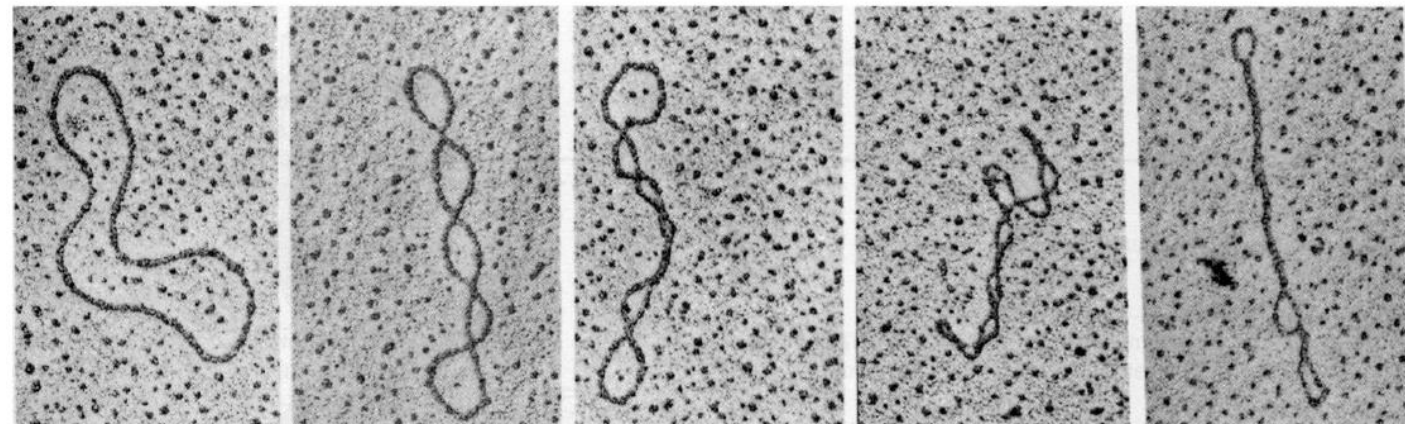
❖ Ordres de grandeur :

Chromosome circulaire (*E. coli* 1 mm, de  $4,6 \times 10^6$  pb à  $5,5 \times 10^6$  pb)

❖ Composition : **Protéines (Polymérase)**



Topoisomères  
de l'ADN

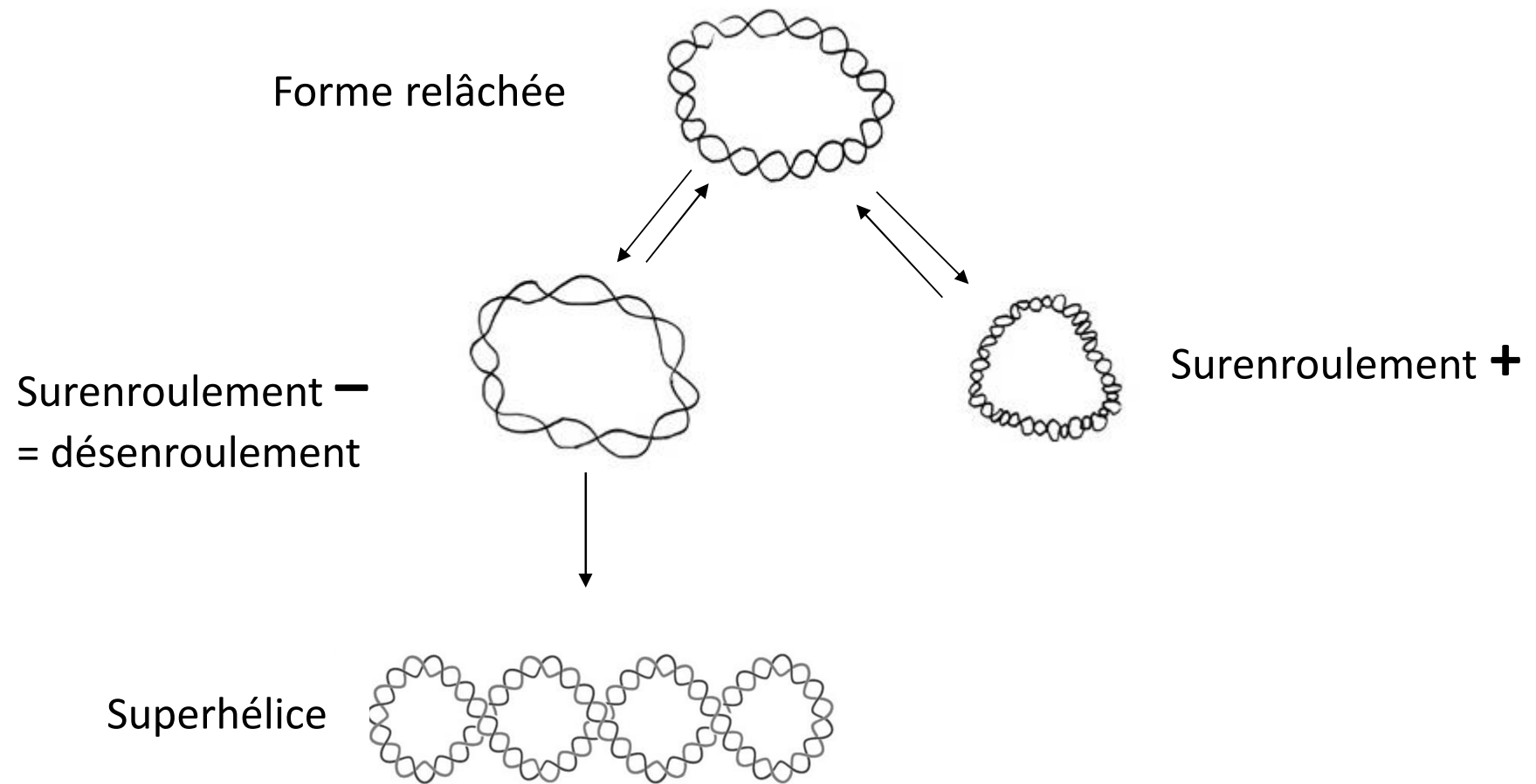


Relaxée

Torsadée

Super torsadée

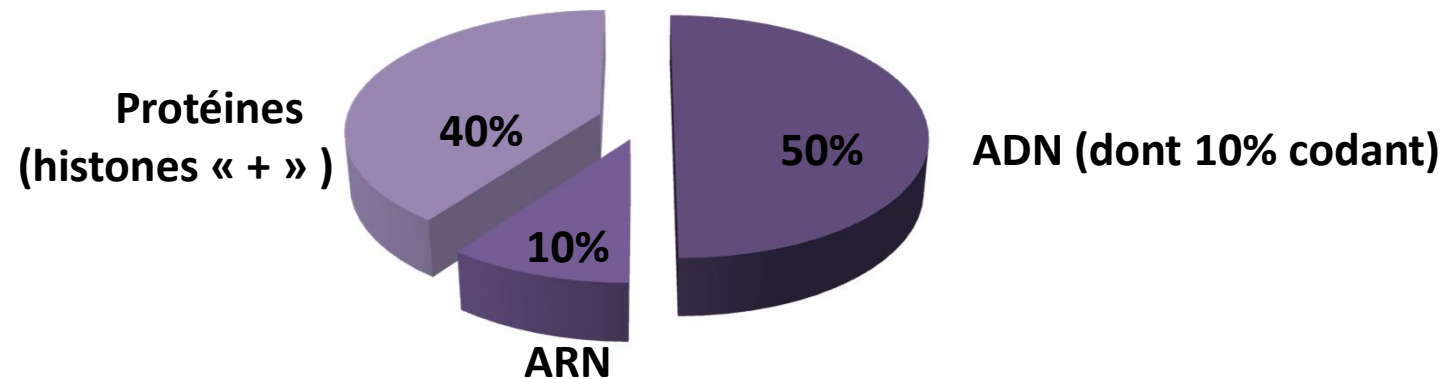
# Topoisomères de l'ADN



## 1.2.2 Matériel génétique des eucaryotes

❑ Plusieurs chromosomes dans le noyau (dble hélice enroulée en superélice)

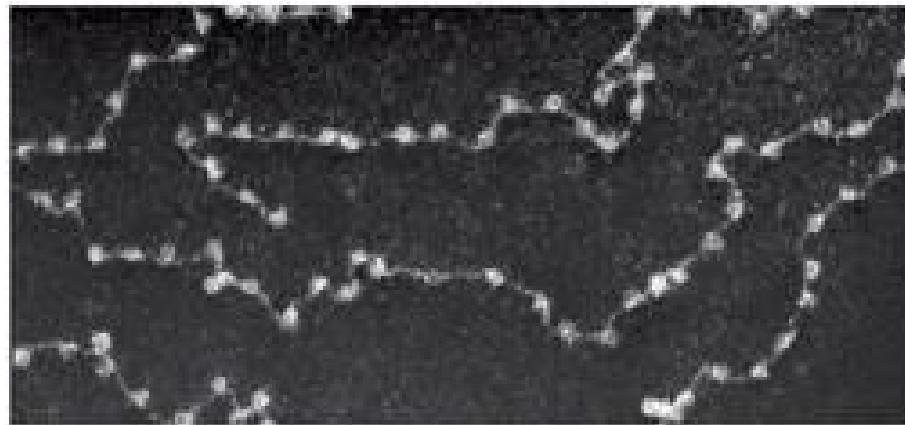
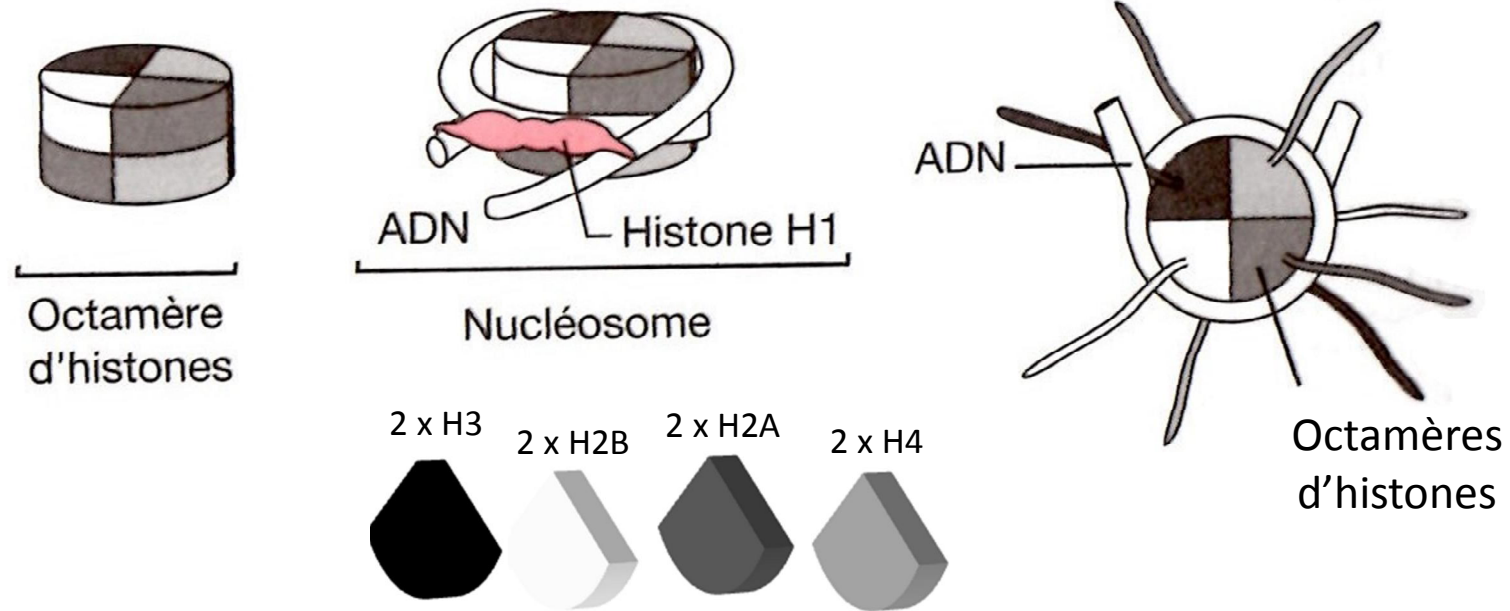
- ❖ Ordres de grandeur : 1 chromosome  $\sim 10^8$  pb
- ❖ Composition :



### Quelques définitions :

- ❑ Haploïdie / Diploïdie : 1 exemplaire / une paire de chaque chromosome (homme : 23 paires, levures : 16-18 chromosomes (ou paires), moisissures : 4-8 chromosomes)
- ❑ Nucléole : structure riche en ARN, site de synthèse de l'ARNr
- ❑ Division cellulaire végétative : bourgeonnement ou scissiparité

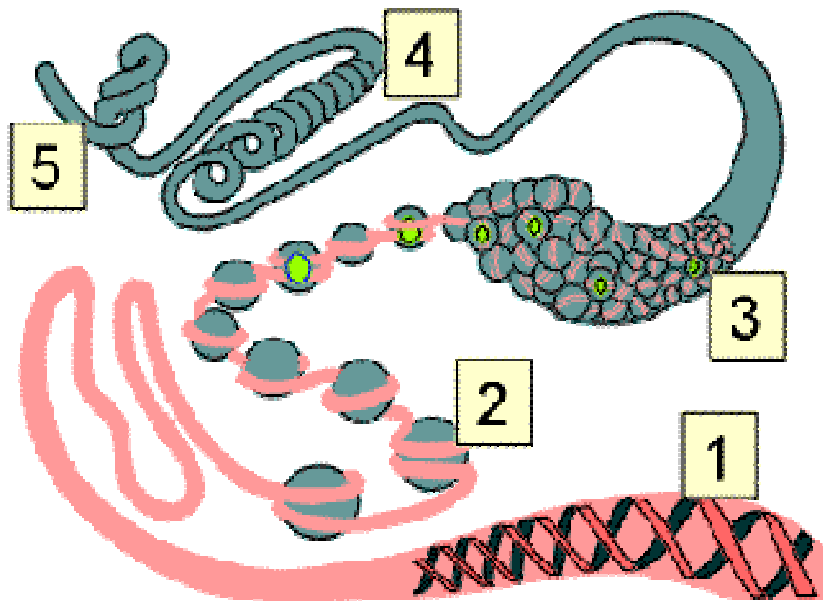
- Le nucléosome est l'élément structural de base  
(fibre de base = 200 pb et 10 nm de diamètre)



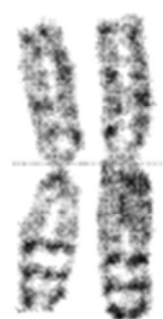
100 nm

## □ Structuration de l'ADN chromosomique (Représentation schématique)

Succession de déroulements qui caractérisent la molécule d'ADN au niveau des chromosomes.



1. **Double hélice classique : 10 pb / tour.**
2. **Nucléosome : l'hélice se déroule autour des histones (200 pb / nucléosomes ; 10 nm de diamètre)**
3. **Nucléosomes se regroupent en se déroulant sous forme de fibres de 30 nm de diamètre (6-7 nucléosomes/tour) = solénoïde.**
4. **Chacune de ces fibres se déroulent sous forme de spirale (superhélice de 50 fibres, 500 nm de diamètre).**
5. **Spirales se regroupent sous forme de mini-bandes visibles en microscopie optique (environ 18 spirales ou > 1 M de pb).**



1



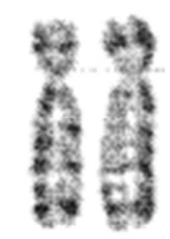
2



3



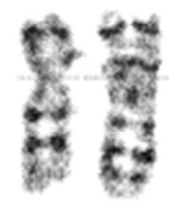
4



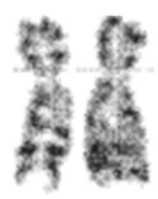
5



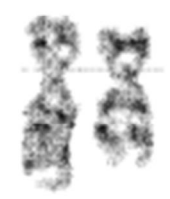
6



7



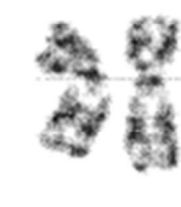
8



9



10



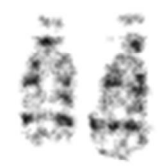
11



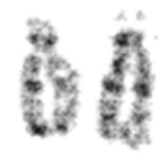
12



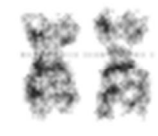
13



14



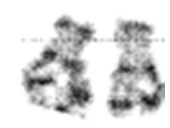
15



16



17



18



19



20



21



22



X



Y

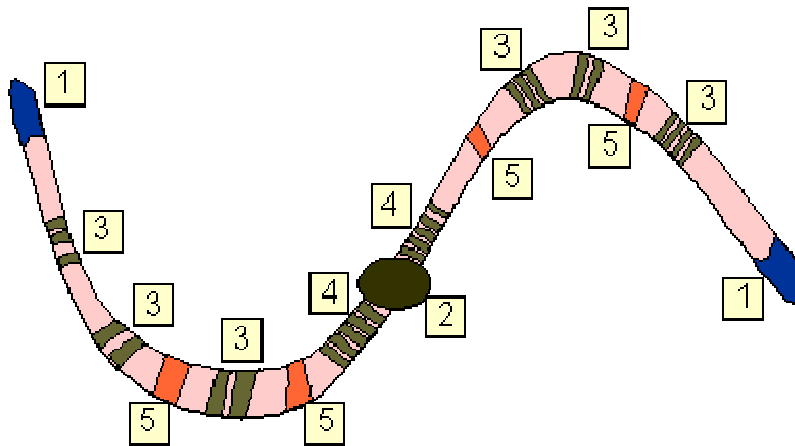


## □ Les chromosomes eucaryotes

- ❖ Chromatine hétérogène : hétérochromatine et euchromatine (code des histones)
- ❖ Gènes = exons + introns
- ❖ Gènes répétés en tandem : gènes qui codent pour l'ARN ribosomique (250 copies/cellule), l'ARN de transfert (50 copies/cellule) et les histones (20 à 50 copies par cellules)
- ❖ Séquences non codantes = extragénique (98% du génome humain)
- ❖ Séquences répétées : env. 50% du génome.  
Séquences répétitives en tandem ou dispersées

	Longueur	Motif (pb)	Localisation
Satellites	100 kb - 1 Mb	>100	Centromères
Minisatellites	1 - 30 kb	10 - 100	Télomères
Microsatellites	<150 pb	1 - 10	

❑ Les zones composant un chromosome eucaryote (Représentation schématique)



1. Télomères (aux deux extrémités du chromosome), localisation des séquences minisatellites
2. Centromère, localisation des séquences satellites
3. Séquences non codantes modérément répétitives;
4. Zones très répétitives non codantes
5. Gènes, plus ou moins spécifiques, certains gènes (comme ceux codant pour l'ARN ribosomal étant codés des centaines de fois).

## 1.3 Matériel extrachromosomique

### 1.3.1 ADN mitochondrial

- ❖ ADN DS et circulaire (sauf exception), plusieurs copies / mitochondries
- ❖ Les mitochondries sont des organites présents dans les cellules eucaryotes qui seraient issues de l'endosymbiose d'une alpha-protéobactérie (2 Mds d'années). Les mitochondries ont conservé leur propre génome.

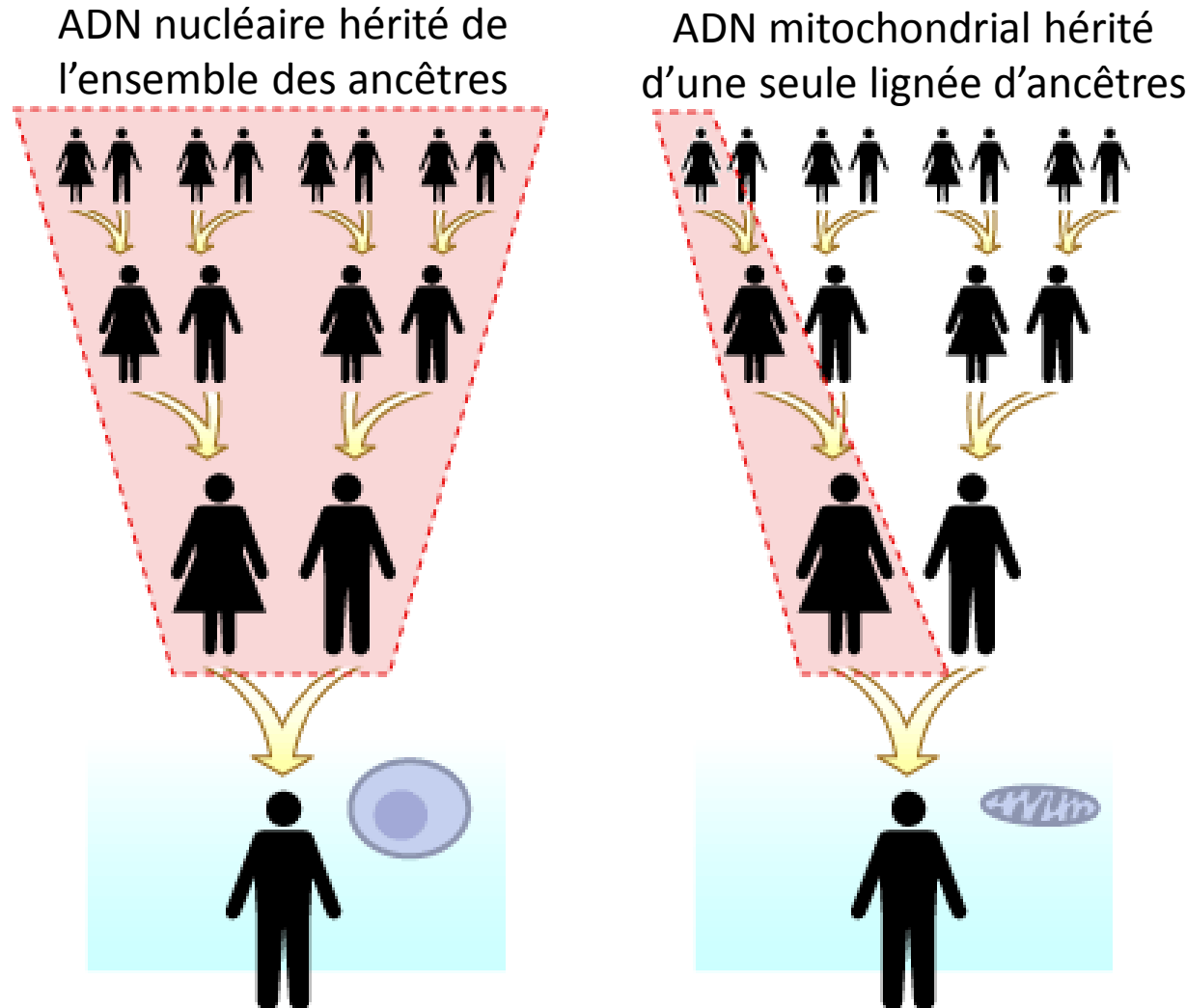
Chez l'homme : 16 kb (37 gènes séparés par des régions non codantes très courtes)

*Saccharomyces cerevisiae* 86 kb (43 gènes)

*Arabidopsis thaliana* 367kb (60 gènes)

□ Les mitochondries de la cellule-œuf proviennent exclusivement de l'ovocyte.

➔ Transmission de l'ADN mitochondrial par la mère (génétique des population)



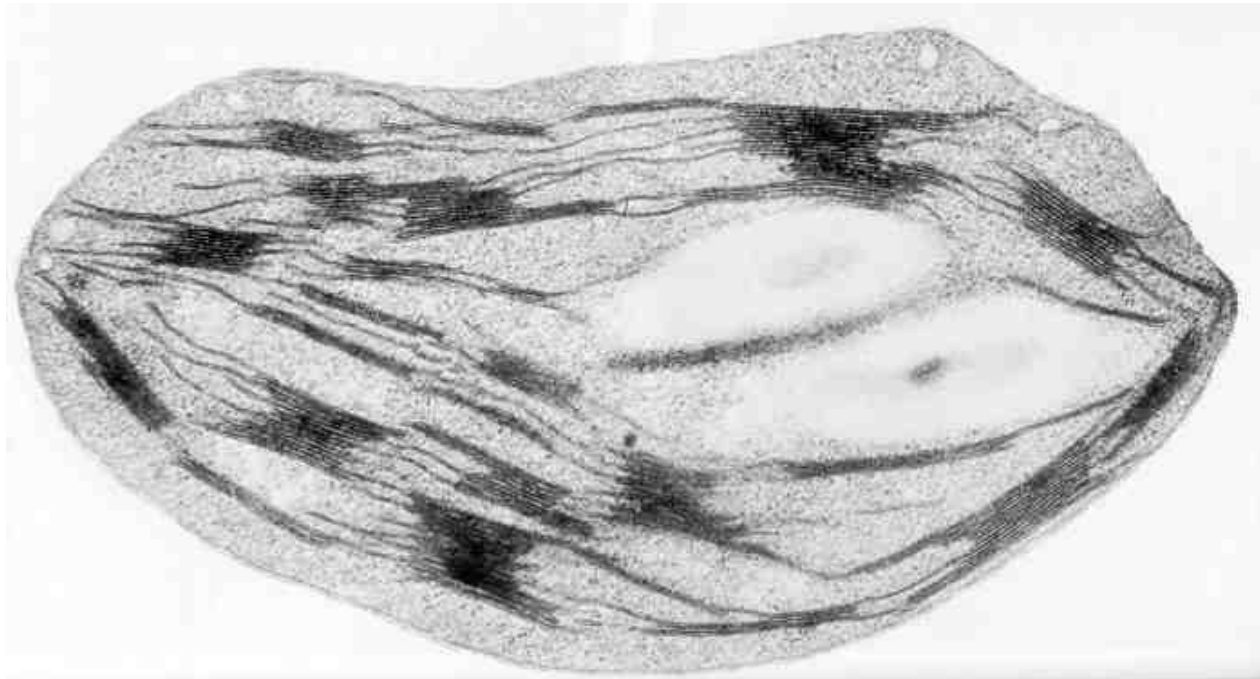
### 1.3.2 ADN Chloroplastique

Taille de 120 à 160 kb

Plusieurs copies / chloroplaste

ADN circulaire et non associé à des histones

Code pour une partie des protéines chloroplastiques (organites semi autonomes),

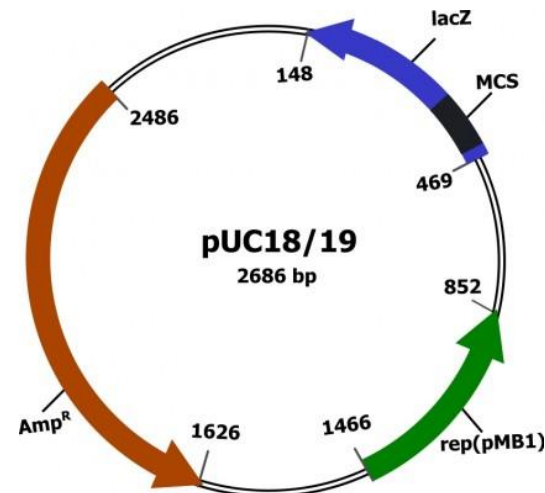


Chloroplaste (microscopie électronique)

### 1.3.3 Plasmides (principalement chez les procaryotes)

- Copies : 5 à 30 / cellules
- ADN DS circulaire
- Taille  $10^3$  à  $2 \times 10^5$  pb
- Réplication autonome, non indispensable à la cellule
- Transmission de cellule à cellule par conjugaison
- « Parfois intégrés au chromosome, épisome (cas des phages) »

❖ Très utilisés en génie génétique  
(plasmides commerciaux)



### Plasmides eucaryotes

Moins nombreux, même structure et taille

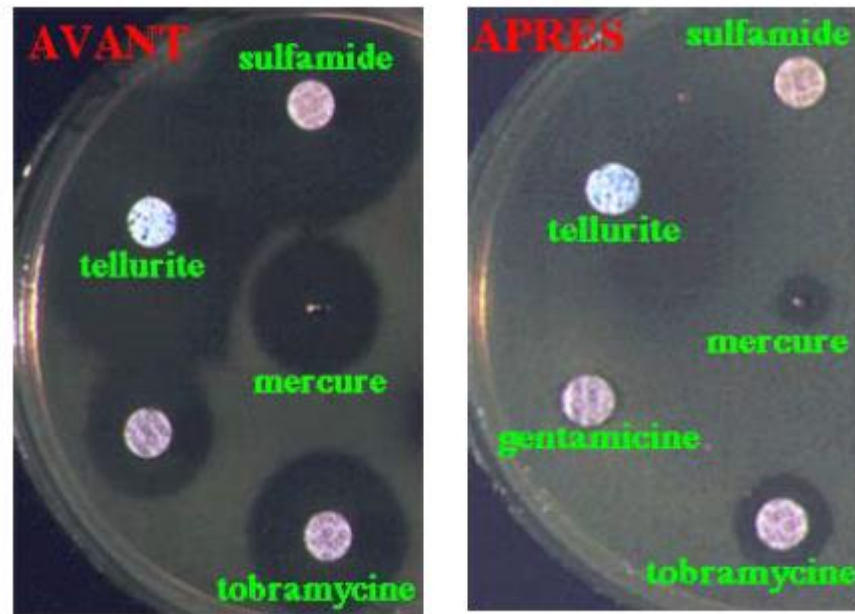
*S. cerevisiae* : taille 2  $\mu$ m ; 50-100 copies / cellule

- ❖ Ils sont médiateurs de nombreuses propriétés → meilleure adaptation des bactéries :

Port de gènes d'attaque (AB, toxines...) ou de résistance (AB, métaux lourds...) ou de synthèse d'enzymes métaboliques spécifiques (phénols, camphre...)

➔ Plasmides quasi-ubiquitaires chez les bactéries pathogènes

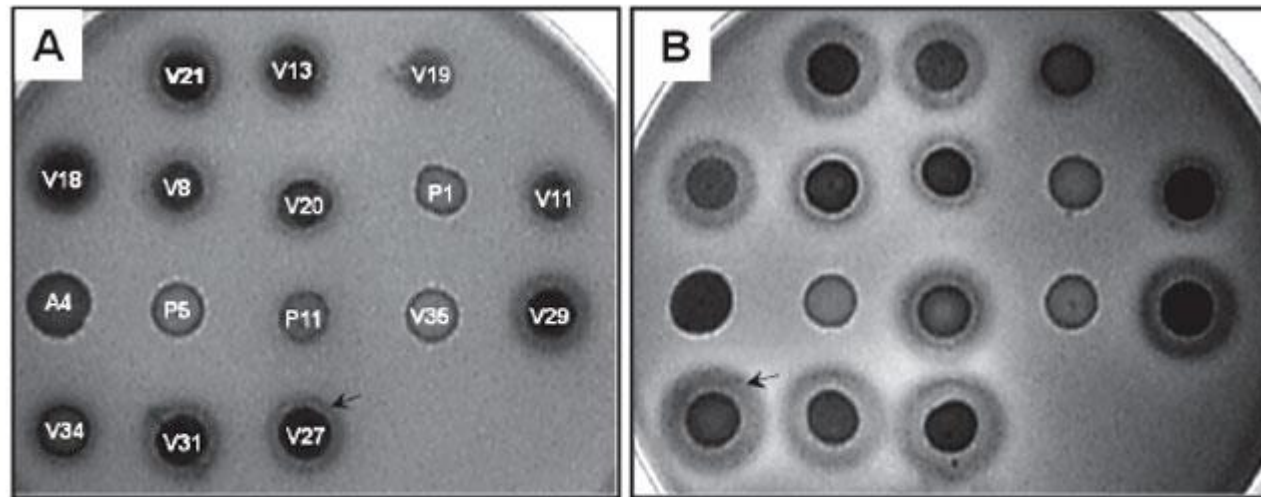
- ❖ Exemple : acquisition de différentes propriétés par des bactéries :





### 1.3.4 Facteur « killer »

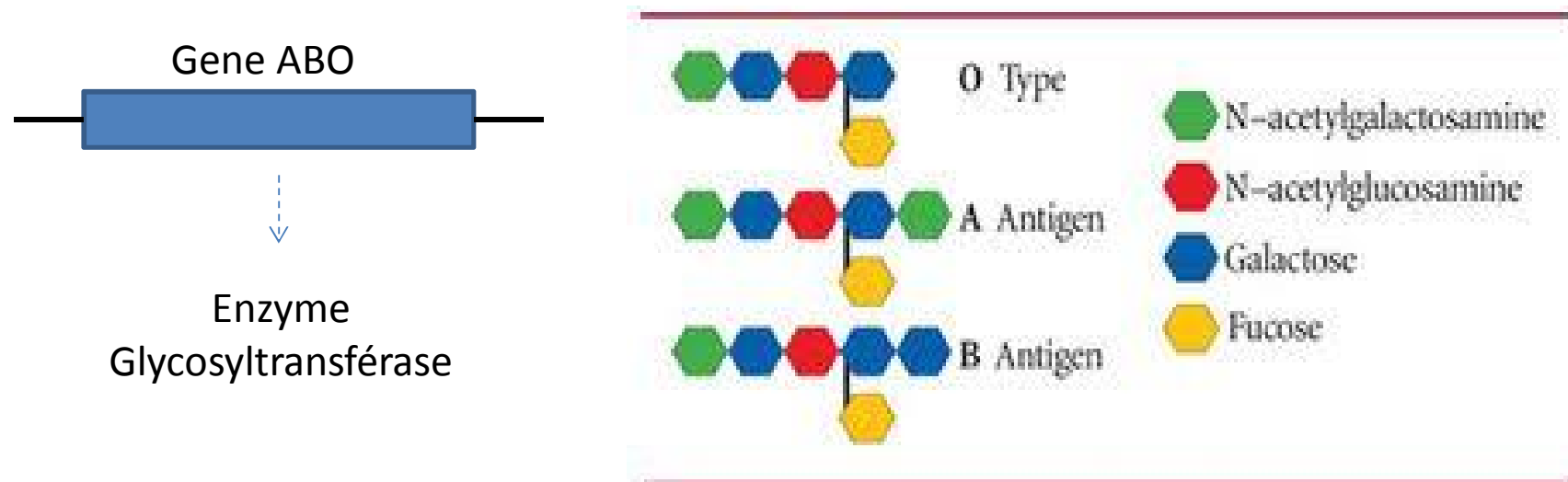
- ❑ Certaines souches de *Saccharomyces sp.* ont la propriété de libérer des substances toxiques.
- ❑ Certaines levures contiennent 2 virus à ARN.
  - ” Le « petit virus » code une protéine toxique exo-cellulaire capable de tuer d'autres levures et une protéine de résistance à cette même toxine pour empêcher les levures « killer » de se tuer entre elles.
  - ” Le « grand virus » est nécessaire à la multiplication et au maintien du « petit virus » dans le cytoplasme, il contient, entre autres, les gènes codants pour la capsid.



Mise en évidence de levures « killer »

## 1.4 Les gènes

- ❑ Les gènes sont les entités contenant l'information nécessaire à la synthèse de macromolécules
- ❑ Un gène est une portion d'ADN.
- ❑ Un gène peut coder pour une fonction / un caractère (ex : l'insuline, groupe sanguin).



- ❑ Une fonction / un caractère peut être codée par plusieurs gènes (ex : les antibiotiques, les mycotoxines, la couleur de la peau)



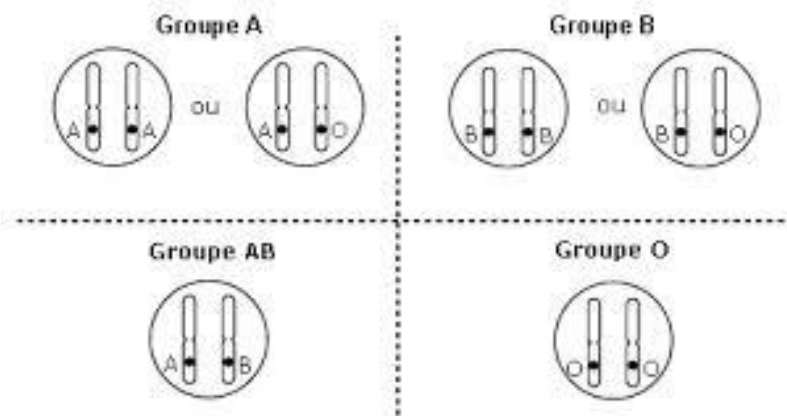
Cluster de gènes codants pour les enzymes responsable de la biosynthèse de mycotoxines (trichotecenes, *Fusarium sp.* sur orge de brasserie)

## Quelques définitions :

- ❑ Le locus est l'emplacement d'un gène sur le chromosome.

ex : locus du gène CFTR responsable de la mucoviscidose sur le chromosome n°7

- ❑ Les allèles correspondent aux différentes versions d'un même gène, Un allèle se différencie d'un autre par une ou plusieurs différences de la séquence de nucléotides. Même locus. (ex : groupes sanguins A, B et O).

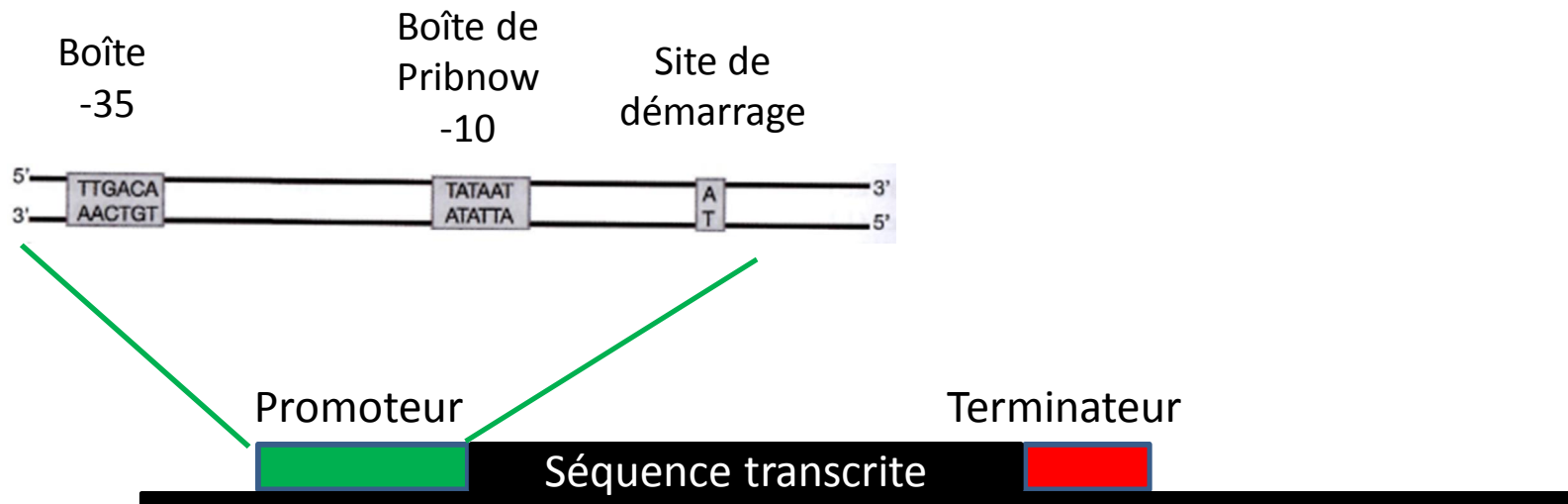


		DONNEUR							
		O-	O+	B-	B+	A-	A+	AB-	AB+
RECEVEUR	AB+	🔴	🔴	🔴	🔴	🔴	🔴	🔴	🔴
	AB-	🔴		🔴		🔴		🔴	
	A+	🔴	🔴			🔴	🔴		
	A-	🔴				🔴			
	B+	🔴	🔴	🔴	🔴				
	B-	🔴		🔴					
	O+	🔴	🔴						
	O-	🔴							

- ❑ Génome = ensemble des gènes
- ❑ Génotype = combinaisons d'allèles
- ❑ Phénotype = caractère observable d'un individu (il correspond à une interaction génotype / environnement (ex : un jumeau développe un cancer et pas l'autre))

## 1.4.1 Structure du gène procaryote

- Unité transcriptionnelle
- Pas d'intron
- Promoteur : Unité reconnue par l'ARN polymérase
- Organisés en opéron



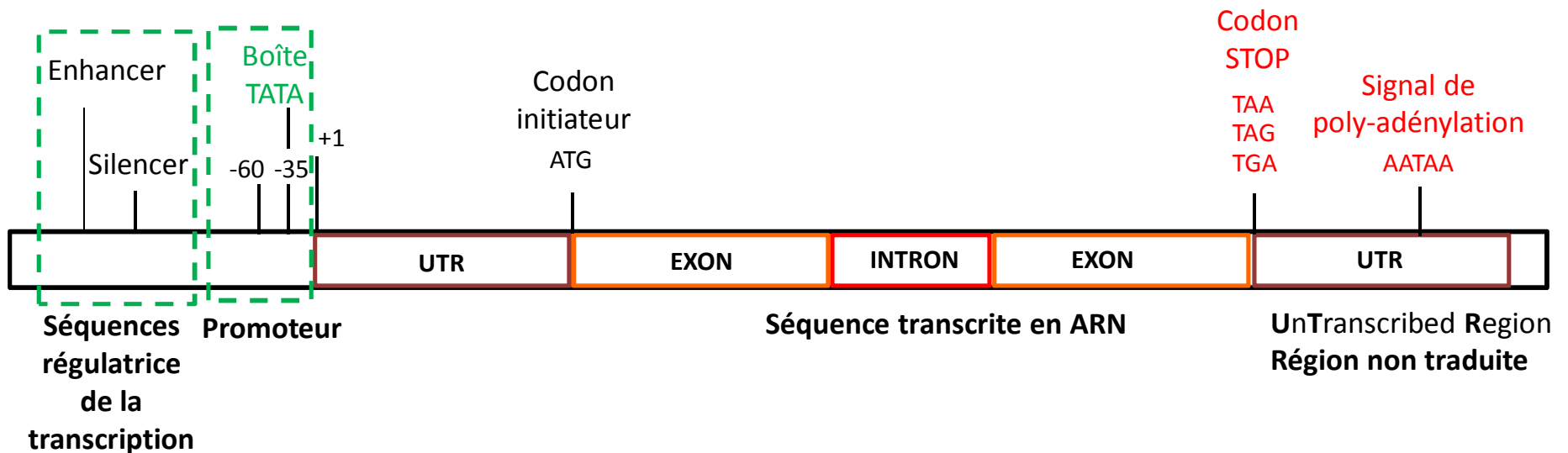
Séquences répétées riches en GC

```
5' cgcatttgccctccggtaggaggctttt 3'
3' gcgaaacggaggccatccaccgaaaaa 5'
```

Répétition de A

## 1.4.2 Structure du gène eucaryote (unité transcriptionnelle)

- Présence d'intron(s) - épissage
- TATA Box (20% des cas aux alentours de la position 40)
- Initiation, élongation et terminaison
- Promoteur plus étendu que chez les procaryotes (plusieurs centaines de nucléotides)
- En aval du codon stop, pas de terminateur = décrochage de la polymérase mais une séquence de clivage/polyadénylation = maturation (coupure + ajout queue PolyA)



# TECHNIQUES DE GENIE GENETIQUE

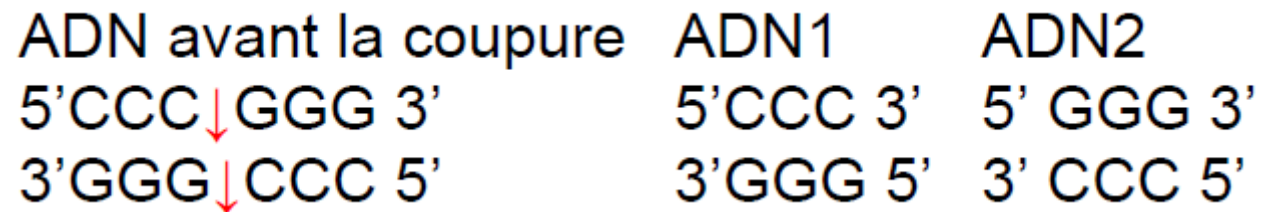
**I Clonage** Isolement d'un fragment d'ADN à partir d'un génome et obtention d'une population de molécules d'ADN identiques par insertion de ce fragment dans un vecteur.

Découpage : Utilisation des enzymes de restriction (origine microbienne).

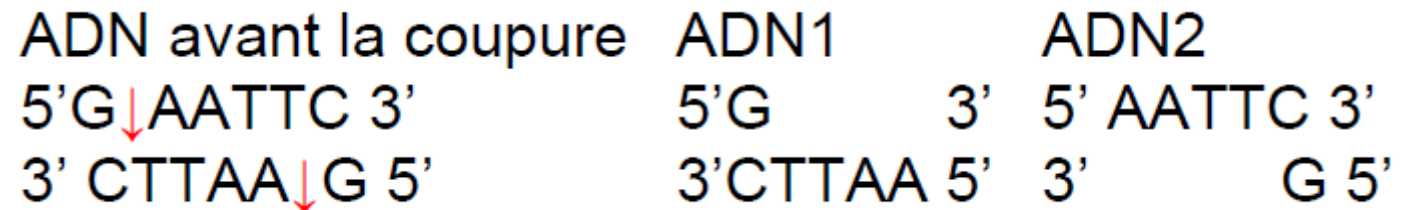
Reconnaissance de sites ayant une séquence déterminée (entre 4 et 15 pb)  
= **site de restriction** (séquence palindromique ERII)

La fixation de l'ER sur l'ADN permet une double coupure = **endonucléase**

*Bouts francs*



*Bouts sortants*



Les parties 5S peuvent s'apparier



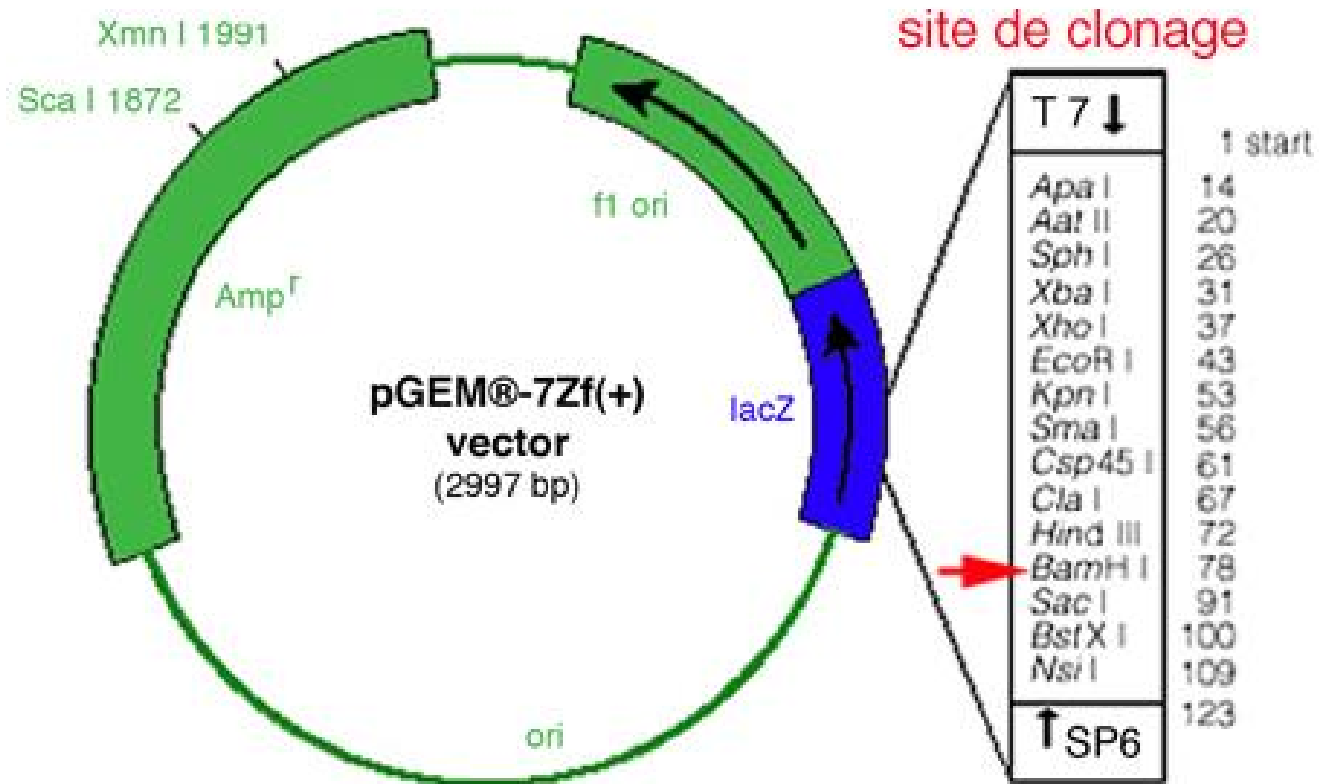
## Nomenclature des enzymes de restriction

- La première lettre, en majuscule, représente l'initiale du genre bactérien
- Les deux lettres, minuscules, qui suivent la première sont représentatives de l'espèce.
- Le chiffre romain qui suit ces trois lettres est le numéro d'ordre de découverte de l'enzyme pour la même bactérie source.
- La dernière lettre majuscule n'est pas obligatoire pour toutes les endonucléases de restriction. Elle est représentative de la souche de la bactérie d'où l'enzyme a été isolée.

La bactérie **E**scherichia **coli Ry13**

Les enzymes **EcoR I** : première enzyme isolée chez *Escherichia coli* Ry13  
**EcoR V** : cinquième enzyme isolée chez *Escherichia coli* Ry13

<b>Enzyme</b>	<b>Origine bactérienne</b>	<b>Site de restriction</b>
Alu I	<i>Athrobacter luteus</i>	AG/CT
Bam HI	<i>Bacillus amyloliquefaciens</i>	G/GATCC
Bgl II	<i>Bacillus blobiggi</i>	A/GATCT
Eco RI	<i>Escherichia coli</i>	G/AATTC
Hind III	<i>Haemophilus influenzae</i>	A/AGCTT
Hea III	<i>Haemophilus aegyptus</i>	GG/CC
Hpa II	<i>Haemophilus parainfluenzae</i>	C/CGG
Mbo I	<i>Moraxella bovis</i>	GA/TC
Pst I	<i>Providentia stuartii</i>	CTGCA/G
Taq I	<i>Thermophilus aquaticus</i>	T/CGA
Wba I	<i>Xanthomonas badrii</i>	T/CTAGA



## Polissage

Transformer les bouts sortants en bouts francs

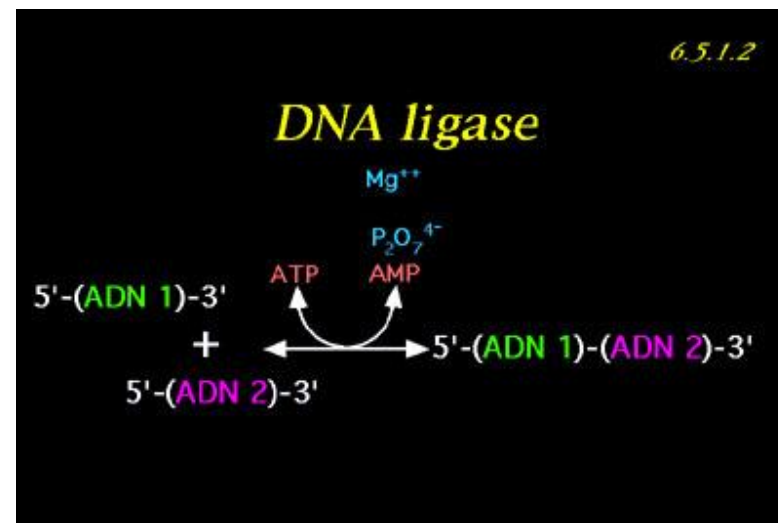
- Endonucléase SS (Haricot *mungo*)
- ADN polymérase du phage T4

Si 5' sortante : Activité polymérase rebouche l'ADN en allongeant l'extrémité 3'OH

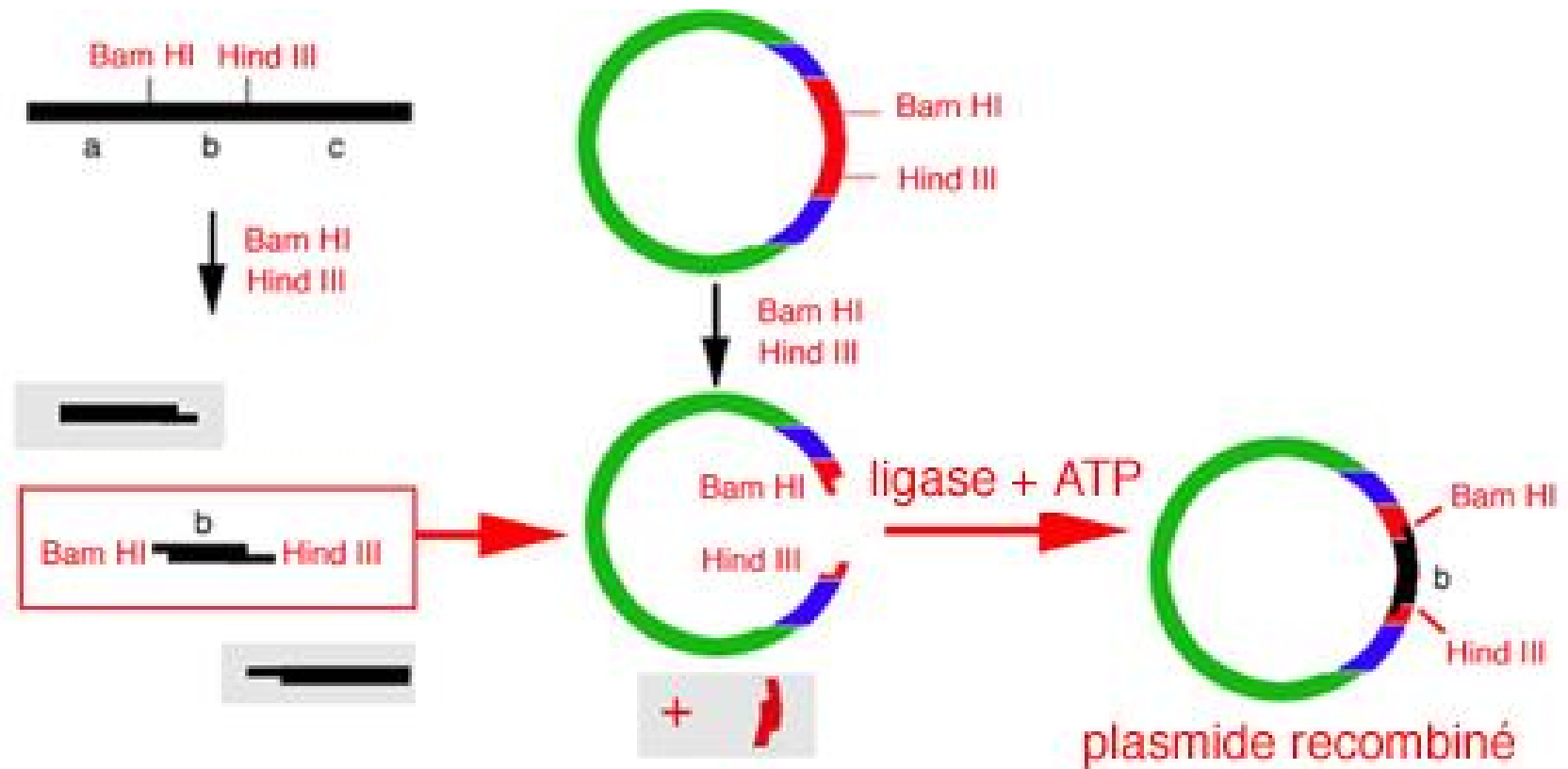
Si 3'OH sortante : activité exonucléase digère de 3' vers 5'

## Ligature

L'ADN ligase est une enzyme qui répare les brins brisés d'ADN.  
Création d'une liaison phosphodiester entre le 5' P et OH 3'.

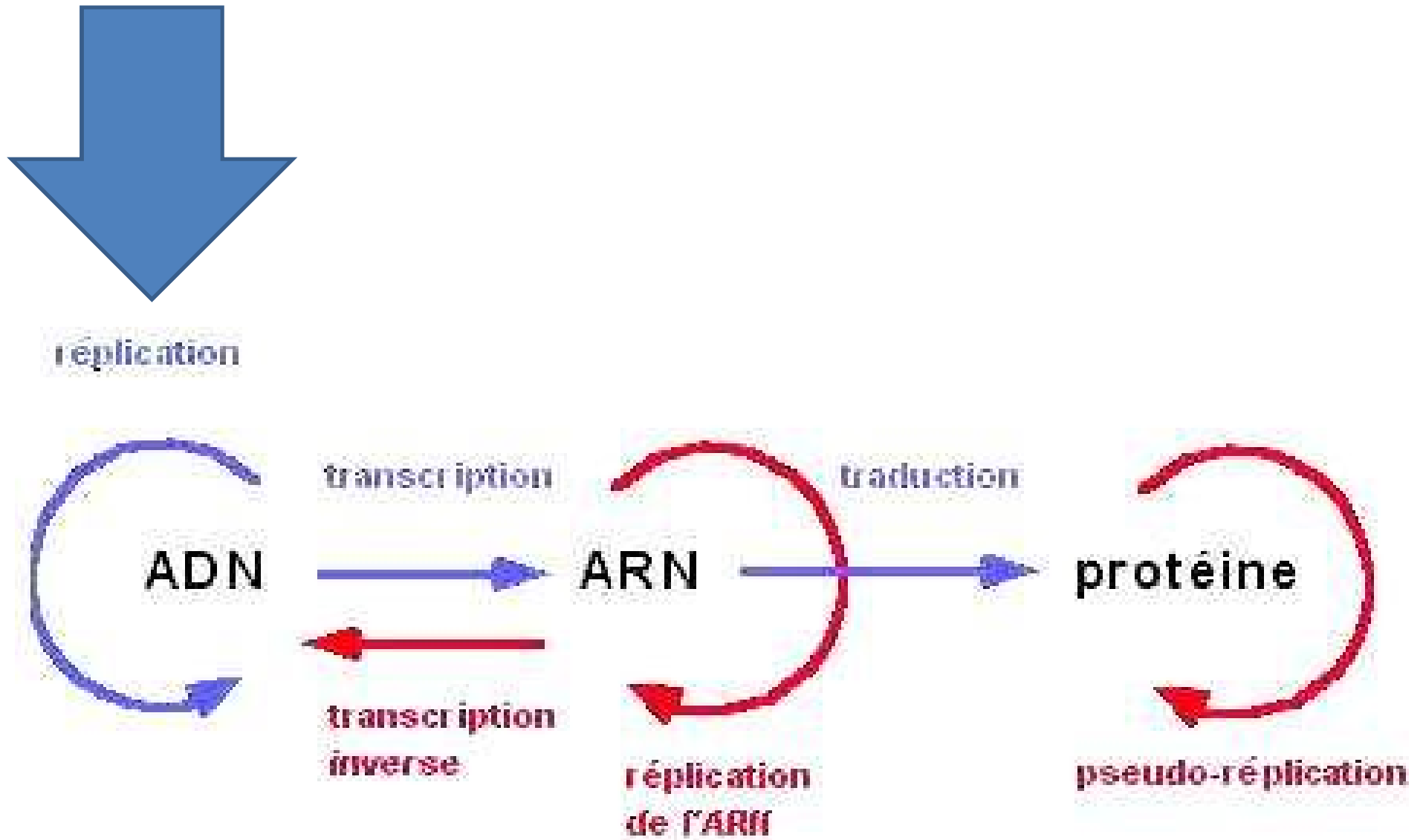


## Clonage d'un fragment d'ADN dans un plasmide



## II - RÉPLICATION

40



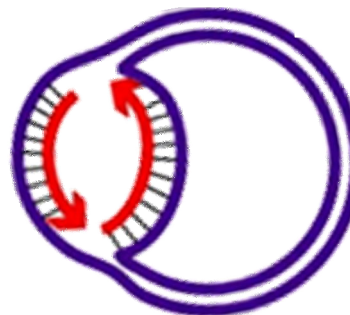
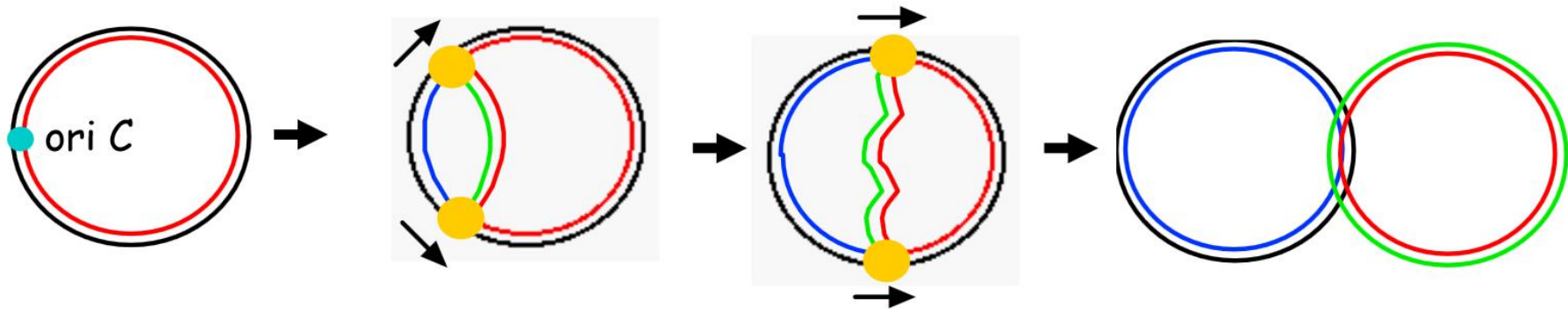
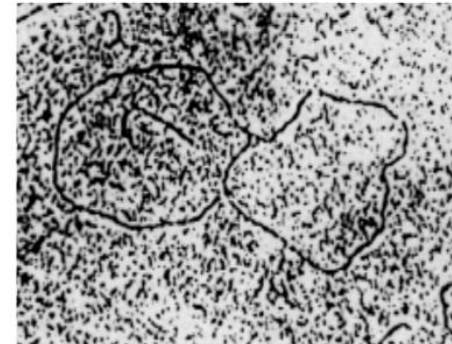
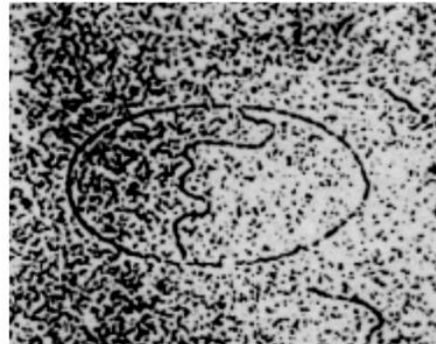
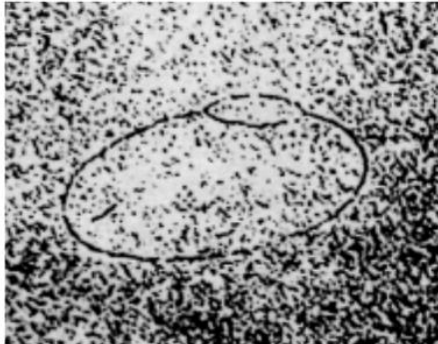
# II - RÉPLICATION

## 2.1 Caractéristiques

- ❑ Les deux brins de la double hélice se déroulent, chacun devient une matrice pour la synthèse d'un nouveau brin : **création de 2 molécules filles d'ADN DS identique à la molécule parentale.**
  
- ❑ La séquence des nouveaux brins est complémentaire à celle des brins matrices.
  
- ❑ Bi-diréctionnelle
  
- ❑ Polymérisation unidirectionnelle (5'P > 3'OH) : formation d'une liaison phosphodiester entre l'extrémité 3'OH du brin en voie d'élongation et l'extrémité 5'P du nucléotide ajouté par l'ADN polymérase
  
- ❑ Semi-conservative : (brin neosynthétisé = brin parent + brin fils)

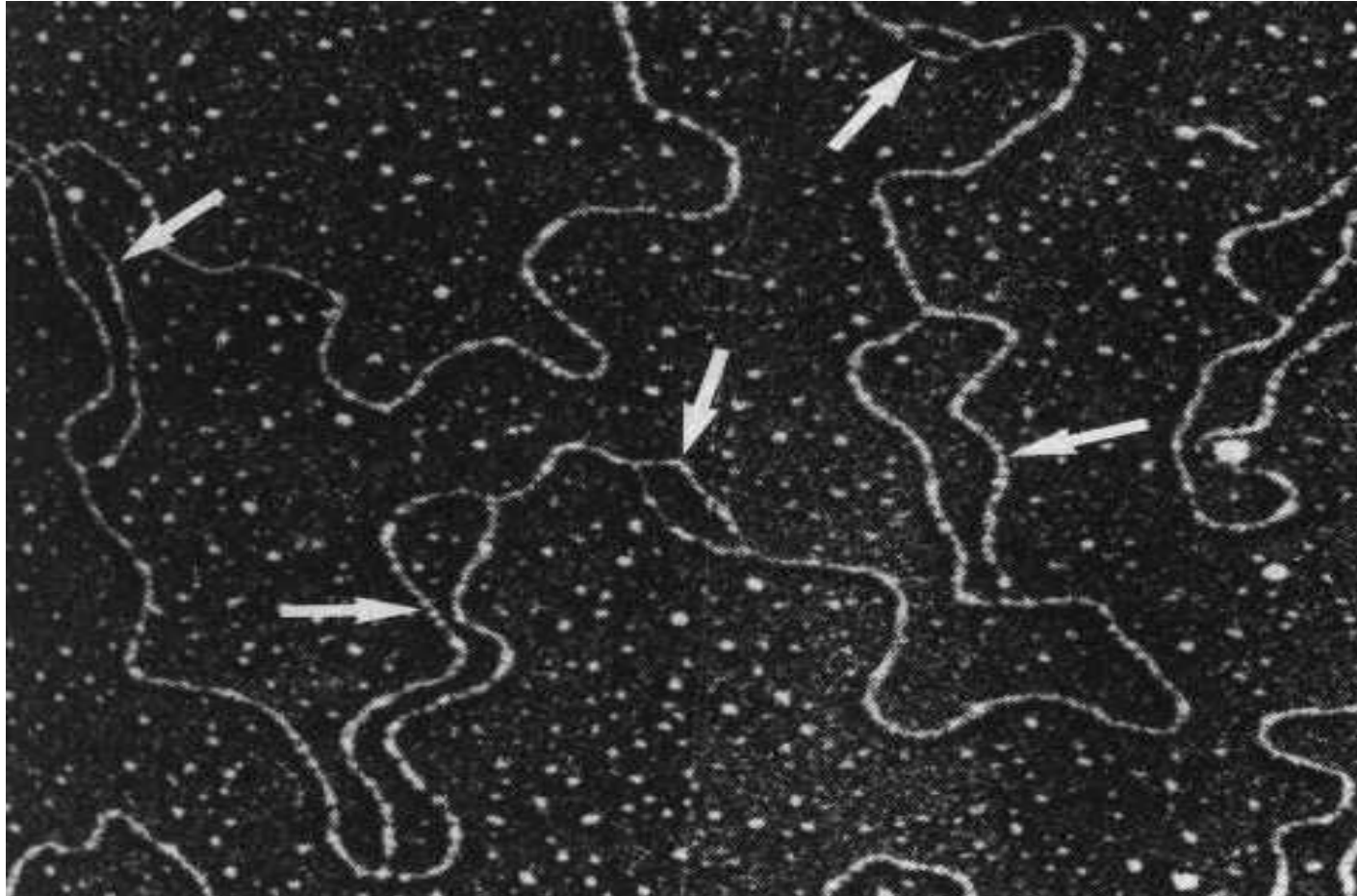
## II - RÉPLICATION

□ Structure de l'œil de répliation des procaryotes :





❑ Photographie au microscope électronique d'ADN eucaryote en répliation.



Plusieurs « yeux » de répliation (les réplicons)

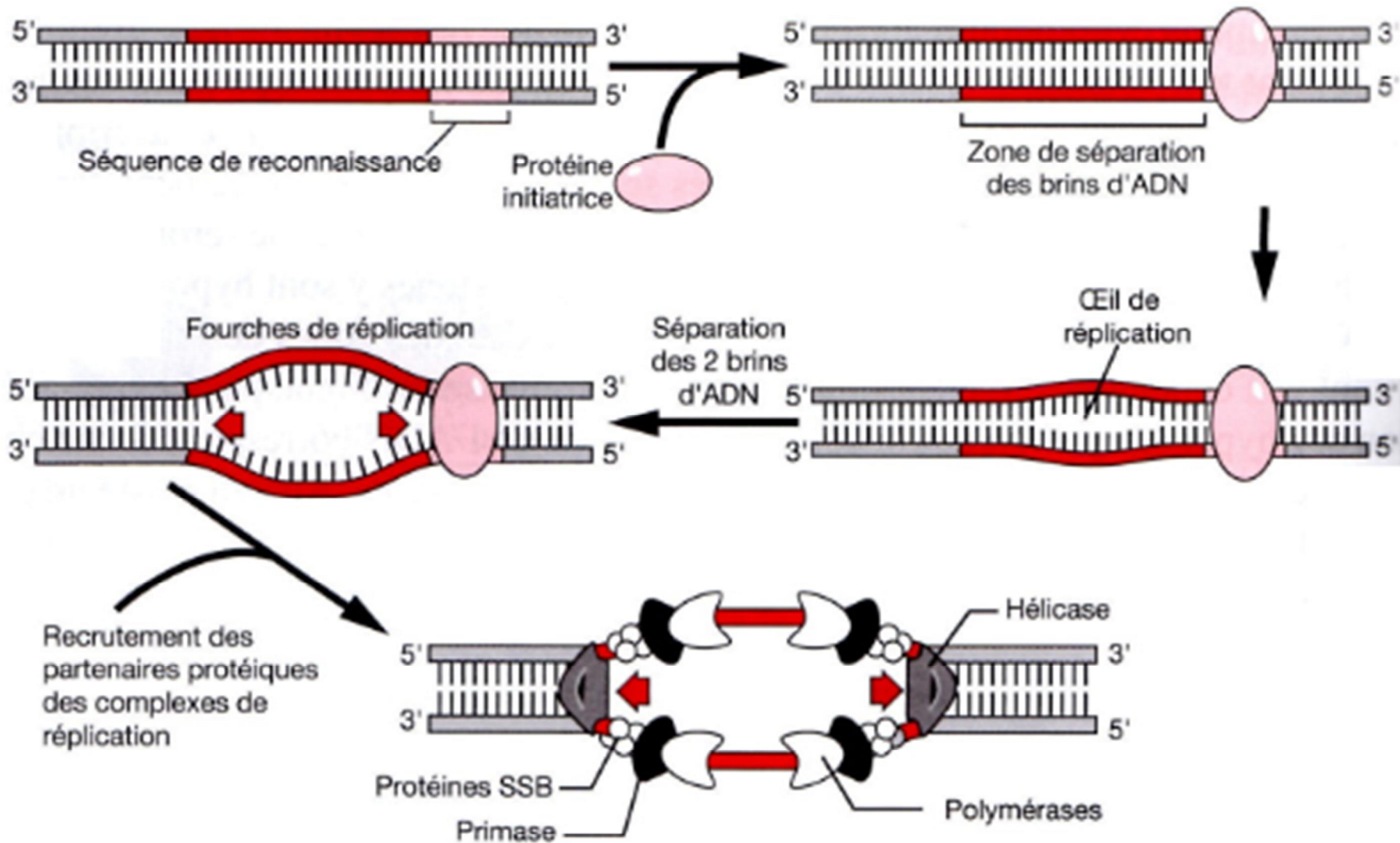
## II - RÉPLICATION

- ❑ **Réplicon** = unité de réplication de l'ADN eucaryote, il contient une origine et une terminaison (ADN procaryote circulaire et une seule origine de réplication).
- ❑ ADN eucaryote peut être répliqué à **plusieurs endroits** en même temps (sinon durée réplication ADN eucaryote = 800 heures)
- ❑ Réplicons = segments de taille variant de 30 kb à 150 kb et jusqu'à 35 000 réplicons. La vitesse de synthèse est de 3000 pb/min pour les eucaryotes et encore plus rapide pour les procaryotes ne présentant pas de chromatine.

**L'origine de réplication** n'est pas localisée au hasard sur la molécule d'ADN, en effet elle est constitué de petites séquences répétées reconnues par des protéines. Elle mesure environ 200 pb pour les procaryotes et 2 kb pour les eucaryotes. Pour l'ADN procaryote, les terminaisons sont doubles pour une origine et sont riches en A et T.

## II - RÉPLICATION

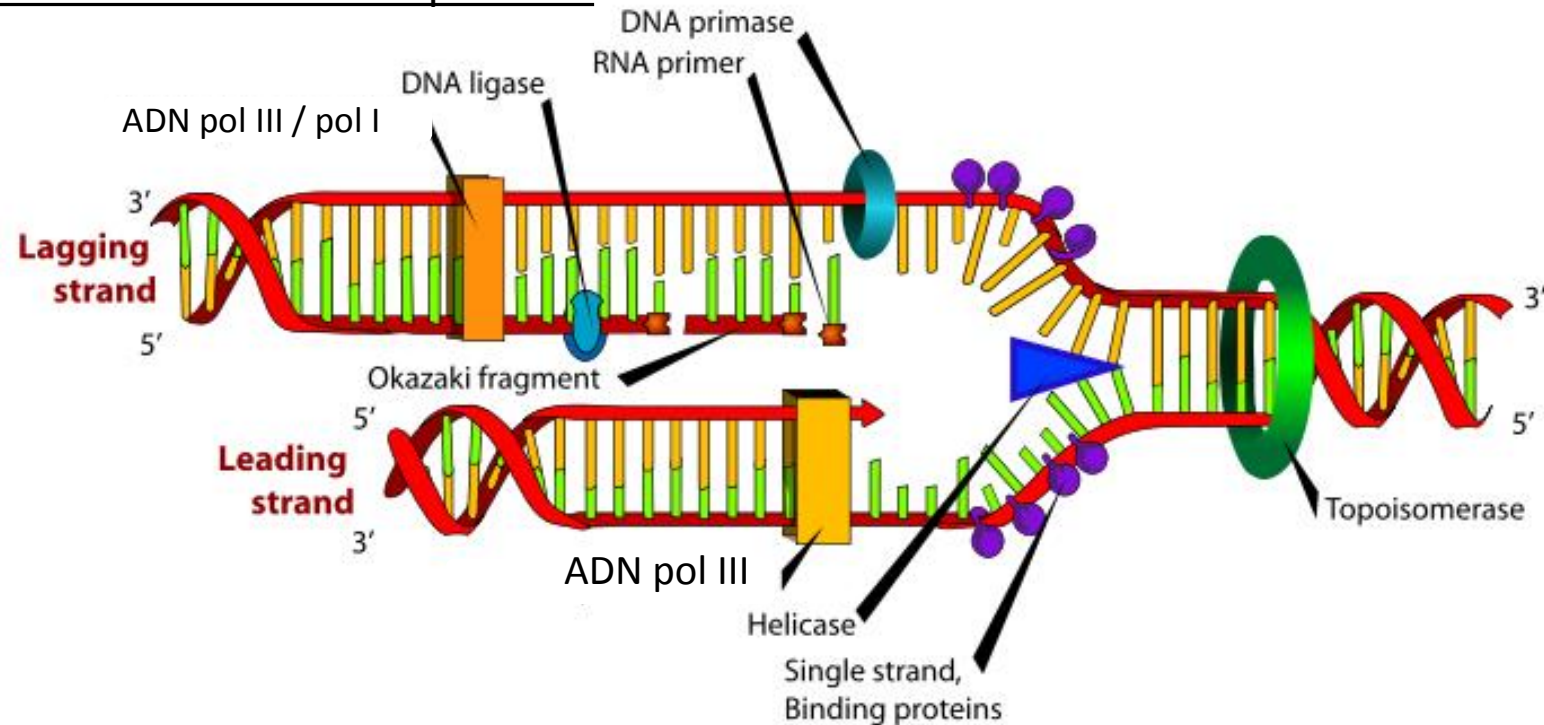
### 2.2 Initiation de la réplication



Fixation du complexe de réplication appelé **le réplisome**

## II - RÉPLICATION

### 2.3 Déroulement de la réplication



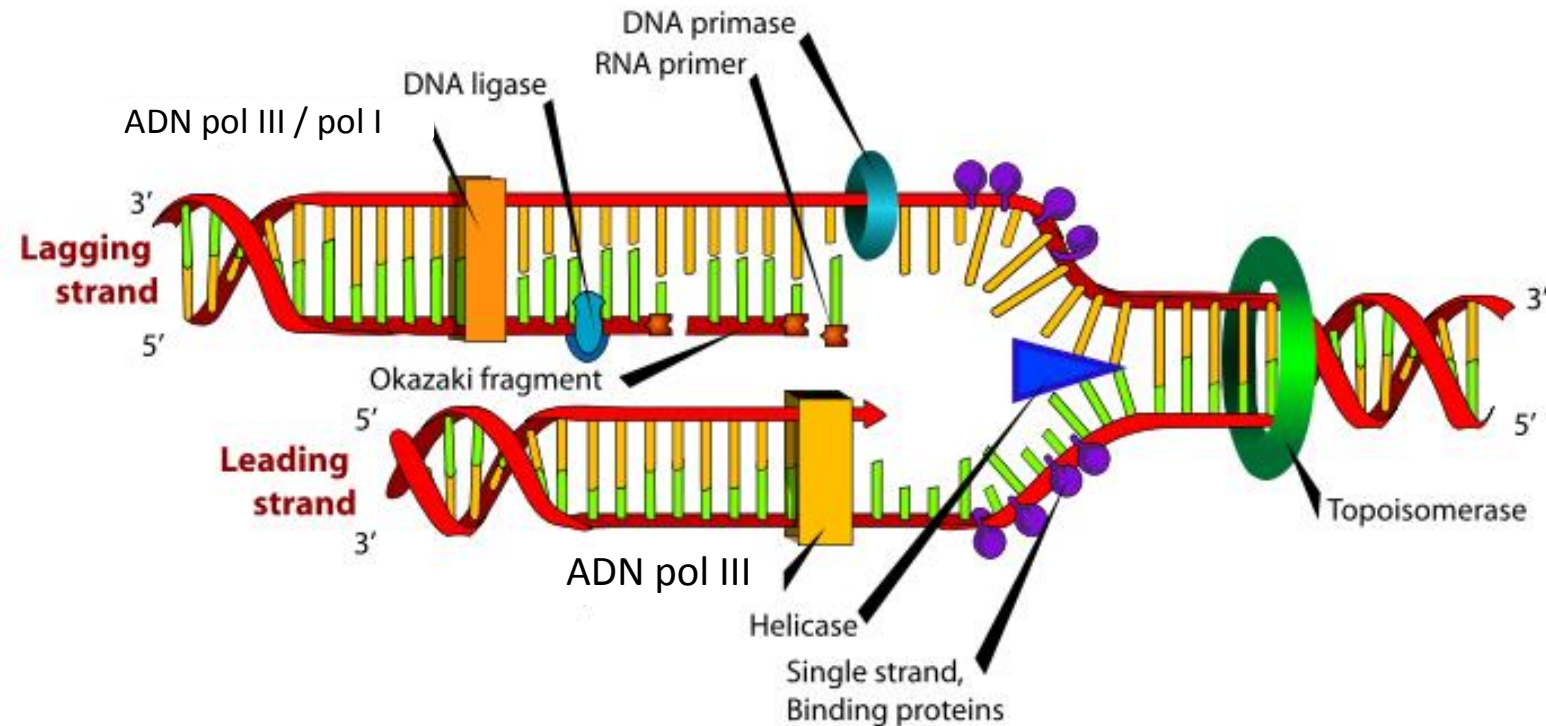
#### 1 Ouverture de la double hélice et formation de la fourche répliative:

Reconnaissance de l'origine de réplication par la protéine initiatrice (DNA A): Formation de l'œil de réplication et des deux fourches de réplication.

Hélicases (DNA B) : déroulement des deux brins ; les topo-isomérases, en aval de la fourche : permettre l'avancée des l'hélicases.

Les protéines SSB protègent les ADN SS pour les empêcher de se réenrouler.

## II - RÉPLICATION

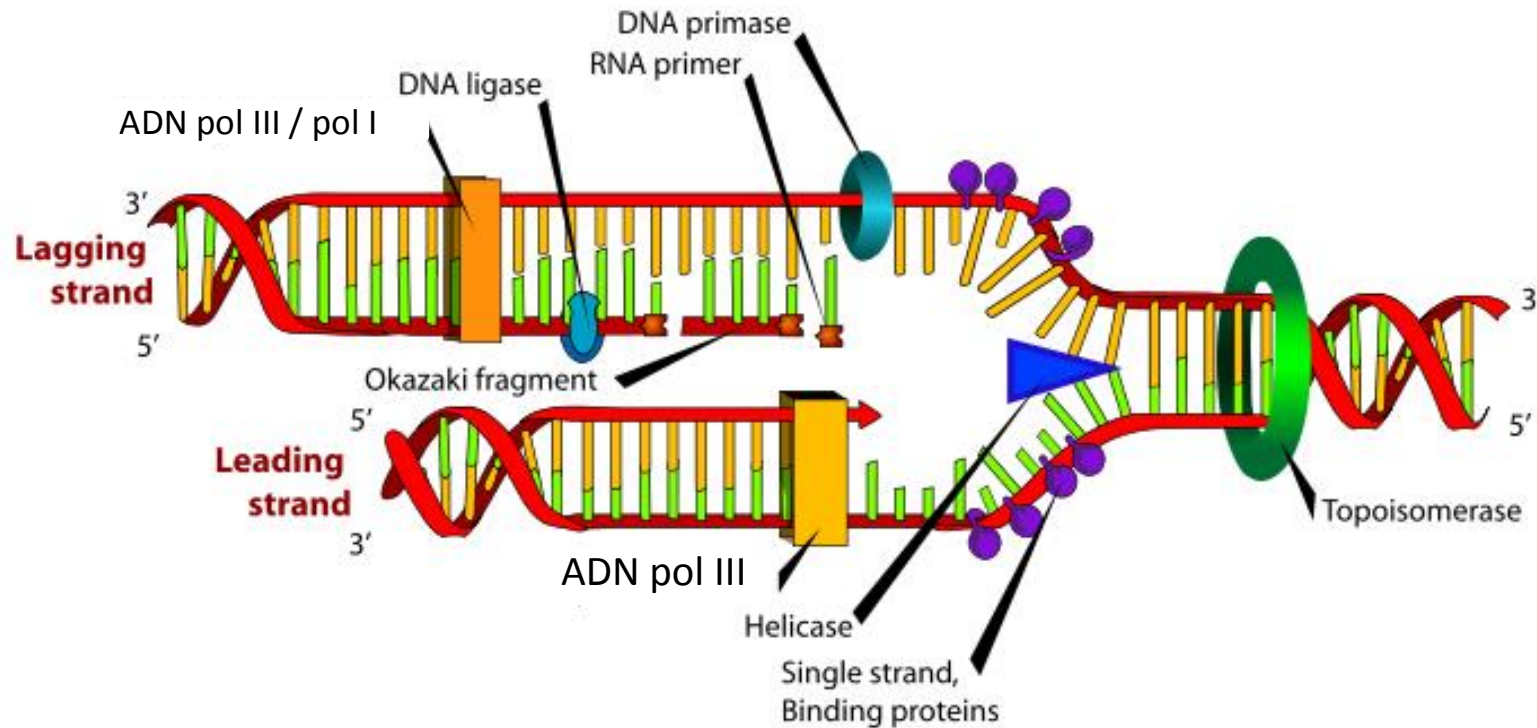


### 2 Elongation du brin précoce dans le sens de déplacement de la fourche :

Le brin qui servira de matrice au brin précoce est lu dans le même sens que l'avancée de la fourche (de 3' vers 5'). Au niveau de l'origine de réplication les ADN polymérases nécessitent une amorce qui sera mise en place par les primases (ARN). L'ADN polymérase III sera responsable de l'initiation et de l'élongation du brin précoce.



## II - RÉPLICATION



### 3 Elongation du brin tardif dans le sens inverse du déplacement de la fourche :

Sa synthèse est segmentée en fragments de taille relativement constante à chaque fois que le brin matriciel sera assez « découvert » et ainsi on respectera le sens d'élongation  $5' > 3'$ .

**Fragments d'Okasaki** (eucaryotes de 100 à 200 pnb et procaryotes de 1000 à 2000 pnb). A chaque segment : recrutement d'une primase pour la synthèse d'une amorce d'ARN constituée de 10 à 50 nucléotides selon l'espèce. Par la suite les amorces vont être détruites par l'activité nucléase de l'ADN pol I, qui va ensuite compléter la brèche. La dernière liaison phosphodiester entre l'extrémité 5' du premier fragment et l'extrémité 3' du deuxième fragment sera réalisée par la ligase.

## ❖ ADN polymérases procaryotes

**Pol I :** réparation et réplication de l'ADN.

- activité polymérase 5' → 3', (synthétise l'ADN à la place des amorces ARN pendant la réplication)

- activité exonucléase 5' → 3' (élimination des amorces ARN pendant la réplication)

- activité exonucléase 3' → 5' (proof reading)

peu processive : se dissocie de la matrice après l'ajout de 20 nt

« **Fragment de Klenow** » (polymérase 5' → 3' + activité exonucléase 3' → 5') utilisé en génie génétique.

**Pol II :** réplication de l'ADN endommagé (activité polymérase 5' → 3' et exonucléase 3' → 5').

peu processive.

**Pol III :** Principale polymérase bactérienne qui intervient dans la réplication (brin précoce + fragments d'Okazaki du brin tardif)

Très processive.

**Pol IV :** ADN polymérase de la famille Y = Translétion synthèse (TLS) polymérases.

**Pol V :** ADN polymérase de la famille Y.



## Structure de la DNA Pol III holoenzyme : dimère asymétrique

> 10 sous-unités (core + clamp + protéines associées).

**Core enzyme** = Assemblage des 3 sous unités principales (catalytiques) , , .

= activité polymérase ( 5' 3'),

= exonucléase 3' 5'

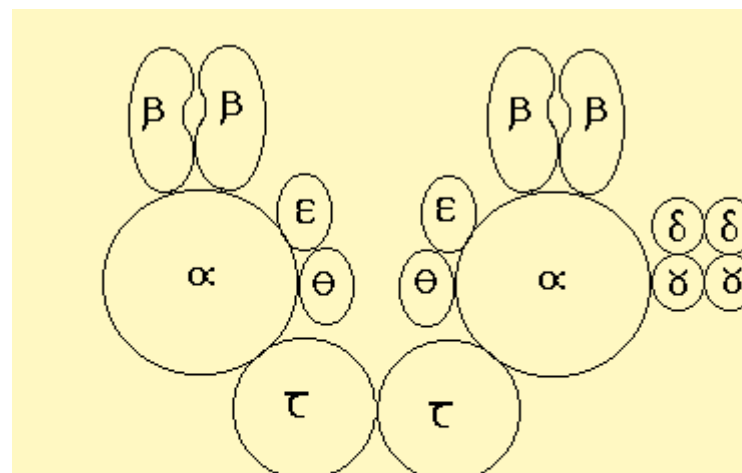
augmente l'affinité entre les sous unités du core

=> 1 core enzyme / brin matrice.

2 se fixent fortement au core de façon à le lier au brin

(psi = activité inconnue)

Partie très processive  
=> brin précoce



Partie moins processive  
=> brin tardif

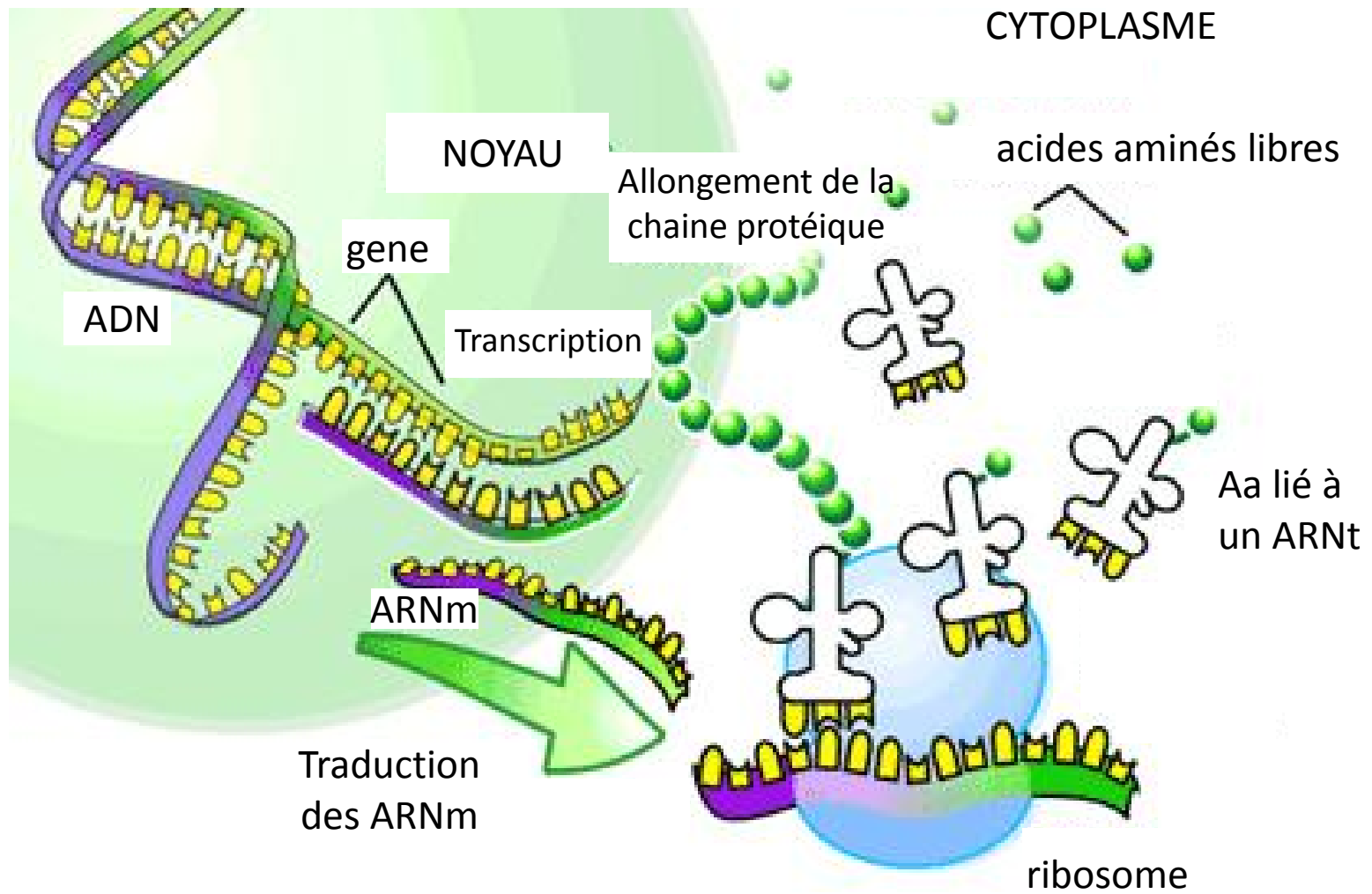
## ❖ ADN polymérase eucaryotes

ADN pol.	Nombre de sous unités	Activités
<b>Alpha</b>	4	<p>1 polymérase 2 primases 1 pour le maintien de la structure</p> <p>alpha n'a pas d'activité exonucléasique 3'→5'</p>
<b>Delta</b>	2	<p>1 activité exonucléasique 3'→5' pas d'activité primase, processivité faible</p>
<b>Epsilon</b>	2	<p>1 activité exonucléasique 3'→5' Pas d'activité primase, forte processivité</p>
<b>Béta</b>	1 >	<p>impliquée dans la réparation de courts fragments d'ADN</p>
<b>Gamma</b>	1	<p>1 activité exonucléasique 3'→5' responsable de la réplication de l'ADN mitochondriale dont la réplication est indépendante de l'ADN nucléaire</p>

## ❖ Différences au niveau de la réplication

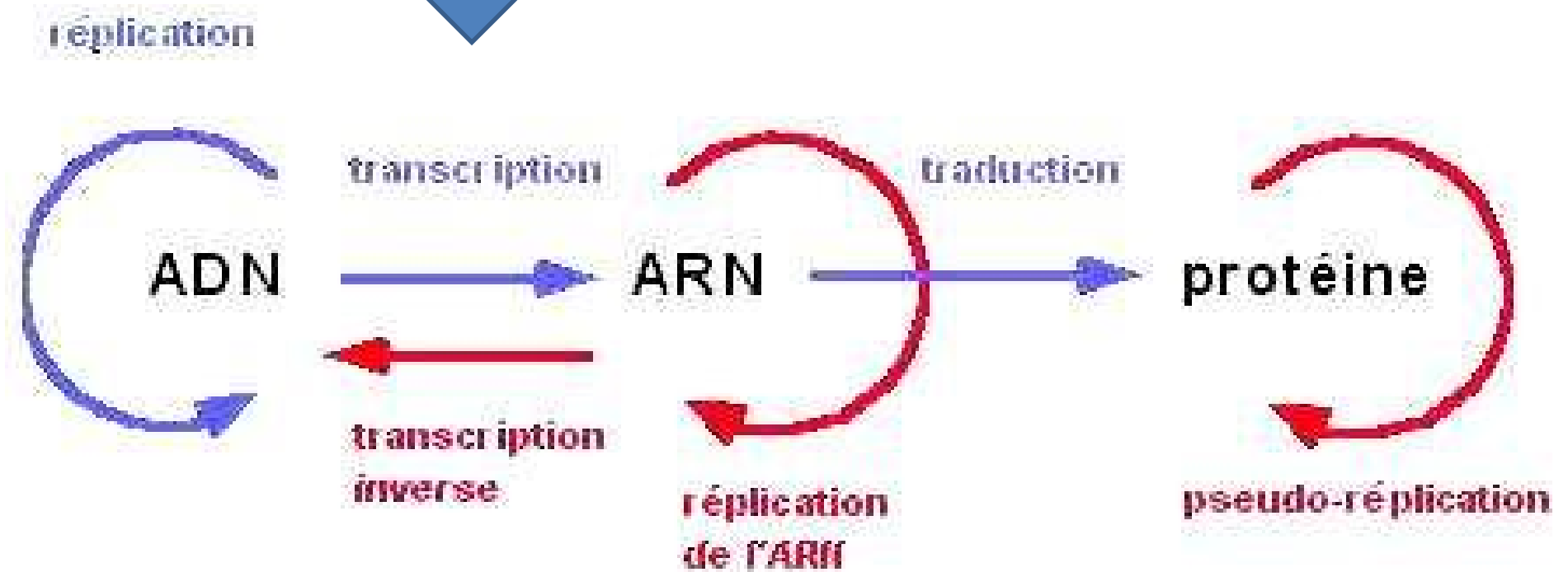
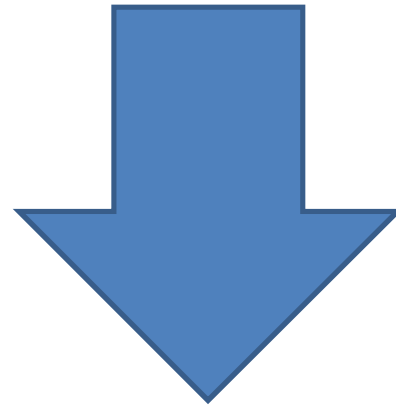
	<b>Eucaryotes</b>	<b>Procaryotes</b>
ORI	Multiple (2 kb)	Unique (200 pb)
Vitesse de réplication	50 nt/ sec	500 nt/ sec
Protéine stabilisatrice de l'ADN SS	RP-A ou RF-A	SSB
Amorce ARN	oui synthétisée par alpha suit par alpha ou bêta ou epsilon dégradé par RNase H trou bouché par alpha ou bêta	oui synthétisée par une primase suit par ADN pol III dégradé par l'ADN pol I trou bouché par l'ADN pol I
ADN pol	<b>alpha:</b> 1 polymérase 2 primases pas d'activité exonucléasique 3'→5' <b>delta:</b> activité exonucléasique 3'→5' pas de primase <b>epsilon:</b> activité exonucléasique 3'→5' pas de primase <b>bêta:</b> répare les courts fragments d'ADN <b>gamma:</b> réplique l'ADN mitochondrial activité exonucléasique 3'→5'	<b>ADN pol III:</b> 1 polymérase activité exonucléasique 3'→5' <b>ADN pol I:</b> activité exonucléasique 5'→3' 1 polymérase activité exonucléasique 3'→5'
Fragments d'Okazaki	200 nt	2000 nt
ADN ligase	I, II ou III	ADN ligase

# III - DU GÈNE À LA PROTÉINE



# III - DU GÈNE À LA PROTÉINE

## III - 1 TRANSCRIPTION



# III - DU GÈNE À LA PROTÉINE

3.1 Transcription = Copie des régions codantes de l'ADN en ARN

## 3.1.1 Caractéristiques

- La transcription (ou synthèse de l'ARN) est le processus par lequel l'information contenue dans la séquence des nucléotides de l'ADN est transférée à l'ARN
  
- ARN polymérase = enzyme qui synthétise de l'ARN par copie de l'ADN ou ARN (ARN polymérase ADN (ou ARN) dépendante) .
  
- Vitesse de l'ARN polymérase : 30 b / sec  
Erreur : 1 / 250 pb  
Possibilité de réparer les erreurs
  
- ATCG > AUCG

- ❑ **ARNm** ARN messenger code les protéines
  
- ❑ **ARNr** ARN ribosomal forme la structure de base du ribosome et catalyse la synthèse de protéine
  
- ❑ **ARNt** ARN de transfert, à la base de la synthèse des protéines en tant qu'adaptateurs entre ARNm et les acides aminés
  
- ❑ **ARN régulateurs :**
  - ARNsn** Petits ARN nucléaires (small nuclear) agissent sur différents processus nucléaires (dont épissage des pré-ARNm)
  - ARNsno** Petits ARN nucléolaires (small nucleolar), agissent pour traiter et modifier chimiquement les ARNr (maturation), ARNt , ARNsn
  - ARNmi** Petits ARN (20 à 30 nt) souvent complémentaires de la partie 3' des ANRm, inhibiteurs post-transcriptionnels
  - ARNsi** Petits ARN (2 à 25 nt DS) interférents qui arrêtent l'expression d'un gène (application : étude de la fonction d'un gène)



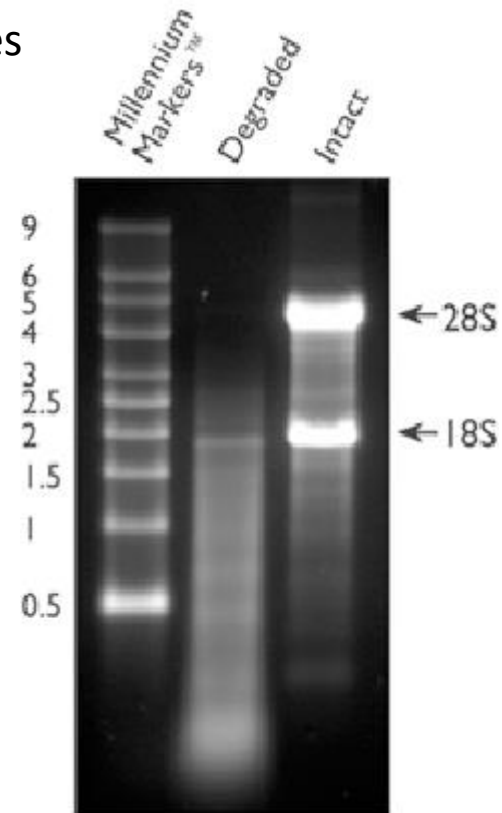
□ Abondance relative des ARN chez les bactéries

ARNr 82%

ARNt 16%

ARNm 2%

Autres ARN < 2%

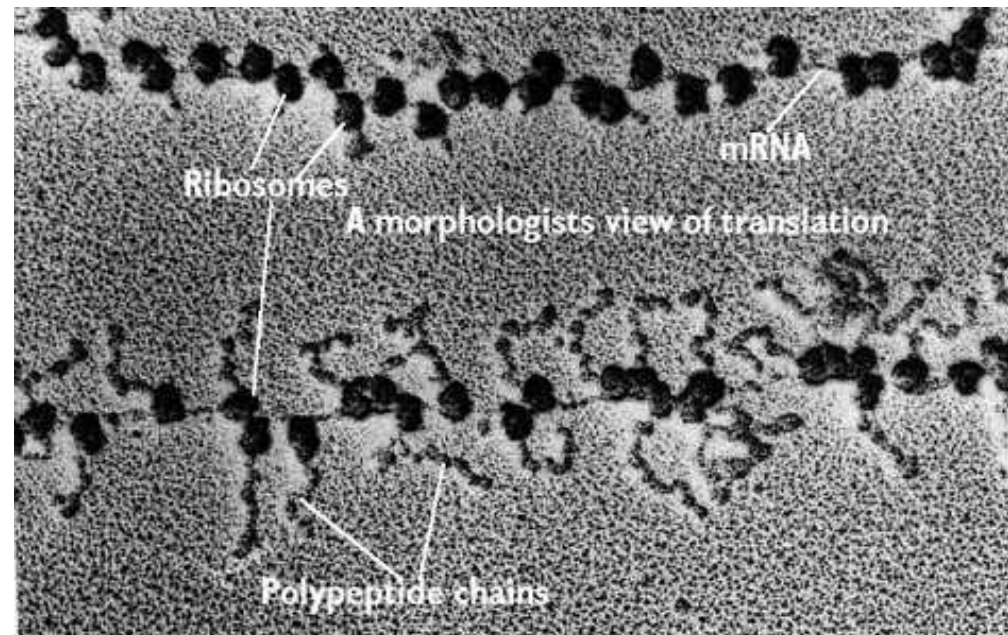
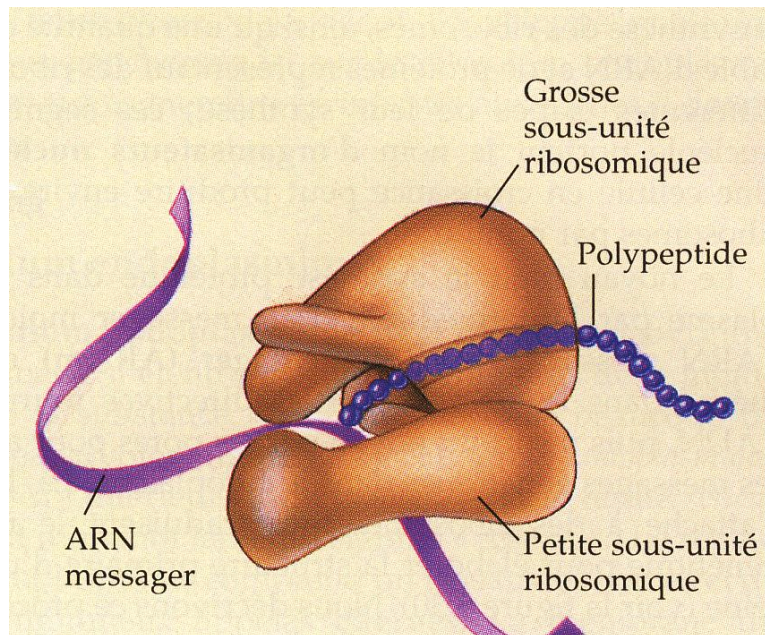


Migration d'ARN eucaryotes sur gel d'agarose

## ARN ribosomaux

	Éléments du ribosome	ARNr
Procaryotes	Petite sous unité 30S Grande sous unité 50S	16S 23S ; 5S
Eucaryotes	Petite sous unité 40S Grande sous unité 60S	18S → ARN pol I 28S ; 5S ; 5,8S

↙ ARN pol I      ↓ ARN pol III      ↘ ARN pol I



- ❑ ARN polymérase procaryote unique
- ❑ ARN polymérases chez les eucaryotes (gros complexes enzymatiques) composés de plusieurs chaînes protéiques. Elles catalysent la formation de la liaison phosphodiester.

**ARN pol I** synthétise les ARN ribosomiques

**ARN pol II** réalise la synthèse des ARN messagers, précurseurs des protéines, ARNsno et certains ARNsn

**ARN pol III** synthétise les ARN de transfert et l'ARN ribosomique 5S, certains ARNsn, autres petits ARN

Autres ARN polymérases (IV et V) semblent impliquées dans le mécanisme d'ARN interférence.

- ❑ Similitudes entre ARN polymérase procaryote et l'ARN polymérase II eucaryote

### 3.1.4 Transcription procaryote

#### □ ARN polymérase procaryote

Pas besoin d'amorces pour initier la synthèse

Grosse molécule : *E.coli* PM 480 kDa

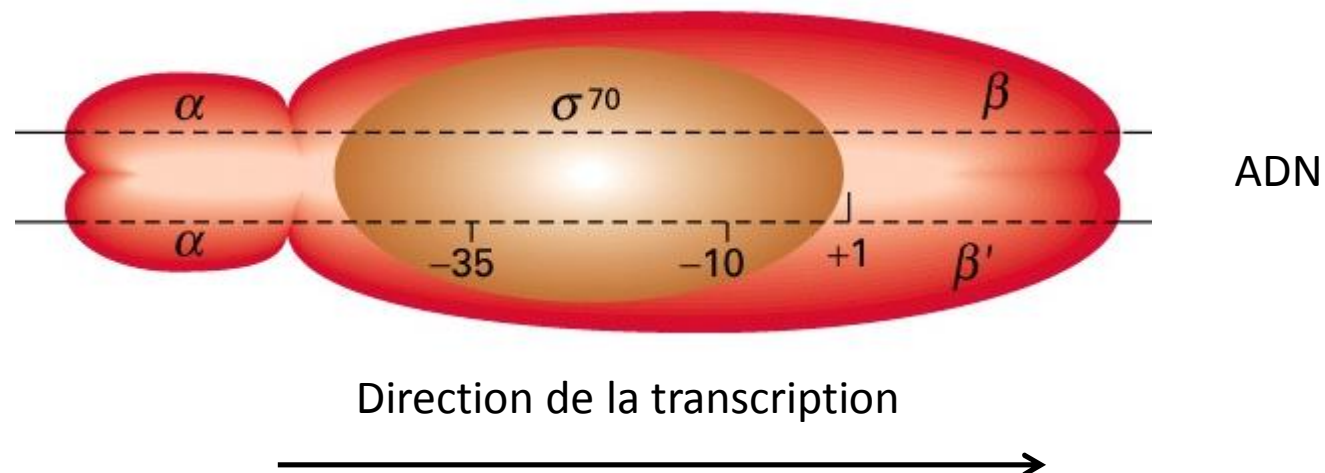
5 sous-unités polypeptidiques (cœur tétramérique) :

$2\alpha$  : reconnaissance des promoteurs

$\beta$  : fixe l'ARN

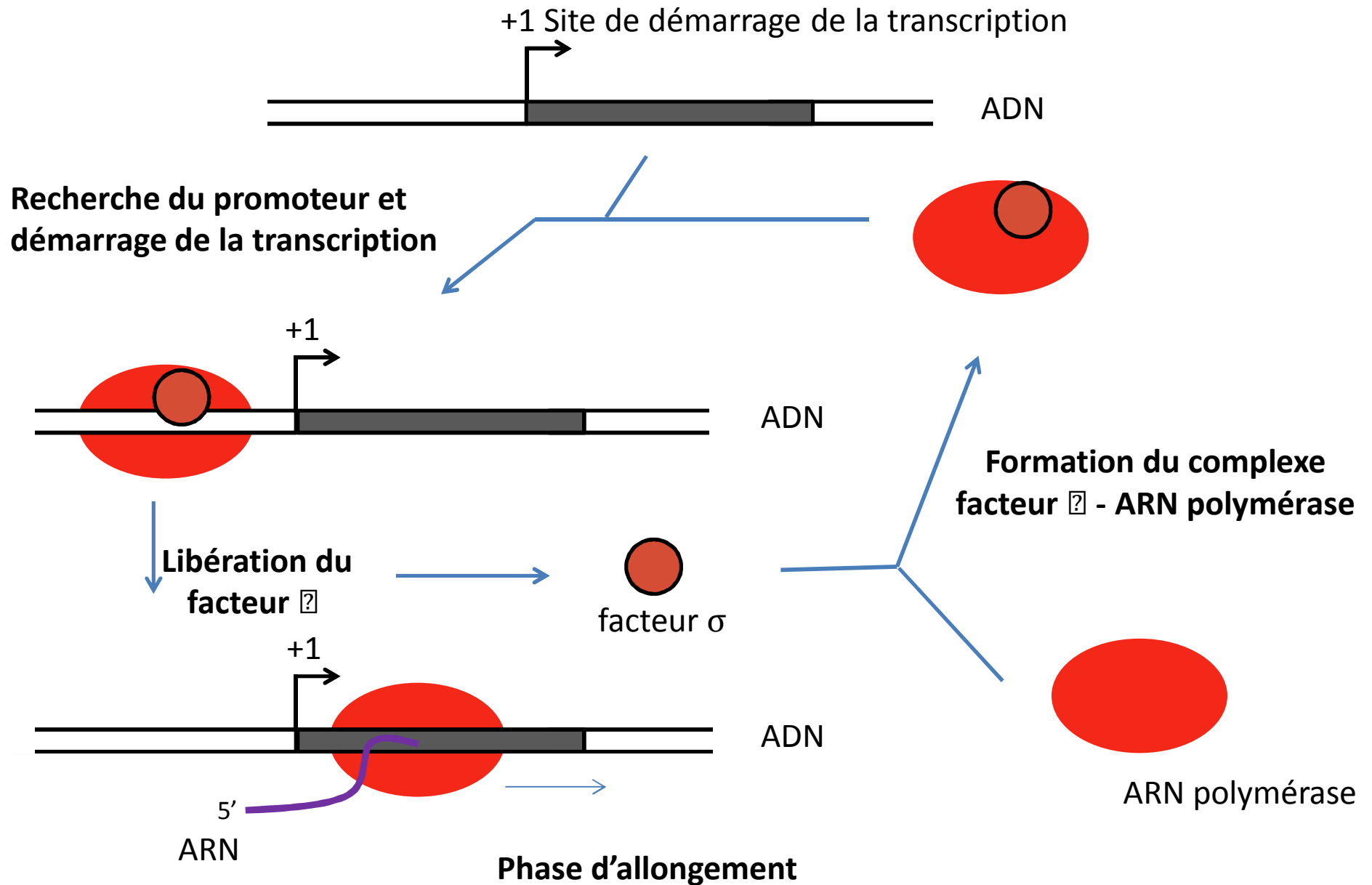
$\beta'$  : fixe l'ADN

Facteur  $\sigma$  : Pas d'activité catalytique, contrôle de la liaison ADN/ARN



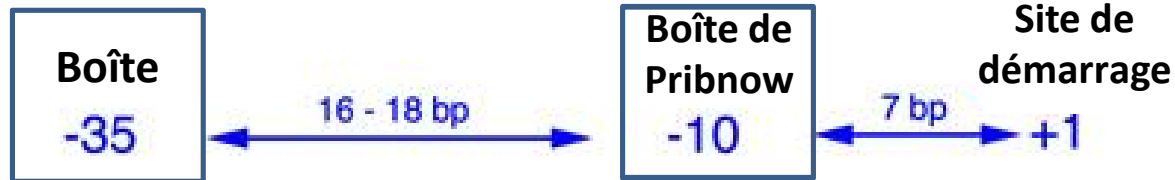
□ Initiation de la transcription chez les procaryotes :

Association / Dissociation de l'ARN polymérase et du facteur  $\sigma$  chez les procaryotes



□ Structure du gène procaryote

Env. 1 tour d'hélice => Les 2 boites interagissent en même temps avec sigma



XX TTGACAXXXXXXXXXXXXXXXXX TATAAT:XXXXXXXXX

82 84 78 65 54 45

80 95 45 60 50 96

% de conservation des bases

% de conservation des bases

Promoteur

Termineur

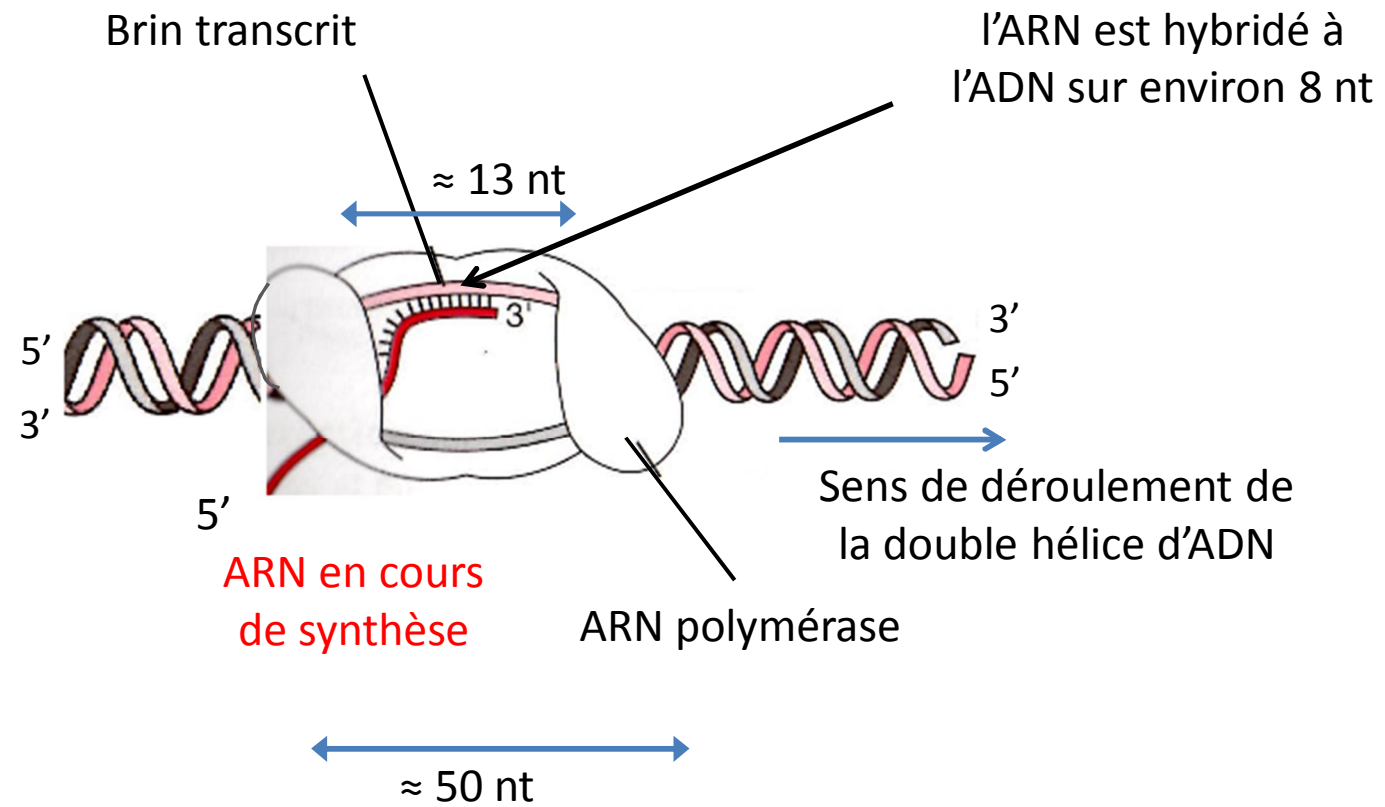


Séquences répétées riches en GC

5'cgcatttgccctccggtaggaggctttt 3'  
3'gcgaaacggaggccatccaccgaaaaa 5'

Répétition de A

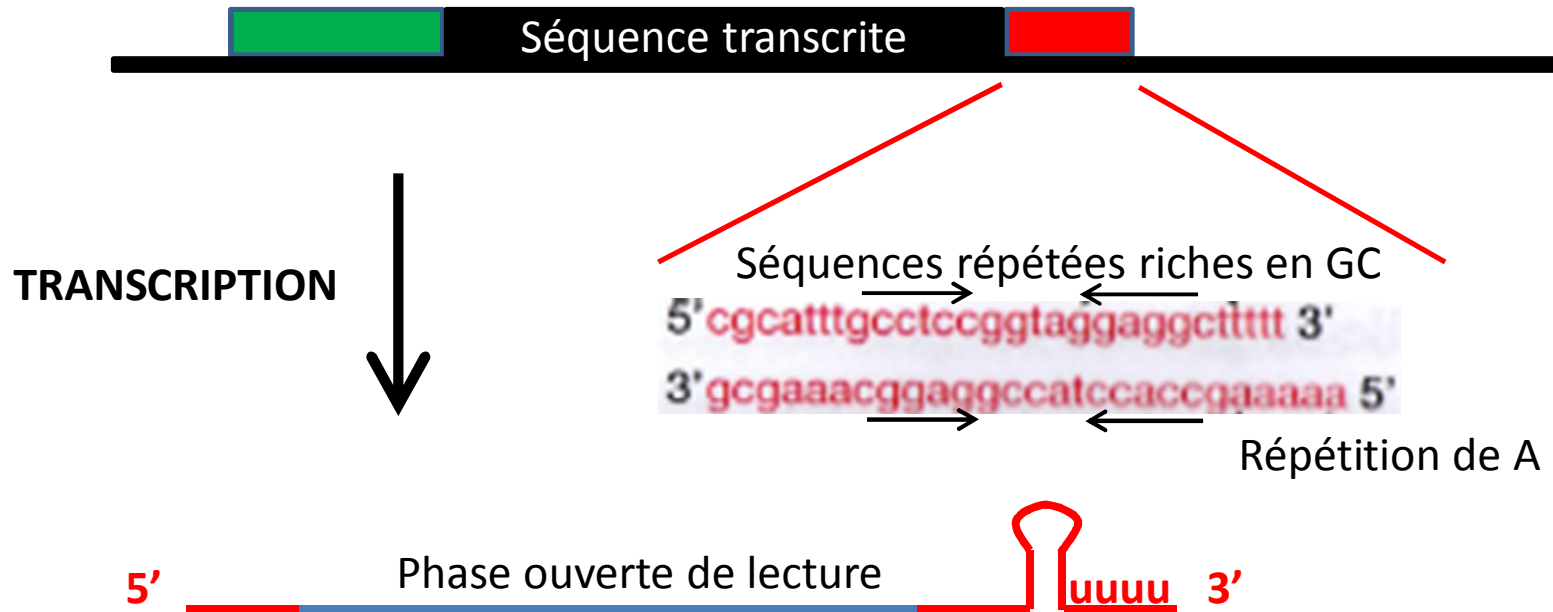
## □ Elongation de la transcription chez les procaryotes





## ❑ Terminaison de la transcription chez les procaryotes

Fin de la transcription => Dissociation du complexe ADN/ARN/ARN polymérase



### ❖ Terminaison $\rho$ indépendante

Formation d'une épingle à cheveux, formation de 4 paires A-U dans l'hybride ADN/ARN (au lieu de 8pb)

### ❖ Terminaison $\rho$ dépendante

Pas de série de A à la suite des séquences palindromiques.  
Protéine  $\rho$  a activité hélicase permet de dissocier l'ARN de l'ADN

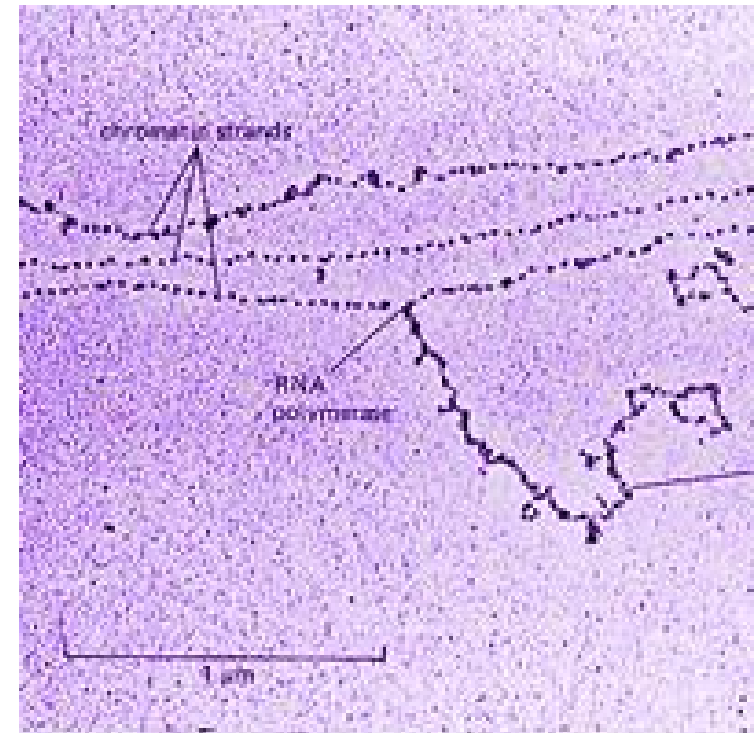
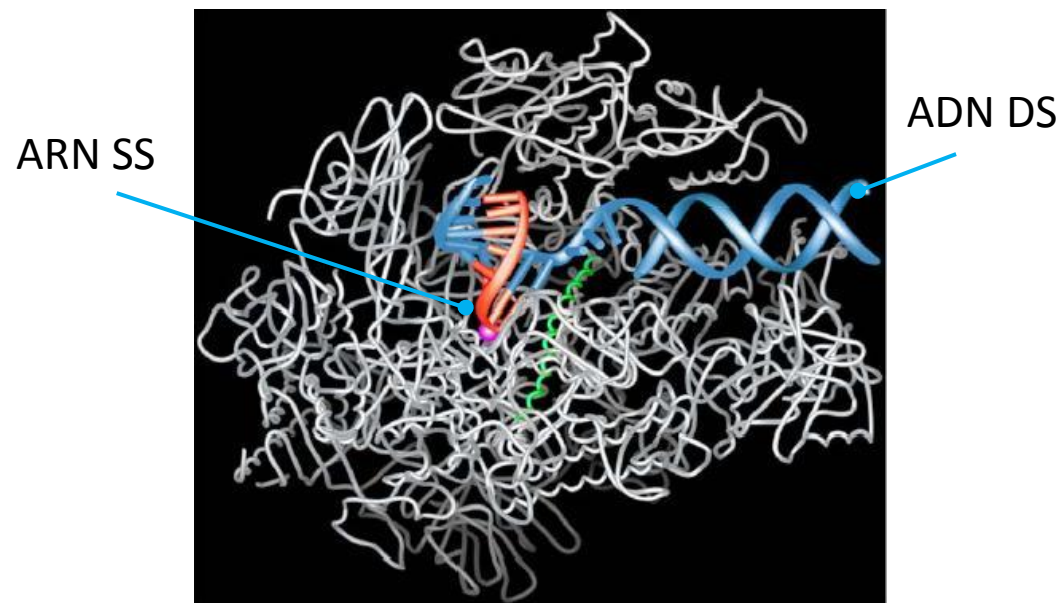
### 3.1.5 Transcription eucaryote

Fonctionnement similaire de la transcription chez les eucaryotes et les procaryotes

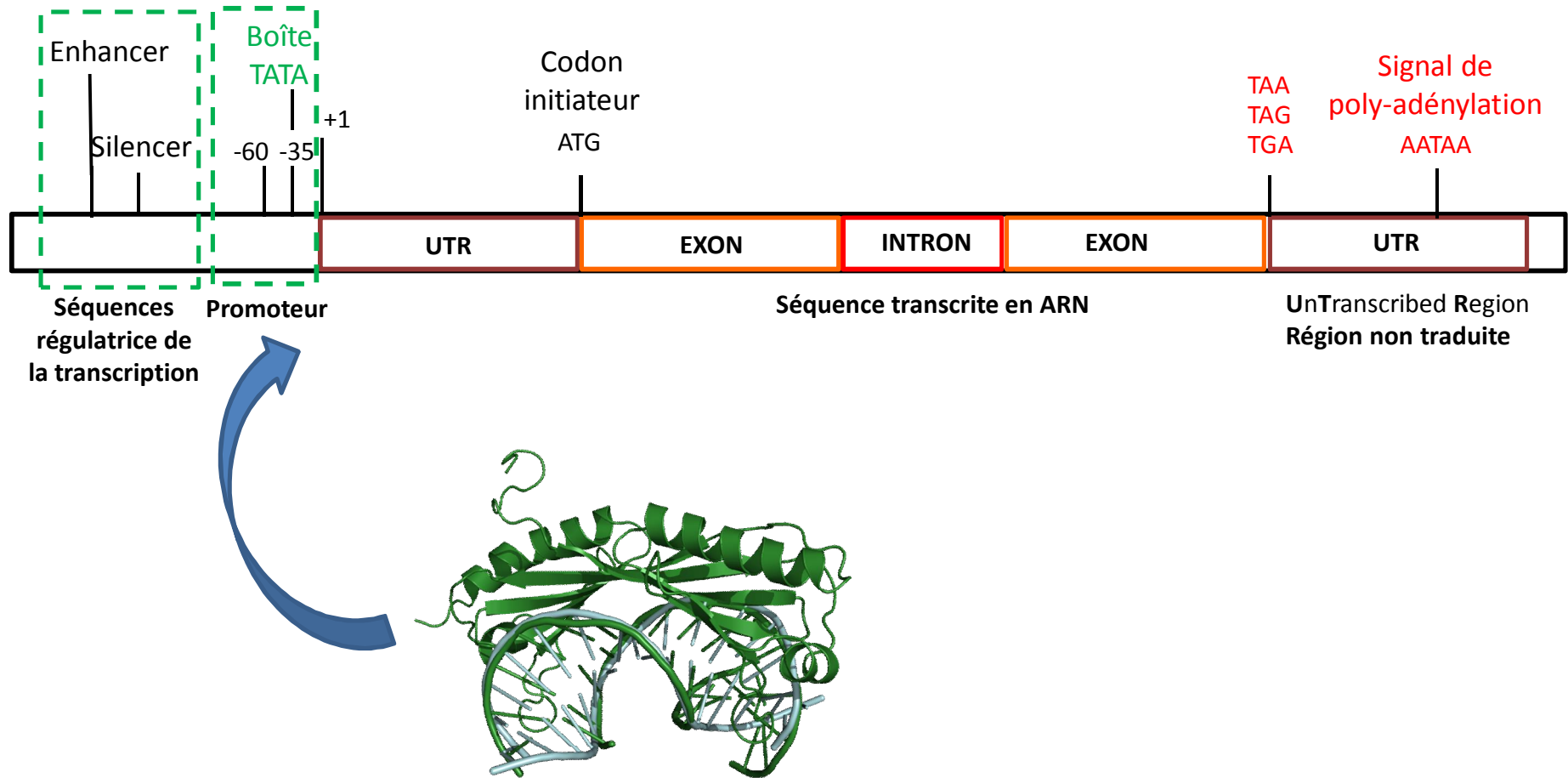
- ❖ 3 ARN polymérase
- ❖ Le nucléosome (empaquetage de l'ADN) complique la transcription

*Dans ce cours, c'est la transcription par l'ARN polymérase II qui sera présentée.*

- ❖ L'ARN polymérase II intervient dans le processus de synthèse des protéines  
Prix Nobel de Kornberg 2006  
(Modèle du complexe ARN pol II + imagerie de l'ARN pol II)



## □ Initiation de la transcription chez les eucaryotes



❖ Reconnaissance de la boîte TATA par la protéine TBP (TATA Binding protein)

Mise en place du complexe de pré-initiation de la transcription

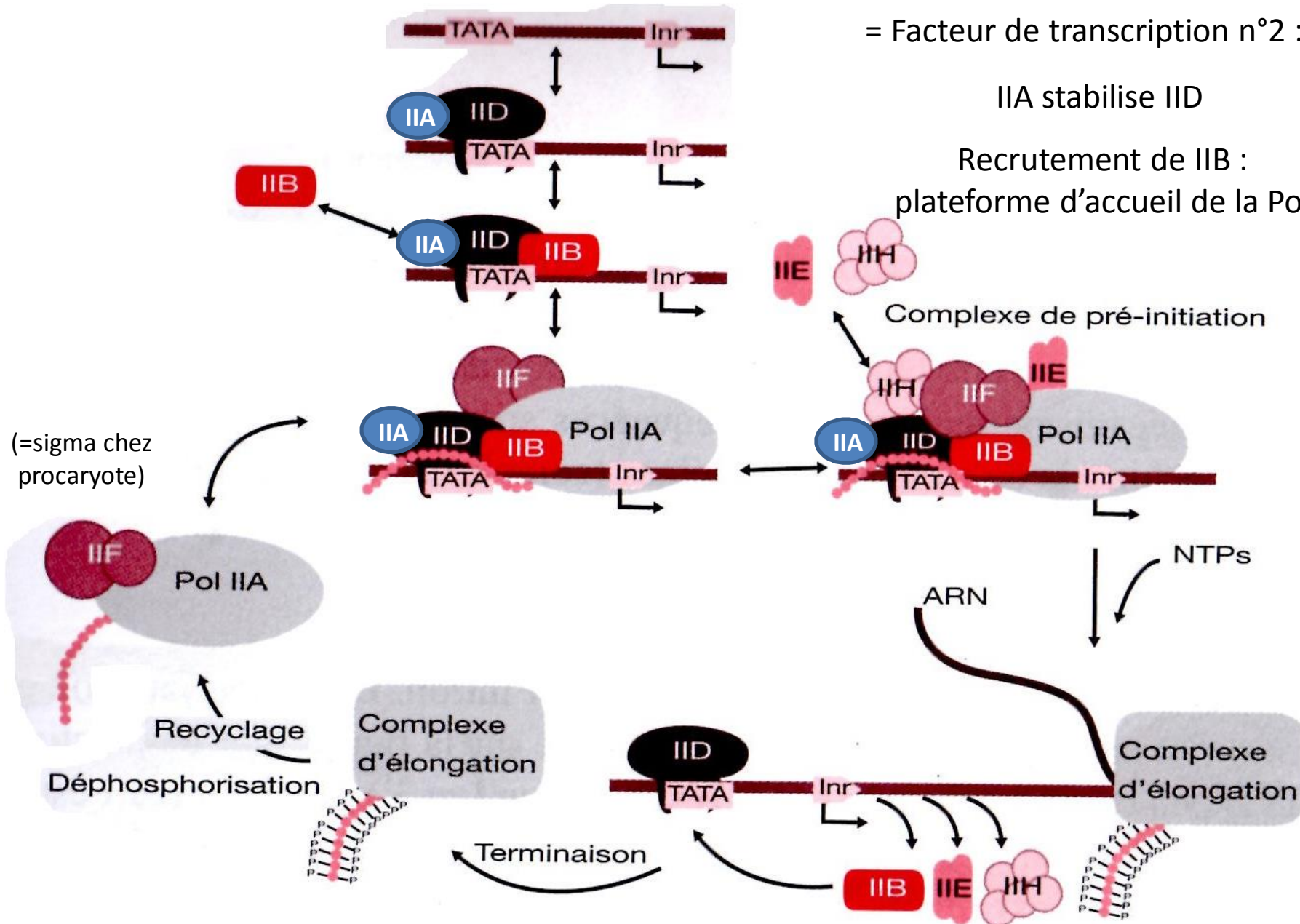
□ Formation du complexe de pré-initiation

TBP + TAF

= Facteur de transcription n°2 : IID

IIA stabilise IID

Recrutement de IIB :  
plateforme d'accueil de la Pol



(=sigma chez prokaryote)

Complexe de pré-initiation

Recyclage

Déphosphorisation

Complexe d'élongation

Terminaison

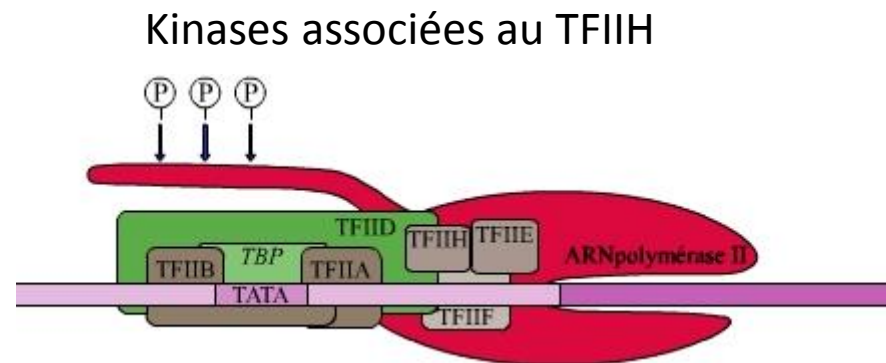
Complexe d'élongation

NTPs

ARN

## □ Elongation et terminaison de la transcription chez les eucaryotes

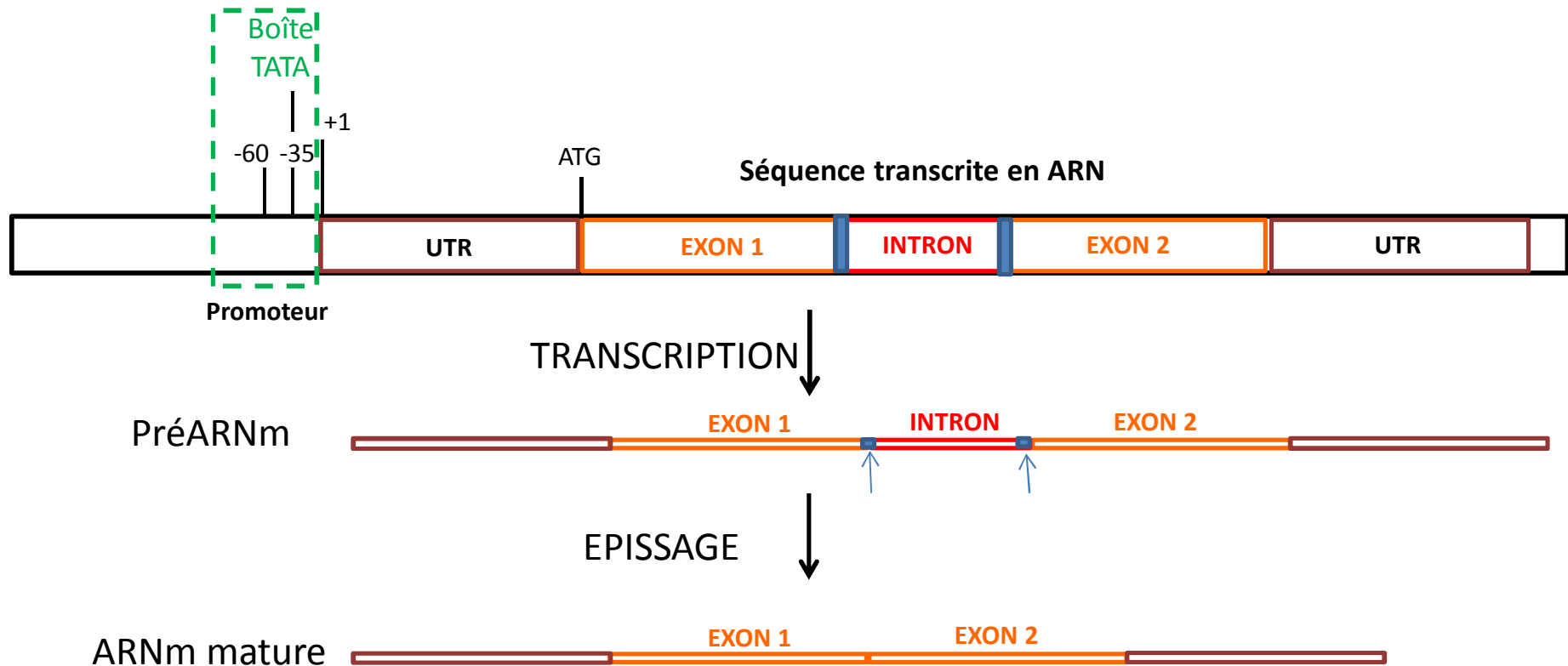
- ❖ Phosphorylation du C-terminal de la polymérase  
(répétition de motif protéique particulier)



- ❖ Pas de « terminateur ». L'ARN pol décroche plusieurs centaines de nucléotides après ce qui sera l'extrémité 3'OH des ARN

## □ Epissage des ARNm

L'épissage est l'étape de maturation des ARN au cours de laquelle une séquence interne à l'ARN est enlevée

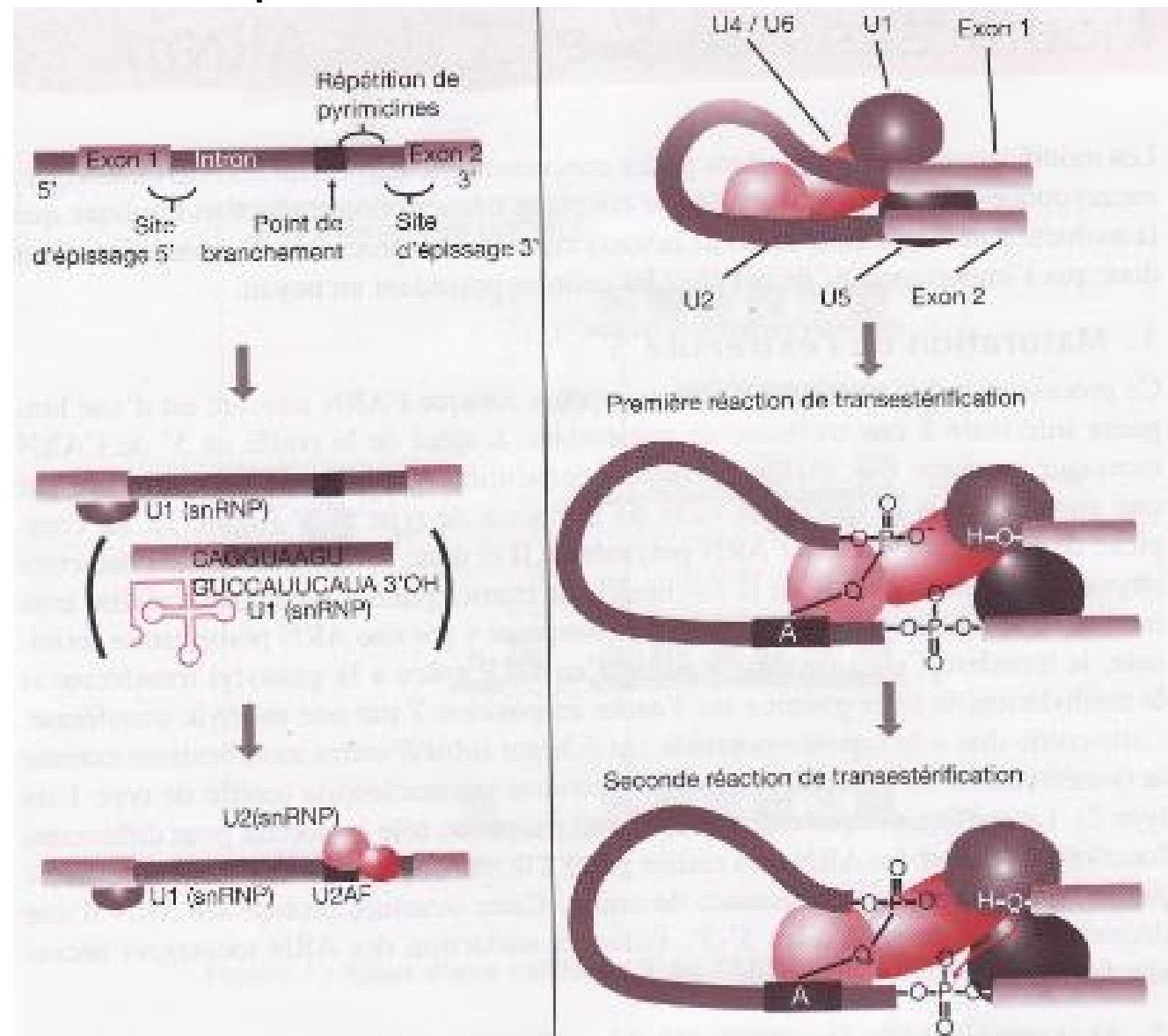


Reconnaissance de **séquences spécifiques** de l'ARN (jonction intron - exon)

par 5 ribonucléoprotéines (snRNP) : U1, U2, U4/U6 et U5

Spliceosome : complexe d'ARN et de protéines

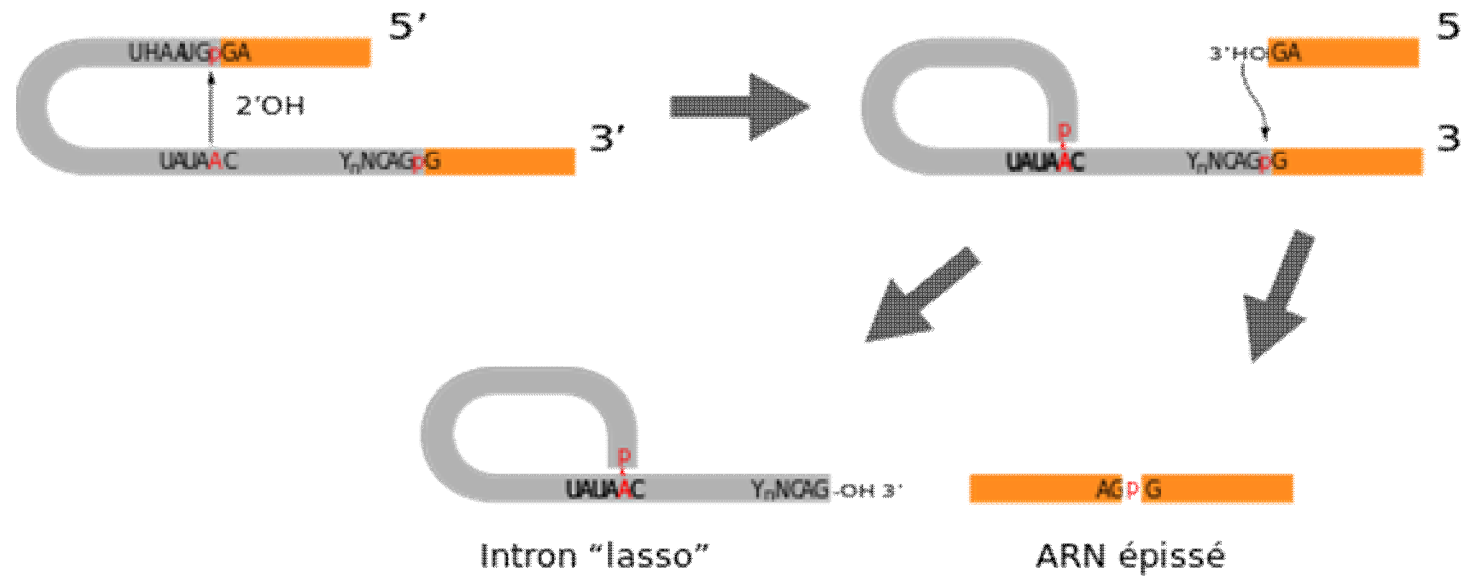
## □ Formation et action du **spliceosome**



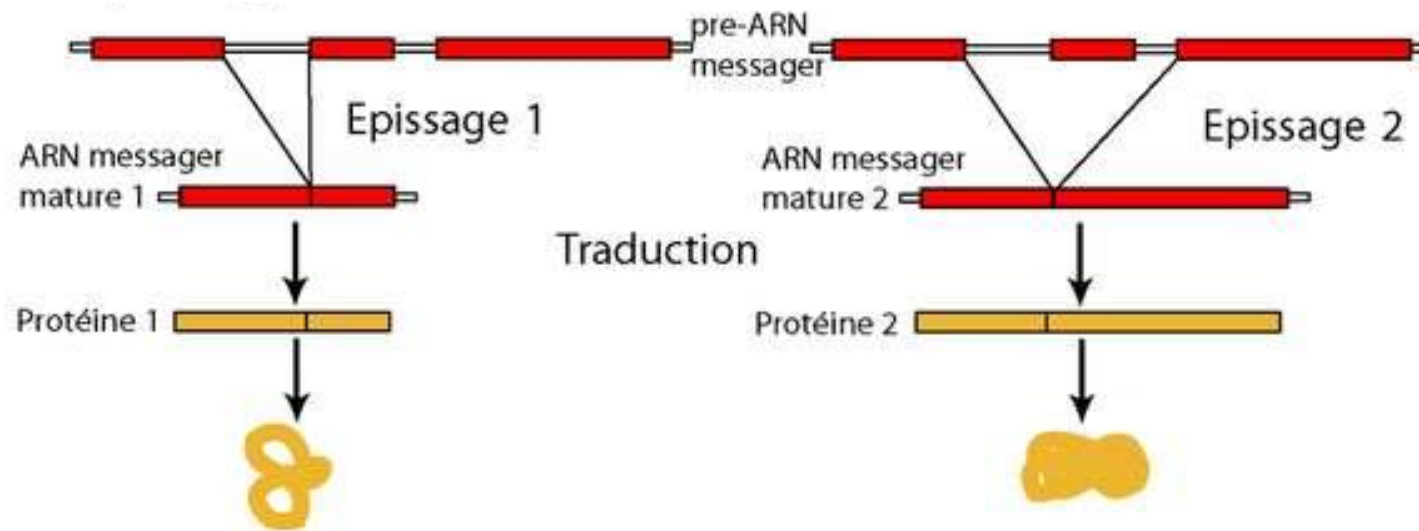
U2 (snARN) contient une pseudo-uridine qui se positionnera en face d'une adénine intronique = point de branchement



## □ Mécanisme de l'épissage



## □ Epissage alternatif des ARNm



L'intérêt pour la cellule c'est de pouvoir faire plusieurs protéines différentes à partir d'un même gène en fonction des conditions de régulation.

## ☐ Modifications post-transcriptionnelles des ARNm

### Coiffage des ARN

= Maturation de l'extrémité 5'

ARN < 30 nt

Complexe de coiffe lié à l'ARN pol

ARN phosphatase (Py 5')

Guanylyl transférase (dGTP)

Méthyl transférase (CH<sub>3</sub> sur N<sub>7</sub>

Guanine)

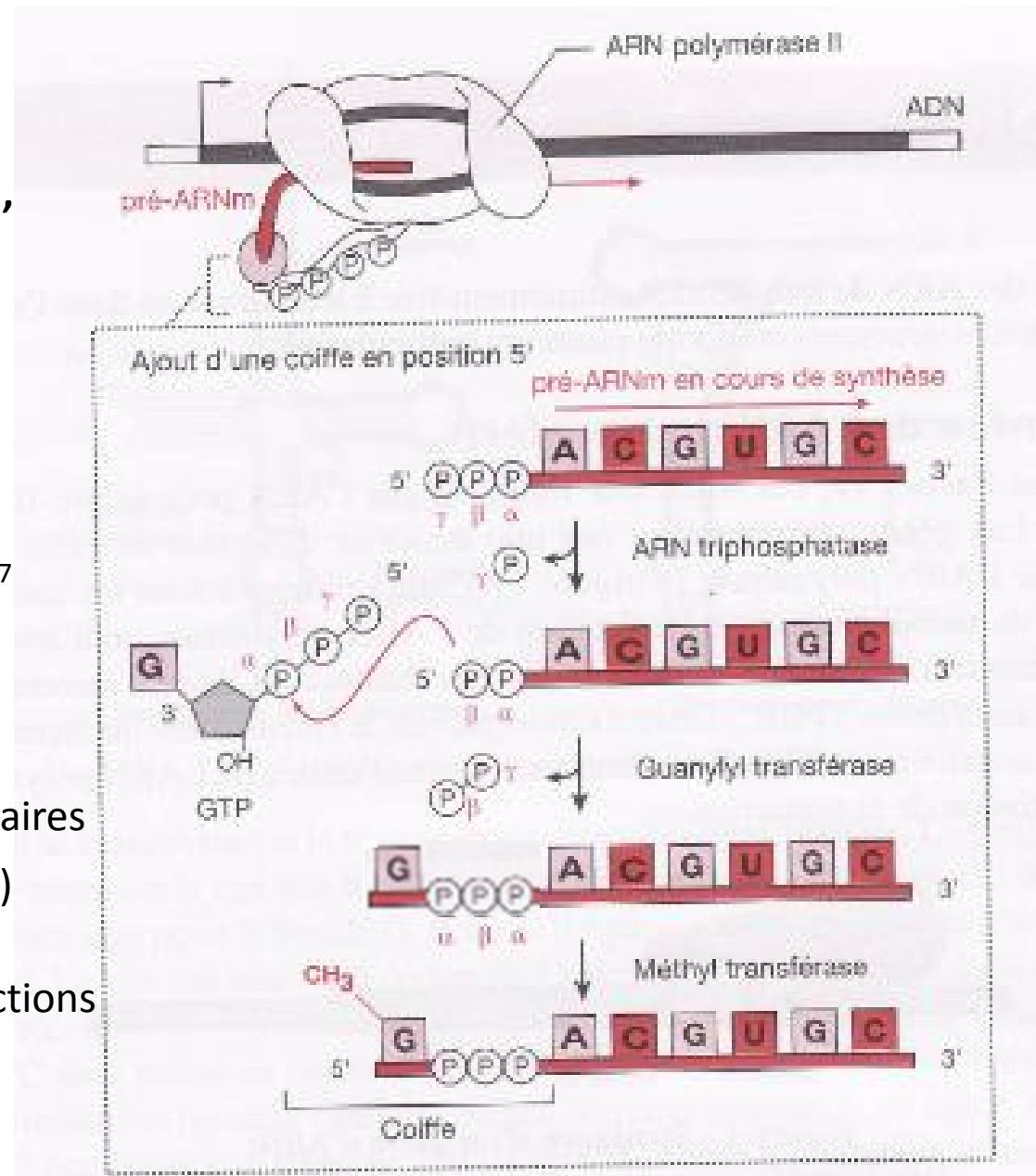
Liaison 5' - 5'

= coiffe de type 0

Possibilité de modifications supplémentaires

(coiffes de type 1 et de type 2)

- ❖ Rôle important dans différentes fonctions (export, traduction...)

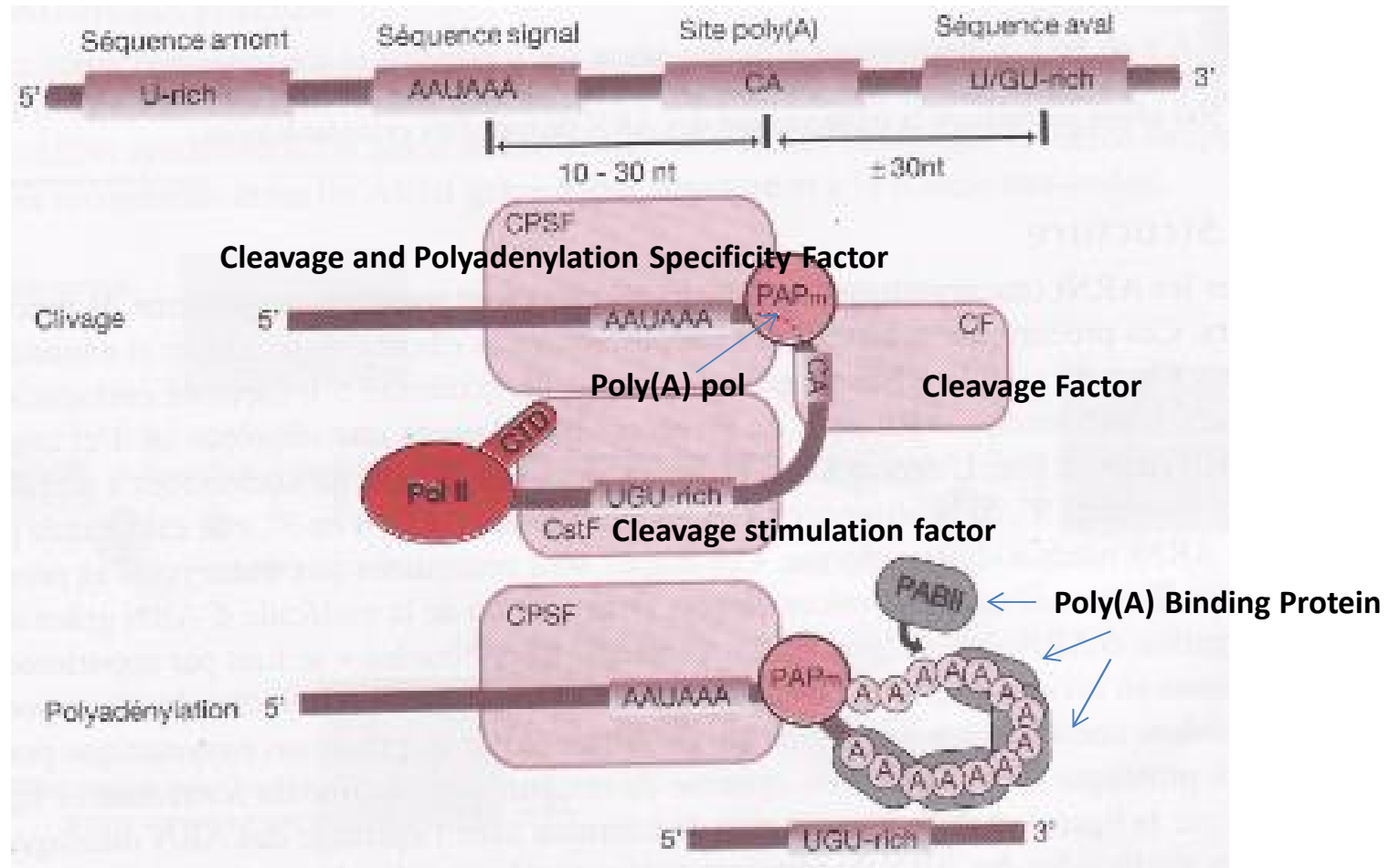


## ☐ Modifications post-transcriptionnelles des ARNm

**Polyadénylation** Ajout d'une queue poly(A)

= **Maturation de l'extrémité 3'**

Mécanisme couplé à la terminaison de la transcription



❖ Rôle important dans la stabilisation de l'ARNm et sa traduction

# **TECHNIQUES DE GENIE GENETIQUE**

## II La réaction de polymérisation en chaîne PCR

### Introduction :

Méthode basée sur la réplication de l'ADN, utilisée dans tout les laboratoires pour amplifier de l'ADN

Inventeur : K Mullis 1983, brevet 1985, prix Nobel 1993

Cette technique a révolutionné le génie génétique, elle demeure INCONTOURNABLE

+ méthode très sensible

- travailler dans des conditions de propreté extrême afin d'éviter la contamination des échantillons

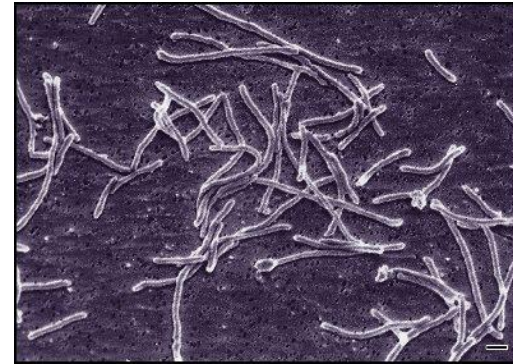
### Applications :

- Manipulation et ingénierie de l'ADN.
- Diagnostic de maladies (HIV, *Mycobacterium tuberculosis*...)
- Détection de microorganismes pathogènes
- Distinction d'espèces (bœuf / cheval)
- Détection de plantes OGM
- ...

Enzyme clé : la **Taq polymérase**

*Thermus aquaticus* :

bactérie thermostable isolée de source chaude Yellowstone Park (USA) 1969 : cocktail enzymatique adapté aux températures élevées (50-80°C)



## Principe : DÉNATURATION DE L'ADN DS

- ❑ Dénaturation par chauffage
- ❑ Plus l'ADN est riche en Guanine et Cytosine (%GC), plus la température de dénaturation est importante
- ❑ La température de fusion  $T_M$  (melting temperature) de l'ADN DS est la température pour laquelle 50 % des molécules d'ADN sont dénaturées (sous forme simple brin ADN SS (single strand)).
- ❑ La température de fusion d'une séquence nucléotidique permet de déterminer sa température d'hybridation.

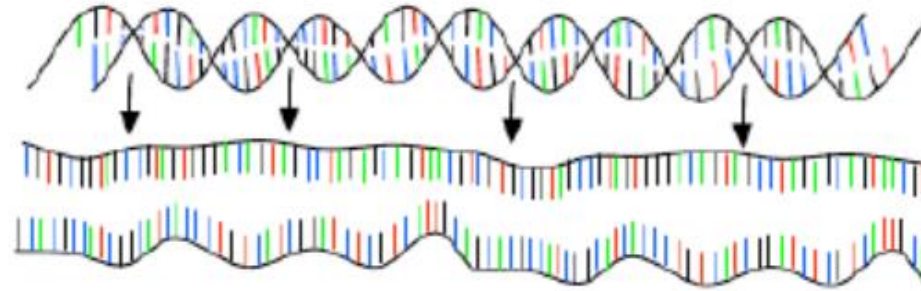


pcr.exe

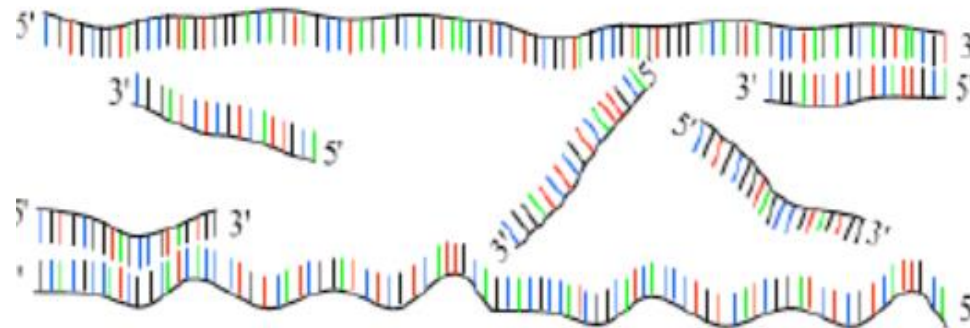


## Principe

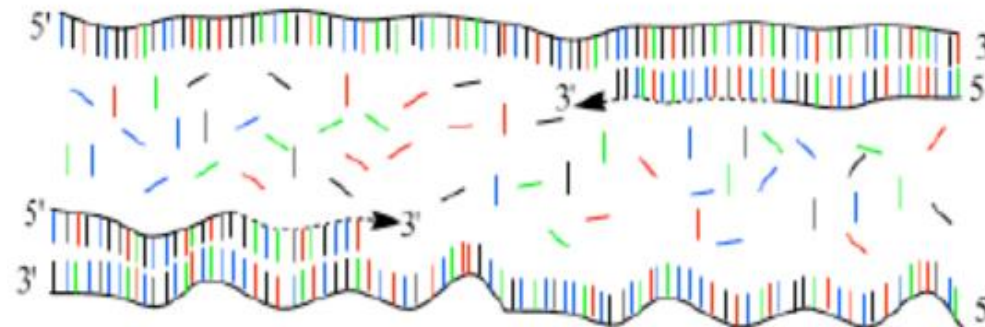
*Illustrations (Ecole de l'ADN - Christian Siatka)*



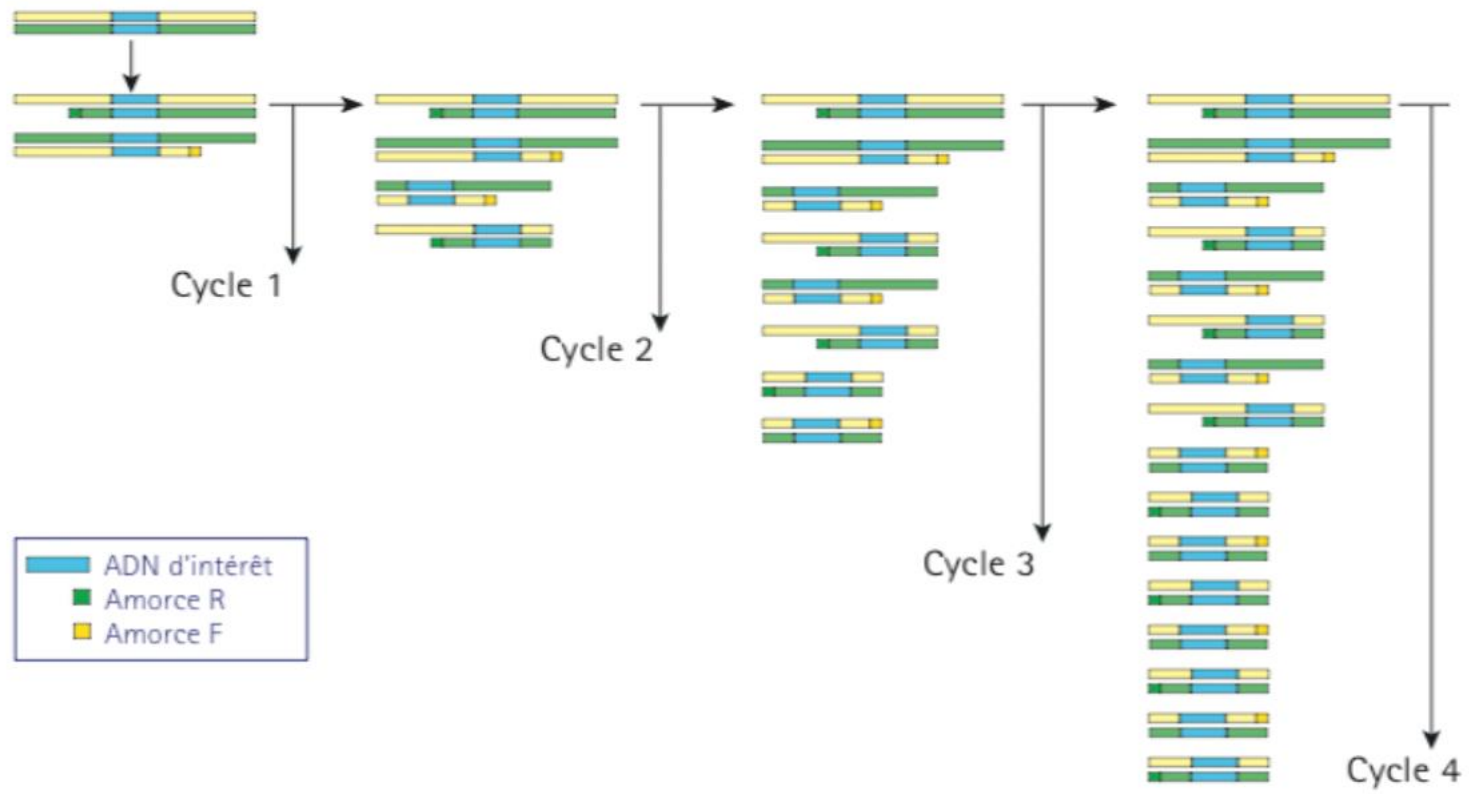
Dénaturation  
94°C



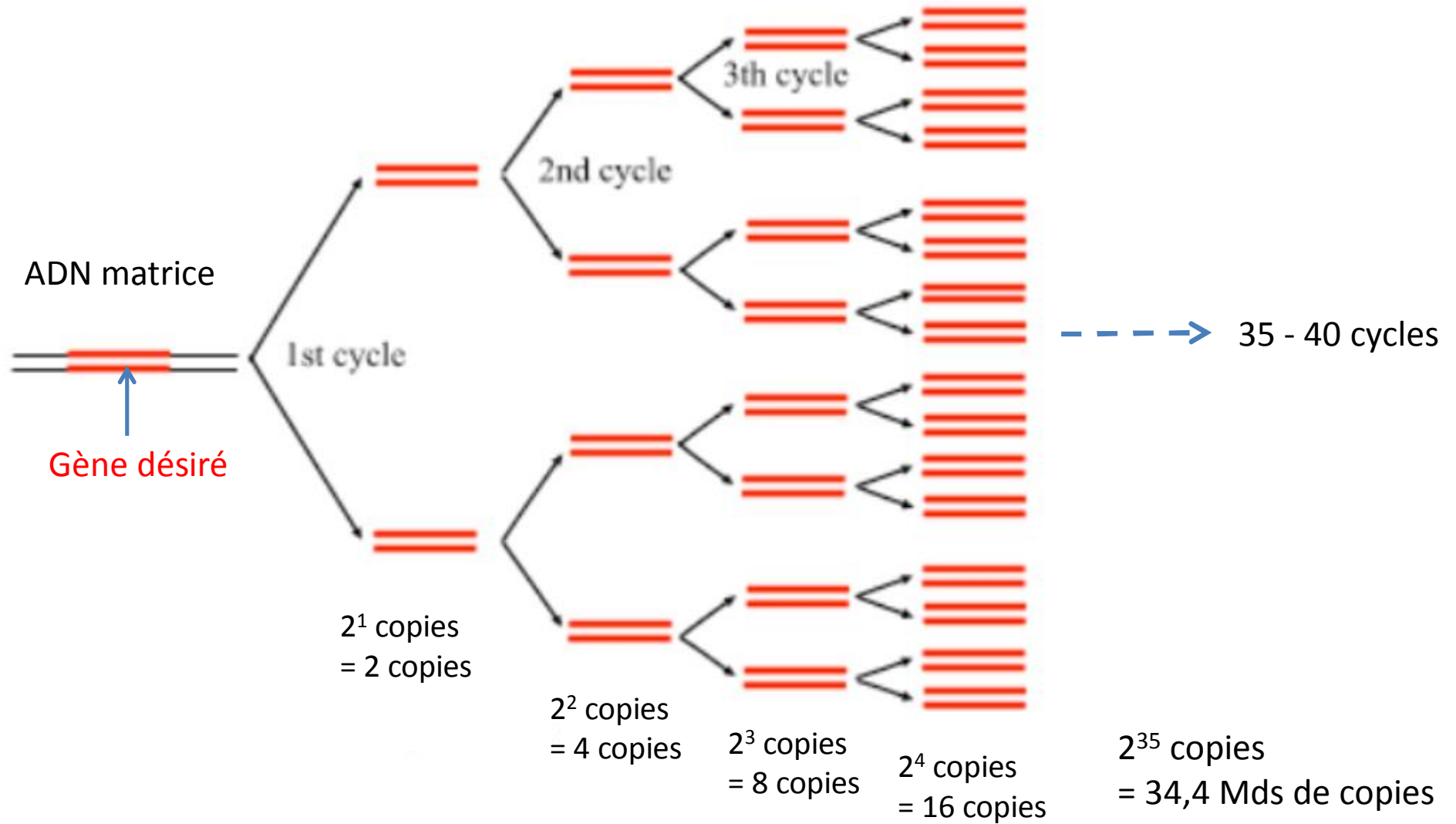
Hybridation  
55 - 65°C



Extension  
72°C



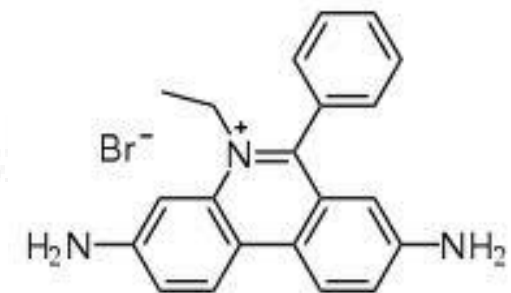
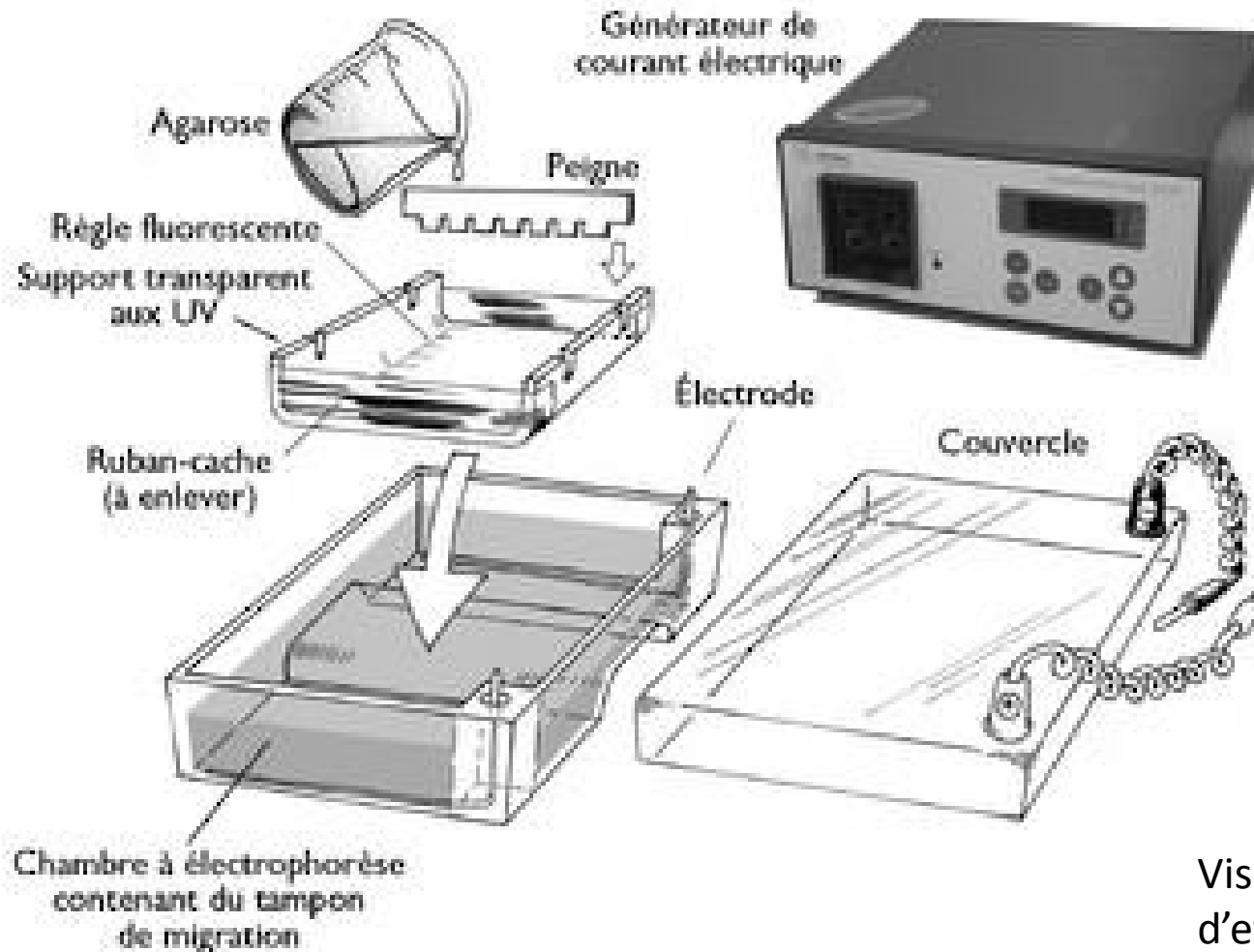
## Principe (suite)



## Principe de la séparation des fragments sur gel d'agarose

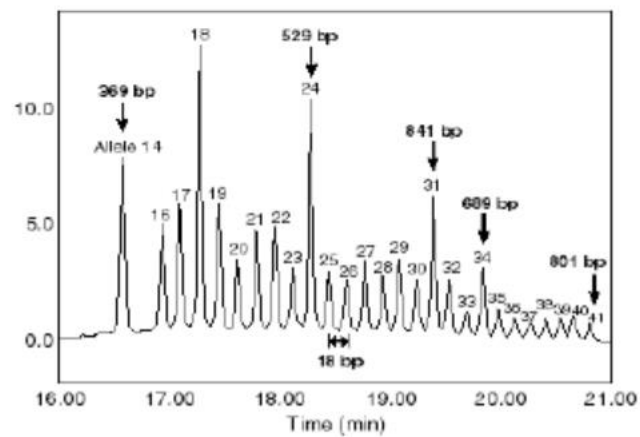
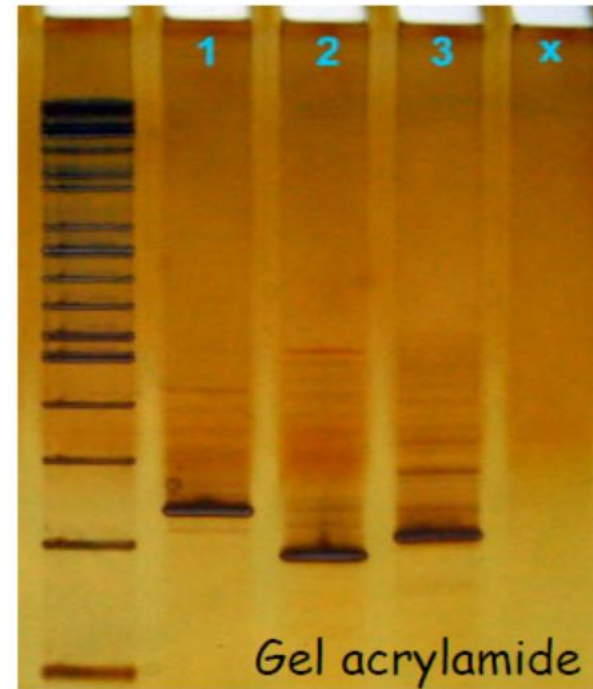
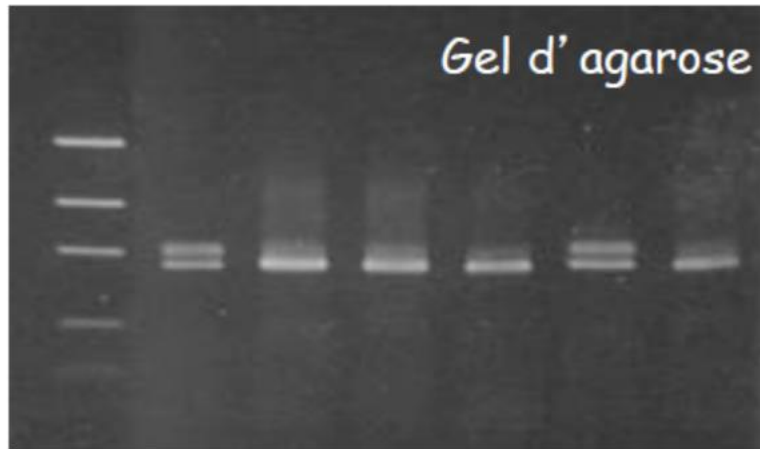
ADN chargé négativement migre vers l'électrode positive

Séparation des fragments selon leur taille (les plus petits migrent le plus loin)



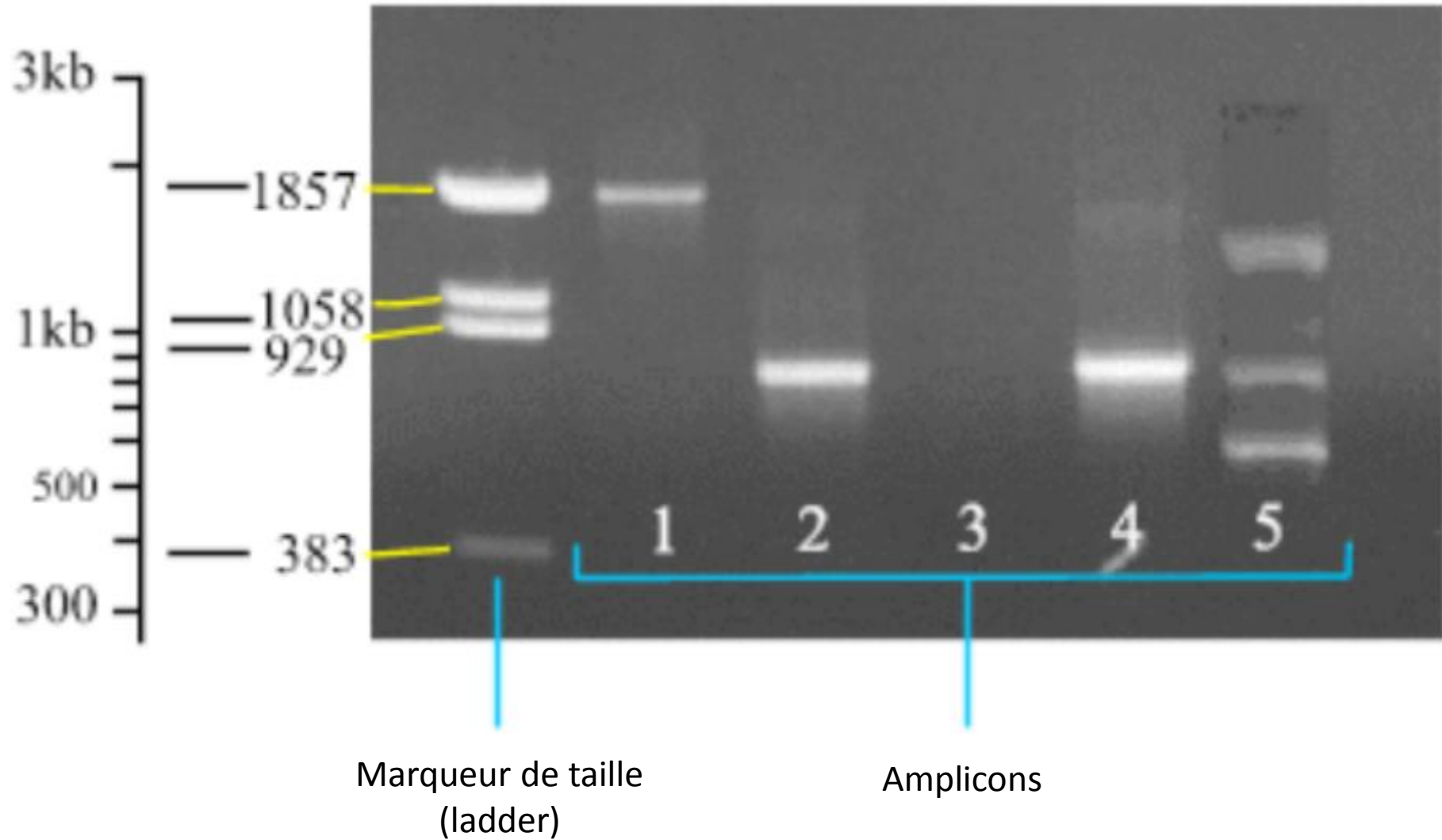
Visualisation grâce au bromure d'ethidium (BET) : intercalent de l'ADN

## Exemples de résultats (gel d'agarose, gel acrylamide et électrophorèse capillaire)



Analyseur de fragments

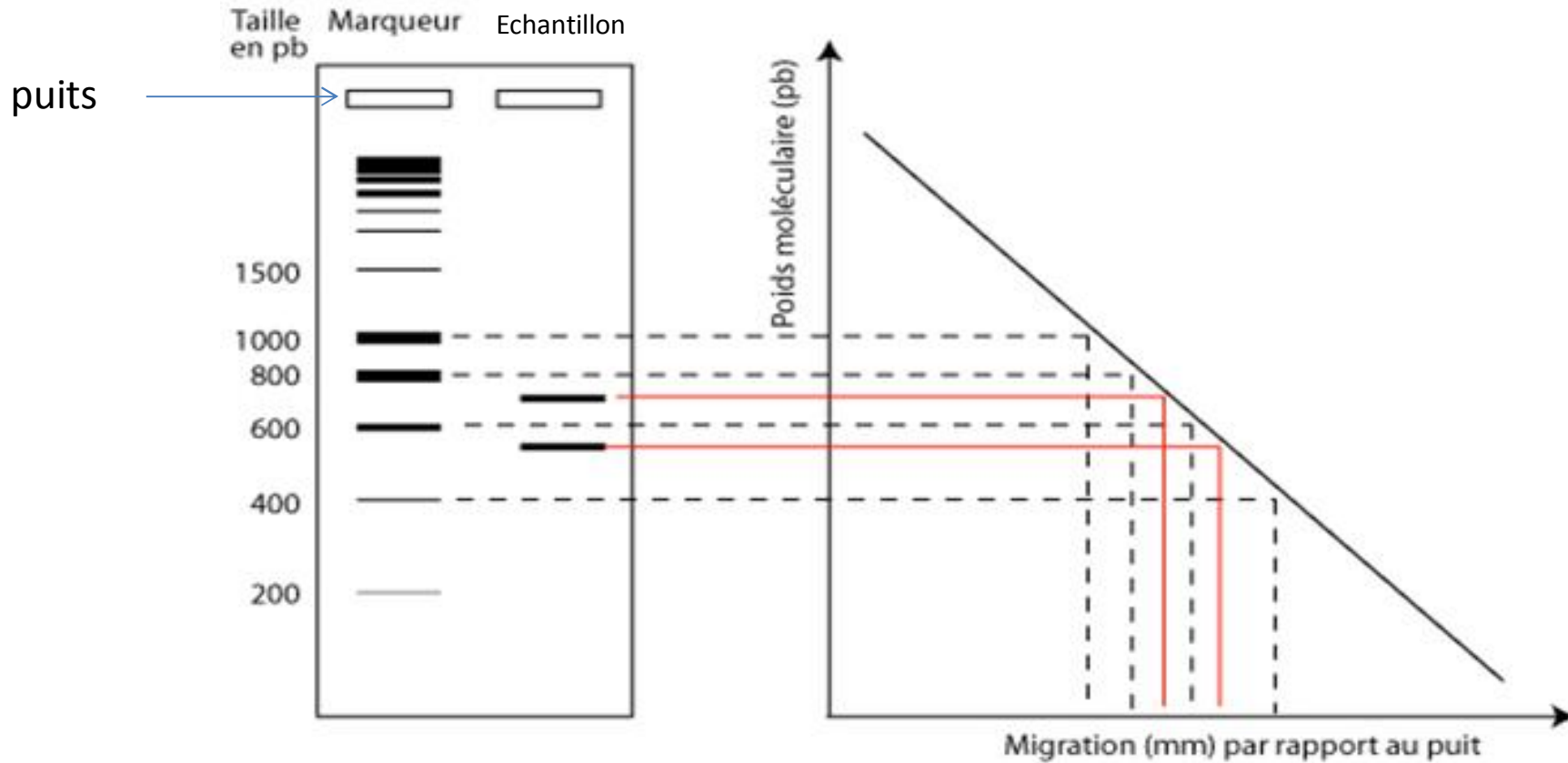
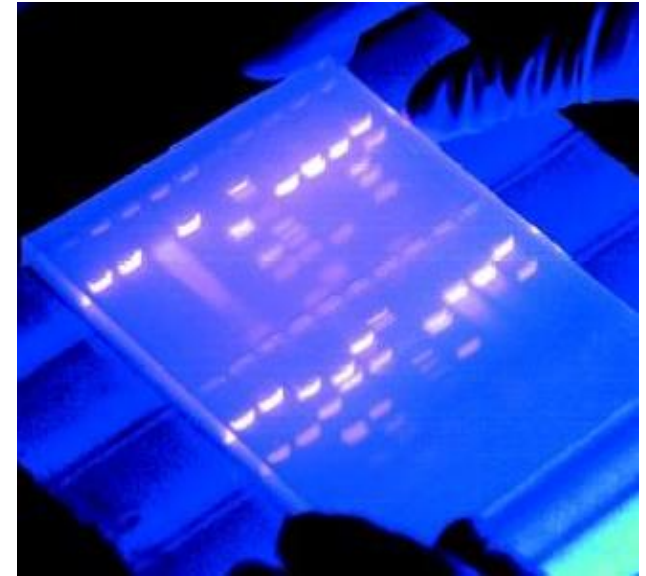
## Vérification des produits obtenus par PCR sur gel d'agarose





puits

- Dépôt de l'ADN dans les puits
- Migration de l'ADN
- Révélation de l'ADN aux UV (BET)



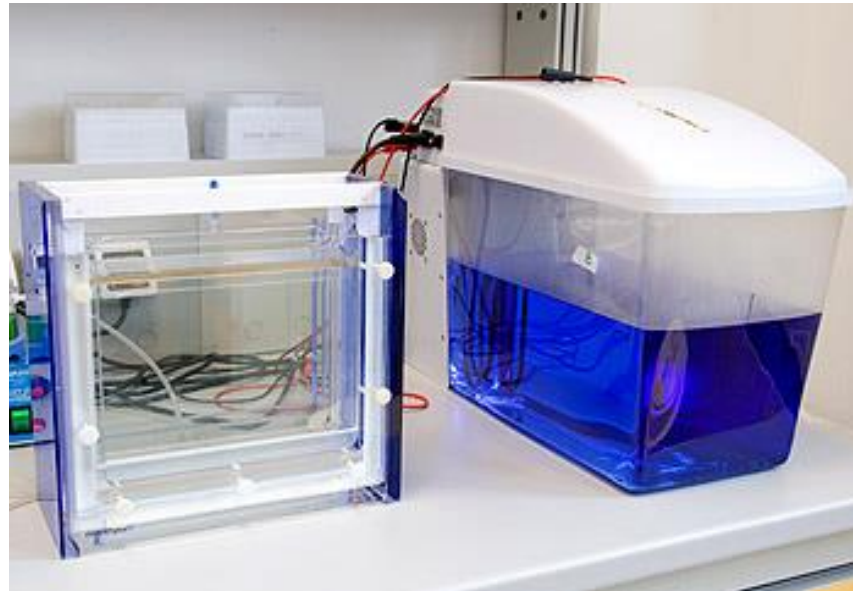
## Optimisation de la PCR

- ✓ Choix des amorces (primers), séquences et concentrations
- ✓ Température d'hybridation (en fonction de la température de fusion ( $T_M$ ) des amorces)
- ✓ Nature de la Taq polymérase
- ✓ Concentration en  $MgCl_2$
- ✓ Quantité d'ADN matrice et nombre de cycles



## ❑ PCR-DGGE : Principe

- Réalisation d'une PCR « universelle »
- Migration de produits PCR en conditions dénaturantes



Gradient de concentration (agent dénaturant)

DGGE (Electrophorèse sur **G**el en **G**radient **D**énaturant)  
ou utilisation de la température : TGGE

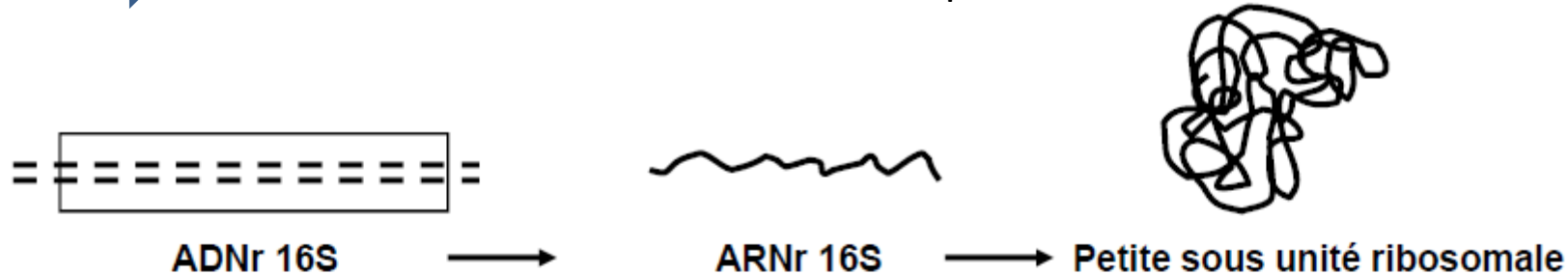
## PCR-DGGE

## 1 - PCR universelle :

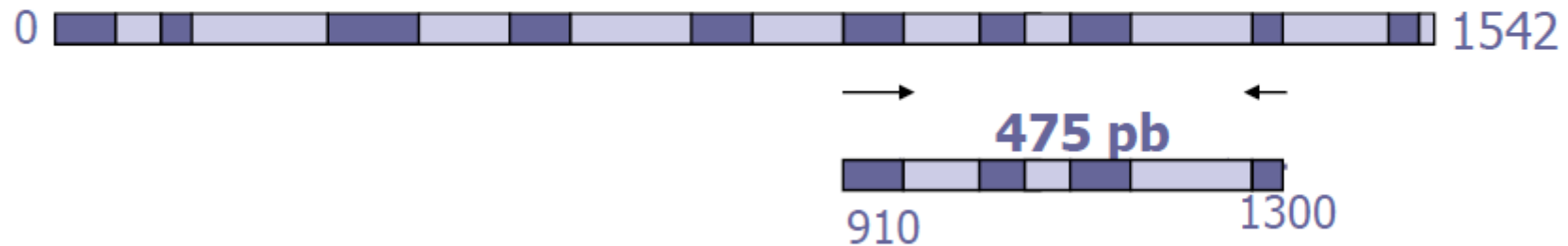
Amorces nucléotidiques positionnées sur les régions conservées de l'ADNr16S et enjambant au moins une région non conservée (domaines V1, V2 ...)



Produits PCR de même taille mais de séquences différentes

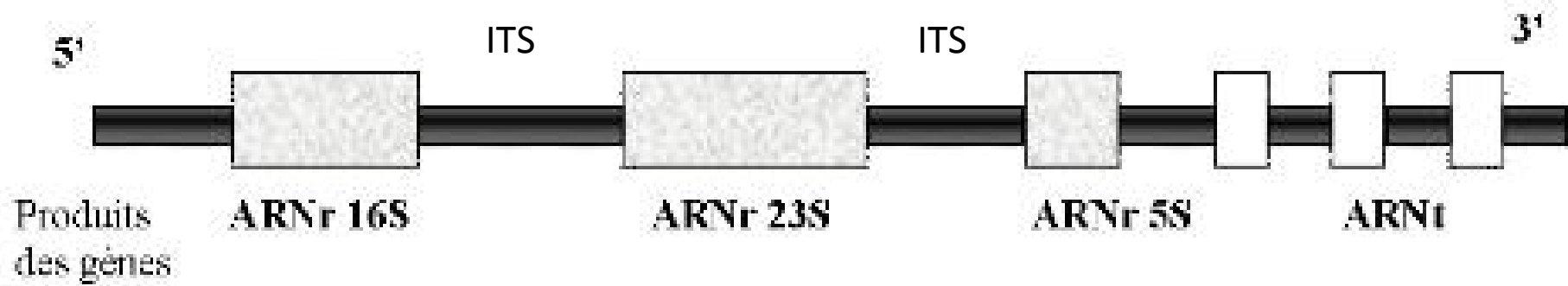


## ADNr 16S

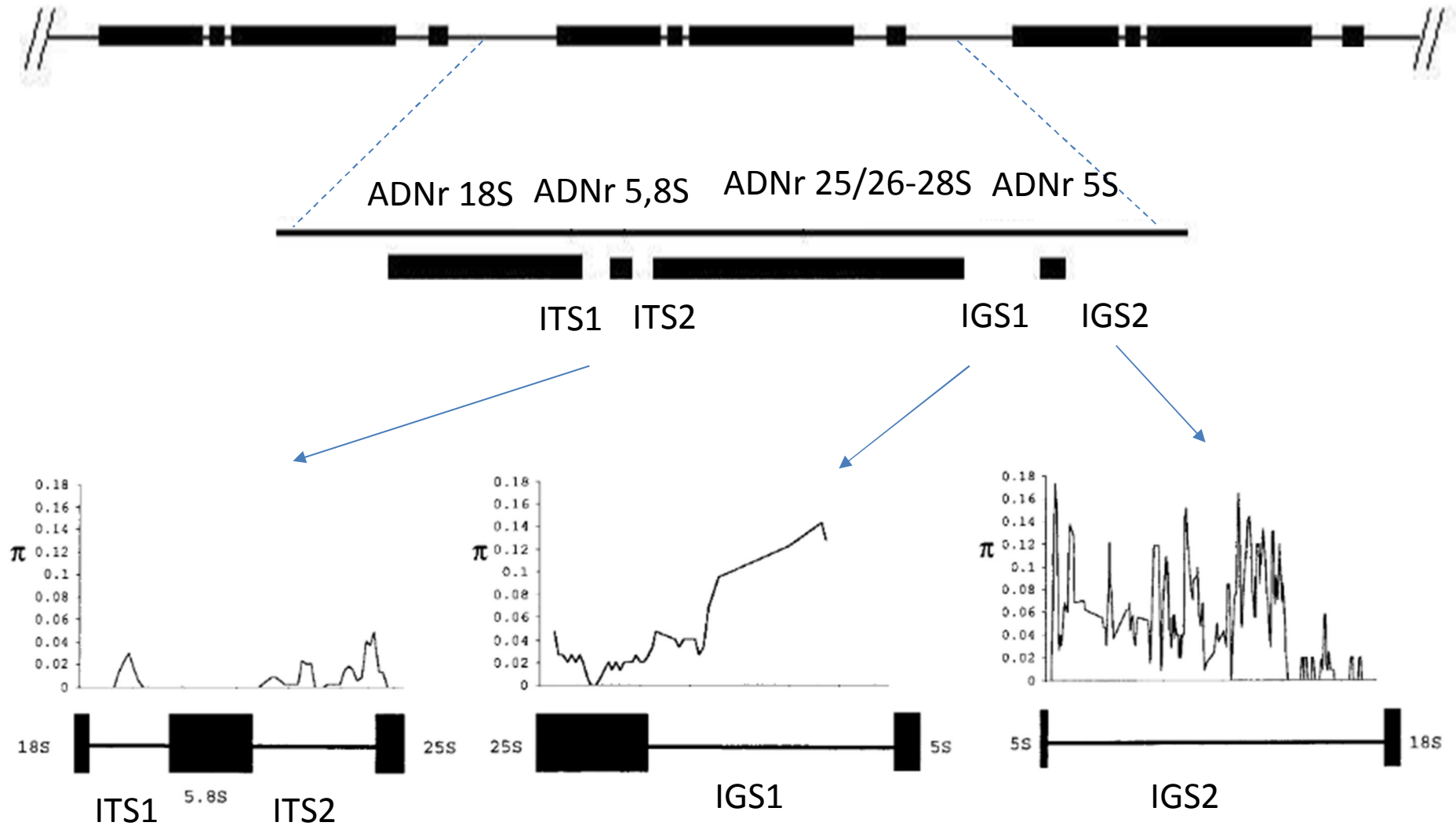


- Régions conservées : séquences nucléotidiques identiques pour toutes les bactéries
- Régions non conservées : séquences nucléotidiques spécifiques du genre et de l'espèce bactérienne

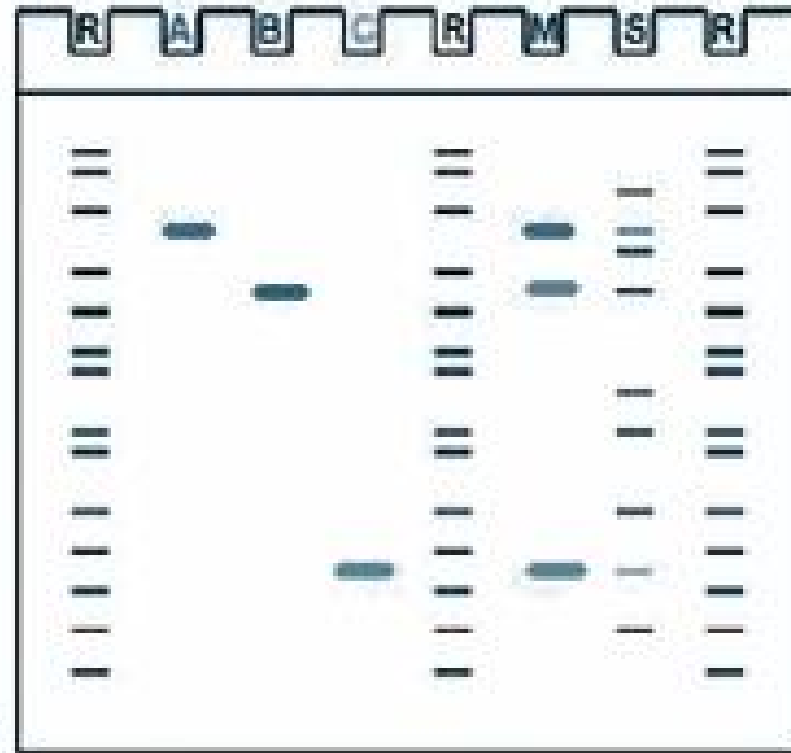
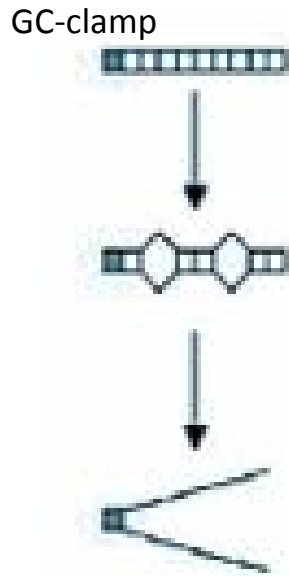
# ADNr des procaryotes



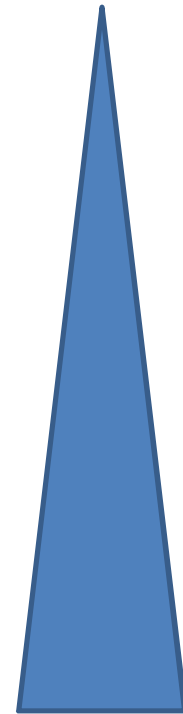
## ADNr des eucaryotes



➔ Produits PCR de même taille séparés en fonction de leur séquence

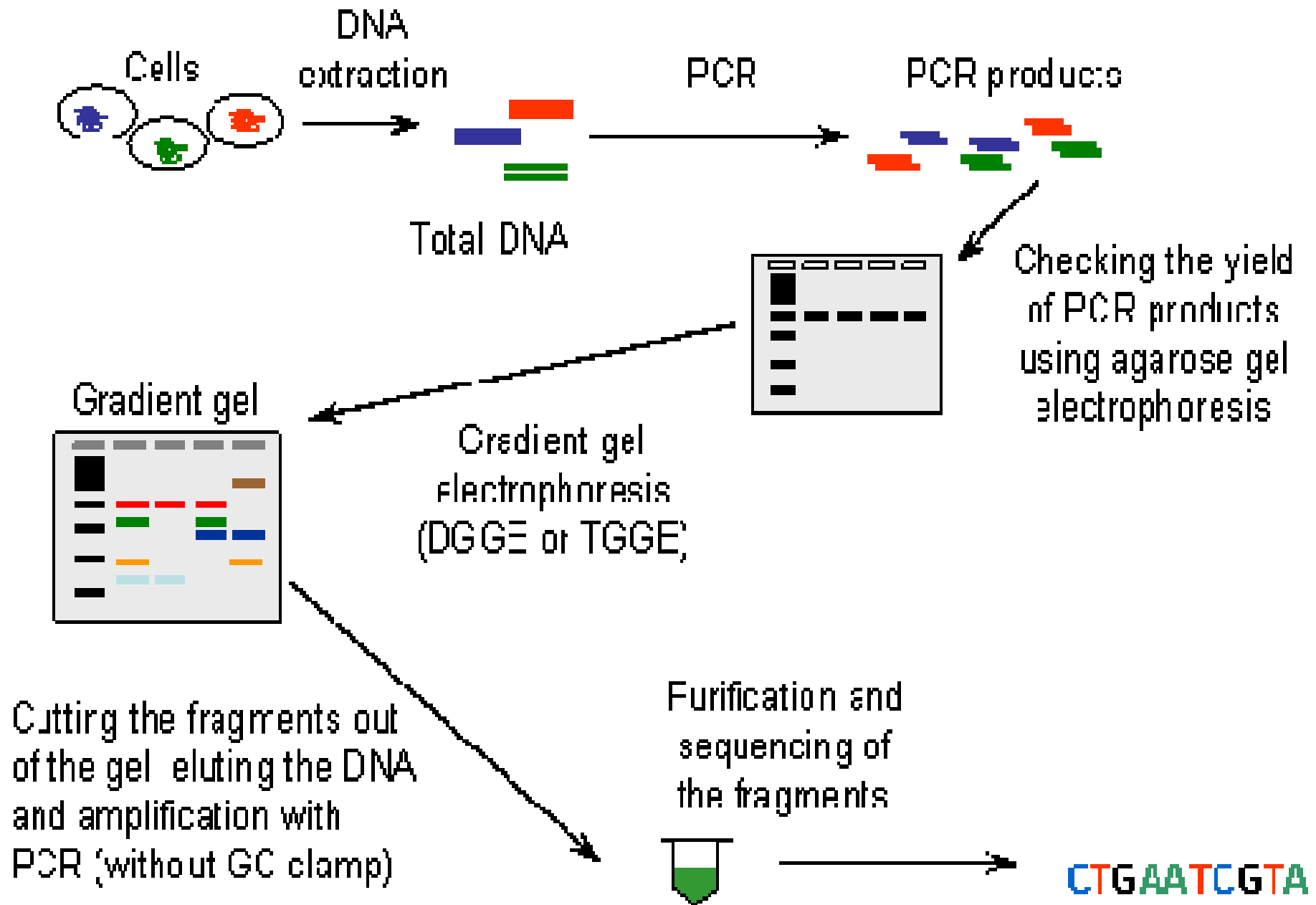


Migration



[Agent dénaturant]

## PCR DGGE

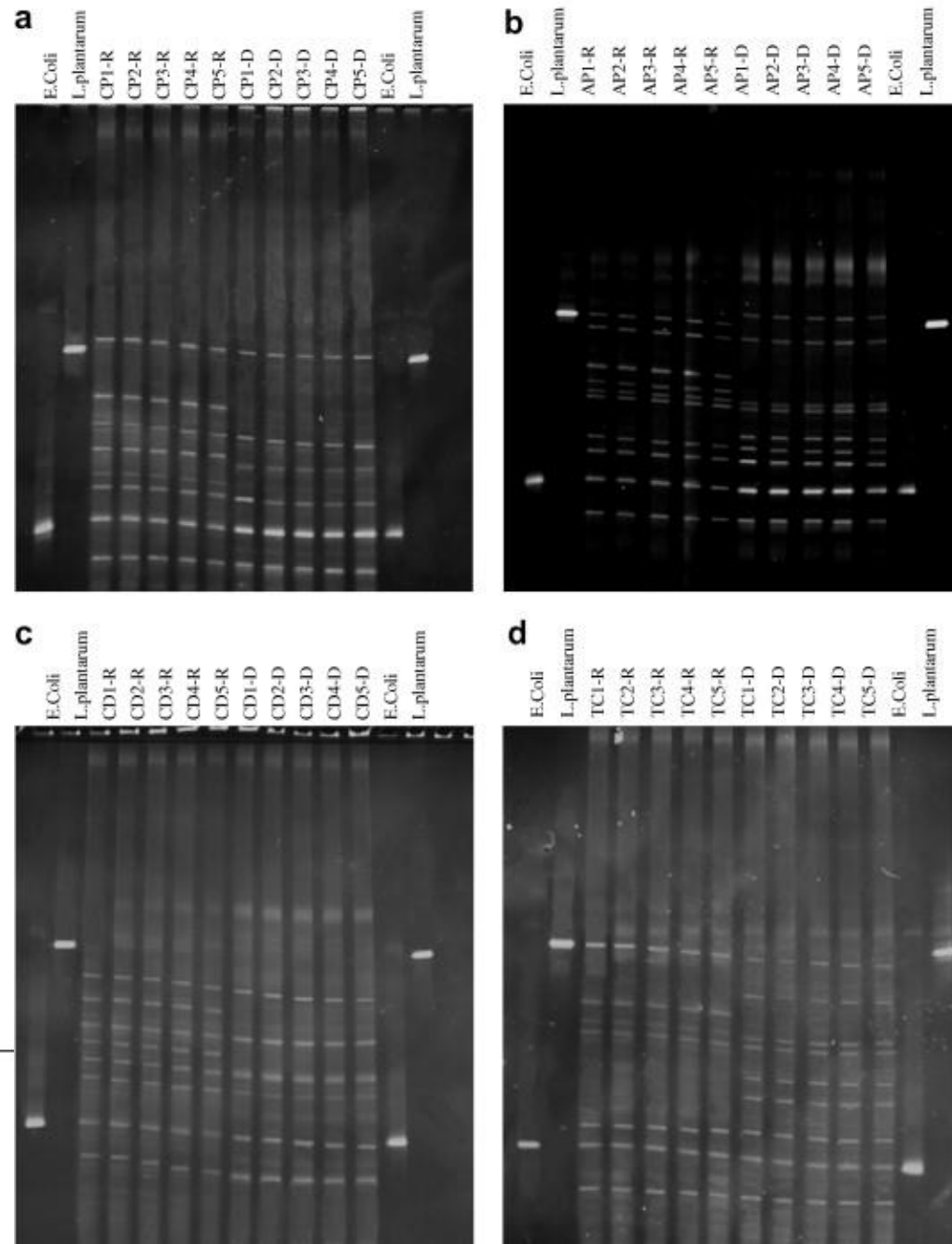


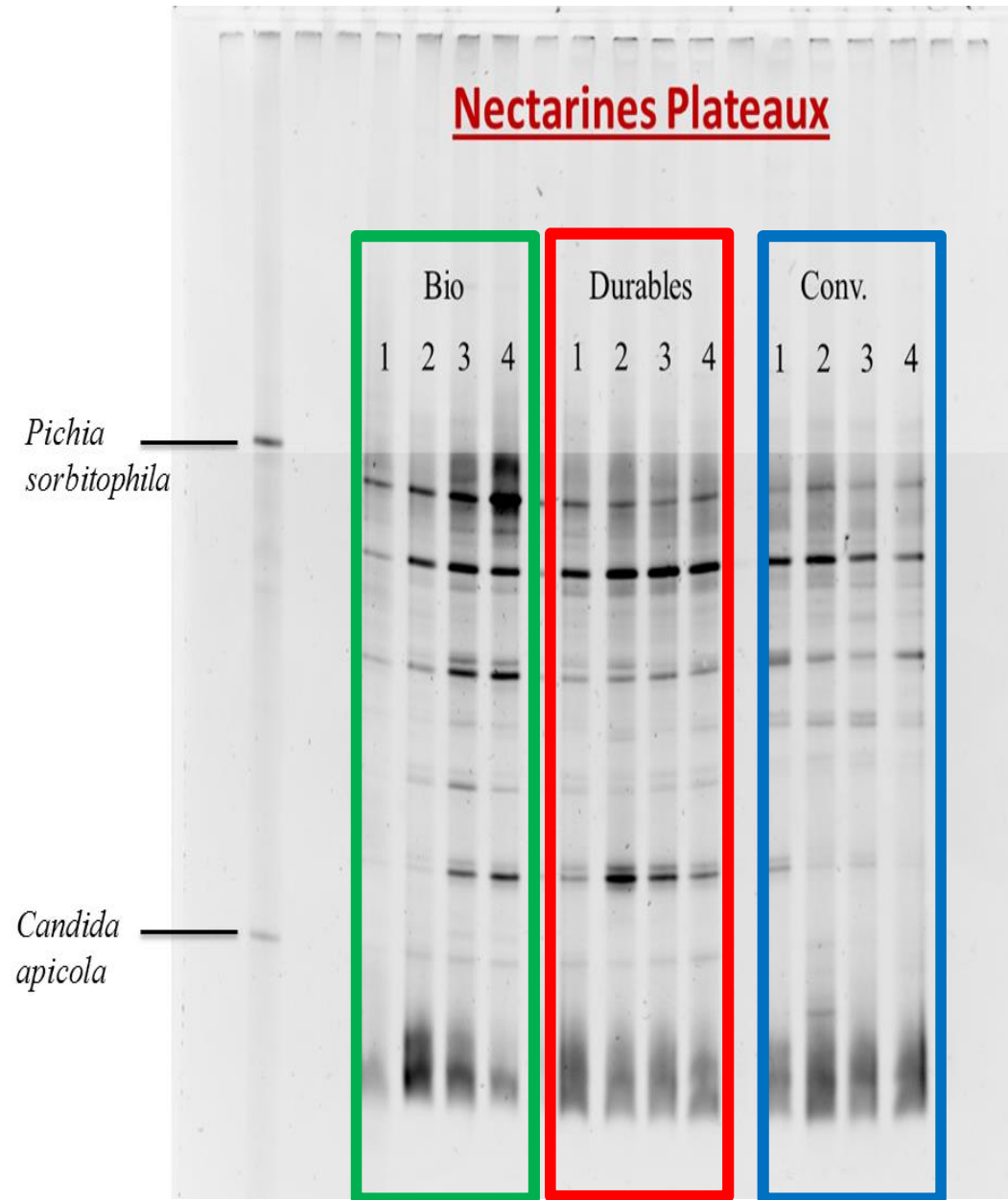
PCR-DGGE

Exemple d'application

Traçabilité

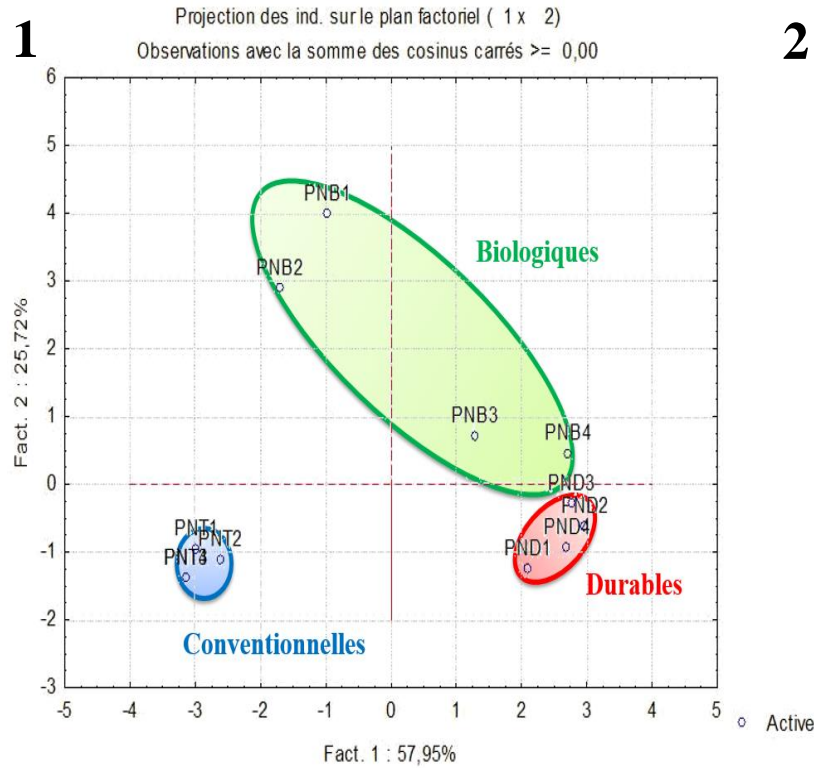
Flore bactérienne issue de poissons de différentes régions



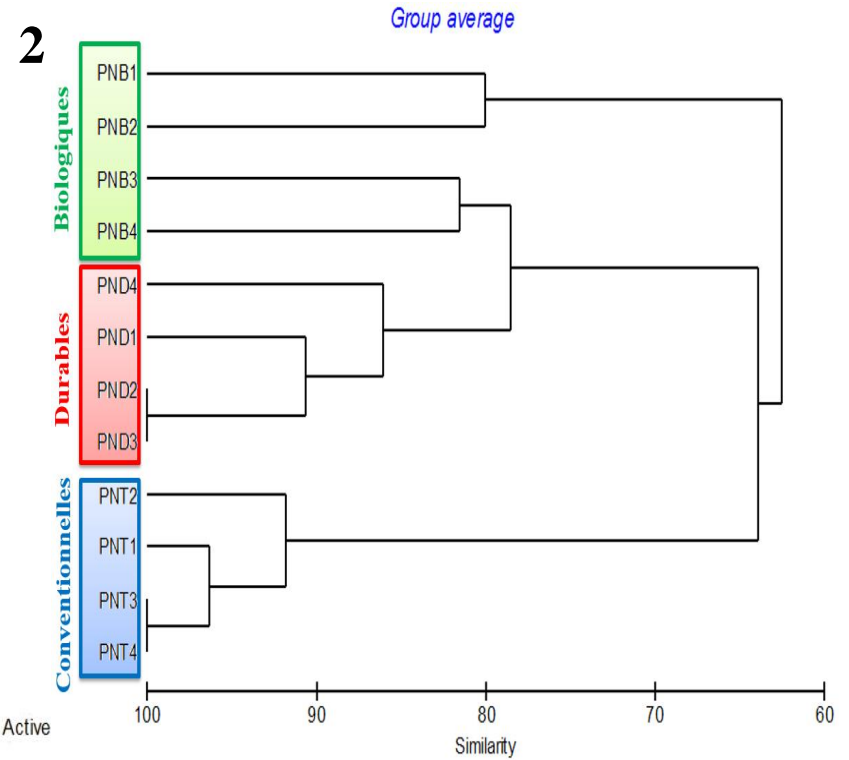


**Profils DGGE de l'ADN levuriens extraits de nectarines récupérées dans des plateaux**

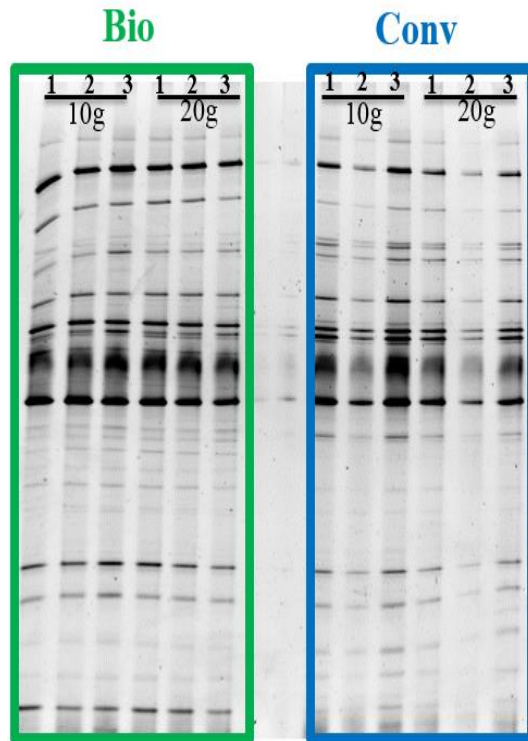




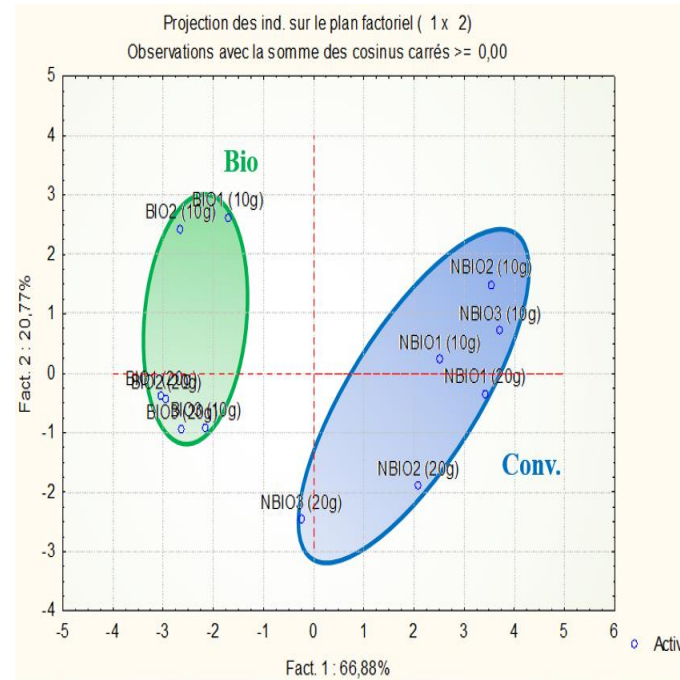
**Analyse factorielle des profils DGGE de levures (ADNr 26S) à partir de nectarines récupérées dans des plateaux de la coopérative et issues de trois types d'agriculture.**



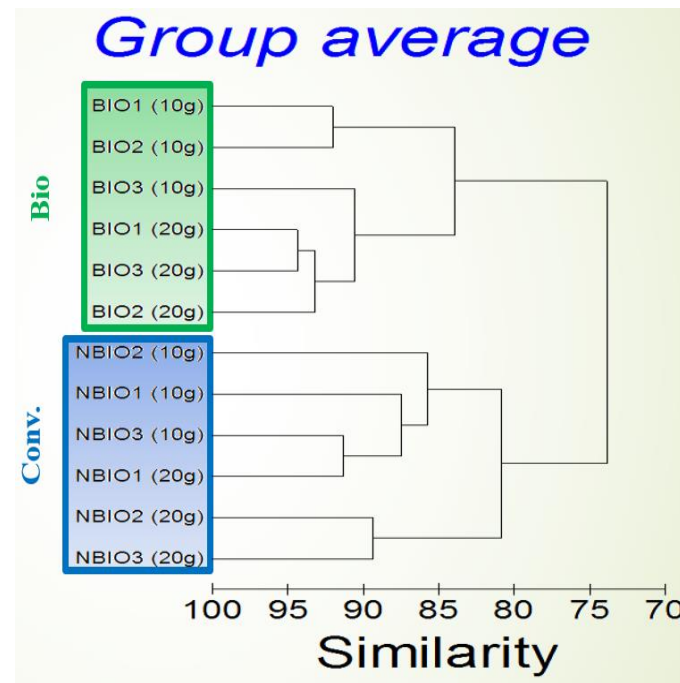
**Analyse par Cluster des profils génétiques levuriens (ADNr 26S) de nectarines récupérées dans des plateaux de la coopérative et issues de trois types d'agriculture.**



**Profils DGGE d'ADN levuriens extraits de graines de cafés verts du Mexique (Biologiques et conventionnelles).**



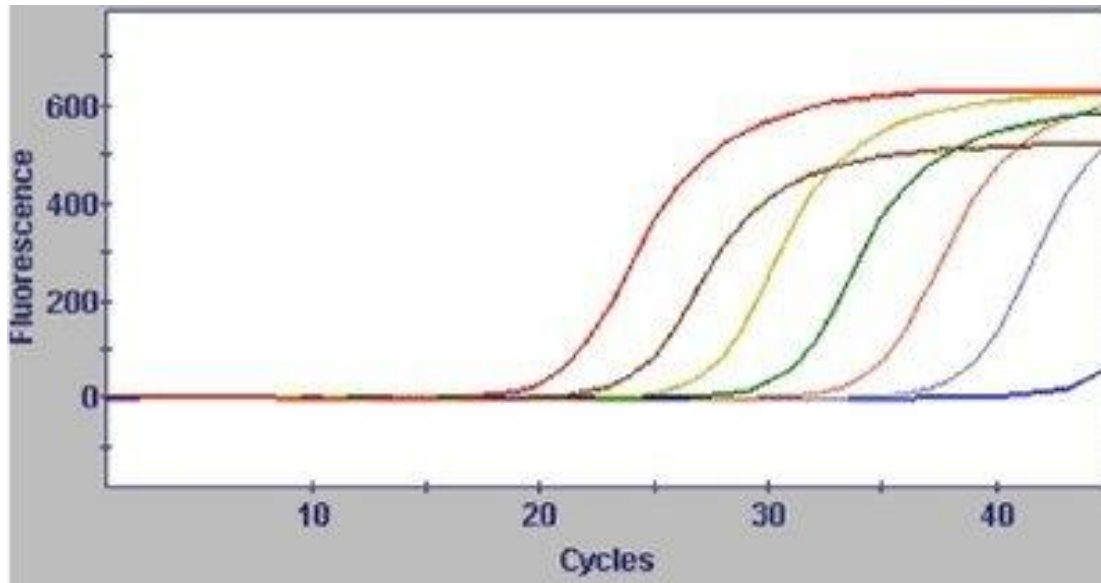
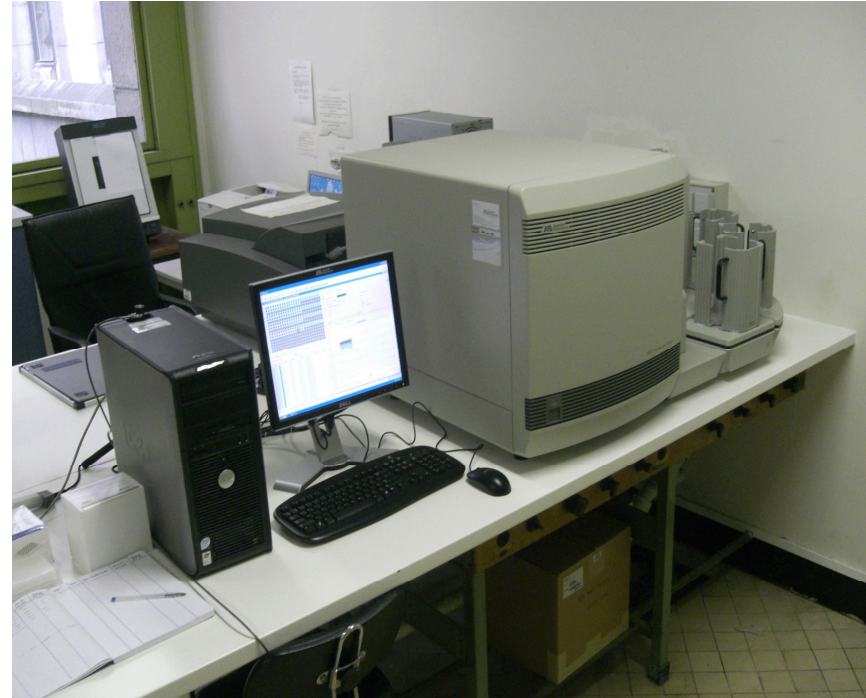
**Analyse factorielle des profils DGGE de levures (ADNr 26S) à partir de cafés du Mexique (Biologiques et conventionnelles).**



**Analyse par Cluster des profils génétiques levuriens (ADNr 26S) à partir de cafés du Mexique (Biologiques et conventionnelles).**

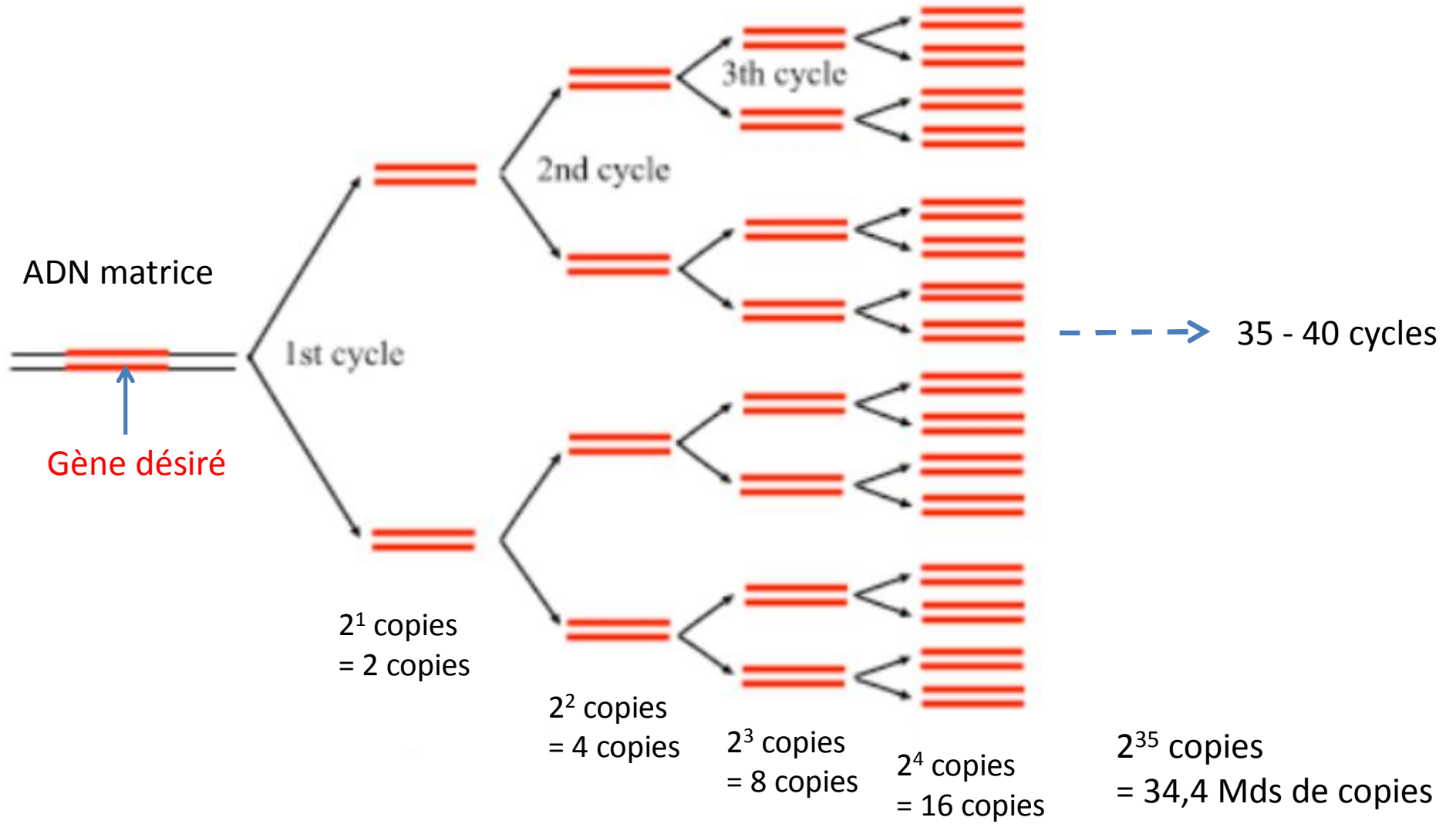
## ☐ Q-PCR

Appareil à Q-PCR



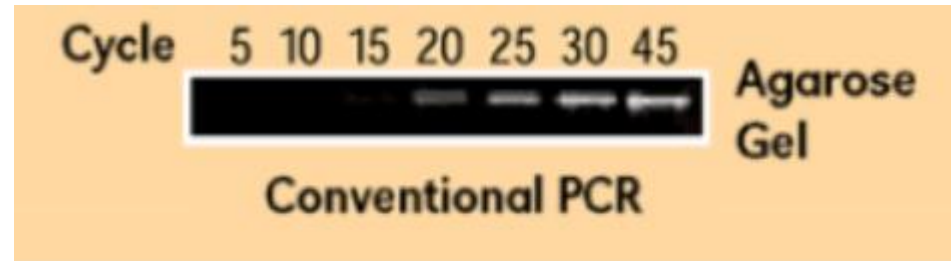
Exemple de résultats

# Principe PCR



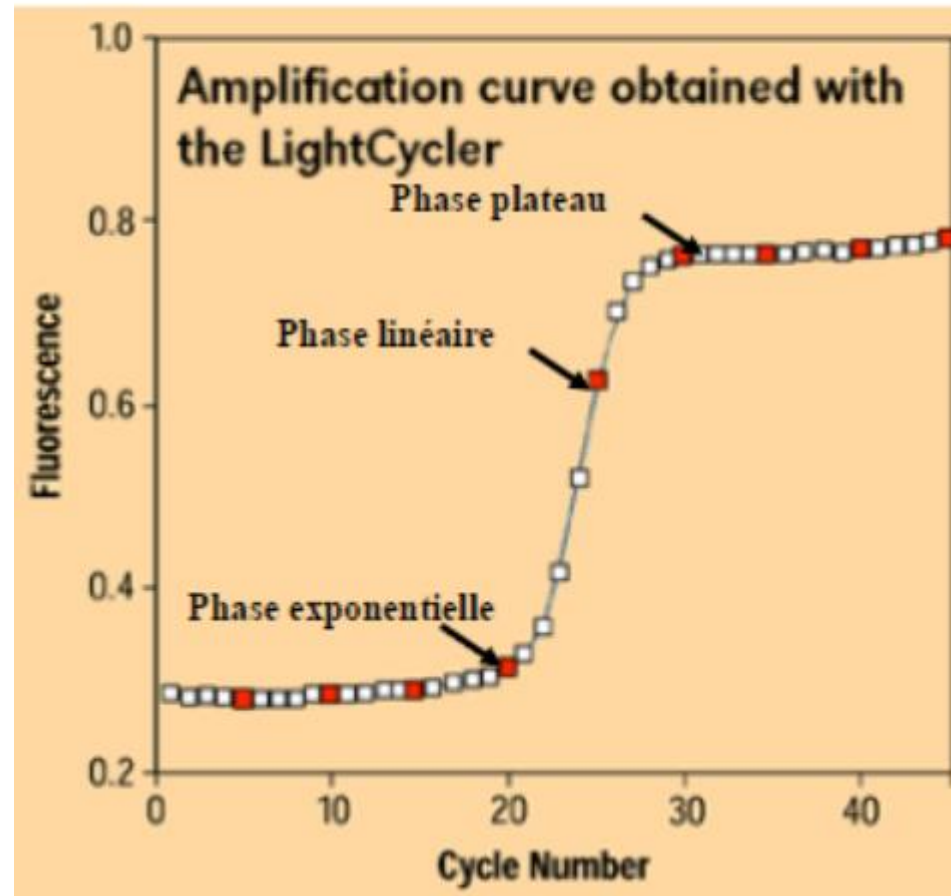
## PCR conventionnelle vs PCR en temps réel

Analyse en point final

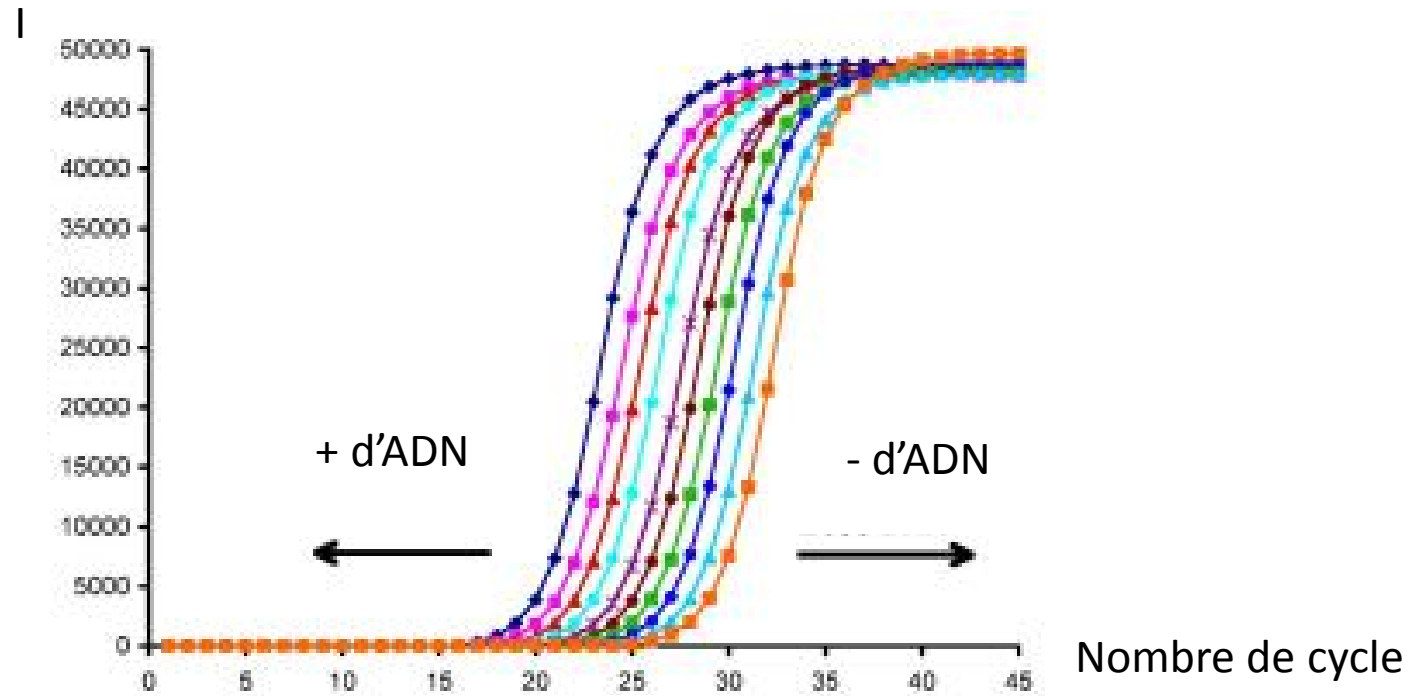


Analyse dynamique

« en temps réel »



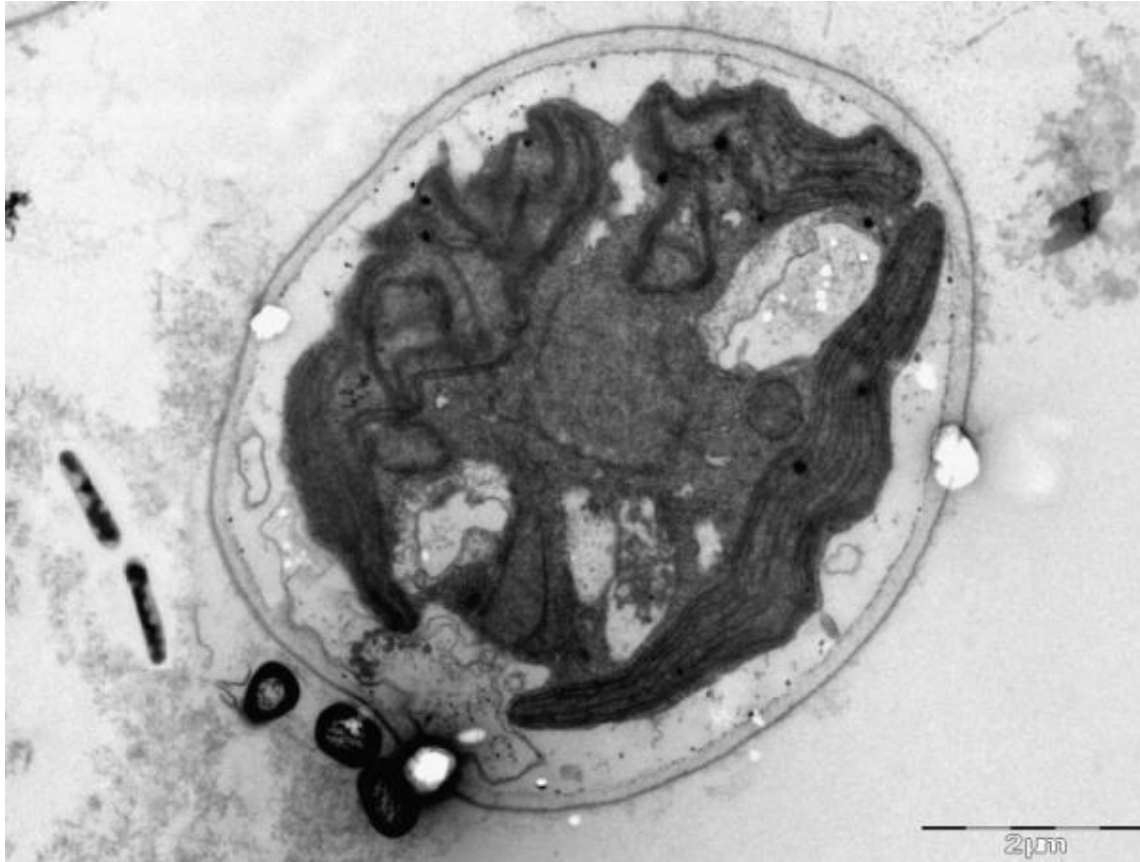
## □ Principe de la Q-PCR



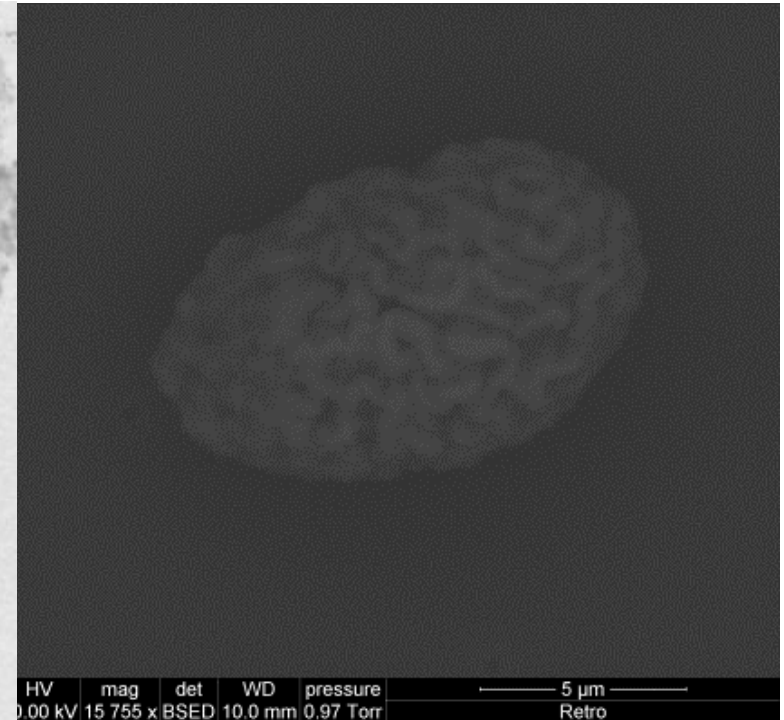


□ Exemple d'application de la Q-PCR

**Suivi de la population de bactéries associées à des microalgues d'intérêt**



Observation au MET de microalgues après 14 jours de culture



Observation au MEB de microalgues après 8 jours de culture

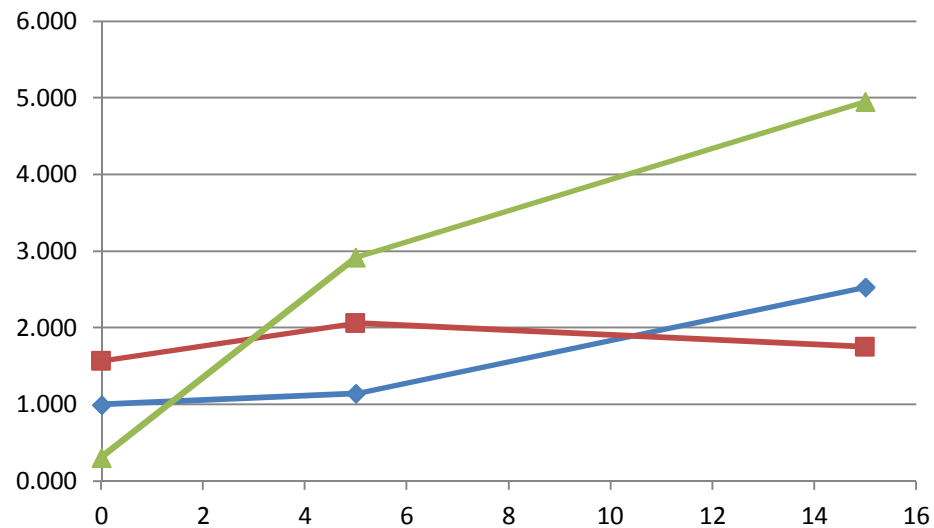
## ❑ Exemple d'application de la Q-PCR

Pourquoi utiliser la Q-PCR pour suivre l'évolution de la population ?

➔ Dénombrement par méthode de microbiologie classique impossible car :

- Ecosystème particulier : Présence de bactéries viables mais non cultivables
- Bactéries indissociables physiquement du glycoalyx de la microalgue

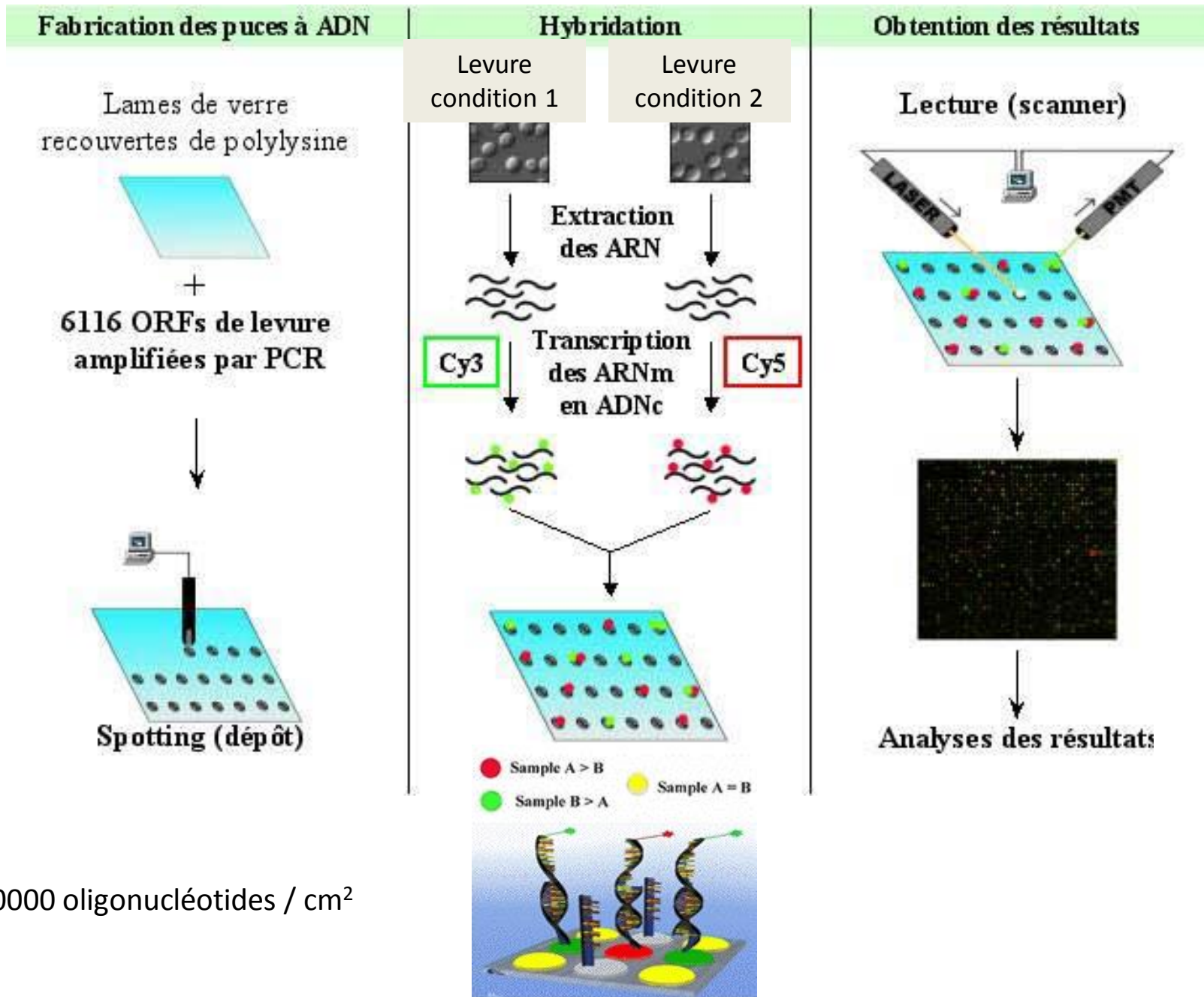
Abondance relative de bactéries



Durée de la culture de microalgues (jours)



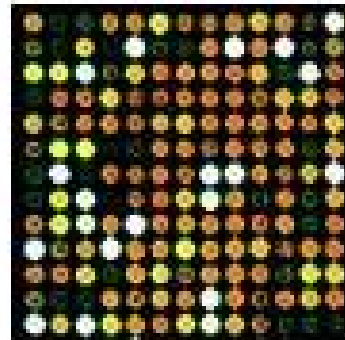
# ❑ Puces à ADN : Application étude du transcriptome (ARNm)



# Microarrays



Détail :



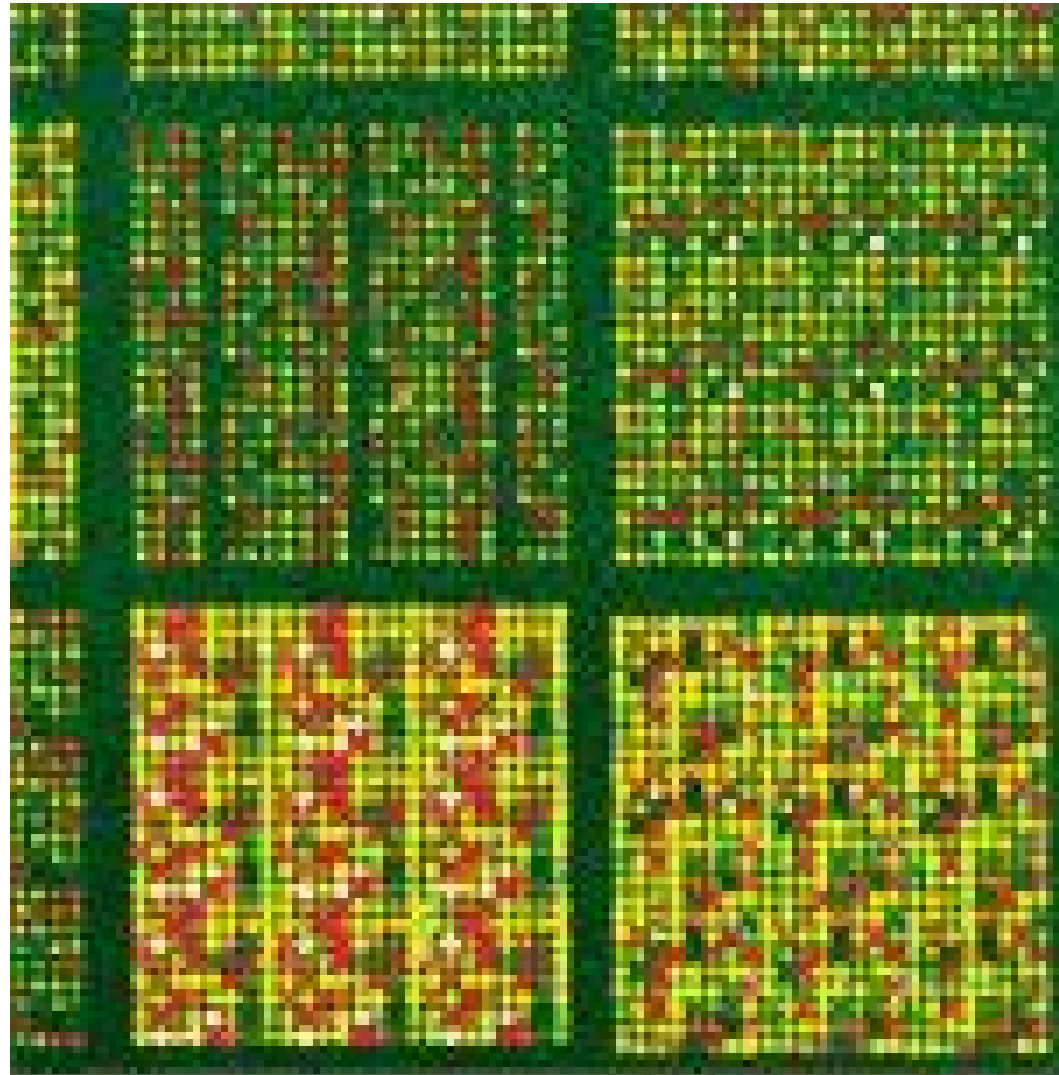
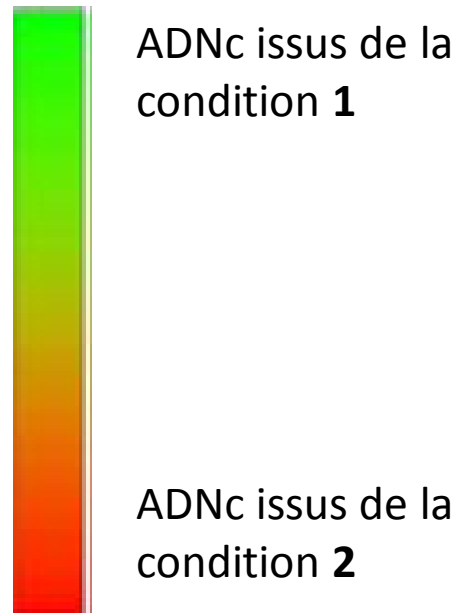
Taille : 5,4cm x 0,9cm

- 10000 clones par lame
- marquage fluorescent
- 2 conditions expérimentales par lame

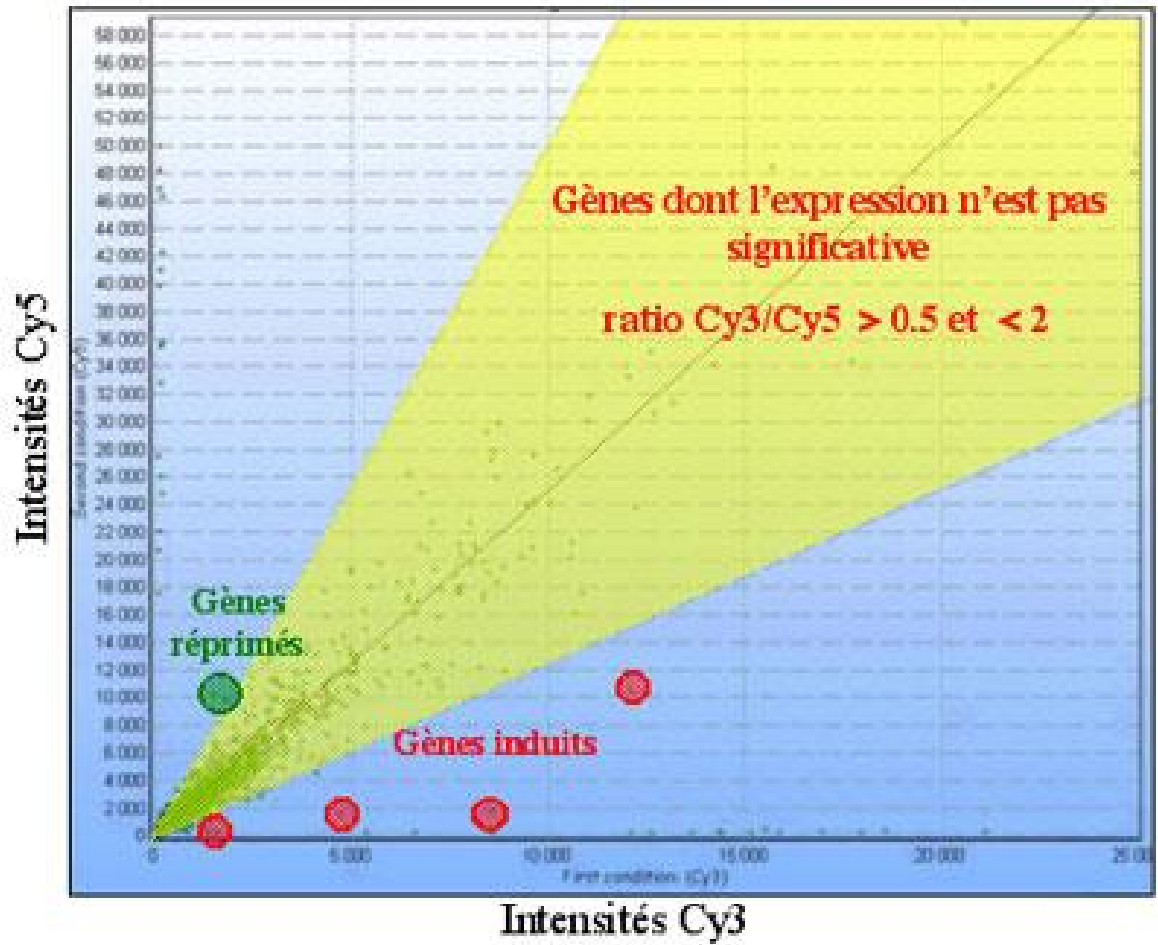
## Lecture des données

Excitation des spots par un laser

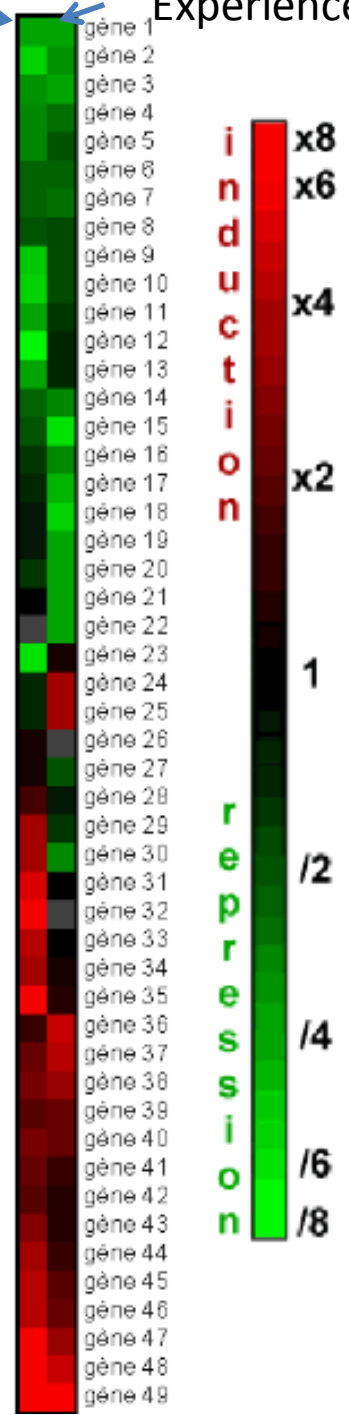
Récupération de la fluorescence émise via un photomultiplicateur



Analyse des données :

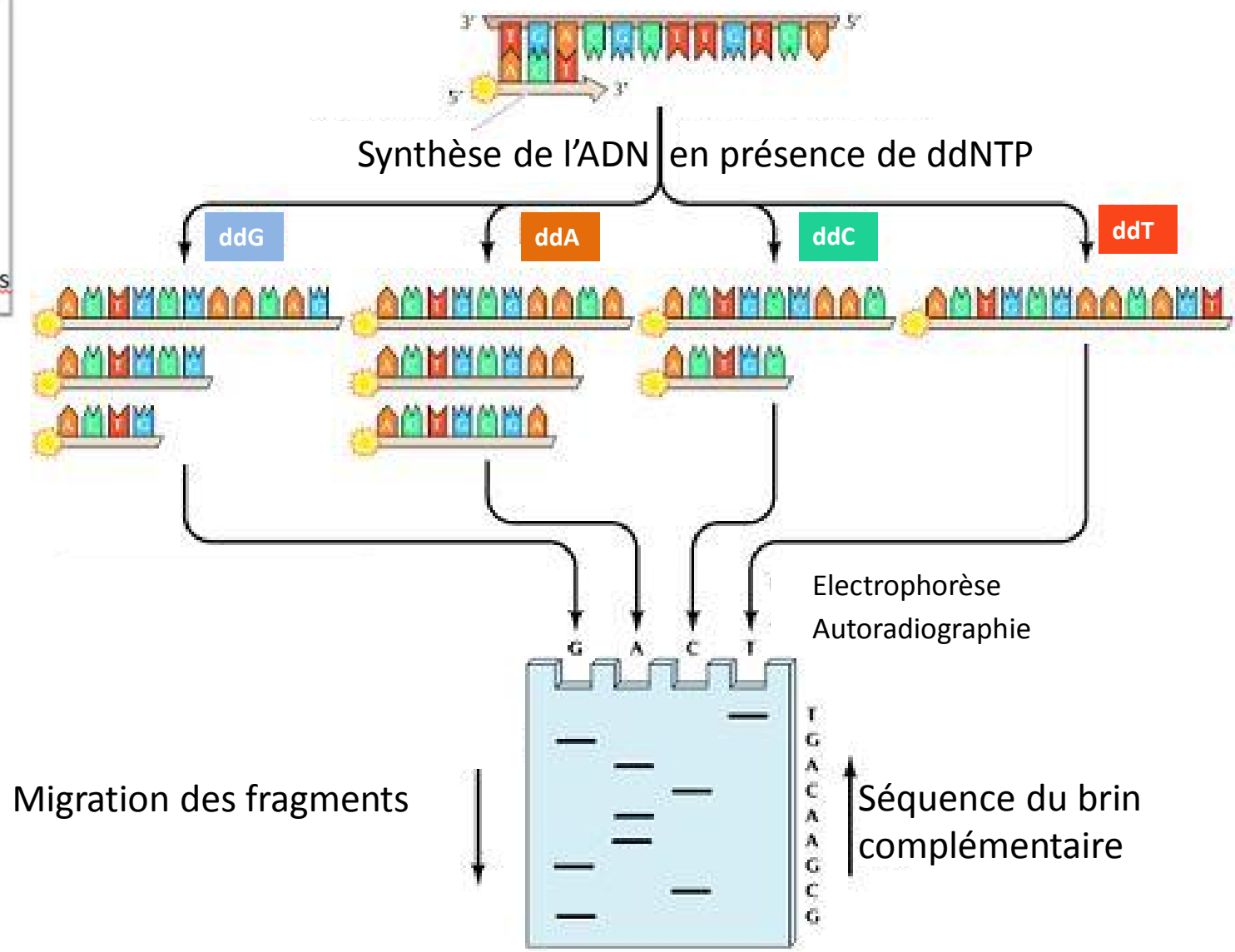
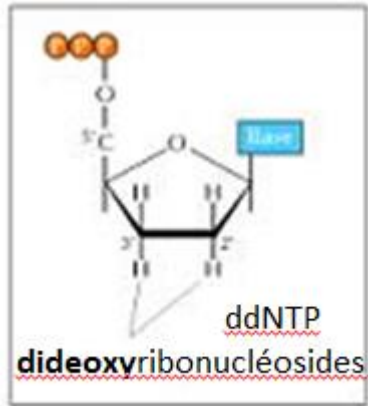


Expérience A → Expérience B



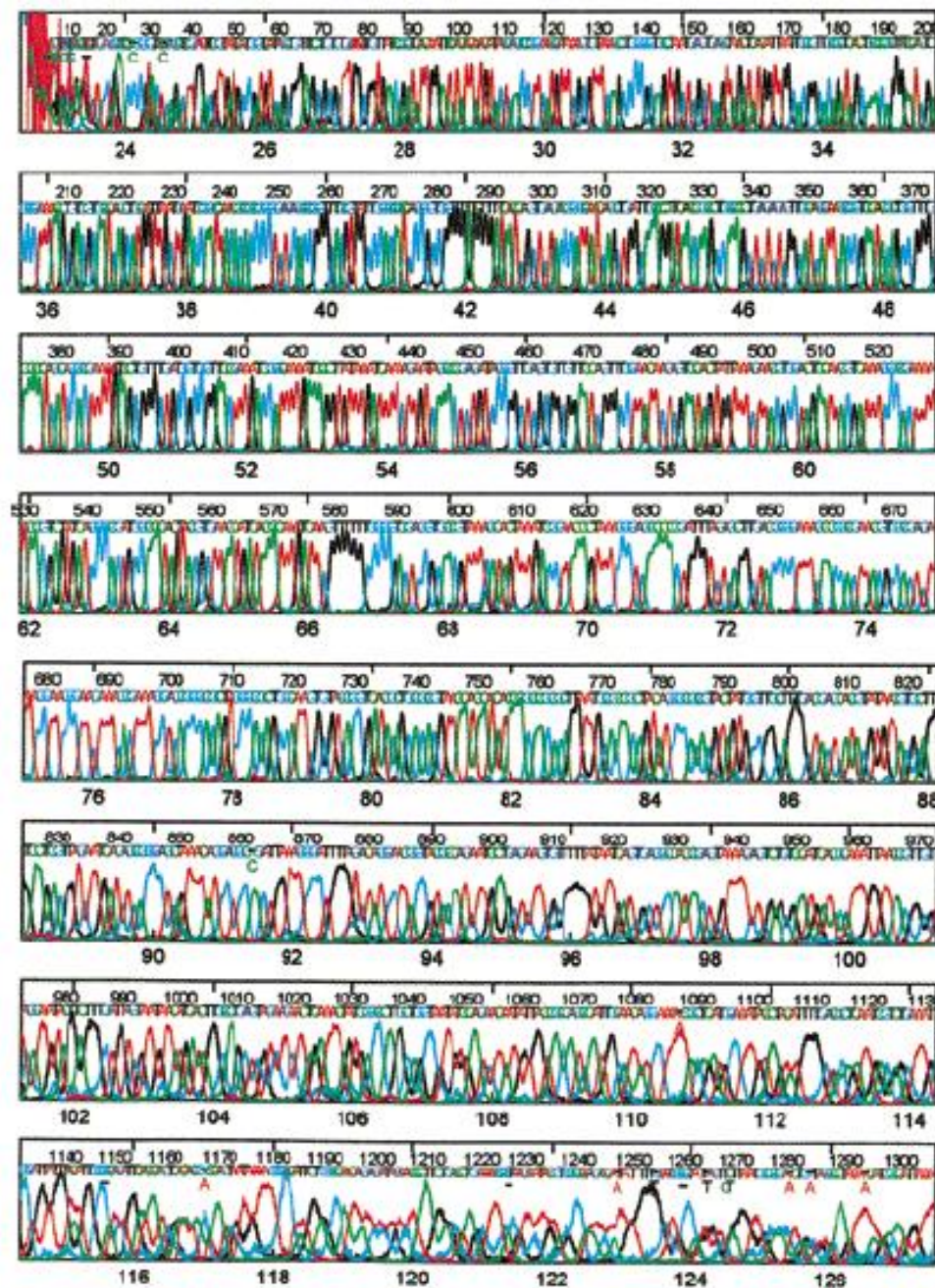
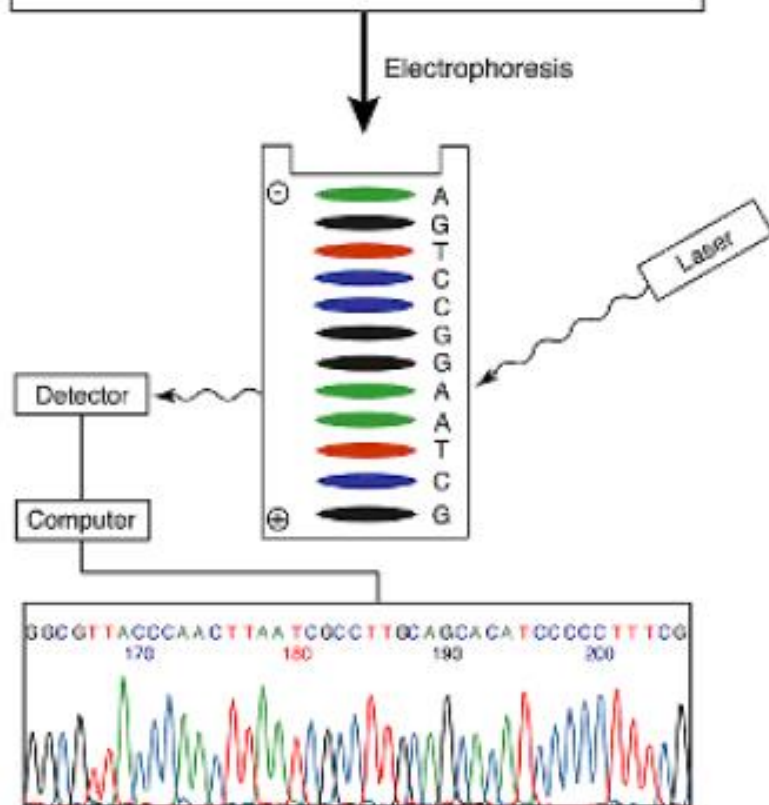
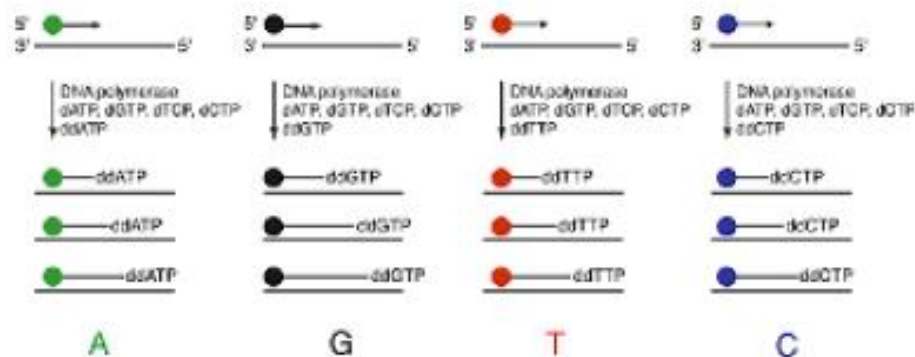
104

❑ Séquençage de l'ADN par la méthode de Sanger (1918-2013)

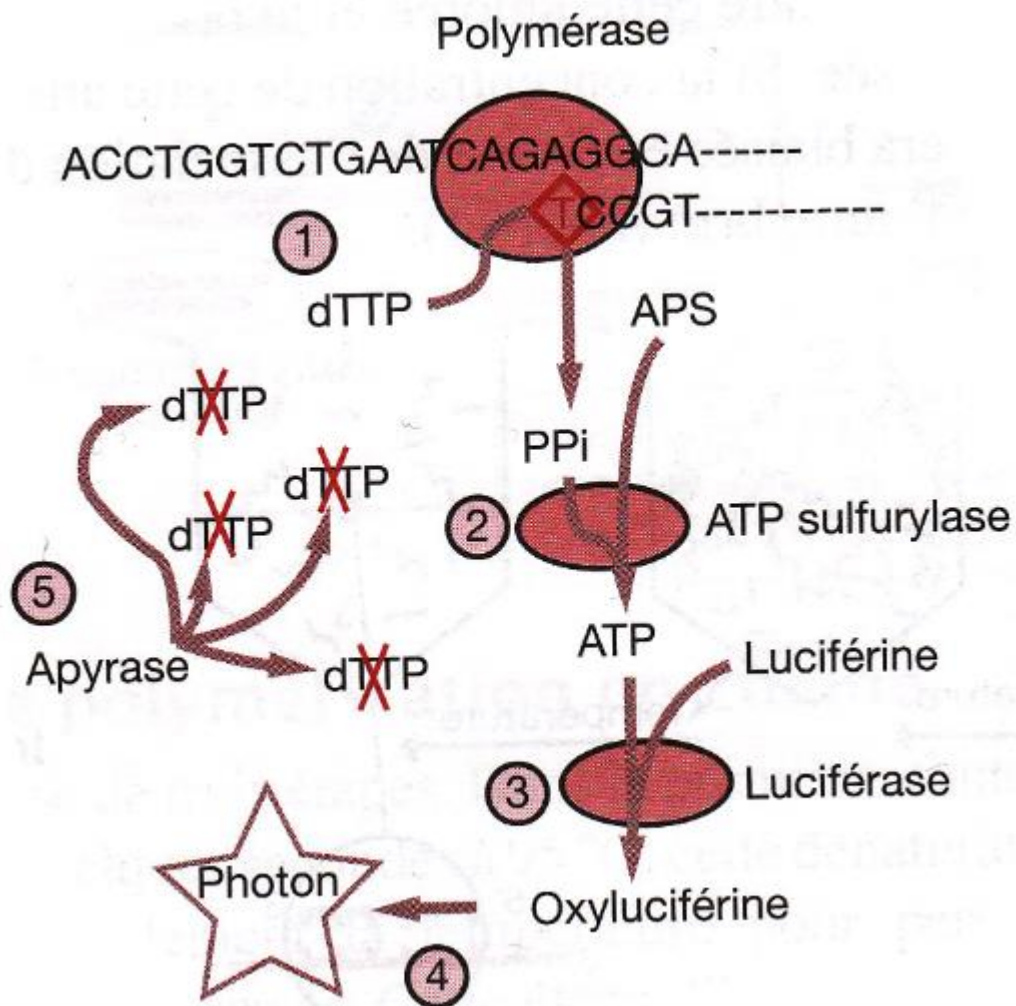




❑ Séquençage de l'ADN par la méthode de Sanger modernisée (BigDye Terminators 1986)

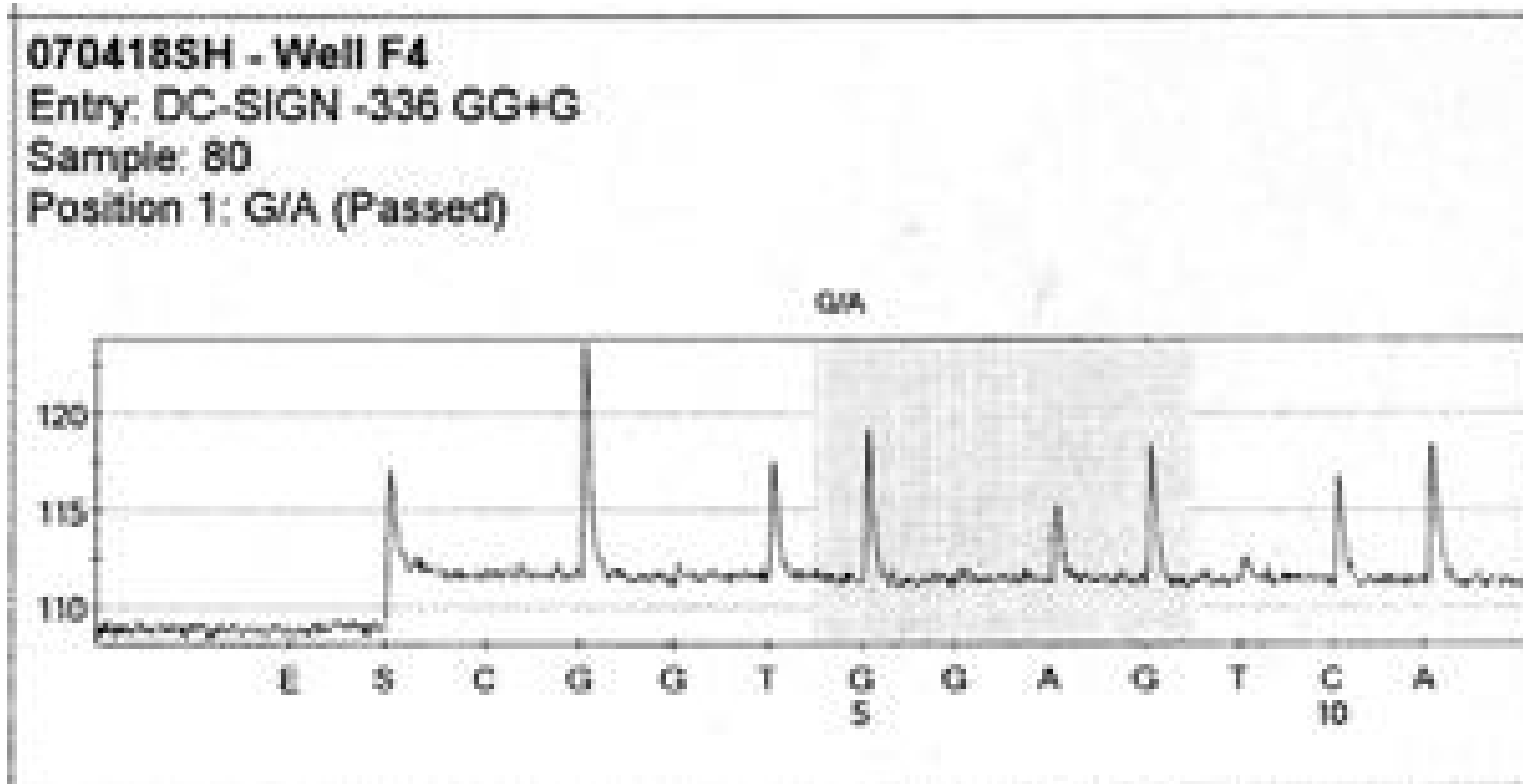


❑ Séquençage 454 : principe du pyroséquençage



1. La polymérase incorpore le dTTP
2. Libération d'un PPI utilisé pour générer 1 ATP
3. La luciférase utilise un ATP
4. La luciférase va émettre un photon
5. Nt non incorporés + ATP dégradés par l'apyrase

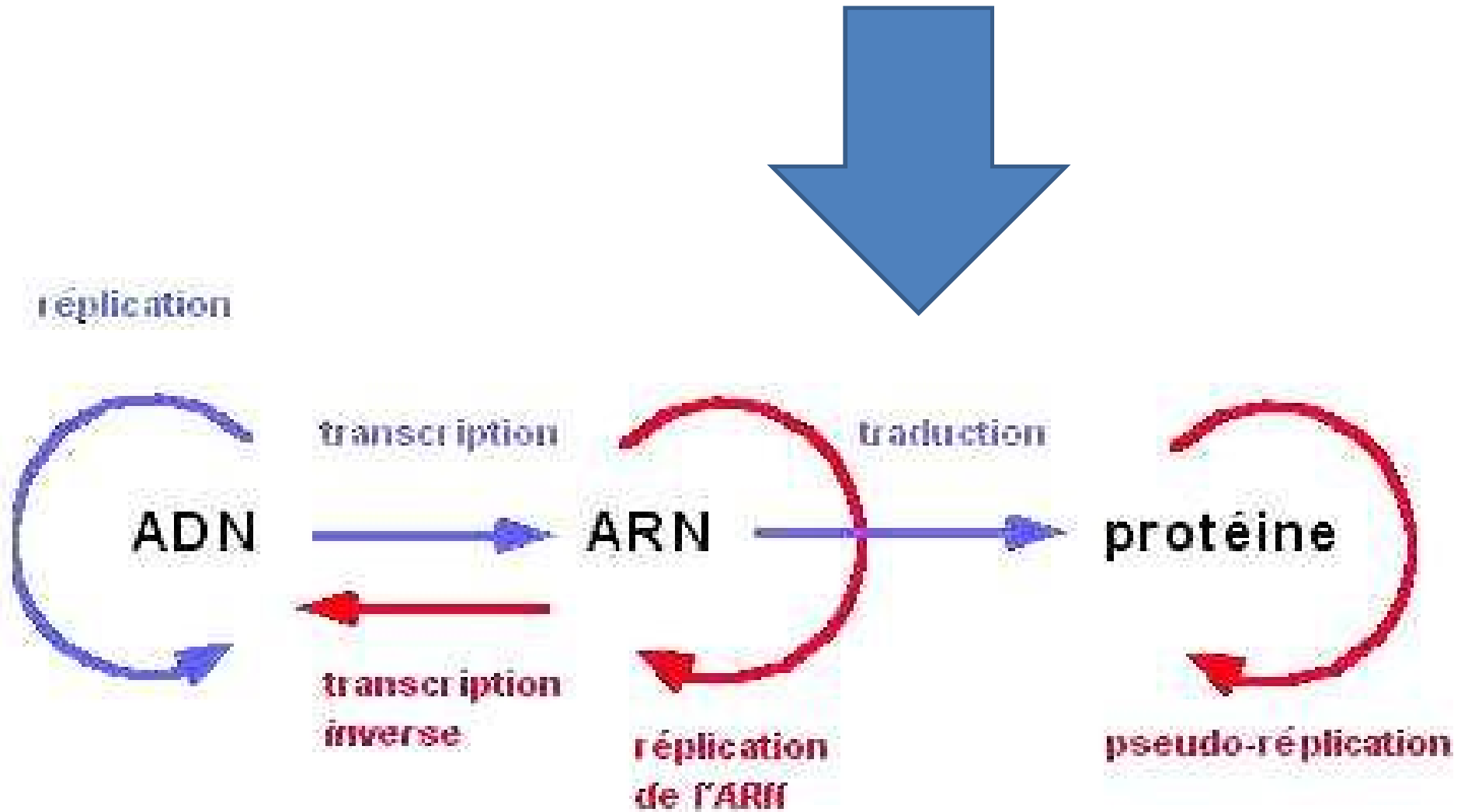
Exemple de résultat de pyroséquençage (Pyrogramme)





# III - DU GÈNE À LA PROTÉINE

## III - 2 TRADUCTION



## 3.2 Traduction

- ❖ La synthèse des protéines est très complexe
- ❖ L'ARNm contient le code et joue le rôle de matrice pour la synthèse des protéines
- ❖ Les ARNt sont des molécules d'adaptation qui transportent les aa vers l'ARNm
- ❖ Les ARNr forment la partie du ribosome qui réunit tout les composant nécessaire à la synthèse de la protéine

<b>Acide aminé</b>	<b>Abréviation à 3 lettres</b>	<b>Abréviation à 1 lettre</b>
<b>Alanine</b>	<b>Ala</b>	<b>A</b>
<b>Arginine</b>	<b>Arg</b>	<b>R</b>
<b>Asparagine</b>	<b>Asn</b>	<b>N</b>
<b>Aspartic Acid</b>	<b>Asp</b>	<b>D</b>
<b>Cysteine</b>	<b>Cys</b>	<b>C</b>
<b>Glutamine</b>	<b>Gln</b>	<b>Q</b>
<b>Glutamic Acid</b>	<b>Glu</b>	<b>E</b>
<b>Glycine</b>	<b>Gly</b>	<b>G</b>
<b>Histidine</b>	<b>His</b>	<b>H</b>
<b>Isoleucine</b>	<b>Ile</b>	<b>I</b>
<b>Leucine</b>	<b>Leu</b>	<b>L</b>
<b>Lysine</b>	<b>Lys</b>	<b>K</b>
<b>Methionine</b>	<b>Met</b>	<b>M</b>
<b>Phenylalanine</b>	<b>Phe</b>	<b>F</b>
<b>Proline</b>	<b>Pro</b>	<b>P</b>
<b>Serine</b>	<b>Ser</b>	<b>S</b>
<b>Threonine</b>	<b>Thr</b>	<b>T</b>
<b>Tryptophan</b>	<b>Trp</b>	<b>W</b>
<b>Tyrosine</b>	<b>Tyr</b>	<b>Y</b>
<b>Valine</b>	<b>Val</b>	<b>V</b>

☐ Le code génétique

“ Quasi-universel

“ Dégénéré (codons synonymes)

“ Préférence codonique

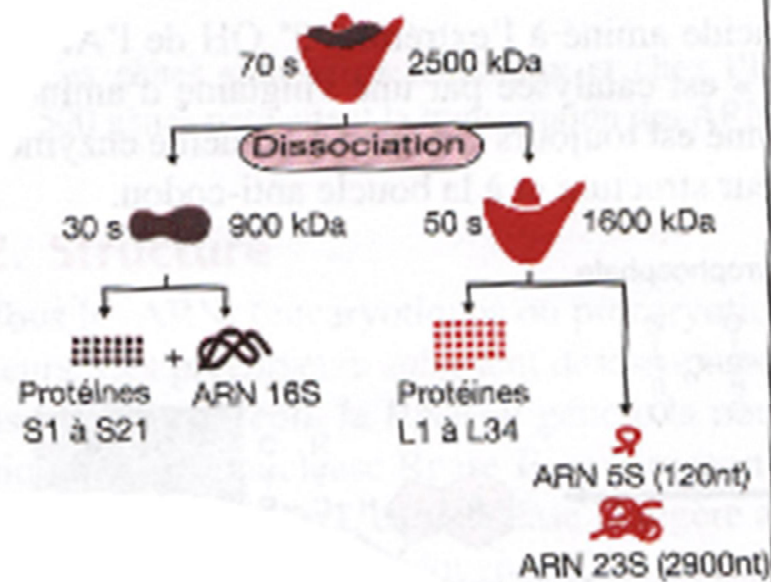
1<sup>ère</sup> base du codon

	U	C	A	G
U	UUU Phe	UCU Ser	UAU Tyr	UGU Cys
	UUC Phe	UCC Ser	UAC Tyr	UGC Cys
	UUA Leu	UCA Ser	UAA Stop	UGA Stop
	UUG Leu	UCG Ser	UAG Stop	UGG Trp
C	CUU Leu	CCU Pro	CAU His	CGU Arg
	CUC Leu	CCC Pro	CAC His	CGC Arg
	CUA Leu	CCA Pro	CAA Gln	CGA Arg
	CUG Leu	CCG Pro	CAG Gln	CGG Arg
A	AUU Ile	ACU Thr	AAU Asn	AGU Ser
	AUC Ile	ACC Thr	AAC Asn	AGC Ser
	AUA Ile	ACA Thr	AAA Lys	AGA Arg
	AUG Met	ACG Thr	AAG Lys	AGG Arg
G	GUU Val	GCU Ala	GAU Asp	GGU Gly
	GUC Val	GCC Ala	GAC Asp	GGC Gly
	GUA Val	GCA Ala	GAA Glu	GGA Gly
	GUG Val	GCG Ala	GAG Glu	GGG Gly

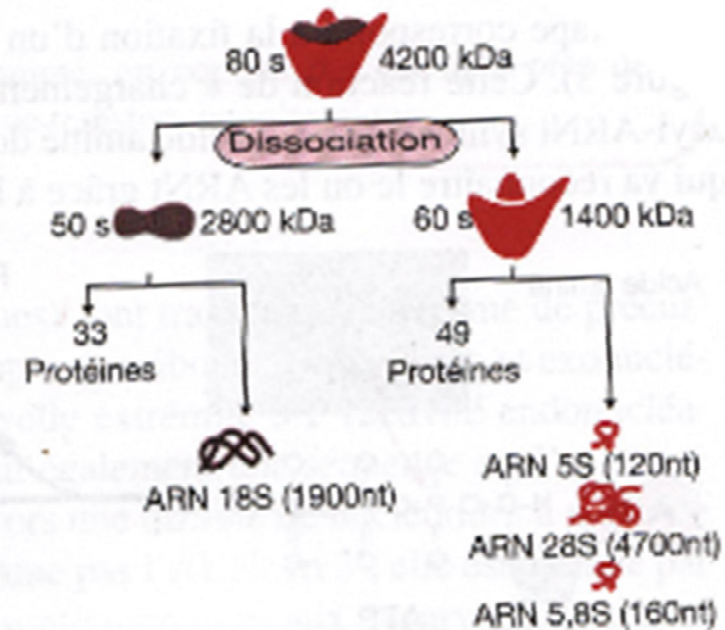
Codon	<i>E.coli</i>	<i>D.melanogaster</i>	<i>H.sapiens</i>	<i>S.cerevisiae</i>
AGT	3	1	10	5
AGC	20	23	34	4
TCG	4	17	9	1
TCA	2	2	5	6
TCT	34	9	13	52
TCC	37	48	28	33

## ☐ Ribosomes

### Ribosome procaryote



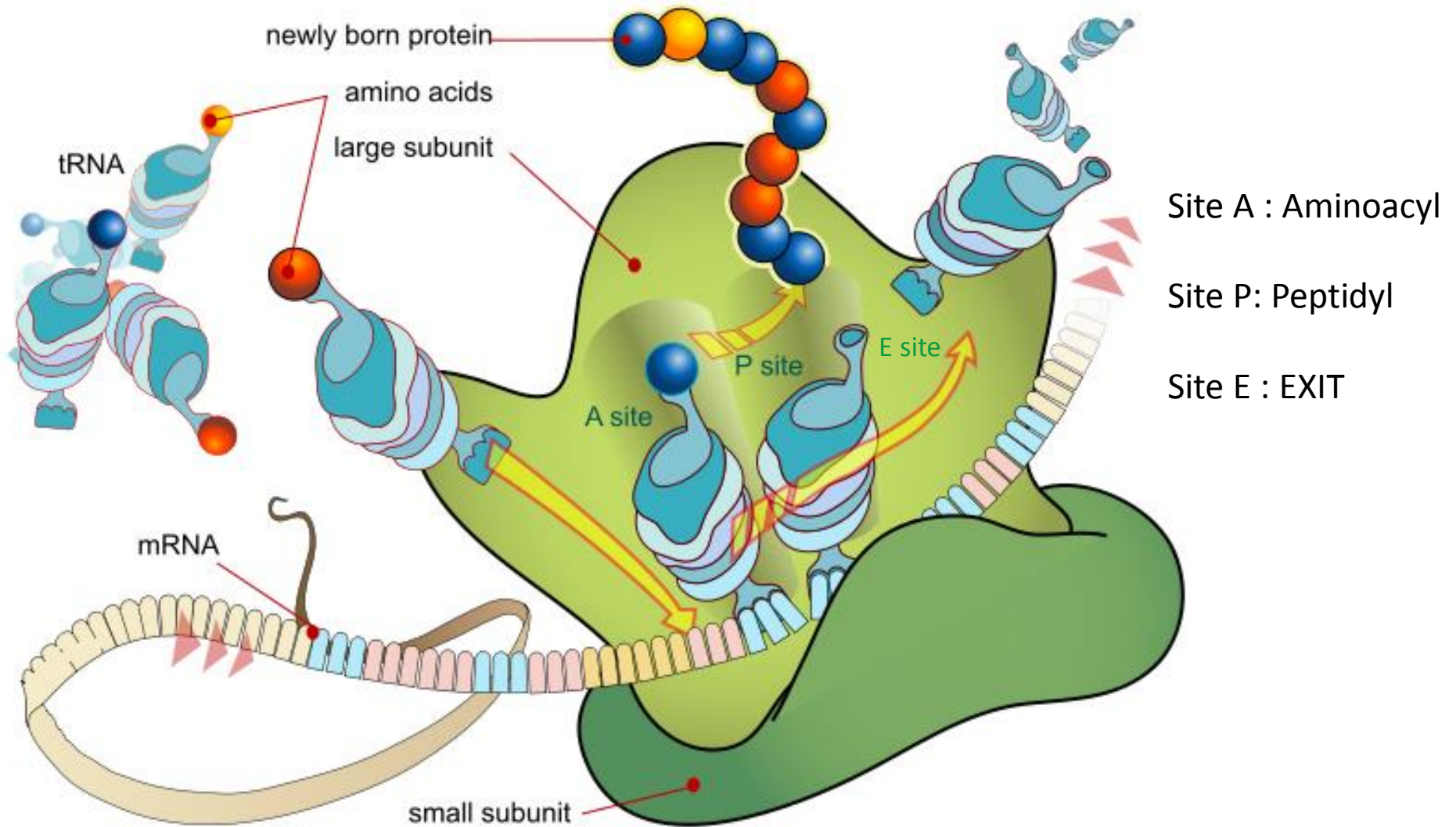
### Ribosome eucaryote



Petite sous-unité : décode les ARNm

Grande sous-unité : forme la liaison peptidique entre les aa portés par les ARNt

❑ Mécanisme de base de la synthèse des protéines

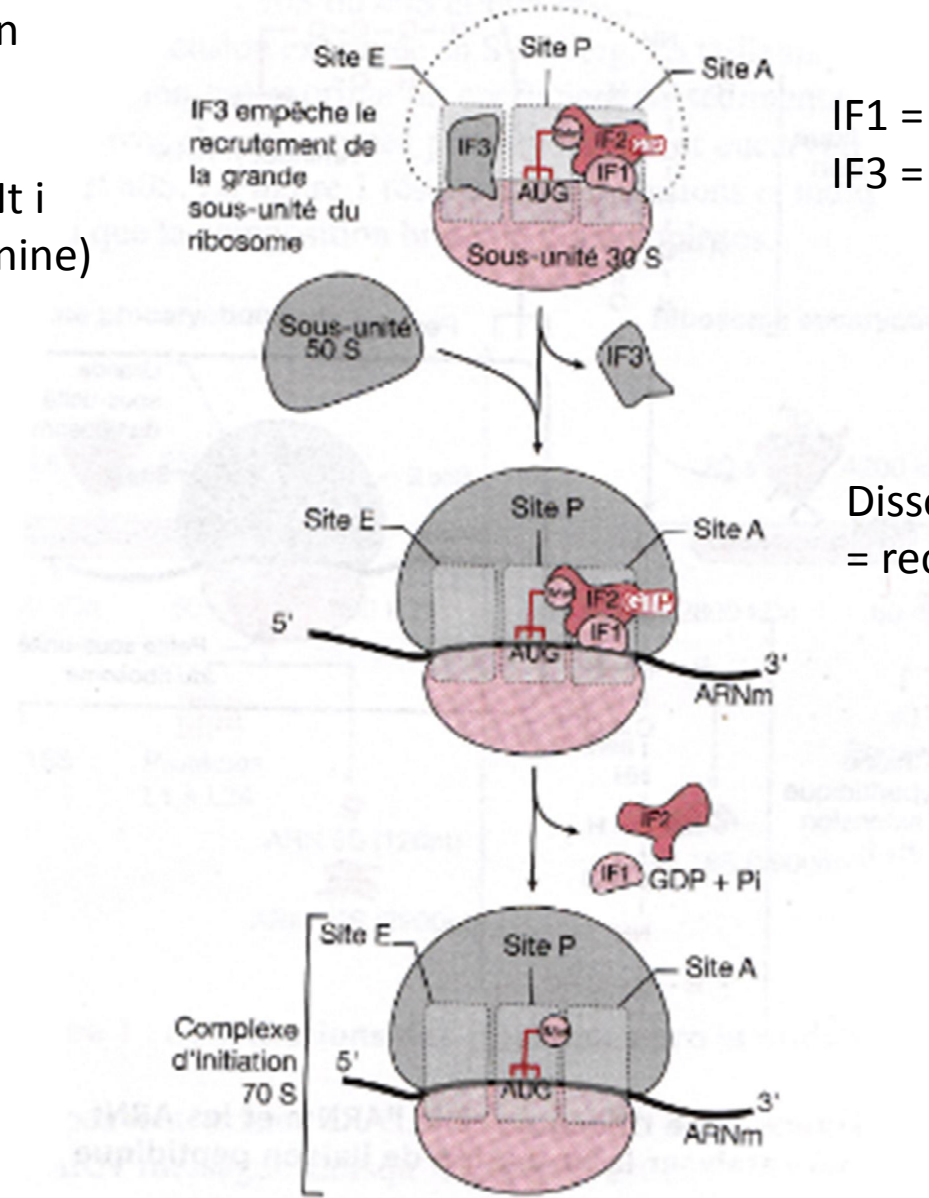


## □ Initiation de la traduction chez les procaryotes

Hybridation de l'ARNr 16S et de la séquence RBS (ou Shine Dalgarno)

3 facteurs d'initiation

IF2 reconnaît ARNt<sup>i</sup>  
(N-formyl méthionine)  
=> site P



IF1 = stabilisateur

IF3 = maintient 30S dissociée

Dissociation de IF3  
= recrutement de 50S

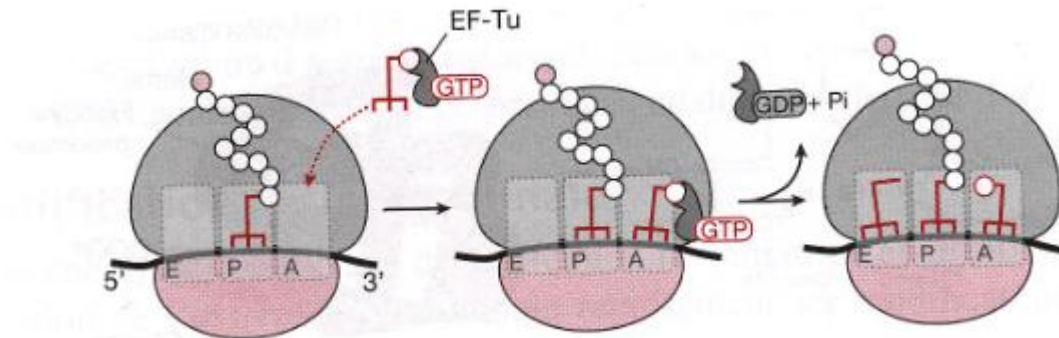


## ❑ Initiation de la traduction chez les eucaryotes

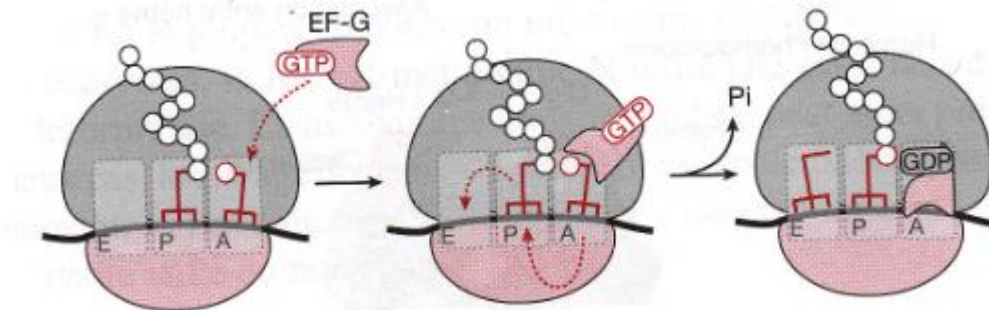
- “ 40S se fixe en 5' des ARNm sur la coiffe (Facteurs eIF)
- “ Fixation de l'ARNt initiateur (Met)
- “ 60S s'assemble dans un 2<sup>ème</sup> temps au niveau du codon initiateur (Séquence de Kozac)

- Elongation de la traduction chez les procaryotes (facteurs EF-Tu et EF-G)  
chez les eucaryotes (facteurs eEF-1 et eEF-2)

- Positionnement de l'aa-ARNt au niveau du site A : Hydrolyse du GTP lié à **EF-Tu**



- Transfert de l'aa-ARNt du site A au site P : **EF-G** (GTP hydrolysé)



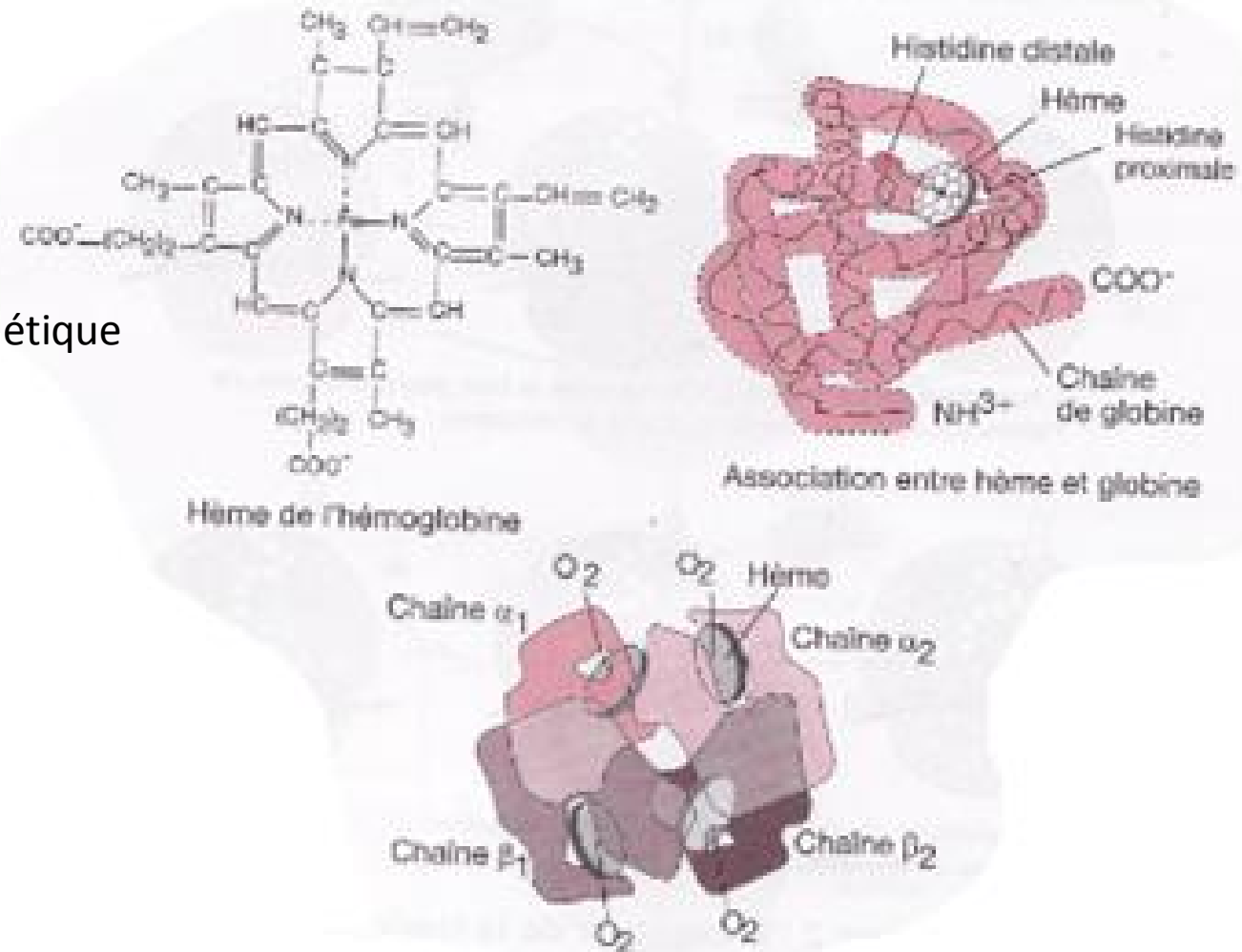
Codon STOP : ARNt chargé remplacé par facteur de terminaison :  
hydrolyse liaison ARNt - Protéine

## ❑ Modification post-traductionnelles

❖ Modification par l'ajout d'un groupe réactif :

### Hémoglobine

Porphyrine + Fe  
= groupement prosthétique



Groupement prosthétique = molécule non protéique liée à la protéine et indispensable à son fonctionnement

## Glycosylation

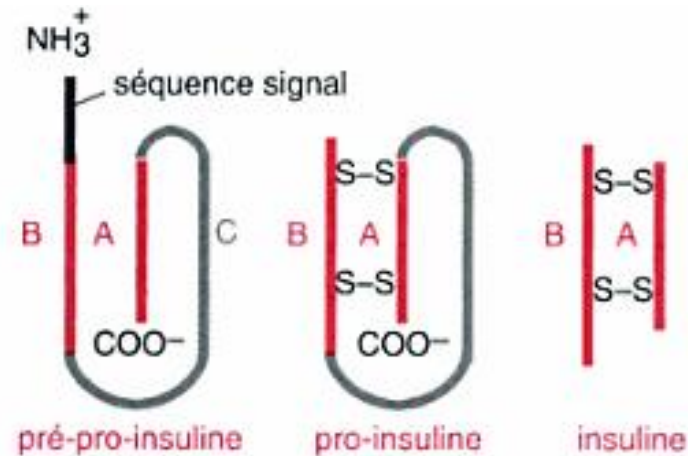
- Ajout de sucres sur des acides aminés particuliers
- Rôle dans le repliement protéique et Contrôle qualité

N-glycosylation (Asparagine) : mécanisme co-traductionnel, 14 sucres  
O-glycosylation (Sérine ou Thréonine) : mécanisme post-traductionnel

- Importance en production de protéines recombinantes

## □ Modification post-traductionnelles

- ❖ Modification par protéolyse : exemple la maturation de l'insuline



- ❖ Modification aminoterminale

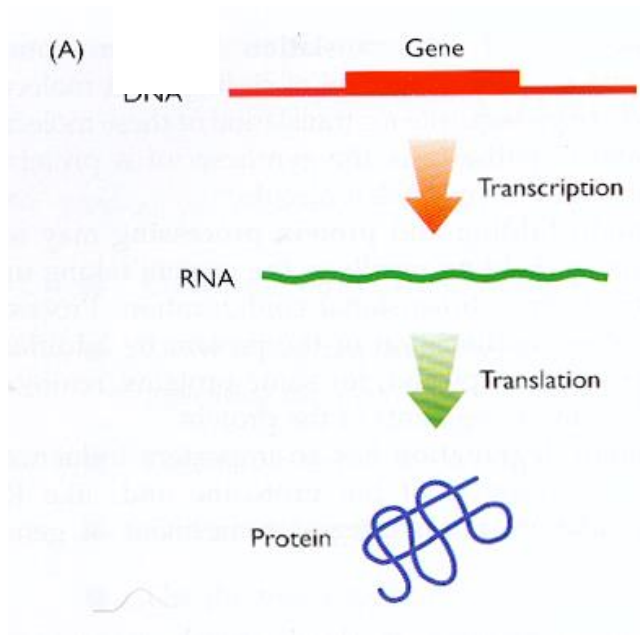
Clivage de la (Formyl-)méthionine N-terminale

Méthionine aminopeptidase

Élimination de la méthionine quand l'aa suivant est petit ou non chargé

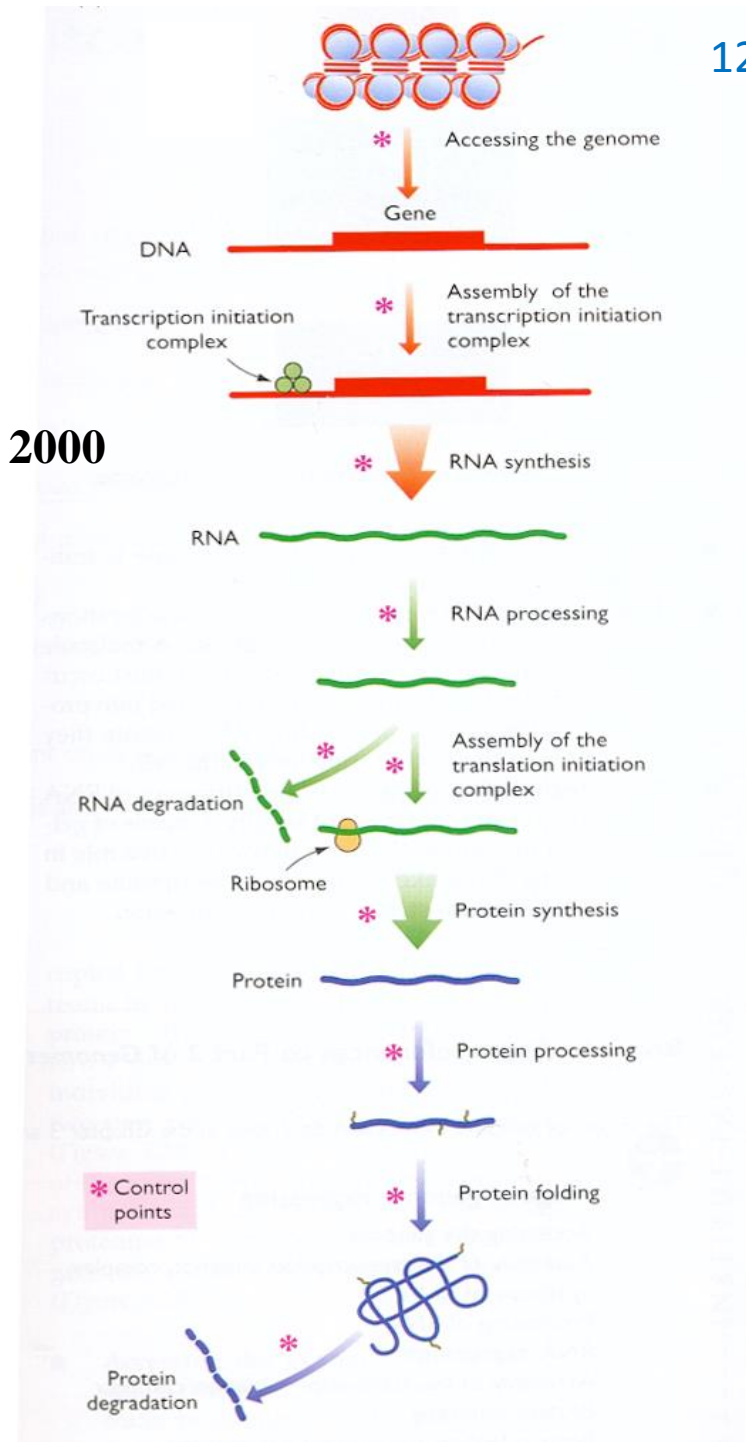
### 3.3 Bilan des connaissances « du gène à la protéine »

□ En 1960



(d'après Brown, 2002, in *genomes*)

□ En 2000

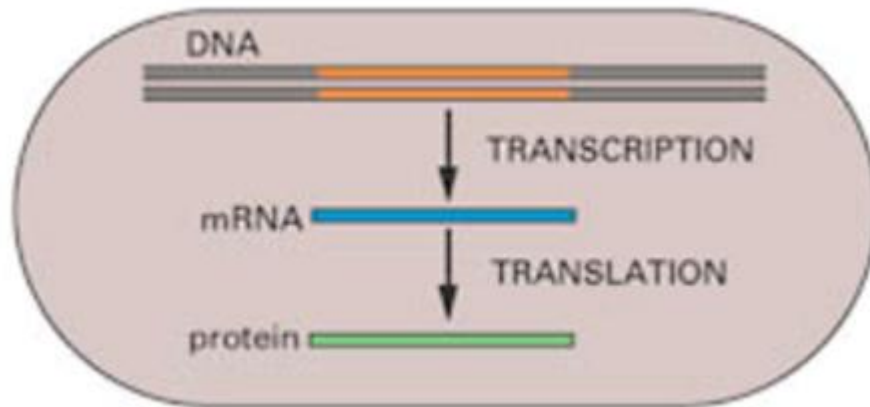


Transcription

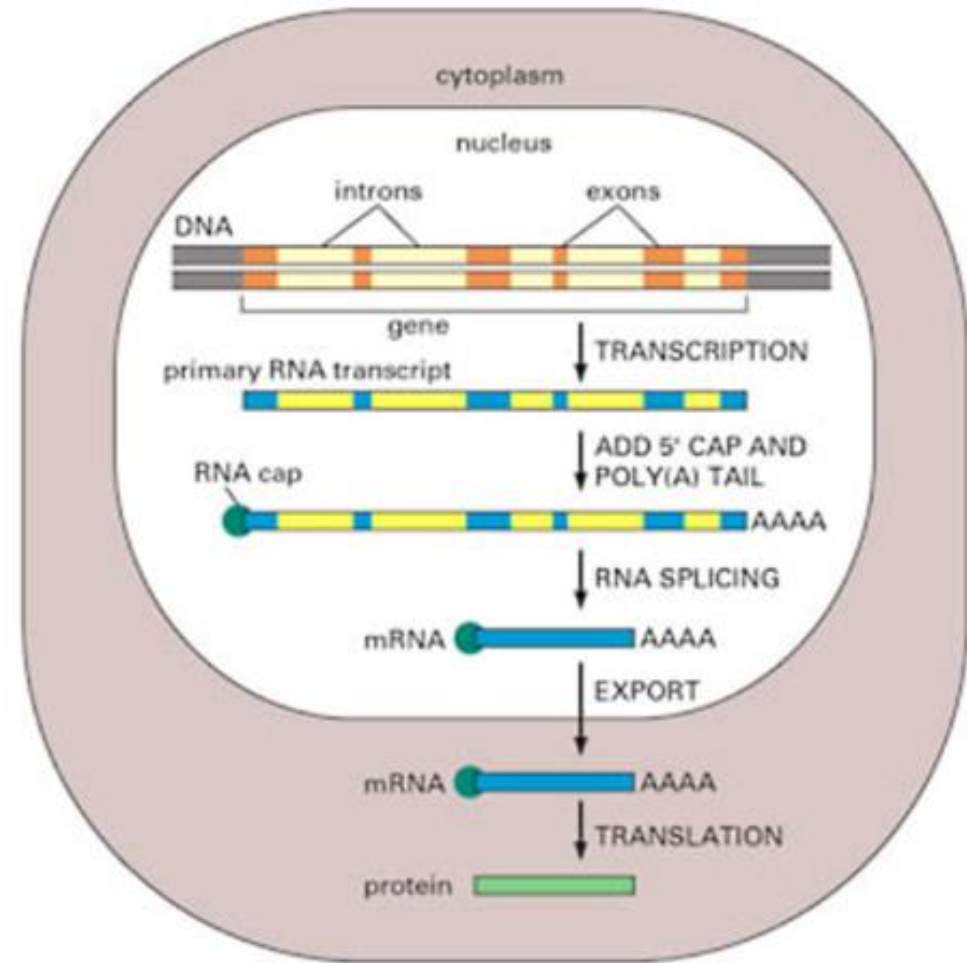
Traduction

□ Du gène à la protéine : les différences eucaryote - procaryote (Schéma simplifié)

### PROCARYOTE



### EUCARYOTE

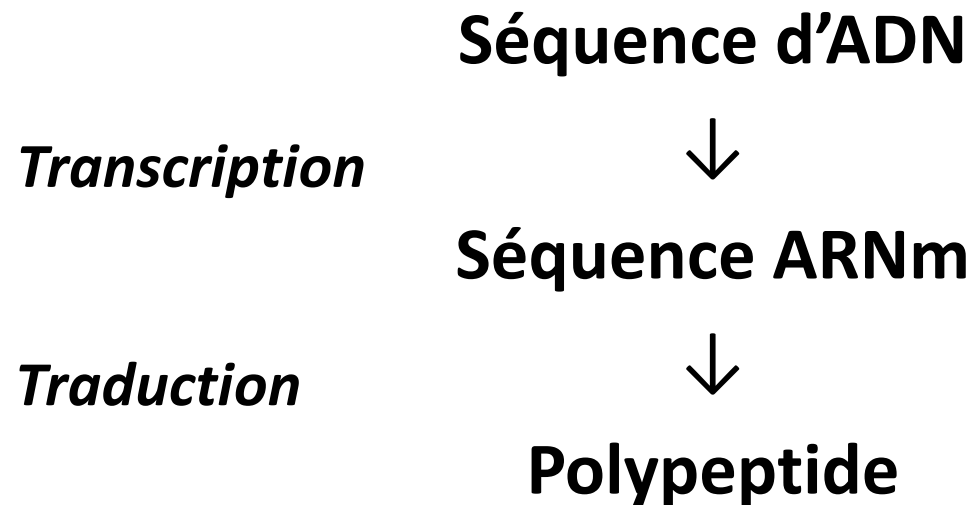


# IV Mutations, Réparation ADN et Régulations

123

## 4.1 Mutations

- ❖ Au cours de la réplication du génome ou indépendamment
- ❖ Mutation avec ou sans conséquence sur le phénotype





Mutations

## Gene normal

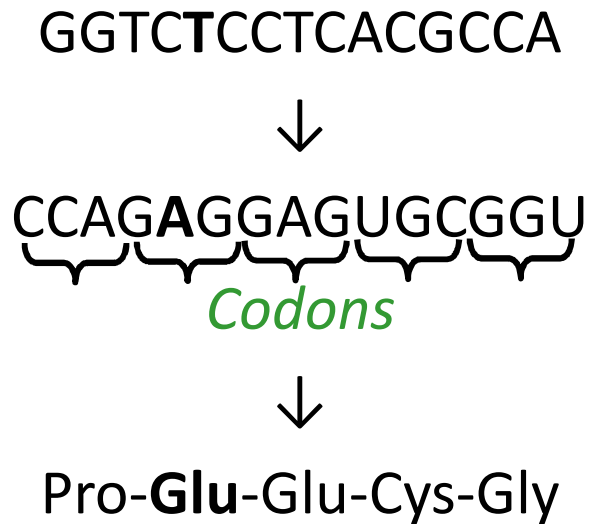
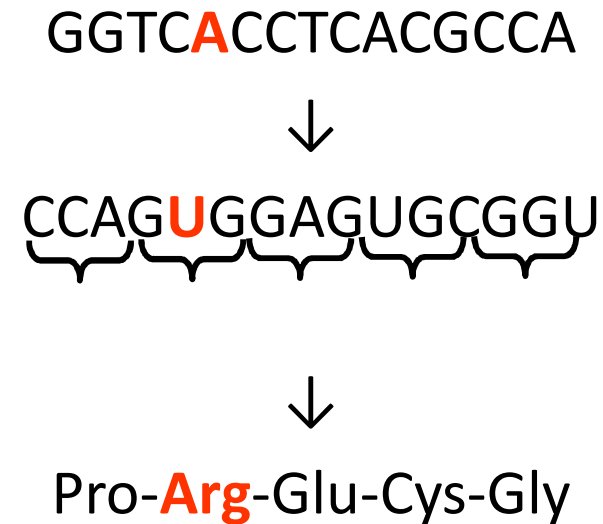
*ADN (brin anti-sens)*

GGTCTCCTCACGCCA

*ARNm*CCAGAGGGAGUGCGGU*Codons**Polypeptide*

Pro-Glu-Glu-Cys-Gly

*Acides aminés*

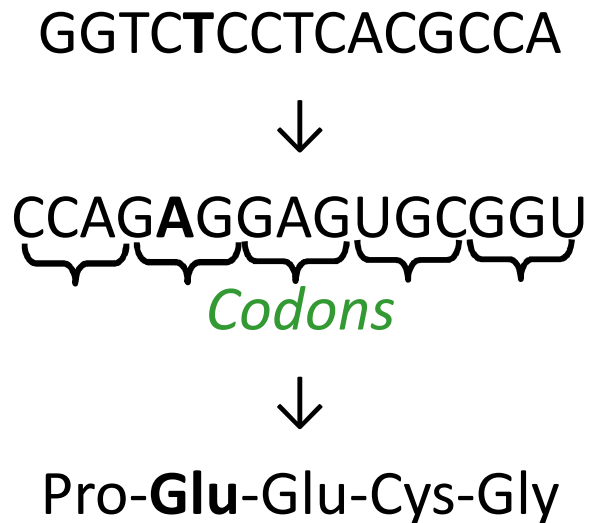
Mutations : Substitution**Gène normal****Substitution**

- Un seul codon affecté
- La gravité de la mutation dépend de l'acide aminé touché (rôle essentiel dans la structure et la fonction de la protéine?)

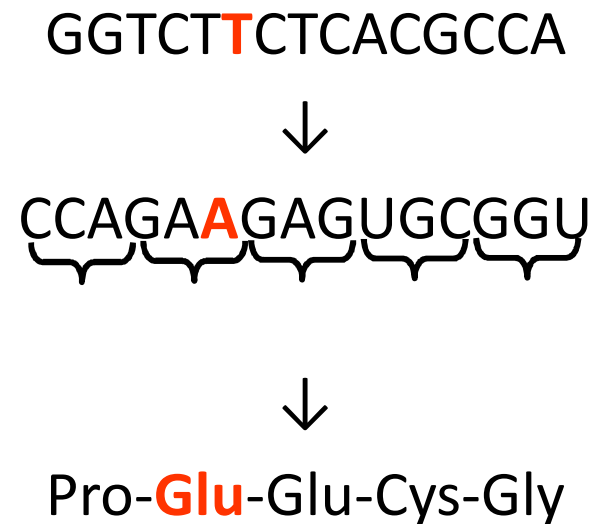
## Mutations : Substitution

- Code génétique dégénéré : Dans une grande partie des cas, si la mutation touche la 3<sup>ème</sup> base du codon, elle est sans conséquence sur la protéine finale

### Gène normal



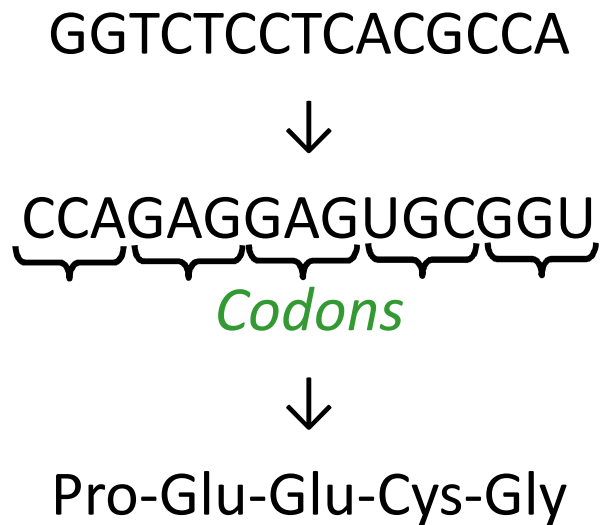
### Substitution



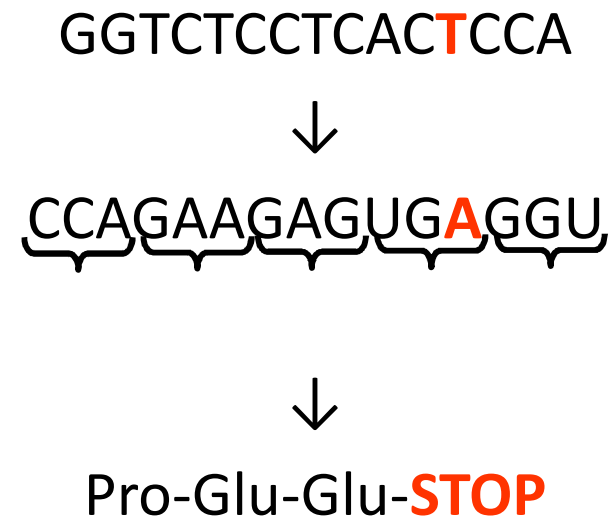
## Mutations : Substitution

- Le code génétique contient des codons STOP : terminaison précoce de la traduction

### Gène normal



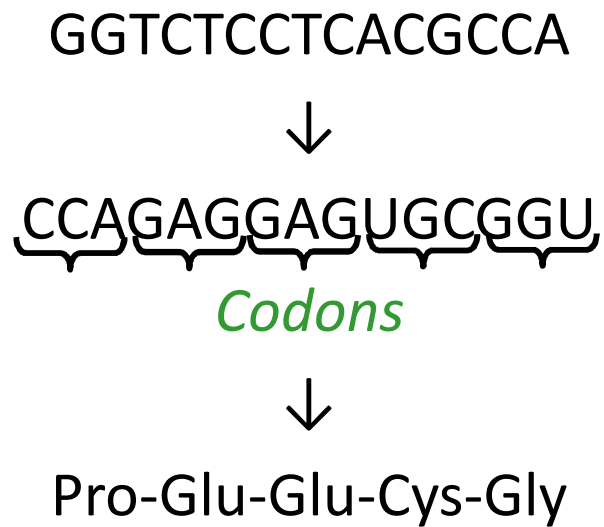
### Substitution



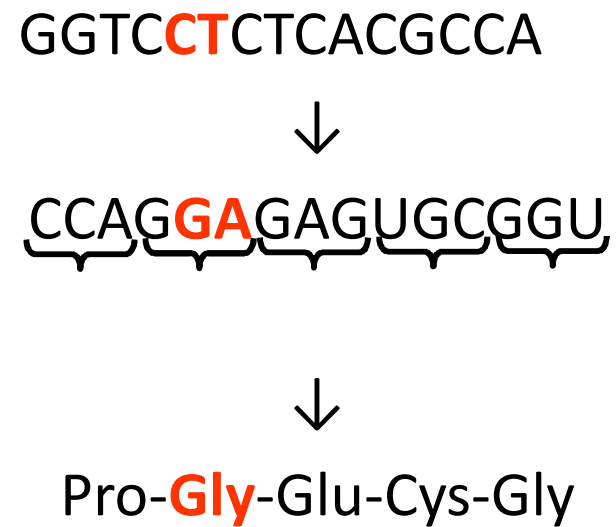
## Mutations : Inversion

- Mutations minoritaires

### Gène normal



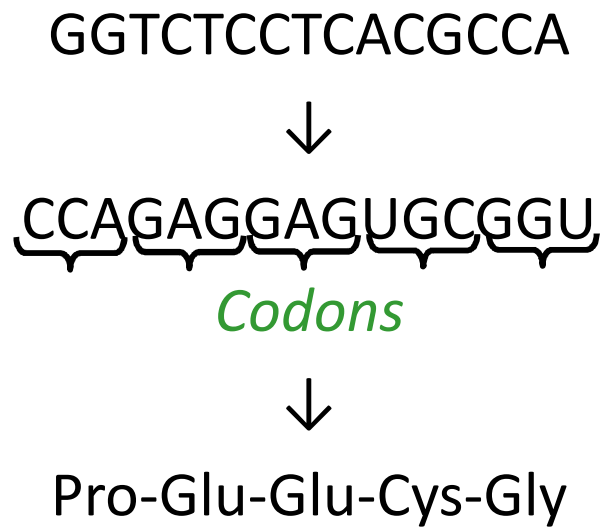
### Inversion



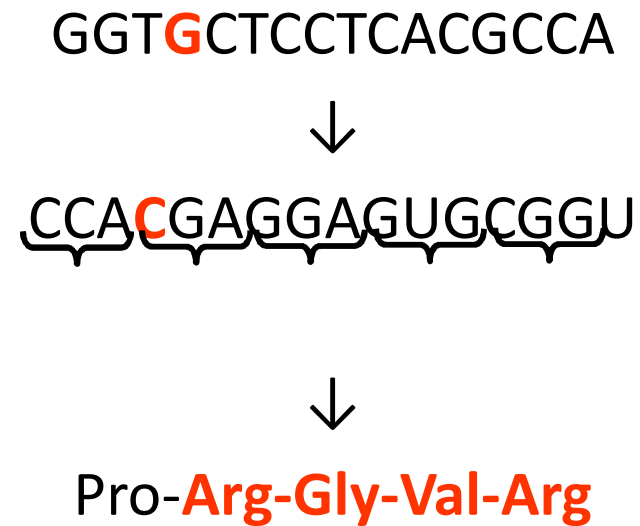
## Mutations : Insertion

- Décalage du cadre de lecture de la traduction

### Gène normal



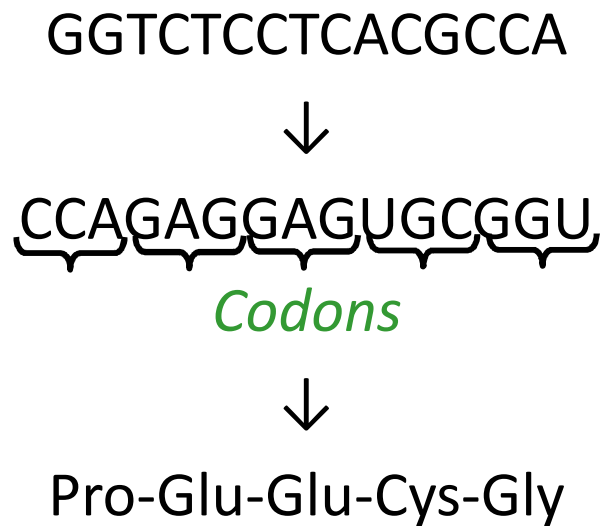
### Insertion



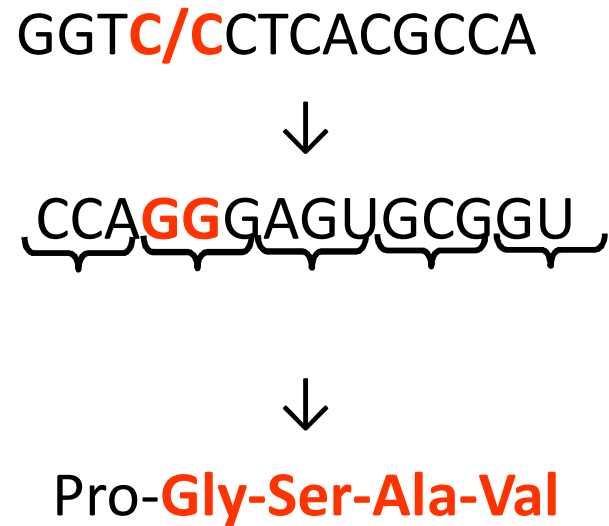
## Mutations : Délétion

- Décalage du cadre de lecture de la traduction

### Gène normal



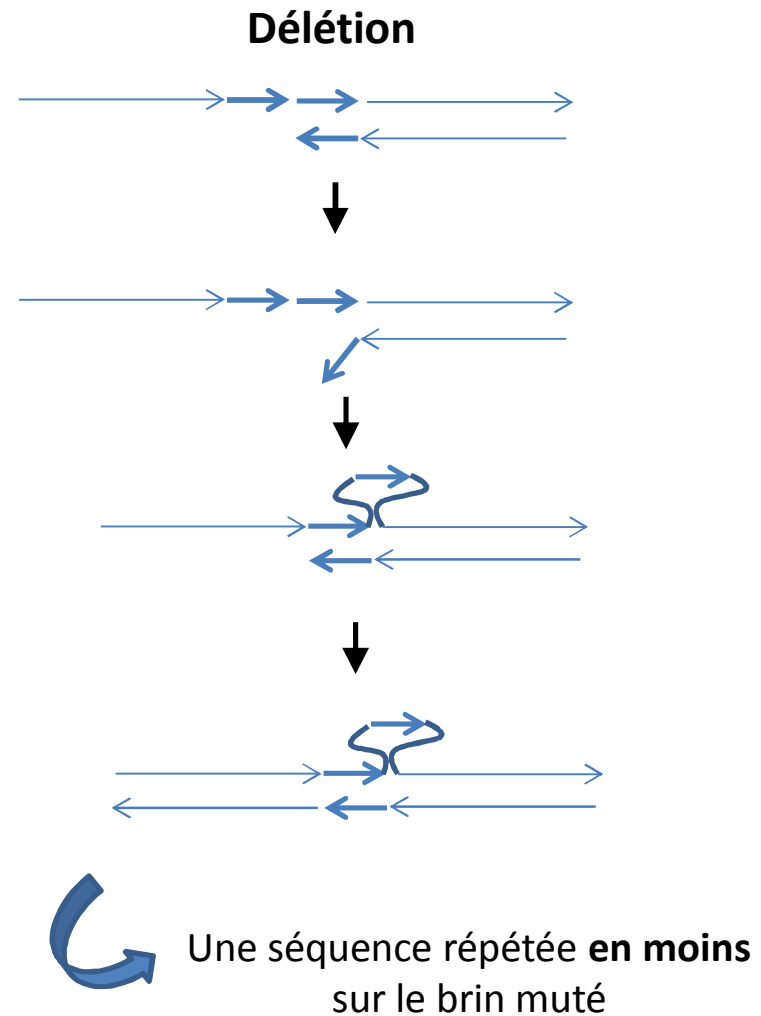
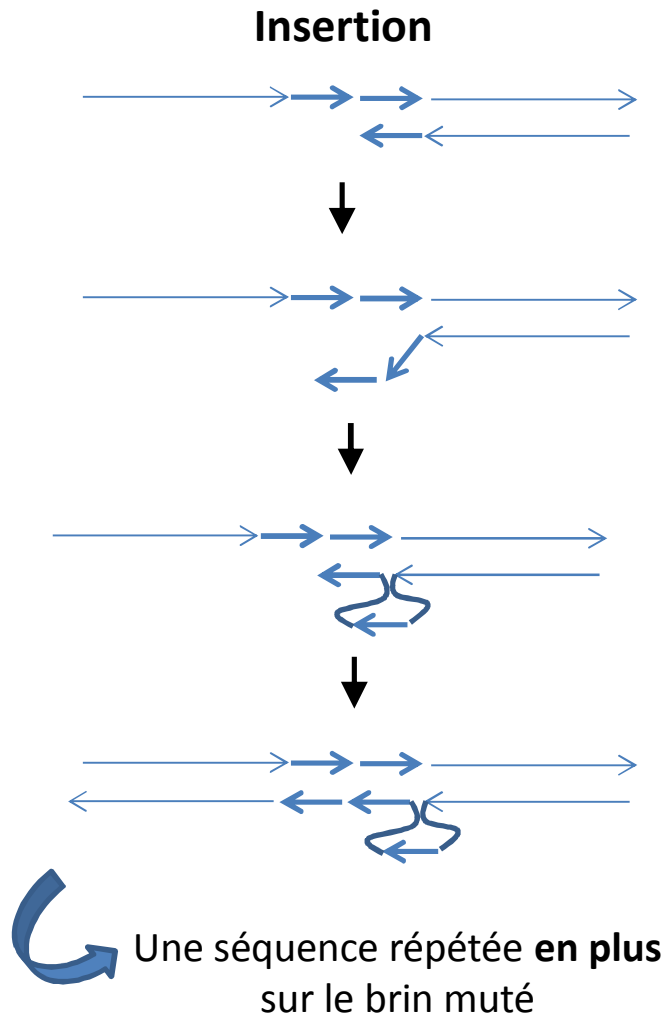
### Deletion



## Mutations : Insertion / Délétion

Glissement de la chaîne transcrite ou pendant la réplication (Modèle de **Streisinger**)

=> Variation de taille des microsatellites

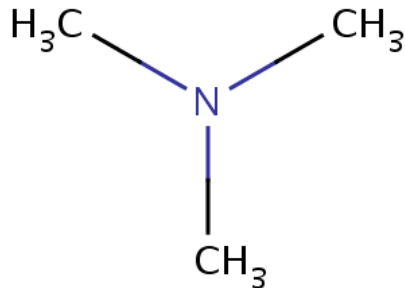




## Mutations : Exemple

**Triméthylaminurie** (Maladie héréditaire : syndrome de l'odeur de poisson pourri)

➔ Grande quantité de triméthylamine dans l'haleine, l'urine et la sueur



❖ Mutations du gène FMO3 sur le chromosome 1

➔ Flavine mono-oxygénase 3 anormale

Triméthylamine (odeur de poisson) => triméthylamine N-oxyde (inodore)

- Mutation faux-sens P153L (CCC muté en CTC)

- Mutation non-sens 148G => STOP

- Mutation délétion d'une G : décalage du cadre de lecture (séquence d'aa changée entre les positions 188 et 192 + codon STOP précoce)

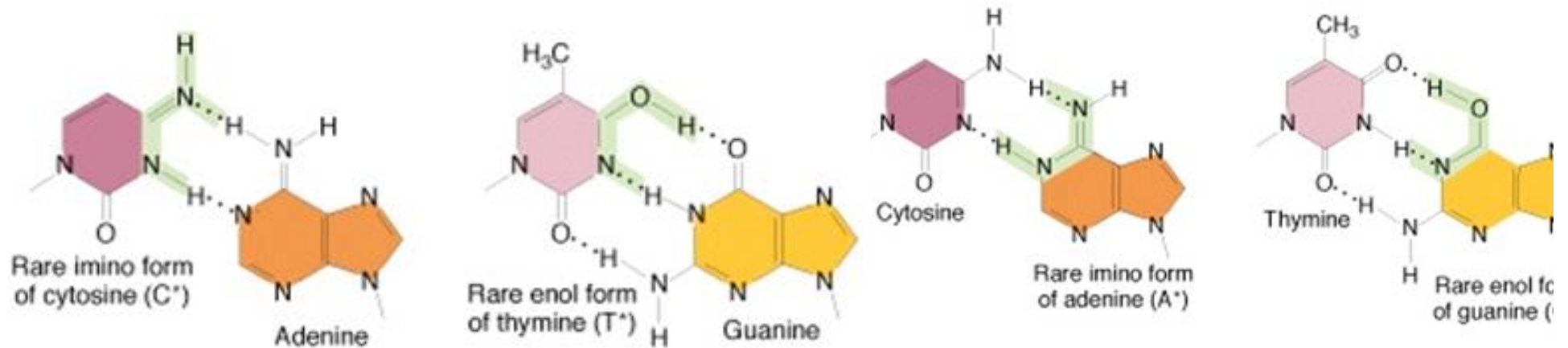
❖ Mutations différentes conduisent à cette maladie, impossible de proposer un test de diagnostic général

❖ Solution : Régime pauvre en choline (sans oeufs, foie, soja, choux, poisson ...) afin d'éviter que la choline soit transformée en TMA par les bactéries de l'intestin

## Mutations :

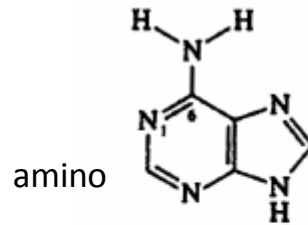
### Mécanismes naturels

- ❖ Tautomérisation : changement d'un groupe donneur de liaison H en un groupe accepteur  
=> changement de nt au cours de la réplication

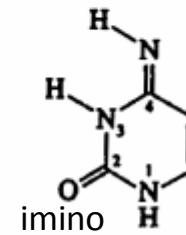
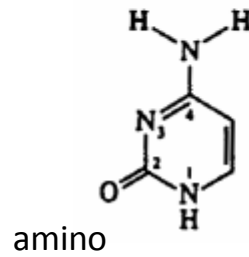


- ❖ Dépurination => Rupture de la liaison entre le désoxyribose et une base purique (A ou G)  
Désamination des cytosines => uraciles
- ❖ Insertion, délétion : glissement de la chaîne en cours de synthèse (ou copiée)  
taille des séquences microsatellites

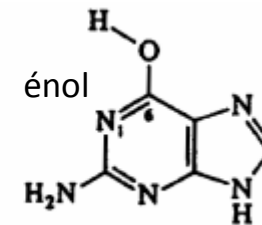
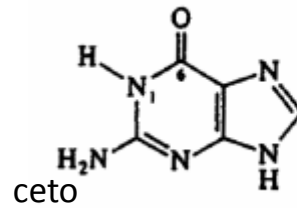
Adénine



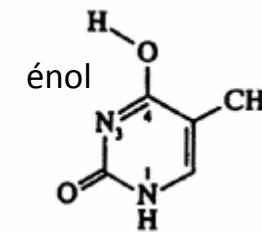
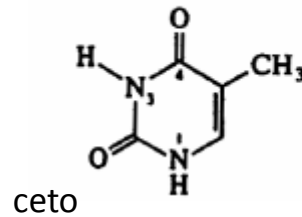
Cytosine



Guanine



Thymine



Formes majoritaires

Formes minoritaires

## Induction des mutations

### ❖ Agents chimiques

- Substances chimiques (acide nitreux ...) : mutations ponctuelles
- Analogues des bases (5BU, aminopurine...) : tautomérisation
- Intercalents de l'ADN (BET, acridine...) : étirement de l'ADN (insertion)

### ❖ Agents physiques

- Rayons X et gamma : coupures ou modifications de bases
- Rayons UV : formation de dimères de pyrimidines (liaison cyclobutane) : blocage de la fourche de réplication

### ❖ Agents biologiques

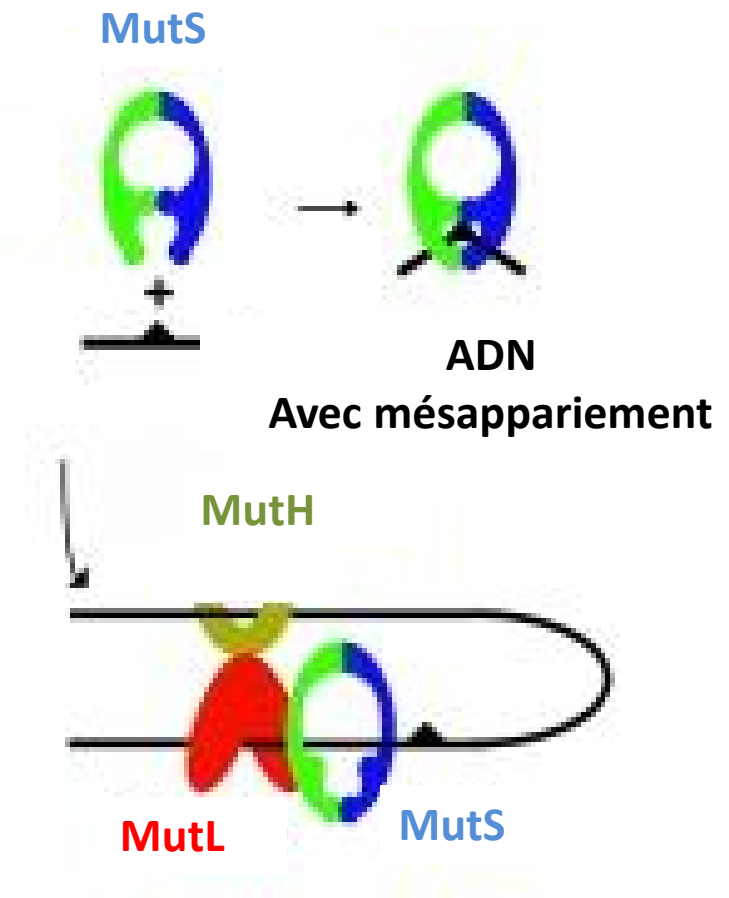
- Pas de mutations ponctuelles
- Mutation provient de l'insertion d'un fragment d'ADN (transposons, virus...) : disruption éventuelle de gènes

## 4.2 Réparation de l'ADN

### mécanisme **co-réplicatifs**

- ❖ Activité 3' - 5' de l'ADN polymérase
- ❖ Le système de réparation **MMR** (Mismatch Mutation Repair) :

- 1- **MutS** (dimère) reconnaît le mésappariement
- 2- MutS recrute **MutL** qui active MutH (endonucléases)
- 3- **MutH** reconnaît le brin non méthylé et le clive au niveau d'un site méthylé proche)  
(sites palindromiques méthylés sur le brin matrice)
- 4- Brin clivé dégradé par **exonucléase** 5'>3'
- 5- **ADN pol III** comble le trou : synthèse du brin correct



## Réparation de l'ADN

### Mécanismes **post-réplicatif**

❖ Réparation sans excision de base (minoritaire):

Rupture de la liaison phosphodiester : action de la ligase

O<sub>6</sub>-CH<sub>3</sub>(G)-ADN méthyltransférase : désalkylation des bases modifiées

Réparation des T-T : photolyase

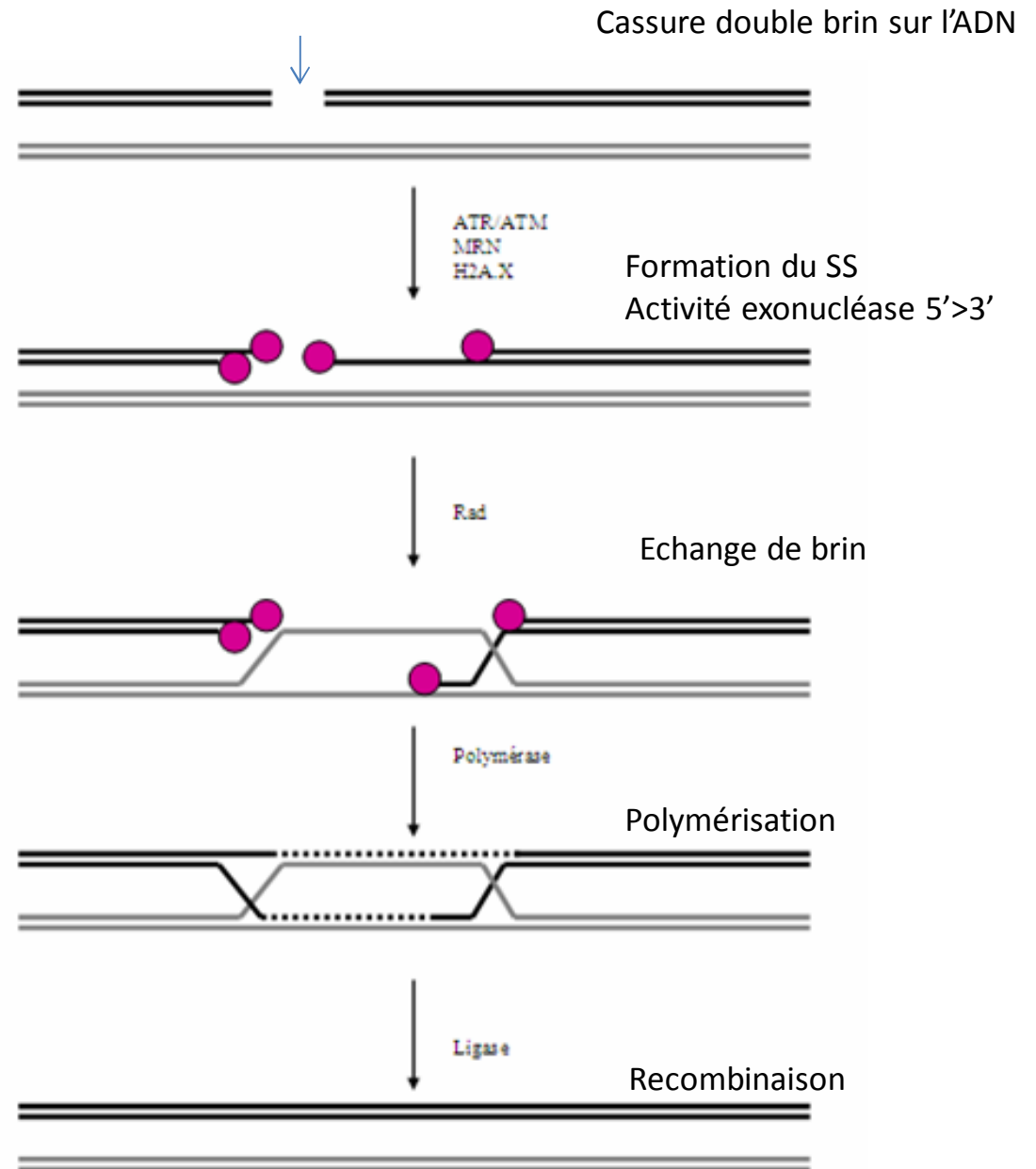
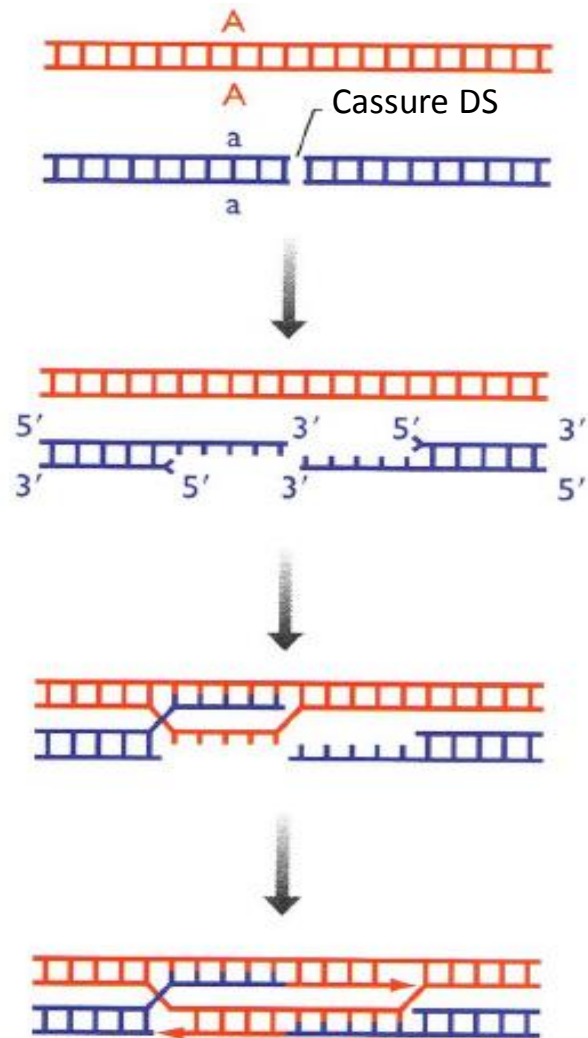
❖ Excision des bases

Action de l'ADN glycosidase => site abasique reconnu par AP endonucléase

❖ Excision des nt

Complexe protéique UvrA - UvrB recrute Uvr C (départ d'UvrA) :  
trou de 12 nt comblé par ADN pol + ligase

## ❖ Réparation par recombinaison homologue



## ❖ Transposition = mécanisme de recombinaison

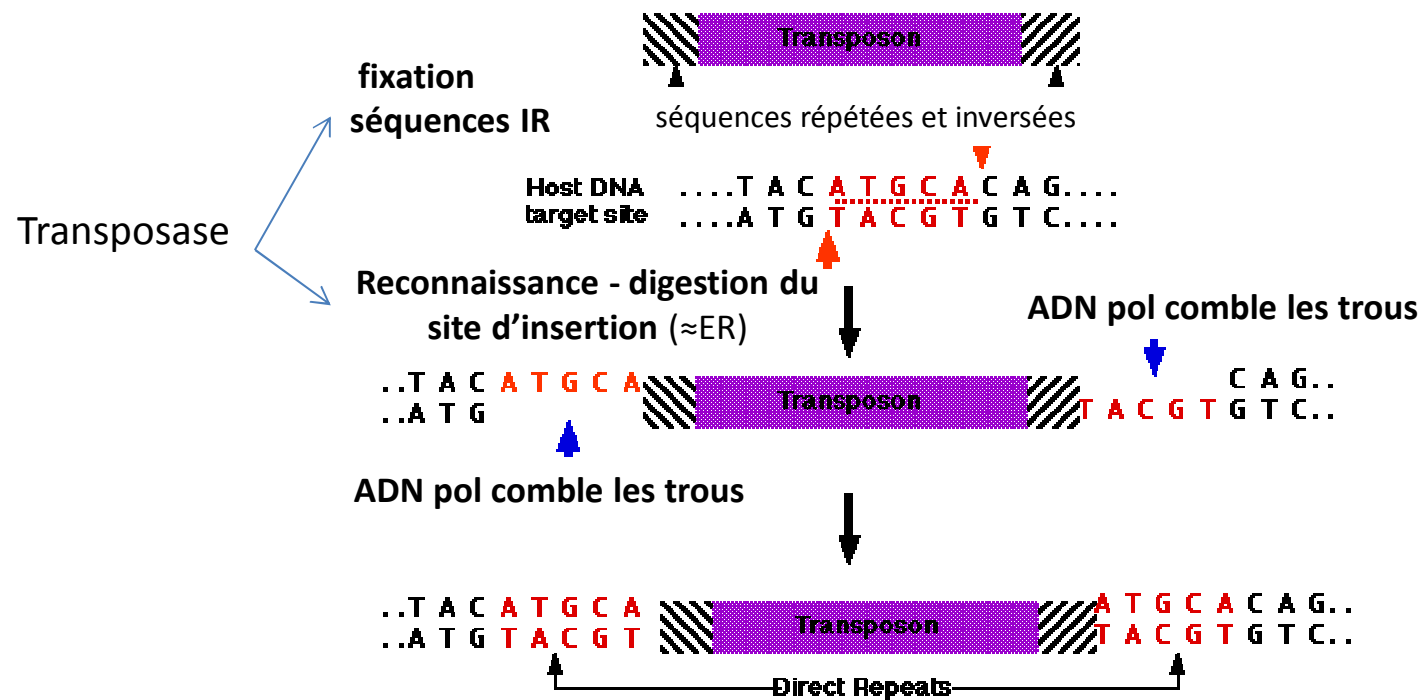
Transposon = élément mobile du génome qui possède les informations nécessaires à sa mobilité  
(Prix Nobel McClintock 1983)

### ❖ Transposons de classe II (environ 1 kb)

(Couper - Coller)

Transposons à ADN conservatifs

Transposase (gène inclus dans le transposon) ; parfois : résolvase (Copier - Coller)



Les transposons de classe II => un **gène de résistance** à un antibiotique.

Possibilité d'insertion sur un **plasmide**

Donc **possibilité de transfert de bactérie à bactérie** du gène de résistance à l'antibiotique



❖ Transposons de classe I (Copier-Coller)

Rétrotransposons (ARN),

Rétrotranscriptase,

longues séquences terminales inversées (LTR)

- LINEs (Long Interspersed element) endonucléase, rétrotranscriptase (6,5 kb)

Transcription du transposon

Traduction : association protéines synthétisées et ARN => retour dans le noyau

Endonucléase : coupure de l'ADN

Reverse transcriptase ARN => ADN : insertion du nouveau transposon dans l'ADN

- SINEs (0,3 kb)

Utilisation d'enzymes fournis par d'autres transposons

❖ Transposons de classe III (0,4 kb)

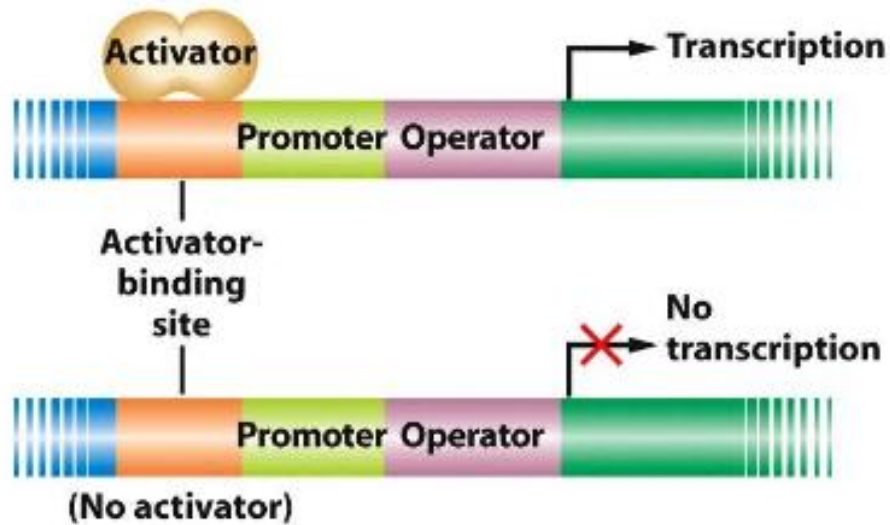
**Miniatures Inverted Repeats Transposable Element**

Séquences IR courtes (Couper - Coller)

## 4.3 Régulations de l'expression génique

### □ Contrôle de la transcription chez les procaryotes

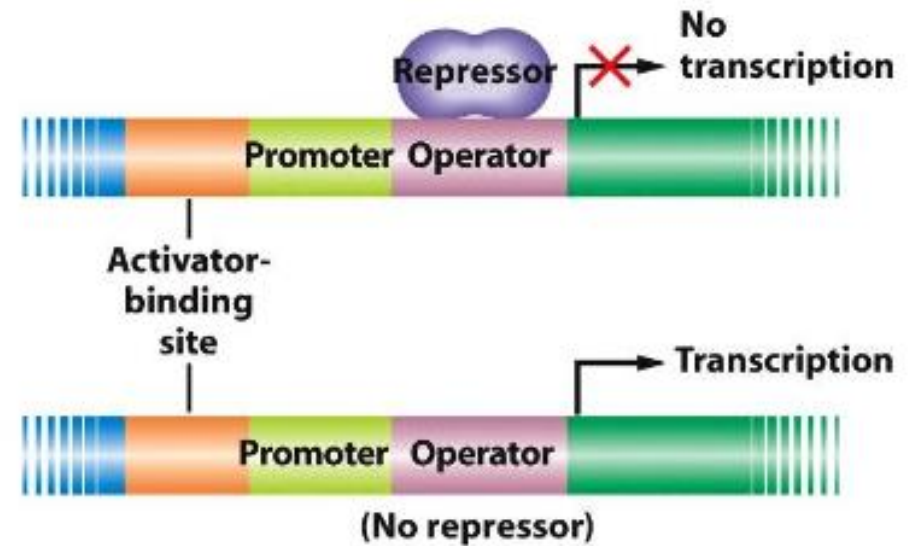
#### Régulation positive



*Introduction to Genetic Analysis, Ninth Edition*  
© 2008 W. H. Freeman and Company

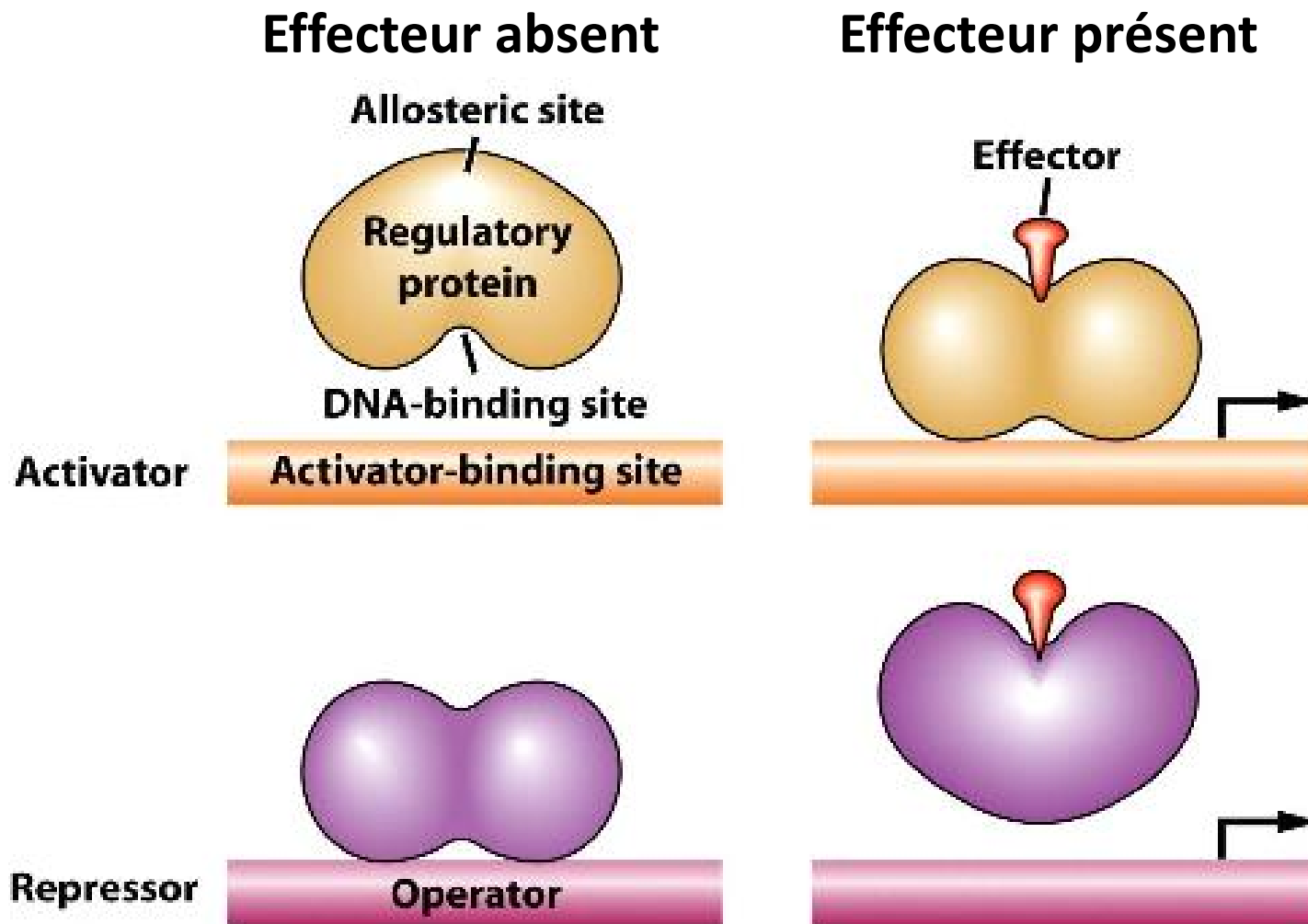
+ FR => stimulation

#### Régulation négative



- FR => inhibition

## Effecteurs allostériques



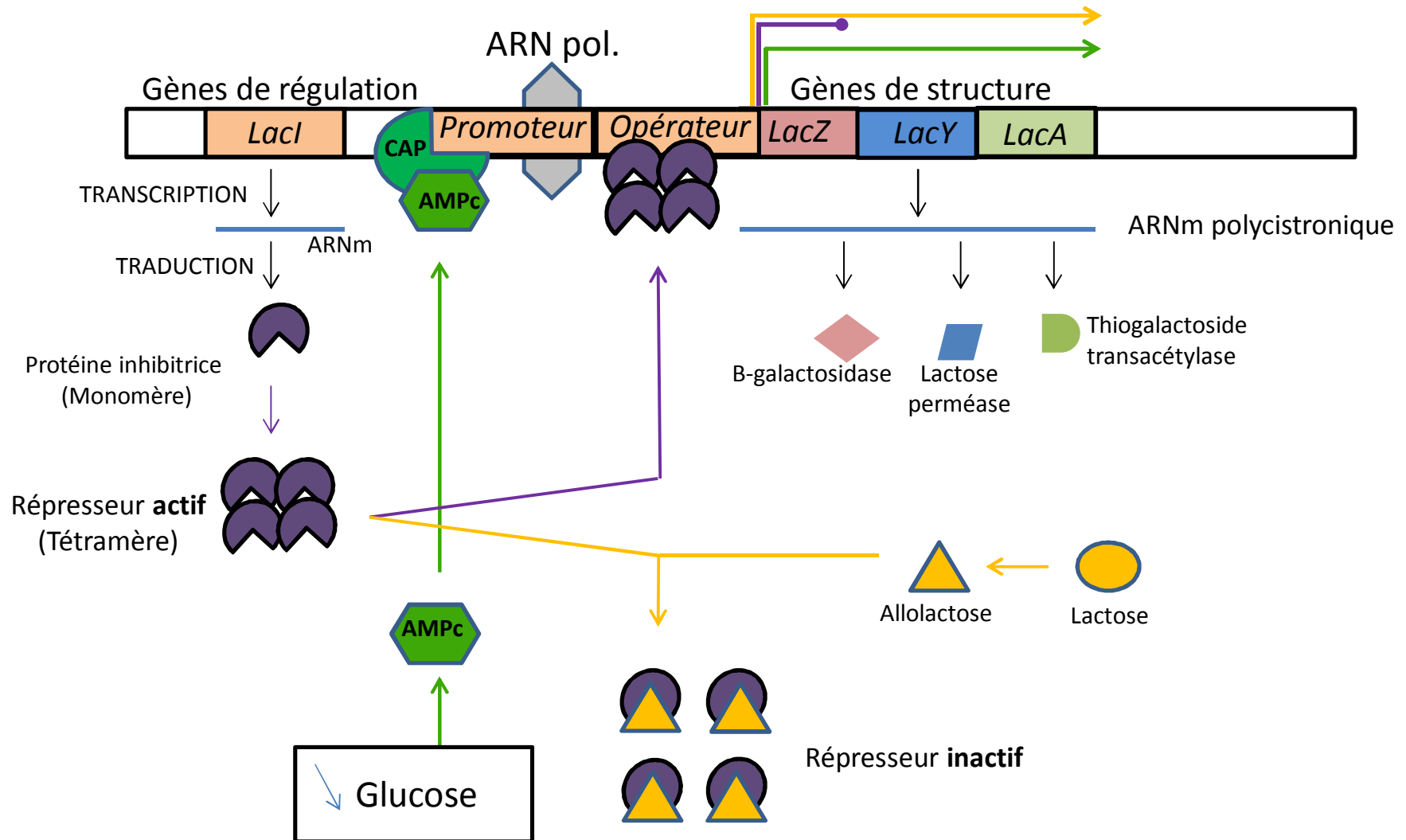
*Introduction to Genetic Analysis, Ninth Edition*  
© 2008 W. H. Freeman and Company

**Fixation des Effecteurs allostériques => Modification de la capacité des activateurs et répresseurs à se lier à l'ADN**

### Définition d'un opéron :

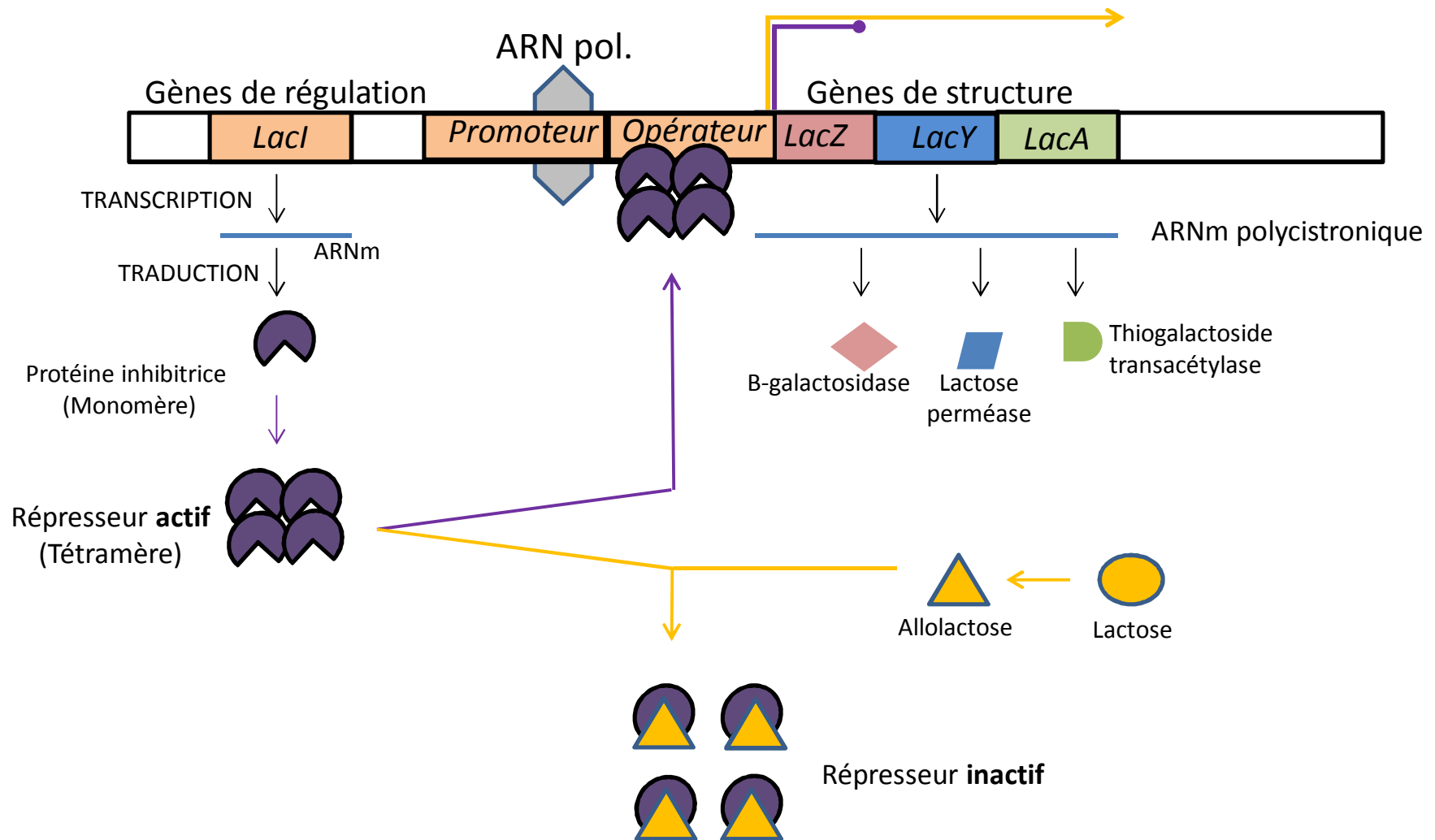
- Chez les bactéries, les gènes sont souvent groupés et organisés en unité de transcription (= opéron).
- contrôle coordonné de l'expression de ces gènes possible grâce à des protéines régulatrices. Elles se lient à l'ADN (séquences spécifiques).
- ARNm polycistronique : information nécessaire à la synthèse des différentes protéines.

# Opéron lactose : régulation positive et négative



## Opéron lactose :

- régulation négative

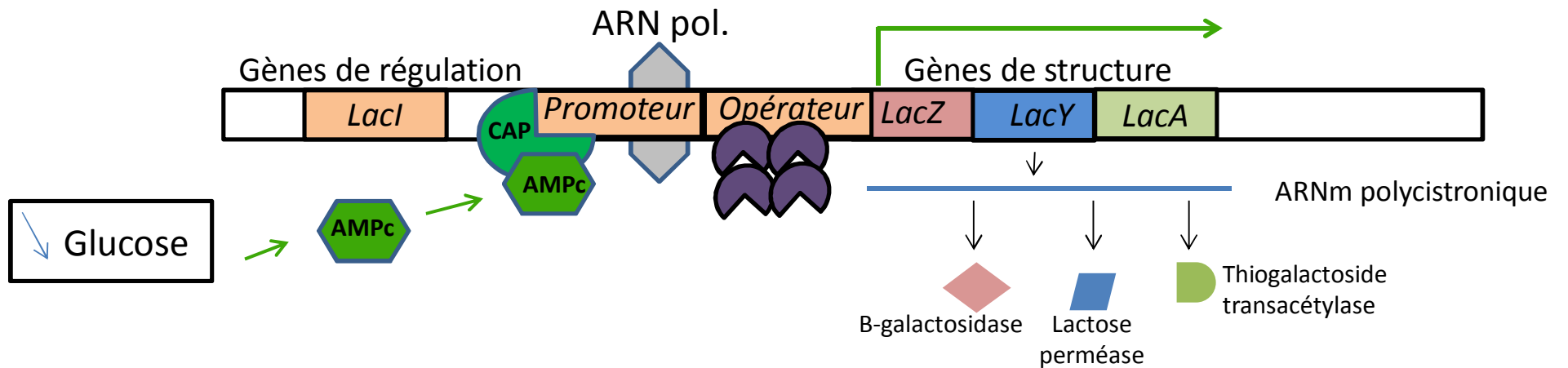
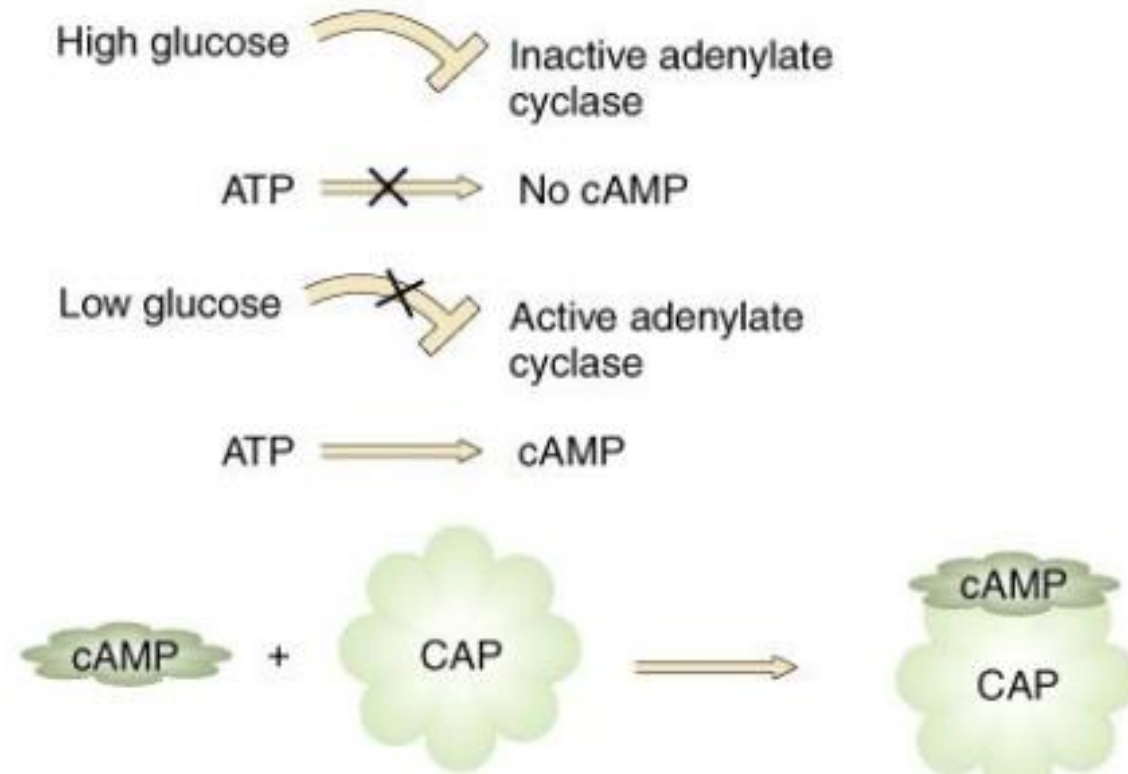


Opéron lactose :

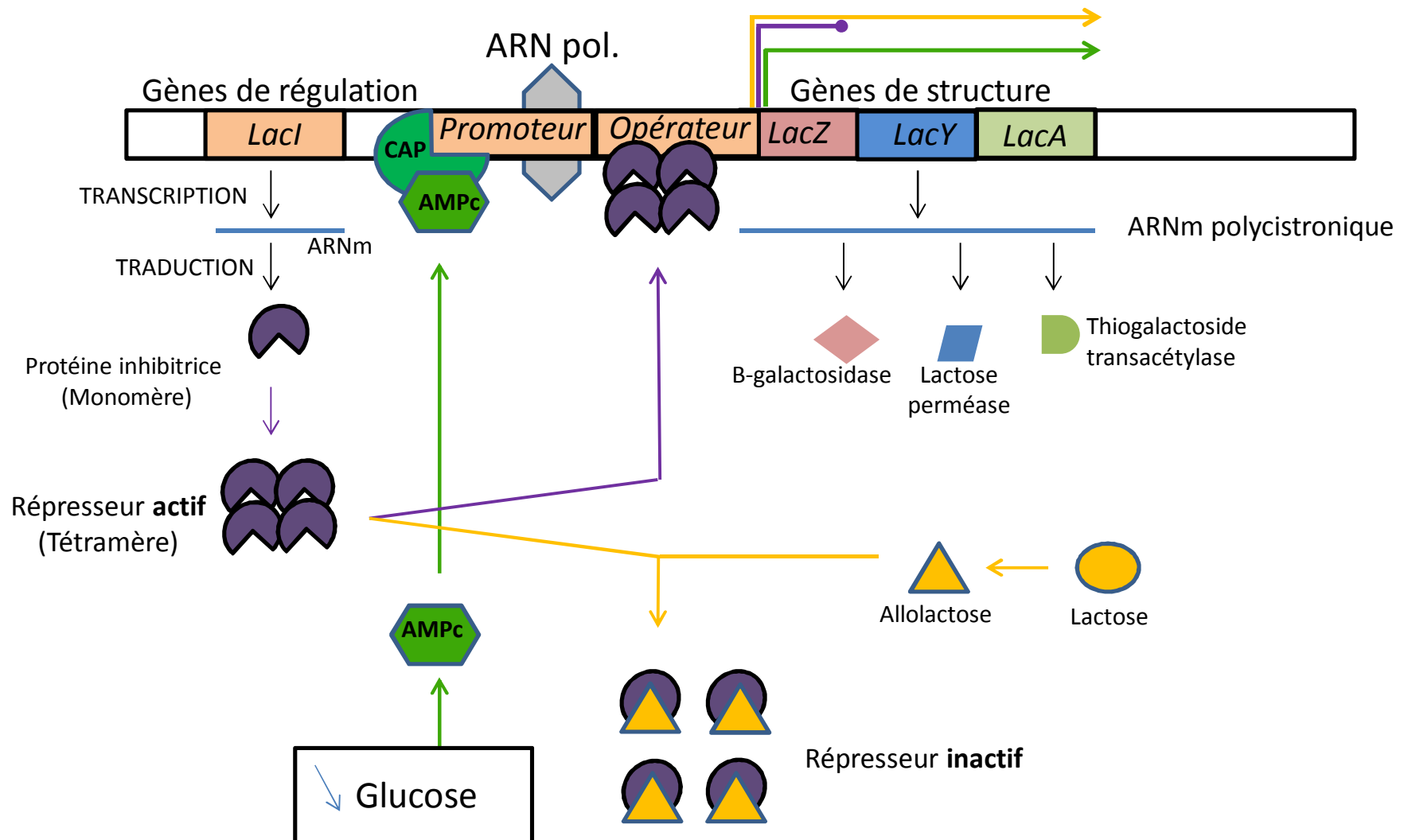
- régulation positive

**CAP**

(Catabolic activator protein)



## Opéron lactose : régulation positive et négative



**CAP** (Catabolic activator protein) Utilisation du glucose favorisée par rapport à d'autres sucres moins « rentables »

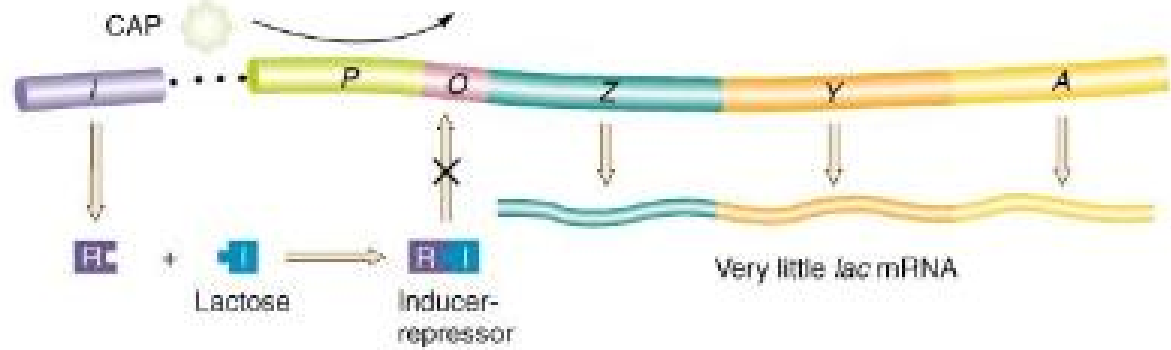
**LacI** Pas de production inutile d'enzyme en l'absence de lactose



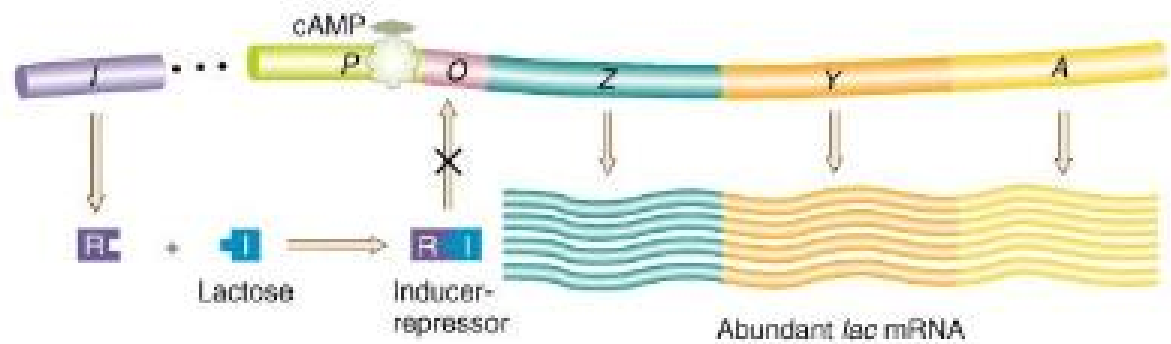
Glucose  
Pas de lactose



Glucose  
Lactose

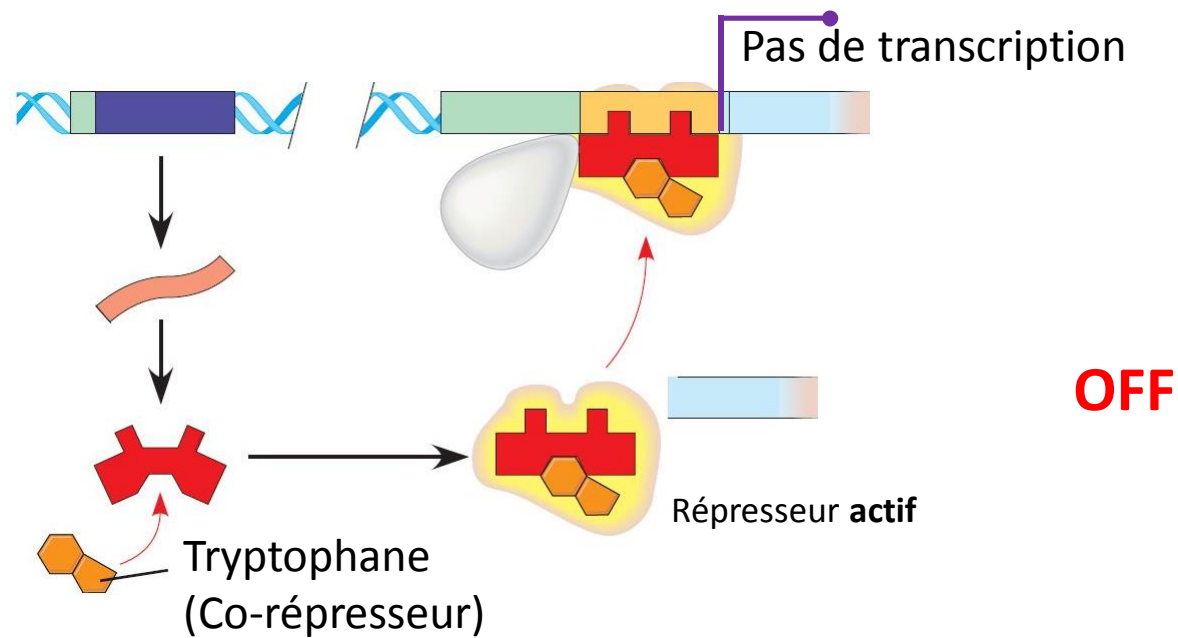
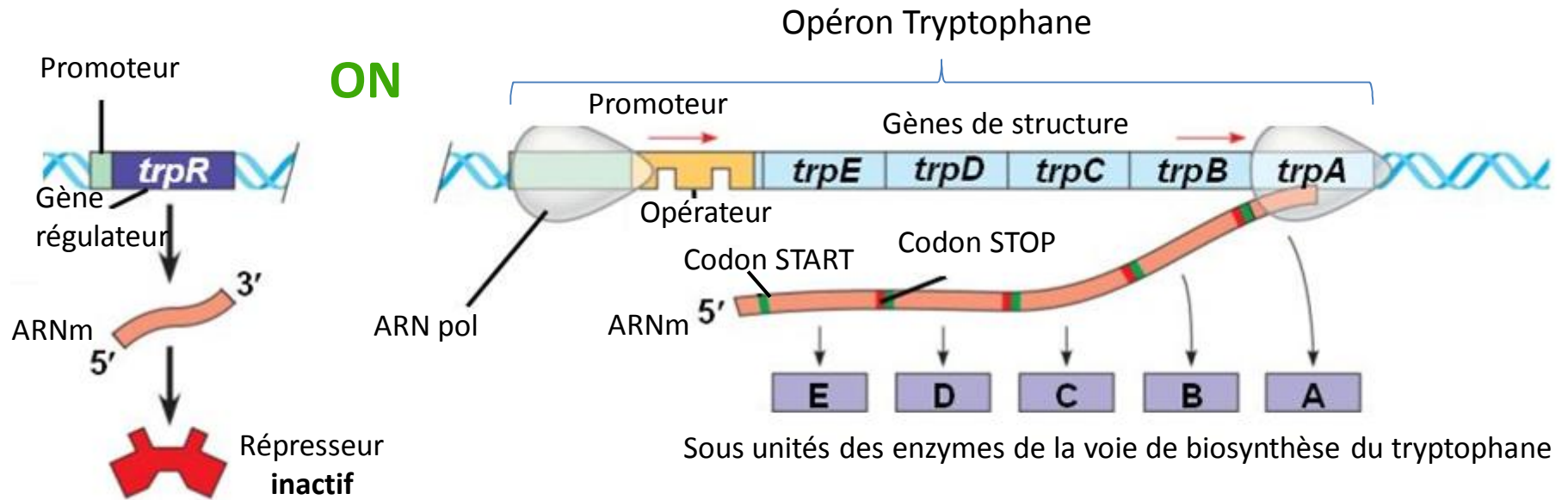


Pas de Glucose  
Lactose



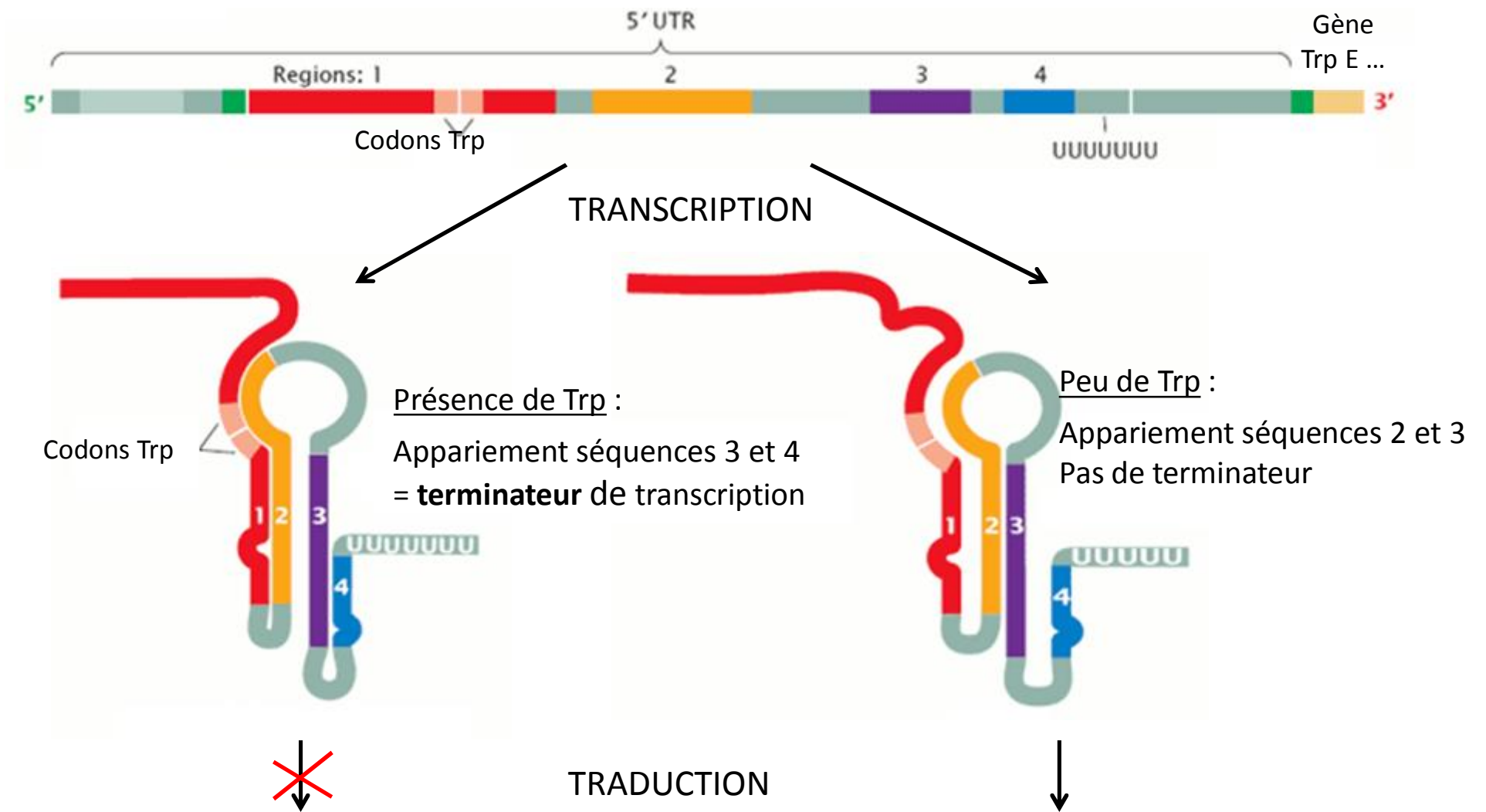
Opéron tryptophane : deux niveaux de régulation différents

❖ 1<sup>er</sup> niveau de régulation



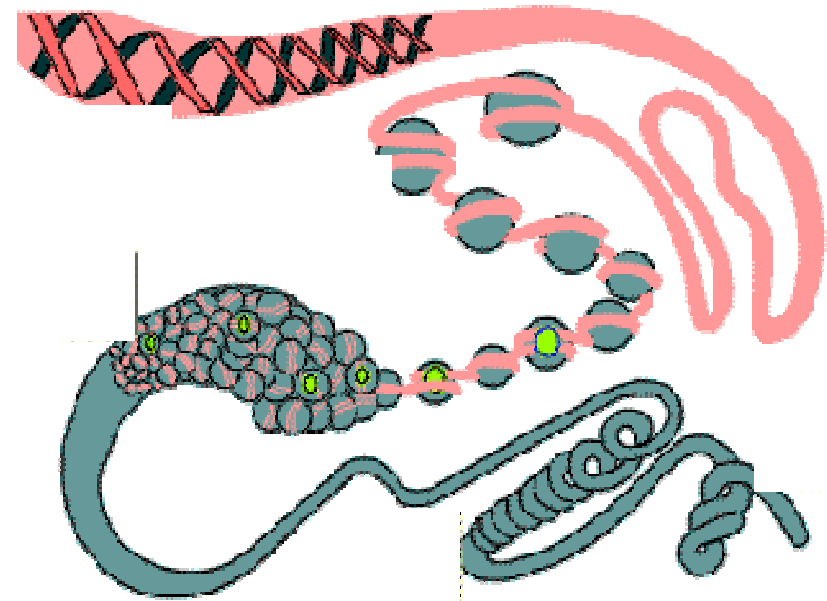
❖ 2<sup>ème</sup> niveau de régulation Mécanisme d'atténuation

(Couplage transcription - traduction)



Rappel :

- ❖ L'ADN chromosomique est empaqueté à l'intérieur du noyau à l'aide des nucléosomes [ = ADN - / Histones + ] ayant une structure en collier de perles  
< = > **euchromatine**
- ❖ Les nucléosomes se replient pour former une fibre de 30 nm qui forme des boucles de 500 nm  
<=> **hétérochromatine**



Les différences entre euchromatine et hétérochromatine ne se limitent pas qu'au niveau du degré de la compaction de l'ADN

# V Epigénétique

- Contrôle de la transcription chez les eucaryotes

## ☐ Contrôle de la transcription chez les eucaryotes

### Différences entre Euchromatine et Hétérochromatine

euchromatine : peu ou pas méthylée

hétérochromatine : méthylée (H3K9me ; H3K9me3 et H3K27me3) + HP1

“ hétérochromatine facultative : condensées < = = > décondensé permettant la transcription.

“ l'hétérochromatine constitutive : gènes silencieux (centromères et télomères).

Les histone-lysine **N-méthyltransférases** (qui méthylent K4, K9 et K27) : G9a (=EHMT2), SUV39H1, SUV39H2 et SETDB1 => **régulation +**

Les histones **déméthylases** : JMJD2C « Jumonji domain-containing 2C » (= KDM4C) et JMJD1A (= KDM3A) => **régulation -**

TIP60 (aussi appelée KAT5) : "TAT-Interacting protein of 60 kDa" (**acétylation des histones**)  
 NuRD ("Nucleosome-Remodeling and Histone Deacetylase chromatin-remodeling complexes") : **Histone désacétylases 1 et 2 (HDAC1 and HDAC2)** - "Chromodomain-Helicase-DNA-binding proteins CHD3 and CHD4" (ATPase)

DNMT3A : **DNA (cytosine-5)-méthyltransférase 3A**

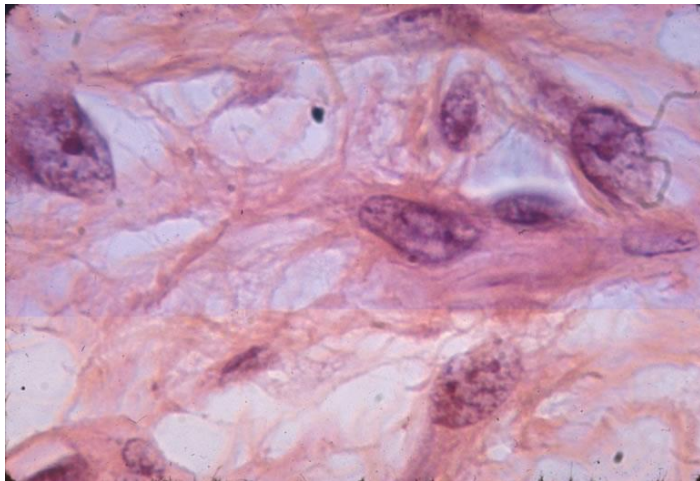
Epigénétique (épi = autour, en plus) : régulation de la transcription sans modification de la séquence de l'ADN. (Une définition parmi d'autres – en fonction des avancées technologiques)

Les facteurs épigénétiques responsables du processus de remodelage de la chromatine sont :

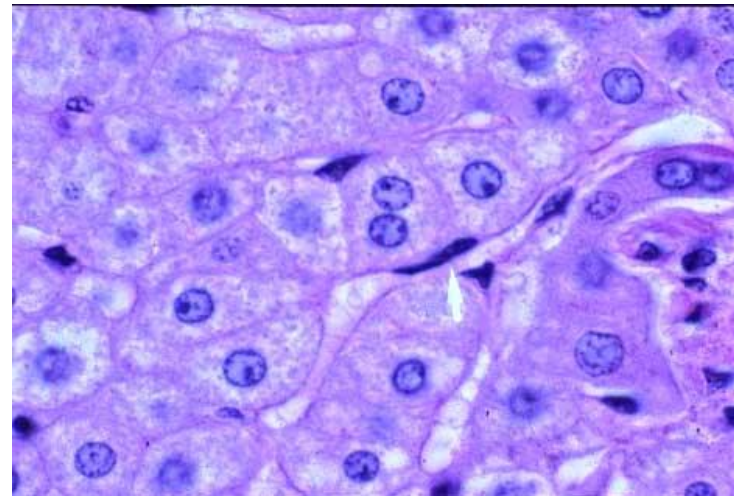
- ❖ les modifications post-traductionnelles des histones
- ❖ la méthylation de l'ADN
- ❖ l'action des ARN non codants

- ❖ Epigénétique : tous mécanismes permettant à la cellule d'avoir un phénotype différent d'une autre cellule alors qu'elle a exactement la même séquence génomique

Contrairement à la séquence de l'ADN, qui est en grande partie inchangée au cours de la vie, les profils épigénétiques varient non seulement d'un **tissu** à un autre mais changent avec l'**âge** et sont sensibles aux influences comportementales et **environnementales** (protection maternelle, pollution, qualité de la nutrition, diète, alcool, cigarette, stress, ...).



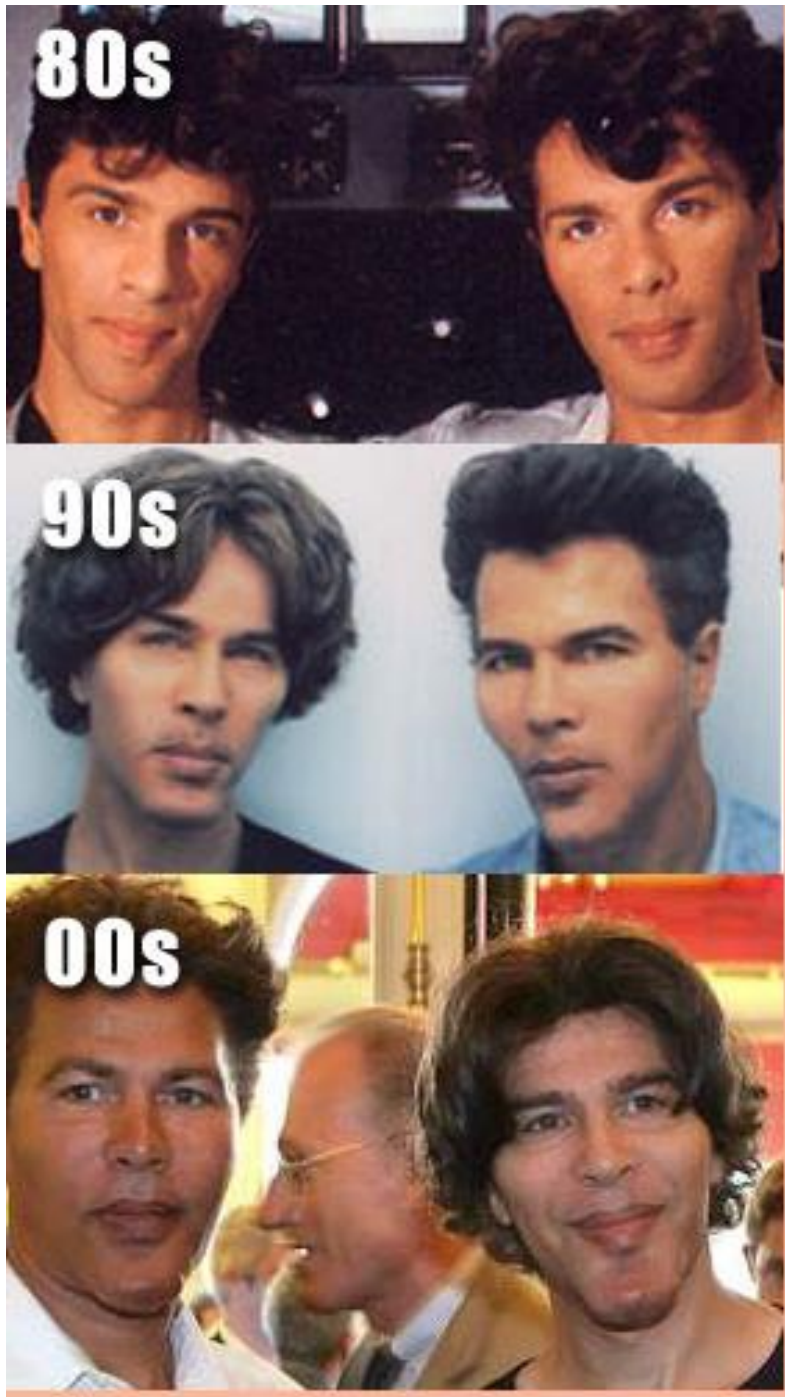
Fibroblastes



Hépatocytes

**Même génome mais cellules différentes**





Le profil d'expression génique des jumeaux homozygotes montrent des différences qui s'accroissent avec l'âge

Même génome mais différents...

Les modifications épigénétiques semblent spécifiques :

- “du stade du cycle cellulaire (interphase en particulier)
- “du stade de développement cellulaire (cellule somatique ou germinale)
- “du type de tissus
- “du stade de développement de l'individu (fœtus, embryon, adulte)

Les mécanismes épigénétiques contribuent donc à l'identité de chaque type de cellule (plus de 200 types de cellule chez l'homme).

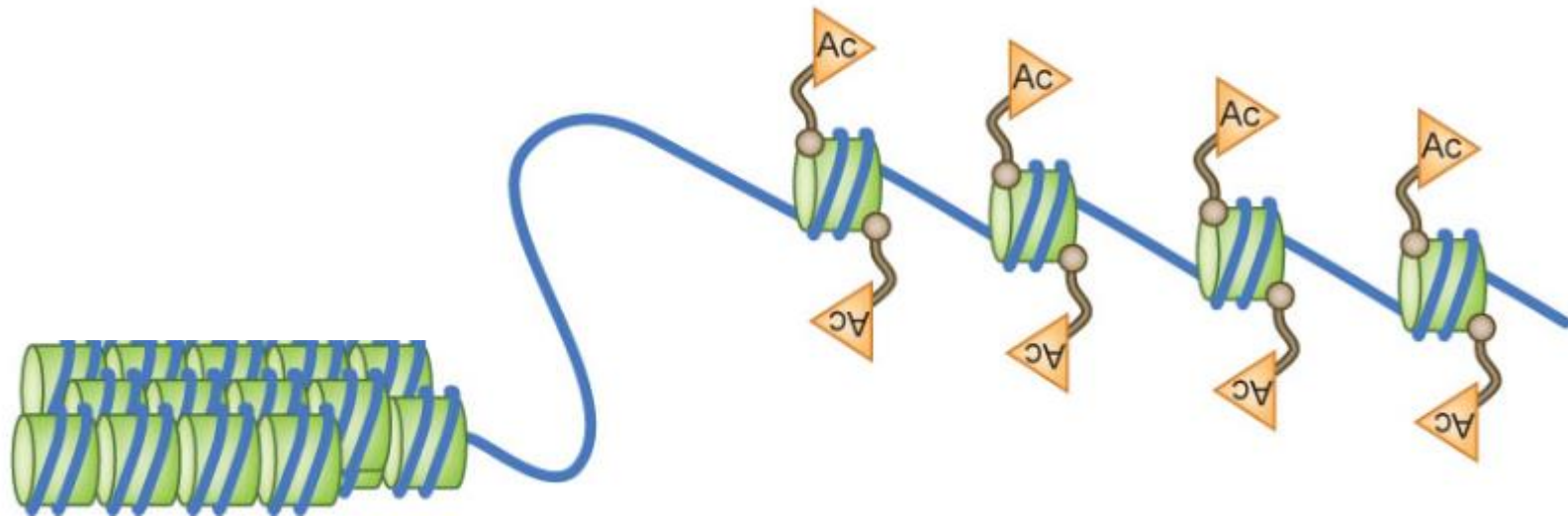


☐ Contrôle de la transcription chez les eucaryotes : Modifications post-traductionnelles des histones

Condensation / Décondensation de la chromatine :

Acétylation (K) : HAT / HDAC

Méthylations (K, R) : méthylase / histone déméthylase



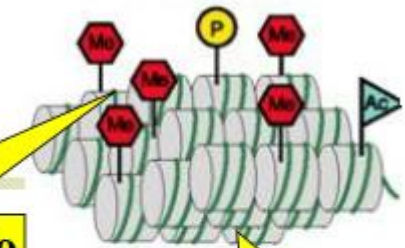
Recrutement d'HAT par des les domaines activateurs de FT :

les promoteurs des gènes sont accessibles et ils peuvent être transcrits

# Code des histones

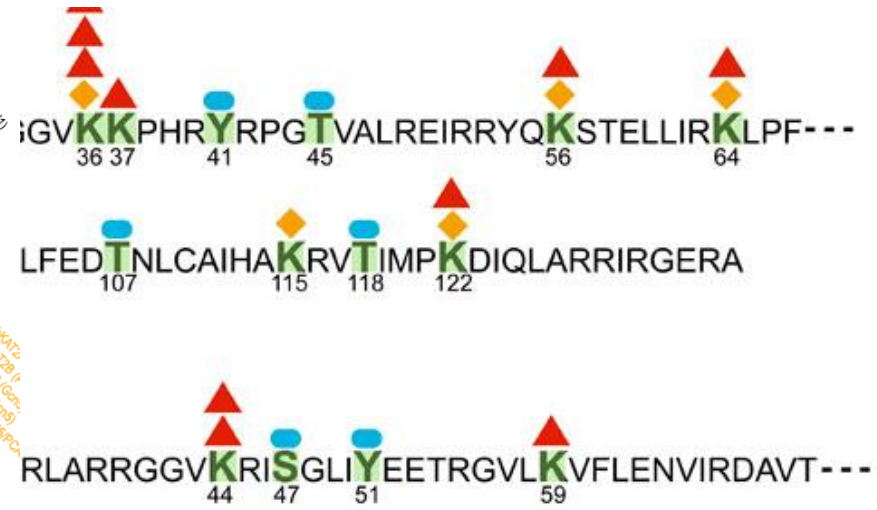
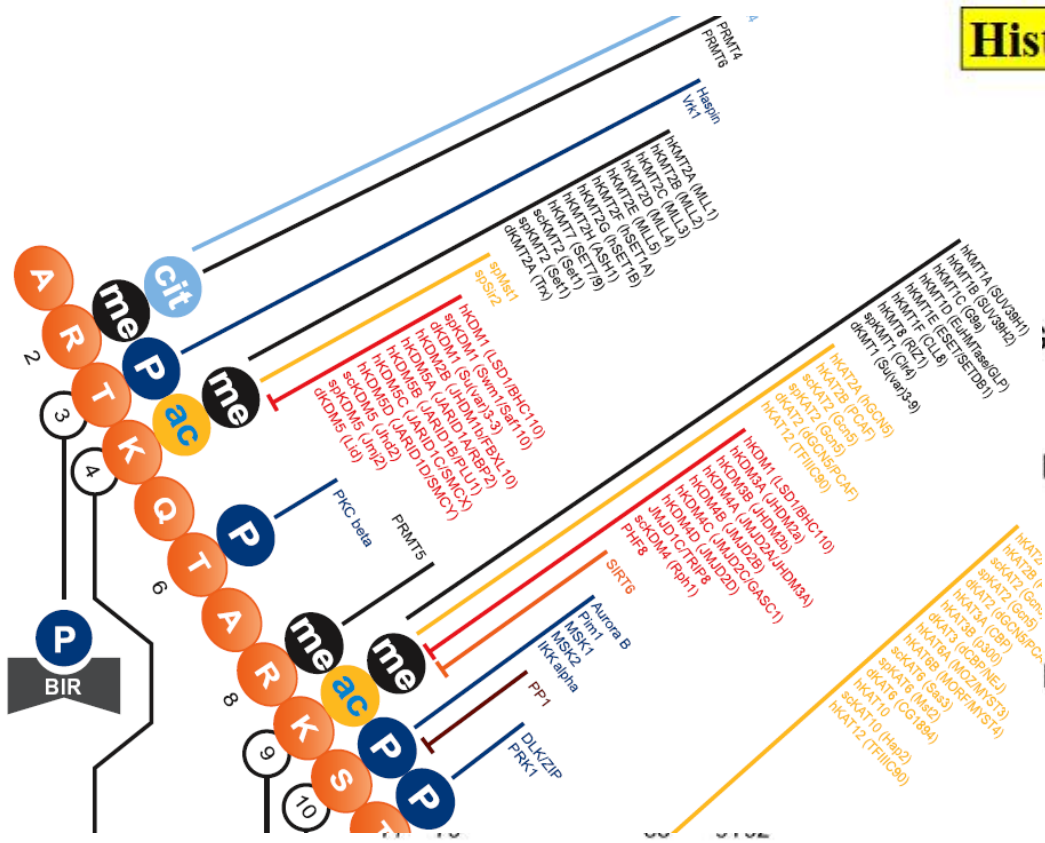
- ❖ les lysines peuvent être mono-, di - ou tri-méthylées
- ❖ les arginines peuvent être mono-méthylées ou di-méthylées de manière symétrique ou asymétrique
- ....

Les histones peuvent subir de nombreuses modifications, chaque modification pouvant influencer les autres



Histone encoded info

DNA encoded info



<span style="color: red;">▲</span> methylation	<span style="color: orange;">◆</span> acetylation
<span style="color: blue;">■</span> citrullination	<span style="color: cyan;">●</span> phosphorylation



❖ Différents niveaux de régulation qui sont interconnectés ... méthylation de l'ADN

▪ Méthylation des bases aux niveaux des îlots CpG :

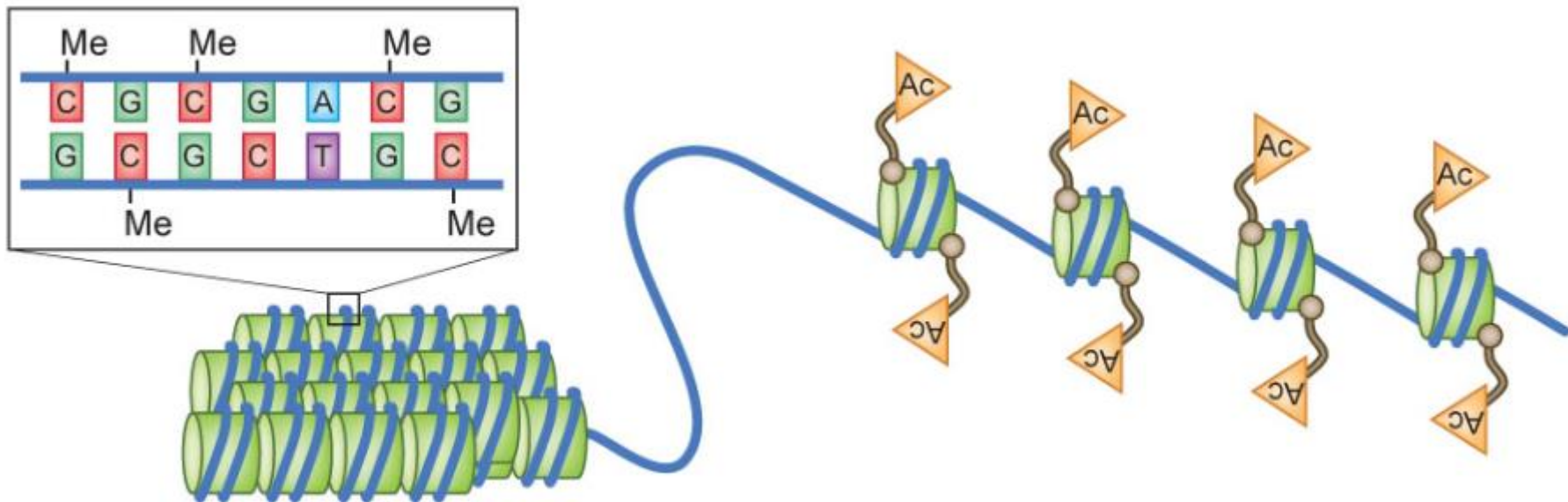
Cytosine => 5-méthylcytosine

méthylation => extinction du gène aval

▪ Recrutement de facteurs qui modifient les queues des histones (méthylation, acétylation ...)

=> condensation ou décondensation de la chromatine

=> activation ou non de la transcription



- ❖ Différents niveaux de régulation qui sont interconnectés ... rôle **des ARN non codants**

### Les petits ARN ...

Certains ARN pol sont spécialisés dans la méthylation de l'ADN dirigée par les ARN ("RNA-directed DNA methylation" - RdM).

#### Mécanisme RdM (résumé)

ARN spéciaux => ARN DS => siARN par enzyme DICER-like 3 => cytoplasme.

Protéine AGO4 => siRNA => complexe siRNA => noyau : interaction avec longs ARN non-codants ("long non-coding RNA" - lncRNA)



Action ADN méthyltransférases => méthylation de novo des cytosines  
=> Répression de la transcription

### ... mais aussi les grands ARN

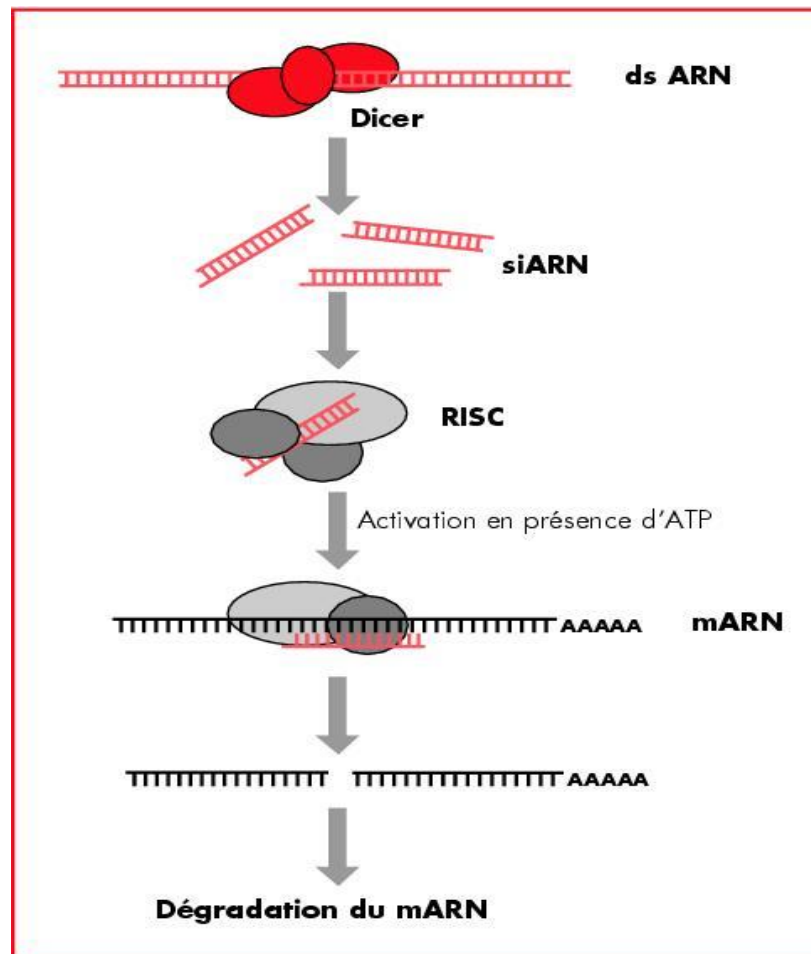
Un lncRNA peut servir d'échafaudage pour des enzymes de modification des histones afin de contrôler la transcription de gènes cibles.

## ❑ Contrôle post-transcriptionnel

Régulation par les ARN interférents (ARNsi ou ARNmi) :

= Mécanisme naturel de régulation (1990)

➡ **Outil pour comprendre la fonction des gènes (1998)**  
**Outils thérapeutique**

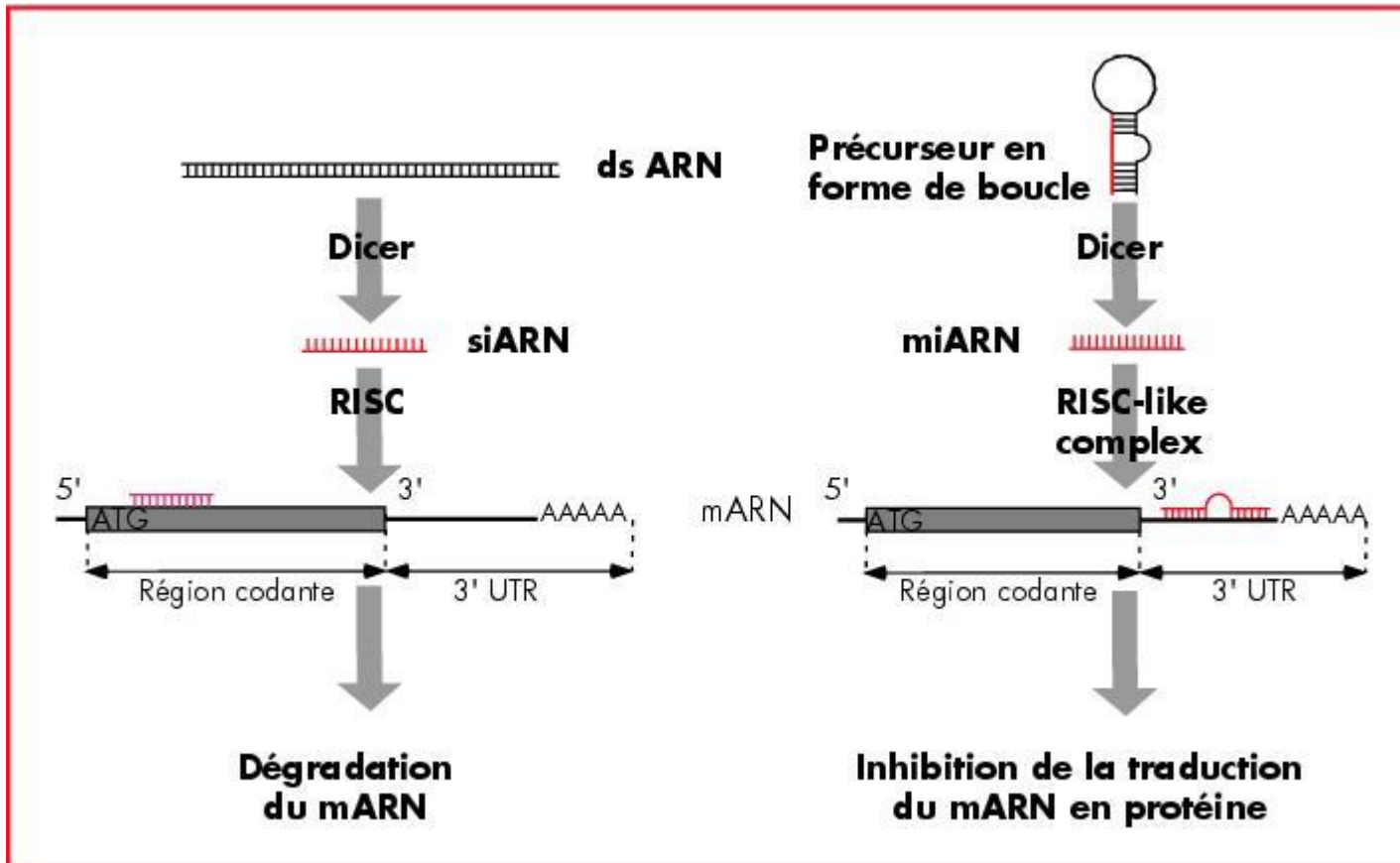


Clivage des ARN DS en ARNsi par l'enzyme DICER

Prise en charge du ARNsi par RISC

Reconnaissance de la séquence homologue sur l'ARNm

## Comparaison des mécanismes d'action des siARNs et des micro ARNs.





# **TECHNIQUES DE GENIE GENETIQUE**

# PCR et Q-PCR : Approfondissement

PCR

❖ Longueur des amorces = fonction de la « complexité » du génome

❖ Séquences des amorces (5' > 3')

Mésappariement possible, mais pas n'importe où

Eviter séquences complémentaires entre amorces et à l'intérieur de l'amorce (=> formation de structures secondaires et de dimères d'amorces)

❖ Température d'hybridation

Température d'hybridation des amorces dépend de la température de fusion des amorces

*Rappel* : La température de fusion  $T_M$  (melting temperature) de l'ADN DS est la température pour laquelle 50 % des molécules d'ADN sont dénaturées sous forme simple brin ADN.

Calcul du  $T_M$  :  $T_M = 4(G+C) + 2(A+T)$

Température d'hybridation =  $\min(T_{M1}, T_{M2}) - 4^\circ\text{C}$

$T_M$  diminue de 1 à 1,5°C par % de différence entre amorce et ADN cible.

(remarque : 1/20 ----> 5%)

$T_M$  proches entre amorces (<5°C)

Exemple : Amplification d'une séquence microsatellite

```
1  TTTTAAGTTA  CTGTGTGCTT  GTTGCAGGAT  CTGTAACTAA  TTCCTATGCG  ATTCTCTTGT
61  TTGTAGGGCG  AAGATGAGGG  AGATCCTGCA  CATCCAGGGA  GGGCAATGTG  GCAACCAGAT
121 TGGCGCCAAG  TTCTGGGAGG  TGGTGTGCGA  TGAACATGGC  ATTGACCACA  CACACACACA
181 CACACACACA  CACACACACA  CACACACACA  CACACACACA  CACACACACA  CACACACACA
241 CACACACACA  CACACACACA  CACACACACA  CACACACACA  CACACACACA  CACACACACA
301 CACACACACA  CACACACACA  CACACACACA  CACACACACA  CACACAACCG  CCTCGGCCAT
361 TGCTGGTAAC  AATTGGGCTA  AGGGCCACTA  CACCGAGGGT  GCTGAGCTCA  TTGACTCTGT
421 TCTGGATGTT  GTGAGGAAGG  AAGCTGAGAA  CTGTGACTGC  TTGCAAGGAT  TCCAAGTATG
481 CCACTCCCTT  GGTGGTGGTA  CTGGATCTGG  TATGGGTACG  CTGTTGATCT  CAAAGATCAG
```

Figure 1 : séquence nucléotidique d'une portion d'ADN génomique

- ☐ [MgCl<sub>2</sub>]

Si [Mg<sup>2+</sup>] trop faible, pas de produit PCR

Si [Mg<sup>2+</sup>] trop élevé, apparition de produits PCR non désirés

- ☐ [dNTP]

[dNTP] élevée : augmentation du rendement de PCR mais risque d'apparition de produits PCR non désirés

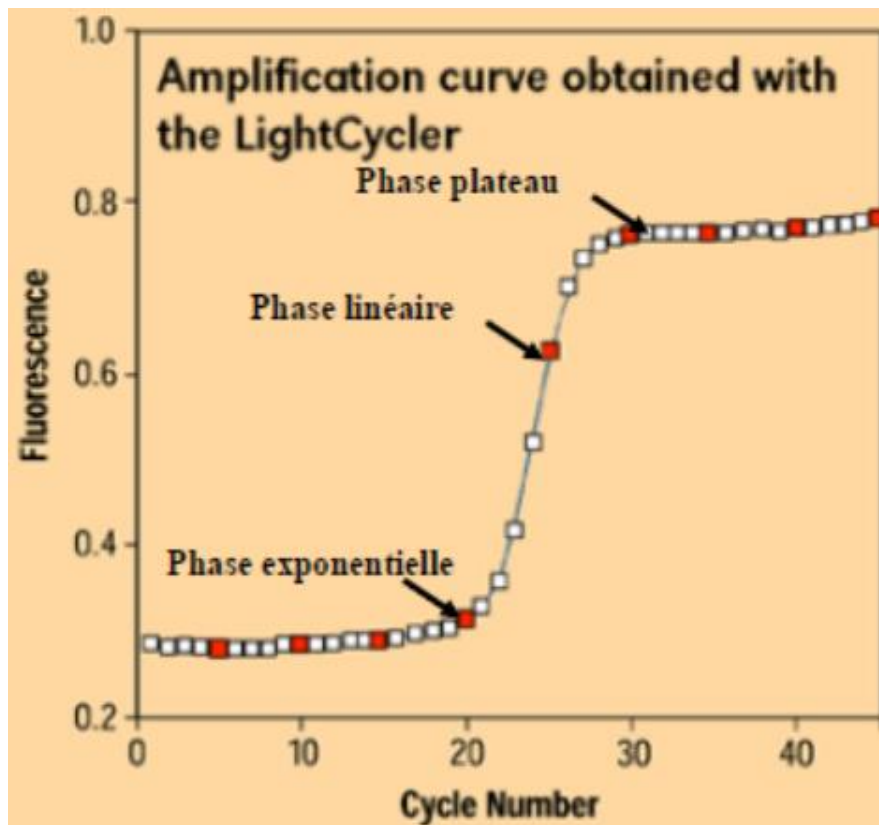
- ☐ Durée de la phase d'élongation : Vitesse d'élongation: -->1 Kb/min

qPCR

## Q-PCR

### Introduction :

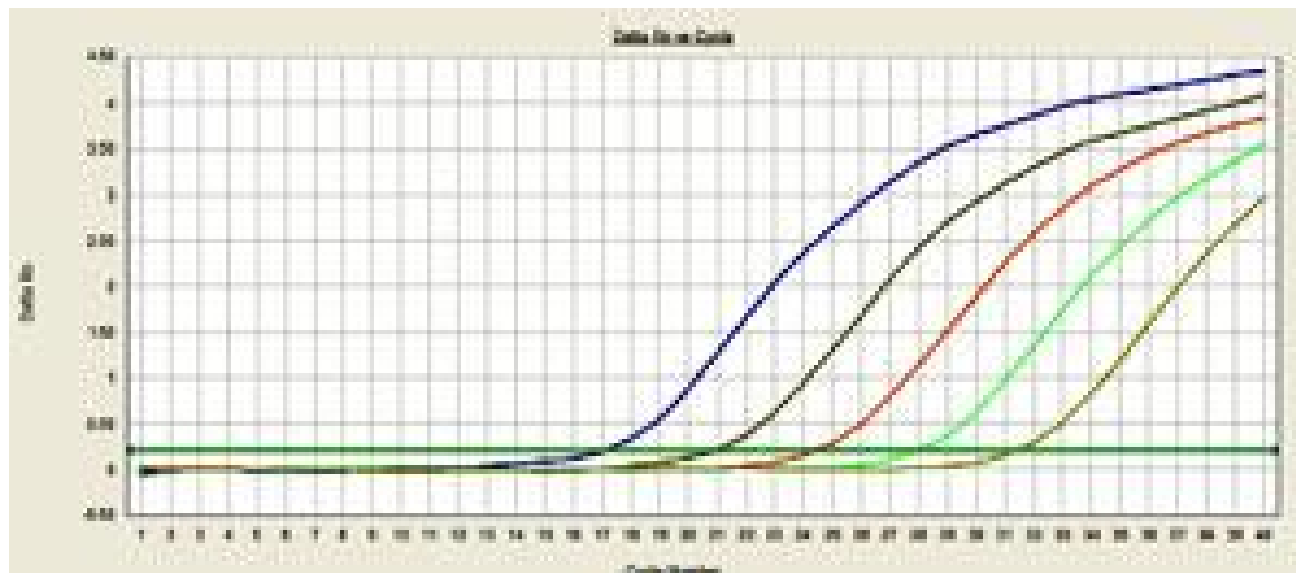
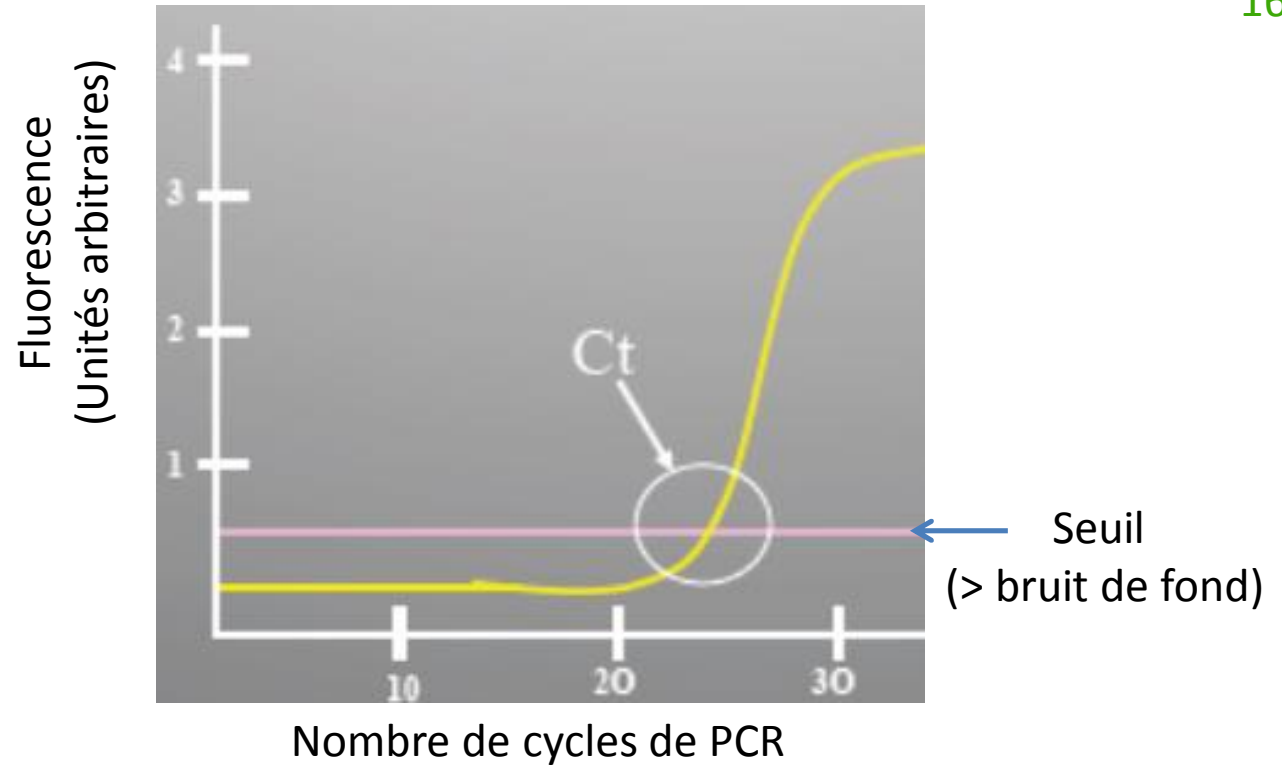
La **PCR quantitative en temps réel** est basée sur la détection et la quantification d'un **reporter fluorescent** dont l'émission est directement proportionnelle à la **quantité d'amplicons** générés pendant la PCR, donc au final permet de quantifier l'ADN de départ.



Cette quantification ne peut se faire avec précision que lors de la **phase exponentielle** de l'amplification. (C'est la phase la plus reproductible de la réaction de PCR).



# Cycle threshold (Ct) [ou Cycle seuil (Cs)]



**Q-PCR**

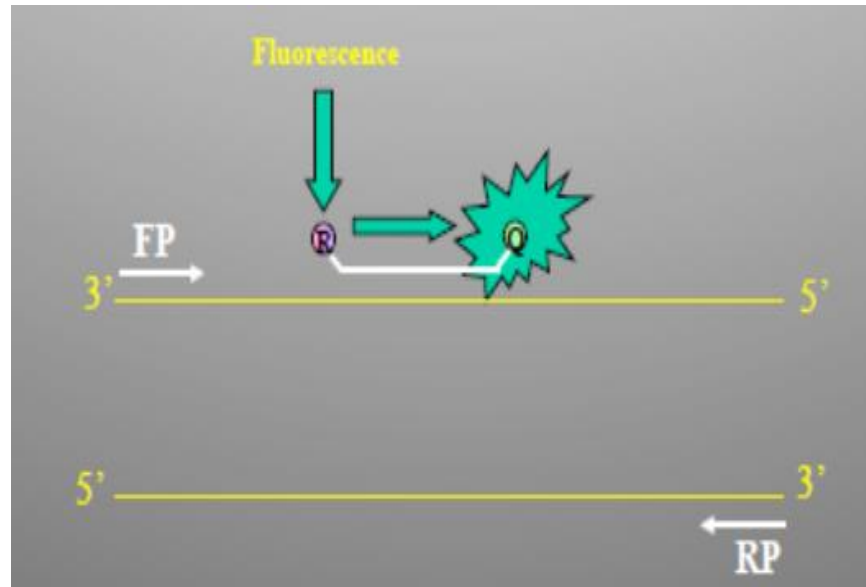
**Fluorescence ?**

# Q-PCR

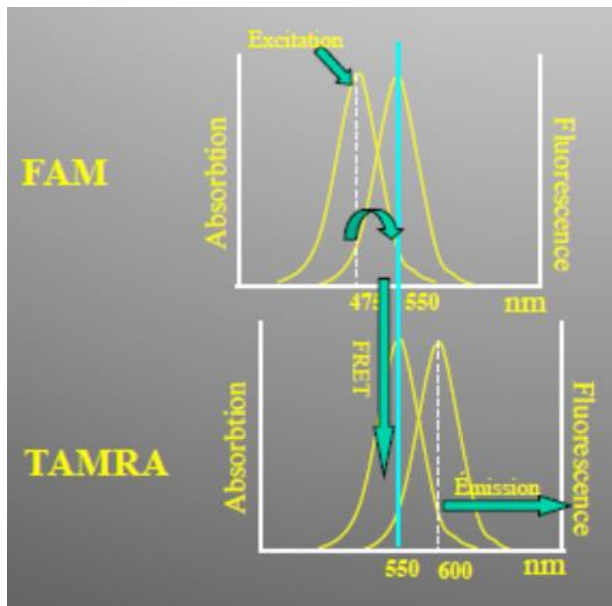
FRET Fluorescence resonance energy transfert

depuis le reporter (haute energie) vers le quencher (faible energie)

Chimie Taq man



Pas de signal fluorescent émis par le reporter quand la sonde est intacte

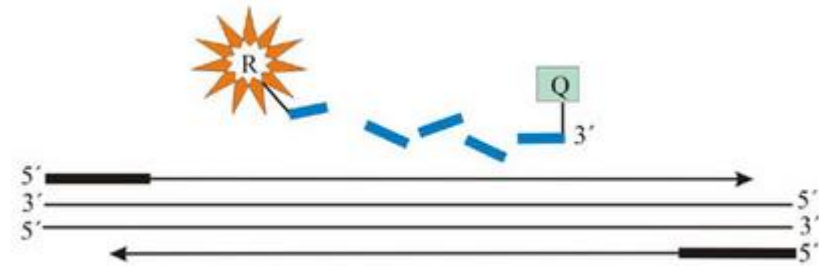
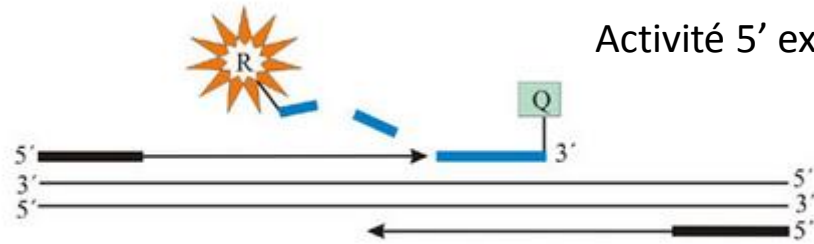
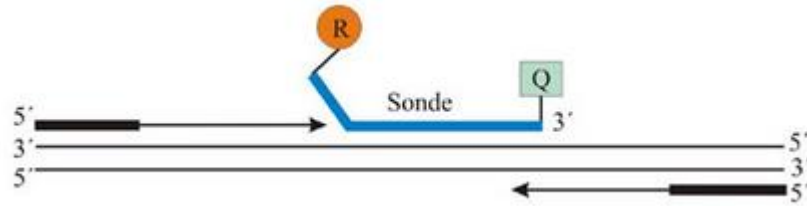
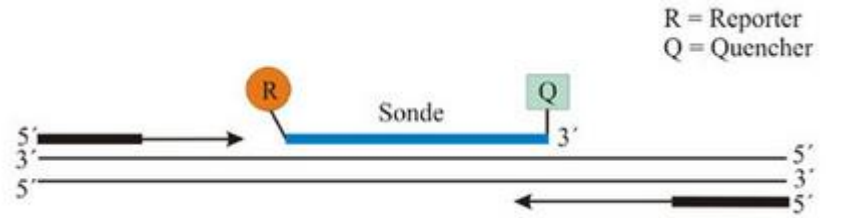


## Mécanisme de quenching

Le spectre d'émission du reporter recouvre le spectre d'absorption du quencher

FAM : 6 carboxy fluoresceine - TAMRA : 6 carboxyl tetraméthyl rhodamine

# Q-PCR Chimie Taq man



Dye	Absorbance Wavelength	Emission Wavelength	Visible color
Hydroxycoumarin	325	386	blue
methoxycoumarin	360	410	blue
Alexa fluor	345	442	blue
aminocoumarin	350	445	blue
Cy2	490	510	green (dark)
FAM	495	516	green (dark)
Alexa fluor 488	494	517	green (light)
Fluorescein FITC	495	518	green (light)
Alexa fluor 430	430	545	green (light)
Alexa fluor 532	530	555	green (light)
HEX	535	556	green (light)
Cy3	550	570	yellow
TRITC	547	572	yellow
Alexa fluor 546	556	573	yellow
Alexa fluor 555	556	573	yellow
R-phycoerythrin (PE)	480;565	578	yellow
Rhodamine Red-X	560	580	orange
Tamara	565	580	red
Cy3.5 581	581	596	red
Rox	575	602	red
Alexa fluor 568	578	603	red
Red 613	480;565	613	red
Texas Red	615	615	red
Alexa fluor 594	590	617	red
Alexa fluor 633	621	639	red
Allophycocyanin	650	660	red
Alexa fluor 633	650	668	red
Cy5	650	670	red
Alexa fluor 660	663	690	red
Cy5.5	675	694	red
TruRed	490;675	695	red
Alexa fluor 680	679	702	red
Cy7	743	770	red

PCR simplex : en général FAM et TAMRA

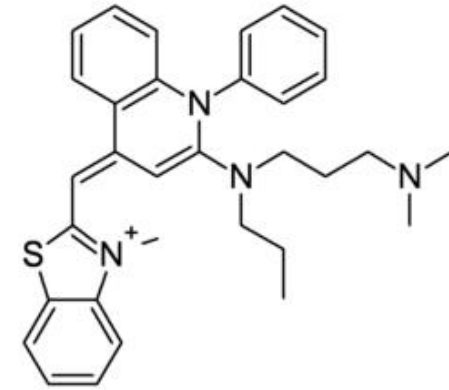
PCR multiplex : attention au choix des fluorophores !

## Q-PCR Chimie Sybrgreen

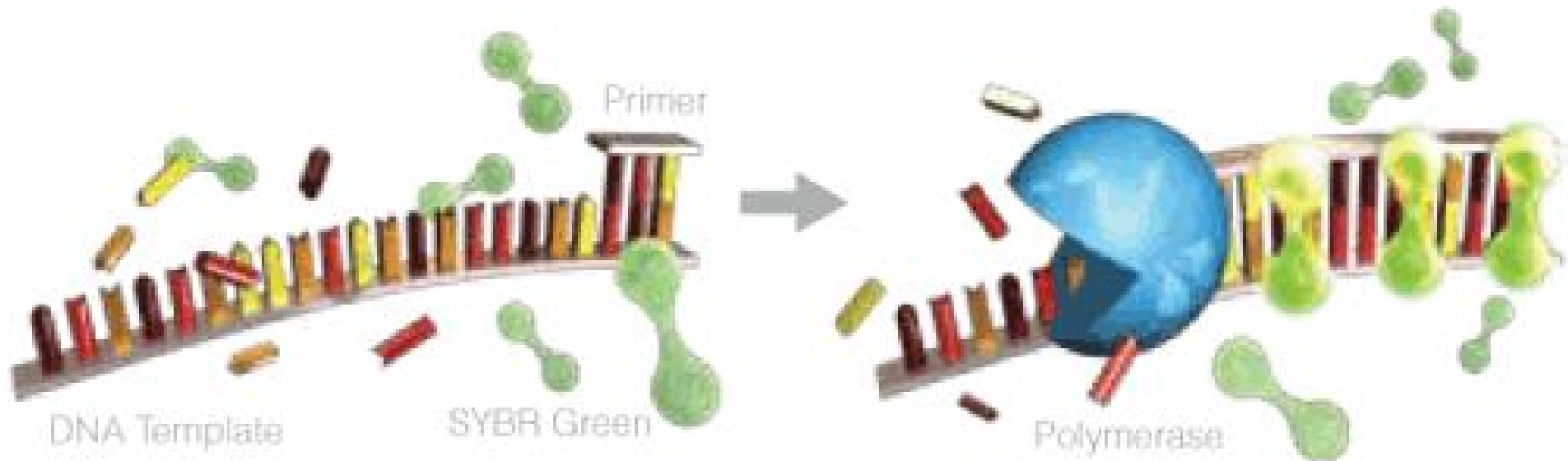
Cyanine asymétrique

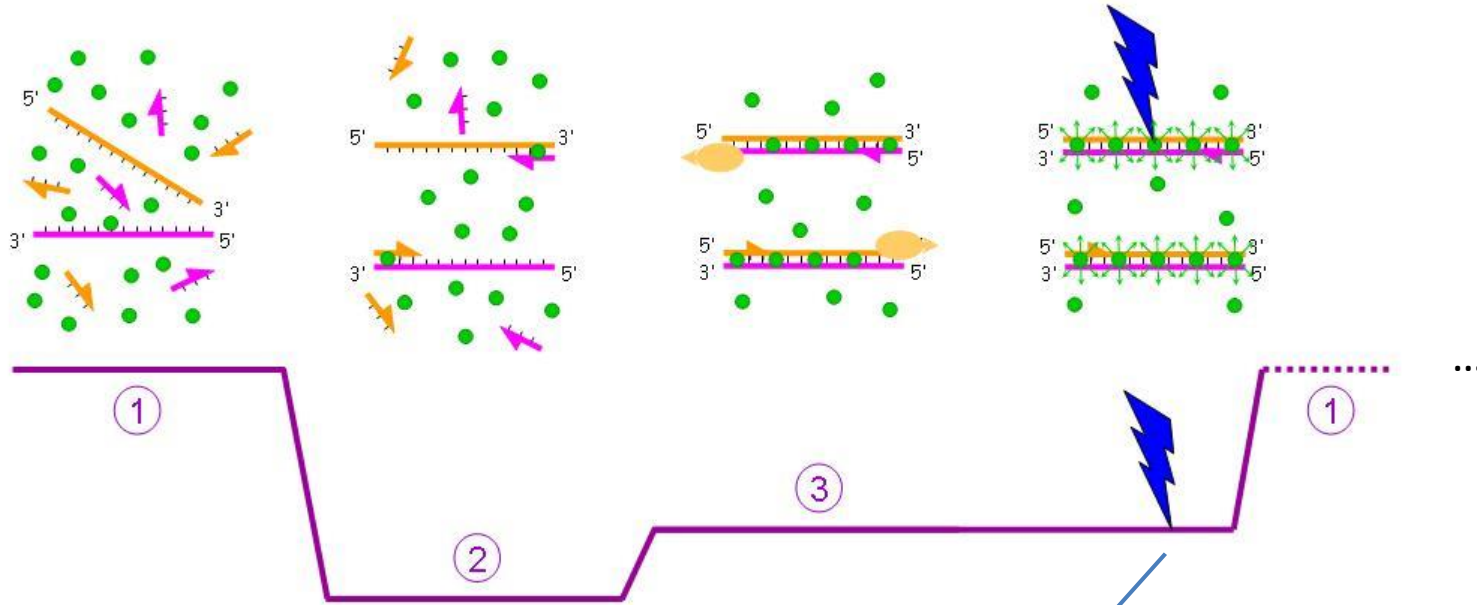
Agent intercalent de l'ADN

Molécule alternative au BET (coloration de gels)

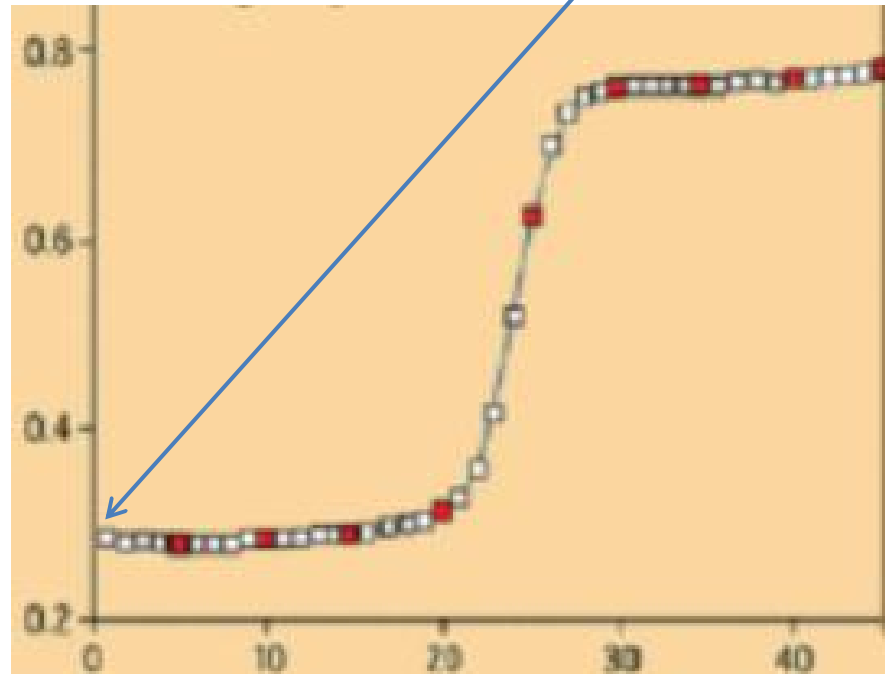


Longueur d'onde absorption 485 nm  
Longueur d'onde émission 520 nm





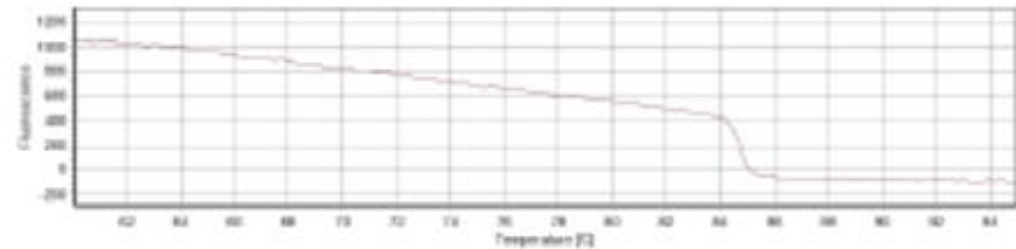
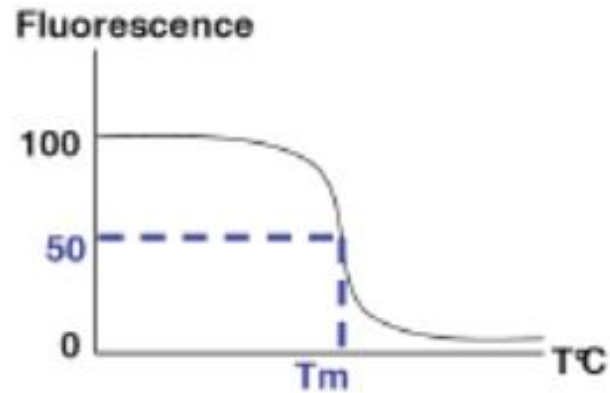
Fluorescence (UA)



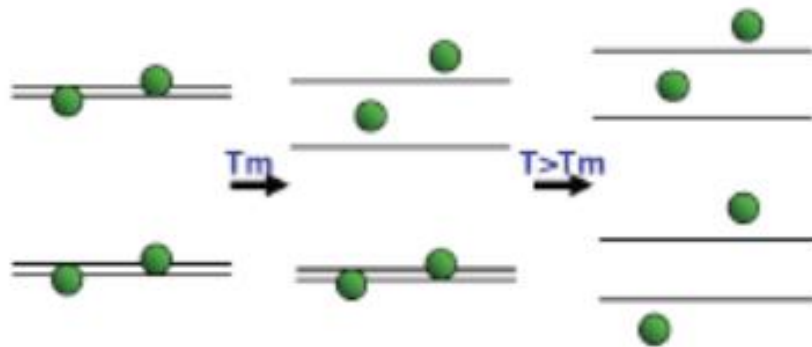
Nombre de cycles

## Courbe de fusion (Melting curve) Estimation de la spécificité des amplicons

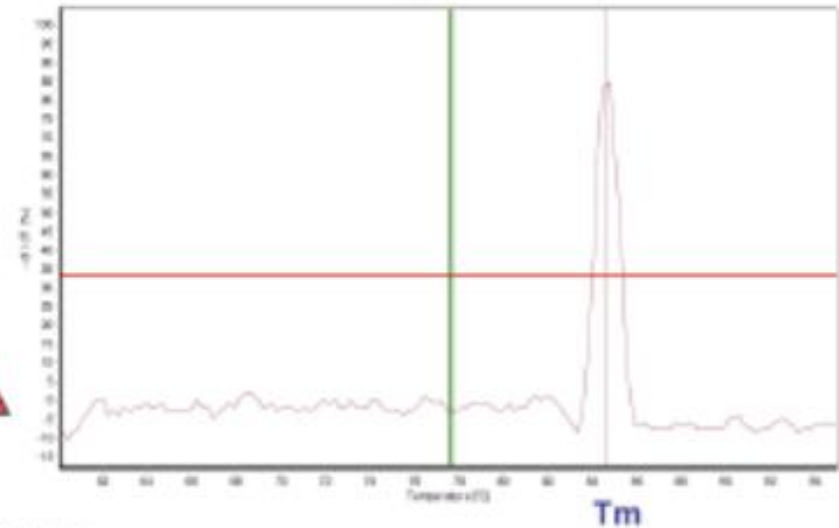
En q PCR la seule variable observable est la fluorescence totale de l'échantillon.  
A priori, on ne peut pas discriminer les amplicons aspécifiques des amplicons spécifiques



↓  
Fonction dérivée



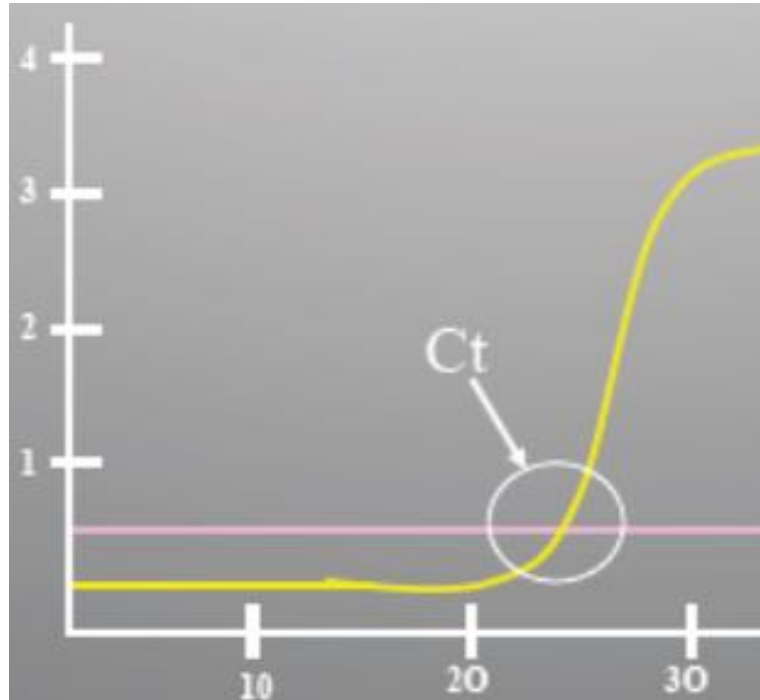
En **Sybr green** on peut suivre la décroissance de la fluorescence à mesure que l'on atteint puis dépasse le  $T_m$





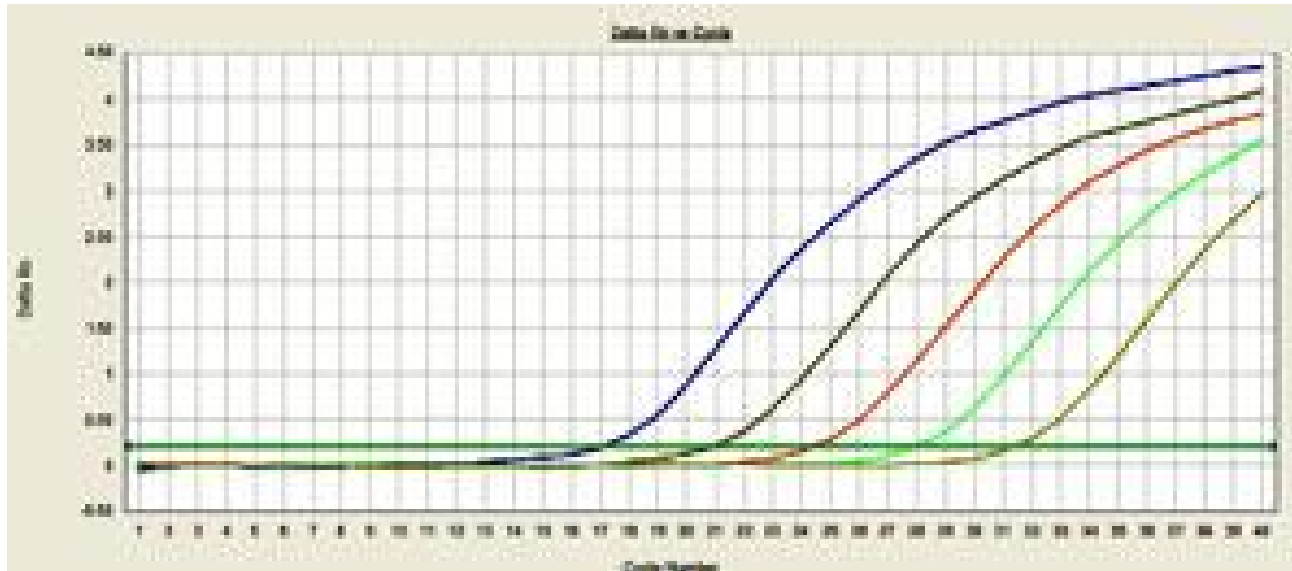
# Cycle threshold (Ct) [ou Cycle seuil (Cs) ou Cq]

Fluorescence  
(Unités arbitraires)



Seuil  
(> bruit de fond)

Nombre de cycles de PCR



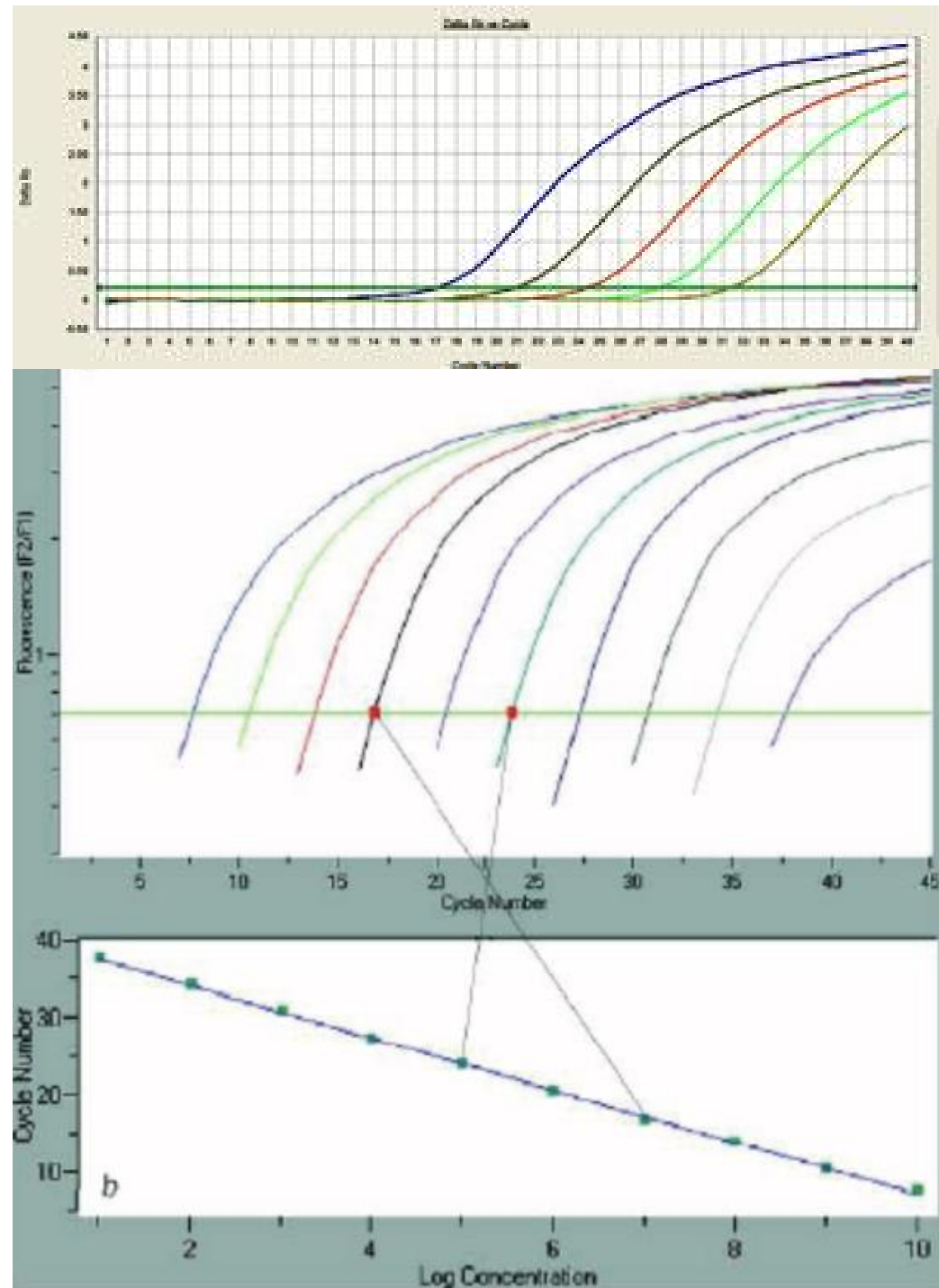
## Droite standard

La droite standard représente les valeurs  $C_t$  obtenues expérimentalement en fonction du log des concentrations en molécules issues de série de dilutions d'un échantillon standard.

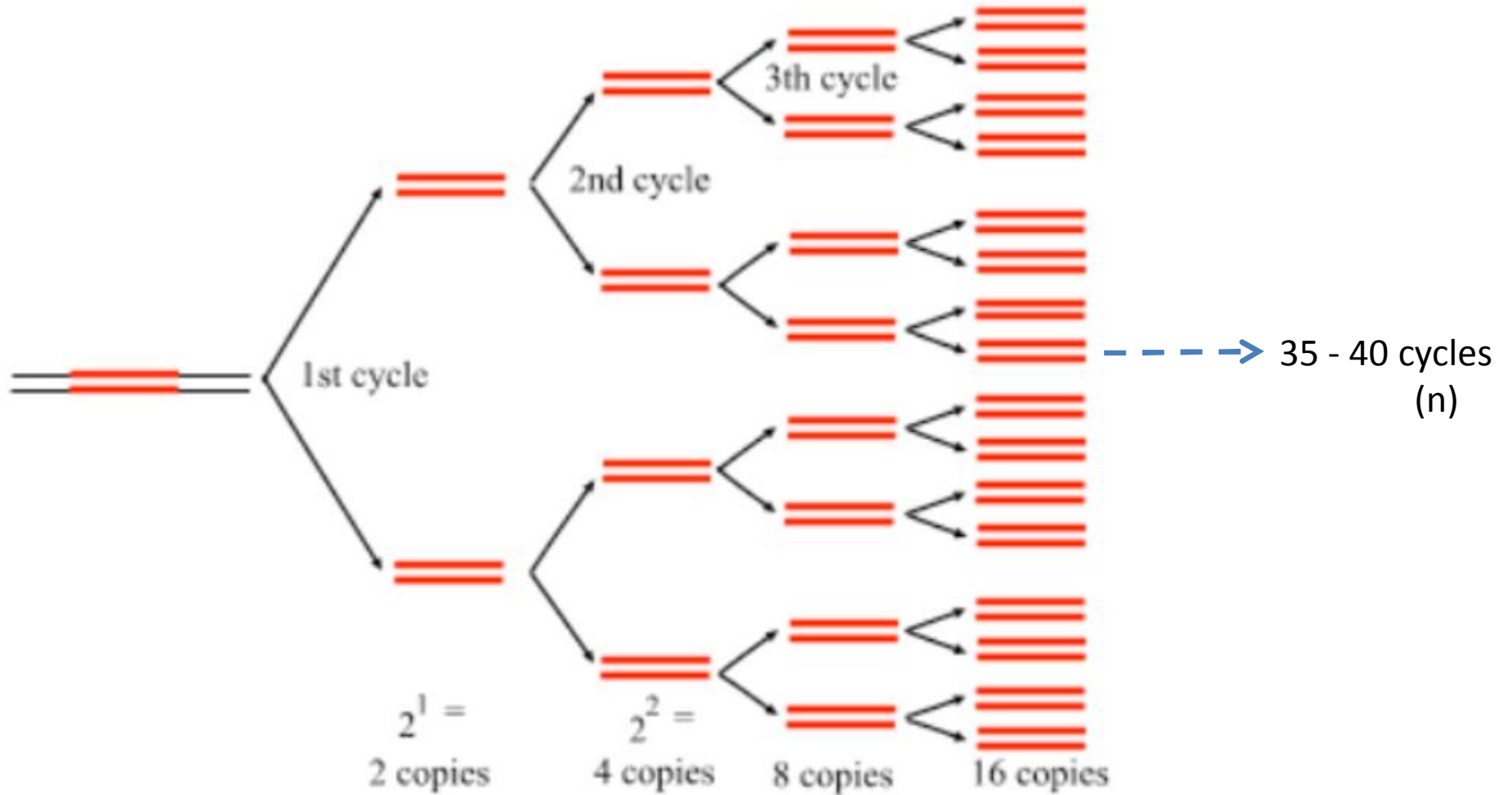
La valeur de  $C_t$  est inversement proportionnelle au log de la concentration initiale en molécules cibles.

Le  $C_t$  obtenu à partir d'un échantillon de concentration inconnue va être ainsi traduit en concentration en molécules cibles grâce à la droite standard.

Y : échelle logarithmique <= échelle linéaire



Principe de la PCR



$N_c = 1 \times 2^n$



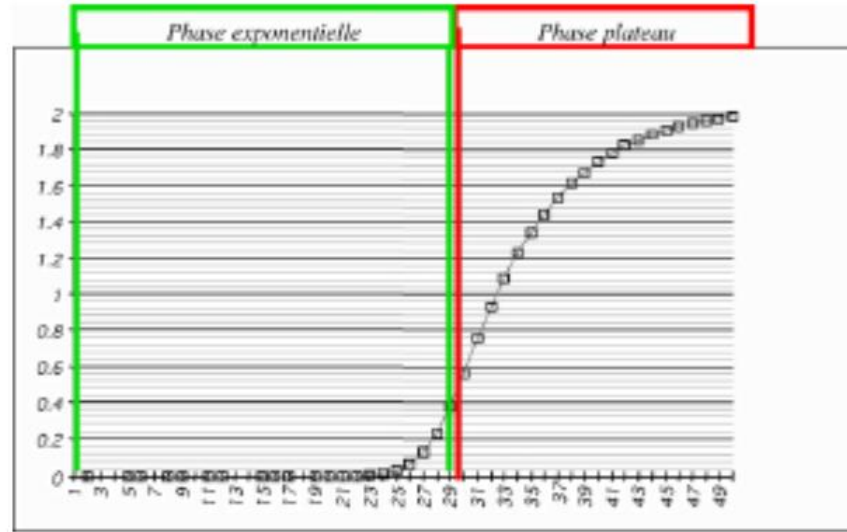
$1 = N_0$



$N_c = N_0 \times 2^n$

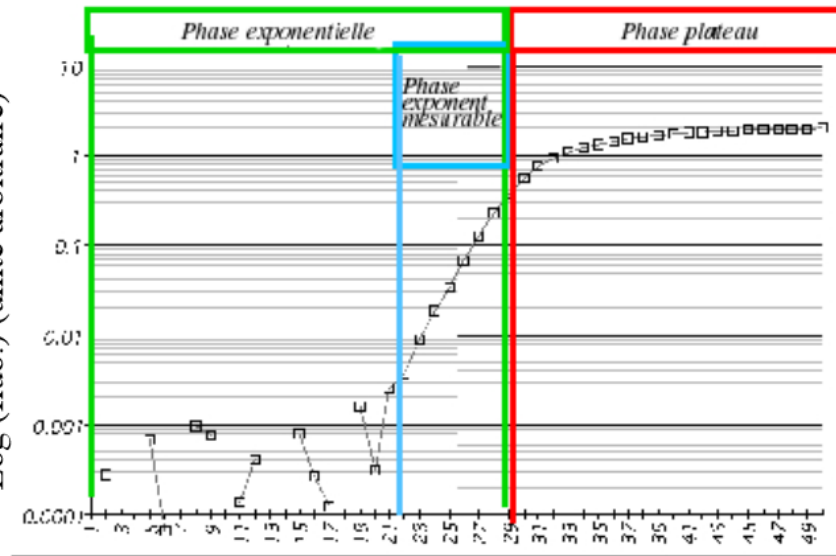
*échelle linéaire*

Fluorescence (unité arbitraire)

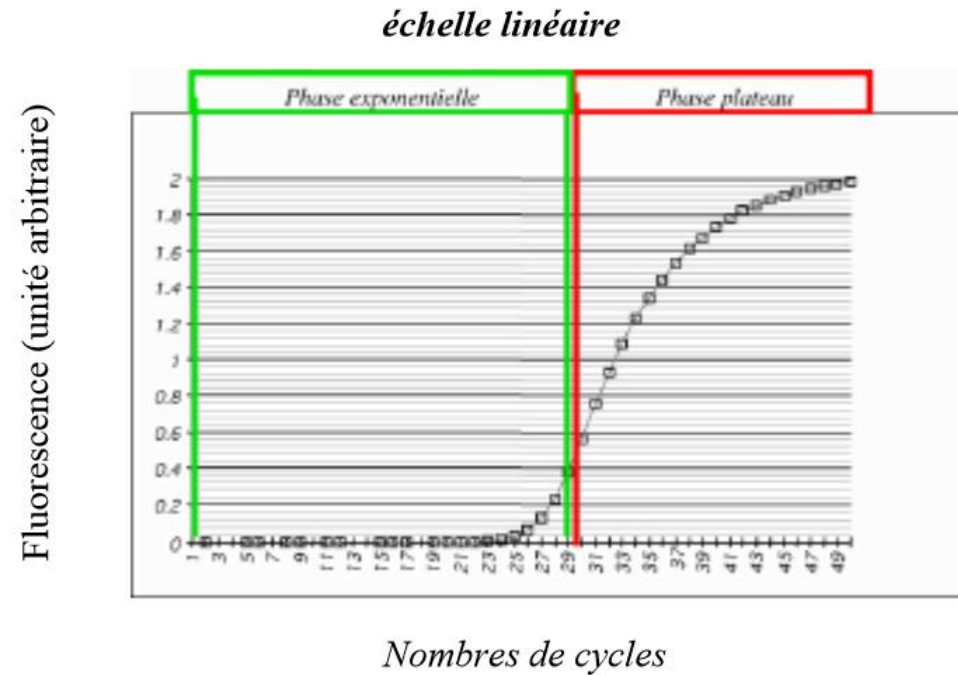


*échelle logarithmique*

Log (flu.) (unité arbitraire)



Nombres de cycles



☐ La phase exponentielle est modélisable.

$$N_1 = N_0 \cdot (1 + E)$$

$$N_2 = N_1 \cdot (1 + E)$$

...

$$N_c = N_{(c-1)} \cdot (1 + E)$$

---


$$N_c = N_0 \cdot (1 + E)^c$$

- $N_0$  : le nombre de copies initiales,
- $N_c$  : le nombre de copie au cycle c
- E : l'efficacité de la PCR

Remarque : Si E = 100%,  $N_c = N_0 \cdot 2^c$

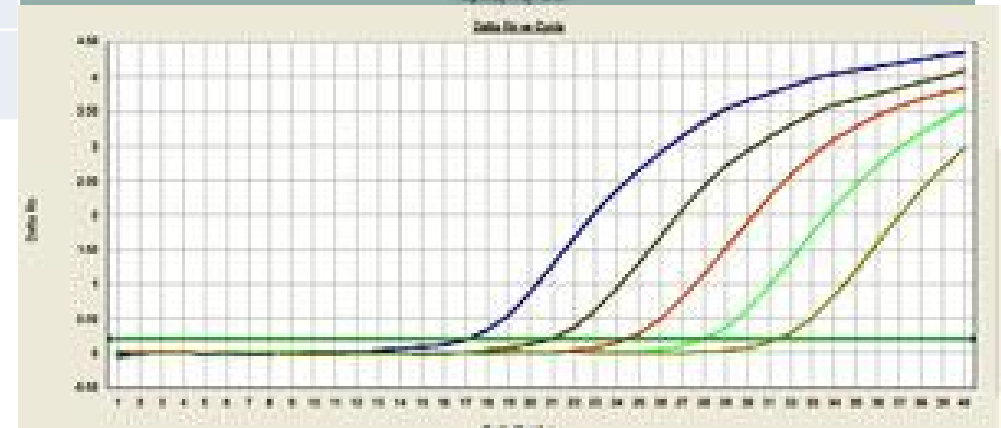
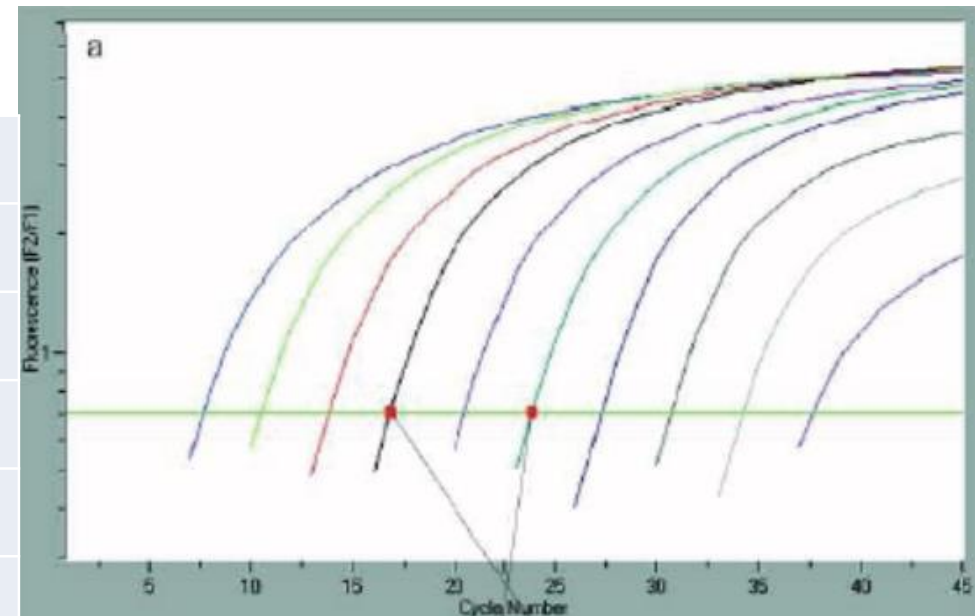
❑ Calcul de l'efficacité d'une PCR

A partir des données suivantes, calculer l'efficacité de la PCR.

Images d'illustration

Données

Concentration en ADN (ng/μL)	Nb de Cycles
1,0E+08	5
1,0E+07	9
1,0E+06	13
1,0E+05	15
1,0E+04	19
1,0E+03	23



## Bases de calculs

$N_{Ct}$  : le nombre de produits PCR présents après « Ct » cycles

$N_0$  : le nombre de molécules cibles présentes à l'origine (zéro cycle)

E : Efficacité de la PCR

Ct : Cycle seuil

- Equation décrivant la PCR

$$N_{Ct} = N_0 (1+E)^{Ct}$$

- Pour simplifier :  $Ef = 1 + E$

$$N_{Ct} = N_0 (Ef)^{Ct}$$

- Linéarisation

$$\log N_{Ct} = \log N_0 + Ct \log Ef$$

- Réarrangement de l'équation

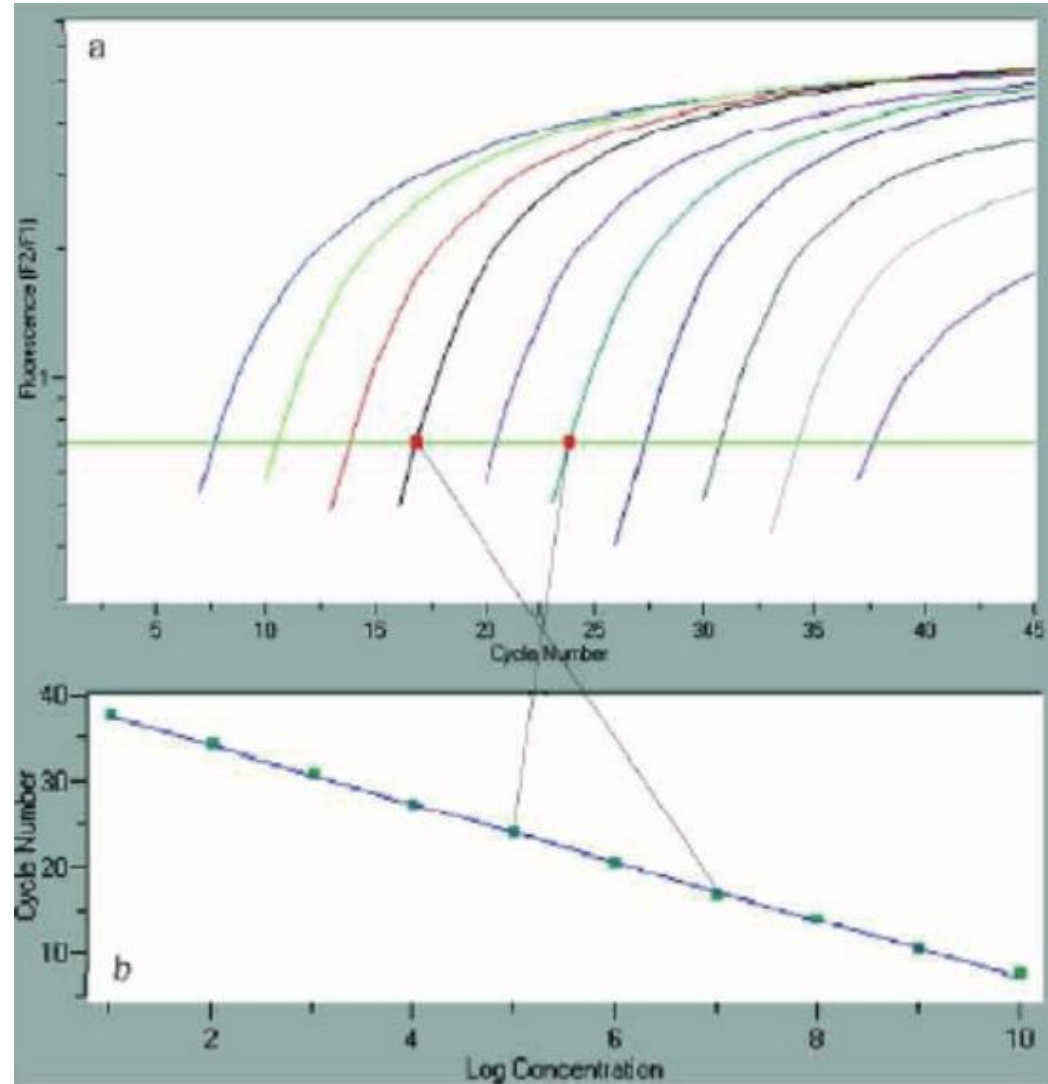
$$Ct = -\left(\frac{1}{\log Ef}\right) \log N_0 + \left(\frac{\log N_{Ct}}{\log Ef}\right)$$

↑  
y

↑  
a

↑  
x

↑  
b




La conversion « pente de courbe standard / efficacité PCR » est donnée par la relation :

$$\text{pente} = - (1/\log \text{Ef})$$

Ainsi :

- une pente de -3,32 représente un coefficient Ef de 2 (= 1 + 1) soit une efficacité E de 100%
- une pente de -3,4 représente un coefficient Ef de 1,97 (= 1 + 0,97) soit une efficacité E de 97%
- une pente de -3,5 représente un coefficient Ef de 1,93 (= 1 + 0,93) soit une efficacité E de 93%
- une pente de -3,6 représente un coefficient Ef de 1,9 (= 1 + 0,9) soit une efficacité E de 90%
- une pente de -3,7 représente un coefficient Ef de 1,86 (= 1 + 0,86) soit une efficacité E de 86%
- une pente de -3,92 représente un coefficient Ef de 1,8 (= 1 + 0,8) soit une efficacité E de 80%

 Pente connue, efficacité inconnue

$$\text{Ef} = 10^{-1/\text{pente}}$$

*Exemple :* pente = -3,7  $\Rightarrow$  Ef =  $10^{-1/-3,7} = 1,86 \Rightarrow E = 86\%$



❑ Calcul de l'efficacité d'une PCR

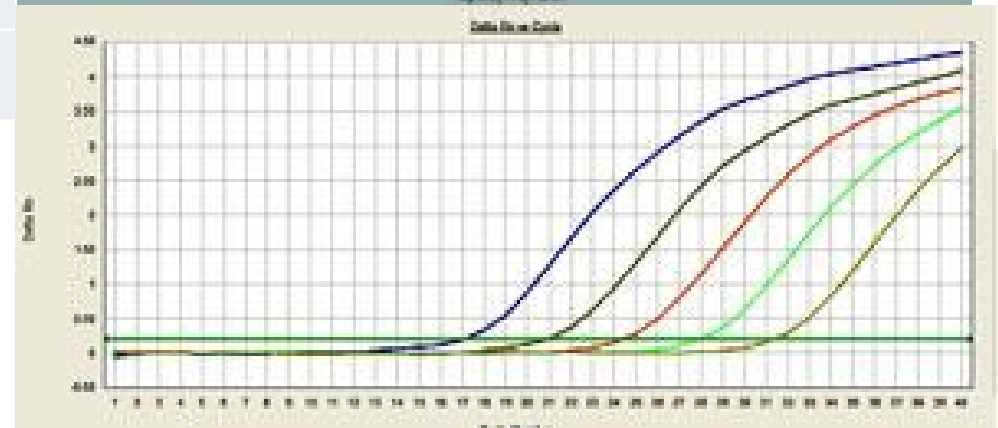
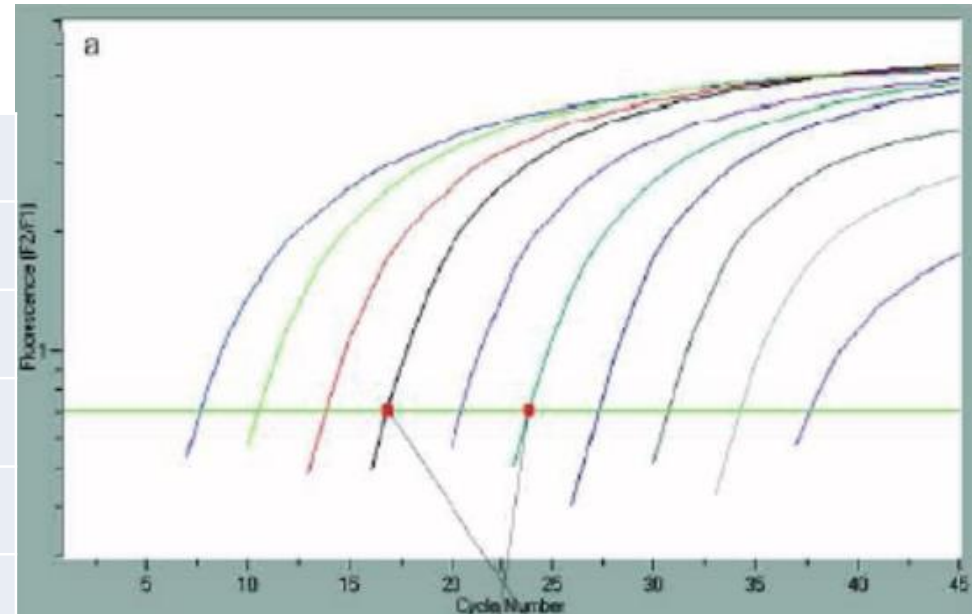
A partir des données suivantes, calculer l'efficacité de la PCR.

Données

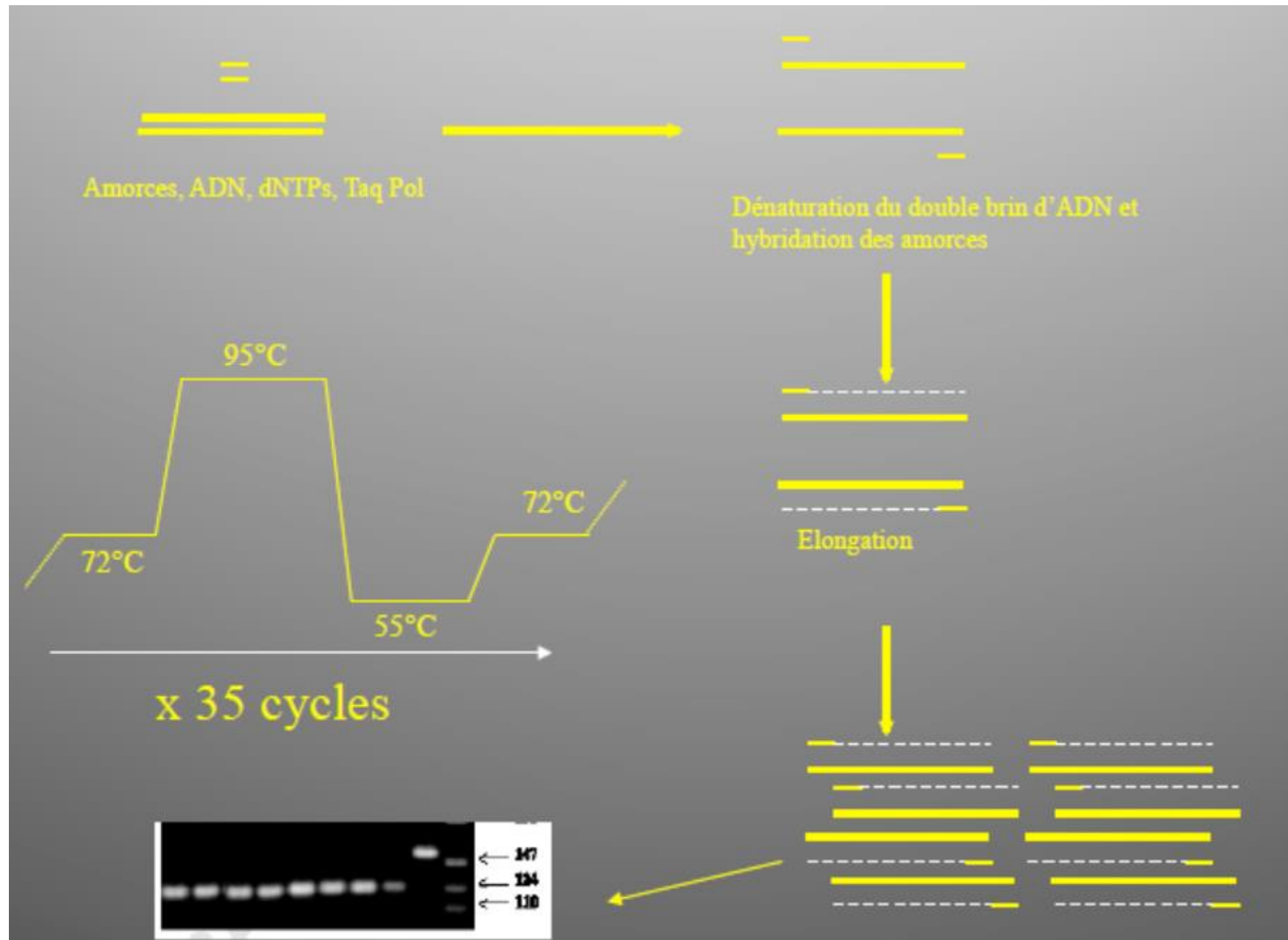
Concentration en ADN (ng/μL)	Nb de Cycles
1,0E+08	5
1,0E+07	9
1,0E+06	13
1,0E+05	15
1,0E+04	19
1,0E+03	23

E ?

Images d'illustration

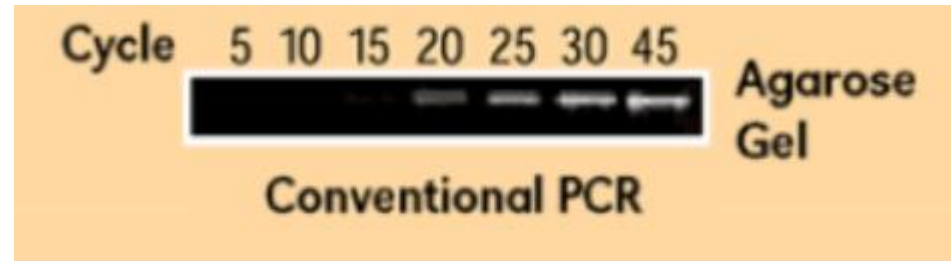


## PCR conventionnelle vs PCR en temps réel



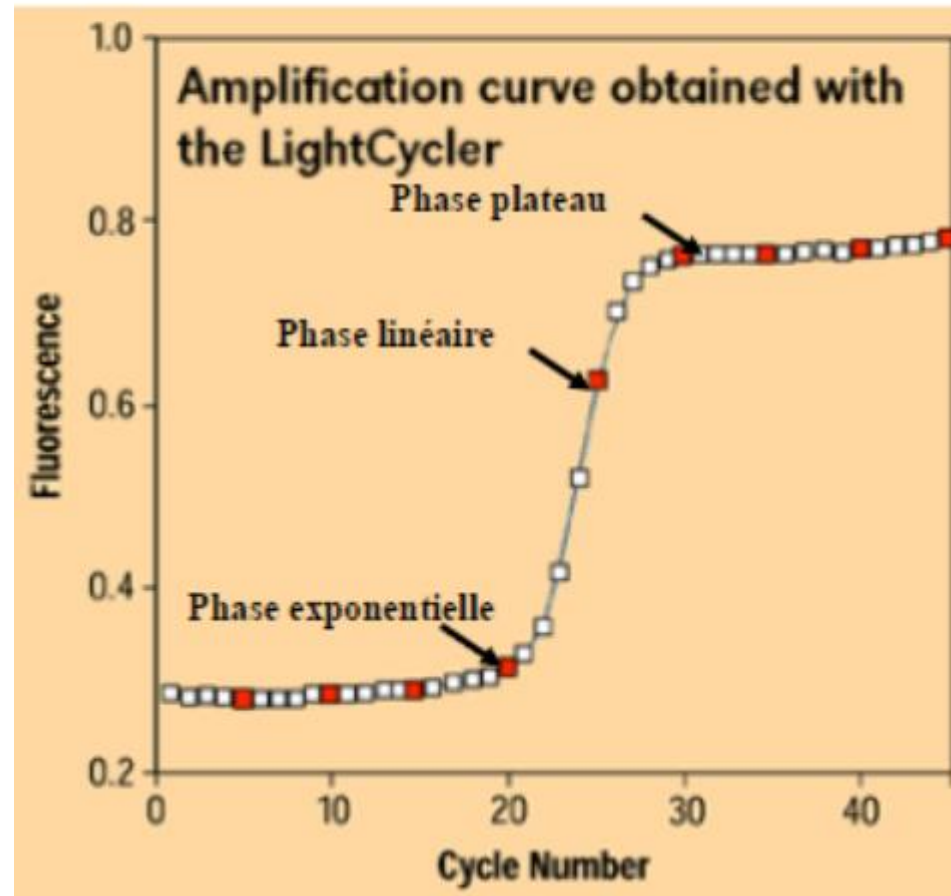
## PCR conventionnelle vs PCR en temps réel

Analyse en point final












Analyse dynamique

« en temps réel »



## □ Les autres type de PCR

- PCR emboîtée ou PCR gigogne (Nested PCR)
-  PCR en gradient de température
-  PCR multiplexe
-  PCR par essais (Touchdown PCR)
-  PCR sur colonie (Colony PCR)
-  RT-PCR (RT-PCR pour Reverse Transcriptase PCR)
-  RT-PCR en une étape (single step RT-PCR)
-  RT-PCR *in situ* (*in situ* RT-PCR)
-  RT-PCR sur une cellule (single-cell RT-PCR)
-  PCR quantitative (q-PCR pour quantitative PCR)

# **Amplifications isothermes de l'ADN**

## ❑ Amplifications isothermes de l'ADN (spécificité de séquence)

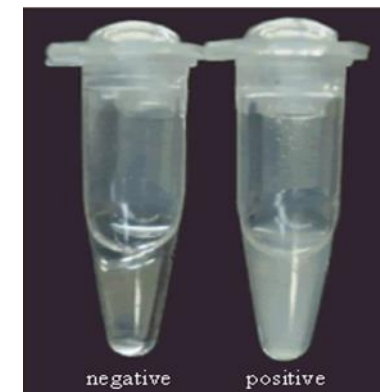
Pas besoin de thermocycleur.

Méthode plus rapide.



**LAMP** (Loop-mediated isothermal amplification : 4 à 6 amorces qui reconnaissent 6 à 8 régions distinctes.

Exemple : détection de *Listeria monocytogenes*

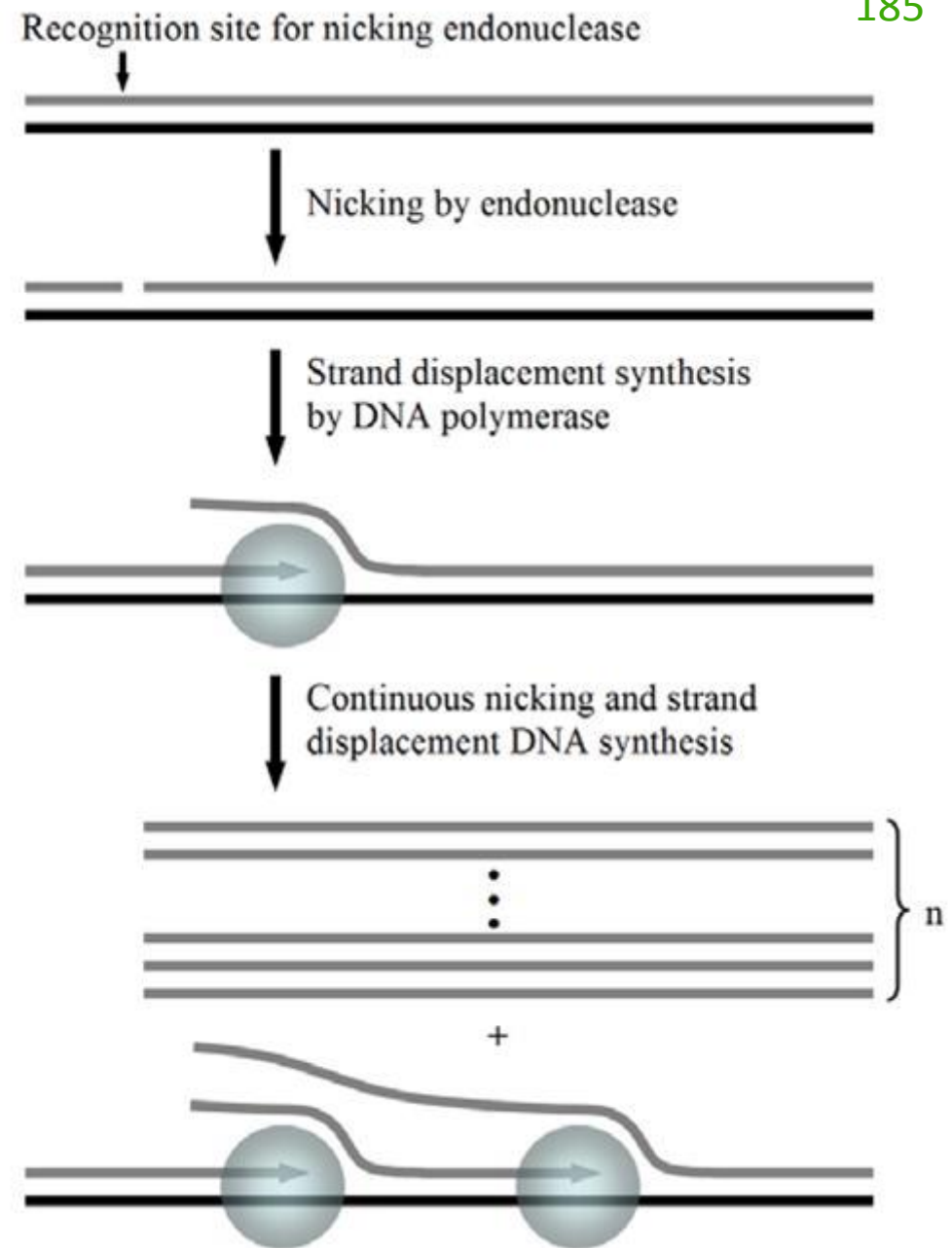


Amplification de l'ADN cible :  
obtention de multi-copie

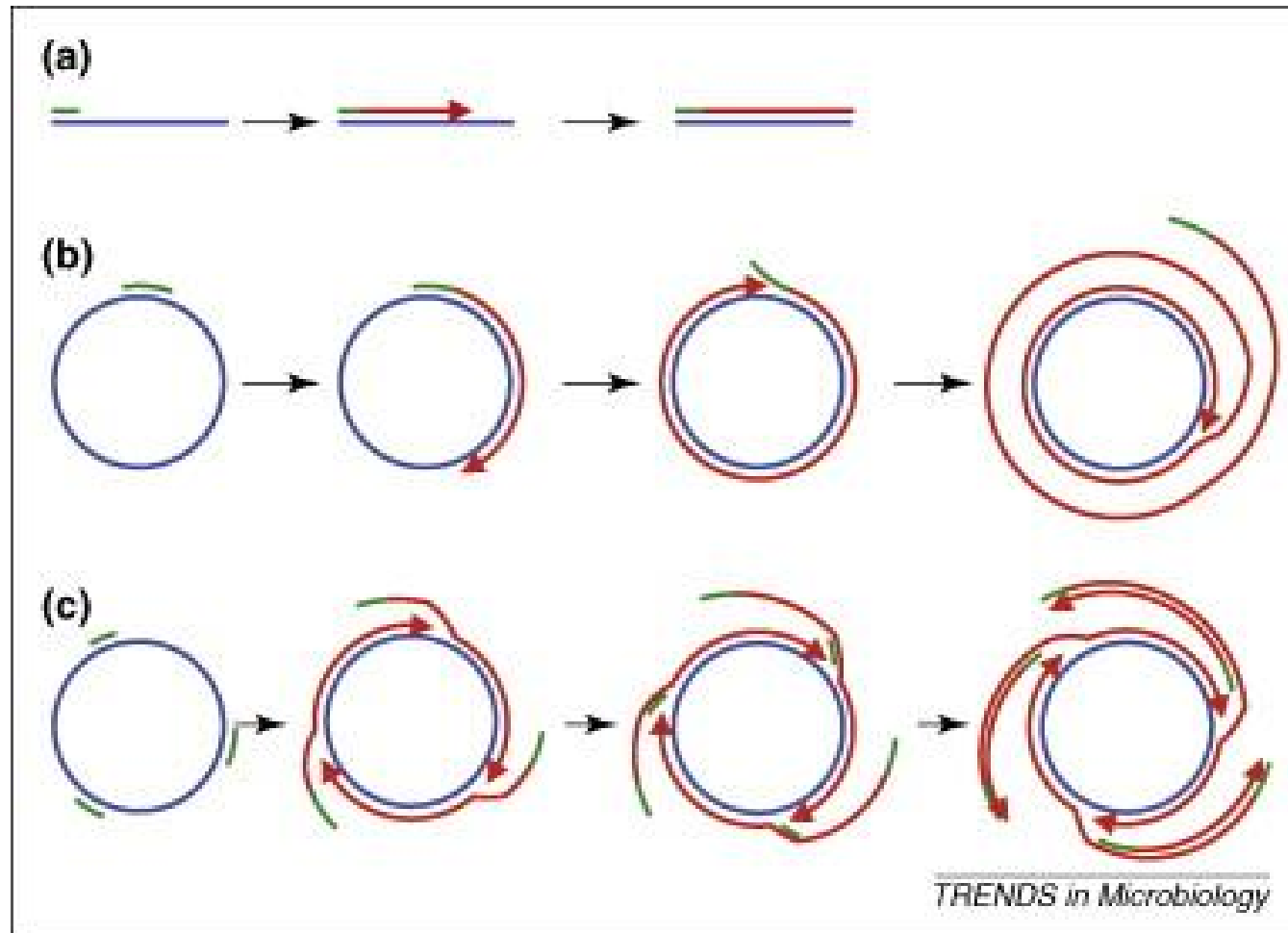
**SDA** Strand Displacement Amplification :  
ADN polymérase (de type Bst , NEB ou  
Klenow fragment) déplace le brin d'ADN pour  
initier la réplication au niveau de trous créés  
par une endonucléase (SS) ou nicking  
enzyme sur un site contenant une amorce.

= NEAR (Nicking enzyme amplification  
reaction ) (ANSR de chez Neogene)

Coupe simple brin par nicking enzyme  
Polymérase qui déplace le brin : obtention  
de multiple copies cibles



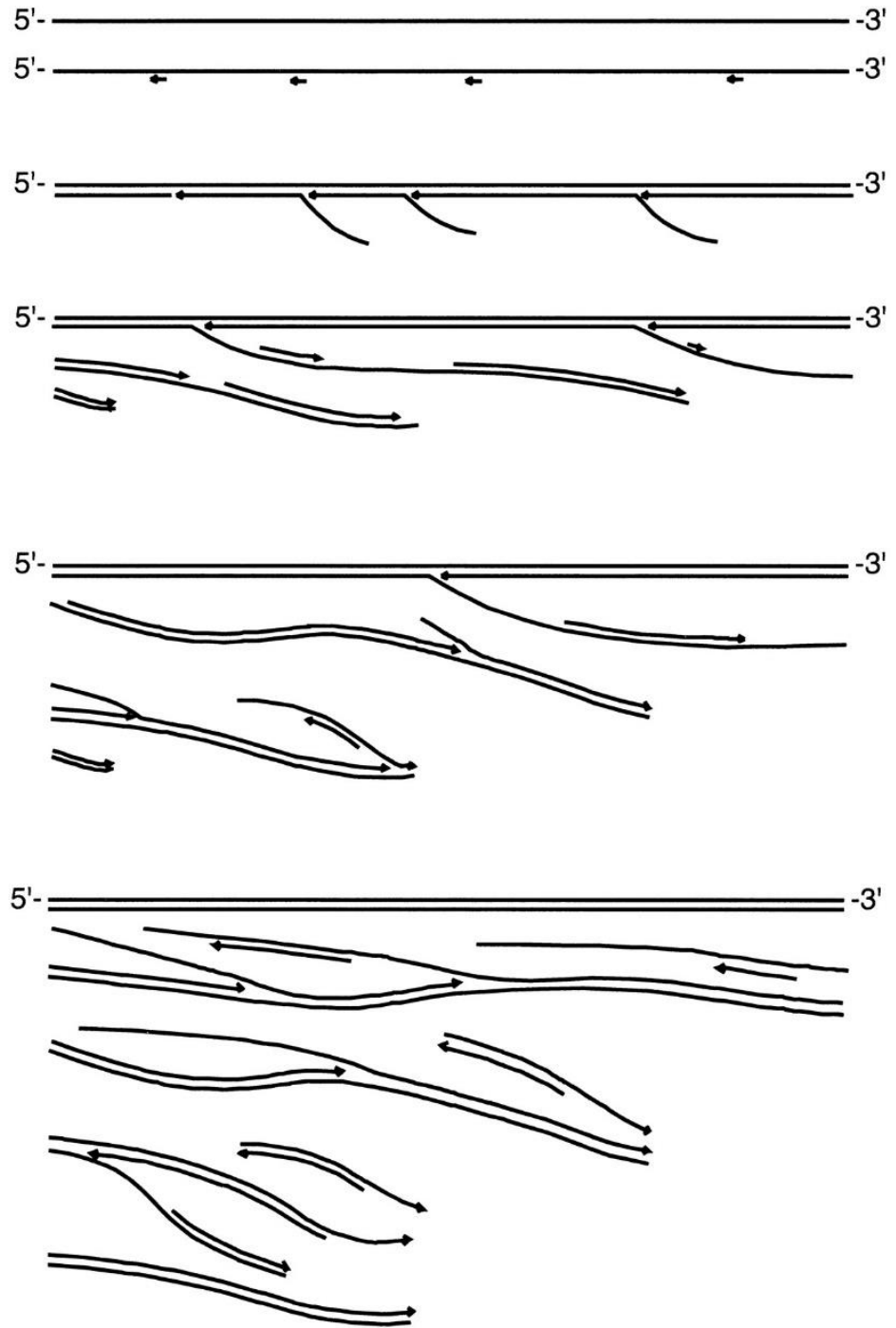
RCA rolling circle amplification  
(à partir de matrices circulaires)





**MDA**

**Multiple  
Displacement  
Amplification**



**HDA** Helicase Dependant Amplification : méthode approuvée par la FDA (utilisation d'une hélicase, une polymérase, et seulement 2 amorces).

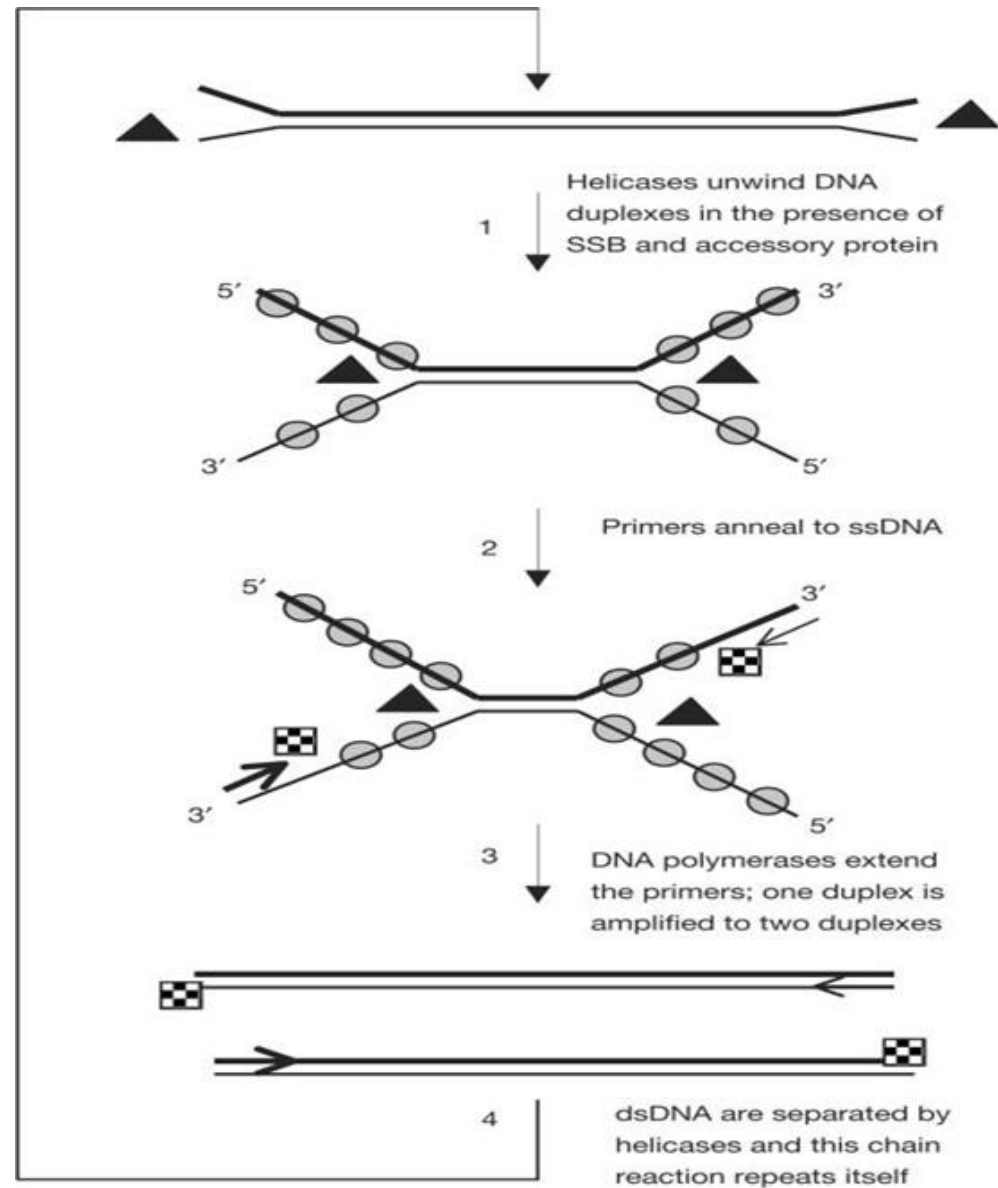
Vincent et al., Embo reports, 2004

(tHDA)

(avec RT possibilité d'appliquer à ARN)

pWGA primase-based amplification

EXPAR Chain reaction isothermal  
EXponential Amplification Reaction

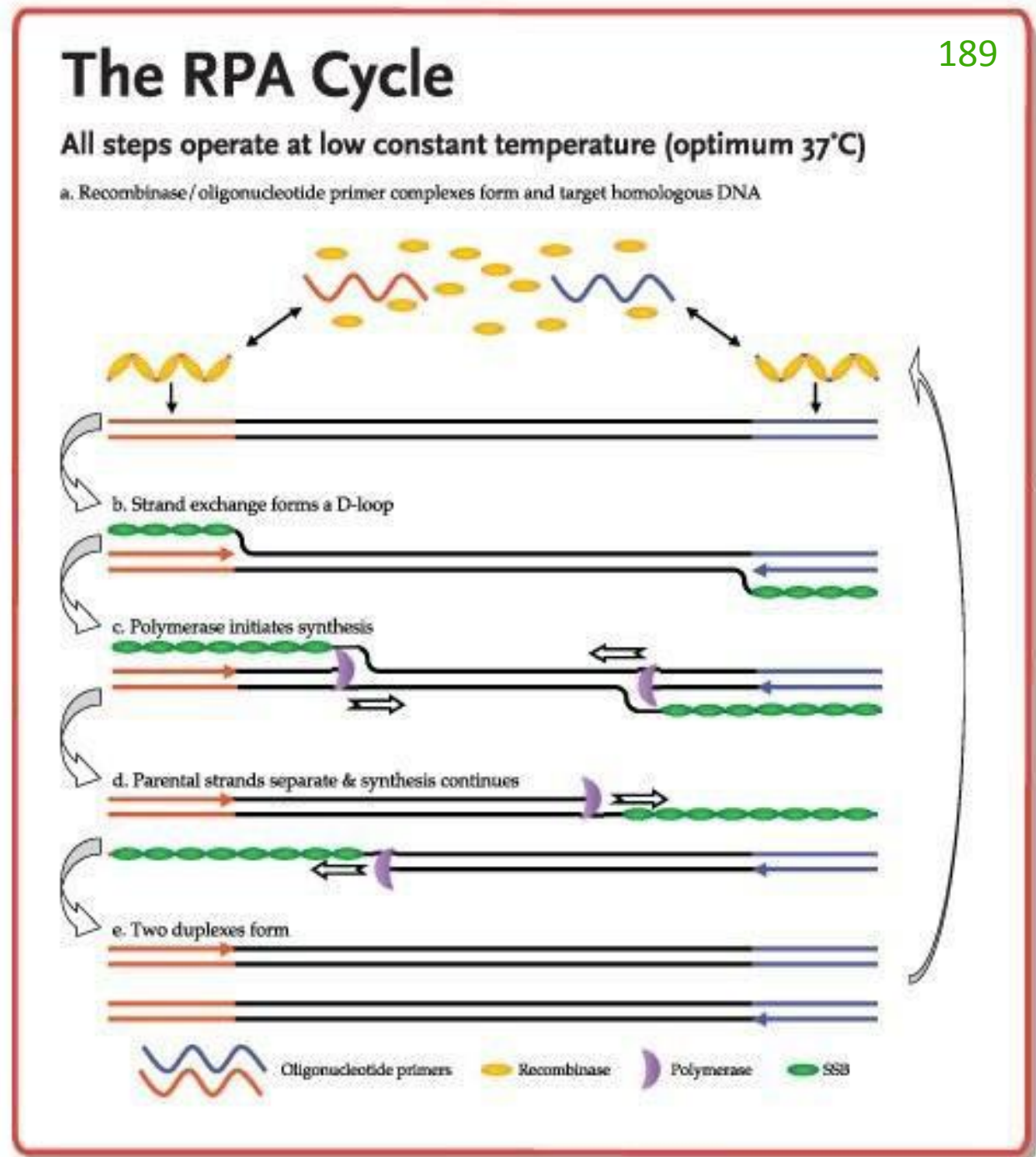


# RPA Recombinase Polymérase Amplification

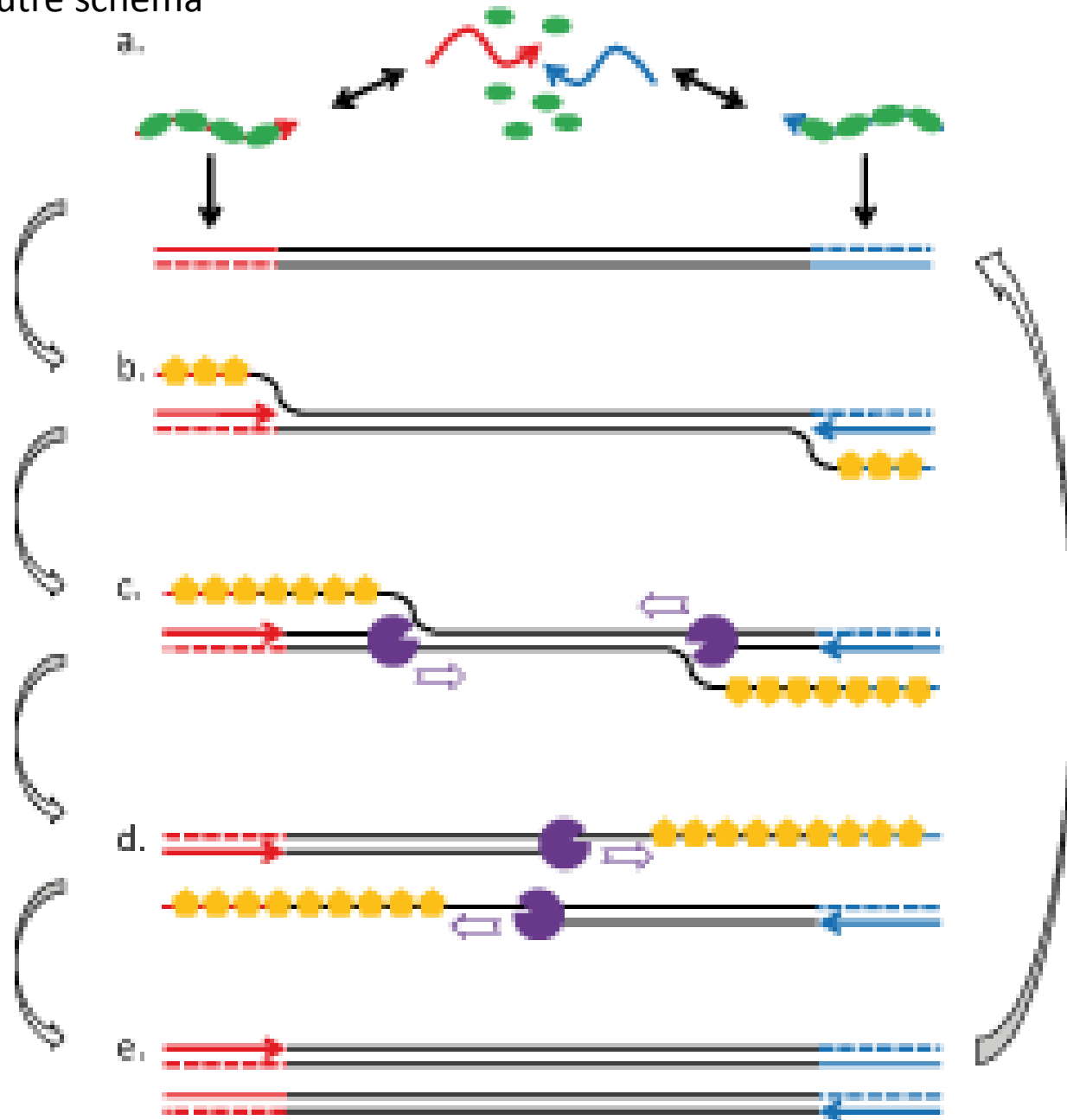
a recombinase,

a single-stranded DNA-  
binding protein (SSB)

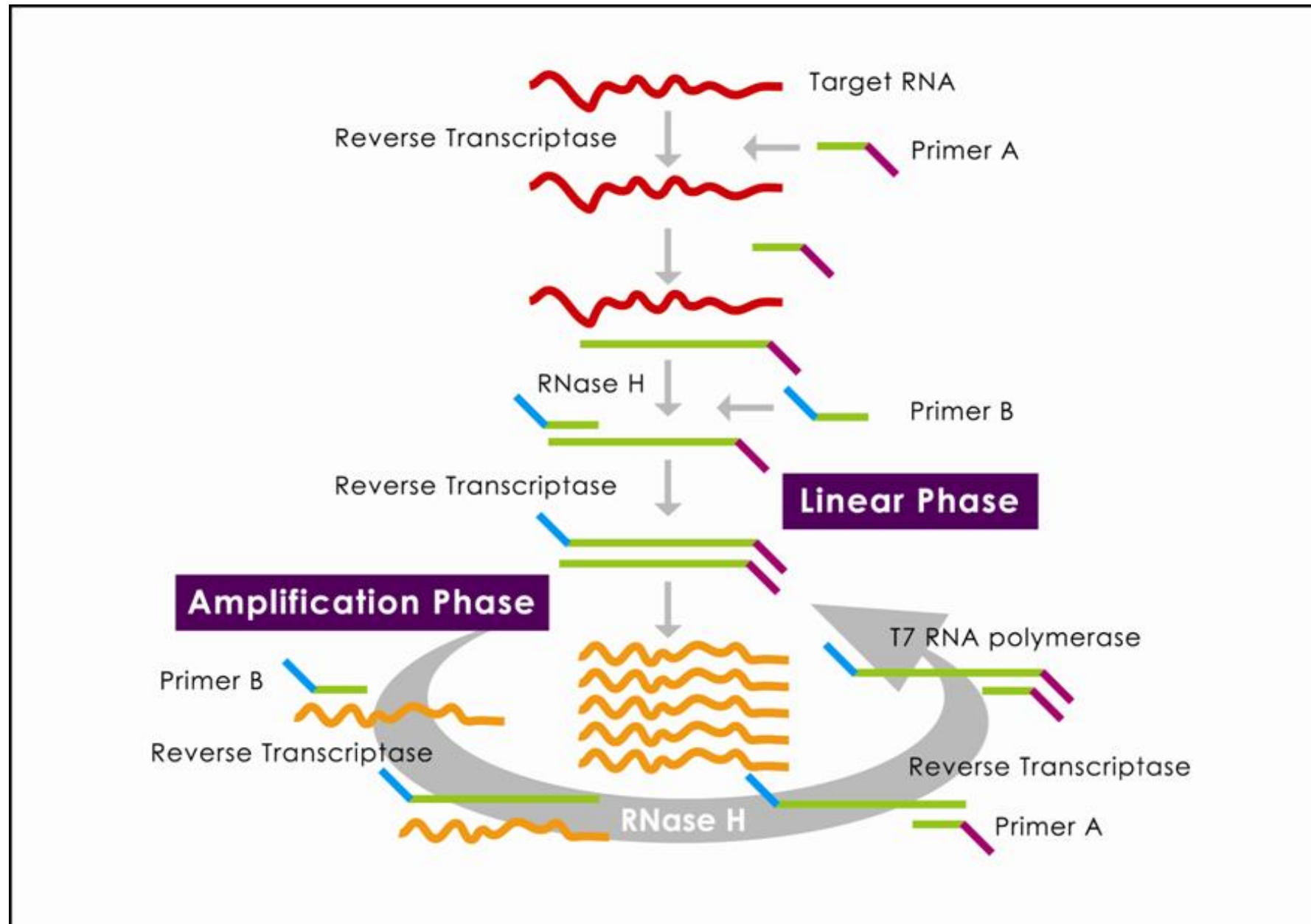
strand-displacing  
polymerase.



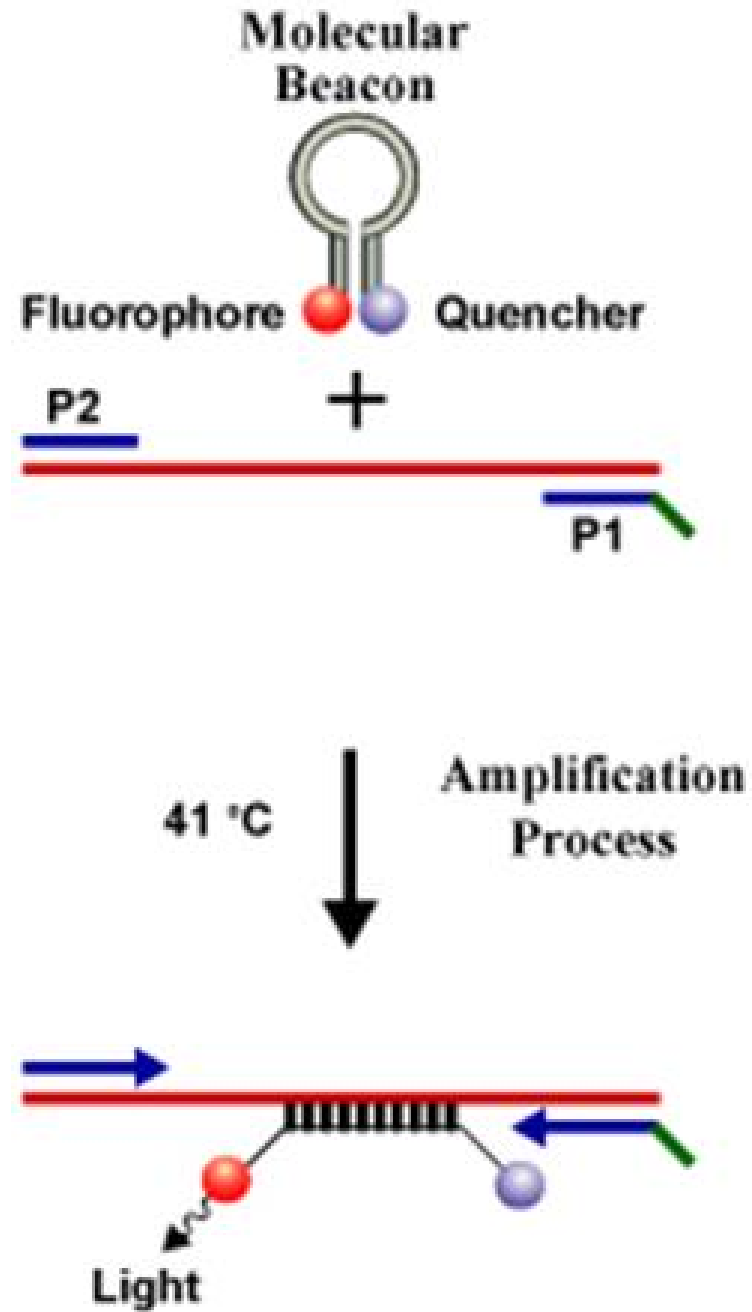
## RPA autre schéma



## NASBA amplification de l'ARN



## NASBA détection du signal



# **Production de protéines recombinantes**

Définition : Protéines produites par des cellules productrices qui sont modifiées génétiquement (transgène)

Objectif : Obtenir une grande quantité de protéine et pouvoir les purifier

-Etude in vitro, étude structurale

-Usage médicale

-Usage industriel

- Transgène placé dans vecteur d'expression souvent sous le contrôle d'un promoteur inductible (toxicité liée à la surexpression de la protéine réduite)
- Choix du système d'expression dépend de la quantité et de la qualité des protéines que l'on souhaite produire (Système d'expression hétérologue)

Marché des protéines recombinantes : 100 Mds d'€ (2012)



## □ Production de protéines recombinantes

### ❖ Production en *E.coli*

Économique

Transformation d'un plasmide avec un gène d'intérêt

Gène de résistance aux AB : sélection

OK pour protéine de taille moyenne avec peu de modifications post-traductionnelles

(Pas de glycosylations possible, Peu de ponts S-S, Phosphorylation possible (certaines))

Optimisation : choix de la souche bactérienne, du promoteur, des conditions de culture et d'induction + Stratégies de purification

### ❖ Production en cellule eucaryotes

-Levures (*Saccharomyces cerevisiae* ; *Pichia pastoris*, économique)

Transgène intégré au chromosome (Pas de plasmide dans *Pichia*)

OK pour grande protéine avec modifications post-traductionnelles (pas de carboxylation ni de sulfatation)

Stratégies de purification

-Cellule d'insectes (Papillon Sf9 Sf21 ; Drosophiles S2)

Système du baculovirus (expression transitoire)

Expression de système protéique complexe (co-infection avec plusieurs baculovirus)

Modifications post-traductionnelles mais rendement plus faible

-Cellule de mammifères : Rdt faible, cout élevé OK pour conserver activité biologique

Production industrielle : transformation stable : utilisation de promoteur fort, lignées CHO ou HEK293

### ❖ Production en système acellulaire

OK pour protéine toxique (extrait de germe de blé, de réticulocytes...)

# **Q-PCR**

## **Complément**

Méthode du  $\Delta\Delta Ct$  (Livak et Schmittgen, 2001)

Contrôle endogène : gène de ménage (GAPDH, tubuline...)

Equation de la qPCR

$$N_C = N_0 (Ef)^C$$

Pour les deux gènes, on a :

$$\text{Gène cible } N_{CT} = N_{OT} (Ef)^{CT}$$

$$\text{Gène de ménage } N_{CR} = N_{OR} (Ef)^{CR}$$

### GENES

T : Target (gène cible)

R : Reference  
(gène de ménage)

### ECHANTILLONS

S : Sample (échantillon)

C : Calibrateur  
(échantillon de référence)

En résumé,

**1** Normalisation par rapport au contrôle endogène

$$Ct T - Ct R = \Delta Ct$$

**2** Normalisation par rapport à l'échantillon standard

$$\Delta Ct S - \Delta Ct C = \Delta\Delta Ct$$

**3** Variation du nombre de copie du gène cible

$$R = 2^{-\Delta\Delta Ct}$$

... en détail

$$\text{Normalized Ratio} = \frac{\frac{N_{T(S)}}{N_{R(S)}}}{\frac{N_{T(C)}}{N_{R(C)}}} = \frac{\frac{N_{T(S)} \times E_T^{CpT(S)}}{N_{R(S)} \times E_R^{CpR(S)}}}{\frac{N_{T(C)} \times E_T^{CpT(C)}}{N_{R(C)} \times E_R^{CpR(C)}}} = \frac{K_{\text{Sample}}}{K_{\text{Calibrator}}} = 1 \quad (\text{for } K_{\text{sample}} = K_{\text{calibrator}}) \Rightarrow$$

$$\text{Normalized Ratio} = \frac{\frac{N_{T(S)}}{N_{R(S)}}}{\frac{N_{T(C)}}{N_{R(C)}}} = \frac{\frac{E_T^{CpT(C)}}{E_R^{CpR(C)}}}{\frac{E_T^{CpT(S)}}{E_R^{CpR(S)}}} = \frac{E_T^{CpT(C)}}{E_R^{CpR(C)}} \times \frac{E_R^{CpR(S)}}{E_T^{CpT(S)}} = \boxed{E_T^{CpT(C) - CpT(S)} \times E_R^{CpR(S) - CpR(C)}}$$

Si l'efficacité des 2 PCR est identique ( $E = E_T = E_R$ ),

Le facteur R représente le facteur d'expression du gène cible dans l'échantillon par rapport à son expression dans le calibrateur.

$$R = E^{\Delta\Delta Ct} = E^{CtT(C) - CtT(S) + CtR(S) - CtR(C)} = E^{(CtT(C) - CtR(C)) - (CtT(S) - CtR(S))} = E^{\Delta CtC - \Delta CtS} = E^{-(\Delta CtS - \Delta CtC)}$$

Avec  $\Delta Ct = C_{tT} - C_{tR}$

Si  $E = 100\%$

Alors  $Ef = 2$  et  $R = 2^{-\Delta\Delta Ct}$

**Application :**

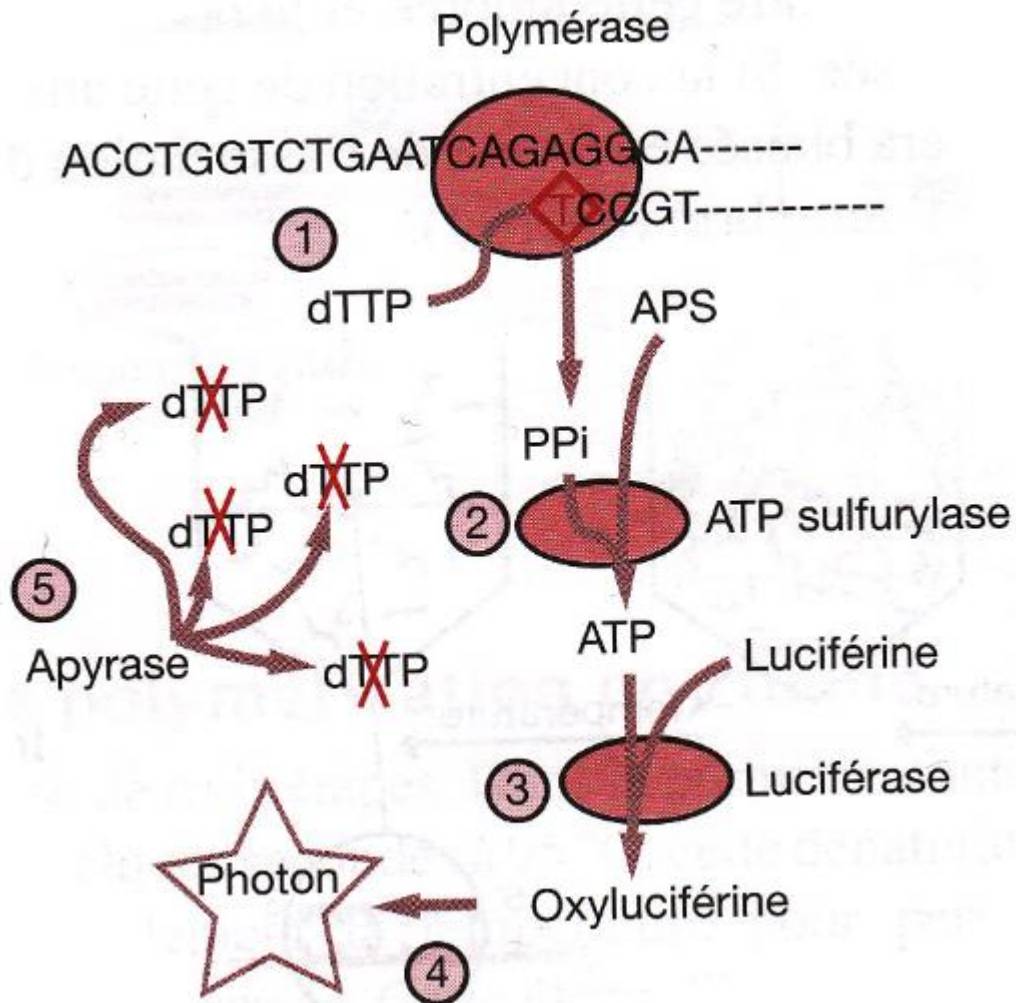
Variation de l'expression d'un gène (*Xfac*) chez une levure cultivée sur milieu YEG + antifongique au cours du temps (*hkg*: gène de ménage).

T 0h	<i>hkg</i>	<b>Ct 14</b>
	<i>Xfac</i>	<b>Ct 32</b>
<hr/>		
T 24h	<i>hkg</i>	<b>Ct 16</b>
	<i>Xfac</i>	<b>Ct 29</b>

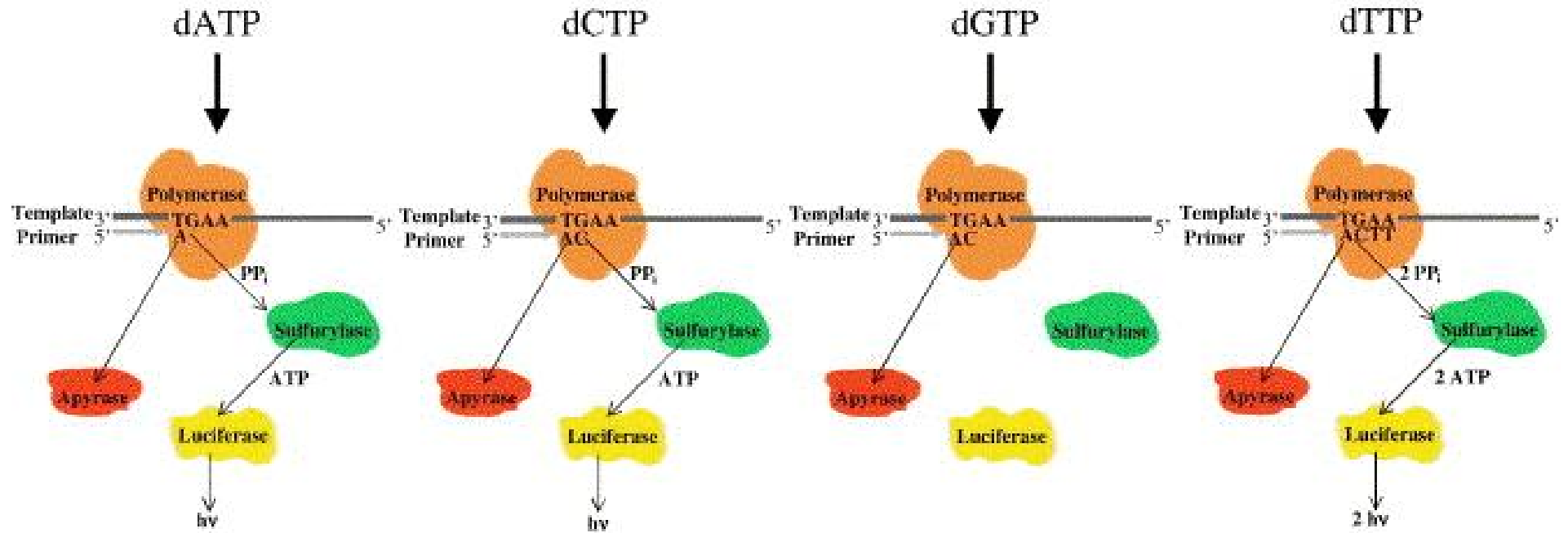
## ❑ Nouvelles technologies de séquençages

- ✓ 454 Pyrosequencing / Roche 2005
- ✓ Illumina / Solexa 2006
- ✓ SOLiD systems 2007
- ✓ Ion Torrent 2010
- ✓ PacBio 2011
- ✓ Futures technologies

❑ Séquençage 454 : principe du pyroséquençage



1. La polymérase incorpore le dTTP
2. Libération d'un PPI utilisé pour générer 1 ATP
3. La luciférase utilise un ATP
4. La luciférase va émettre un photon
5. Nt non incorporés + ATP dégradés par l'apyrase





## ❑ Séquençage 454 : principe du pyroséquençage

❖ Séquence déterminée en direct lors de la synthèse de l'ADN

❖ Procédure :

- Brins d'ADN fixés individuellement à des billes microscopiques
- Chacune de ces billes est placée séparément dans des puits sur une plaque (1,7 M de puits / microplaque)
- ADN amplifié (PCR en émulsion) pour donner un très grand nombre de molécule
- dNTP ajoutés séquentiellement (cad les uns après les autres) (à chaque incorporation du nucléotide complémentaire , libération d'un PPI, réaction enzymatique via luciférase, libération d'un photon).
- Détection du photon par un capteur colorimétrique (la succession des signaux détectés permet d'interpréter quelles bases ont été ajoutées et dans quel ordre).

❖ Taille des lectures (reads) très longue : 250 pb jusqu'à 900 pb (technologie GS-FLX++) : facilitation de l'assemblage de séquence (assemblage de génome)

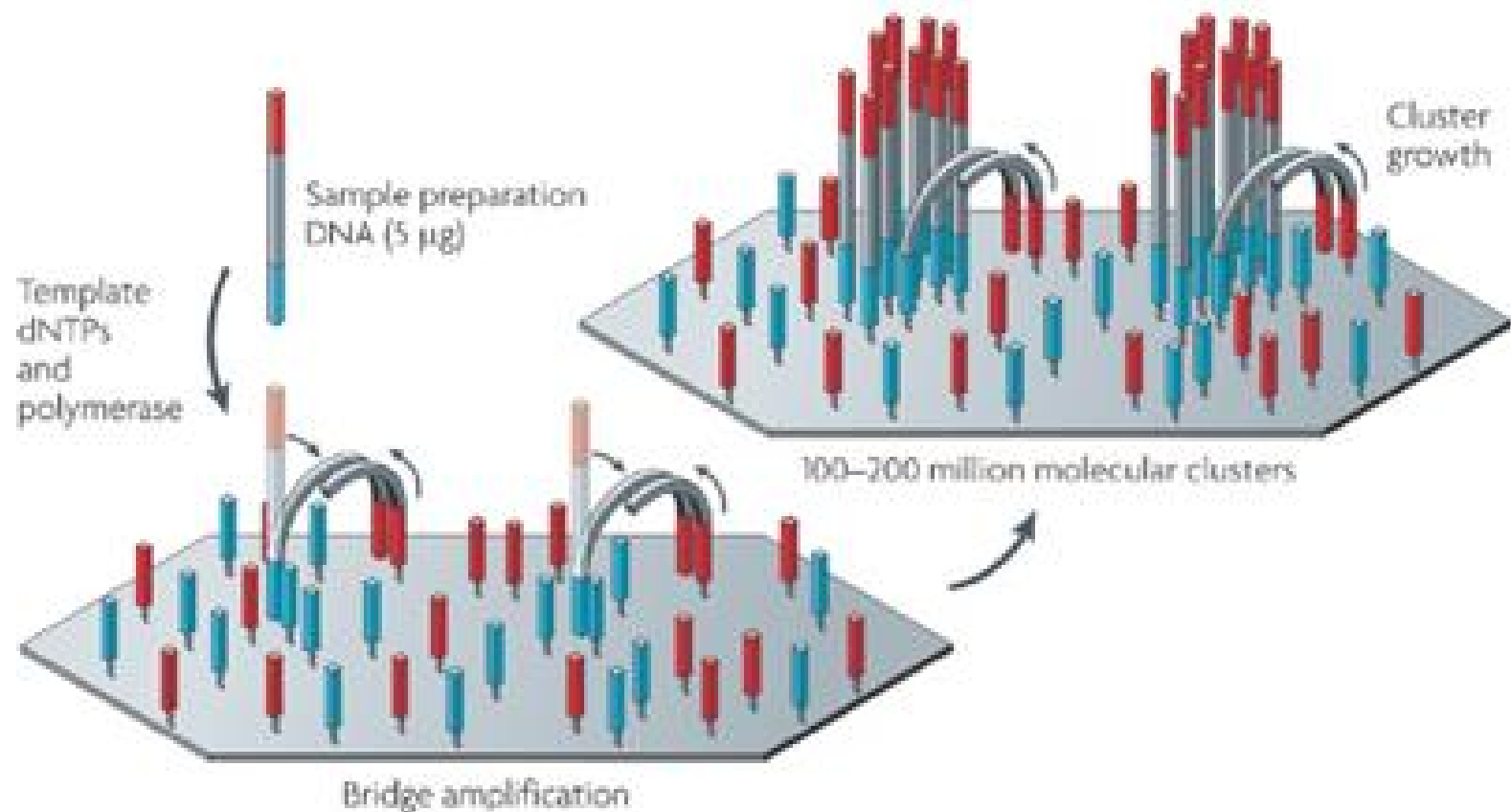
❖ Nombre de lecture limité (1,5 M max)

❖ Prix élevé

❖ Erreurs : erreurs homopolymériques (saturation du signal colorimétrique), erreurs d'insertion /délétion => inflation de la diversité taxonomique

❑ Illumina/Solexa (Env. 56% du marché des NGS).

- ❖ Basée sur la synthèse de l'ADN sur support solide : formation de clusters
  - ❖ Utilisation de dNTP terminateur de synthèse réversible marqués
  - ❖ Reads plus courts (100 – 300 pb)
  - ❖ Puissance de séquençage supérieure (150 M de reads)
  - ❖ Prix très bas
- ❖ « *single-end sequencing* » Séquençage à partir d'une extrémité  
« *paired-end sequencing* » Séquençage à partir des 2 extrémités

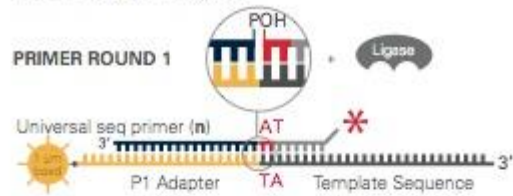


❑ Solid système

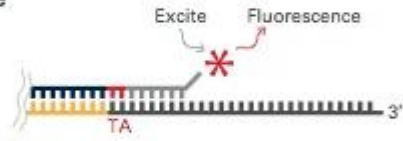
<http://seqanswers.com/forums/showthread.php?t=10>

- ❖ Amplification de l'ADN sur microbille placée ensuite sur surface de verre
- ❖ Ligation de sondes marquées sur l'ADN
- ❖ Biais sur séquences palindromiques

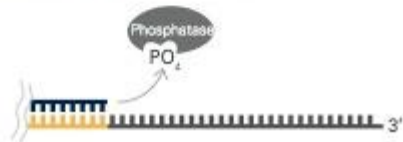
### 1. Prime and Ligate



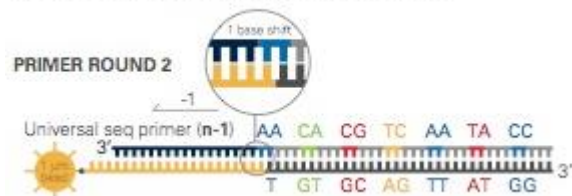
### 2. Image



### 3. Cap Unextended Strands



### 7. Repeat steps 1-5 with new primer



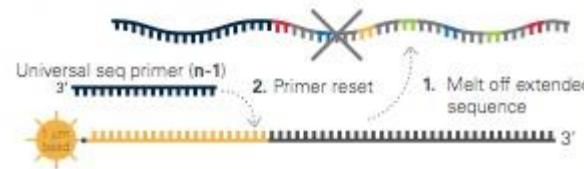
### 4. Cleave off Fluor



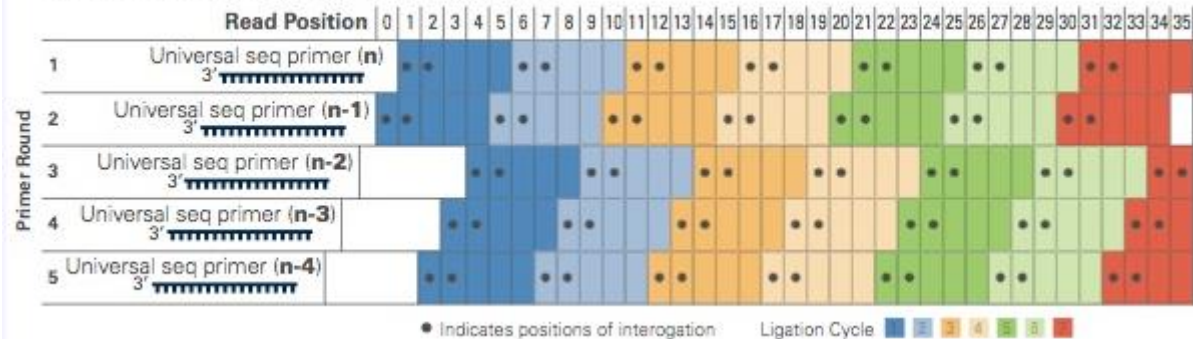
### 5. Repeat steps 1-4 to Extend Sequence



### 6. Primer Reset



### 8. Repeat Reset with , n-2, n-3, n-4 primers



Sondes de 8 nt comportant des marquages différents en fonction des 2 premières bases.

Hybridation à l'amorce à l'aide de la ligase et si la séquence est complémentaire => Détection fluorescence ;

Régénération (élimination fluorescence + les 3 derniers nt de l'oligomère de 8 nt).

Opération répétée avec des amorces de longueurs n-1; n-2; n-3 et n-4, les 5 premiers nt de la séquence peuvent être déterminés et ainsi de suite.

## ☐ Ion Torrent

Technologie proche du pyroséquençage

Signal de polymérisation de l'ADN détecté par la libération du  $H^+$  grâce à une nano électrode à pH

Taille des lectures (reads) relativement longue : 250 pb

Faible coût, appareil miniaturisé

## ❑ PacBio RS /Pacific Biosciences

Séquençage réalisé sur une molécule unique d'ADN avec une ADN polymérase

Structure nanophotoniques 70-100 nm (Zero-mode Waveguide)  
dNTP fluorescents, signal capturé par ZMW

Reads : taille jusqu'à 30 kb

Nombre de reads : 50000

Assemblage de génome complexe / correction des erreurs  
d'assemblage

### *Technologies nouvelles génération :*

Séquençage en temps réel

Sans étape de pré-amplification

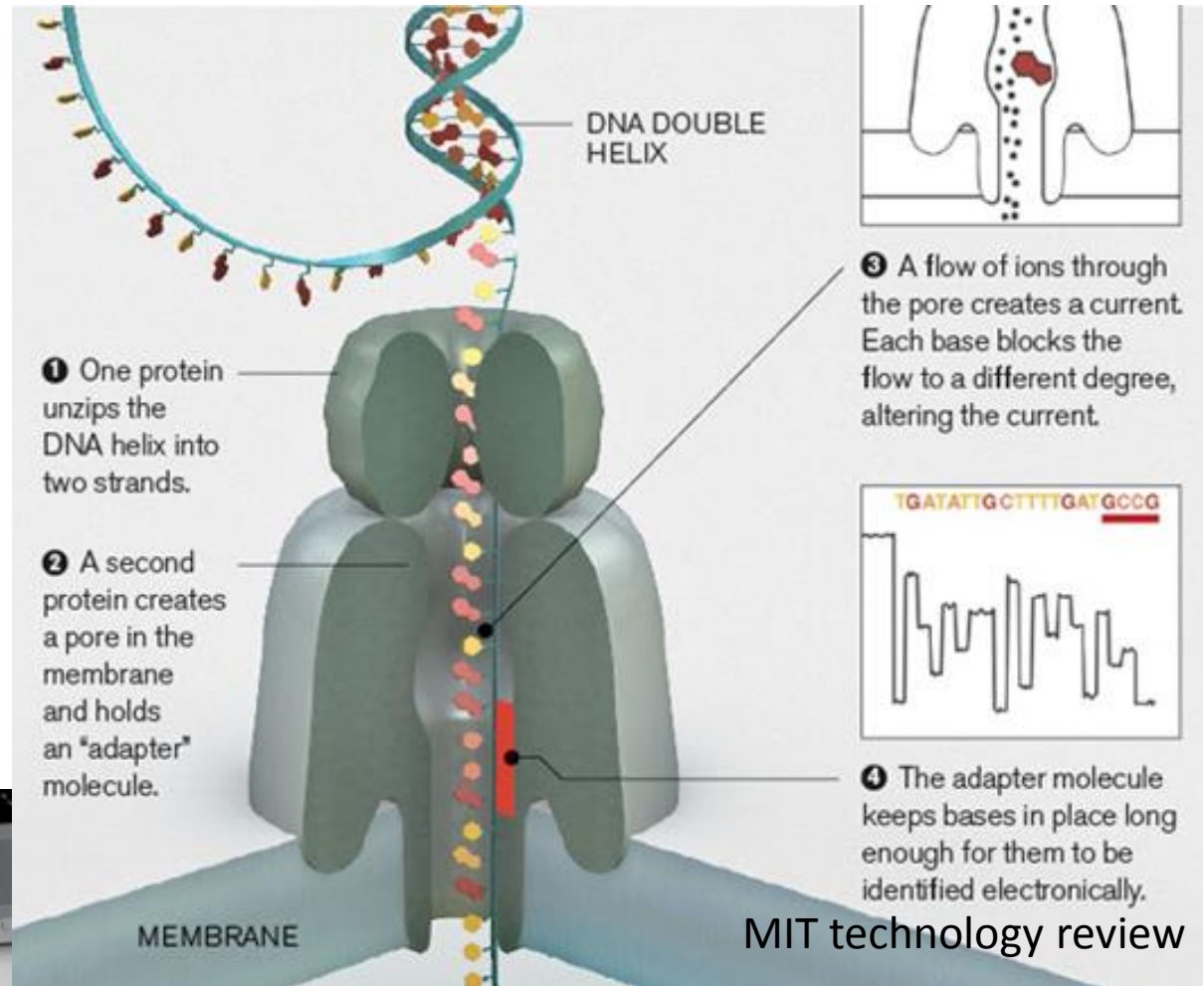
Analyse de la transcription sans passer par une  
étape de rétrotranscription

❑ Oxford Nanopore : Séquençage par transit d'un brin d'ADN au travers de nanopores (alpha-hémolysine / cyclodextrine) 205

L'acide nucléique qui transite par le pore change son courant d'ion. (f. forme, taille et longueur de la séquence)

Chacun des nt bloque le flux d'ions par le pore pendant un temps qui lui est spécifique.

**Technologie MinIon™**



Marketing:

appareil de la taille d'une clé USB  
Coût < 900\$

☐ Autres technologies / entreprises NGS nouvelles générations

Helicos (faillite 2012),

Noblegen,

Starlight,

Cracker Bio,

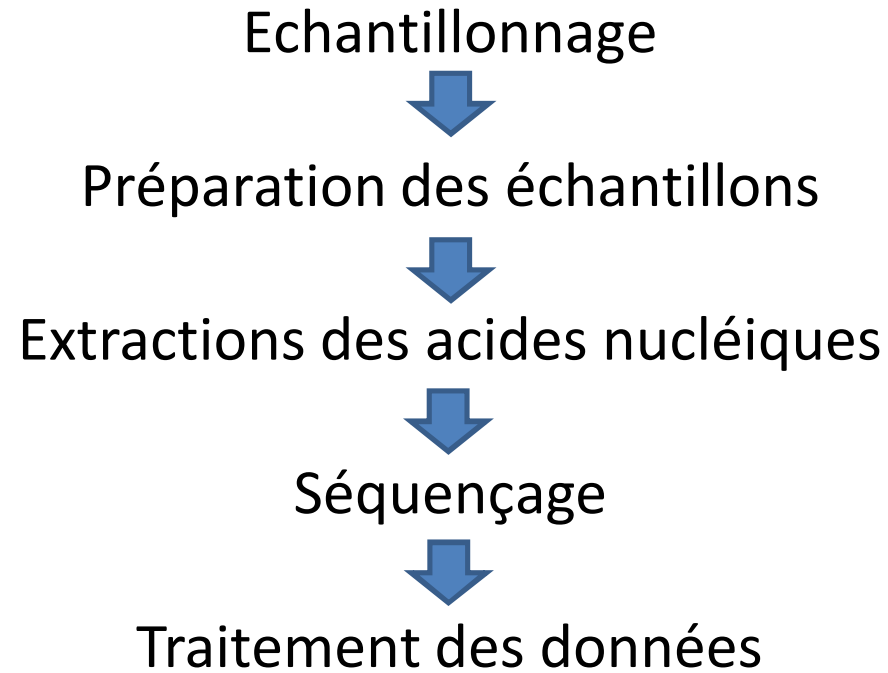
NABSys,

Halcyon

Illumina Tru-seq Synthetic Long-Read™



☐ NGS : méthodologie générale



❑ Les NGS permettent d'aborder :

-le séquençage d'un génome inconnu (séquençage de novo )

ou le re-séquençage d'un génome connu

-l'annotation (ou la ré-annotation) de plus en plus précise et exhaustive d'un génome

-l'étude du transcriptome, analyse « quantitative » du niveau de transcription des gènes

-étude des évènements de l'épissage

-quantification et détection d'ARN rares,

identification de régions dont on ne savait pas au préalable qu'elles sont transcrites

étude du profil en petits ARN non codants, découverte de gènes codant ces types d'ARN

l'édition des ARN : modification post- ou co-transcriptionnelle

-Épigénétique

étude du profil de méthylation de l'ADN et de la structure de la chromatine

étude des interactions ADN / protéines

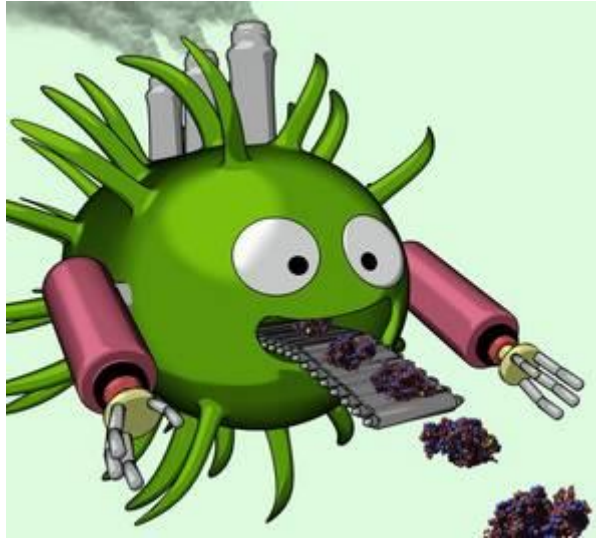
étude des modifications post-traductionnelles des histones

-la métagénomique : étude du génome d'un organisme prélevé directement dans un environnement complexe

l'étude du microbiome. Exemple : "Human Microbiome Project"

le séquençage simultané d'un très grand nombre d'échantillon ("Next-generation DNA barcoding")

...

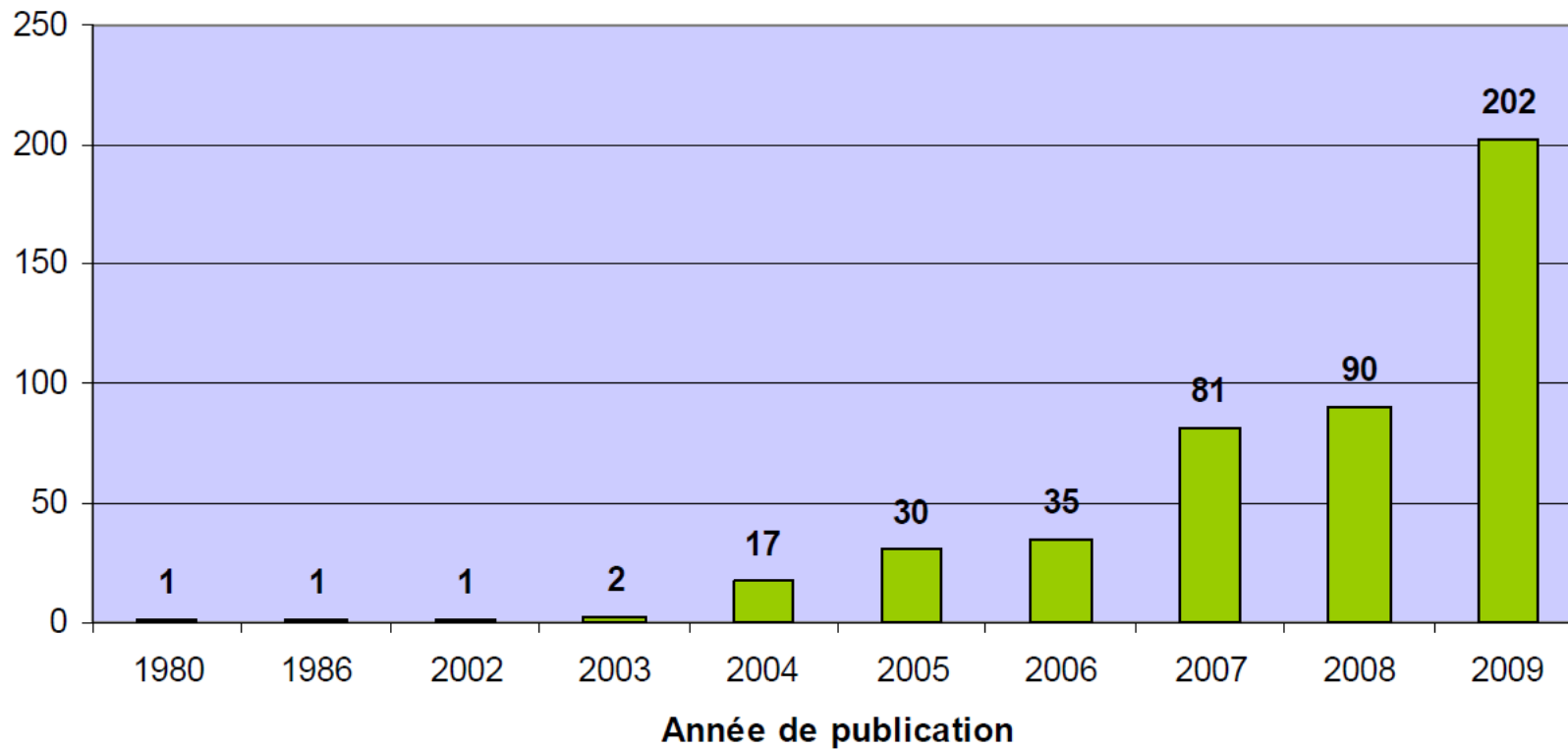


# Biologie synthétique

**POURQUOI ?**

## Intérêt croissant de la communauté scientifique

Nombre de références comportant le terme exacte "synthetic biology" recensées, par année, dans PubMed



## Intérêt de la classe politique

**Rapport parlementaire** demandé à l'OPECST **présentant les enjeux de la biologie de synthèse** , daté du mois de février 2012.

OPECST = Office parlementaire d'évaluation des choix scientifiques et technologiques

**auteur** : Geneviève Fioraso (Députée).

Ministre de l'enseignement supérieur et de la recherche (2012 – 2015).

<http://www.assemblee-nationale.fr/13/rap-off/i4354.asp>

Discours E. Macron octobre 2015 : économie numérique et « bigdata »

## Des enjeux économiques considérables

De nombreuses applications dans différents secteurs :

- ❖ **Santé** : outils de diagnostics améliorés, médicaments et vaccins nouveaux, thérapeutique moins intrusive
- ❖ **Chimie et ressources naturelles** : Secteur de **bio-production**, avec de nouvelles techniques moins polluantes : bio-fuels, matériaux issus de la pétrochimie, ingrédients alimentaires...
- ❖ **Bio-matériaux ou matériaux intelligents**, grâce aux outils issus de ces recherches
- ❖ **Environnement** : Bio-remédiation pour traiter les milieux contaminés ou les eaux usées, Bio-senseurs : spectre d'application potentiel très large, qui inclut la lutte contre le bio-terrorisme, traitement et valorisation des déchets...

# PLAN

- ❑ Contexte
- ❑ Repères historiques
- ❑ Les définitions et les différentes approches
- ❑ Technologies clés
- ❑ Applications
- ❑ L'origami ADN
- ❑ Développement économique, Recherche et Enseignement
- ❑ Les entreprises, les acteurs
- ❑ L'éthique et les risques
- ❑ Perspectives



## L'ère de la biologie synthétique

### RESEARCH ARTICLE

# Creation of a Bacterial Cell Controlled by a Chemically Synthesized Genome

Daniel G. Gibson,<sup>1</sup> John I. Glass,<sup>1</sup> Carole Lartigue,<sup>1</sup> Vladimir N. Noskov,<sup>1</sup> Ray-Yuan Chuang,<sup>1</sup> Mikkel A. Algire,<sup>1</sup> Gwynedd A. Benders,<sup>2</sup> Michael G. Montague,<sup>1</sup> Li Ma,<sup>1</sup> Monzia M. Moodie,<sup>1</sup> Chuck Merryman,<sup>1</sup> Sanjay Vashee,<sup>1</sup> Radha Krishnakumar,<sup>1</sup> Nacyra Assad-Garcia,<sup>1</sup> Cynthia Andrews-Pfannkoch,<sup>1</sup> Evgeniya A. Denisova,<sup>1</sup> Lei Young,<sup>1</sup> Zhi-Qing Qi,<sup>1</sup> Thomas H. Segall-Shapiro,<sup>1</sup> Christopher H. Calvey,<sup>1</sup> Prashanth P. Parmar,<sup>1</sup> Clyde A. Hutchison III,<sup>2</sup> Hamilton O. Smith,<sup>2</sup> J. Craig Venter<sup>1,2\*</sup>

We report the design, synthesis, and assembly of the 1.08–mega–base pair *Mycoplasma mycoides* JCVI-syn1.0 genome starting from digitized genome sequence information and its transplantation into a *M. capricolum* recipient cell to create new *M. mycoides* cells that are controlled only by the synthetic chromosome. The only DNA in the cells is the designed synthetic DNA sequence, including “watermark” sequences and other designed gene deletions and polymorphisms, and mutations acquired during the building process. The new cells have expected phenotypic properties and are capable of continuous self-replication.

Science, Volume 329 no. 5987 pp. 52-56, 2 July 2010

## Création d'une cellule synthétique

- ❑ Le **20 mai 2010**, le John Craig Venter Institute annonce la «Création d'une cellule "synthétique" » dans la revue **Science** :

« **Creation of a Bacterial Cell Controlled by a Chemically Synthesized Genome** ».

- ❑ L'article décrit la transplantation d'un génome artificiel de *Mycoplasma mycoides* dans une cellule de *Mycoplasma capricolum*, de façon à créer artificiellement des cellules de *Mycoplasma mycoides* contrôlées uniquement par l'ADN synthétique (*Mycoplasma genitalium* 580kb, 485 gènes )



Controverses sur le plan **éthique**



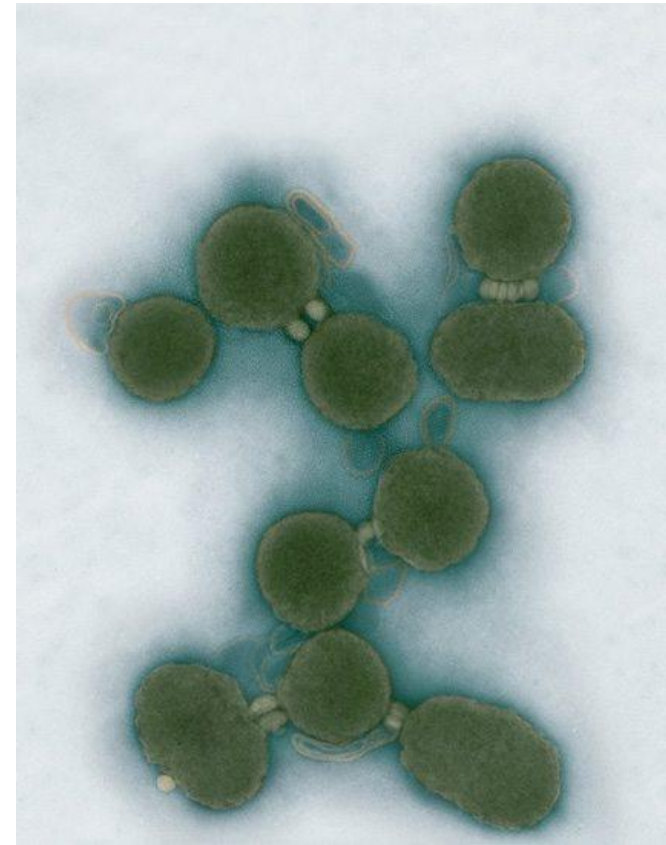
Intérêts des **applications** potentielles

❑ 1<sup>ère</sup> bactérie synthétique développée par une vingtaine de chercheurs du **J. Craig Venter Institute**.

❑ La cellule synthétique a été baptisée :

***Mycoplasma mycoides* JCVI-syn1.0**

❑ Elle prouve que les génomes peuvent être dessinés par ordinateur, fabriqués artificiellement dans un laboratoire et transplantés dans une cellule receptrice pour produire une nouvelle cellule autorépliquative uniquement contrôlée par le génome synthétique.



Negatively stained transmission electron micrographs of aggregated *M. mycoides* JCVI-syn1.0

Nombreuses applications : biocarburants, médicaments, vaccins, traitement des eaux et des aliments, dépollution des marées noires

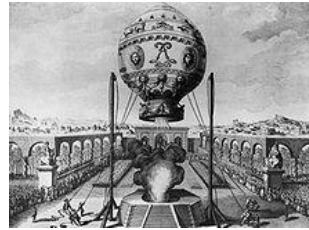
# REPERES HISTORIQUES



Antiquité



1500 De Vinci



1783 Frères Mongolfier



1903 Frères Wright



1961 Gagarine



1969 Amstrong

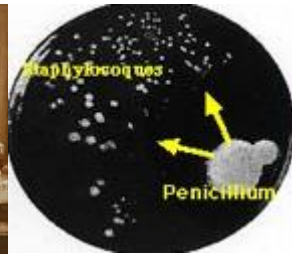


1861 Pasteur



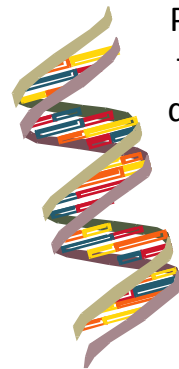
Réfutation de la génération spontanée

1928 Fleming



Découverte des antibiotiques

1953 Watson & Crick



Découverte de la structure de la double hélice d'ADN

1972 Berg

Première tentative transgénèse gène du virus SV40 dans *E. coli*

1982



Première application commerciale du génie génétique  
Approbation de l'insuline recombinante

1988

Utilisation de la Taq polymérase dans la PCR

2003

Séquençage complet du génome humain

2010



1<sup>er</sup> microorganisme synthétique

## DEFINITION(S)

Il existe différentes **démarches** qui recoupent différentes **approches** ainsi que différents **objectifs** de la biologie synthétique, d'où la difficulté d'en donner une seule définition claire

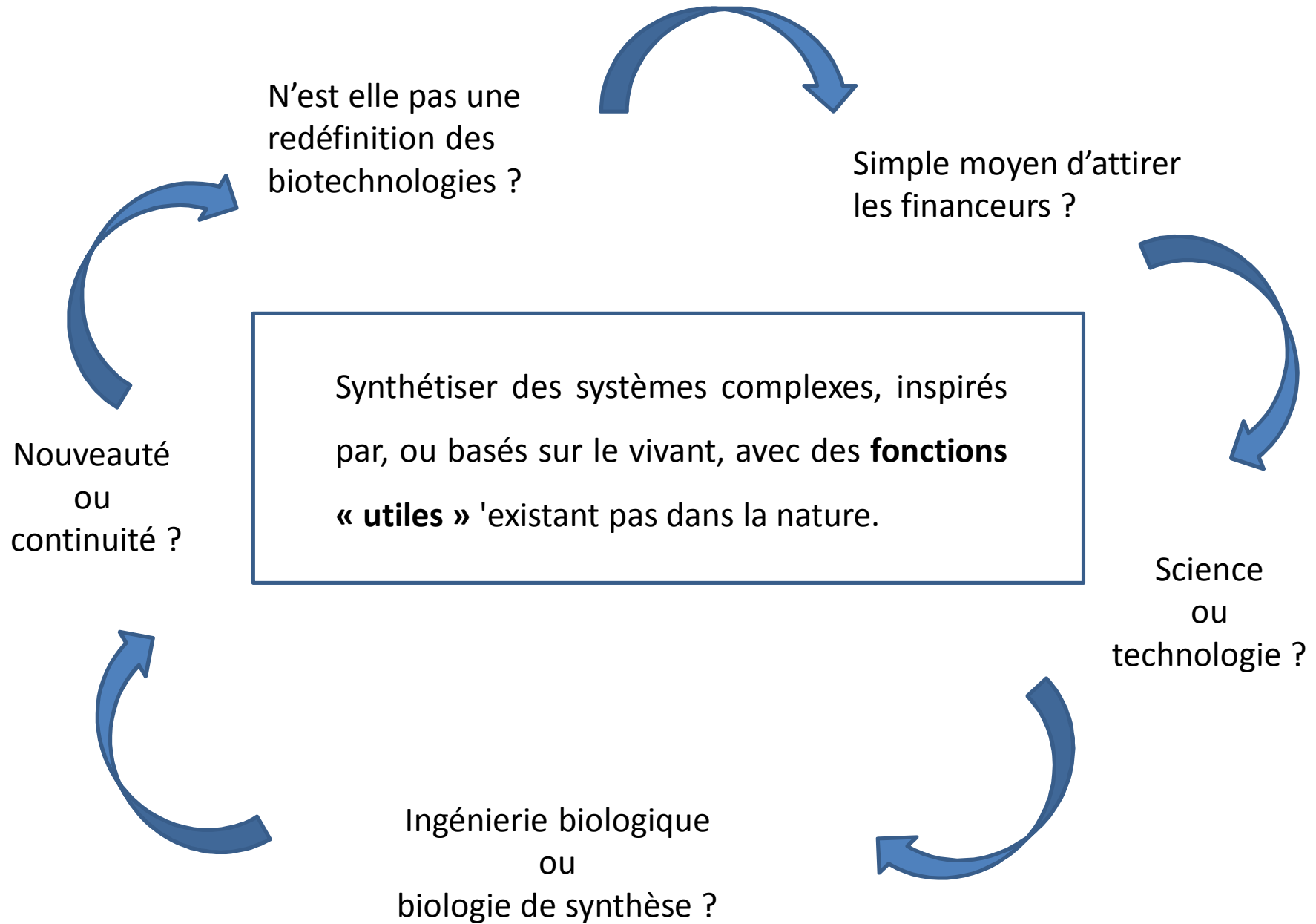
## DEFINITION(S)

La définition la plus consensuelle semble être celle proposée par le consortium européen de recherche Synbiology (6<sup>e</sup> PCRD, avec des partenariats entre la France, l'Allemagne, les États-Unis, la Grèce et le Portugal) :

*« La BS est l'ingénierie de composants et systèmes biologiques qui n'existent pas dans la nature et la réingénierie d'éléments existants ; elle porte sur le design intentionnel de systèmes biologiques artificiels, plutôt que sur la compréhension de la biologie naturelle. »*

Définition de la commission américaine de bioéthique

*« Un champ de recherche émergent, combinant des éléments de la biologie, de l'ingénierie, de la génétique, de la chimie et de l'informatique. Elle souligne qu'elle est profondément enracinée dans la biologie moléculaire et qu'elle a pu émerger grâce aux progrès réalisés dans cette discipline »*



## DEFINITION(S)

- ❑ La biologie synthétique est un domaine scientifique combinant **biologie** et principes d'**ingénierie** dans le but de concevoir et construire de nouveaux systèmes et fonctions biologiques.
- ❑ L'application efficace de ces principes d'ingénierie nécessite le développement de méthodes numérique et d'outils mathématiques venant en support.
- ❑ En fonction des différents points de vue, la biologie synthétique apparaîtrait comme l'état actuel d'**évolution des biotechnologies** ou comme une discipline nouvelle à part entière.
- ❑ Le *rubicon* à franchir de la biologique synthétique n'est pas d'engendrer des organismes que l'évolution n'a pas produits, mais des **organismes que l'évolution n'aurait pu spontanément produire**, sans intervention de la raison.



## DIFFERENTES APPROCHES

- ❑ **Approche génétique** : des réseaux génétiques aux systèmes : Créer de nouveaux composants pour les systèmes biologiques afin d'y faire apparaître de nouvelles fonctionnalités (Cette approche inclue les « biobricks »)

**Approche Bottom-up** : assemblage construction du vivant

- ❑ **Encapsulation des réseaux** : de la cellule minimale au nano réacteur : **cellules minimales** (fonctions métabolisme, reproduction et évolution - création d'une cellule semi-synthétique qui est un liposome dans lequel est encapsulé un génome) et vésicules artificielles (cell free system)

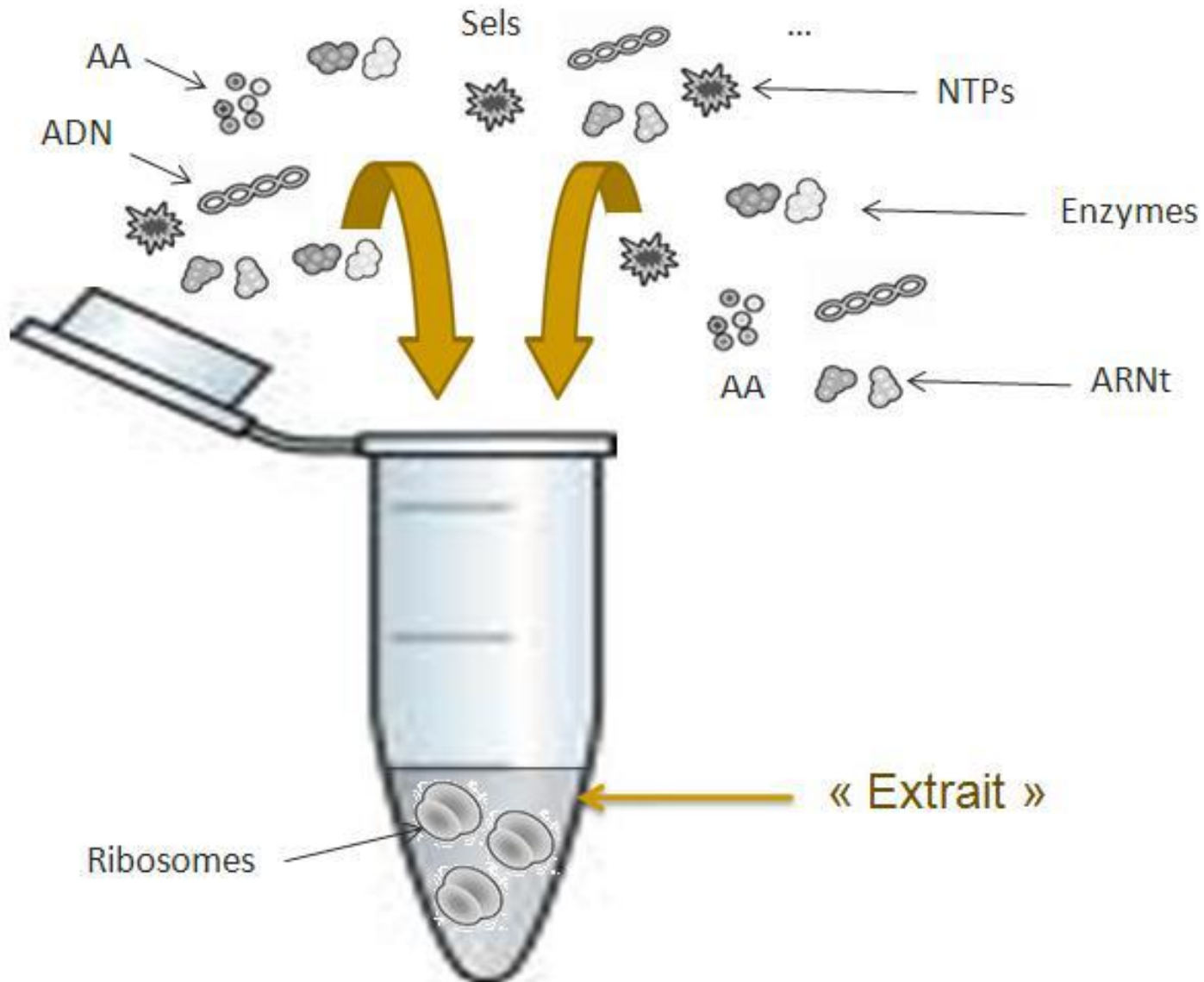
**Approche Top-down** : déconstruction du vivant

Bottom-Up



Top Down

## Protocell : système d'expression acellulaire



# HISTORIQUE

## ❑ Apparition du terme **biologie synthétique** :

Stéphane Leduc 1912, selon lui, la SCIENCE suit 3 étapes qui sont l'observation, l'analyse et la synthèse. Au début du XX<sup>e</sup> siècle, la biologie est en retard sur les autres sciences car il lui manque la synthèse.

« Méthode synthétique => compréhension de la vie »

## ❑ Découverte des Enzymes de Restriction et application aux problèmes d'ingénierie génétique (1978) : **Réapparition de la notion de biologie synthétique** .

# LA BIOLOGIE SYNTHETIQUE EST PLURIDISCIPLINAIRE

- ❑ Thermodynamique : Energétique microbienne
- ❑ Biologie métabolique
- ❑ Informatique – algorithmique
- ❑ Génie génétique, biotechnologie
- ❑ Modélisation moléculaire, prédiction de structure moléculaire
- ❑ Biochimie structurale : Relation structure - fonction
- ❑ ...



Pasteur (et d'autres) étaient des physiciens avant tout.  
1<sup>er</sup> en France à établir un lien entre monde académique  
et industriel (brasseur lillois).

# TECHNOLOGIES CLES et LEUR AUTOMATISATION

## ❑ Séquençage de l'ADN

Acquisition information,

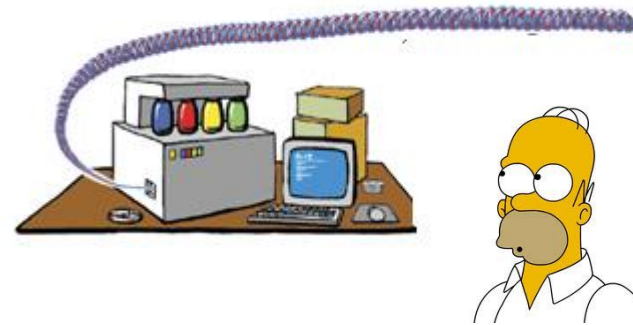
Vérification des séquences élaborées,

Détection et identification des organismes / systèmes de synthèse.


## ❑ Synthèse d'oligonucléotides

Synthèse de gènes artificiels, les « biobricks ».

## Synthèse de l'ADN AUTOMATISATION



- Augmentation de la capacité de synthèse
- Réduction des coûts

 Jusqu'à 50000 pb et 0,39€/pb

George Church, de la Harvard Medical School, et Xialian Gao, de l'Université de Houston (Texas) ont quant à eux utilisé des puces à ADN programmables pour synthétiser des oligonucléotides, les amplifier, les sélectionner par hybridation afin de réduire le taux d'erreur et enfin les assembler.

- ❑ Une biobrick est une séquence d'ADN dont l'expression réalise une fonction « biologique » particulière connue .
- ❑ Intérêt : assemblage de ces séquences à l'image d'un Lego et création d'un système biologique qui réalisera une tâche préalablement définie.
- ❑ C'est exactement ce qu'ont fait Elowitz et Leibler avec leur repressilateur :  
Une combinaison de 4 séquences (3 promotrices et 1 protéine fluorescente)



# Synthèse de l'ADN

## Biobricks

- ❑ Mises en place d'un registre de pièces biologiques par le MIT

[http://partsregistry.org/Main\\_Page](http://partsregistry.org/Main_Page)



### Welcome to the Registry of Standard Biological Parts.

The Registry is a [continuously growing](#) collection of genetic parts that can be mixed and matched to build synthetic biology devices and systems. Founded in 2003 at MIT, the Registry is part of the Synthetic Biology community's efforts to make biology easier to engineer. It provides a resource of available genetic parts to [iGEM](#) teams and academic labs. You can [register a new lab here](#).

The Registry is based on the principle of "get some, give some". Registry users benefit from using the parts and information available from the Registry in designing their engineered biological systems. In exchange, the expectation is that Registry users will, in turn, contribute back information and data on existing parts and new parts that they make to grow and improve this community resource.



Catalog of parts & devices



Help



Users & groups



DNA repositories

#### Registry tools

- [Search parts \(?\)](#)
- [Add a part](#)
- [Request a part](#)
- [Send parts to the Registry](#)
- [Sequence analysis](#)



You'll notice some significant changes to the Registry recently. In particular, the Registry [catalog of parts](#) has been entirely redesigned to allow for easier browsing of the available parts and devices. You can now browse parts and devices by type, by function, by chassis and by standard. You'll also notice that the documentation and help pages for each class of parts have been greatly enhanced.

The Registry of Standard Biological Parts is \*always\* a work in progress. Please browse the new catalog and let us know what you think, or feel free to edit and improve the pages further.

# Synthèse de l'ADN

## Biobricks

Biobricks regroupées selon leur type et leur niveau de complexité.

- « Parts » : séquences régulatrices, terminales, RBS, séquences élémentaires...
- « Devices » : système rapporteur (séquence régulatrice, RBS, GFP, terminatrice).
- « Systems » : ensemble de système rapporteurs
- « Chassis » : Systemes incorporés dans les cellules hotes

### Browse parts by type

Catalog List



**Promoters (?)**: A promoter is a DNA sequence that tends to recruit transcriptional machinery and lead to transcription of the downstream DNA sequence.



**Ribosome Binding Sites (?)**: A ribosome binding site (RBS) is an RNA sequence found in mRNA to which ribosomes can bind and initiate translation.



**Protein domains (?)**: Protein domains are portions of proteins cloned in frame with other proteins domains to make up a protein coding sequence. Some protein domains might change the protein's location, alter its degradation rate, target the protein for cleavage, or enable it to be readily purified.



**Protein coding sequences (?)**: Protein coding sequences encode the amino acid sequence of a particular protein. Note that some protein coding sequences only encode a protein domain or half a protein. Others encode a full-length protein from start codon to stop codon. Coding sequences for gene expression reporters such as LacZ and GFP are also included here.



**Translational units (?)**: Translational units are composed of a ribosome binding site and a protein coding sequence. They begin at the site of translational initiation, the RBS, and end at the site of translational termination, the stop codon.



**Terminators (?)**: A terminator is an RNA sequence that usually occurs at the end of a gene or operon mRNA and causes transcription to stop.



**DNA (?)**: DNA parts provide functionality to the DNA itself. DNA parts include cloning sites, scars, primer binding sites, spacers, recombination sites, conjugative transfer elements, transposons, origami, and aptamers.



**Plasmid backbones (?)**: A plasmid is a circular, double-stranded DNA molecules typically containing a few thousand base pairs that replicate within the cell independently of the chromosomal DNA. A plasmid backbone is defined as the plasmid sequence beginning with the BioBrick suffix, including the replication origin and antibiotic resistance marker, and ending with the BioBrick prefix.



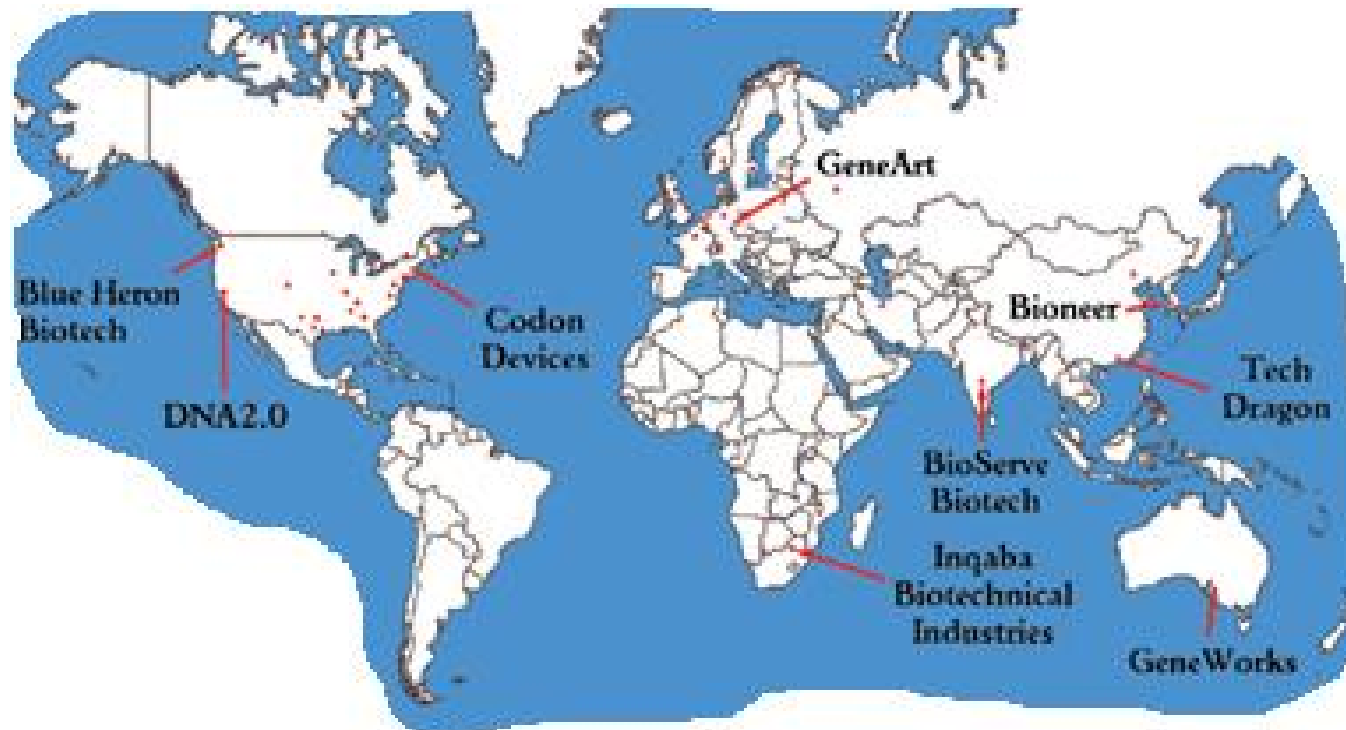
**Plasmids (?)**: A plasmid is a circular, double-stranded DNA molecules typically containing a few thousand base pairs that replicate within the cell independently of the chromosomal DNA. If you're looking for a plasmid or vector to propagate or assemble plasmid backbones, please see the set of [plasmid backbones](#). There are a few parts in the Registry that are only available as circular plasmids, not as parts in a plasmid backbone, you can find them here. Note that these plasmids largely do not conform to the BioBrick standard.

- ❑ Création d'un concours de BS réservé aux étudiants : le concours iGEM  
International Genetically Engineered Machine



## Synthèse de l'ADN

Gene Foundries : Quelques entreprises de synthèse d'ADN



# STRATEGIES DE CONFINEMENT

- ❑ **physique** : culture dans des endroits fermés, création de « Tchernobyl biologique ».
- ❑ **nutritionnel** : créer un organisme nécessitant des substances rares ou inconnues dans la nature pour survivre (vitamines artificielles, catalyseurs n'intervenant pas ou peu dans le vivant comme la silice);
- ❑ **évolutif** : créer un organisme mal adapté à la survie autonome dans la nature. Par exemple, une bactérie peu compétitive comparée aux formes « naturelles ». Elle ne survivrait que cultivée en condition optimale, entretenue par l'homme;
- ❑ **sémantique** : créer un organisme ne pouvant pas échanger de gènes avec les autres organismes; ou créer un nouveau code génétique avec des gènes ou des bases synthétiques ne pouvant pas se mélanger avec l'ADN des autres organismes.

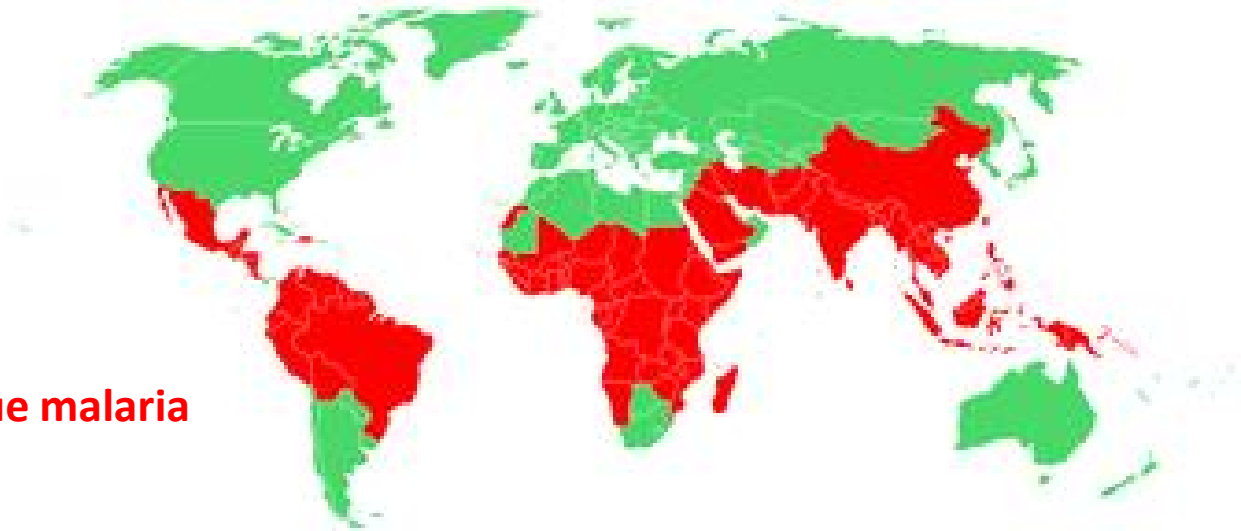
Actuellement, aucun de ces 4 « verrous » n'offre à lui seul de protection absolue. (Possibilité d'adaptation inattendues provoquées par des mutations spontanées). Renfort de la sécurité envers des phénomènes d'adaptation biologique.

 Combinaison de ces différentes stratégies



❖ Synthèse d'artémisine

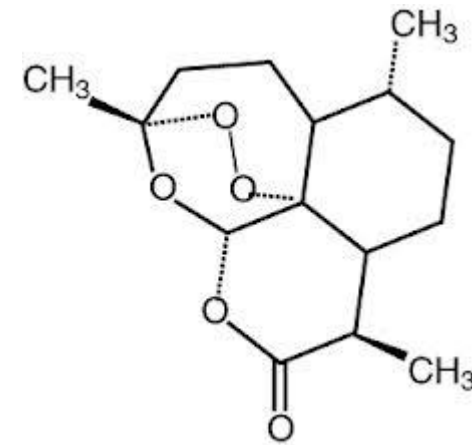
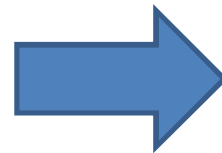
Zones à risque malaria



*Artemisia annua*



L'artémisine est une molécule antimalarique



Petite molécule à la structure complexe.

## ❖ Synthèse d'artémisine

En 2003, Jay **Keasling** (Université de Berkeley) ont réussi l'implantation de la voie de biosynthèse partielle de l'artémisinine chez *E.coli* : production d'un précurseur de l'artémisine.

⇒ Projet 42,5 M\$ (Bill and Melinda Gates fondation) :

Implantation de la voie de biosynthèse complète de l'artémisine chez la levure

Résultat : Production de l'acide artémisinique par la levure (Ro et al., 2006, Nature)

Etape de synthèse chimique :

Transformation en artémisine par catalyse et photoréaction.

❖ Synthèse d'artémisine

Industrialisation :  
Sanofi (Usine en Italie)  
en partenariat avec Amyris

50 t d'artémisine / an



- ❑ Apparition de *Plasmodium falciparum* résistants à l'artémisine (Phyo et al., 2012)
- ❑ Développement de **dérivés fluorés** de l'artémisinine (partenariat laboratoire public avec Pierre Fabre)



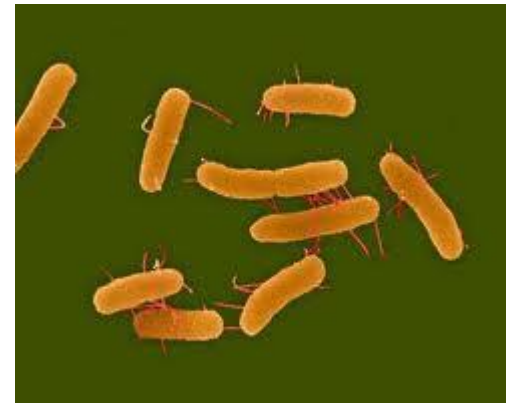
❖ Production de protéine de soie d'araignée



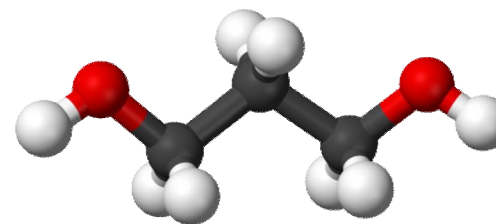
Transgénèse : Chèvre araignée, produisant les protéines de soie dans leur lait (entreprise Nexia).  
Problème de scaling-up et de coût



Production de protéines de soie d'araignée (biomatériau élastique) par *Salmonella typhimurium*, à la place de sa protéine infectieuse. (Chris Voigt)



❖ Synthèse du 1,3 propanediol (PDO)



Le 1,3 propanediol sert à former du triméthylène glycol (polyester aromatique).

Le PTT (polytriméthylène téréphtalate) est une fibre de polyester à base de triméthylène glycol.

Utilisation du PTT dans l'industrie textile (grande élasticité, facilité de teinture et meilleure résistance), dans l'industrie des détergents, ingrédients (industrie agroalimentaire et pharmaceutique)...



Metabolic Explorer (France) : lancement d'une unité de production de PDO par fermentation en Malaisie (2013) : 50000 t/an.

Utilisation d'une souche microbienne : DG1 développé par un laboratoire public.

- ❖ Synthèse du 1,3 propanediol (PDO) - Travaux de l'équipe PEEP (**P**athways **E**volution and **E**ngineering in **P**rocaryotes) du LISBP Toulouse :

### **Développement d'une stratégie d'ingénierie métabolique chez *Clostridium acetobutylicum***

Isolement et caractérisation d'une nouvelle enzyme aux propriétés originales, la **glycérol déshydratase vitamine B12 indépendante** de *Clostridium butyricum*.

La voie du 1,3 propanediol (composée de deux enzymes, la glycérol déshydratase et la PDO déshydrogénase) de *C. butyricum* a été introduite dans une souche mutante de **C. acetobutylicum (DG1)** naturellement incapable de dégrader le glycérol et de produire du PDO. L'introduction de cette voie métabolique chez *C. acetobutylicum* lui permet de convertir le glycérol en PDO (avec co-production d'acétate et de butyrate).

En **culture fed-batch** alimentée en glycérol, cette souche DG1 est capable de produire du PDO à une concentration et une productivité supérieures à celles obtenues avec la souche de *C. butyricum* naturellement productrice de PDO. Possibilité de faire de la culture continue.

❖ Développement d'un cocktail d'enzymes efficaces dégradant la biomasse végétale

Etude du LISBP (Toulouse) en partenariat avec la société Libragen (Tasse et al., 2010)

*Un des obstacles à la production rentable de biocarburants est l'accès des microorganismes producteurs aux substrats issus de la biomasse végétale*

Dans **l'intestin humain**, de nombreux micro-organismes dégradent les fibres alimentaires grâce à **des cocktails d'enzymes parfaitement adaptés**. L'équipe du LISBP cherche à reproduire ces cocktails pour dégrader de la même manière la biomasse végétale

Etude à partir d'excréments d'un individu végétarien contenant des milliards de microorganismes.

Or, il est aujourd'hui impossible de cultiver en laboratoire toutes ces espèces, leurs gènes ont été **identifiés** par « metagénomique » et utilisés pour **transformer *E. coli***.

Puis **criblage automatisé de l'activité des enzymes** produites par ces bactéries modifiées, pour isoler les plus efficaces.

Pistes : Microorganisme du tube digestif des termites, de celui des pandas...

## ❖ Biocapteurs

- ❑ Bactérie témoin de condition d'expérience

<http://www.igem-paris-saclay.u-psud.fr/>

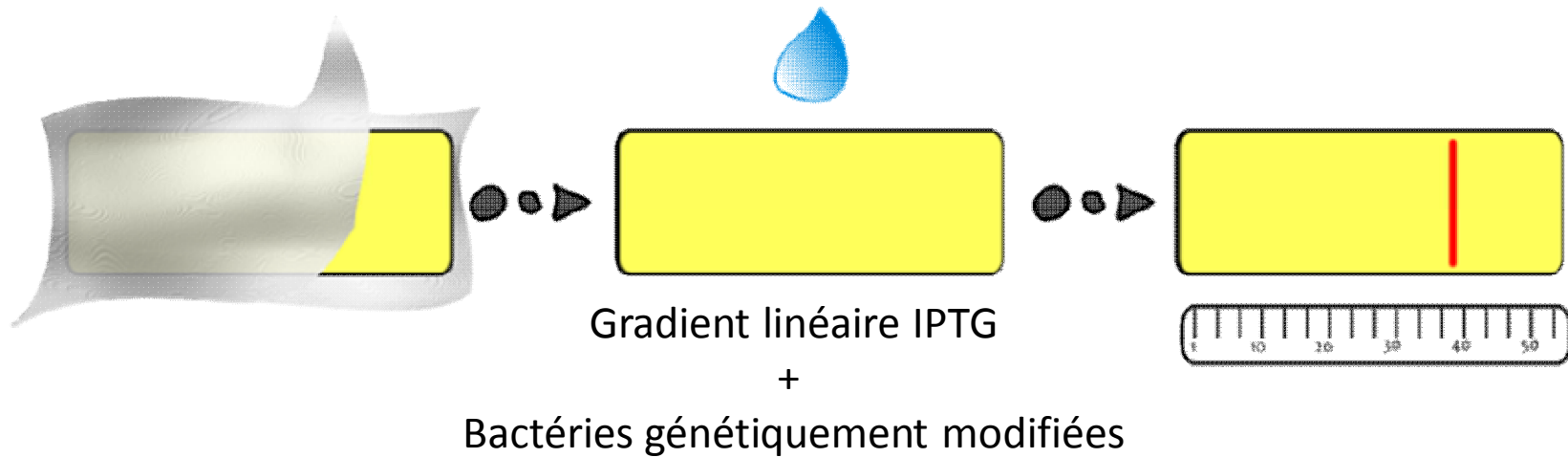
- ❑ Bactérie qui induit un changement de couleur de l'eau lorsque celle-ci contient un parasite. Evaluation de la potabilité de sources d'eau éloignées, méthode facile à mettre en œuvre : résultat clairement visible, changement de couleur

- ❑ *E. Chromi* : bactérie qui change de couleur si elle est en contact avec une bactérie pathogène, un virus, un polluant



❑ Le projet Mercurio-coli

<http://2011.igem.org/Team:Grenoble/Projet/Intro>



Si IPTG < Mercure, les bactéries libèrent un colorant rouge

# STRATEGIE DE RECHERCHE

- ❑ Dans le monde, environ 1000 chercheurs recensés « biologie synthétique »  
Publications scientifiques : US 68% Europe 17% (dont Allemagne 8%)

- ❑ **En France,**

- Institut de biologie synthétique et systémique à l'Université d'Evry (2010)
- Genoscope d'Evry
- Le TWB (Toulouse White Biotechnology) : un démonstrateur préindustriel pour les biotechnologies industrielles (P. Monsan)  
7 entreprises dédiées dont un département de Total

- ❑ **Les Etats Unis loin devant ...**

Dotation publique : 430 M€/an depuis 2005 du gouvernement US (1 Md de \$ si on ajoute les investissements privés),  
Institut de biologie synthétique et systémique Genoscope 1 M€/an depuis 2010 (Europe 30 M€ financé dans le cadre du 6<sup>ème</sup> PCRD).

# DEVELOPPEMENT ECONOMIQUE

GlobalBioenergies (Isobutène, France , 2008)

Deinove



Amabiotics (biorémédiation, France, 2010)

Metabolic explorer (L-méthionine, acide glycolique, butanol, PDO (1,3 propanediol), MPG (1,2 propanediol), France, 1999)

GTPtechnology (expression de protéines recombinantes, France, 2000)  
(Microme)

Synthetic Genomics Inc. (fondé par JCV + Hamilton Smith, USA, 2005)

Metabolix (production de PHA, USA, 1992)

Genecor (DuPont Industrial Biosciences, Danemark, 1982)

Iogen (Canada, 70s)

Solazyme (microalgue, USA, 2003)

Fermentalg



USA Genencor, LS9, Gevo et Amyris ...



GLOBAL BIOENERGIES





## DEVELOPPEMENT ECONOMIQUE

- ❑ De nombreux grands groupes investissent dans les laboratoires de recherche et développe des partenariat avec les start-up de type «BioSyn Tech » (= entreprise qui construisent des microorganismes d'intérêts)

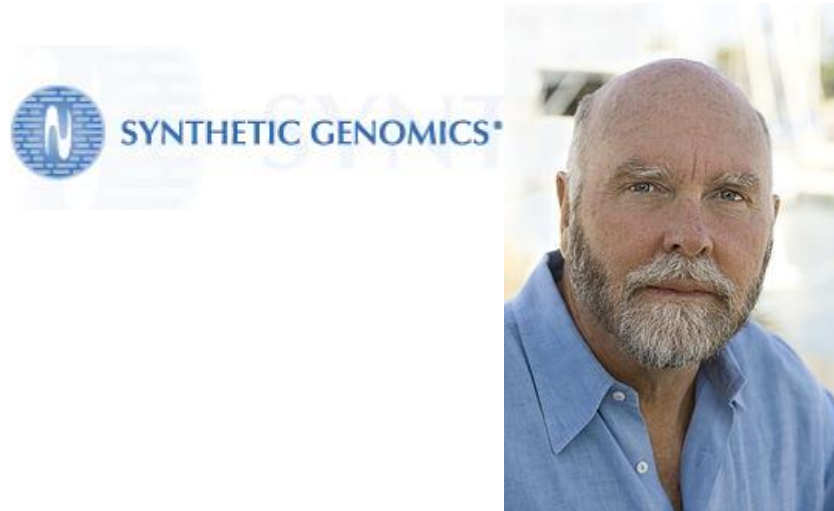
=> La quantité d'argent investie est énorme !



**Cargill**  
**BP**  
**Exxon**  
**DuPont**  
**Pfizer**  
**Virgin ...**

- ❑ Business club, investment club, strategic club fleurissent ...

# DEVELOPPEMENT ECONOMIQUE



- ❑ Craig Venter a signé deux accords importants (avec Exxon et BP) pour concevoir des micro-organismes censés fabriquer en continu des hydrocarbures (éthanol et butanol).

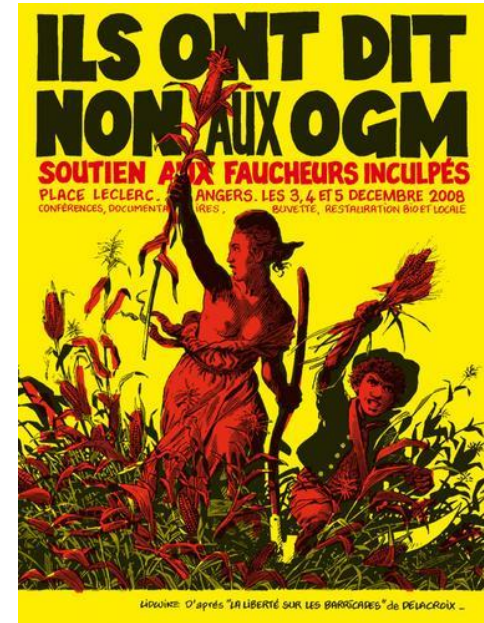
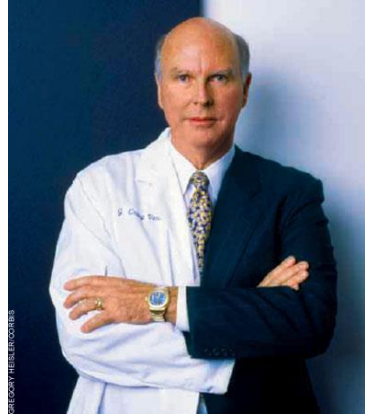
**ExxonMobil**

Taking on the world's toughest energy challenges.™



- ❑ Partenariat Total / Amyris (développement du Biofene développement de biodiesel)

## Et l'éthique ?





## Et les risques ?

Du bioterrorisme au bioratage ...

Risque lié au « garage developments »



Risque radioactif décroît après l'explosion nucléaire

Risque biologique **croît** après dissémination d'un gène dans l'environnement

Quel principe de précaution ?

➔ Stratégies de confinements



1958