

1 ère partie

**Du génome aux phénotypes :
comment les caractères héréditaires sont-ils contrôlés ?**

ou

le gène est une unité de fonction qui peut muter.

Chapitre 2 : Les facteurs imaginés par Mendel sont constitués d' ADN et sont maintenant appelés gènes. Leur nature moléculaire permet la stabilité au cours des générations et le contrôle de la synthèse des polypeptides, via la transcription et le code génétique.

1.L'ADN programme les êtres vivants :c'est le matériel génétique.

1.1. la transformation du pneumocoque.

1.1.1. Les pneumocoques et les souris (Griffith 1928).

1.1.2. L'agent transformant est l' ADN (Avery, Mac Carthy, MacLeod 1944).

1.2 . généralisation.

1.3. hypothèse : les « facteurs » de Mendel sont faits d' ADN .

2. La structure de l' ADN .

2.1. la séquence des bases d'un polynucléotide paraît quelconque.

2.1.1. les constituants sont le phosphate, un sucre et quatre bases.

2.1.2 le polynucléotide possède un contenu informatif qui peut être considérable.

2.2. l' ADN des êtres vivants est double à l'état natif.

2.2.1. Le modèle de Watson et Crick.

2.2.2. Une conséquence technique : dénaturation- renaturation.

3. Les propriétés de l' ADN sont dues à sa structure.

3.1. la réplication semi conservative assure la continuité du message.

**3.2. la transcription et la traduction font passer du message génétique
aux polypeptides, acteurs du métabolisme et de la structure des cellules .**

**4. Conclusion : le gène peut être défini à partir de sa nature
et de sa structure moléculaires, car ses propriétés en découlent.**

Chapitre 2 : Les facteurs imaginés par Mendel sont constitués d' ADN et sont maintenant appelés gènes. Leur nature moléculaire permet la stabilité au cours des générations et le contrôle de la synthèse des polypeptides, via la transcription et le code génétique.

Les travaux de Mendel doivent être considérés avec le plus grand intérêt, comme nous venons de le faire dans le chapitre précédent. Cependant, l'essor véritable de la génétique ne date pas de ce précurseur (1) : la discipline naît réellement avec le XX^{ème} siècle. Plus ou moins indépendamment, en 1900, De Vries, Correns et Van Seysenegg font l'hypothèse selon laquelle chaque cellule sexuelle porterait un seul déterminant pour chaque caractère (2) : on reconnaît dans cette proposition l'une des idées de Mendel.

En 1909, Johannsen propose de nommer *gène* un tel déterminant, et c'est Bateson qui crée le terme de *génétique* pour la nouvelle discipline. Ensuite, pendant une cinquantaine d'années, la génétique se développe d'une manière assez étonnante : on définit de nombreuses propriétés des gènes mais l'on ne connaît pas leur nature. Les généticiens ont coutume de dire que le gène en est alors au stade de la « **boîte noire** ». Il quitte ce statut à la suite de deux séries de travaux que nous allons étudier ici.

Nous allons tout d'abord décrire les expériences permettant de passer de l'idée abstraite des « facteurs » de Mendel à une réalité chimique (1944). Puis nous présenterons un condensé des propriétés moléculaires des gènes et de leurs fonctions (3), de 1953 aux années 1960. **Tout cela permettra de donner une première définition du gène.**

1. L'ADN programme les êtres vivants : c'est le matériel génétique.

Admettons avec Mendel que des « facteurs » contrôlent les caractères et qu'ils sont transmis de génération en génération. Nous nommerons « **matériel génétique** » l'ensemble de ces agents assurant le contrôle du métabolisme et la reproduction quasi à l'identité (chapitre 1). Les expériences permettant d'envisager la nature de ce matériel génétique sont déjà anciennes.

1.1. La transformation du pneumocoque

1.1.1. Les pneumocoques et les souris (Griffith 1928).

Les pneumocoques sont des bactéries qui sont responsables de la pneumonie chez les mammifères, lorsqu'ils sont **virulents**. Il existe également des pneumocoques **non virulents**, qui sont inoffensifs (figure 5).

Griffith constate que la différence entre les deux types est **héréditaire**. Il observe que les pneumocoques virulents, notés S, ont une **capsule polysaccharidique** épaisse alors qu'elle est pratiquement inexistante chez les pneumocoques non virulents, notés R (4).

Il injecte à des souris **un mélange de R vivants et de S tués** par la chaleur. Les souris ainsi traitées meurent de **pneumonie**. Leur autopsie montre qu'elles possèdent des R et **des S vivants (figure 6)**.

Griffith conclut qu'il y a eu **transformation** des R, qui ont reçu *quelque chose* des S tués : il définit ainsi un **agent transformant**.

Cette transformation est **héréditaire** : les S apparus à la suite de la transformation donnent des S possédant des capsules et qui sont virulents

Figure 5 : deux types de pneumocoques et leurs conséquences chez la souris (5).

L'aspect des bactéries, vu au microscope, est indiqué dans le cercle.

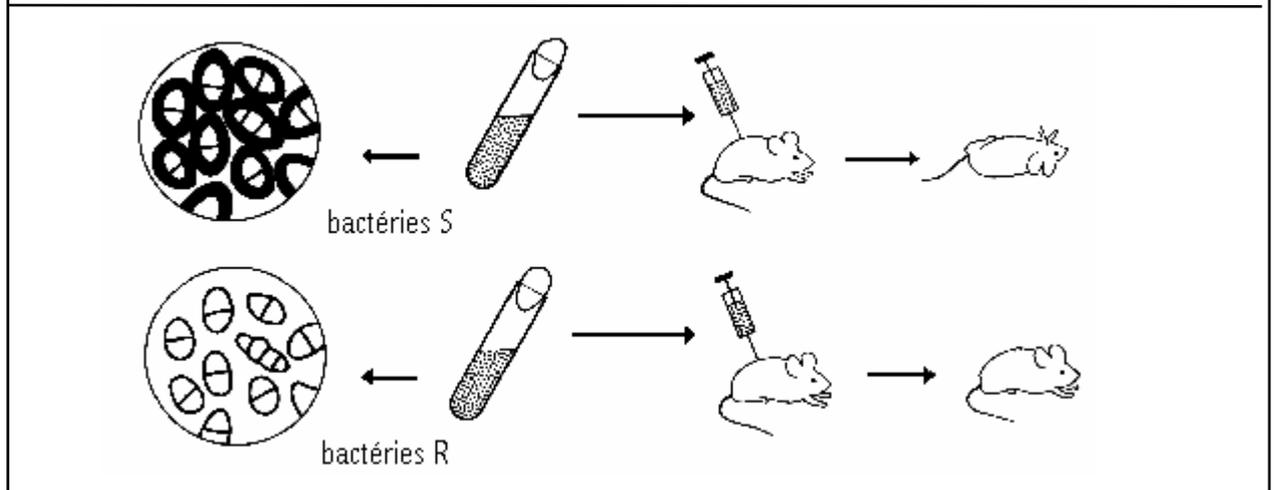
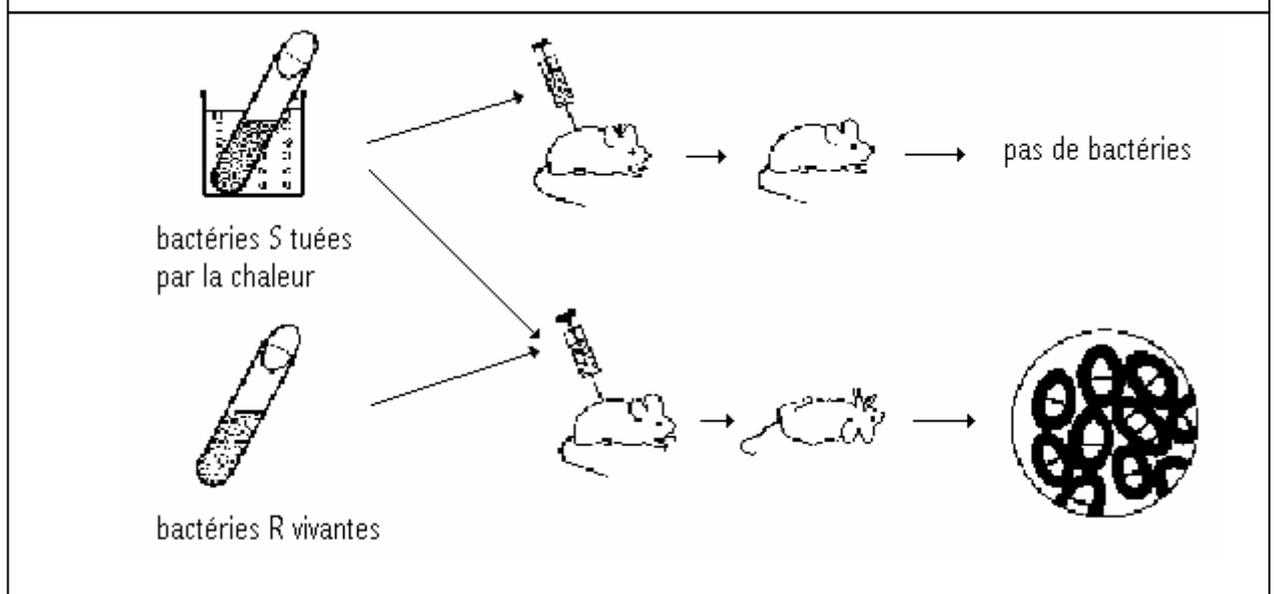


Figure 6 : transformation de bactéries R non virulentes en bactéries virulentes S (5).



(1) : sur « l'oubli » de Mendel, voir chapitre 1 , note 15.

(2) : attention, nous l'avons déjà dit (chapitre 1, note 14), cette définition « historique » n'est pas entièrement correcte.

(3) : pour plus de précision, voir un cours de biochimie. Si l'on veut atteindre un niveau plus élevé, consulter un classique : *Biologie moléculaire de la cellule*, Boeck University , en réclamant la dernière version de l'ouvrage.

(4) : S= smooth = lisse ; R = rough = rugueux. Cette nomenclature correspond à l'aspect des colonies provenant de bactéries cultivées sur des boîtes de pétri. Les R sont facilement phagocytés par les souris alors que les S le sont difficilement à cause de leur capsule.

(5) : les figures 5 et 6 ont été modifiées de Prévost, Hermann 1967. Nous rendons ainsi hommage à la mémoire de l'auteur : il y a trente ans il a permis à la génétique d'entrer dans les disciplines de base du premier cycle de l'Université devenue Paris 6 , puis dans les autres universités françaises. Pour la petite histoire, cela veut dire qu'il a fallu attendre presque 15 ans après Watson et Crick pour que cela se produise... Heureusement l'Université française ne refuse plus les progrès de la biologie : elle y participe même très activement !!

1. 1. 2. L'agent transformant est l' ADN (Avery, Mac Carthy, MacLeod 1944).

Les expériences sont ensuite simplifiées. La transformation est réussie *in vitro* (6), en cultivant des R en présence de S tués : cela élimine le paramètre « souris » de l'analyse des expériences. Puis on réalise la transformation en cultivant des R en présence d'un extrait acellulaire (7) de S tués : la réussite de l'expérience montre que l'intégrité cellulaire est inutile. On en conclut que **l'agent transformant est de nature chimique.**

On s'en assure en fractionnant l'extrait acellulaire puis en purifiant l'agent transformant : on montre successivement qu'il est de nature **macromoléculaire** (8), qu'il n'est constitué ni de polysaccharides, ni de protéines, ni d'acide ribonucléique (ARN). Par contre, une fraction ne contenant que l'acide désoxyribonucléique (**ADN**) est capable de réaliser la transformation.

Cette conclusion doit cependant être confirmée, car lorsque l'on effectue une purification, il est possible que le produit obtenu comporte des impuretés. Afin de s'assurer que ces éventuelles impuretés ne sont pas responsables de la transformation l'expérience suivante est réalisée.

Un extrait acellulaire est partagé en trois lots. Le premier est traité par une protéase, enzyme qui hydrolyse spécifiquement les protéines, le deuxième par une ribonucléase (ARN-ase) qui hydrolyse spécifiquement les ARN. Ces deux premiers extraits ainsi traités permettent la transformation.

Le troisième lot est traité par une désoxyribonucléase (**ADN-ase**) qui hydrolyse spécifiquement les ADN. Cette fois, **la transformation n'a pas lieu.**

L'activité enzymatique étant hautement spécifique on en conclut que les protéines et l'ARN ne sont pas impliqués dans la transformation. **La transformation héréditaire est une propriété de l'ADN puisqu'elle ne se produit plus lorsque cette molécule est spécifiquement dégradée** (figure 7).

Cette série d'expériences permet d'énoncer deux conclusions:

- l'ADN de S permet à lui seul de provoquer la synthèse de polysaccharides de capsule chez des bactéries R qui en étaient incapables : donc **l'ADN programme le métabolisme.**

- les bactéries R transformées en S ont pour descendance des bactéries S. La transformation est héréditaire : donc **l'ADN contrôle la multiplication à l'identique (9).**

Les **deux caractéristiques** des êtres vivants que nous analysons depuis le chapitre 1 (le métabolisme, la multiplication quasi à l'identique) sont donc contrôlées par **un seul type moléculaire (10), l'ADN.**

1. 2. Généralisation

Le seul reproche que l'on puisse faire à la démonstration d'Avery est qu'elle concerne un seul caractère, chez un seul organisme. Au cours du demi siècle qui suit l'expérience sur le pneumocoque, la transformation est réussie chez de nombreuses bactéries, pour des caractères très variés. Puis elle est réalisée chez des organismes supérieurs, des champignons jusqu'à la souris en passant par la mouche du vinaigre, la drosophile (11). L'observation de la transformation chez le pneumocoque est donc généralisée, ce qui permet de conclure définitivement que

l'ADN remplit les deux fonctions du matériel génétique, le contrôle du métabolisme et la continuité au cours des générations

1.3. Hypothèse : les « facteurs » de Mendel et l'ADN.

En deux séries d'expériences, nous avons progressé de manière spectaculaire : les expériences de Mendel indiquent que les **caractères** sont contrôlés par des **facteurs** de nature inconnue ; les expériences d'Avery montrent que les caractères sont contrôlés par l'ADN de manière héréditaire. Le rapprochement entre les deux conclusions faites par Mendel puis par Avery est plus que tentant ! Faisons donc une **hypothèse** qui relie l'ensemble des expériences :

les facteurs mendéliens sont constitués d'ADN !

Nous allons voir que cette hypothèse est confirmée par l'étude de la structure et des fonctions de la molécule.

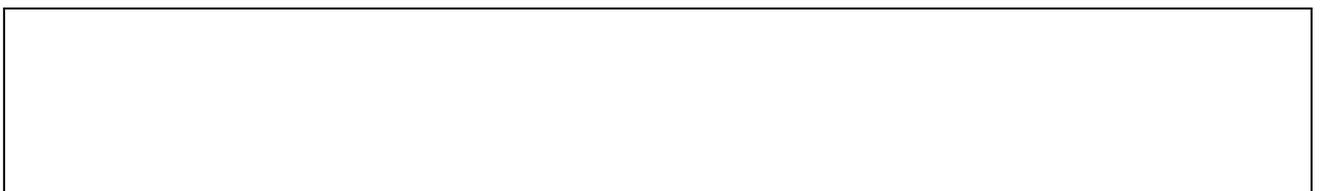


Figure 7 : mise en évidence de la nature du principe transformant , grâce à la spécificité enzymatique.

lot n°1	lot n°2	lot n°3
+protéase	+ARN-ase	+ ADN-ase
il y a transformation	il y a transformation	il n'y a pas de transformation

Encart 3 : deux remarques à propos des expériences de transformation chez le pneumocoque

1. La démonstration historique de Avery, Mac Carthy, MacLeod est totalement démonstrative : pourtant le rôle de l'ADN comme matériel génétique n'est vraiment accepté par la communauté scientifique que lorsque Watson et Crick proposent leur modèle célèbre.

C'est un deuxième exemple (après le cas Mendel) d'un décalage entre une conclusion parfaitement fondée et son acceptation dans un contexte non préparé à l'accepter et à en profiter.

2. Ces expériences illustrent parfaitement la manière de travailler des biologistes : ils utilisent des méthodes et des raisonnements qui ne se distinguent pas de ceux des autres sciences exactes.

(6) : « in vitro » : en dehors de l'organisme. S'oppose à « in vivo » : chez l'organisme vivant.

(7) : un extrait est appelé acellulaire, lorsqu'il ne contient pas de cellule entière.

(8) : les macromolécules sont des molécules de grande taille. Chez les êtres vivants, il s'agit des polysaccharides, des protéines et des acides nucléiques. Ces macromolécules sont constituées d'un grand nombre d'éléments plus simples, respectivement ici de monosaccharides (sucres simples), d'acides aminés ou de nucléotides. Voir un cours de biochimie.

(9) : il ne s'agit pas d'un caractère acquis de manière transitoire. La transformation par l'ADN est ici héréditaire.

(10) : ce résultat est spectaculaire, voire inattendu : on aurait très bien pu penser que chacune des deux fonctions soit déterminée par deux types moléculaires différents. Si l'on fait de la fiction, on peut imaginer que Mendel n'aurait cependant pas été surpris par le résultat, puisque les facteurs qu'il a définis contrôlent et transmettent les caractères. Même si l'on ne doit pas réinterpréter les résultats anciens avec les connaissances actuelles, tout cela reste cohérent.

(11) : la transformation par l'ADN est même possible entre espèces différentes, très éloignées : par exemple Escherichia Coli peut fabriquer des protéines « humaines », lorsqu'on transforme ces bactéries par de l'ADN humain (voir 3^{ème} partie).

2. Structure de l'ADN.

2.1. La succession des bases d'un polynucléotide paraît quelconque.

2.1.1. les constituants élémentaires sont le phosphate, un sucre et quatre bases, agencés en 4 monomères.

↓

Lorsqu'on étudie un ADN chauffé au delà de 80 degrés (12), on observe de longues molécules. Un traitement chimique approprié de ces molécules libère quatre types de constituants que l'on appelle des **nucléotides monophosphates**. Un nucléotide mono phosphate est lui-même séparable en trois éléments, un groupe **phosphate**, un **sucre** (le désoxyribose, d'où le nom d'acide **désoxyribonucléique** pour l'ADN) et une **base azotée** (figure 8).

On nomme **monomères** les nucléotides et **polymère** la molécule qui leur a donné naissance.

L'examen des bases azotées des monomères (figure 9) montre qu'elles se ressemblent deux à deux : les purines (**adénine et guanine**) et les pyrimidines (**cytosine et thymine**).

Figure 8 : les six constituants de l'ADN.

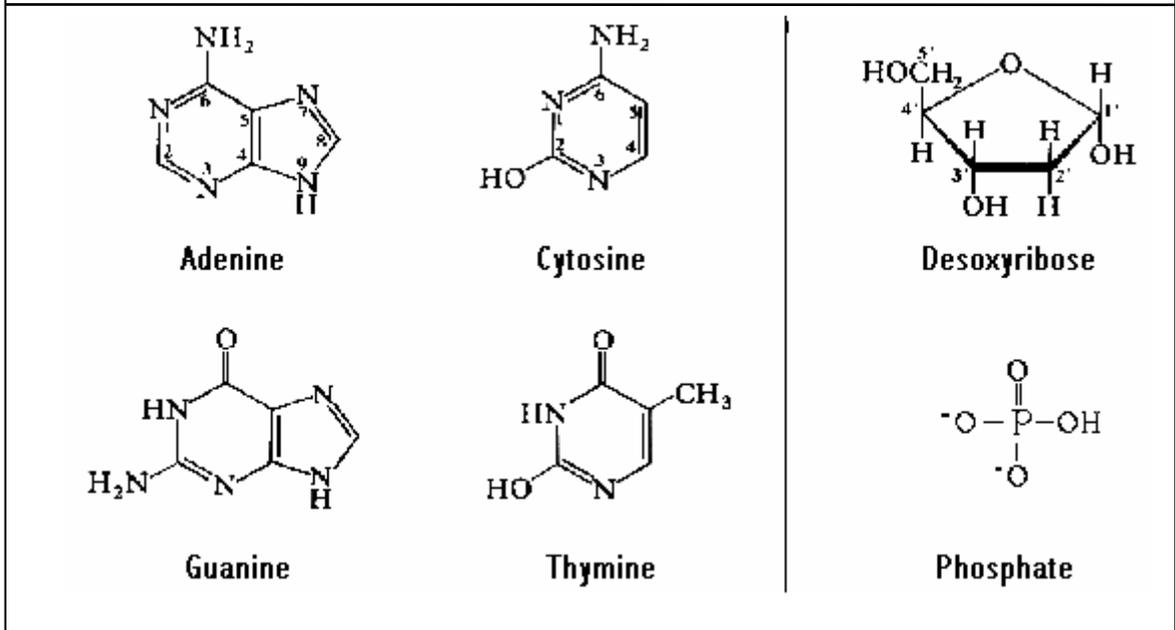
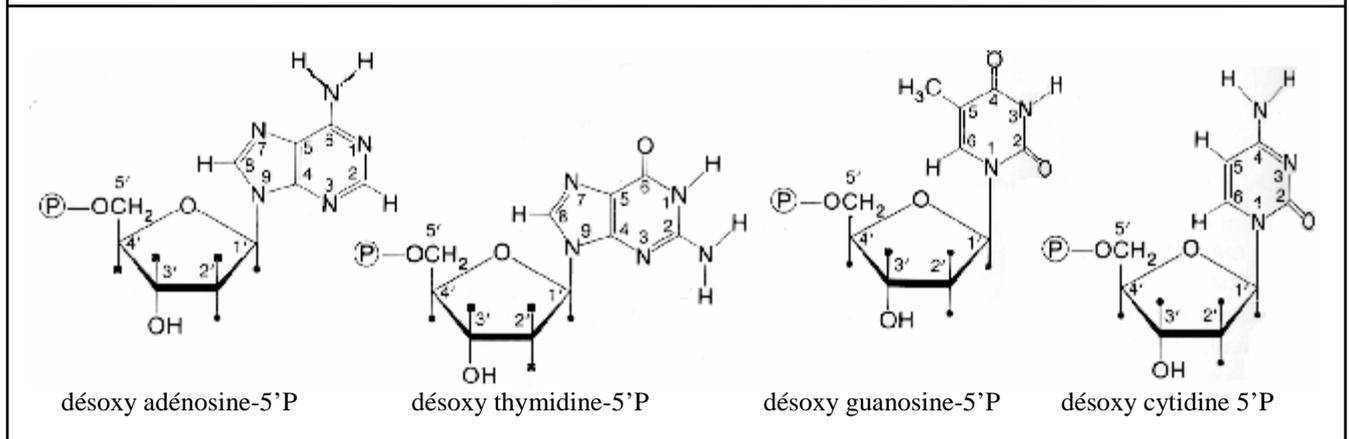


Figure 9 : les quatre nucléotides monophosphates issus de l'ADN .

P = phosphate ■ = H



(12) : on verra plus loin dans ce chapitre (2.2.2.) la raison de cette précision.

2.1.2. le polynucléotide possède un contenu informatif qui peut être considérable.

La figure 10 montre que le polymère est fait d 'un squelette ... *phosphate-sucre-phosphate- sucre....* sur lequel sont branchées les bases azotées : c 'est la structure primaire.

Il possède une extrémité dite **5'** qui comporte un groupement phosphate et une extrémité dite **3'** qui n'en comporte pas (13). La chaîne de monomères est donc **orientable** : par convention on place l'extrémité 5' à gauche dans toutes les séquences présentées horizontalement.

La succession des nucléotides est appelée séquence.

Dans le langage courant, on parle simplement de la **séquence des bases**.

Dans un fragment de polymère (figure 11) **aucune logique particulière ne paraît présider à cette succession.**

Et il ne s 'agit pas d 'un hasard, car **il en est de même si on étudie l 'ADN complet du virus Φ X 174** (figure 12).

Mais si aucune logique particulière n 'apparaît, il n 'en reste pas moins que la succession des bases, dans un ordre donné ressemble d 'assez près à différents systèmes de transfert d 'informations dont le plus connu est le système binaire, rencontré dans l 'alphabet morse ou dans la succession des valeurs 0 ou 1, chères à l 'informatique par exemple. Les séquences de nucléotides ont donc un potentiel informatif considérable.

Figure 10 : fragment de polymère, dont la séquence est TACGC

(Les séquences se lisent à partir de l'extrémité 5').

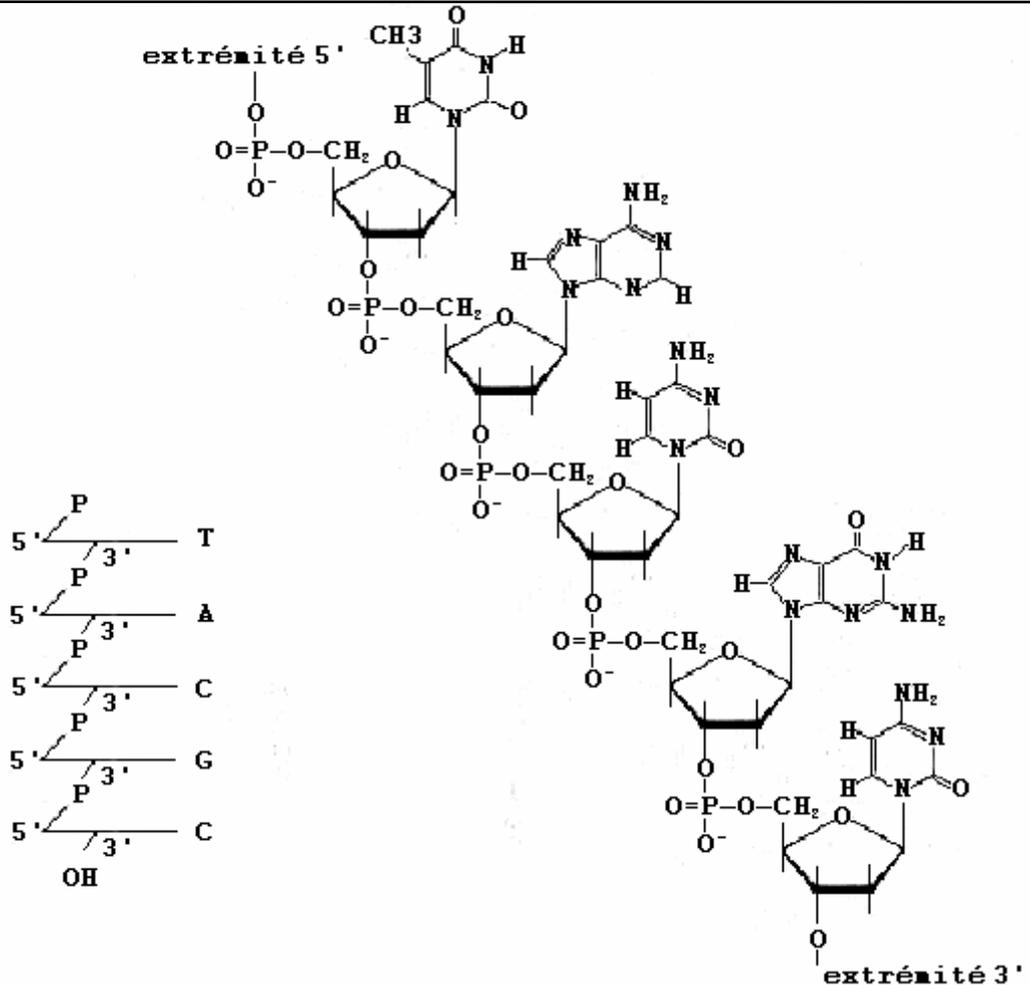


Figure 11 : exemple d'une partie de séquence issue de l'ADN de l'Homme (14).

....5'- GCTGGTTGTAAGAACTTCTTTTGAAGACTTTCACCTTCGTGT-3'....

(13): la notation 5'/3' provient de la numérotation conventionnelle des carbones et des azotes. « Prime » permet de distinguer ceux des bases azotées et ceux du sucre.

(14) : pour les curieux, avouons qu'il s'agit d'une partie de l'un des brins du gène de l'hormone somatostatine : cela n'a aucune importance pour nos propos actuels.

Encart n°4 : Arrêtons nous un moment devant la **page** de nucléotides représentant le génome de Φ X 174 .

Φ X 174 peut se perpétuer de génération grâce à cette page d 'instructions génétiques faites d'environ 5000 lettres (et grâce à la cellule bactérienne qui est parasitée par le virus).

Quand on sait que les cellules de l '**homme** disposent de plus de 3 milliards de bases, on a un peu le vertige lorsqu 'il faut bien admettre que cela représente $3 \cdot 10^9 / 5 \cdot 10^3 = 6 \cdot 10^5$ pages, soit **mille livres de 600 pages...**

Par ailleurs, cela nous conduit à nous émerveiller devant les progrès de la génétique, qui, tranquillement mais sûrement, déchiffre les rapports entre cette bibliothèque d 'instructions et les résultats biologiques: le métabolisme (figure 1), le développement, le cerveau

Encart 5 : les **virus** sont des êtres organisés qui ne possèdent pas de métabolisme propre et sont des parasites endocellulaires obligatoires. On ne les range donc pas chez les êtres vivants car leur matériel génétique n 'assure que la fonction de contrôle de la multiplication. Ce matériel génétique peut être soit de l 'ADN soit de l 'ARN. Leur unique acide nucléique peut être fait d 'un seul brin (c 'est le cas de Φ X 174) ou de deux (très souvent) .

2.2. L' ADN des êtres vivants est double à l'état natif.

2.2.1. Le modèle de Watson et Crick.

Les acides nucléiques ont été découverts en 1868 par Mischer qui les a appelés ainsi parce qu'il les a isolés à partir de *noyaux* cellulaires.

Il a pourtant fallu attendre 1953 pour que Watson et Crick découvrent leur structure grâce des progrès techniques, en particulier grâce à des clichés pris aux rayons X (15).

Il a fallu également résoudre une contradiction entre deux observations qui apparaissent contradictoires :

-**des quantités relatives apparemment quelconques des quatre bases dans un polymère** : par exemple , 8A , 17T , 9G , 7C dans un fragment (figure 11) ; 1291 A , 1684 T, 1254 G , 1157 C dans la **séquence complète** du génome du virus Φ X 174 (figure 12).

-**des quantités comparables de A et de T, de G et de C** lorsqu 'on étudie la composition en bases de l 'ensemble de l 'ADN des êtres vivants (figure 13).

(15) : il ne s 'agit pas du tout d'une technique ressemblant à une radio pulmonaire !! Voir un ouvrage de biochimie.

Figure 12 : les 5386 bases de Φ x174 : 1291 A, 1684 T, 1254 G, 1157C .

La lecture du texte se fait ligne par ligne , de gauche à droite.

GAGTTTTATCGCTTCCATGA CGCAGAAGTTAACTTTTCG GATATTTCTGATGAGTCGAA AAATTATCTTGATAAAGCAG GAATTACTACTGCTTGTFTA
 CGAATTAATCGAAGTGGAC TGCTGGCGGAAAATGAGAAA ATTCGACCTATCTTGGCAG GCTCGAGAAGCTTACTTTT GCGACCTTTCGCCATCAACT
 AACGATTCTGTCAAAAACGTG ACGCGTTGGATGAGGAGAAG TGGCTTAATATGCTTGGCAC GTTCGTCAAGGACTGGTTTA GATATGAGTCAACTTTTGT
 CATGGTAGAGATTCTCTTGT TGACATTTAAAAGAGCGTG GATTAATCTCTGAGTCCGAT GGTGTTCAACCACTAATAGG GCTGTTTAAAGAAATCATGAGTCAAGT
 TACTGAACAATCCGTACGTT TCCAGACCGCTTTGGCCTCT ATTAAGCTCATCAGGCTTC TGCCGTTTTGGATTTAACCG TGCCGTTTTATGGTACGCTGGA TAAGAAATCATGAGTCAAGT
 ACGAGTAAACAAAGTTGGAT TGCTACTGACCGCTCTCGTG CTGCTCGCTGCGTTGAGGCT CACGCTTTATGGTACGCTGGA AAAAAACGTTCTGGCGCTCGC CTTTGTGGGATACCCCTCGCT
 TTCTGCTCCTGTGAGTTT ATFGCTGCCGTCAATGCTTA TTAGTTCATCCCGTCAACA TTCAAACGCGCTGTCTCATC TCAAACGCGCTGTCTCATC ATGGAAGGCGCTGAATTTAC
 GAAAAACATTATTAATGGCG TCGAGCGTCCGGTTAAAGCC GCTGAATTGTTCCGCTTTAC CTGCGTGTACGCGCAGGAA CTGCGTGTACGCGCAGGAA ACACTGACGTTCTTACTGAC
 GCAGAAGAAAACGTGCGTCA AAAATTACGTGCGGAAGGAG TATGTAATGTCTAAAAGGTA AAAAAACGTTCTGGCGCTCGC AAAAAACGTTCTGGCGCTCGC CACTGCTGCTCCGACGCCGT
 GCGAGGTAATAAGGCAAGC GTAAAGGCGCTCGCTTTTGG TATGTAGGTGGTCAACAATT TTAATTGACGGGCTTCCGG TTAATTGACGGGCTTCCGG CCTTACTTGAGGATAAATT
 ATGTCTAATATCAAACGTG CGCCGAGCGTATGCCCATG ACCTTTCCCATCTTGGCTTC CTGTGGTCCAGATGGTTCG CTGTGGTCCAGATGGTTCG TCTTATFACCACTTCAACTA
 CTCCGGTTATCGTGGCGAC TCCTTCGAGATGGACGCCGT TCCCTGAGATGGACGCCGT CATGTGTAATGTCTAAAAGGTA TAACCGTATGAGCTTAAATC AAGATGATGCTCGTTAGGT
 TTTTACTTTTTATGTCCCTC ATCGTACGTTTTATGGTGAA CAGTGGATTAAGTTTATGAA CCCTTTCCCATCTTGGCTTC TACCACATCTATTGACATTA TGGGTCTGCAAGCTGCTTAT
 ACTGGTTATATTGACCATG CGCTTTTCTTGGCAGGATTA AGCGTTACCATGATGTTATT TCTTCAATTTGGAGGTA AAAAC CCAAGCATTGTTTTCAGG TATTATTTGAATCTATAACA
 ACTATTTTAAAGCGCCGTTG TCCTTAATCTTGGGCATCTGG TAACCTGATGAGCTTAAATC TCTCGCCAAATGACGACTTC ACCTGATATGAGCTTAAATC AAGATGATGCTCGTTAGGT
 AAACATTTGGACTGCTCCGC TTCTCCTGAGACTGAGCTT AGCGTTACCATGATGTTATT TCTTCAATTTGGAGGTA AAAAC CCAAGCATTGTTTTCAGG TATTATTTGAATCTATAACA
 GCTAATTTGCATACTGACCA AGAACGTGATTACTTCAATG AGCGTTACCATGATGTTATT TCTTCAATTTGGAGGTA AAAAC CCAAGCATTGTTTTCAGG TATTATTTGAATCTATAACA
 TCTCCCTTACTTGCATGCGC TCTTAATCTTGGGCATCTGG TAACCTGATGAGCTTAAATC TCTCGCCAAATGACGACTTC ACCTGATATGAGCTTAAATC AAGATGATGCTCGTTAGGT
 GACCTATAAACATTTCTGTG AGAACGTGATTACTTCAATG AGCGTTACCATGATGTTATT TCTTCAATTTGGAGGTA AAAAC CCAAGCATTGTTTTCAGG TATTATTTGAATCTATAACA
 CAGTACCTTAACGCTAAAGG TGCTTTGACTTATACCATA TACCTGATGAGCTTAAATC TCTCGCCAAATGACGACTTC ACCTGATATGAGCTTAAATC AAGATGATGCTCGTTAGGT
 TTTCCGTTATCGTGGTATTC AACCOCCTTCTGGTGAATTTG TATGCTGATATGCTTTTGGT GATGCTGATATGCTTTTGGT GATGCTGATATGCTTTTGGT TGGGTCTGCAAGCTGCTTAT
 AGGCTTCCCATTTACAGG TAAATTTAATGTGACCCGTTT TAAAAAATTTAAATTTTGGC TAATTTCAAACTTTTCTG TATTTGATAGTTTGGCTT TATGCTGATATGCTTTTGGT TGGGTCTGCAAGCTGCTTAT
 TTGCAGTGAATCATTCAGT AGGCTTAAATGTGACCCGTTT TAAAAAATTTAAATTTTGGC TAATTTCAAACTTTTCTG TATTTGATAGTTTGGCTT TATGCTGATATGCTTTTGGT TGGGTCTGCAAGCTGCTTAT
 AGGTTATAACCGCGAAGCGG TAAAAAATTTAAATTTTGGC TAATTTCAAACTTTTCTG TATTTGATAGTTTGGCTT TATGCTGATATGCTTTTGGT TGGGTCTGCAAGCTGCTTAT
 CAGACTTTTAAATTTCTCGCA TAATTTCAAACTTTTCTG TATTTGATAGTTTGGCTT TATGCTGATATGCTTTTGGT GATGCTGATATGCTTTTGGT TGGGTCTGCAAGCTGCTTAT
 AAGCTACATCGTCAACGTTA TATTTGATAGTTTGGCTT TATGCTGATATGCTTTTGGT GATGCTGATATGCTTTTGGT GATGCTGATATGCTTTTGGT TGGGTCTGCAAGCTGCTTAT
 TAATCAGTTGTTTCTGTTG TATTTGATAGTTTGGCTT TATGCTGATATGCTTTTGGT GATGCTGATATGCTTTTGGT GATGCTGATATGCTTTTGGT TGGGTCTGCAAGCTGCTTAT
 CTCCGACTGCTATGATGT TATTTGATAGTTTGGCTT TATGCTGATATGCTTTTGGT GATGCTGATATGCTTTTGGT GATGCTGATATGCTTTTGGT TGGGTCTGCAAGCTGCTTAT
 CGGGCAATAACGTTATGTT TATTTGATAGTTTGGCTT TATGCTGATATGCTTTTGGT GATGCTGATATGCTTTTGGT GATGCTGATATGCTTTTGGT TGGGTCTGCAAGCTGCTTAT
 TTGTCTTACGCCCATTTAGT TATTTGATAGTTTGGCTT TATGCTGATATGCTTTTGGT GATGCTGATATGCTTTTGGT GATGCTGATATGCTTTTGGT TGGGTCTGCAAGCTGCTTAT
 GCGGTCAAAAAGCGCCCTCC TATTTGATAGTTTGGCTT TATGCTGATATGCTTTTGGT GATGCTGATATGCTTTTGGT GATGCTGATATGCTTTTGGT TGGGTCTGCAAGCTGCTTAT
 AGGCTCTAATGTTTCAACC TATTTGATAGTTTGGCTT TATGCTGATATGCTTTTGGT GATGCTGATATGCTTTTGGT GATGCTGATATGCTTTTGGT TGGGTCTGCAAGCTGCTTAT
 GGCACCTTCTGCCGTTTCTGA TATTTGATAGTTTGGCTT TATGCTGATATGCTTTTGGT GATGCTGATATGCTTTTGGT GATGCTGATATGCTTTTGGT TGGGTCTGCAAGCTGCTTAT
 CATTTCTGAGCTTAATGCT TATTTGATAGTTTGGCTT TATGCTGATATGCTTTTGGT GATGCTGATATGCTTTTGGT GATGCTGATATGCTTTTGGT TGGGTCTGCAAGCTGCTTAT
 GCAACTGACAAATCAGAAAG AGATTGCCGAGATGCAAAAT TATGCTGATATGCTTTTGGT GATGCTGATATGCTTTTGGT GATGCTGATATGCTTTTGGT TGGGTCTGCAAGCTGCTTAT
 CAGGTATATGCACAAAATGA GATGCTTGTCTTATCAACAGA TATGCTGATATGCTTTTGGT GATGCTGATATGCTTTTGGT GATGCTGATATGCTTTTGGT TGGGTCTGCAAGCTGCTTAT
 AGGTTTCCGAGATATATGCGC CAAATGCTTACTCAAGCTCA TATGCTGATATGCTTTTGGT GATGCTGATATGCTTTTGGT GATGCTGATATGCTTTTGGT TGGGTCTGCAAGCTGCTTAT
 TGAGGTTGACTTAGTTCATC AGCAAACGCAGAAATCAGCGG TATGCTGATATGCTTTTGGT GATGCTGATATGCTTTTGGT GATGCTGATATGCTTTTGGT TGGGTCTGCAAGCTGCTTAT
 GCTTCTGGTGTGGTTGATAT TTTTCATGGTATTGATAAAG TATGCTGATATGCTTTTGGT GATGCTGATATGCTTTTGGT GATGCTGATATGCTTTTGGT TGGGTCTGCAAGCTGCTTAT
 ATTTGCTAGGAAATAACCG TCAGGATGACACCCCTCCA TATGCTGATATGCTTTTGGT GATGCTGATATGCTTTTGGT GATGCTGATATGCTTTTGGT TGGGTCTGCAAGCTGCTTAT
 TTCTGAATGCATGCTGATT ATTTGACTTTGAGCGTATC GGATAACCGCATCAAGCTCT TATGCTGATATGCTTTTGGT GATGCTGATATGCTTTTGGT GATGCTGATATGCTTTTGGT TGGGTCTGCAAGCTGCTTAT
 CTGGCTTCCATAAGCAGAT GGATAACCGCATCAAGCTCT TATGCTGATATGCTTTTGGT GATGCTGATATGCTTTTGGT GATGCTGATATGCTTTTGGT TGGGTCTGCAAGCTGCTTAT
 TTGACGGCCATAAGGCTGCT TCTGACGTTCTGATGAGTT TAGACCACCGCCCGAAGG TATGCTGATATGCTTTTGGT GATGCTGATATGCTTTTGGT GATGCTGATATGCTTTTGGT TGGGTCTGCAAGCTGCTTAT
 ACTTATGATTAATAACACTA TAGACCACCGCCCGAAGG TATGCTGATATGCTTTTGGT GATGCTGATATGCTTTTGGT GATGCTGATATGCTTTTGGT TGGGTCTGCAAGCTGCTTAT
 GCTGAACGCCCTCTAAGGA TATTCGCGATGAGTATAAAT TATGCTGATATGCTTTTGGT GATGCTGATATGCTTTTGGT GATGCTGATATGCTTTTGGT TGGGTCTGCAAGCTGCTTAT
 AATCGCGTAGAGGCTTTGCT ATTCAGCGTTTGTGATGATGC AATTTTGGTTCGTCGGGTACG TATGCTGATATGCTTTTGGT GATGCTGATATGCTTTTGGT TGGGTCTGCAAGCTGCTTAT
 CGACGATTAGAGCGTTTTT ATGATAATCCCAATGCTTTG TATGCTGATATGCTTTTGGT GATGCTGATATGCTTTTGGT GATGCTGATATGCTTTTGGT TGGGTCTGCAAGCTGCTTAT
 GATTACACGCGGACTGCTA TCAGTATTTTGTGTGCTG TATGCTGATATGCTTTTGGT GATGCTGATATGCTTTTGGT GATGCTGATATGCTTTTGGT TGGGTCTGCAAGCTGCTTAT
 CTACAGGTAGCGTTGACCT AATTTTGGTTCGTCGGGTACG TATGCTGATATGCTTTTGGT GATGCTGATATGCTTTTGGT GATGCTGATATGCTTTTGGT TGGGTCTGCAAGCTGCTTAT
 CGCAGTTCTGCTACACGCGG ACCTTTTTCAGTTCTGTT TATGCTGATATGCTTTTGGT GATGCTGATATGCTTTTGGT GATGCTGATATGCTTTTGGT TGGGTCTGCAAGCTGCTTAT
 GGTTCATATGTGCTAAATA CGTTAACAAAAGTCAAGATA TATGCTGATATGCTTTTGGT GATGCTGATATGCTTTTGGT GATGCTGATATGCTTTTGGT TGGGTCTGCAAGCTGCTTAT
 TGTGCTACTTCCAAAGAAG CTGTTTCAAGATCAGAATGAG TATGCTGATATGCTTTTGGT GATGCTGATATGCTTTTGGT GATGCTGATATGCTTTTGGT TGGGTCTGCAAGCTGCTTAT
 ACTTACCAAGCTGGTTACG ACAGGACCGCTTCAACCAG TATGCTGATATGCTTTTGGT GATGCTGATATGCTTTTGGT GATGCTGATATGCTTTTGGT TGGGTCTGCAAGCTGCTTAT
 GACGTTTTGGCGCGCAACC TGTGACGACAAATCTGCTCA AATTTATGCGCGCTTCGATA AATTTATGCGCGCTTCGATA AAAATGATTGGCGTATCCAA CCTGCA

Figure 13 : composition de l 'ADN de quelques êtres vivants

organisme	tissu	Adénine	Thymine	Guanine	Cytosine
Mycobacterium		15,1	14,6	34,9	35,4
Levure		31,3	32,9	18,7	17,1
Oursin	sperme	32,8	32,1	17,7	18,4
Hareng	sperme	27,8	27,5	22,2	22,6
Rat	moëlle osseuse	28,6	28,4	21,4	21,5
Homme	thymus	30,9	29,4	19,9	19,8
Homme	foie	30,3	30,3	19,5	19,9
Homme	sperme	30,7	31,2	19,3	18,8

Pour résoudre cette contradiction, Watson et Crick ont eu un trait de génie. Ils ont tranché le noeud gordien (16) pour réaliser une des découvertes scientifiques majeures de ce siècle, malgré son extraordinaire simplicité apparente:

**Une molécule d' ADN des êtres vivants n'est pas constituée d'un polymère
mais de deux polymères .**

Ces deux polymères appelés **brins** sont des reflets l'un de l'autre : quand sur un brin la base est A, l'autre brin porte un T, quand sur un brin il y a un G, un C est présent sur l'autre brin (figure 14).

Les **règles de complémentarité des bases** émises par Watson et Crick reposent sur des considérations que nous ne détaillerons pas (3). Il nous suffira de remarquer que les couples envisagés comportent chacun une base de type **purine** (A ou G) et une base de type **pyrimidine** (T ou C). Des liaisons chimiques faibles (dites hydrogène) s'échangent entre A et T (deux) et entre G et C (trois).

On remarquera que lorsque l'on connaît la séquence d'un brin, on peut en déduire celle de l'autre. Par ailleurs on observe que les deux brins sont orientés de manière opposée, si l'on considère les extrémités 3' et 5' (on dit que les brins sont **antiparallèles**).

2.2.2. Une conséquence technique de la structure double brin

Lorsque l'on chauffe un ADN natif, on constate que les deux brins se séparent. Les liaisons faibles sont détruites, tandis que les liaisons fortes (covalentes) de chacun des brins restent intactes : il y a **dénaturation**. C'est d'ailleurs un ADN dans cet état qui a été envisagé lors de la présentation de la notion de séquence (2.1.)

Si l'on refroidit rapidement la solution cette séparation se maintient. Par contre, un retour à une température basse dans des conditions ménagées particulières permet une **renaturation** : de manière assez étonnante les séquences complémentaires se réajustent et reforment la structure double brin initiale (17). La figure 15 illustre cette expérience.

Ces phénomènes de dénaturation et renaturation seront mis à profit lors des expériences d'**hybridation** moléculaire que nous envisagerons plus loin.

Figure 14 : représentations d 'un fragment d 'ADN .

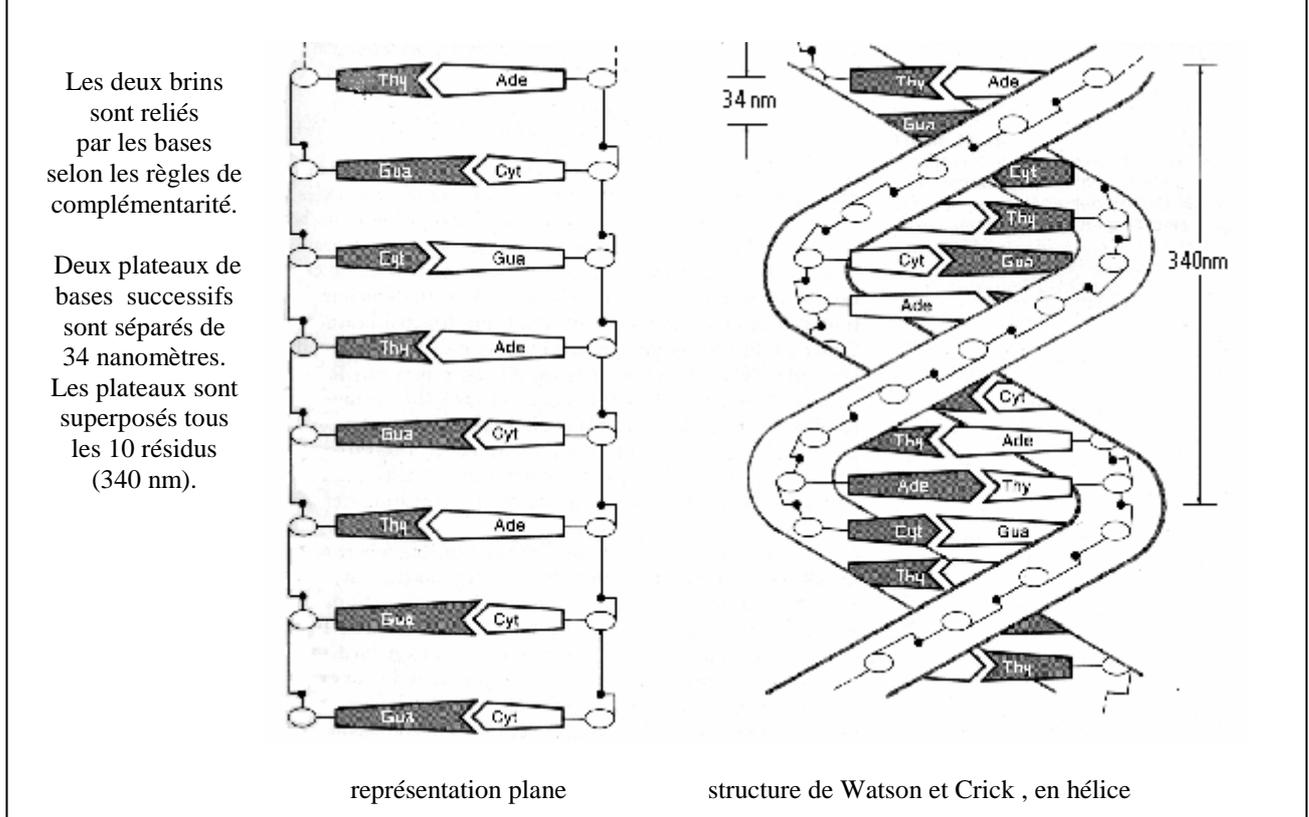
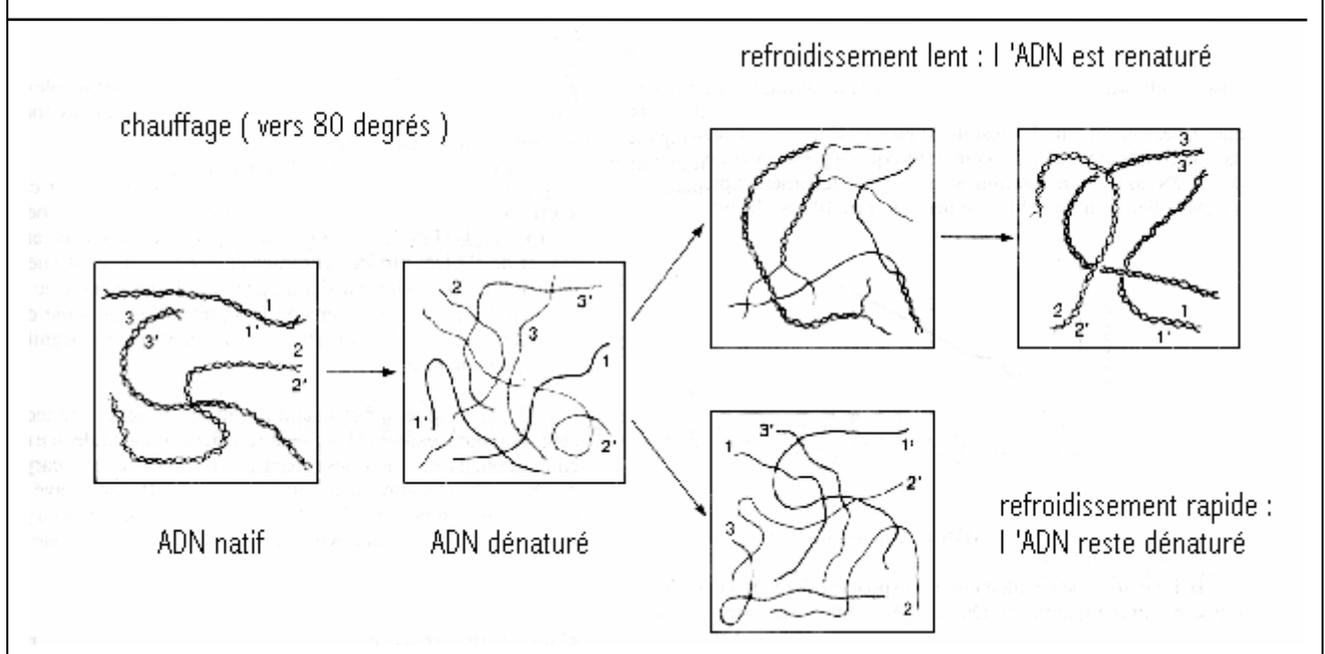


Figure 15 : effets de la température sur la structure de l 'ADN .



(16) : Si l 'on ne connaît pas l 'Histoire Ancienne , voir la vie d 'Alexandre le Grand , dans un ouvrage spécialisé !

(17) : On suit la renaturation par la simple observation de l'absorption de la lumière ultraviolette de 260 nanomètres de longueur d'onde. Cette lumière est plus absorbée par les simples brins que par les doubles brins

3. Les propriétés de l'ADN sont dues à sa structure.

3.1. La réplication est semi-conservative, ce qui assure la continuité du matériel génétique.

Nous avons vu que, dès 1944, Avery et ses collaborateurs ont montré très clairement que l'ADN est le matériel génétique. Il a pourtant fallu attendre le modèle de Watson et Crick pour que la collectivité scientifique en soit vraiment convaincue, 9 ans plus tard.

En effet, elle a tout de suite compris que l'aspect le plus spectaculaire du modèle est de permettre d'envisager la **pérennité** du matériel génétique d'une génération à une autre.

Nous ne rentrerons pas dans les mécanismes de la synthèse de l'ADN, aujourd'hui remarquablement connus. Il nous suffira de reprendre la proposition initiale des auteurs du modèle pour que l'on comprenne le principe de la synthèse de deux molécules identiques à partir d'une seule : cela se fait tout simplement grâce à la possibilité de séparation des deux brins et aux règles de complémentarité des bases.

Les brins « anciens » se séparent. Chacun sert de modèle à la fabrication d'un brin nouveau : les deux molécules obtenues sont strictement identiques (**un brin ancien, un brin nouveau**). Les doubles séquences de bases sont donc identiques entre elles et à la séquence de la molécule d'origine :

c'est **la réplication**.

On constate donc que la fonction de doublement de l'ADN est remarquablement inscrite dans sa structure (figure 16).

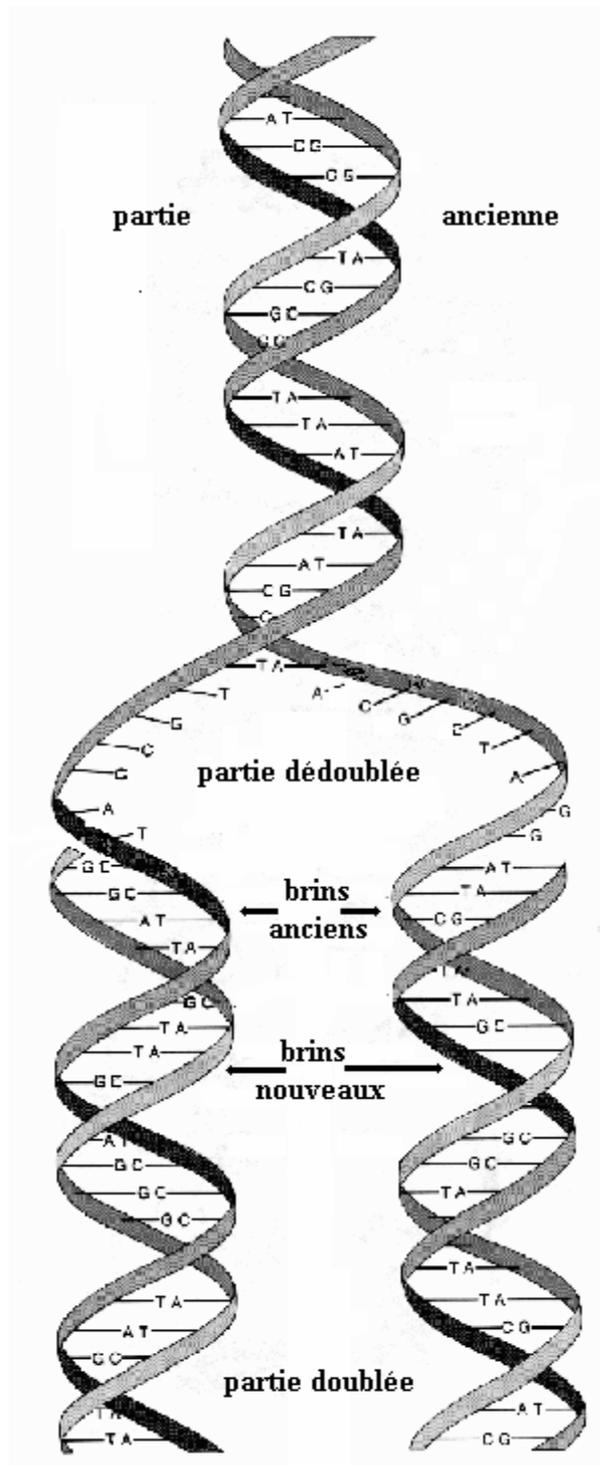


Figure 16 : le modèle de doublement de l'ADN , selon Watson et Crick (1953).

3.2. La transcription et la traduction font passer du message génétique aux polypeptides, acteurs du métabolisme et de la structure des cellules.

Comment l'ADN contrôle-il le fonctionnement cellulaire? Nous allons nous contenter, encore ici, de donner les indications qui nous sont indispensables (3).

La succession des bases ATGC qui forme une séquence porte en elle un contenu informatif, exactement comme un alphabet constitué de lettres permet de fabriquer des mots.

On démontre que ce message n'est pas utilisé tel quel par les cellules : il existe un intermédiaire fait d'acide ribonucléique, qui est synthétisé dans le noyau puis qui est transporté dans le cytoplasme cellulaire: on l'appelle l'**ARN messenger**.

Les acides ribonucléiques ont une structure proche de celle de l'ADN. Le sucre est le ribose (d'où leur nom). Les bases sont au nombre de 4 comme dans l'ADN mais l'une d'entre elles (la thymine de l'ADN) est remplacée par l'uracile dans l'ARN (18). Ces constituants sont illustrés dans la figure 17.

Les ARN sont synthétisés à partir de l'ADN par copie d'**un seul des deux brins** :

cette étape est appelée **transcription** (figure 18).

Figure 17 : constituants des ARN.

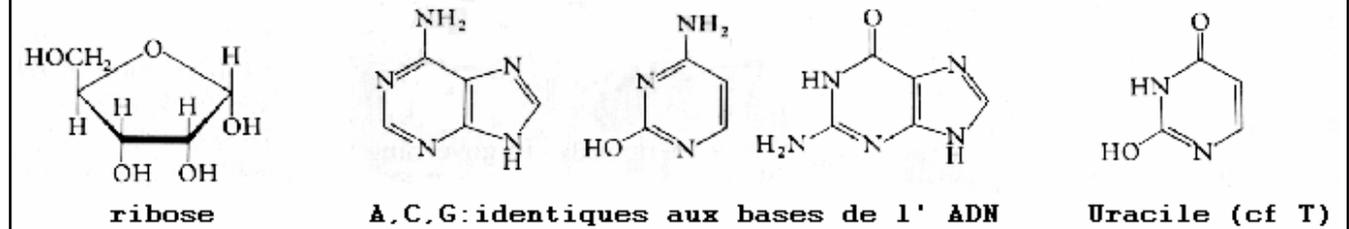
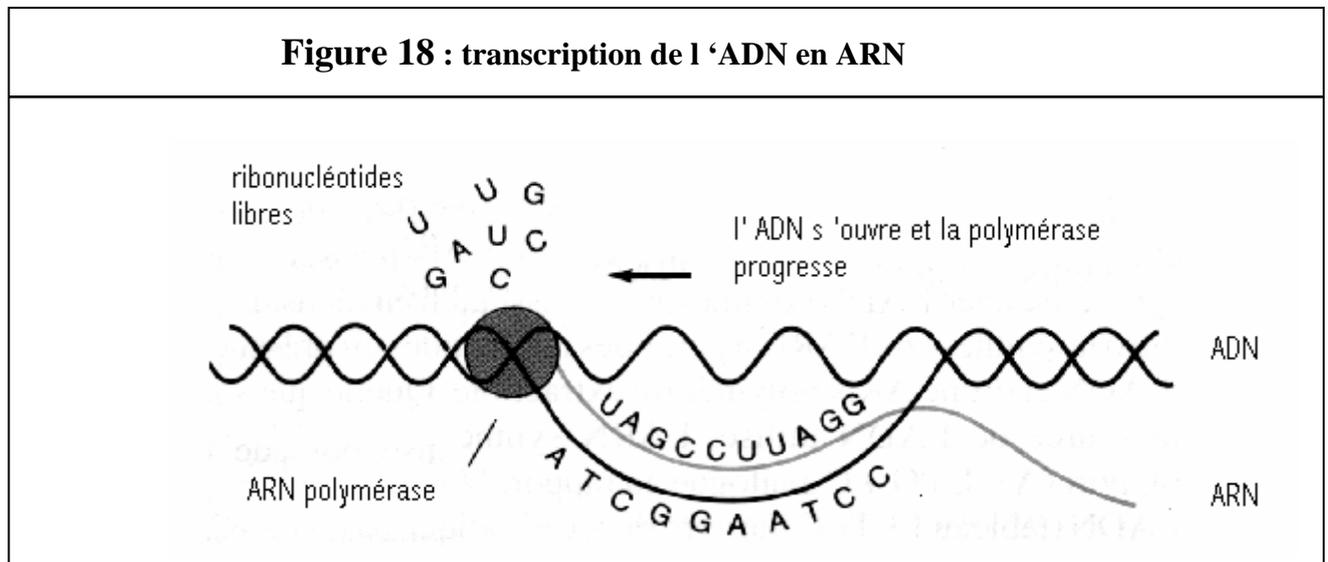


Figure 18 : transcription de l'ADN en ARN



(18) : Les ARN qui nous intéressent ici sont simples brins. Il existe des ARN double brins chez certains virus.

En général, un ARN messenger est responsable de la synthèse d'un polypeptide. Cette synthèse consiste à assembler de petits monomères, les **acides aminés** (figure 19), dans un **ordre particulier**, pour constituer un polymère (figure 20) que l'on appelle **polypeptide** (19). Cet ordre est appelé **séquence** (c'est la structure primaire).

Il existe dans une cellule plusieurs milliers de polypeptides chacun avec sa séquence (20) particulière. On a dû noter que l'on emploie le terme « séquence » aussi bien pour les acides nucléiques que pour les polypeptides. Cette identité va plus loin que le simple vocabulaire : il y a un rapport fonctionnel entre le polynucléotide et le polypeptide. Le polynucléotide est porteur d'un langage, grâce à la **succession de ses bases**, il permet la mise en place successive d'acides aminés du polypeptide. Et le polypeptide est unique parce que « son » ARN messenger est unique.

La mise en place des acides aminés est séquentielle : elle se fait grâce à d'autres intermédiaires, appelés ARN de transfert (ARNt). Ils sont de petite taille (moins de 100 nucléotides).

Un ARN t reconnaît un acide aminé particulier et un codon de l'ARN messenger. Cette reconnaissance se fait grâce à un anticodon, triplet de bases complémentaires de celles du codon.

Après avoir été mis en place, l'acide aminé est soudé au reste de la chaîne polypeptidique en cours d'allongement, par l'établissement d'une liaison peptidique (figures 20 et 21).

Figure 19 : acides aminés. Formule générale, les 20 habituels et leurs abréviations.

$ \begin{array}{c} \text{H} \quad \text{R} \\ \quad \\ \text{N} - \text{C} - \text{C} - \text{OH} \\ \quad \\ \text{H} \quad \text{O} \end{array} $	Alanine	Aln	Glutamine	Gln	Leucine	Leu	Sérine	Ser
	Arginine	Arg	Ac. Glutamique	Glu	Lysine	Lys	Thréonine	Thr
	Asparagine	Asn	Glycine	Gly	Méthionine	Met	Tryptophane	Trp
	Ac. Aspartique	Asp	Histidine	His	Phénylalanine	Phe	Tyrosine	Tyr
	Cystéine	Cys	Isoleucine	Ile	Proline	Pro	Valine	Val

Figure 20 :
mode de formation d'un polypeptide

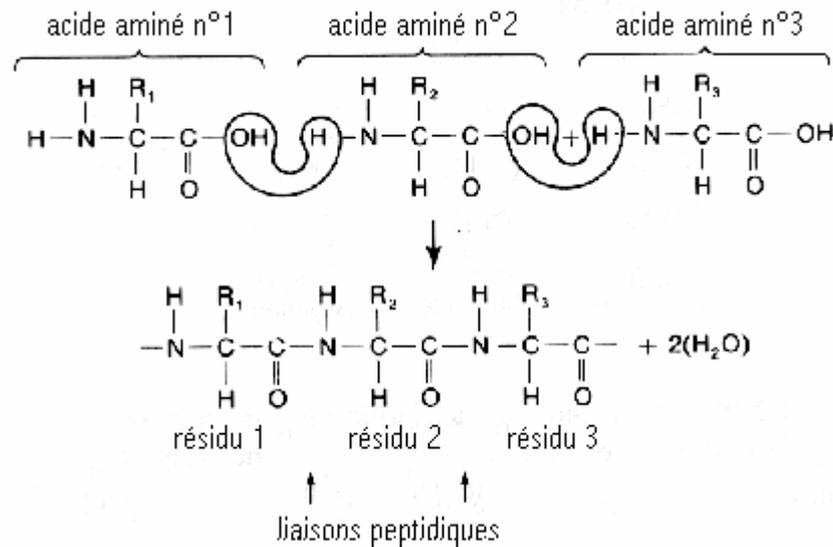
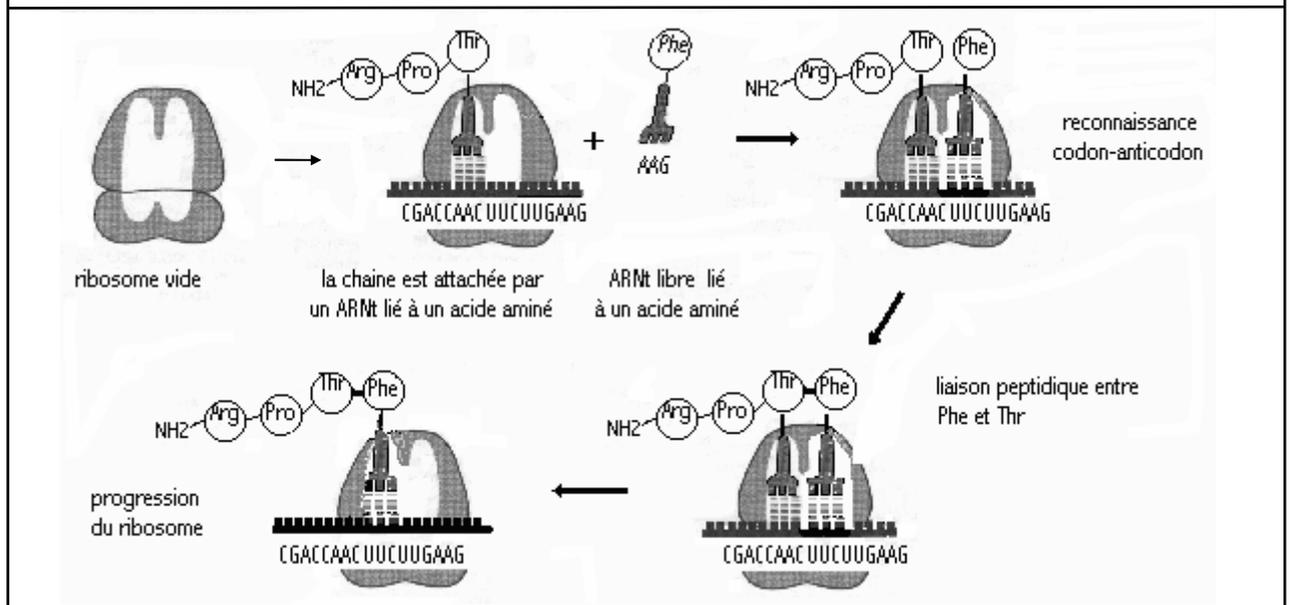


Figure 21 : traduction de l'ARN messager en polypeptide.



(19) : les termes *polypeptide* et *protéine* ne sont pas synonymes : il existe en effet des protéines faites de plusieurs polypeptides. On verra plus loin l'importance de cette remarque dans les études génétiques.

(20) : La succession des acides aminés est donnée de manière conventionnelle de gauche à droite. Cela correspond à une asymétrie du polymère, qui porte un radical NH₂ à l'une de ses extrémités et un radical COOH à l'autre.

Le message est lu **sans chevauchement**.

L'unité élémentaire d'information (le **codon**) est fait de trois nucléotides successifs (triplet).

La mise en place d'un acide aminé se fait grâce à un transporteur qui lui est propre. Il s'agit d'un petit ARN, que l'on appelle **ARN de transfert** (ARN t)

Les correspondances entre codons et acides aminés représentent le **code génétique** (figure 22).

Quatre codons jouent un rôle particulier : **AUG** lorsqu'il est proche du lieu d'accrochage du ribosome et de l'ARN messager est un signal indiquant le début de la synthèse d'un polypeptide (**initiation**) et le choix d'une des trois phases .

UAA, UGA et UAG sont des codons de **terminaisons**, où s'arrêtent la synthèse (on dit souvent qu'il s'agit de codons **stop**).

Trois lectures (trois **phases**) sont possibles selon que la traduction commence à la base n° 1 , 2 ou 3. Une seule de ces trois phases possibles est fonctionnelle (figure 23). Bien entendu , la phase fonctionnelle ne comporte pas de codon stop .

Il y a plus de codons ayant un sens (61, car un acide aminé - met - correspond à AUG), que d'acides aminés (vingt)
En fait le code génétique contient des **synonymes** : on dit qu'il est dégénéré ou **redondant** (21). Cette redondance globale est plus ou moins grande selon l'acide aminé : par exemple , si le tryptophane ne peut être codé que par le seul codon UGG , la sérine peut être codée par l'un des 6 codons : AGC, AGU, UCA, UCC, UCG ou UCU.

Cet ensemble de phénomènes s'appelle la **traduction**

est également un **récepteur** : des molécules diverses peuvent modifier son fonctionnement, dans les phénomènes de régulation.

4. Conclusion : le gène peut être défini à partir de sa nature et de sa structure moléculaires car ses propriétés en découlent.

Le matériel génétique est l'ADN. La structure de cette macromolécule lui confère deux capacités :

1 - la succession des bases (structure primaire ou séquence) possède un **contenu informatif**. La synthèse d'un polypeptide, via la transcription, la traduction et le code génétique, est sous le contrôle d'une portion de cette séquence

Une fois synthétisé, le polypeptide joue un rôle : il peut être doué d'activité catalytique (enzyme) ou bien être un transporteur (hémoglobine...) ou encore être impliqué dans la structure même de la cellule... Nous le considérerons comme l'**unité de fonction biochimique** (encart 9).

Il y a une correspondance entre le polypeptide, et le fragment d'ADN qui contrôle la séquence de ses acides aminés. Ce segment d'ADN est l'unité de fonction génétique et c'est à elle que nous donnerons le nom de gène :

Un gène est un fragment d'ADN qui code pour un polypeptide (22).

2 - la structure double brin (structure secondaire) permet que le message entier soit reproduit de génération en génération, via la réplication, de manière fidèle. Cela permet donc la conservation des gènes au cours des générations.

Ces deux conclusions peuvent être résumées en une seule phrase (23) :

l'élaboration et la continuité des caractères sont la double manifestation fonctionnelle de la structure de l'ADN.

Dans les deux prochains chapitres, nous allons voir que la structure moléculaire des gènes permet également d'expliquer la troisième caractéristique des êtres vivants qui nous intéresse, à savoir la **variabilité**, propriété qui apparaît tout d'abord opposée à la **continuité** que nous venons d'envisager.

Encart 7 : Combien de bases dans un gène , combien de gènes ?

L'ordre de grandeur d'un gène est le millier de paires de nucléotides chez les bactéries et leurs virus .
Certains gènes d'eucaryotes peuvent dépasser 1.000.000 bases.

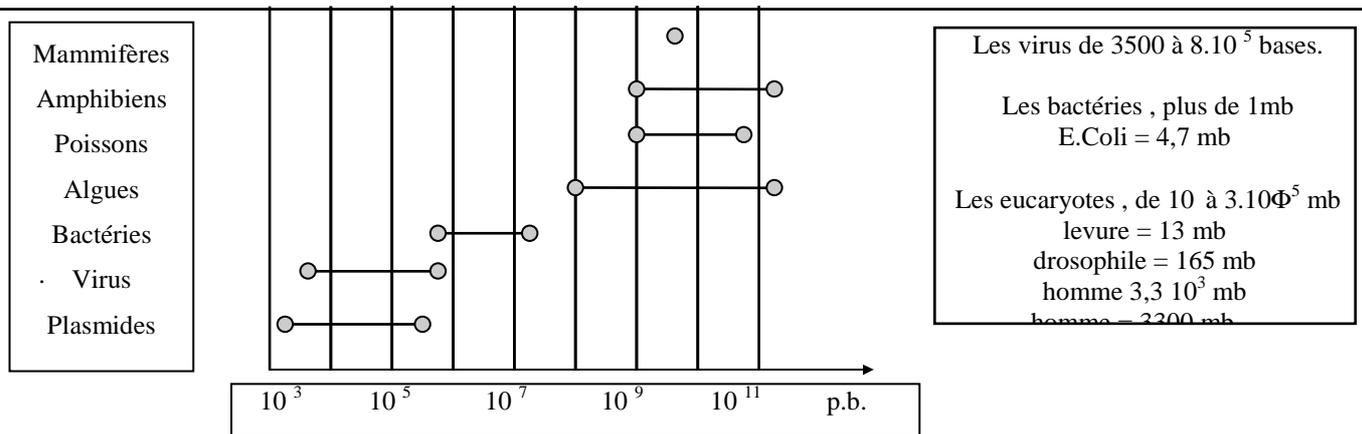
Afin de simplifier les notations,
on a créé les termes de kilo base (1 kb = 1.000 bases) et de méga base (1 mb = 1.000 . 000 bases).

Il existe de 2.000 à 3.000 gènes dans une cellule bactérienne.

Chez l'Homme, l'estimation du nombre de gènes va de 50.000 à 100.000.

Ces gènes contrôlent la synthèse d'autant de polypeptides assurant le métabolisme et / ou participant à la structure de la cellule ou à d'autres fonctions.

Encart 8 : le contenu en bases de l'ADN des êtres organisés est très varié.
Quelques nombres significatifs doivent être retenus.



Encart 9 : à eux tous , les polypeptides sont responsables de l'élaboration des caractères. Nous le verrons en détail dans le chapitre 5. Ce rapprochement devra être précisé : il ne faut pas croire qu'a x caractères correspondent x polypeptides , ou qu'à un caractère correspond un polypeptide. Une remarque semblable a été faite dans le chapitre 1 (note 14). Si l'on se répète , c'est dire l'importance de cette remarque !

(22): cette définition est approximative mais est suffisante à notre niveau . Notons cependant qu'il existe des séquences responsables de la synthèse des ARN ribosomiques (on parle de gènes ribosomiques) et des séquences codant pour les ARN de transfert, que l'on hésite plus ou moins à appeler gènes , car leur taille est au moins 10 fois plus petite que celle des gènes habituels

(23) : Les conclusions les plus simples doivent souvent être soigneusement méditées.

