



# Introduction à la biologie moléculaire

Nadia El-Mabrouk

1

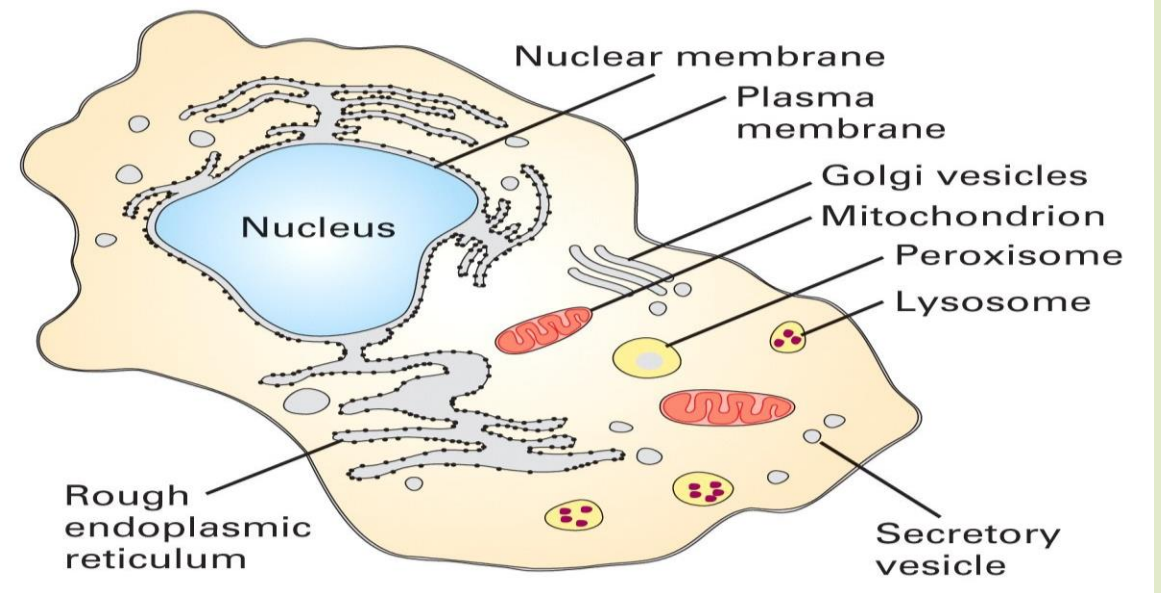
# Biologie moléculaire

- Née de la synthèse de deux approches complémentaires de l'étude de la vie : **la biochimie et la génétique**. Deux perspectives complémentaires sur le même sujet.
  - Biochimie: a permis d'étudier **les protéines**
  - Génétique: a permis de découvrir **les gènes**
- Discipline qui se popularise après la **découverte par James Watson et Francis Crick**, en 1953, de la structure de l'**ADN** (acide désoxyribonucléique).

Ce chapitre introductif présente un bref historique des fondements de la biologie moléculaire moderne, et définit les termes et concepts utilisés dans le cours.

# Les cellules

- Les organismes vivants sont faits de cellules: plus petite unités structurelle capable d'avoir un fonctionnement « autonome »
- Les cellules stockent, utilisent et passent de l'information.
- Cycle de vie d'une cellule: elle naît, se nourrit, se réplique. Meurt.



## Composition d'une cellule:

- 70% d'eau
- 7% de petites molécules (sel, lipides, acides amines, nucleotides, etc)
- 23% de macromolécules (protéines, polysaccharides, etc)

# Les cellules – Deux catégories principales

- Diversité de la vie sur terre : Des organismes aussi petits que le symbiote unicellulaire *Nanoarchaeum equitans* (rayon de 400 nanomètres), et aussi grands que la colonie de peupliers faux-trembles Pando dans l'Utah, un ensemble de tiges d'arbres reliées par un système de racines, pesant environ 6000 tonnes.
- **Eucaryotes – Procaryotes** : Principale subdivision des cellules vivantes.
  - **Procaryotes: Bactéries+ Archaea**. Organismes unicellulaires. Les bactéries représentent ~ 60% de la biomasse terrestre. Il y a 2-3 milliards d'espèces différentes de bactéries, dont celles composant le microbiome humain.
  - **Eucaryotes**: Domaine du vivant regroupant tous les autres organismes vivants (unicellulaires ou multicellulaires). Quatre grands règnes: **les animaux, les champignons, les plantes et les protistes** (eucaryotes unicellulaires).
- Virus: Ne fait pas partie des 3 domaines du vivant. Entité biologique nécessitant un hôte pour se répliquer (parasite). Formé d'ADN ou d'ARN.



Eucaryotes

<https://fr.wikipedia.org/wiki/Eukaryota>

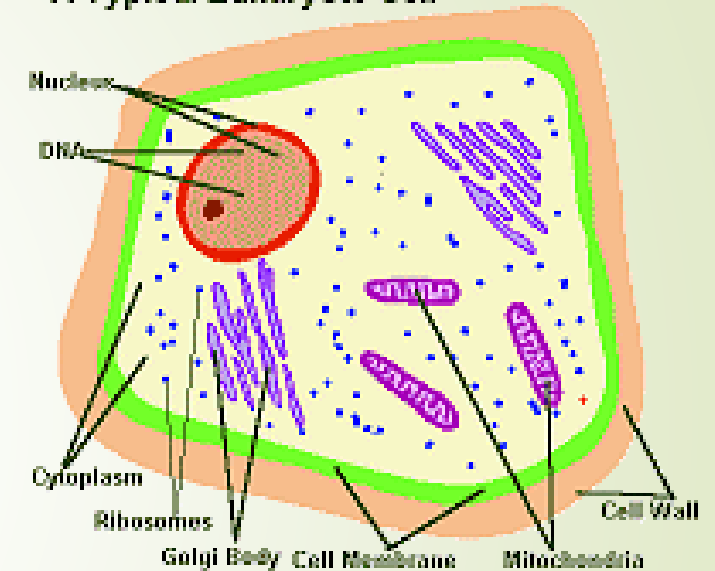
# Les cellules – Deux catégories principales

- **Cellule eucaryote:** Contenu cellulaire réparti dans plusieurs compartiments différents entourés de membranes lipidiques: **organites membranaires**. Trois organites particulièrement importants car ils contiennent le génome cellulaire:
  - **Le noyau:** Contient l'ADN nucléaire.
  - **Les mitochondries:** Centrales énergétiques de la cellule, produisent l'ATP. **Les mitochondries ont leur propre ADN**, héritée uniquement par la lignée maternelle.

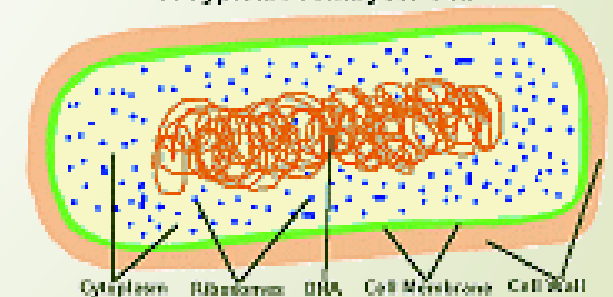
Les ancêtres des mitochondries étaient des cellules indépendantes qui vivaient en symbiose avec l'ancêtre des cellules eucaryotes et qui, avec le temps sont devenues des parties intégrantes des cellules eucaryotes. C'est l'hypothèse de l'*endosymbiose*.

  - **Chloroplastes et plastides:** Autres types d'endosymbiotes qu'on retrouve dans les plantes. Responsables de la photosynthèse.
- **Cellule procaryote:** Pas de noyau, pas d'organelles. ADN libre dans la cellule

A Typical Eukaryote Cell



A Typical Prokaryote Cell



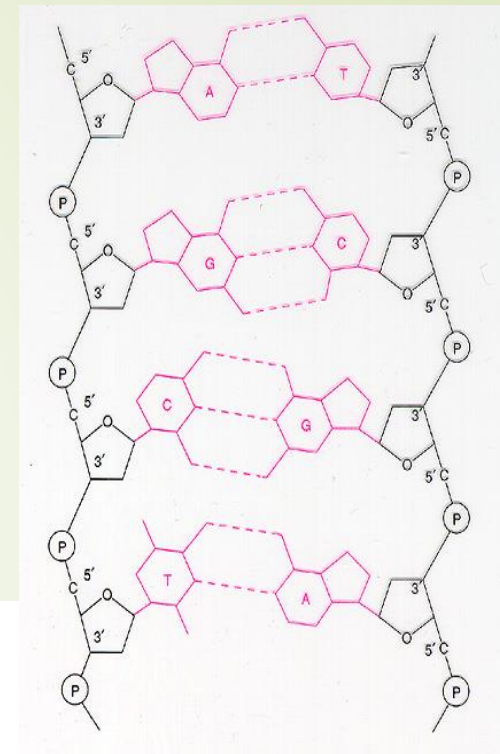


# L'ADN

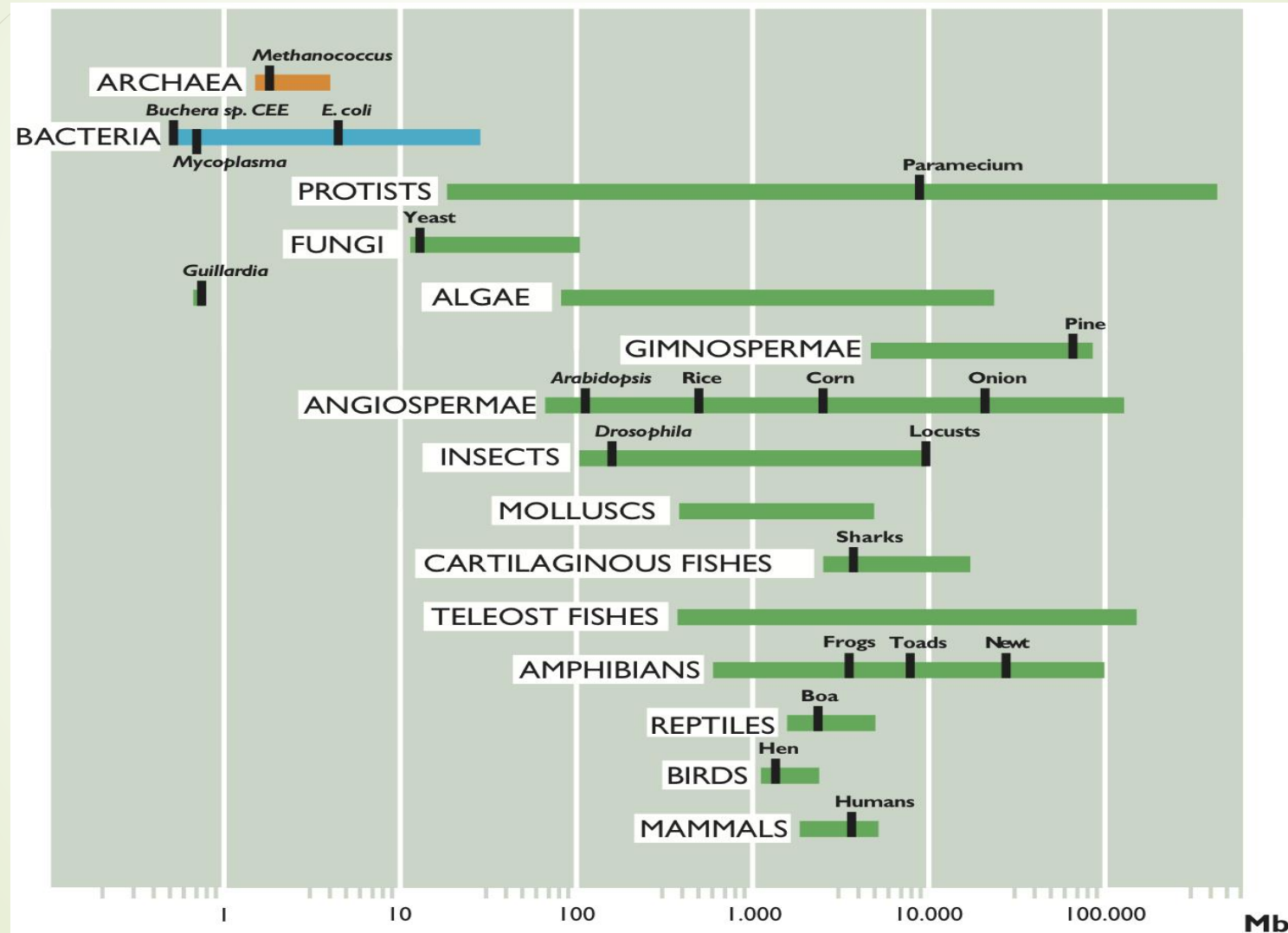
- L'ordre dans lequel les bases nucléotidiques apparaissent dans un polymère est la **séquence du polymère**. Les monomères nucléotidiques sont liés par des liaisons chimiques asymétriques. C'est pourquoi les séquences ont une direction: les deux séquences AGCT et TCGA décrivent deux molécules différentes. Les deux extrémités d'une molécule sont notées 5', 3', et la **lecture se fait toujours de 5' vers 3'**.

```

BPSEUDO.SEQ : SEQUENCE
Position: 1167
1.167kb
10 20 30 40 50 60 70 80
atgctcggaatcaacagcaacatttaactcggttggtcgctcaacagaacctcaacggctcgcaaggcgccctgtcccgaagc 80
gatcaaccgacctgtcgctcgggcaagcgcatacaacagcgcggcggaacgatcgggcggcctcgcgatcgccaccgggatgc 160
aaacgcagatcaacggcctgaaccagggcgtgtcgaaacgcgaaacgacggcgtgtcgatcctgcaaacggcatcgagcggc 240
ctgaacctcgctcaaccaacagcctgcagcgtatccgccagctcgccgtgcaggcctcgaaacggcccctgagcgcgagcga 320
cgcgctcggcgctgcaacaggaagtgcgcagcagatctcggaagtgaaccgtatcgcttcgcagacgaaactacaacggca 400
agaacatctcgacggctcggcaggcacgctgagcttcaggctcggcgcgaaacgtcgggccagacgggtctccgtcgacctc 480
acgcaaacgcatgtcgggggcgaagatcgggcgggcgaatggttcagacggggccagacgctcgggcaacgatcaaggtggcgat 560
cgactcgagcggcgcgccctggctcgctcgggcagcaaccggccaggagacgcagcagatcaacgctcgtgtcggaacggcaagg 640
ggcggcttcacggttcaacgatcagaaacaaccagggcctgtcgctcgacggcctgacccgctggttcggctcgtcgaccgcc 720
ggcaacgggcaacggcggcctcgccgtcgttcagacgctggcgtgtcgacttcggcaaccagcgcgctgttcgcgacggga 800
ccaggcgaacgccacggcgatggttgcgcagatcaacgcggtcaacaagccgcaaacgggtctcgaaacctcgacateagca 880
cgcagacggggcgcgtaccagggcgatggtatcgatcgacaacgcactcgcgacgggtcaacaatctgcaggaacgcctcggc 960
ggcggcgcacaaaccgcttcaaccgcatcgcgacgcagcagcaagccggctcgaaacaacctcgcgcagggcgaatcgcacaa 1040
ccagagcgcggacttgcctcaggaacgcaggaatctctcgagcgcgaagtgtccagcgaacggcgaatctcggtgtctcg 1120
cgcaacgcgaactcgtcgcgcagcgaagtgtcgaagctcctgcaataa 1167
  
```



# Taille des génomes en millions de pb



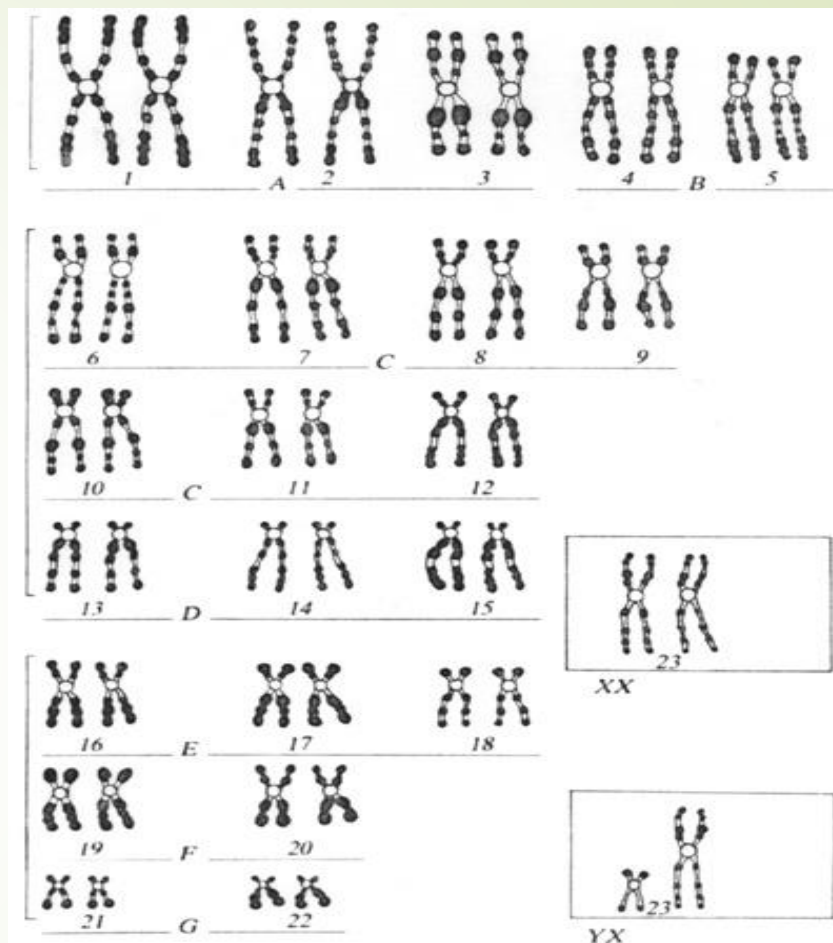


# Les chromosomes

- L'ADN est contenu dans un ou plusieurs **chromosomes**.
- Le génome est l'ensemble des chromosomes
- Chaque chromosome peut-être **linéaire ou circulaire**.
- Pendant la majeure partie du cycle cellulaire, les chromosomes eucaryotes ont une structure filamenteuse, dans laquelle les chromosomes individuels ne peuvent pas être distingués par microscopie optique. Cependant, juste avant la division cellulaire, les filaments chromatiniens y atteignent leur condensation maximale, ce qui rend les chromosomes individuels clairement visibles.
- **Cellule diploïde**: contient deux copies « homologues » de chaque chromosome: une copie provenant de chaque parent. **Le génome nucléaire humain est diploïde.**

**En tout, 46 chromosomes, pour un total d'environ 6,4 milliards de pb.**

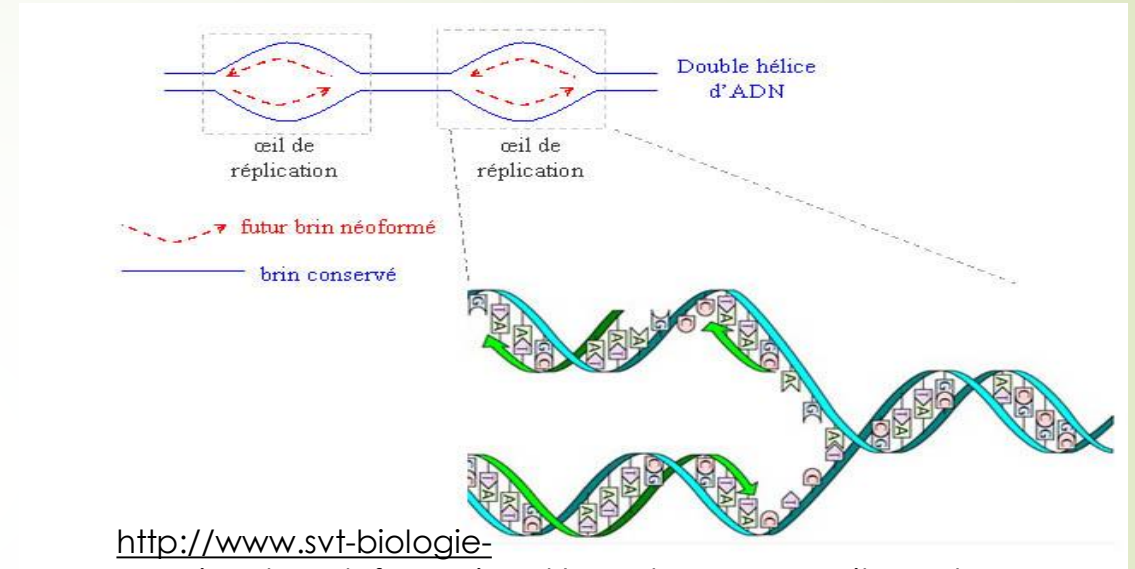
Nadia El-Mabrouk. Bio-informatique génomique.



Caryotype humain: 23 paires de chromosomes homologues. 22 paires d'autosomes et une paire de chromosomes sexuels : XX chez la femme, XY chez l'homme.

# La division cellulaire- La réplication de l'ADN

- La double hélice d'ADN s'ouvre, et chacun des deux brins sert de **matrice** pour le nouveau brin complémentaire, construit par ajout successif de nucléotides selon la complémentarité de séquence.
- La réplication est initiée à des sites appelés **origines de réplication**. Les bactéries ont une seule origine de réplication alors que les eucaryotes en comptent plusieurs. Chez l'humain, plus de 10 000 fourches de réplifications simultanées.
- Réplication assurée par une enzyme appelée l'**ADN polymérase**.

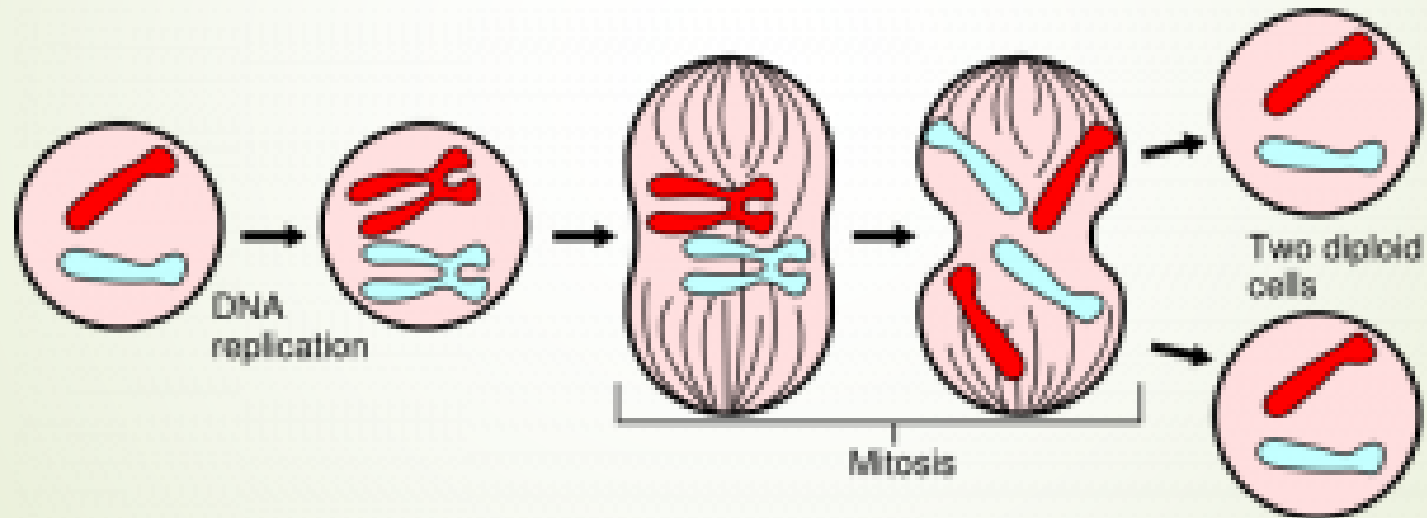


[http://www.svt-biologie-premiere.bacdefrancais.net/morphogenese\\_mitose.php](http://www.svt-biologie-premiere.bacdefrancais.net/morphogenese_mitose.php)

- Objectif du cycle cellulaire : produire deux copies identiques du génome dans les cellules filles. L'incapacité à copier tout l'ADN, la production d'un trop grand nombre de copies de certaines régions ou l'incapacité à séparer le génome de manière égale entre les cellules filles, peut entraîner des maladies (tumeurs cancéreuses, etc).

# La division cellulaire - La mitose

- **La mitose:** Processus de division nucléaire. Un noyau diploïde contenant  $2N$  chromosomes (préalablement répliqués, mais pas séparés), produit deux noyaux diploïdes identiques contenant chacun  $2N$  chromosomes. Suit de la division cellulaire.

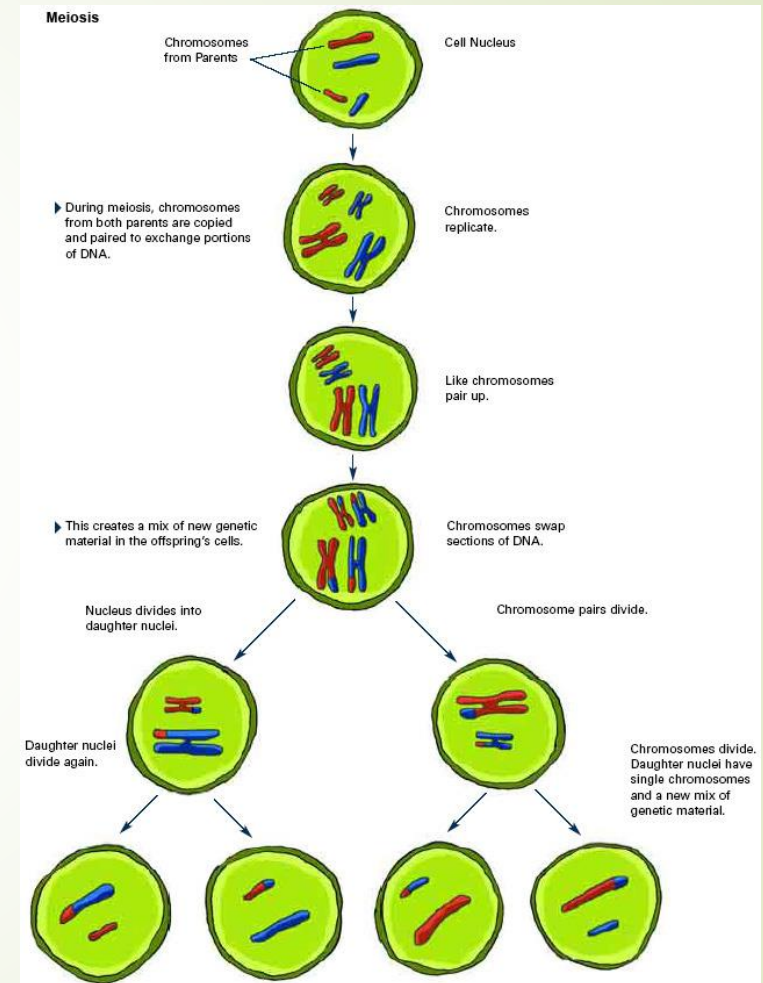


<http://www.wereyouwondering.com/what-is-the-difference-between-mitosis-and-meiosis/>

# La division cellulaire - La méiose

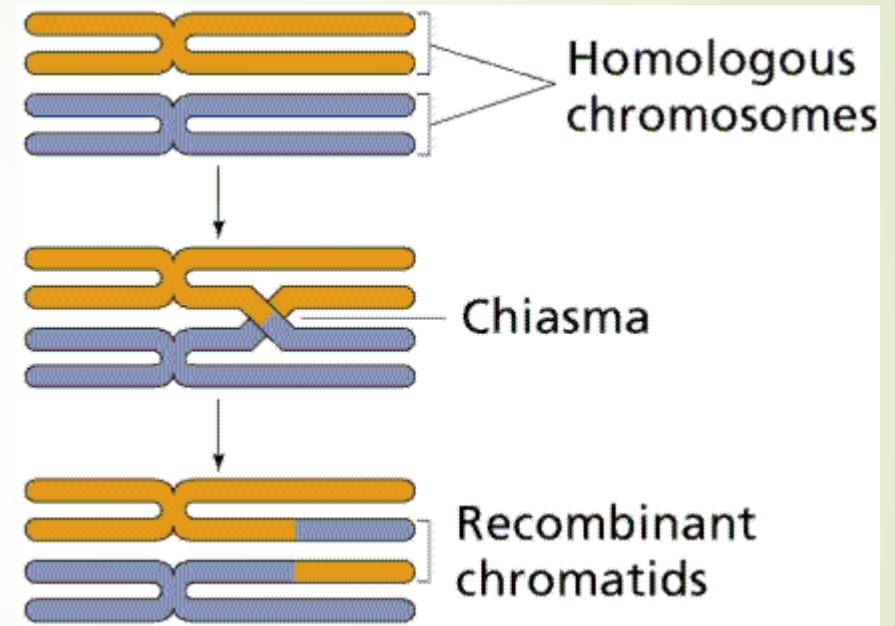
Chez les organismes présentant un dimorphisme sexuel (deux sexes, mâle et femelle):

- **La méiose:** Résulte en la formation de gamètes **haploïdes**. À partir d'un noyau diploïde contenant  $2N$  chromosomes, une réplication et deux divisions nucléaire produisent 4 noyaux haploïdes (**gamètes**) contenant chacun  $N$  chromosomes.
- La méiose, combinée à la fertilisation résulte en la conservation du nombre de chromosomes.



# La division cellulaire- La méiose

- Les chromosomes maternels et paternels ont des histoires évolutives différentes, ont accumulé des mutations différentes.
- La diversité génétique est due à des recombinaisons affectant les chromosomes homologues durant la méiose.



[www.emc.maricopa.edu/faculty/farabee/biobk/biobookmeiosis.html](http://www.emc.maricopa.edu/faculty/farabee/biobk/biobookmeiosis.html)

# Les protéines et les gènes: deux découvertes indépendantes

- ▶ **Biochimie** : Consiste à purifier et à caractériser les composants chimiques responsables de l'exécution d'une fonction particulière. Étude d'un composant purifié d'un organisme.
- ▶ Les organismes vivants sont composés principalement de carbone, d'hydrogène, d'oxygène et d'azote. Ces éléments sont combinés dans des macromolécules complexes: protéines, acides nucléiques, lipides, glucides. De toutes les macromolécules, **les protéines** ont la gamme de fonctions la plus diversifiée. **Assure le fonctionnement de la cellule**. Le corps humain fabrique environ 100 000 protéines distinctes:
  - ▶ **les enzymes**, qui catalysent les réactions chimiques, comme la digestion des aliments
  - ▶ **les molécules structurales**, qui constituent les cheveux, la peau et les parois cellulaires
  - ▶ **les transporteurs de substances** telles que l'hémoglobine, qui transporte l'oxygène dans le sang
  - ▶ **les transporteurs d'informations** tels que les récepteurs à la surface des cellules, l'insuline, les hormones

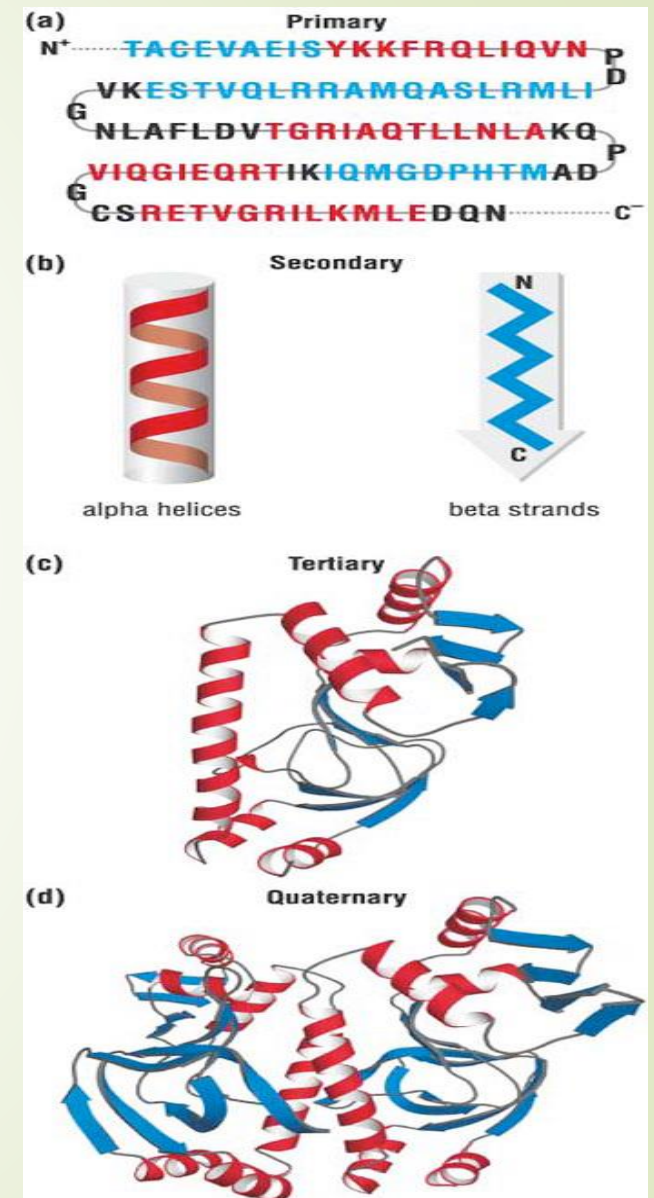
# Les protéines et les gènes: deux découvertes indépendantes

- Une protéine est une macromolécule d'acides aminés.
- Séquences linéaires de 20 (+ 2 rares) acides aminés

*MPKLNSVEGFSSFEDDCPSARGFHHLRTMY...*

- Longueur typique au alentour de 300 AA, étendue: 100-5000 AA
- Structures:
  - Primaire: Séquence linéaire d'AA
  - Structure secondaire: Hélices a et feuilletts b
  - Structure tertiaire: Regroupement en domaines

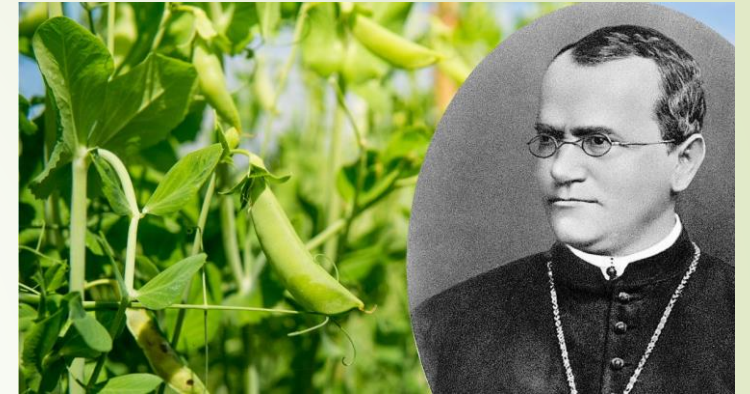
**La structure détermine la fonction d'une protéine**



# Les protéines et les gènes: deux découvertes indépendantes

- ▶ **Génétique** : Les généticiens étudient des organismes mutants qui sont intacts à l'exception d'un seul composant. **Découverte des gènes.**
- ▶ On peut faire remonter la génétique aux expériences pionnières de **Gregor Mendel** en 1865 sur des souches de petits pois. En particulier, il a étudié le trait suivant: avoir des graines rondes ou sinueuses. Mendel croise des souches à graines rondes avec des souches à graines sinueuses. Il note que la génération  $F_1$  est constituée uniquement de pois à graines rondes. En croisant les pois de la génération  $F_1$  entre eux, il obtient 75 % de ronds et 25 % de sinueux. Il déduit de ses expériences que **les traits sont affectés par des facteurs discrets. C'est ce qu'aujourd'hui on appelle des gènes.**

Nadia El-Mabrouk. Bio-informatique génomique.



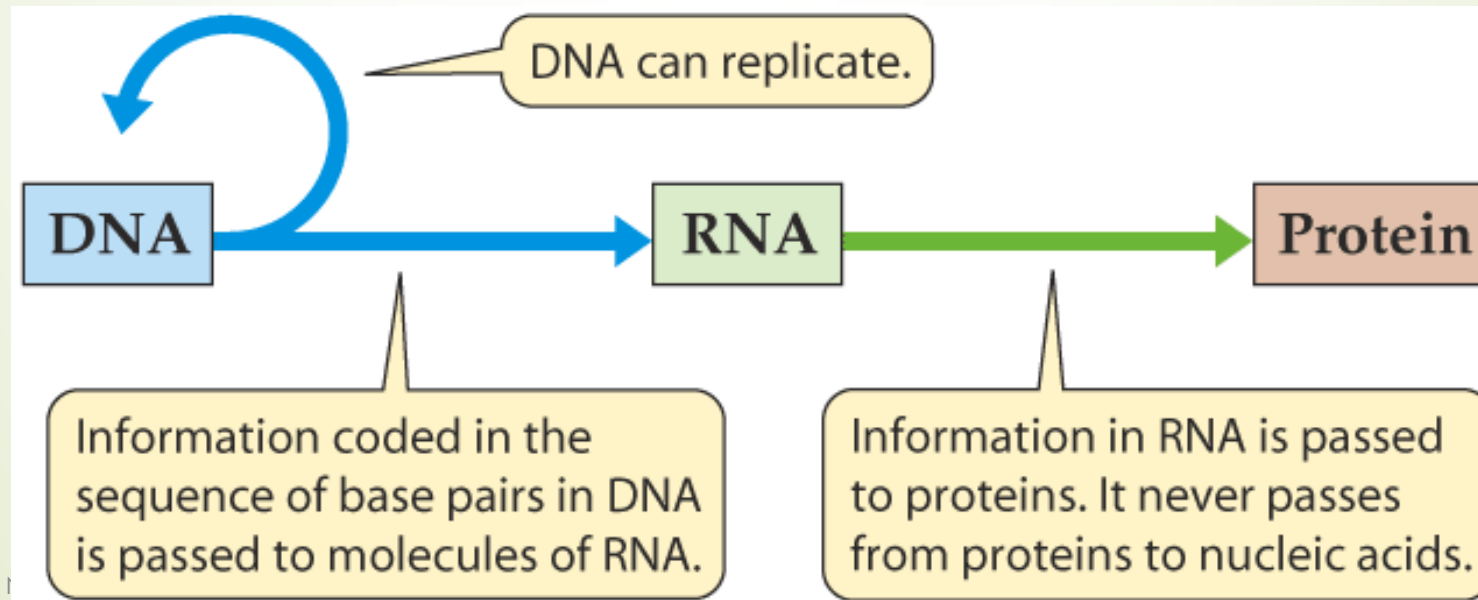
<https://www.thegreatcoursesdaily.com/gregor-mendel-fake-data/>

- ▶ Il a fallu attendre près de 35 ans pour que les biologistes aient une idée de l'endroit où ces gènes hypothétiques se trouvaient dans la cellule (dans les chromosomes) et près de 100 ans pour qu'ils comprennent leur nature biochimique: **Les gènes sont faits d'ADN.**



# Le dogme central de la biologie moléculaire

- ▶ Le processus qui permet d'extraire l'information du génome pour en faire des molécules actives: **c'est le même processus pour TOUS les organismes vivants:**
  - ▶ Code source : ADN
  - ▶ **Transcription** en un code exécutable: l'**ARN messenger**; aussi ARN de transfert, ARN ribosomiques, etc.
  - ▶ **Traduction: Protéines**



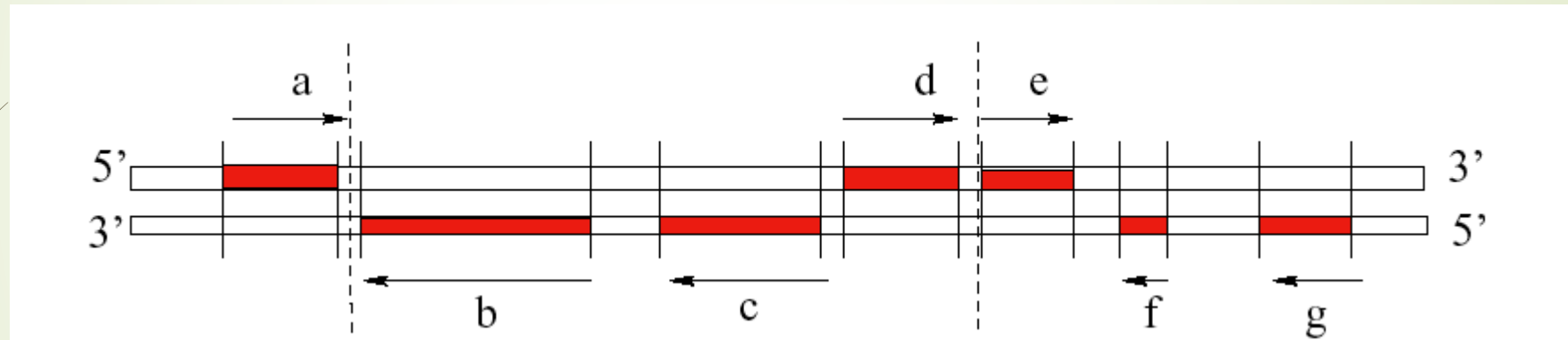
# L'ARN - Acide ribonucléique

- Chaîne d'acide nucléiques dont le sucre est le ribose et les bases possibles sont:
  - Adénine (A), Cytosine (C), Guanine (G), **Uracil (U)**
  - Liens Watson-Crick: **A-U, G-C**. Également fréquent: **G-U**
- Macromolécules qui jouent des rôles essentiels dans la cellule.  
3 types principaux:
  - ARN messenger (mRNA)
  - ARN de transfert (tRNA)
  - ARN ribosomique (rRNA)

# Les gènes

Comment l'alphabet de 4 lettres de l'ADN peut-il encoder les instructions permettant de produire les protéines sur un alphabet de 20 lettres?

- Les gènes correspondent à des segments spécifiques d'ADN le long du chromosome.



- Génotype:** Le structure en nucléotides des gènes
- Phénotype:** Le résultat induit par le gène.

Par exemple, des graines rondes ou sinueuses;  
des yeux verts ou noirs.



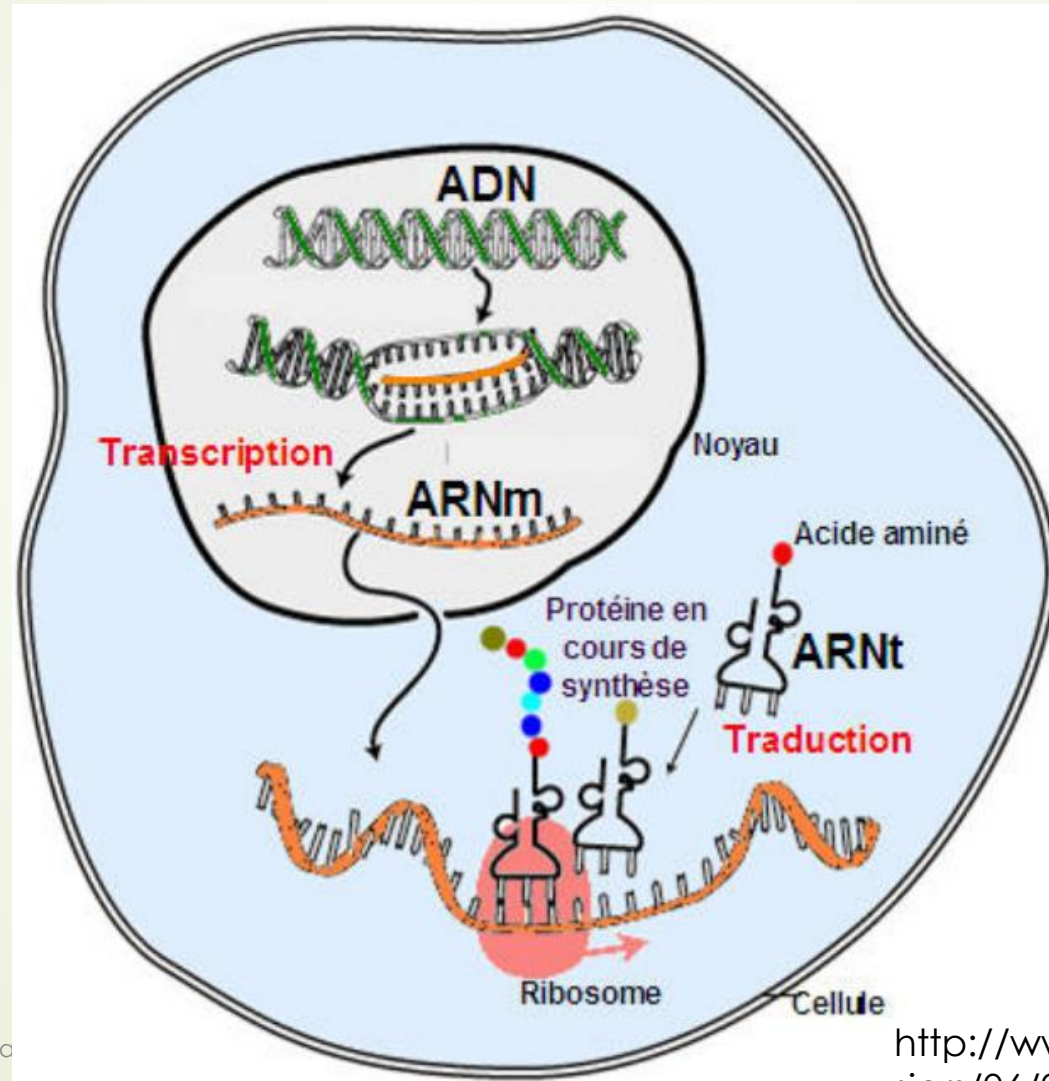
# Les gènes

- Variation génétique=variation allélique
- Un organisme diploïde porte un ou deux allèles pour un même gène.
- Les allèles d'un gène sont souvent désignés par des lettres. Par exemple, l'allèle codant pour la capacité de synthèse de la mélanine dans le cas de l'albinisme est appelé A (**allèle dominant**) et l'allèle codant l'incapacité correspondante (qui aboutit à l'albinisme) est désigné par a (**allèle récessif**).
  - A/A=peau pigmentée
  - A/a=peau pigmentée
  - a/a=peau non-pigmentée (albinisme)



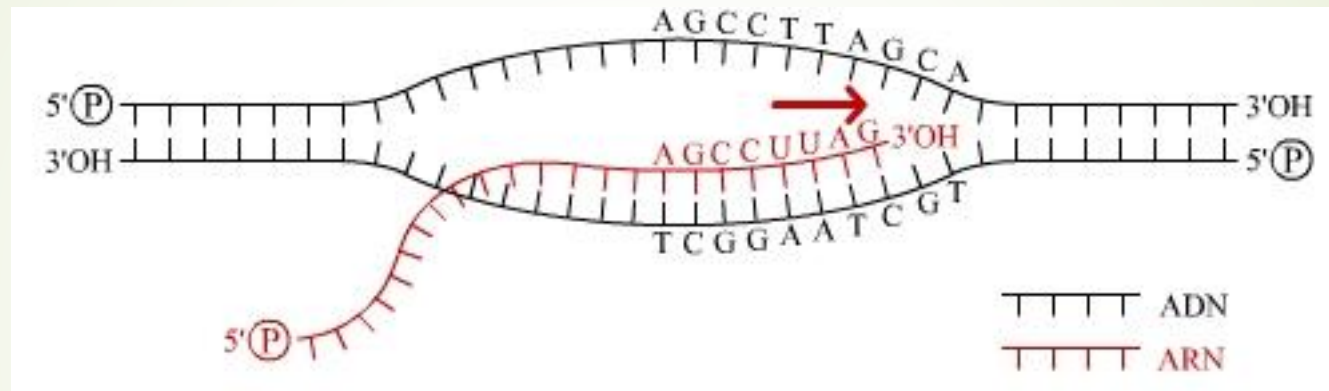
Un albinos. Le phénotype résulte de l'**homozygotie d'un allèle récessif a/a**

# Transcription et traduction



# Transcription de l'ADN en ARNm

- La transcription est la **production d'un transcrit d'ARN (ARNm) à partir d'un gène.**

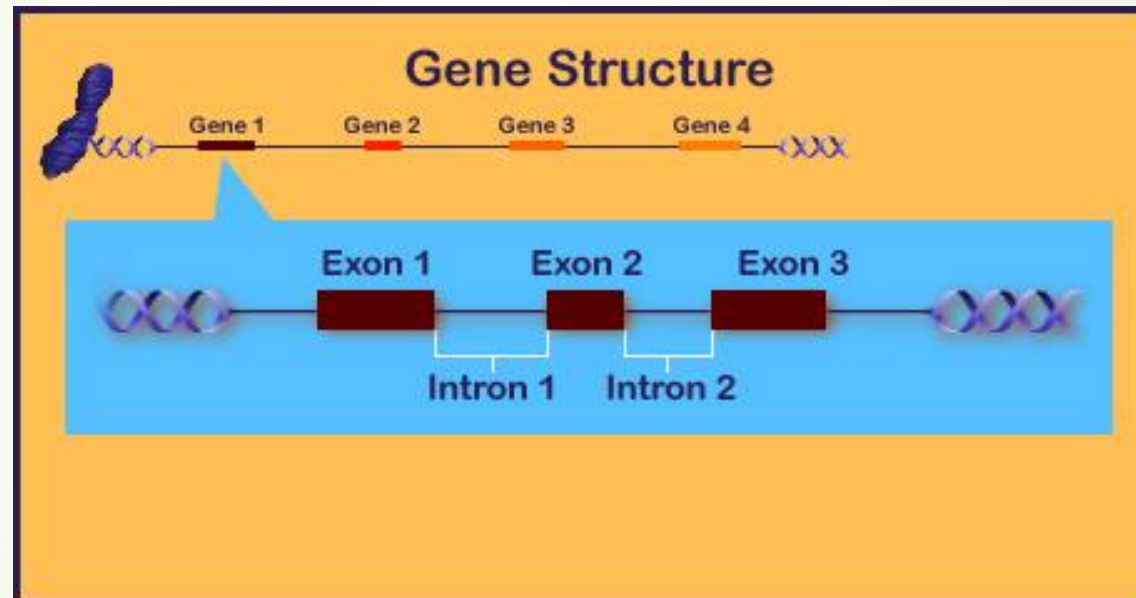


<https://planet-vie.ens.fr/thematiques/cellules-et-molecules/physiologie-cellulaire/la-transcription-chez-les-eucaryotes>

- Effectuée par une enzyme appelée **ARN polymérase** (RNAPol). La RNAPol se lie à la région promotrice du gène par le biais d'interactions avec un ensemble de facteurs de transcription. Après l'**initiation**, les transcrits subissent une **élongation**, après quoi le processus se **termine**.
- L'ARNm produit est simple brin.

# Transcription de l'ADN en ARNm

- Pour les gènes eucaryotes, l'ARNm est ensuite **épissé**: élimination des **introns** du transcrit d'ARN initial. Les séquences restantes sont appelées **exons**.



- **Épissage alternatif** : Un même gène peut produire plusieurs ARNm différents impliquant différents ensembles d'exons. 60 % des gènes humains sont soumis à un épissage alternatif.

# Traduction de l'ARNm en protéine

- ▶ 4 nucléotides vs. 20 acides aminés
  - ▶ Code en triplet: 3 nt (1 codon) = 1 aa
  - ▶ 3 codons stop qui marquent la fin de la traduction
- ▶  $4 \times 4 \times 4 = 64$  codons distincts.
- ▶ Des codons qui encodent le même acide aminé sont appelés **codons synonymes**.
- ▶ Le code génétique peut varier selon les organismes.

Nadia El-Mabrouk. Bio-informatique génomique.

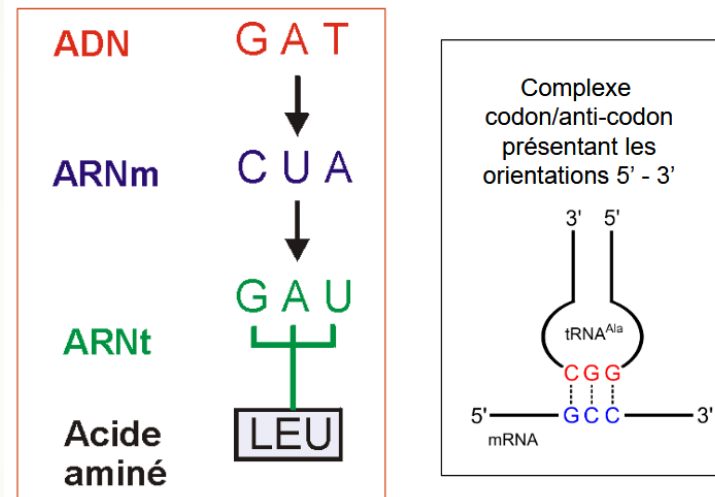
le code génétique									
	Deuxième lettre								ijk
	U		C		A		G		
U	UUU	Phe	UCU	Ser	UAU	Tyr	UGU	Cys	U
	UUC	Phe	UCC	Ser	UAC	Tyr	UGC	Cys	C
	UUA	Leu	UCA	Ser	UAA	Stop	UGA	Stop	A
	UUG	Leu	UCG	Ser	UAG	Stop	UGG	Trp	G
C	CUU	Leu	CCU	Pro	CAU	His	CGU	Arg	U
	CUC	Leu	CCC	Pro	CAC	His	CGC	Arg	C
	CUA	Leu	CCA	Pro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

Troisième lettre (côté 3')



# Traduction de l'ARNm en protéine

- Différents codons sont reconnus par les anticodons de différents ARN de transfert (ARNt)



[http://www8.umoncton.ca/umcm-filion\\_martin/cours/genetique/Chapitre%201.pdf](http://www8.umoncton.ca/umcm-filion_martin/cours/genetique/Chapitre%201.pdf)

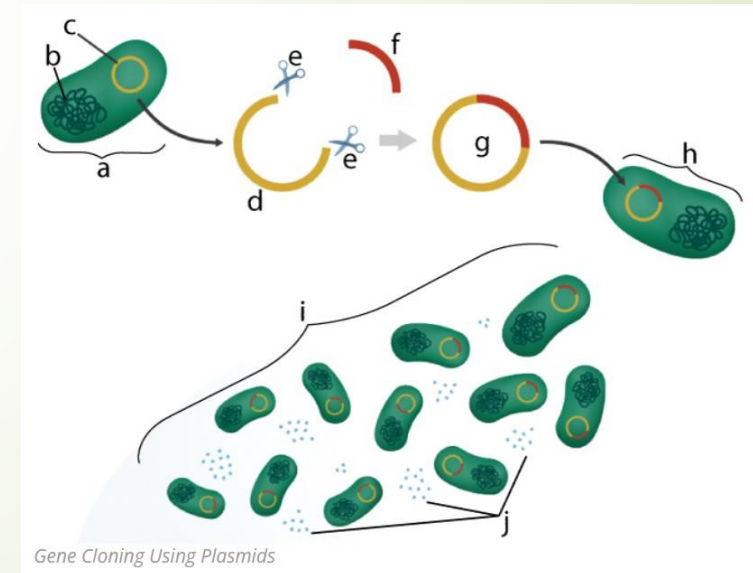
- Un **cadre de lecture** est une région de l'ADN d'une taille suffisante (ex. 90 nuc.) située entre un codon START (ou Méthionine) et un codon STOP.
- La première base à laquelle commence la traduction détermine le **cadre de lecture**.

5' - ... **AUG** ACC GAU GAC CCU ... **UAG** ... - 3'

# La révolution de l'ADN recombinant

- ▶ En 1965, la biologie moléculaire avait élucidé les secrets fondamentaux de la vie. Cependant, sans la possibilité de manipuler les gènes, la compréhension était plus théorique qu'opérationnelle. Dans les années 1970, cette situation se transforme par **la découverte d'une variété d'enzymes fabriquées par des bactéries qui permettaient de manipuler l'ADN à volonté.**
- ▶ Les bactéries fabriquent des **enzymes de restriction** qui coupent l'ADN à des séquences spécifiques (défense contre les virus), et des **ligases** qui relient les fragments d'ADN.
- ▶ Il est devenu possible de **couper et de coller des fragments d'ADN** à volonté et de les introduire dans des cellules vivantes. Ces expériences de **clonage** permettent d'étudier les gènes individuels.

- ▶ Insérer un gène (ou tout autre morceau d'ADN) dans un plasmide. Cela crée un plasmide contenant de l'ADN recombinant. Ce plasmide recombinant est ensuite inséré dans une cellule bactérienne. Chaque fois que la cellule se reproduit, elle réplique également le plasmide recombinant.



Gene Cloning Using Plasmids

# Séquençage de l'ADN

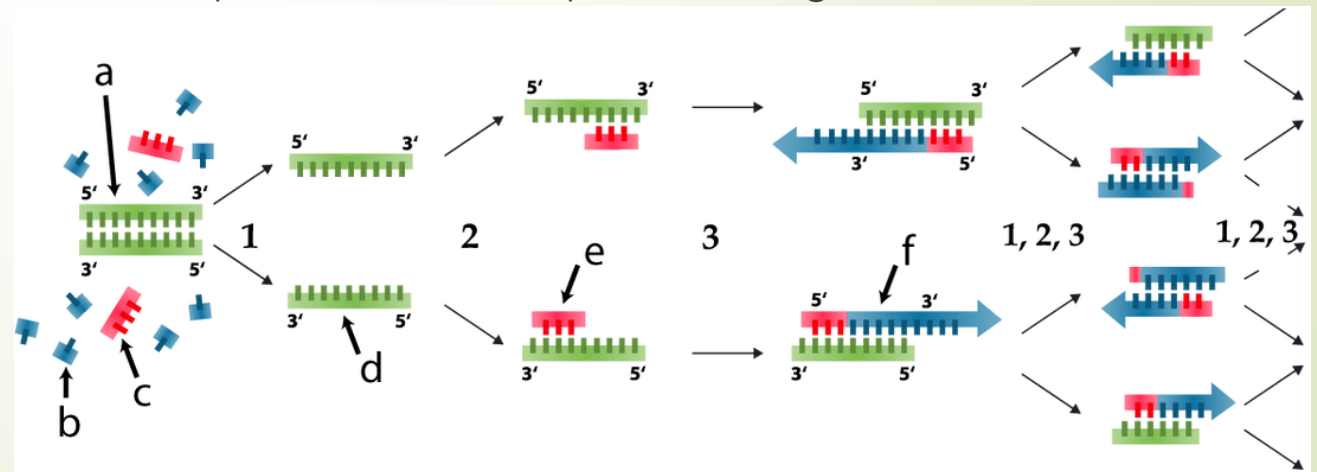
L'invention du séquençage de l'ADN a constitué une avancée majeure.

- A révélé des similitudes frappantes entre des êtres vivants aussi divers que les humains et les levures.
- A conduit à une explosion de l'information en biologie, les bases de données publiques continuant à se développer à un rythme exponentiel.

<https://sciencemusicvideos.com/ap-biology/genetic-engineering-and-biotechnology/polymerase-chain-reaction-pcr-tutorial/>

Nadia El-Mabrouk. Bio-informatique génomique.

- Une nouvelle technique puissante « **Polymerase Chain Reaction** » ou réaction de polymérisation en chaîne (**PCR**) permet d'amplifier une séquence d'ADN sans avoir recours au clonage. On utilise de courtes molécules d'ADN appelées **amorces**, « **primers** » (longueur d'environ 20 bases) qui sont complémentaires des séquences flanquant la région d'intérêt. Chaque amorce peut s'apparier avec une base dans la région complémentaire et est ensuite étendue pour contenir la séquence complète de la région en utilisant l'enzyme **ADN polymérase**. De cette façon, une seule copie de la région donne lieu à deux copies. En répétant cette étape  $n$  fois, on peut obtenir  $2^n$  copies de la région.



# Un exemple de la nouvelle puissance préférée par les outils technologiques : Étude de **La fibrose kystique**

La mucoviscidose est une maladie **récessive** (comme les pois à graines sinueuses de Mandel): si deux porteurs non affectés du gène récessif  $a$  de la mucoviscidose (c'est-à-dire des hétérozygotes de génotype  $Aa$ ) se marient, un quart de leur progéniture sera affectée (c'est-à-dire aura le génotype  $aa$ ).

Mais la base moléculaire de la maladie est restée un mystère jusqu'en 1989.

- 1) **Cartographie génétique** (en 1985): On observe que le profil héréditaire de la maladie est étroitement corrélé au profil d'hérédité d'un polymorphisme particulier de l'ADN (i.e. une variation orthographique commune de l'ADN) situé sur le chromosome 7. Cela indique que **le polymorphisme est situé près du site du gène de la maladie.**
- 2) **Cartographie physique et séquençage de l'ADN du gène de la FK** : On part de la position connue du polymorphisme et on isole séquentiellement les fragments adjacents dans un processus appelé marche chromosomique, jusqu'à ce que le gène de la maladie soit atteint. Ensuite on séquence le gène.

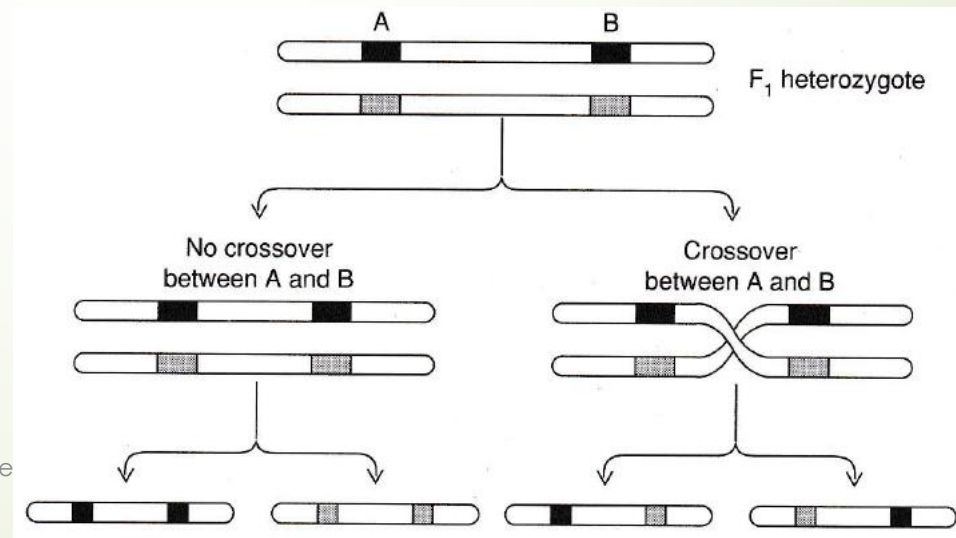
Gène de la FK code pour une protéine de 1480 acides aminés. Anomalie: **suppression d'un codon de la phénylalanine en position 508 de la protéine.** Il est ainsi devenu possible d'effectuer un diagnostic ADN sur des individus pour voir s'ils étaient porteurs de la mutation commune de la FK.

## Un exemple de la nouvelle puissance préférée par les outils technologiques : Étude de **La fibrose kystique**

- 3) **En comparant la séquence de la protéine aux bases de données publiques**, on a constaté qu'elle présentait de fortes similitudes avec une classe de protéines (transporteurs d'ATP qui résident dans la membrane cellulaire). Et donc, par analogie, il a été possible de déduire une forme tridimensionnelle probable pour la protéine de FC. Ainsi, **l'analyse bio-informatique de la séquence a permis d'élucider la structure et la fonction de ce gène.**
- 4) Avec l'avènement de la thérapie génique (possibilité d'utiliser un virus pour implanter une copie fonctionnelle d'un gène dans des cellules portant une version défectueuse), des essais cliniques ont été lancés pour tenter de guérir la maladie dans les cellules pulmonaires des patients atteints de fibrose kystique.

# Cartographie génétique

- ▶ Basé sur l'idée contrintuitive qu'**on peut trouver où est situé un gène sans connaître le gène**. Possible d'identifier l'emplacement d'un gène inconnu causant une maladie en corrélant le modèle d'hérédité de la maladie avec le modèle d'hérédité de marqueurs génétiques connus.
- ▶ La cartographie génétique repose sur la reconnaissance du fait que **la fréquence de recombinaison T entre deux gènes fournit une mesure de la distance qui les sépare**: Si deux gènes sont proches, T sera faible. S'ils sont plus éloignés, T sera grande. Si la fréquence de recombinaison est nettement inférieure à 0,5, on dit que les **gènes sont liés (linked)**.



# Conclusion

- ▶ Cette introduction retrace les grandes découvertes en **biochimie, génétique et biologie moléculaire**, qui ont permis de révéler les processus biochimiques fondamentaux de la vie.
- ▶ Les implications scientifiques et pratiques sont sans précédent : L'histoire de l'évolution de la vie peut être lue à partir des séquences d'ADN; le bétail et les cultures peuvent être modifiés pour de meilleurs rendements; des gènes défectueux peuvent être remplacés par des gènes fonctionnels afin de guérir des maladies graves, etc.
- ▶ La vie sur terre est soumise à des **processus très similaires**, ce qui n'empêche pas **l'énorme diversité des organismes vivants** et des variations génomiques.
- ▶ L'élucidation des mécanismes et des processus qui sous-tendent cette diversité a été, et reste, l'un des principaux objectifs de la recherche biologique.
- ▶ La **bio-informatique** est née du besoin de faire appel à des méthodes mathématiques et informatiques afin de pouvoir organiser, modéliser et analyser les données biologiques.

# Références

1. « Calculating the secrets of life », *E.S. Lander, M.S. Waterman editors*, National Academy Press, 1995.
2. « Evolutionary genomics, Statistical and computational methods », *M. Anisimova editor*, Humana Press, 2012.
3. [http://www8.umoncton.ca/umcm-filion\\_martin/cours/genetique/Chapitre%201.pdf](http://www8.umoncton.ca/umcm-filion_martin/cours/genetique/Chapitre%201.pdf)  
(consulté le 18 juin 2021).