

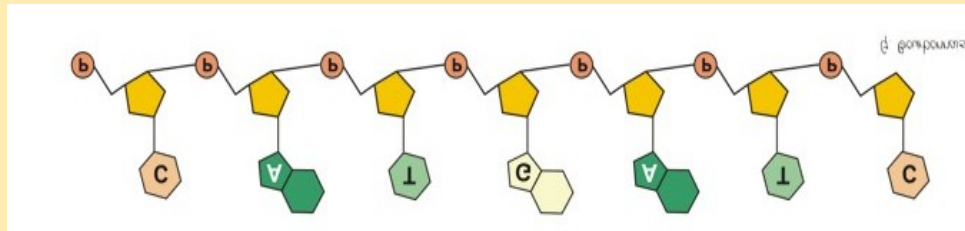
L'ADN, entre identité et différence

2. Les différences



Rappels

1 ADN et ARN



ADN : sucre Désoxyribose
base T (Thymine)

ARN : sucre Ribose
base U (Uracyle)

- ribose – phosphate – ribose – phosphate – ribose –
phosphate -
↓ ↓ ↓
base 1 base 2 base 3

TCACCT GAGAGAGAGCAT

Ces 4 lettres représentent les « bases »

**L'information se loge dans la succession des
4 bases A, T, C et G**

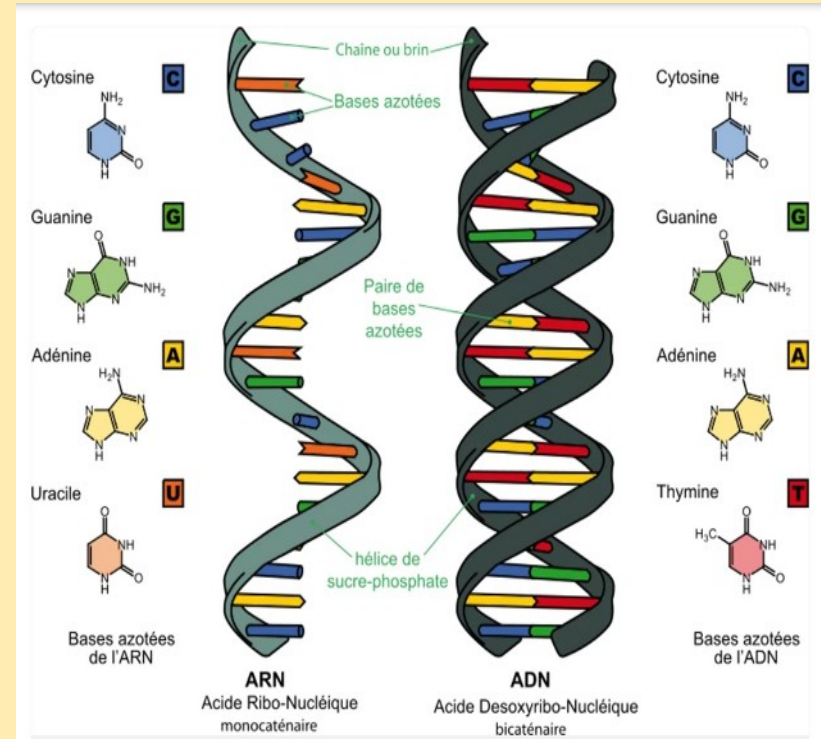
2 L'ADN : la double hélice

Les bases se groupent deux par deux :
« paires de bases » pb ou bp

A est toujours lié à T

C est toujours lié à G

Si l'on connaît un brin, on connaît son complémentaire

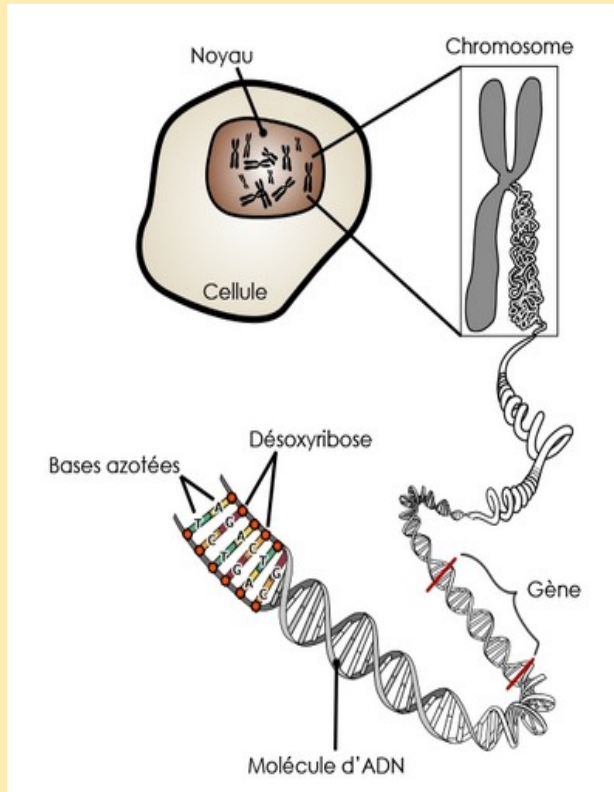


L'ARN est différent :

→ simple brin, en hélice aussi

→ T est remplacée par U, Uracile

3 L'ADN, les gènes et les chromosomes



ADN

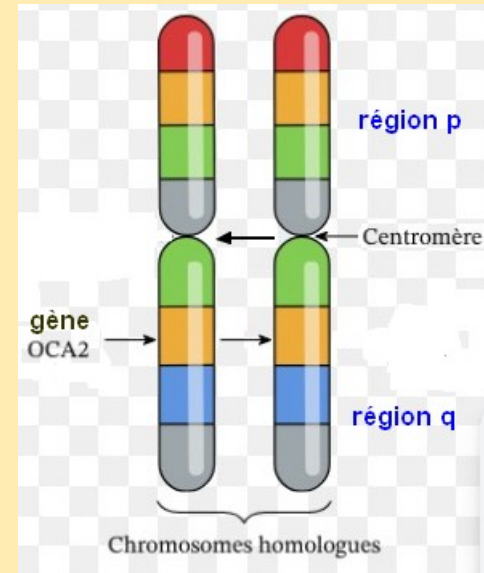
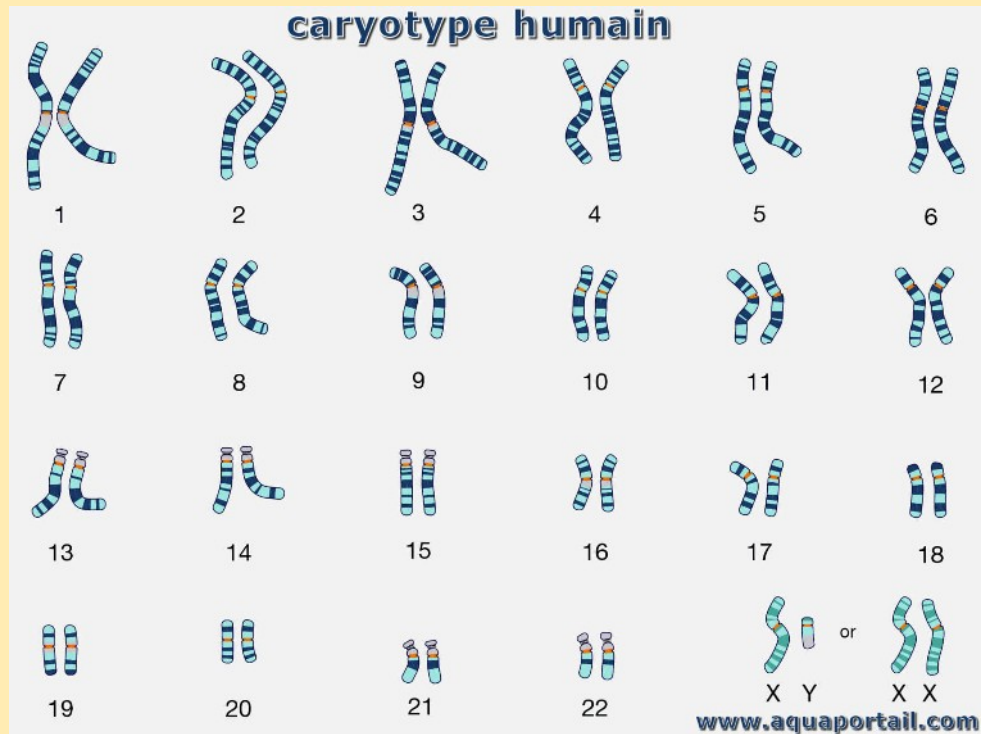
L'hélice à 2 brins

Gènes

*Parties de l'ADN
porteuses d'information*

Chromosome :
ADN embobiné

4 Les chromosomes sont au nombre de 2×23



Les 2 chromosomes portent les mêmes gènes sur le même locus

5 Synthèse des protéines

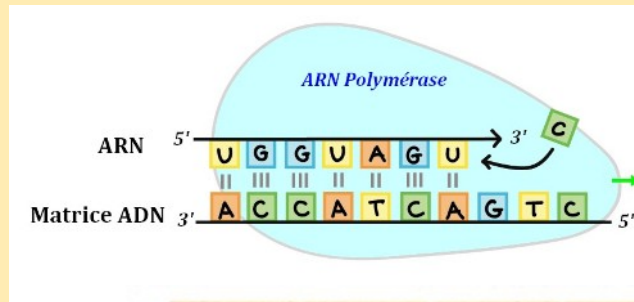
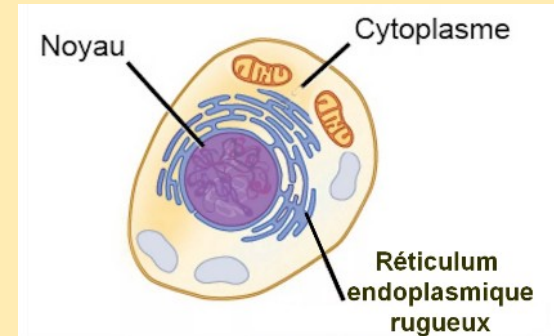
TRANSCRIPTION : Le codage est transmis par l'ARN

L'ADN dispose des bases A T C G le long de sa chaîne

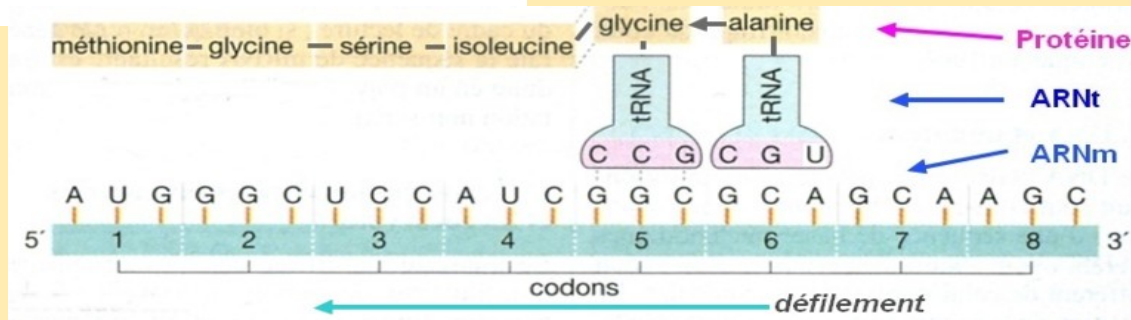
a) L'ARN messager **copie ce code** (cet ordre) sur l'un des brins avec un léger changement :

→ La base T (Thymine) est remplacée par la base U (Uracile)

b) après quelques modifications l'ARNm quitte le noyau pour le RER où se passe la synthèse des protéines



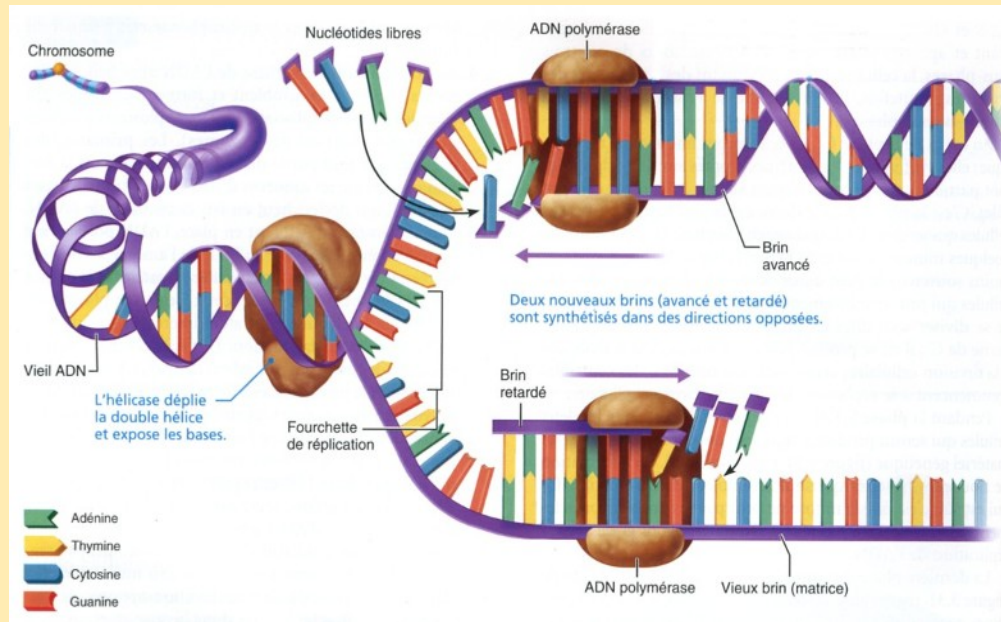
TRADUCTION : Le RER fabrique les protéines



Se construit brique par brique
 Porte un AA précis et reconnaît « son » codon
 Porte l'information

6 Reproduction

RÉPLICATION : Démultiplication de l'ADN

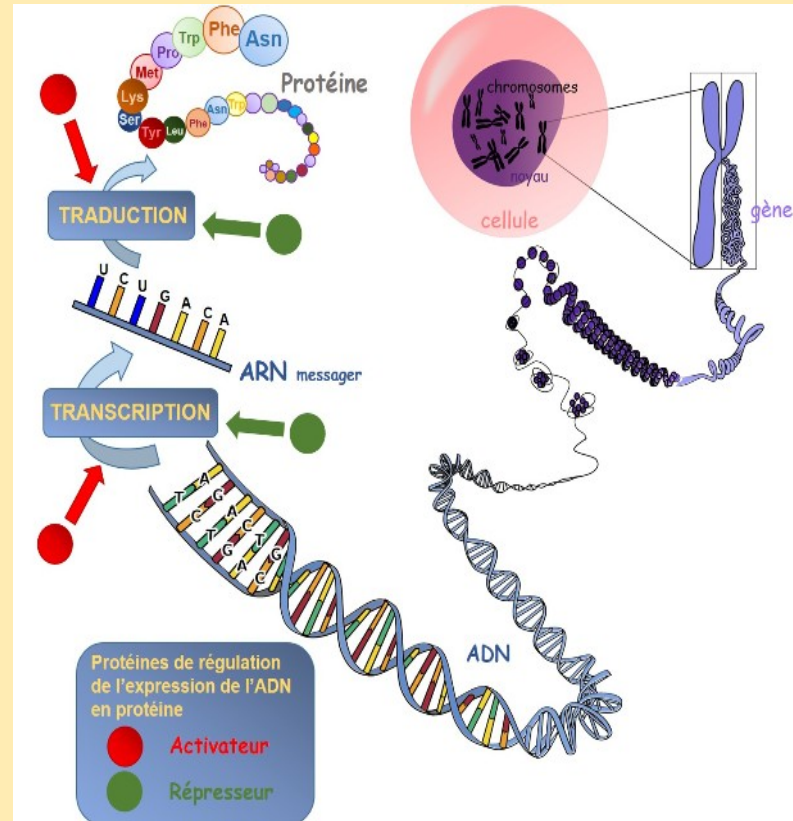


7 Régulation : activation ou répression des gènes



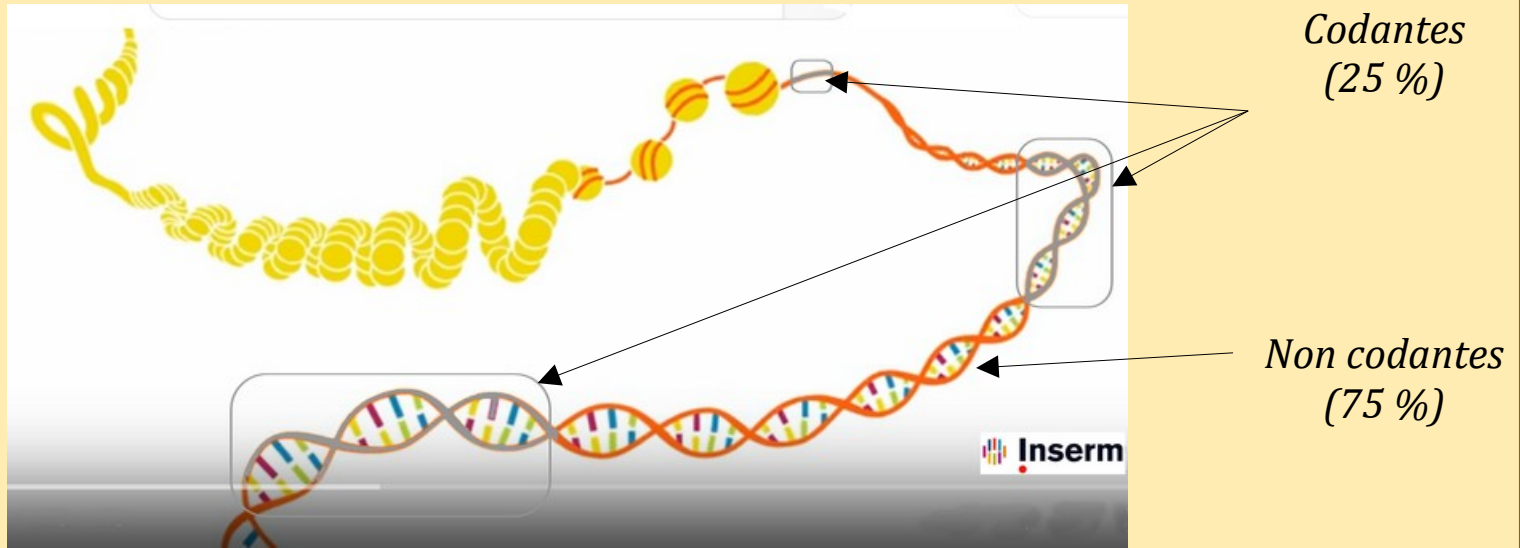
L'activité des gènes est **contrôlée** par des substances qui **activent** ou **répriment** l'expression du gène au moment de la transcription ou de la traduction

« Les gènes ne sont donc qu'une potentialité, une impulsion »



ADN : où se trouvent les différences ?

l'ADN est constituée de différentes zones, codantes ou non



Les parties variables se trouvent dans l'ADN codant comme non codant

1 Les gènes eux-mêmes peuvent coder pour des différences

Gène groupe rhésus : (chr. 1)
Gène groupe sanguin (ABO) : (chr. 9)
Couleur des yeux : *EYCL1, EYCL2, EYCL3* (chr. 13, 15...)
AMELY ou *AMELX* Amélogénine
protéine de l'émail dentaire



2 Mais les différences peuvent provenir d'accidents, de mutations

Nous sommes identiques génétiquement à 99,9 % mais \Rightarrow des différences :

- Mutations lors de la synthèse de l'ADN du spermatozoïde et de l'ovule, (30 sur les 3MMbp)
- Les SNP, trace de mutations très anciennes produisent de très nombreuses variations individuelles
- Les STR sont beaucoup utilisées en identification humaine (FNAEG)

Les deux grands types de différences entre individus

1 Les séquences répétées : STR (short Tandem Repeat)

Ce sont des chaînes courtes répétées à la suite (en *tandem*) dans les séquences non-codantes :

CGTAGCAGCTTCAAGAAGAAGAAGTGCATCCATCATCGGTG
CGTAGCAGCTTCAAGAAGAAGTGCATCCATCATCGGTG

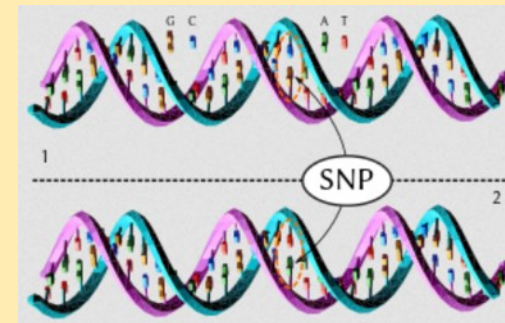
Il y a ici une **chaîne courte** (3 bp A-A-G) répétée 5 fois ou 3 fois selon les individus

2 La variation d'une seule paire de bases : SNP (Single Nucleotide Polymorphism)

Variations fréquentes : 1 sur 1200 environ affectant n'importe quel endroit du génôme

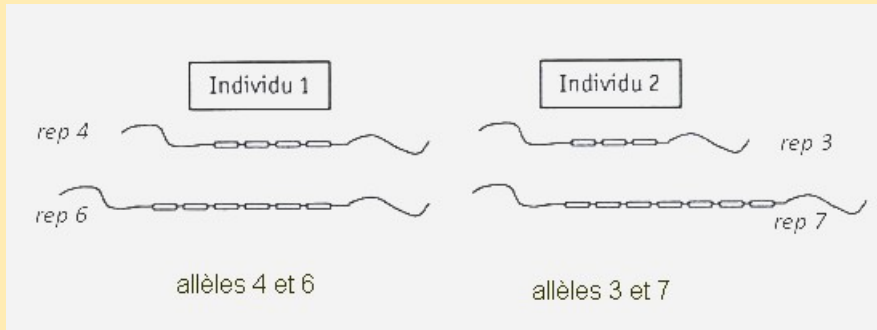
CGTAGCAGCTTCTGCA TCCATCA TCGGTG
CGTAGCAGCTTCTGCT TCCATCA CCGGTG

Ex : sur le chromosome 21 (le plus petit) on a 131 740 SNP pour 47 Mbp



Caractérisation de ces différences. Exemple d'un seul STR

Isoler les allèles



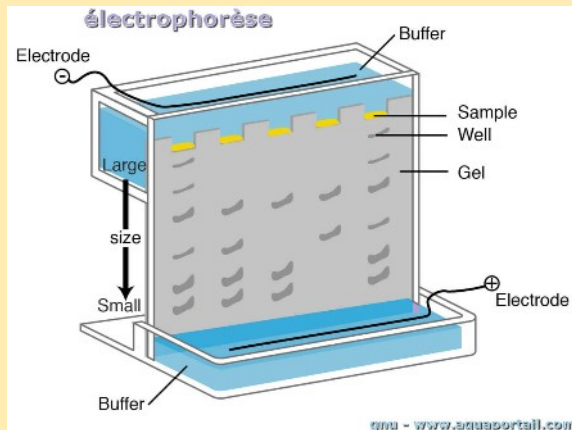
Ces 2 individus comportent 2 STR chacun, de longueur différente :

Individu 1 : 4 et 6 répétitions

Individu 2 : 3 et 7 répétitions

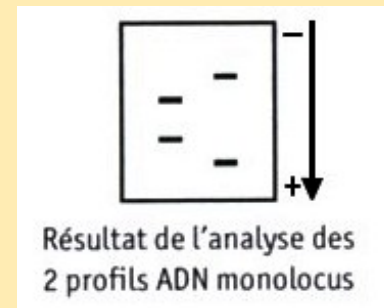
On isole les allèles correspondants et on obtient 4 séquences de longueur différente

Différencier : l'électrophorèse



Les séquences isolées sont entraînées par un champ électrique (ddp env. 100 V)

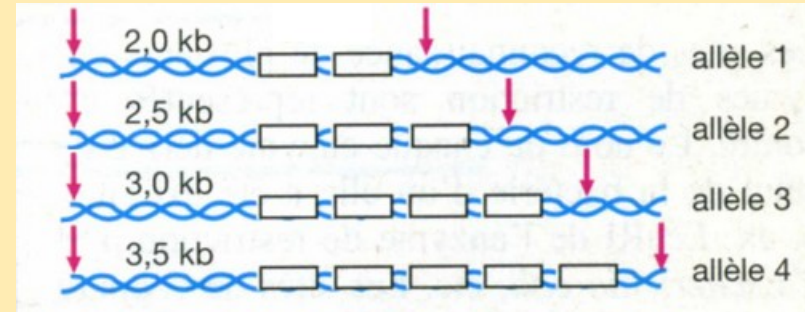
Les plus petites migrent plus vite dans le gel



Allèles multiples : meilleure différenciation

On imagine une séquence répétée qui peut présenter de 2 à 5 répétitions selon les individus (**4 allèles différents**).

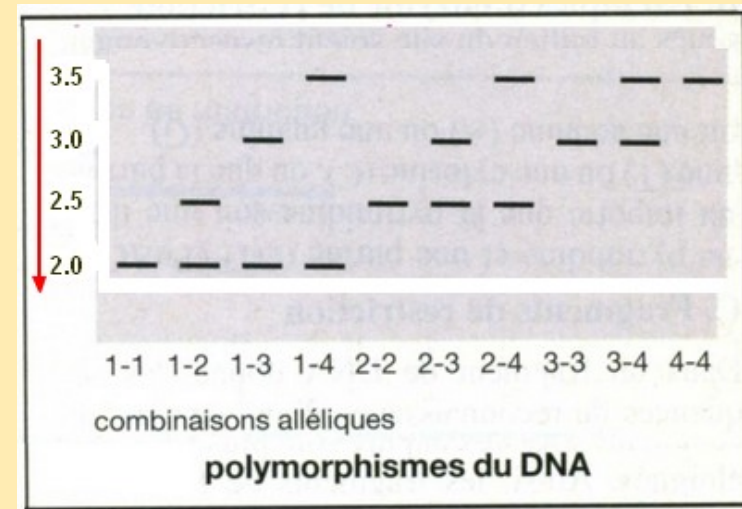
La personne reçoit un allèle de son père et un de sa mère.



En tout cela fait 10 possibilités différentes selon les allèles reçues des 2 parents:

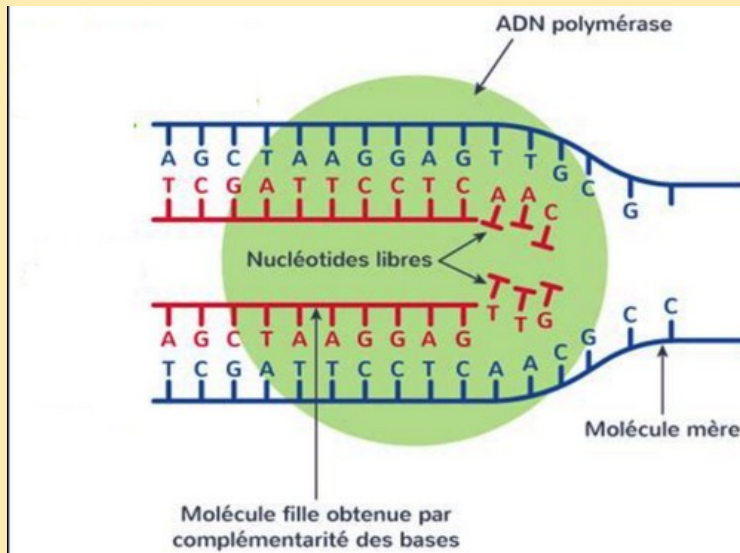
1-1 2-2 3-3 4-4
1-2 1-3 1-4
2-3 2-4
3-4

Si chaque allèle a la même probabilité d'apparition, le pouvoir de discrimination est de **90 %**

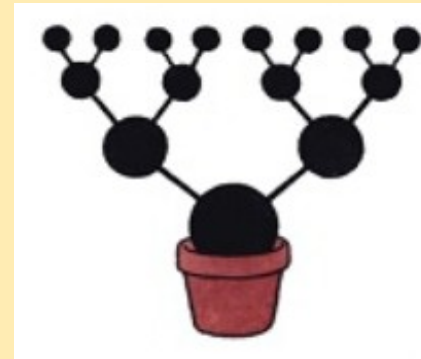


Que faire quand on a très peu d'ADN ?

La réaction de polymérisation en chaîne consiste à fabriquer deux brins d'ADN (ou d'ARN) à partir d'un seul, puis à nouveau deux à partir de chaque nouveau brin etc.



On multiplie de façon exponentielle le nombre de chaînes d'ADN par clonage.
« Polymerase Chain Reaction »



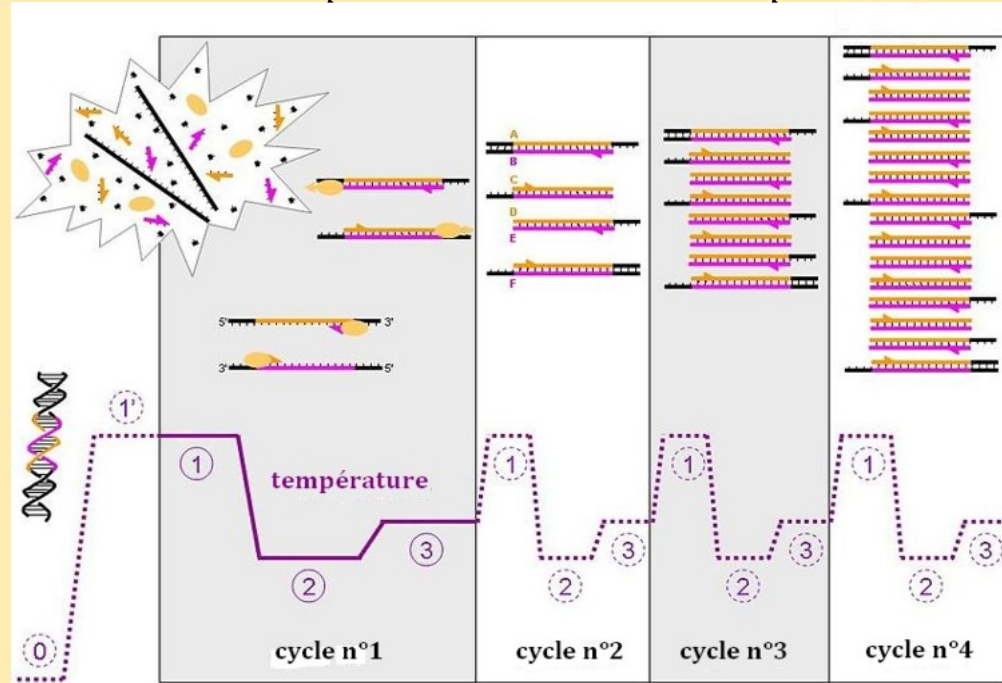
On a réalisé une amplification par PCR

La machine est simple

Machine PCR.



Les 3 séquences successives et leur répétition



① : Dénaturation

② : Amorçage

③ Élongation



La généalogie par ADN

Le principe

Il y a trois types de recherches :

- sur l'ADN autosomal (les 22 premières paires + le X)

→ origines ethniques

→ cousins génétiques jusqu'à (potentiellement) la 8^e génération.

- sur l'ADN du chromosome Y

lignée patrilinéaire

- marqueurs STR (de 25 à 67 différents)

- ou SNP pour les ancêtres lointains

Exemple : le STR DYS 393 dont la séquence est AGAT
(répétition de 9 à 17 fois)

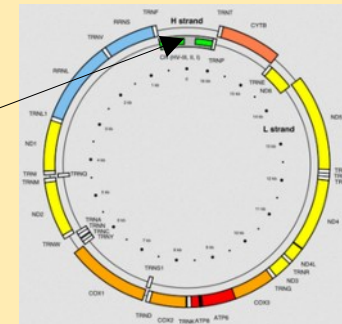
- sur l'ADN mitochondrial

lignée matrilineaire.

- 16 569 paires de bases

- 37 gènes

- mutations sur la région HVR-1 et HVR-2
ou des SNP



La généalogie par examen ADN est-elle fiable ?

Nous partageons exactement 50 % d'ADN avec chacun de nos parents

- 50% = parent ou enfant
- 25% = grand-parent, petit-enfant, oncle ou tante, nièce ou neveu, demi-frère
- 12.5% = cousin germain, arrière-grand-parent, arrière-petit-enfant, grand-oncle ou tante, petit-neveu ou nièce, demi-oncle ou tante, demi-neveu ou -nièce

*Ces proportions sont approximatives :
on peut partager 10 % d'ADN avec son grand-père paternel et 14 % avec son aïeul maternel.*

Les tests montrent l'existence d'un lien de parenté mais ne disent pas sa nature.

On découvre donc des « amis génétiques » et non des parents !



Probabilités géographiques d'origine

Mais il y a d'autres méthodes

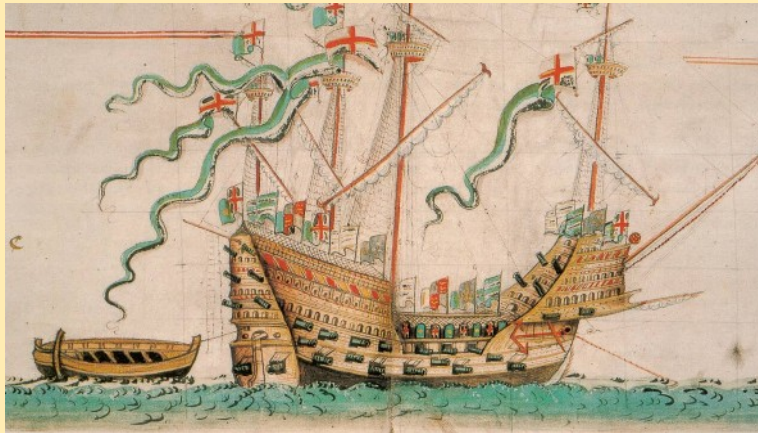
Le naufrage du Mary Rose

Navire anglais qui a coulé en 1545

Des analyses ont montré qu'au moins trois des marins n'étaient pas originaires de Grande-Bretagne, mais probablement d'Italie, d'Espagne et d'Afrique du Nord

Quelques examens

- *examen des artefacts (poteries, armes...)*
- *examen des dents (le taux d'oxygène est plus élevé en pays chaud)*
- *analyse isotopique → identifie le régime alimentaire*
- *examen de l'ADN mitochondrial → origine berbère (Henry)*
- *portrait-robot et appel à la population*



Le navire reposait par 15 m de fond près de l'île de Wright

Redécouvert en 1971

Renfloué en 1982

Exposé à Portsmouth.

En identification judiciaire : les STR

Un marqueur d'identification : STR D3S1358

Marqueur D3S1368 :

- D : situé sur l'ADN
- 3 : sur le chromosome 3 [3q21.31]
- S : séquence unique et non dispersée
- 1358 : numéro d'ordre

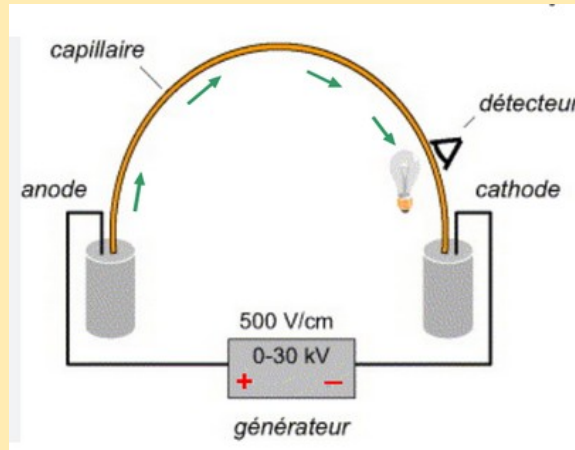
La séquence comporte 4 maillons (nucléotides) (paires de base) qui se répètent entre 11 et 19 fois selon les individus

motif (TCTA) n $11 < n < 19$

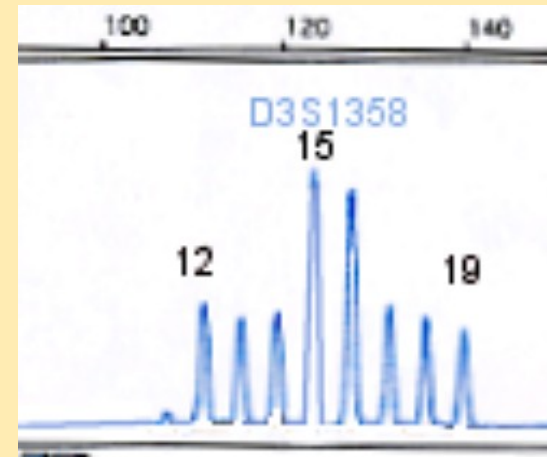
Pouvoir discriminant 92 %

Analyse : par électrophorèse capillaire

Beaucoup plus rapide que sur gel



Électrophorégramme



Le kit multiplex d'identification

En pratique on opère sur **plusieurs STR** simultanément (Multiplex) pour augmenter la probabilité d'identité.

Le résultat est comparé à une gamme étalon connue.

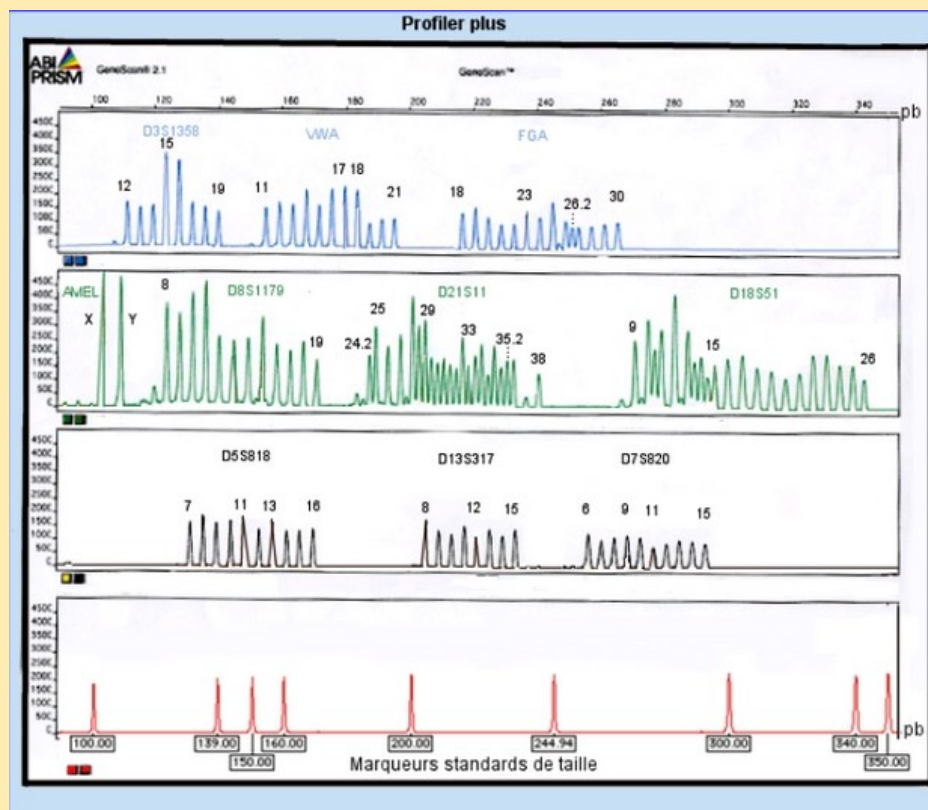
Chaque étalon comporte l'ensemble des allèles possibles

Ici : 9 étalons présentant de 8 à 18 répétitions possibles

+

1 séquence X-Y

Les séquences sont colorées (fluorescence) pour les distinguer

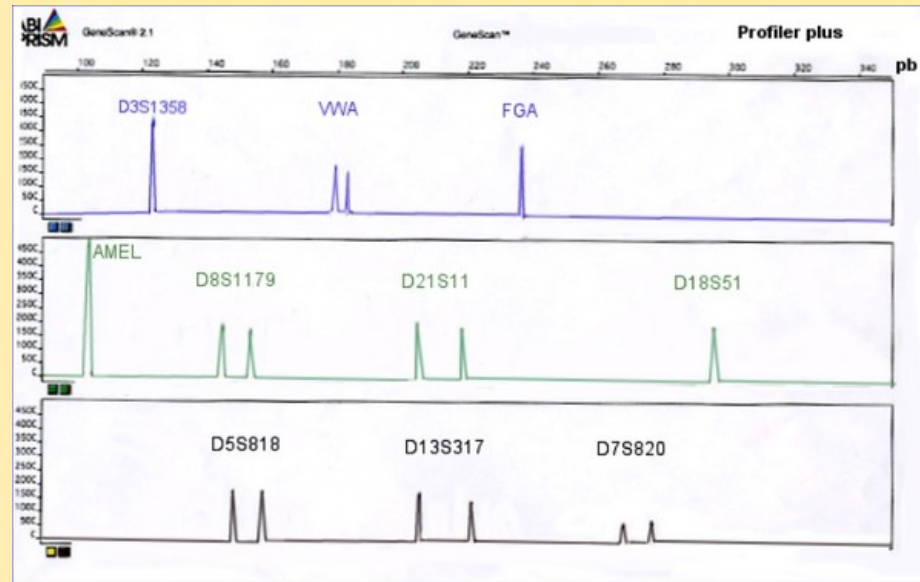


L'échantillon à identifier

Pour chaque séquence l'ADN du suspect montre la présence de deux allèles (papa et maman), éventuellement identiques

L'électrophorégramme ne montre pas directement le nombre de répétitions pour chaque allèle, mais seulement leur position.

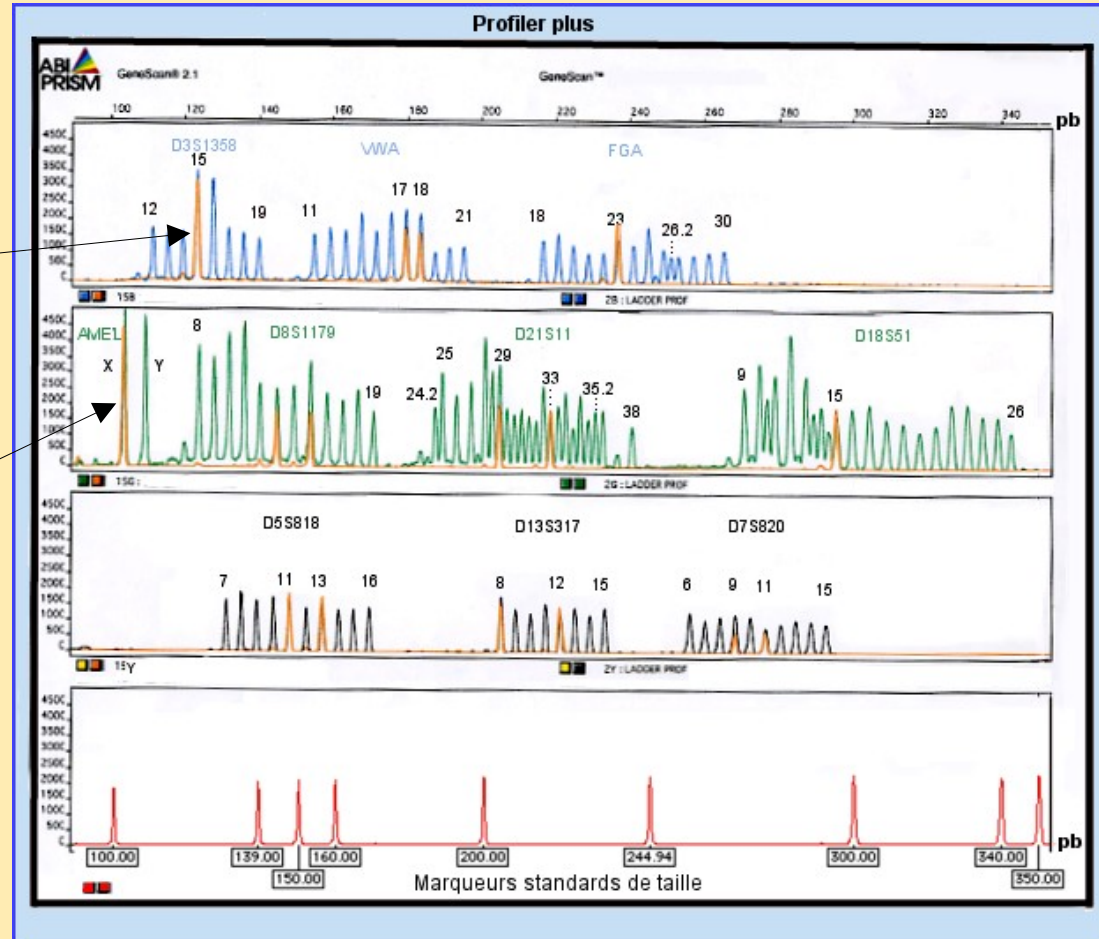
Le nombre de répétitions est déterminé par comparaison avec les marqueurs



La superposition

Pour la première séquence (D3S1358) il y a chez le suspect 15 répétitions

*Et le suspect est une femme
1 seul pic AMEL*



Quel pouvoir de discrimination avec un kit multiplex ?

Ex : Identification avec un kit SGM Plus

D3S1358

Fréquence de l'allèle 15/16 : 1 personne sur 8

WMA

Fréquence de l'allèle 15/16 : 1 personne sur 27

Combinaison des deux : 1 personne sur 200

*Combinaison des 10 STR : **1 pers. sur 1000 MM***

Et si j'analyse des SNP ?

Pour avoir le même pouvoir discriminant qu'un kit STR il faut analyser beaucoup plus de SNP.

Mais il y en a beaucoup et leur analyse est rapide

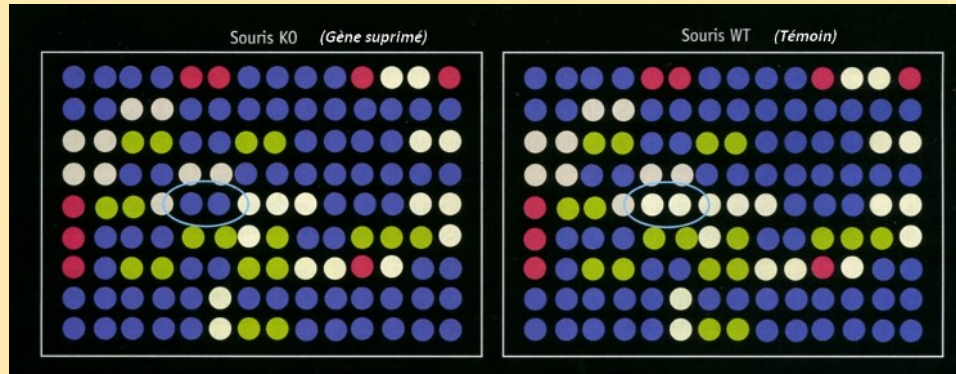
En médecine, deux techniques incontournables

La puce à ADN

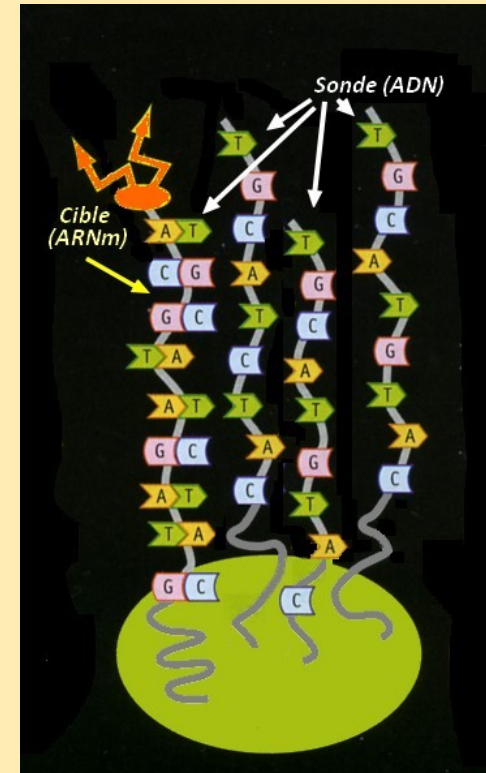
Problème : il s'agit de trouver quels gènes sont exprimés, c'est-à-dire quels brins d'ARNm sont capables de se lier à l'ADN (s'hybrider).

On dépose sur une lame des centaines de brins d'ADN (**sondes**) sur lesquels viennent s'hybrider ou non des morceaux d'ARNm (**cible**).

Une molécule fixée sur la cible devient fluorescente en cas d'hybridation.



Résultat d'analyse : souris « KO » vs souris « WT »



Une puce

• Les ciseaux moléculaires

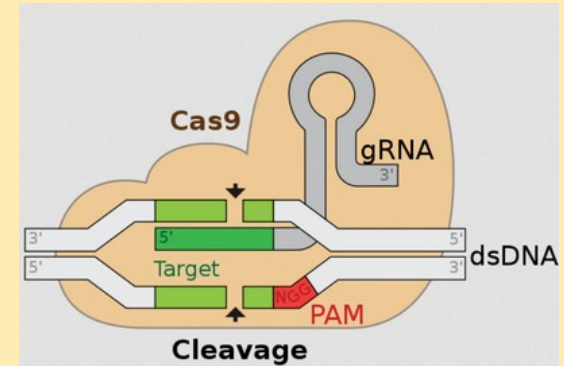
2020 : prix Nobel de chimie pour **Emmanuelle Charpentier** et **Jennifer Doudna** pour la technique de « CRISPR/Cas9 » ou « ciseaux moléculaires »

- **en thérapie génique** remplacement de gènes défectueux chez les patients pour guérir certaines maladies.

- **en oncologie**, recherche de nouvelles cibles thérapeutiques. On découpe des gènes dans certaines cellules pour retracer les étapes de la croissance des tumeurs.

Auparavant, un découpage précis dans l'ADN réussissait dans une cellule sur un million.

Avec le CRISPR, on réussit souvent dans une cellule sur 100



Clustered Regularly Interspaced Short Palindromic Repeats - Crispr Associated Protein 9



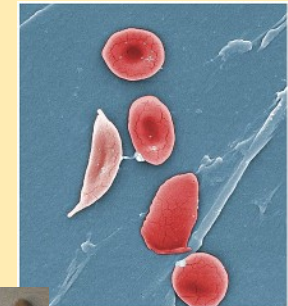
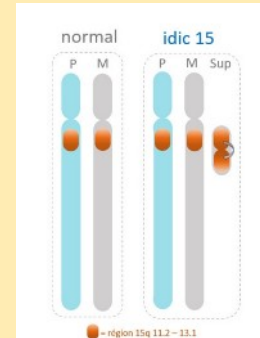
Gènes associés à des maladies

IDIC 15 (Isodicentrique Duplication Inversée, 15q11-13) : anomalie génétique la plus fréquente de **l'autisme** résultant d'une duplication inversée du chromosome 15

BRCA 1 (Breast Cancer, 13q12-13) et **BRCA 2** (17q12-21) : risque accru de **cancer du sein** résultant d'une mutation du gène, qui touche 2 femmes sur 1000. Ces deux gènes participent à la réparation des lésions que l'ADN subit régulièrement.

HBB (Hemoglobin Subunit beta, 11p15-5) donne la drépanocytose (**anémie falciforme**) par remplacement dans l'hémoglobine du glutamate en position 6 par une valine. Risque d'anémie, d'AVC, d'infection bactérienne.

ARN Xist c'est un ARN qui joue un rôle en épigénétique. Il inhibe un chromosome X complet chez les femelles de mammifères. Aléatoire, et responsable en particulier du pelage en « écailles de tortue » des chattes



SNP associés à des maladies

Plus de 13000 variations d'un seul nucléotide ont été associés à des pathologies à ce jour

SNP synonymes : la substitution d'un codon par un autre engendre le même AA du fait de la redondance du code génétique.

Ex : CUU ↔ CUC

SNP non synonymes : ils engendrent un autre AA

Ex : GAA ↔ GUA*

SNP non-sens : provoquent l'arrêt de la traduction

Ex : UUG ↔ UAG

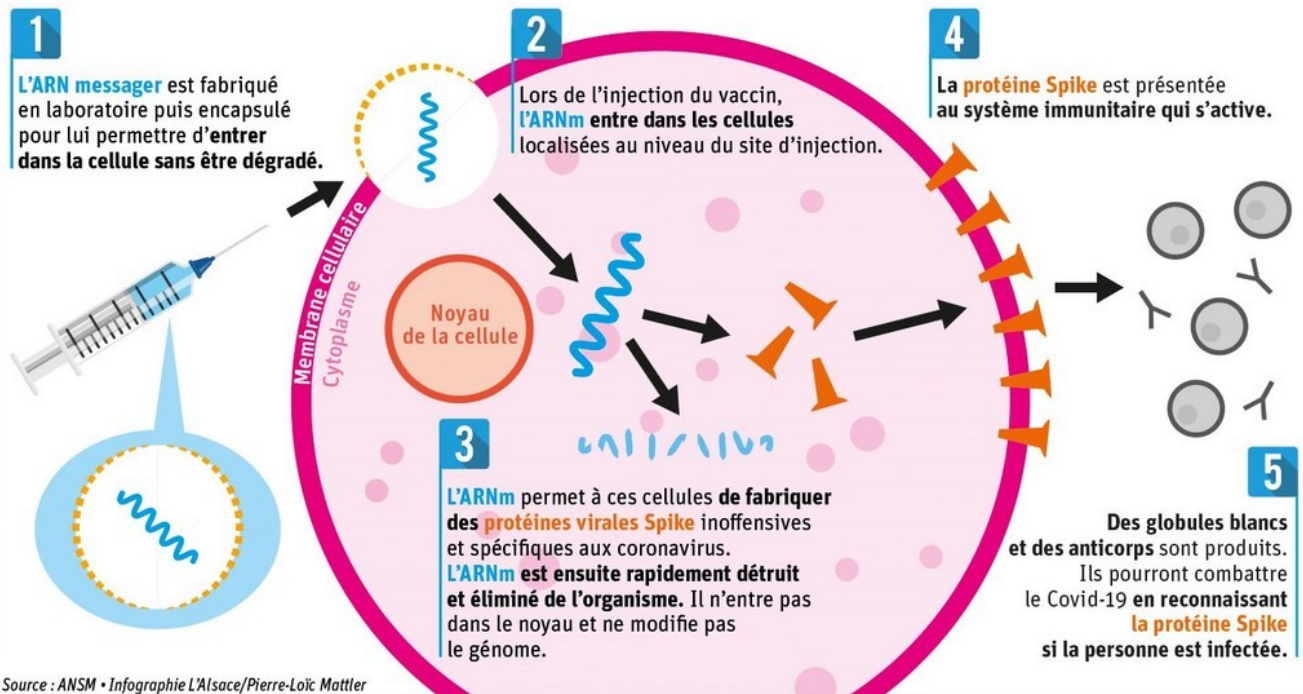
	U		C		A		G		
U	UUU	phénylalanine	UCU	sérine	UAU	tyrosine	UGU	cystéine	U
	UUC		UCC		UAC		UGC		C
	UUA	leucine	UCA		UAA	stop	UGA	stop	A
	UUG		UCG		UAG		UGG	tryptophane	G
C	CUU	leucine	CCU	proline	CAU	histidine	CGU	arginine	U
	CUC		CCC		CAC		CGC		C
	CUA		CCA		CAA	CGA	A		
	CUG		CCG		CAG	CGG	G		
A	AUU	isoleucine	ACU	thréonine	AAU	asparagine	AGU	sérine	U
	AUC		ACC		AAC		AGC		C
	AUA	méthionine	ACA		AAA	lysine	AGA	arginine	A
	AUG		ACG		AAG		AGG		G
G	GUU	valine	GCU	alanine	GAU	acide aspartique	GGU	glycine	U
	GUC		GCC		GAC		GGC		C
	GUA		GCA		GAA	GGA	A		
	GUG		GCG		GAG	GGG	G		

* Dans l'hémoglobine, remplacement du Glutamate en position 6 par la Valine ⇒ Drépanocytose

Les vaccins à ARNm

Comment fonctionne un vaccin à ARN messager ?

L'ARNm (acide ribonucléique messager) permet de transmettre l'information codée d'un gène du noyau vers la cellule.



Dans l'élevage

« 3R » Rencontres Recherches Ruminants, décembre 2022 :

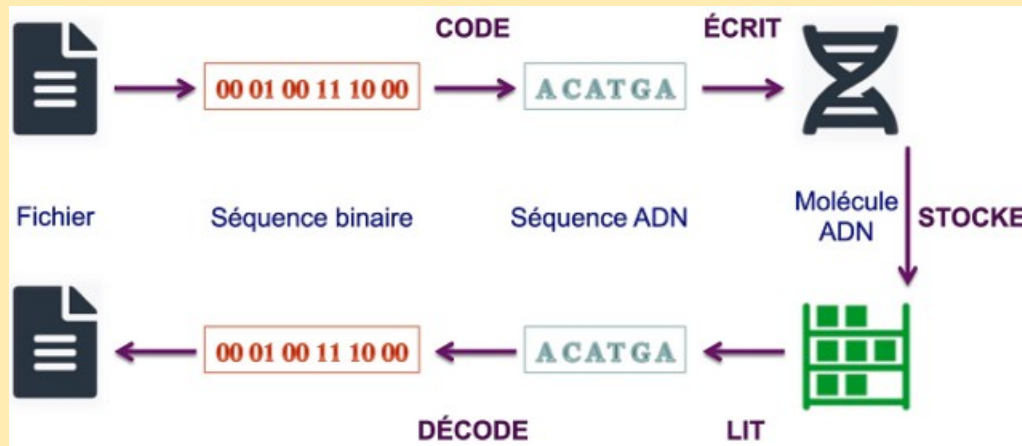
« Utiliser des index génomiques pour sélectionner les femelles de renouvellement parmi les femelles issues de croisements laitiers ? Ce sera bientôt possible grâce à l'INRAE qui met au point une méthode permettant de tracer l'origine raciale des allèles impliqués dans les caractères de production »

*Comparaison entre le génotype d'animaux croisés
(5238 génotypes)
et celui de races pures
Montbéliarde (20000),
Holstein (22265)
Danoise rouge (6866 génotypes)
en utilisant la puce à ADN EuroGMD comptant 45 000
marqueurs SNP.*



En informatique

Codage d'informations sur ADN artificiel



Le programme **MolecularXiv** vise à inventer de nouveaux dispositifs de stockage de données sur ADN et polymères artificiels.



Microcapsules d'ADN synthétique

