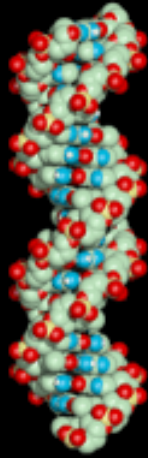


# Cours de Génétique L3



François Bretagnolle

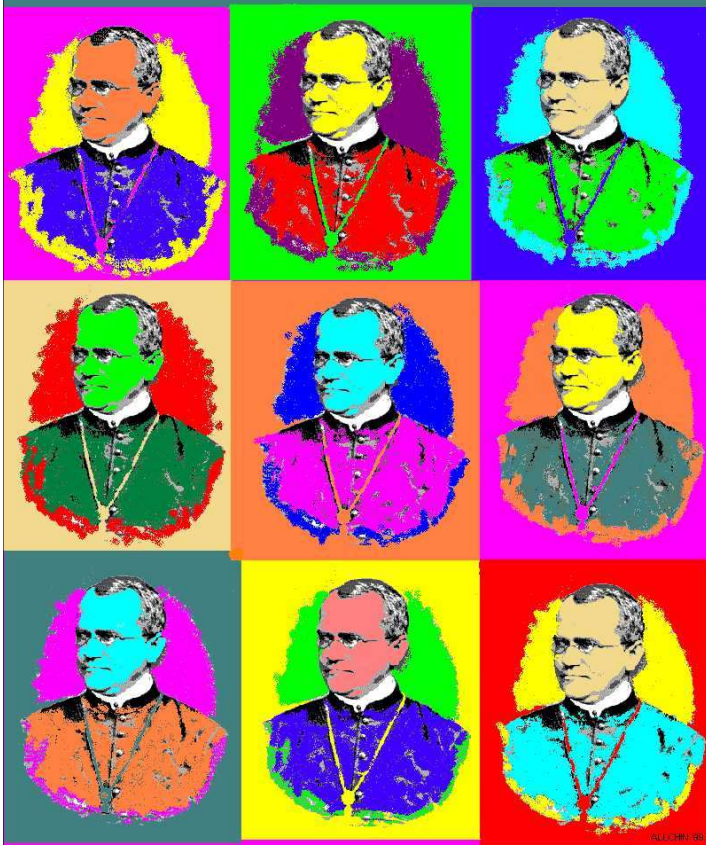
Email: [francois.bretagnolle@u-bourgogne.fr](mailto:francois.bretagnolle@u-bourgogne.fr)

# Plan

- Introduction à la génétique: Quelques rappels
- 1 - Évolution des génomes: Les mutations
- 2 - Quelques notions de biotechnologies et outils de génétique moléculaire
- 3 - Génomique structurale: Structure du génome eucaryote
- 4 - Quelques notions de génomique évolutive

# Introduction: Quelques rappels

# Mendel: le concept de gènes et les lois de transmission des caractères



Les trois lois de Mendel sur la transmission des caractères au fil des générations

# La révolution mendélienne

Caractères héréditaires = Entités discrètes  
transmises indépendamment les unes des autres

Êtres vivants : mosaïques de caractères à  
transmission indépendante

Moins d'un siècle sépare Mendel et  
l'ADN

# Génétique

"Science qui étudie les mystères de l'hérédité et les variations"

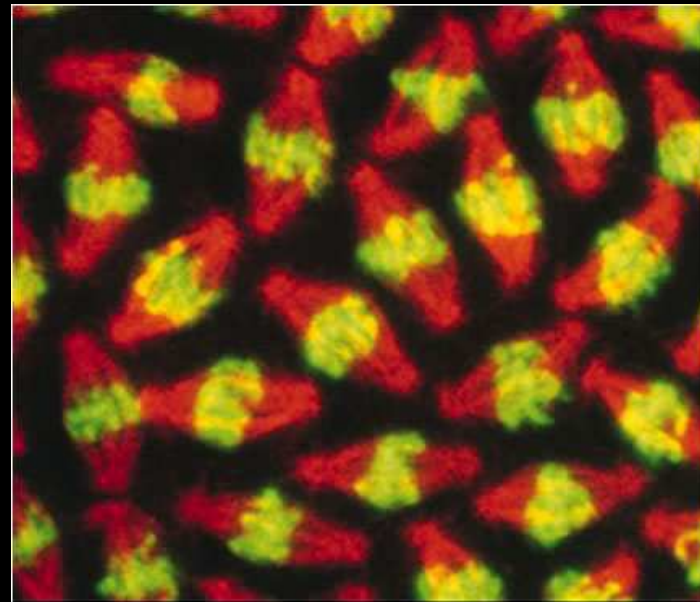
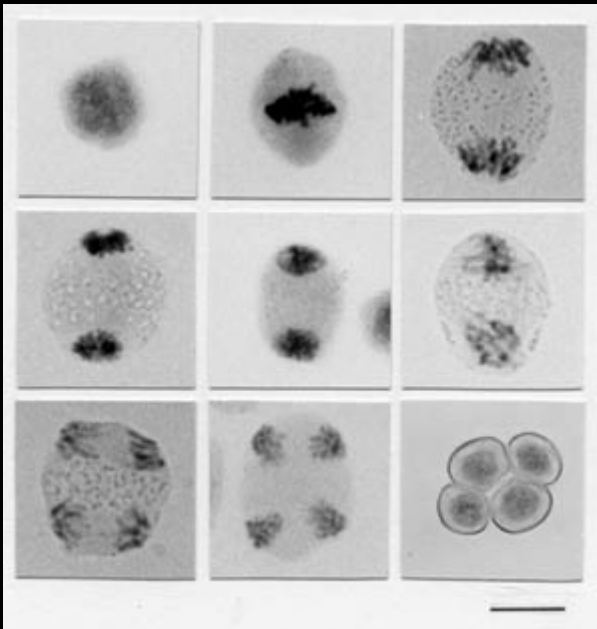
(William Bateson, 1906)

Branche de la biologie qui s'intéresse à l'hérédité et à la variation héréditaire



# Théorie chromosomique de l'hérédité

Sutton (1903) et Boveri (1904)

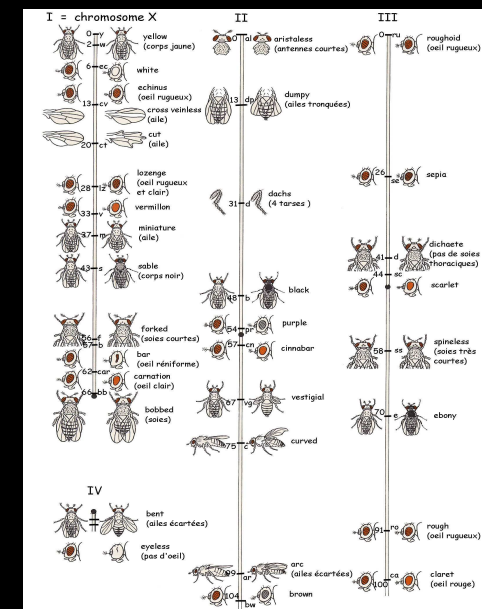


Chromosomes : Support des gènes  
Emplacement exact d'un gène sur le  
chromosome : Locus du gène

# Thomas Morgan, Alfred Sturtevant, Calvin Bridges, Herman Muller



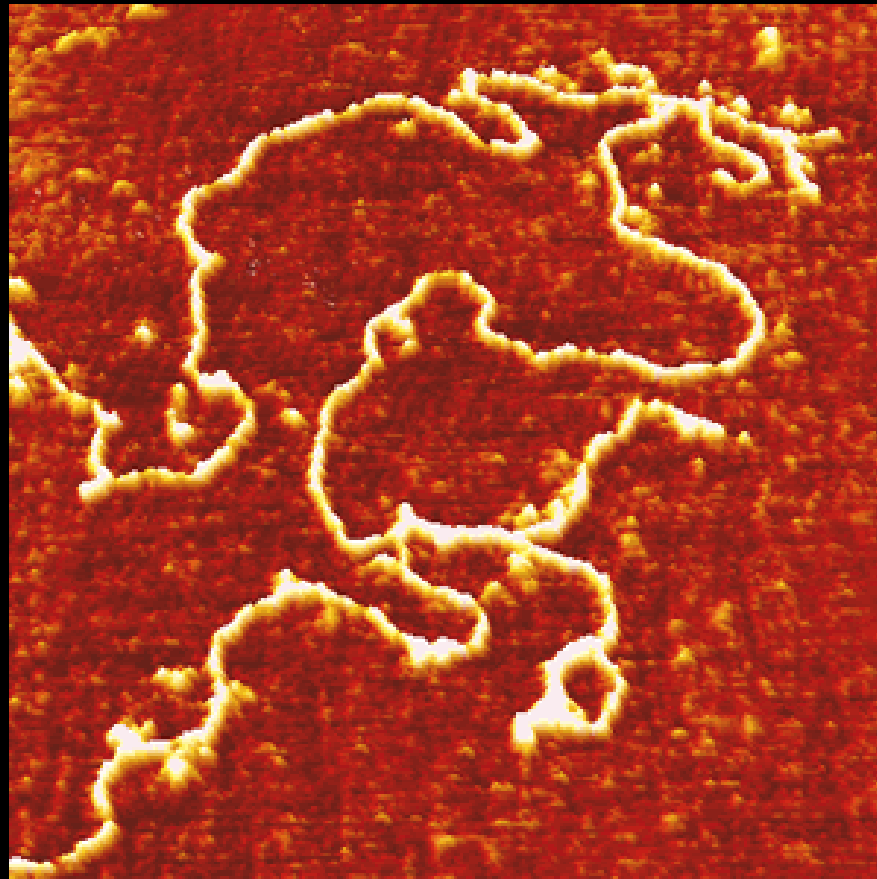
## Les liaisons génétiques et la première carte chromosomique





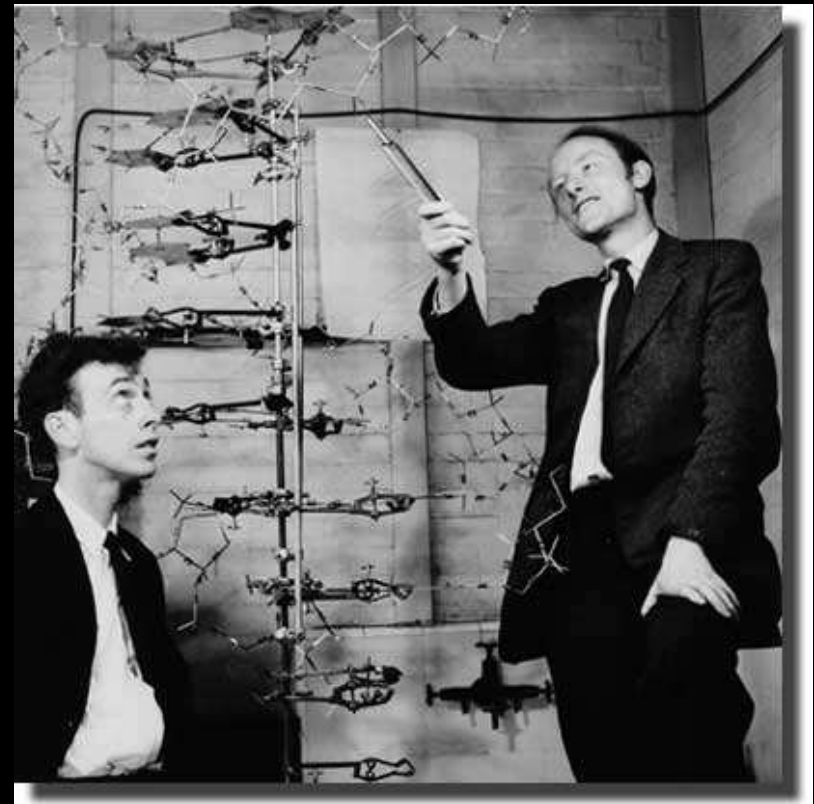
# La nature chimique du gène est l'ADN

*Avery, McLeod & McCarty (1944)*

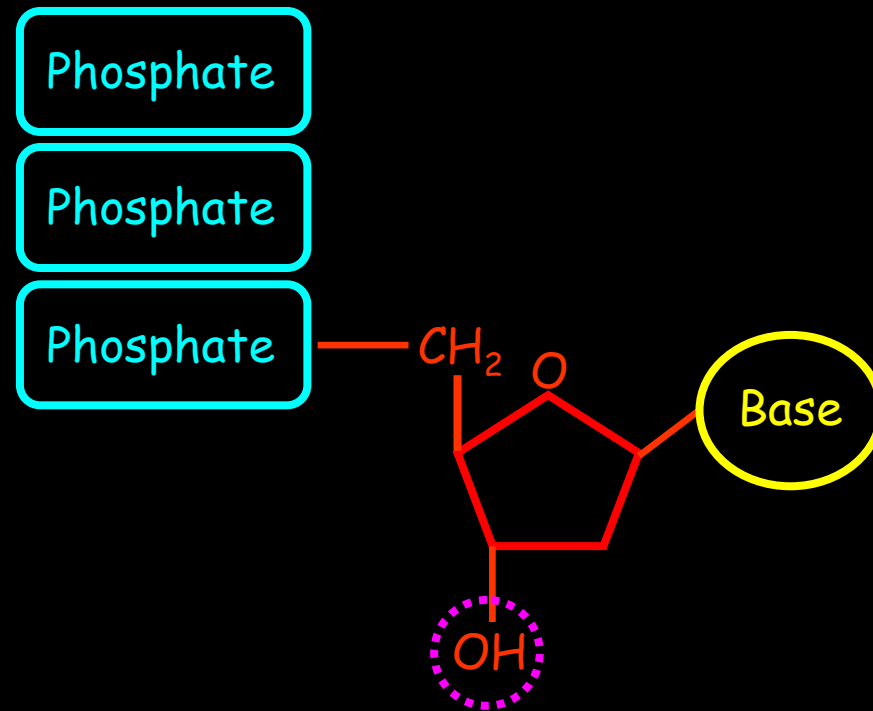


# La structure de l'ADN est une double hélice de brins complémentaires

James Watson & Francis Crick (1953)



# Désoxyribonucléotidetriphosphate dNTP

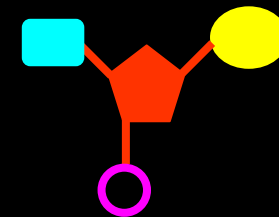


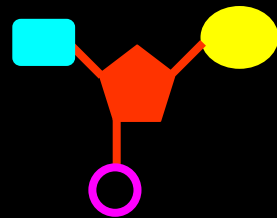
dATP

dGTP

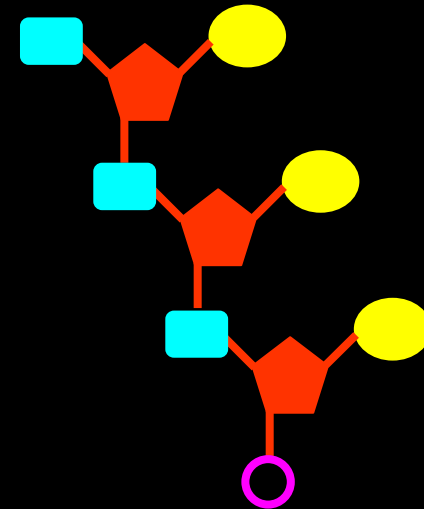
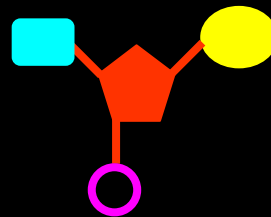
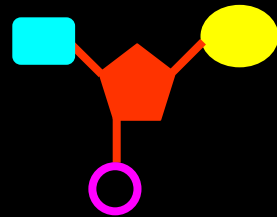
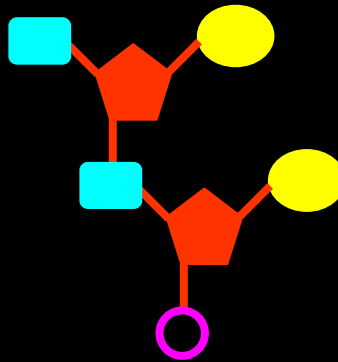
dCTP

dTTP

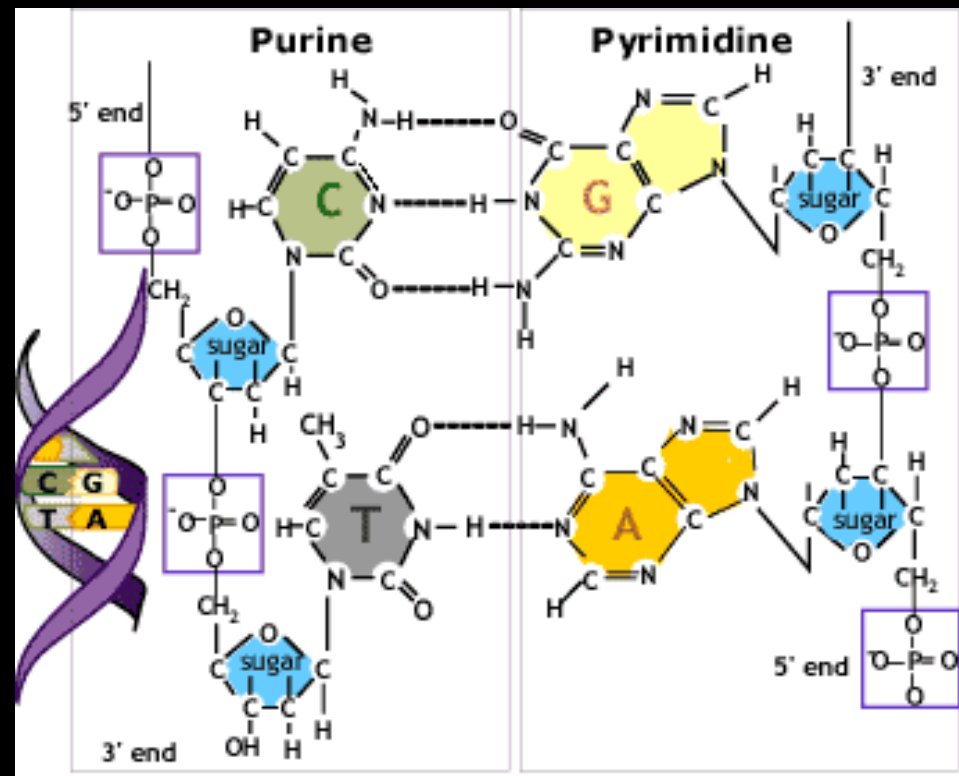
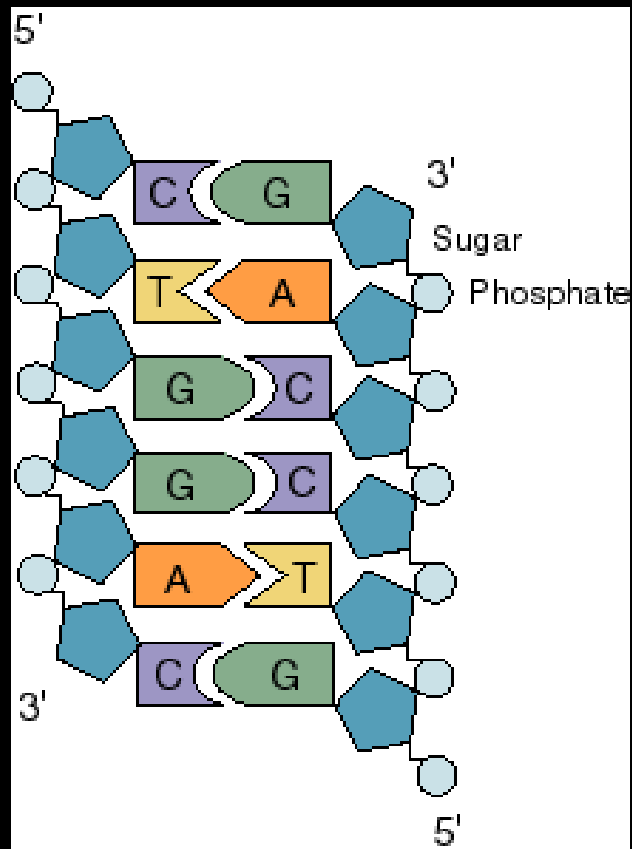




ADN Polase

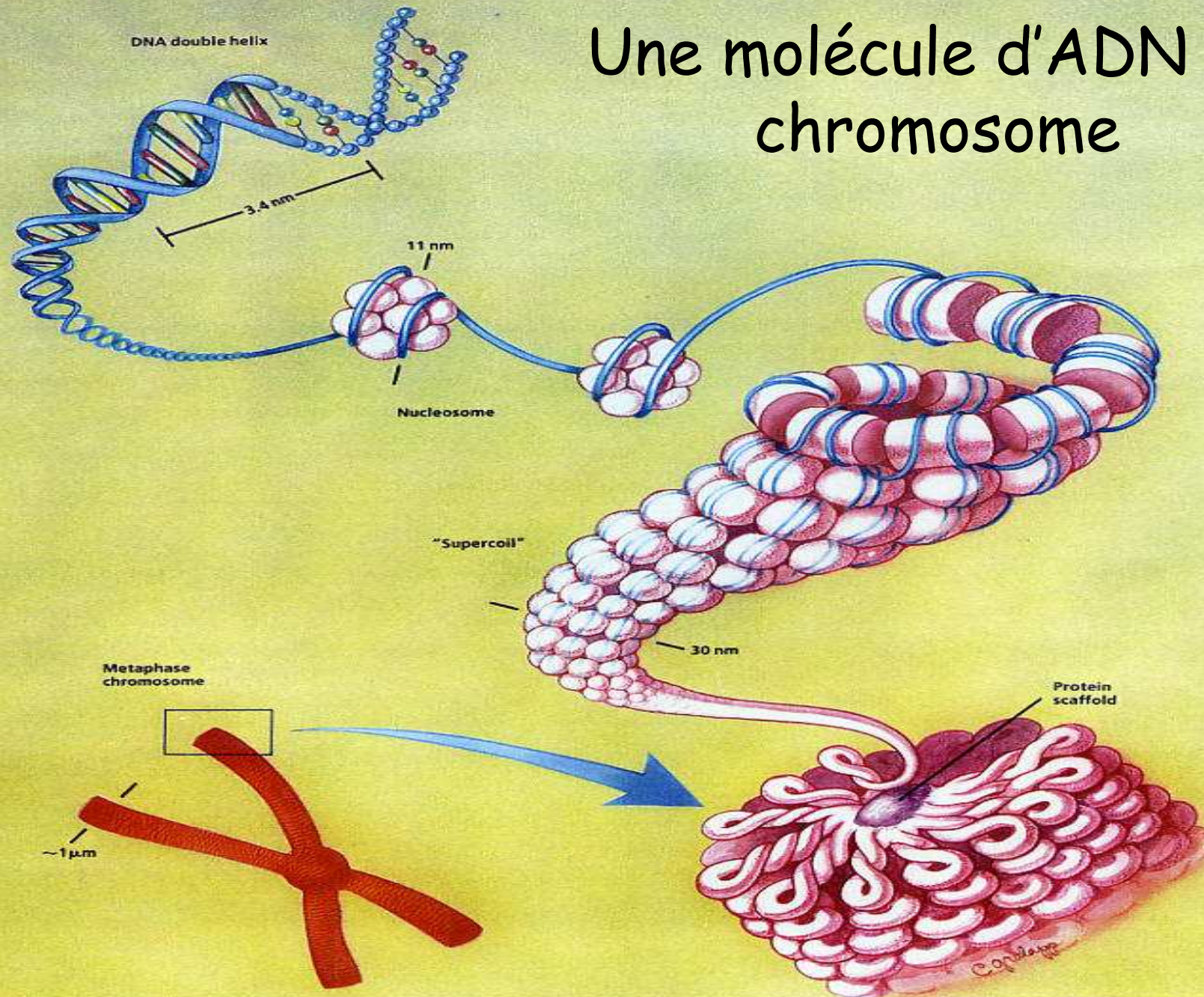


# Ponts hydrogènes



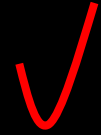


Une molécule d'ADN = 1  
chromosome

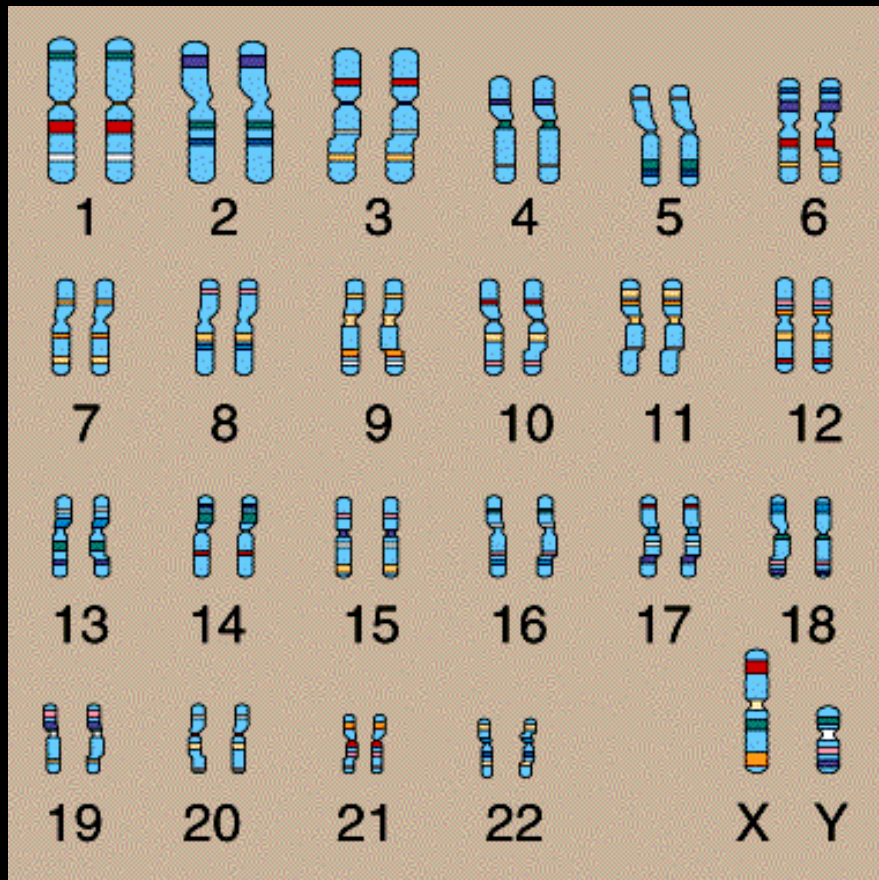




presque



Deux molécules d'ADN identiques en  
séquence = 2 chromosomes homologues



Chromosome 21 : 46,9  
Mb

La succession des 47  
millions de bases est  
(presque) identique  
entre les deux  
homologues

## Début des années 60

G A T T A C A

# Le dogme central de la génétique moléculaire

[illegible]

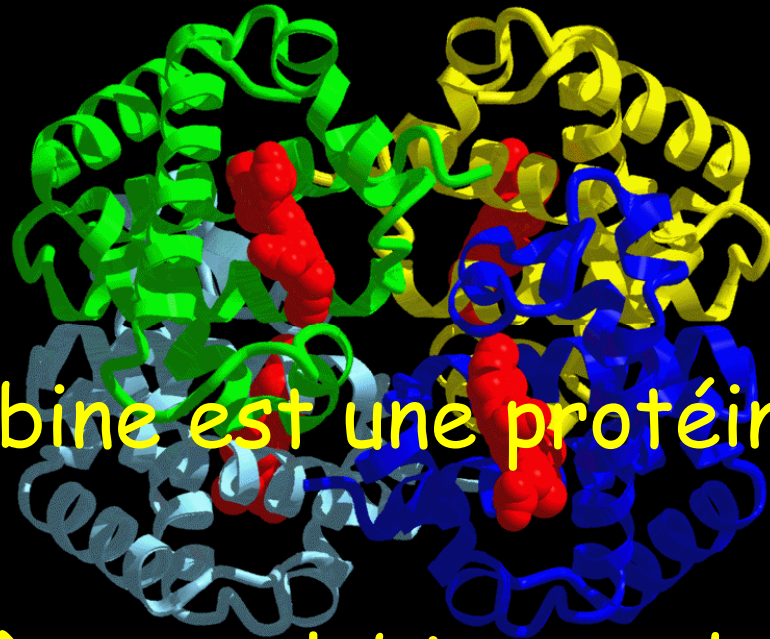


# Code génétique standard

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

# L'expression du gène : de l'ADN au phénotype

Le cas de la drépanocytose (anémie  
falciforme)



L'hémoglobine est une protéine composite

Deux chaînes  $\alpha$ -globine et deux chaînes  $\beta$ -  
globine

## $\beta$ -globine normale

ADN

.....T-G-A G-G-A C-T-C C-T-C.....

ARNm

.....A-C-U C-C-U G-A-G G-A-G.....

a.a.

1

.....

4

thr

5

pro

6

glu

7

glu

146

.....

## $\beta$ -globine mutante

ADN

.....T-G-A G-G-A C-A-C C-T-C.....

ARNm

.....A-C-U C-C-U G-U-G G-A-G.....

a.a.

1

.....

4

thr

5

pro

6

val

7

glu

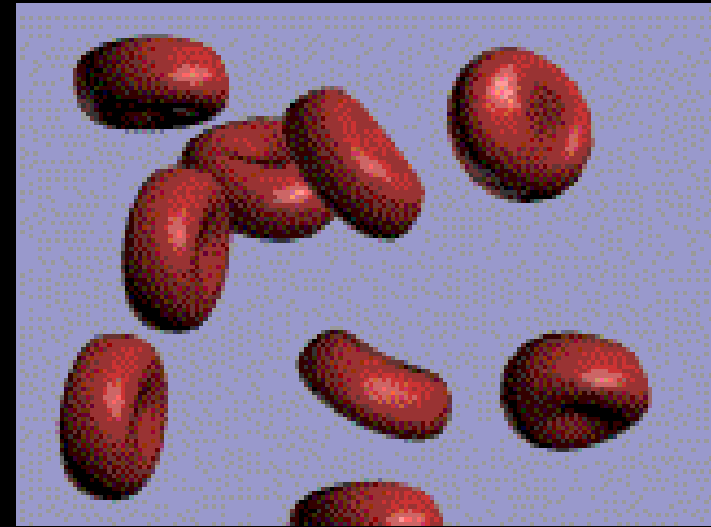
146

.....

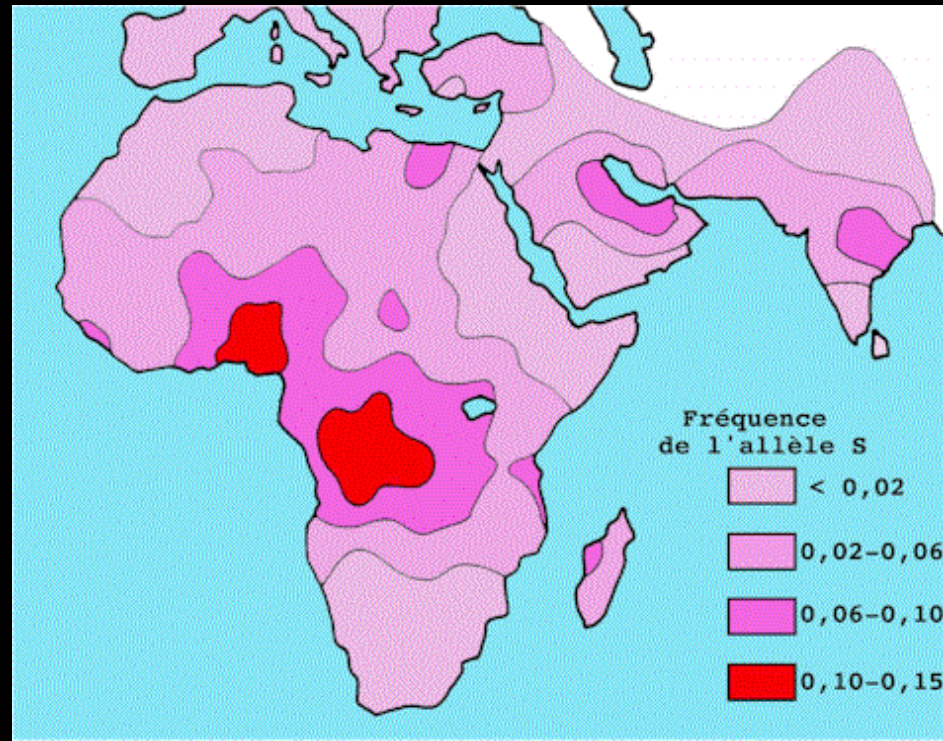
Homozygote mutant : drépanocytose

Molécule d'hémoglobine avec  $\beta$ -globine\* :  
Capable de transporter  $O_2$

Capacité à s'associer en longues fibres : étirent le globule en  
faucille



Blocage de la circulation dans les capillaires, fragilité des  
globules : réduction du n<sup>bre</sup> et anémie



*Population atteinte de l'anémie falciforme en Afrique*

25 % de la population en Afrique centrale est porteuse du trait drépanocytaire, 15 à 20% en Afrique de l'ouest ; 10 à 12 % dans les DOM antillais ; 1 à 15% selon les régions méditerranéennes

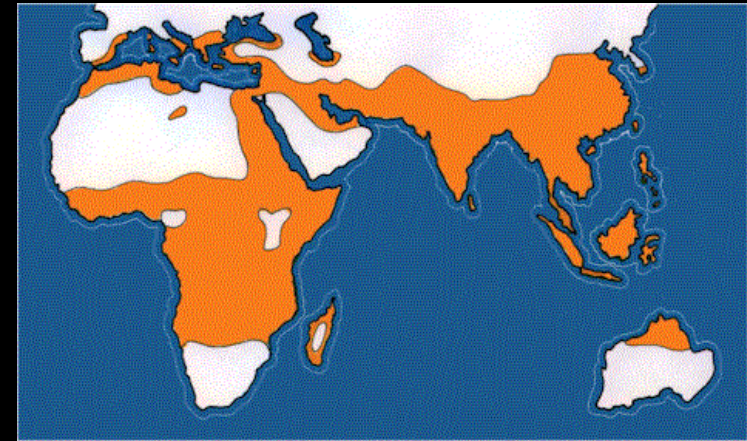
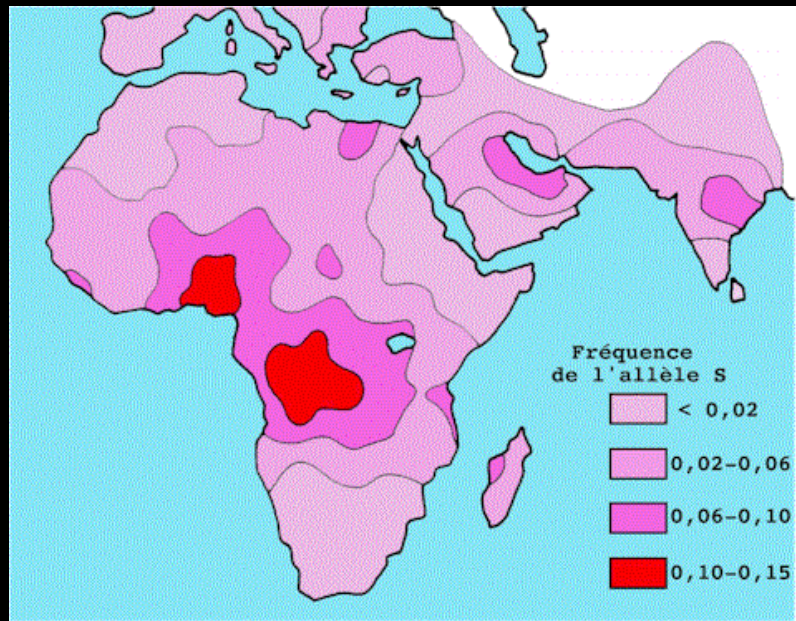
Autrefois : anomalie génétique létale (100 000  
morts / an)

80% des homozygotes HbS/HbS mourraient avant l'âge  
de la reproduction



Pourquoi dans certaines populations humaines la  
fréquence de HbS > 10% ??





*Répartition de la population  
porteuse du parasite du  
paludisme*

**l'hétérozygote Hb/HbS résiste mieux au  
paludisme !**

# A l'ère de l'ADN recombinant et de la génomique

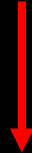
1970 : Découverte des enzymes de restriction chez les bactéries

On peut couper l'ADN de manière spécifique en petits fragments

Polymérase : Le clonage (multiplication) en grand nombre devient possible (PCR)



Une technologie de l'ADN de plus en plus performante



Séquençage des génomes (projet génome humain, 1990)

2001 : Projet génome humain (subvention publique) + projet privé de séquençage (Celera Corporation)



96% des segments génomiques porteurs de gènes

Accélération des projets de  
séquençages



Émergence de la génomique

Analyser les séquences pour étudier la  
structure, la fonction et l'évolution des gènes  
et des génomes

Nouvelle génération de séquenceurs : des  
millions de paires de bases séquencées en un  
seul run



Nouveaux enjeux :  
séquençage de  
génomés de  
centaines d'individus

Importance croissante des  
biotechnologies

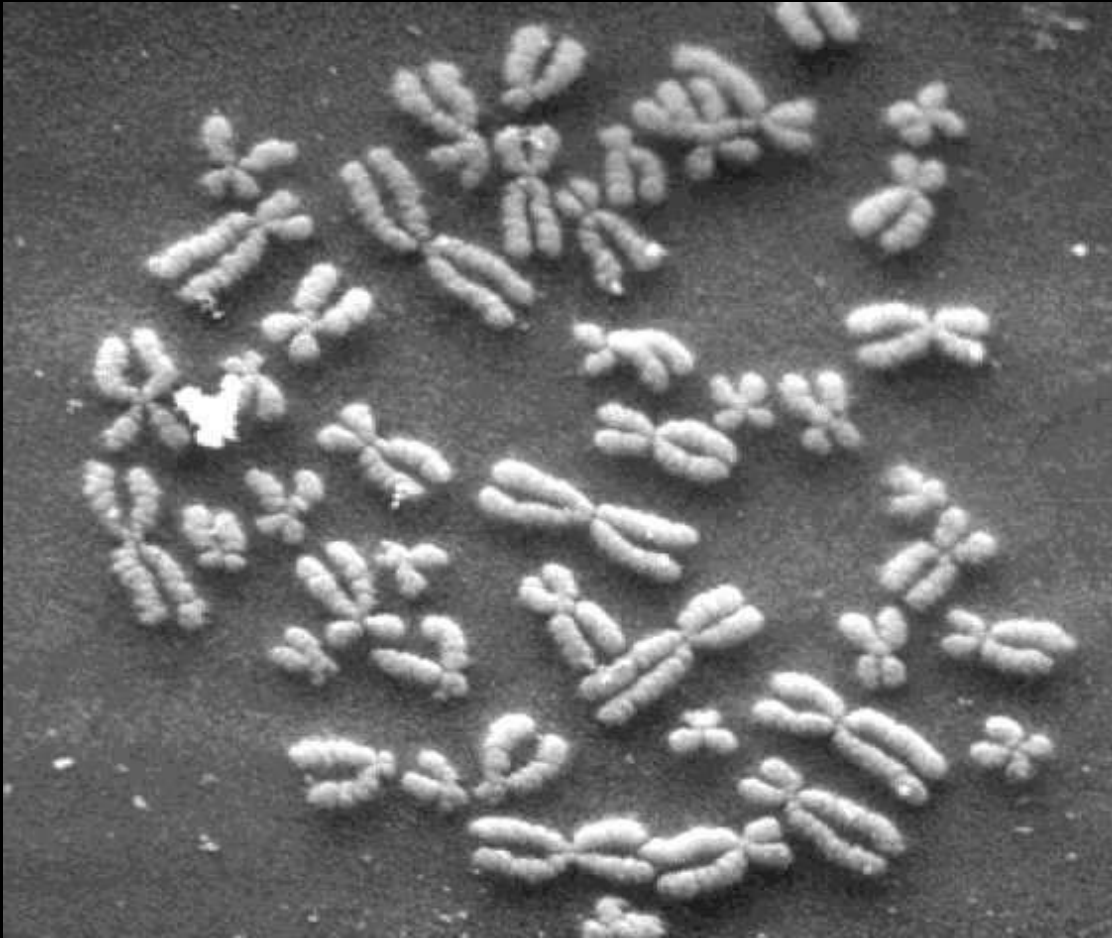


De réels problèmes de société

OGM, brevetage du vivant, clonage, thérapies  
géniques...etc

La technologie génétique avance bcp + vite que  
les lois, les débats publics et les règles  
sociales

# Le génome



Ensemble des  
molécules  
d'ADN présent  
dans une cellule

# Génome

Terme crée en 1920 : Ensemble des gènes contenu dans un organisme

Jeu complet d'ADN d'un organisme

Souvent plusieurs molécules ADN :  
Chromosomes (noyau)

Mitochondries et chloroplaste : Un ADN  
(ADN<sub>m</sub> et ADN<sub>c</sub>)

# Définition moléculaire du gène

La séquence d'acides nucléiques  
entière qui est nécessaire à la  
synthèse d'un polypeptide  
fonctionnel ou d'un ARN

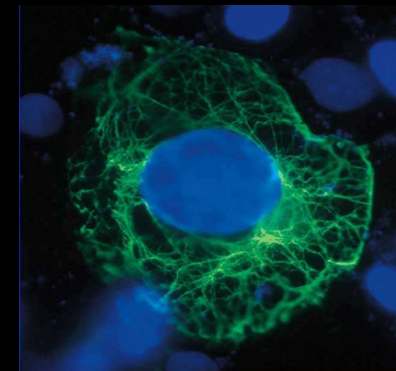
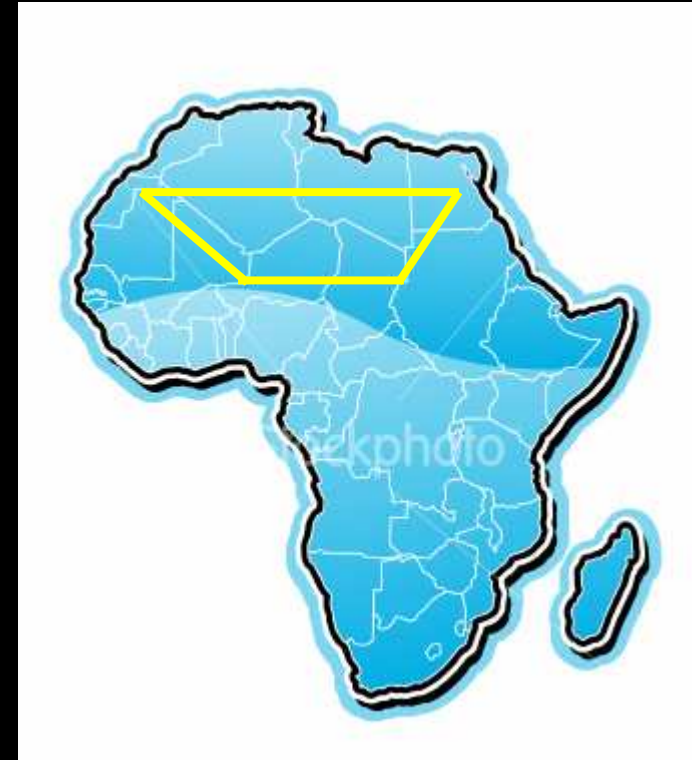


# Génome humain : 3 milliards de bases

Séquence en caractères  
10 : 7 500 km

Peut être stocké sur un  
seul CD-Rom

Biologiquement codé : tient  
facilement dans une  
cellule





# Génomique

Étude des génomes : Caractérisation  
moléculaire de génomes complets



Génomique structurale

Génomique fonctionnelle

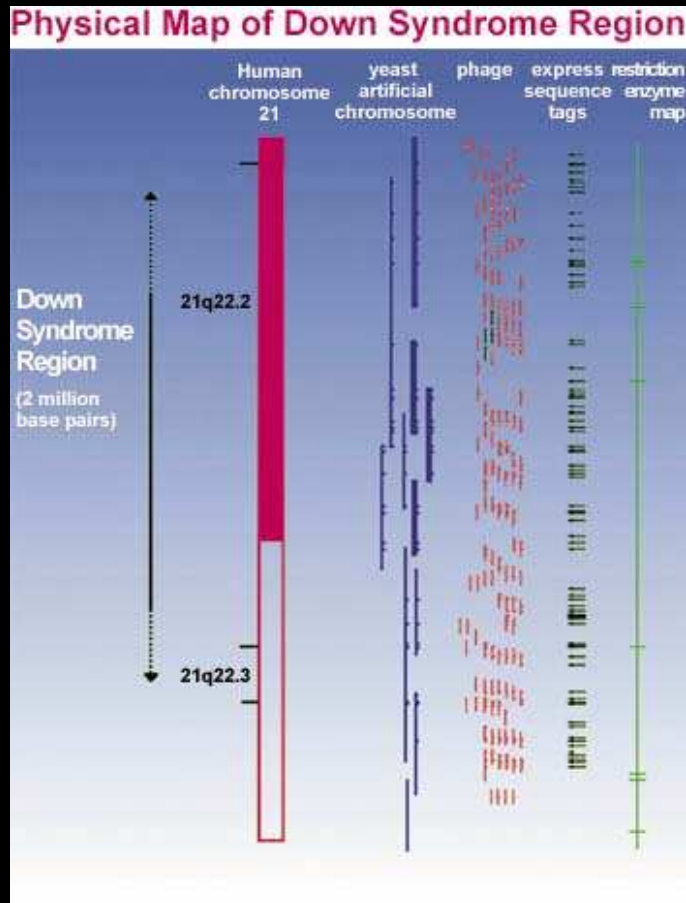
Génomique comparative  
(évolutive)

# Génomique structurale

Caractérise la nature physique des génomes entiers (séquence et identification)

Intègre cartographie physique et génétique ainsi que séquençage des génomes

# Génomique structurale



Organisation des séquences

Localisation des loci sur les chromosomes

Cartes chromosomiques à haute résolution

Séquençage de génomes

Utilisation des cartes génomiques pour analyse génétique

# Génomique fonctionnelle

Protéome et mode d'expression des gènes  
(régulation)

Comprendre l'expression du génome en  
fonction des stades de dév<sup>t</sup> et de l'en<sup>t</sup>

# Quels gènes sont activés (switch on) ou désactivés (off) ?



TGT AAT AGT TAT ATT TTC  
ATT ATA AAT TGT GTT TGT  
AGA CAT CAT AAA TTT AAA  
ACA TGG CTT TTT AAC CTG  
ATA AAT CCT ACG AAT ATT  
TGT AAT AGT TAT GTT ATT  
GCA GTA AGT ACC GTT TGT  
ATT ATA AAT TGT GTT CTG

TGT AAT AGT TAT ATT TTC  
ATT ATA AAT TGT GTT TGT  
AGA CAT CAT AAA TTT AAA  
ACA TGG CTT TTT AAC CTG  
ATA AAT CCT ACG AAT ATT  
TGT AAT AGT TAT GTT ATT  
GCA GTA AGT ACC GTT TGT  
ATT ATA AAT TGT GTT CTG



Young Moo Lee, 2000

# Génomique comparative

Comparaison de séquence et évolution des génomes

Les informations obtenues sur un organisme peuvent avoir des applications sur d'autres organismes

Exemples : relations synténiques

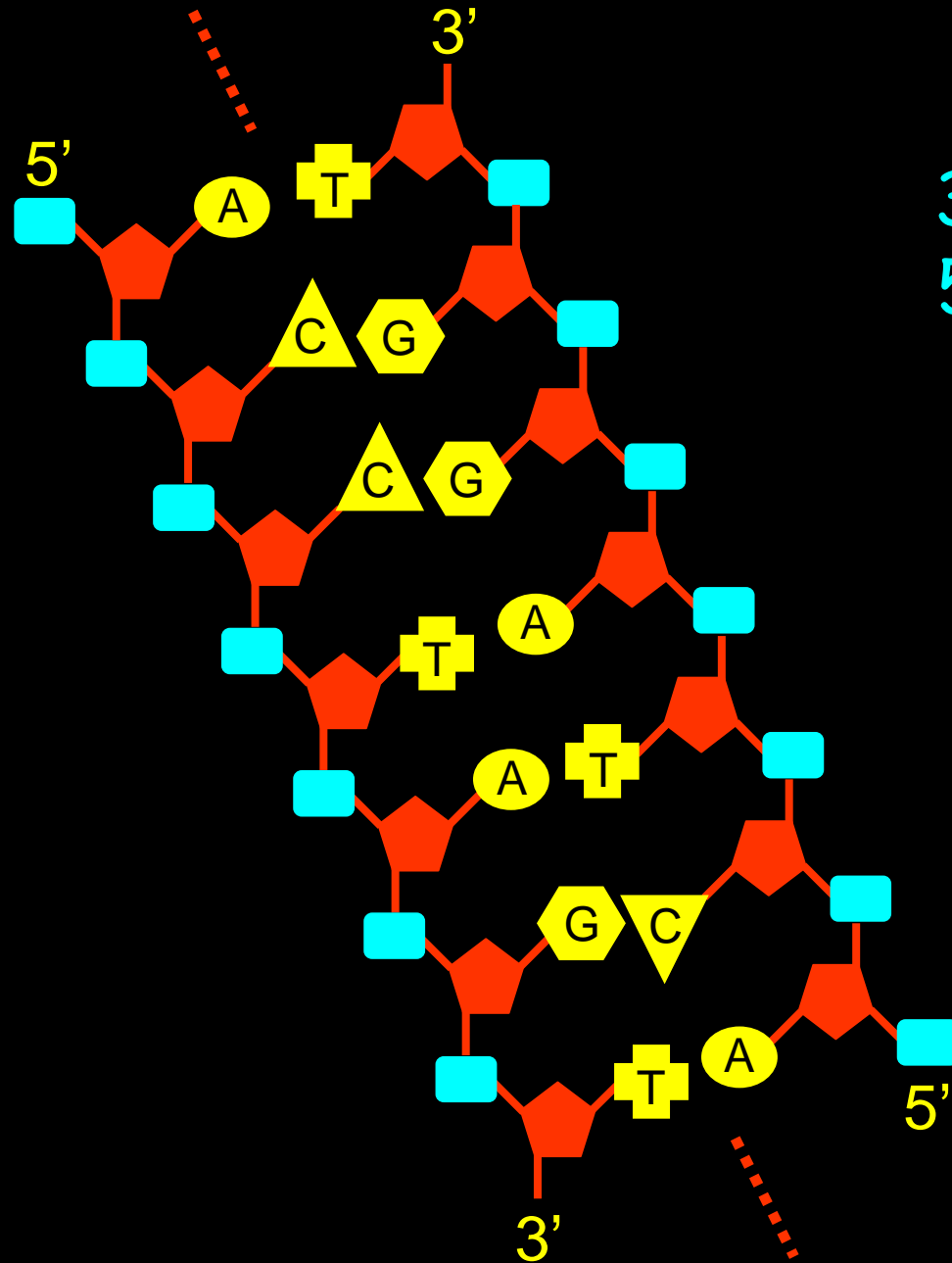


5'

# Notion de séquence d'ADN

AGCTTACCGTAATTTACCGCTTACCGTAATTTACCG  
TGGCCTACTTACCGTAATTTACCGCCTGACCTACCT  
CTACGCTTACCGTAATTTACCGATTACCGTAGGCC  
AATATACTCTTACCGTAATTTACCGGACCTCCTCCGG  
CTTACCGTAATTTACCGTACCTATCCGTTTACT  
ATCGGTCCTACCTACTTACCGTAATTTACCGATTAC  
CTCTACCCTTACCGTAATTTACCGTACGATAAGGCC  
AATACTACTTACCGTAATTTACCGCCTTACCTCCGG  
CTGACTACCTTCTTACCGTACCTATCCGTTTACT  
CGGACTACCCTTACCGTAATTTACCGTCTGATACCC  
AGCTTACCGTAATTTACCGTACCGTAATTTACCG  
TGGCCTACTTACCGTAATTTACCGCCTGACCTACCT  
CTACGCTTACCGTAATTTACCGATTACCGTAGGCC  
AATATACTCTTACCGTAATTTACCGGACCTCCTCCGG  
CTTACCGTAATTTACCGTACCTATCCGTTTACT  
ATCGGTCCTACCTACTTACCGTAATTTACCGATTAC  
CTCTACCCTTACCGTAATTTACCGTACGATAAGGCC  
AATACTACTTACCGTAATTTACCGCCTTACCTCCGG

3'



3'-T-G-G-A-T-C-A-5'  
5'-A-C-C-T-A-G-T-3'

Séquence  
ACCTAGT



# Taille d'une séquence

La taille d'une séquence peut s'exprimer:

- En nombre de bases
  - Milliers de bases → Kilobases *kb*
  - Millions de bases → Megabases *Mb*
  - Milliards de bases → Gigabases *Gb*
- En paires de bases → *pb*

# Séquençage de génomes entiers

- Déterminer la séquence nucléotidique de l'ADN total présent dans chaque  $\phi$  d'un organisme
- Pb : Taille et richesse en séquences répétées  $\uparrow$  difficulté
- Virus, bactéries : Petits génomes et pas de séquences répétées
- Virus : 3 000 à 150 000 pb
- Bactéries : qqs millions pb (Mb)

# Eucaryotes

- Paramécie : 100 Mb
- Mammifères : 2 à 3 milliards pb (Gb)
- Homme : 3 000 Mb (Gb)
- Eucaryotes : Grande taille des génomes
- Eucaryotes : très nombreuses séquences répétées

## Février 2005

- Virus : > 1000 séquencés
- Archaeobactéries : près de 20 souches
- Bactéries : > 125 souches et espèces
- Eucaryotes : 21 espèces

*Saccharomyces cerevisiae* (levure)

*Neurospora crassa* (ascomycète) - ébauche

*Caenorhabditis elegans* (ver nématode)

*Drosophila melanogaster* (mouche du vinaigre)

*Anopheles gambiae* (moustique) - ébauche

*Takifugu rubripes* (fugu : poisson-ballon - ébauche)

*Mus musculus* (souris)

*Homo sapiens* (nous...)

*Plasmodium falciparum* (paludisme)

*Arabidopsis thaliana*,

*Oriza sativa* (ébauches)

# **Complete Published Genome Projects**

**Feb., 2007**

<b>Archaeal</b>	<b>35</b>
<b>Bacterial</b>	<b>421</b>
<b>Eukaryal</b>	<b>47</b>
<b>Organelle</b>	<b>1089</b>
<b>Phage</b>	<b>346</b>
<b>Plasmid</b>	<b>480</b>
<b>Viroid</b>	<b>39</b>
<b>Virus</b>	<b>1260</b>

*Bactérie*

Gène



*Drosophile*

Gène



*Homme*

Gène



ADN non codant

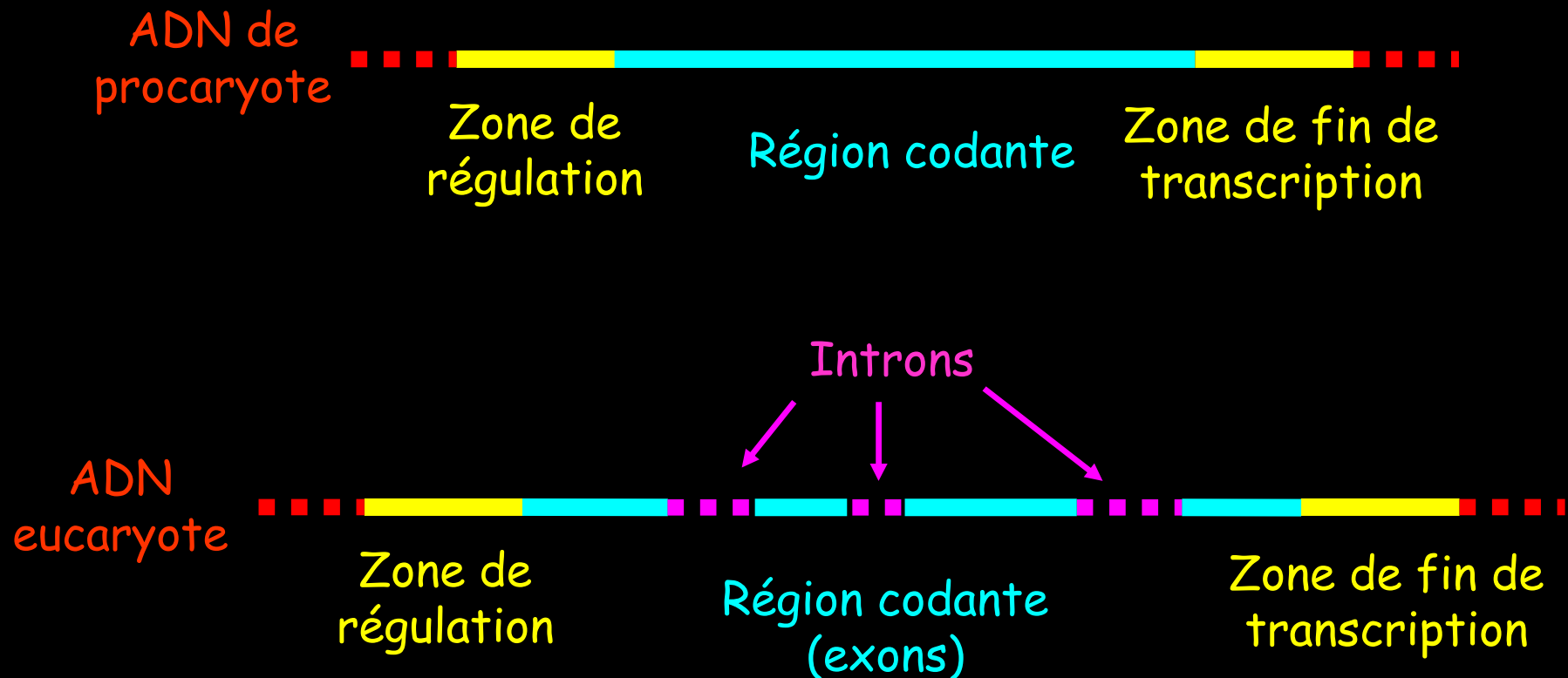
Exon

Intron



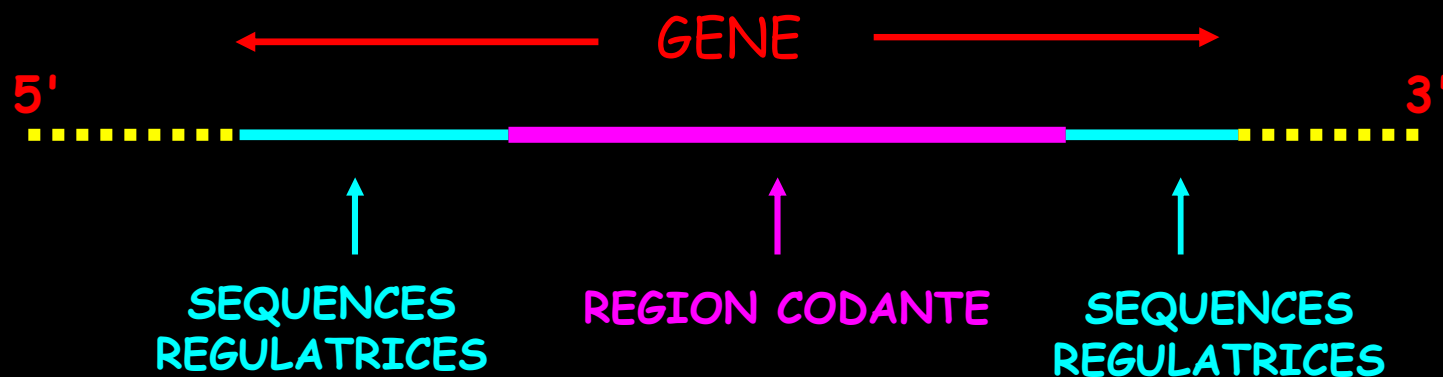
# Nature des gènes

- Région d'ADN transcrite en ARN

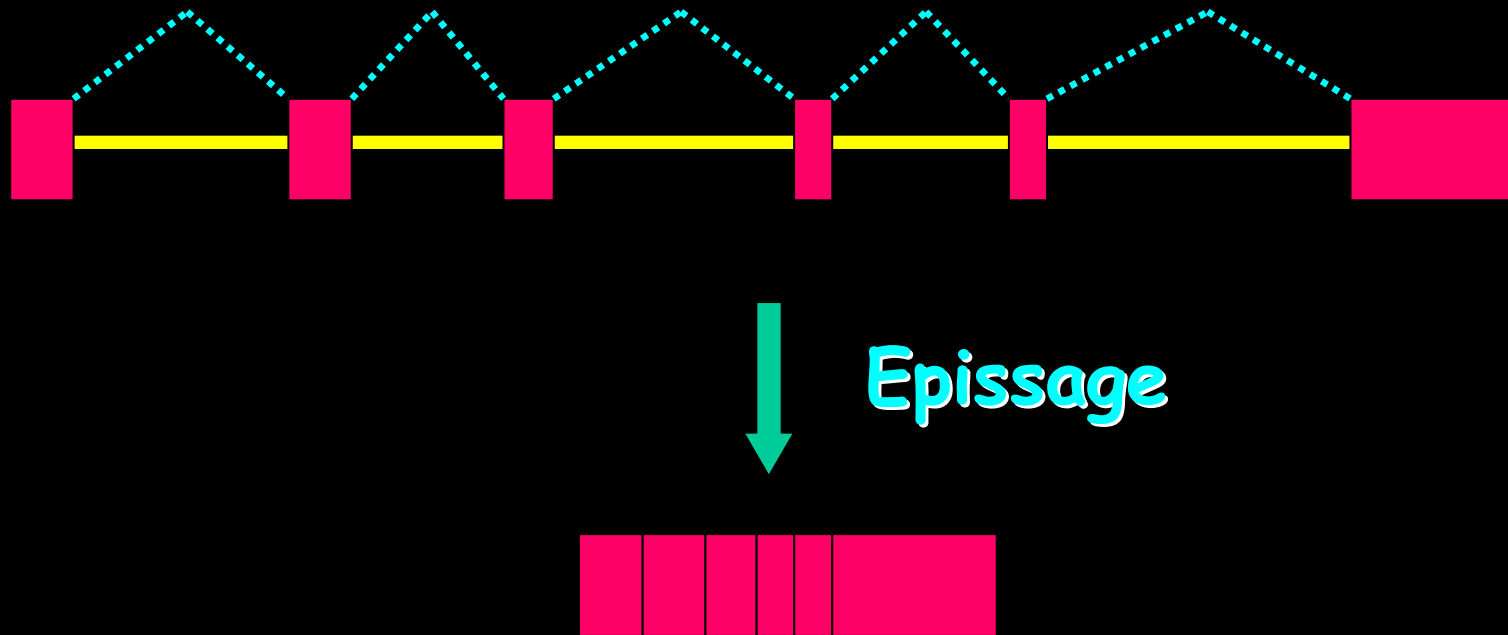


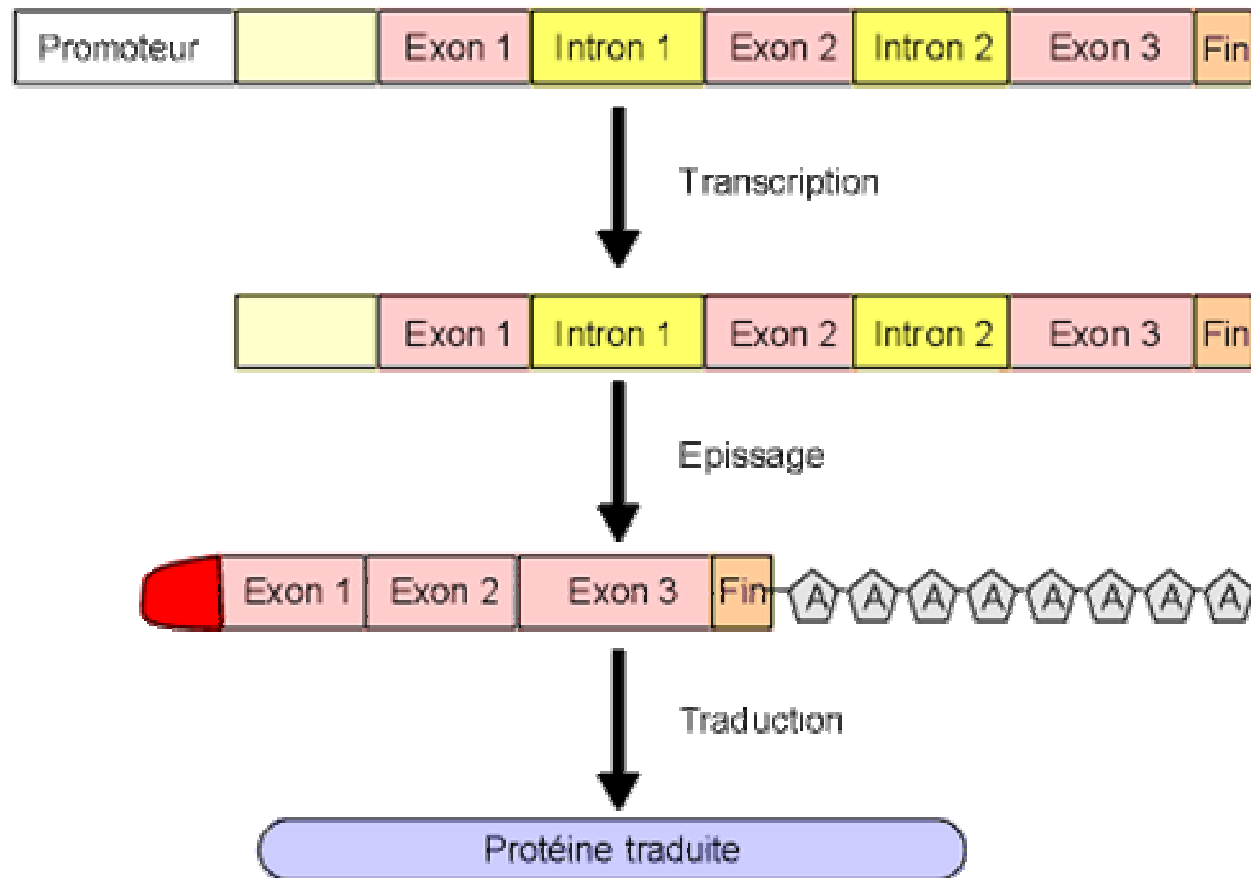
# Structure des gènes

- Gène eucaryote → Région centrale transcrite (région codante)
- Séquences régulatrices → Expression du gène et son rythme



# Gènes eucaryotes (exons & introns)





Pléiotropie

# Quelques personnages célèbres



Amenophis IV



Lincoln



Paganini

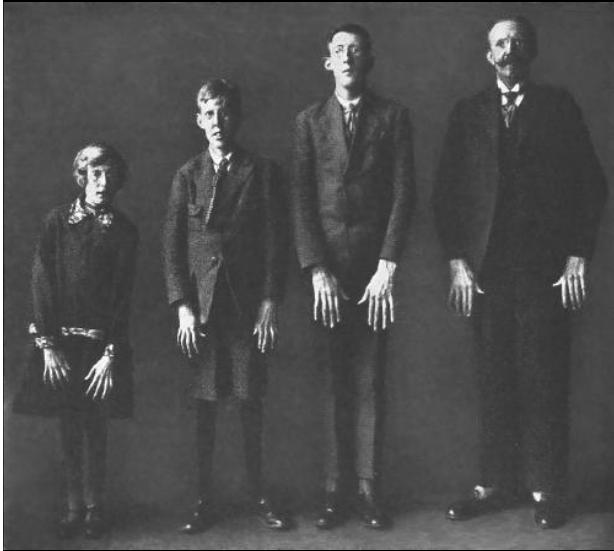


Rachmaninov

Quels liens entre ces personnages ?



# Syndrome de Marfan: Aspects phénotypiques

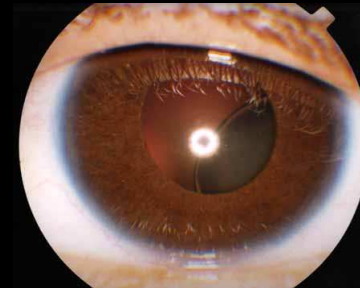


Scoliose



Grande taille

Disfonctionnement  
des valves  
cardiaques



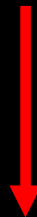
Doigts  
allongés  
Problèmes de  
rétine

Degrés variables de sévérité du  
syndrome !!

# Syndrome de Marfan

Plusieurs caractéristiques  
phénotypiques

Origine ?



Désordre du tissu connectif → Anomalie de distribution  
des fibres élastiques du connectif

# Tissu connectif

Un tissu fondamental !!



Flexibilité et déformabilité des tissus  
dynamiques d'un organisme

Tendons

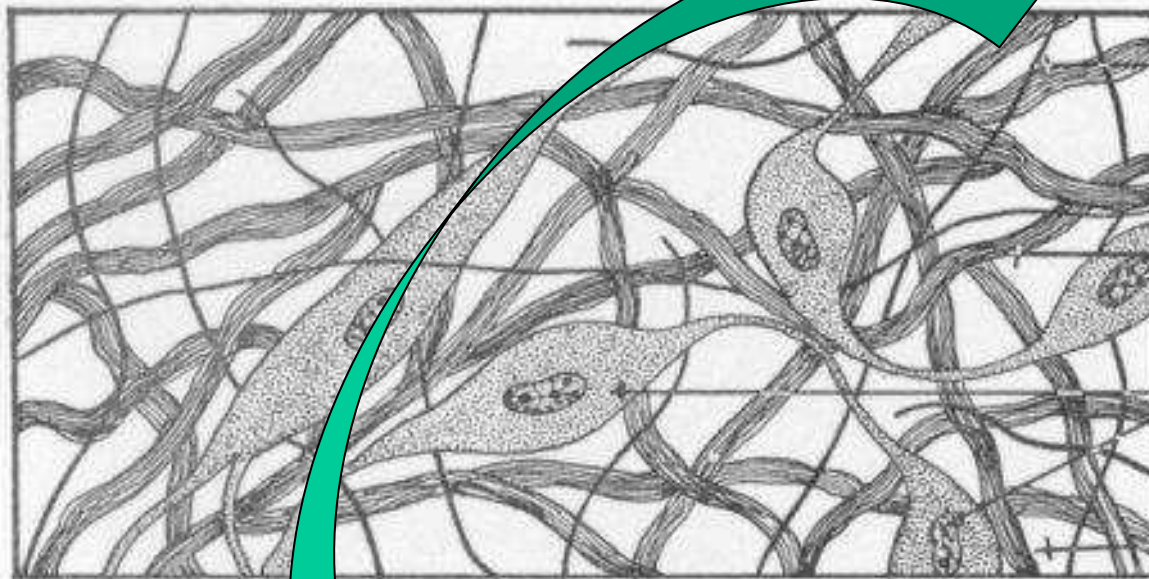
Vaisseaux sanguins

Peau

Poumons

Ligaments

..etc.



Fibre d'élastine

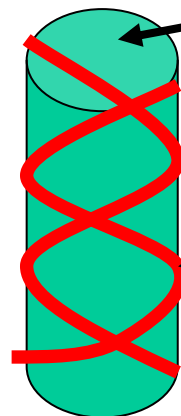
Fibre de collagène

¢ du tissu  
connectif

Gelée

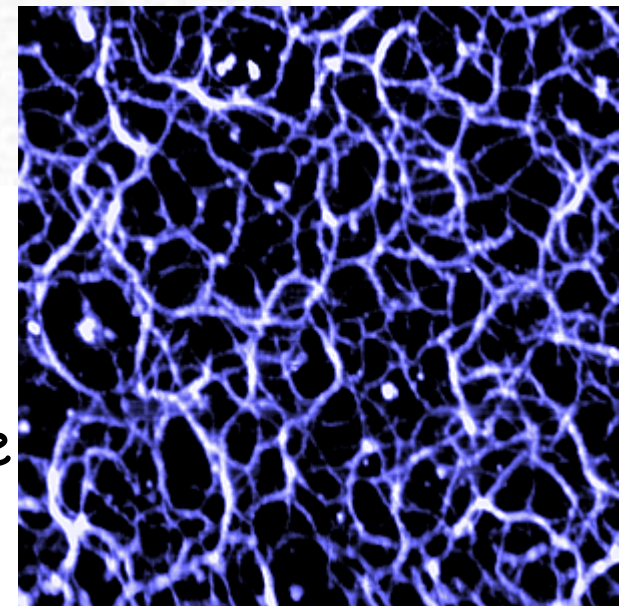
Tissu connectif

Fibre  
d'élastine



Elastine

Microfibrille

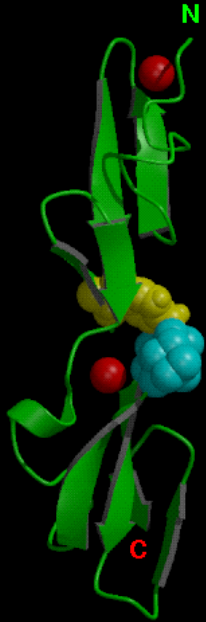


Fibrilline-1 → Protéine majeure des  $\mu$ fibrilles

Protéine EGF (Epidermal Growth Factor) riche en cystéine (14% des a.a.)

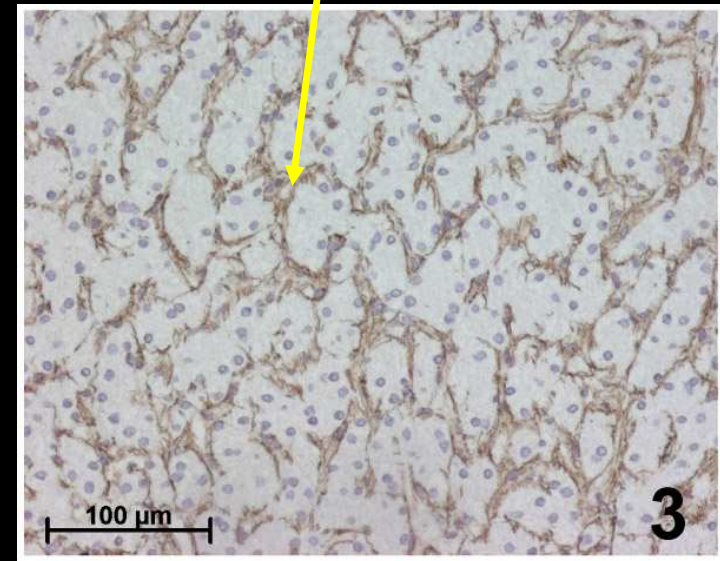
43 sites d'accrochages  $\text{Ca}^{++}$

Cystéines + sites d'accrochages  $\text{Ca}^{++}$  → Agrégation et élasticité



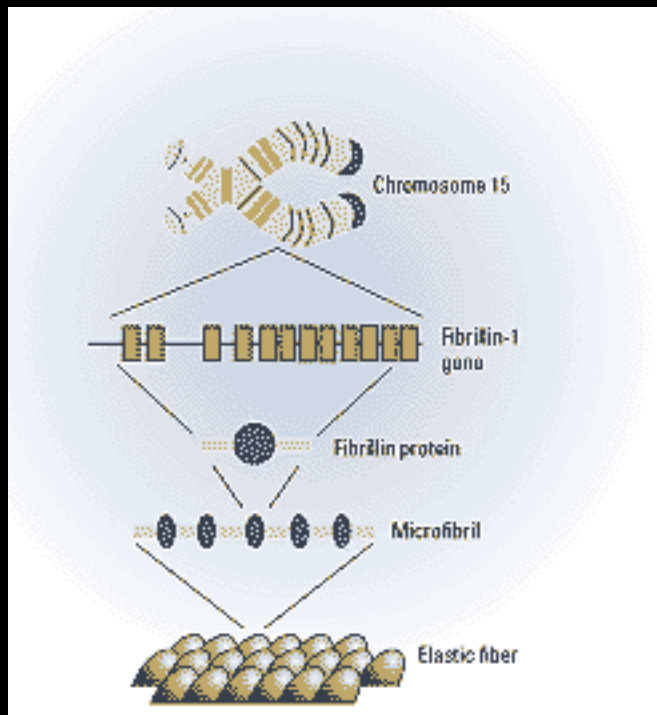
Tissu  
connectif

Fibrilline-1

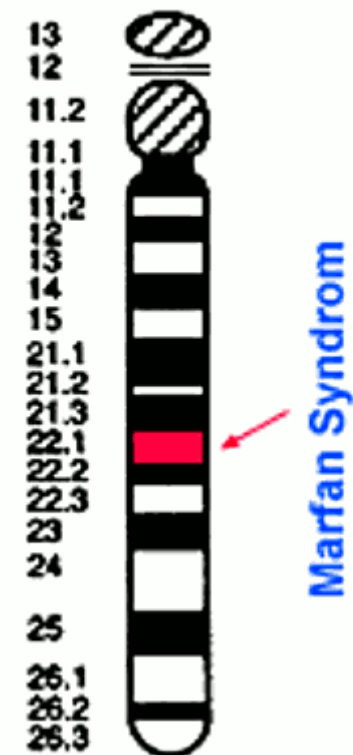


# Gène de la Fibrilline-1 → Gène FBN1

FBN1 → Chromosome 15



## Chromosom 15



Gène = 110 kb avec 65 exons, transcrit de 10 kb



Le syndrome de Marfan est lié à des mutations sur le gène FBN1

1<sup>ère</sup> mutation identifiée en 1991

Transversion G→C au nucléotide 716

Pro→Arg au codon 239

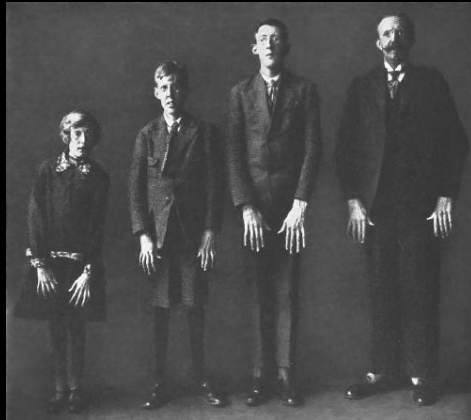
Depuis, des centaines de mutations ≠ identifiées (substitution, délétion)

# Syndrome de Marfan

Un seul gène impliqué

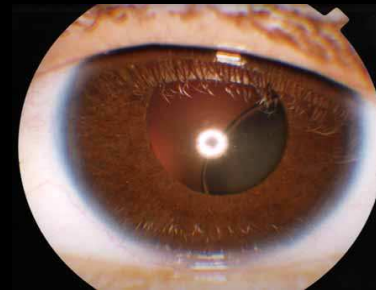
De

multiples

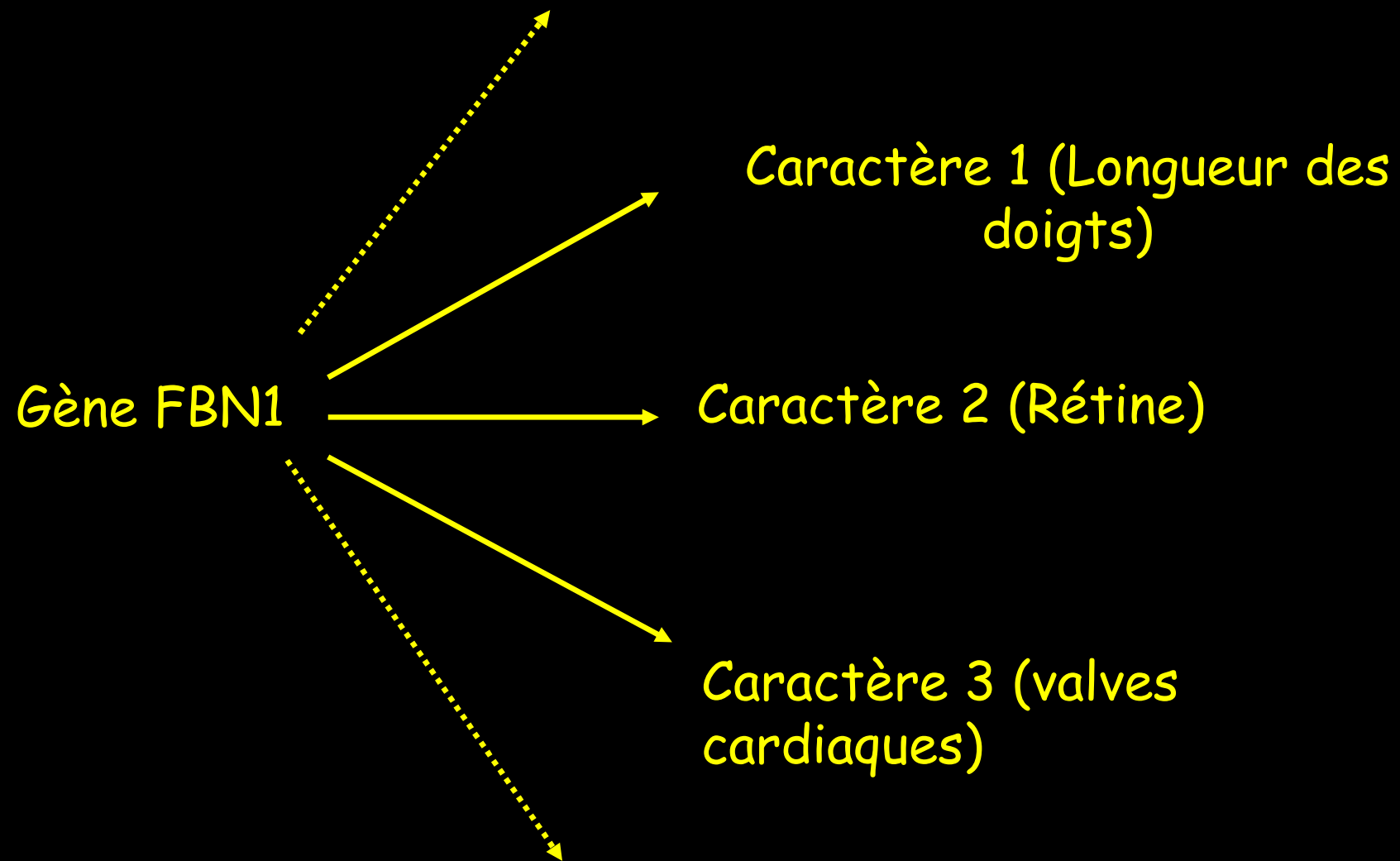


caractères

phénotypiques



touchés



**Relation pléiotropique**

La relation entre gène et phénotype  
peut être plurivalente

### **Pléiotropie**

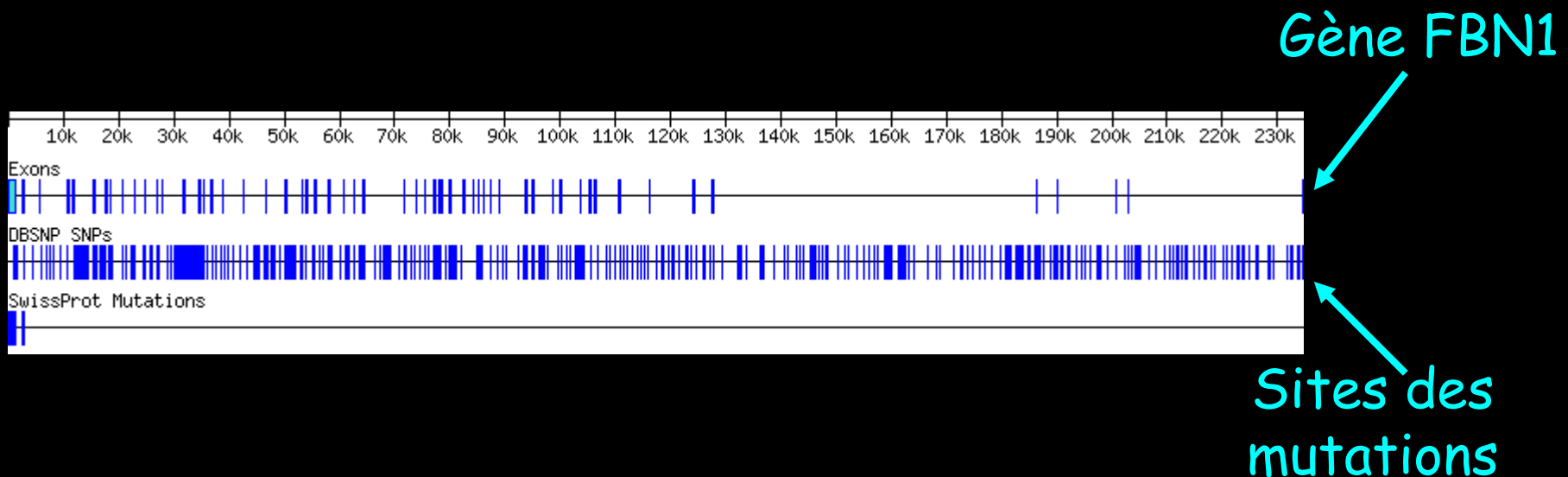
Un gène affecte plusieurs caractères

### **Mutations pléiotropiques**

Une mutation qui affecte plusieurs  
caractères différents

Pourquoi des degrés variables de  
sévérité du syndrome ??

Des centaines de mutations  $\neq$  identifiées



Chaque mutation = 1 allèle  $\neq$

Chaque mutation  $\rightarrow$  Conséquence  $\pm$  importante  
sur la fonctionnalité de la protéine

Bcp de substitutions synonymes

Bcp de mutations sur introns

Délétions

Bcp de substitutions non-synonymes

Substitutions non-sens

Des hétérozygotes → Dominance  
incomplète



# I - Les Mutations

Mutations : source de la variabilité génétique



Fondement de la sélection naturelle



A l'origine des lésions génétiques

Les mutations permettent l'analyse génétique

# Mutation

Une altération de la séquence d'ADN

Peut impliquer de grandes régions d'un ou de plusieurs chromosomes

Peut concerner que qqs nucléotides sur n'importe quelle zone de la molécule d'ADN

# Toutes les séquences sont mutables

Une mutation peut avoir un phénotype si elle touche :

une séquence codante

une séquence régulatrice (promoteur, enhancer...)

une séquence d'épissage ou de polyadénylation

# Deux grandes catégories de mutations

(1) **Mutations géniques** (mutations ponctuelles): Changement dans la séquence

(2) **Mutations chromosomiques**:  
Changements mutationnels : Chromosomes entiers ou des morceaux de chromosomes

Les mutations peuvent affecter les  $\phi$  somatiques ou les  $\phi$  germinales

$\phi$  somatiques  $\rightarrow$  mort  $\phi$ , dysfonc<sup>ment</sup>, tumeurs

$\phi$  germinales  $\rightarrow$  héréditaires