

LE SUPPORT DE L'INFORMATION

1 - NATURE DU MATERIEL GÉNÉTIQUE

L'analyse génétique évoquée dans le chapitre précédent (Notion de gène et Transmission de l'Information Génétique) a permis d'aller très loin dans l'étude du mode de transmission des caractères, dans certains cas d'interactions entre les gènes, leur mode d'expression peut parfois être appréhendé grâce au raisonnement mendélien.

Malgré cela, l'analyse, pourtant relativement simple, du dihybridisme (avec dominance) qui se traduit, en F2, par quatre phénotypes dans des proportions 9-1-3-3 alors qu'il existe neuf génotypes possibles, laisse présager que le petit pois et la drosophile sont des organismes (diploïdes) bien trop compliqués pour saisir les mécanismes fondamentaux de la recombinaison et de l'expression des gènes.

Les génomes haploïdes tels que ceux des bactéries et des virus vont révolutionner la notion de gène.

Des séries de travaux s'étalant de 1928 à 1952 vont permettre d'associer définitivement l'ADN à la notion d'information génétique. Ces travaux sont exposés ci dessous car ils présentent, comme ceux de Mendel, beaucoup plus qu'un intérêt historique : ils sont des modèles d'analyse faisant appel, pour la première fois, à des méthodes de la biologie moléculaire.

1.1 TRAVAUX DE GRIFFITH

Le point de départ est une bactérie pathogène, un pneumocoque, agent de la pneumonie chez l'homme (et, au laboratoire, léthal pour la souris). Cet exemple rappelle que ce sont les pathologistes qui ont amené les généticiens à étudier les procaryotes et les virus.

La virulence de bactéries représente déjà un caractère phénotypique, un autre est apporté par l'aspect des colonies qui se développent à la surface d'un milieu nutritif gélosé lorsque l'on a ensemencé une boîte de Pétri : un mucopolysaccharide sécrété par les bactéries donne un aspect huileux aux colonies, ce phénotype est baptisé S (comme smooth = lisse).

Lors de repiquages et d'étalement successifs, on voit parfois apparaître des colonies présentant un phénotype différent : «rugueux» par opposition à lisse désigné par R (comme rough).

Inoculées à une souris, les bactéries de type R s'avèrent non virulentes. Les deux phénotypes : aspect des colonies et virulence sont liés.

* *Remarque* : L'apparition, rare et spontanée, de variants dans une population bactérienne est le résultat d'une mutation. A l'origine, une seule bactérie est modifiée mais cette modification étant héréditaire, la colonie que l'on observe représente en fait la descendance de la bactérie mutée.

* *Remarque* : Selon des mécanismes qui seront détaillés plus tard, l'accident mutationnel peut, dans certains cas, se produire «dans l'autre sens» : dans une population R peut spontanément réapparaître un individu S, cette modification est, elle aussi héréditaire puisque ce que l'on observe c'est une colonie S c'est à dire une descendance. Une des caractéristiques de ce phénomène de réversion est qu'il se produit avec la même fréquence que la mutation : c'est un événement très rare. Cette notion de réversion, simplement effleurée ici, s'avérera très importante.

Vers 1928, Griffith réalise une expérience fondamentale. Dans une première étape, il inocule à une souris, des bactéries S (de phénotype virulent) tuées par la chaleur, les souris ne présentent aucun trouble.

deuxième étape : il inocule des bactéries S tuées par la chaleur après les avoir mélangées à des bactéries R (phénotype non virulent), des souris meurent de pneumonie. Le prélèvement de bactéries à partir de souris mortes et leur mise en culture révèle un phénotype S (et virulent) pour toute la descendance.

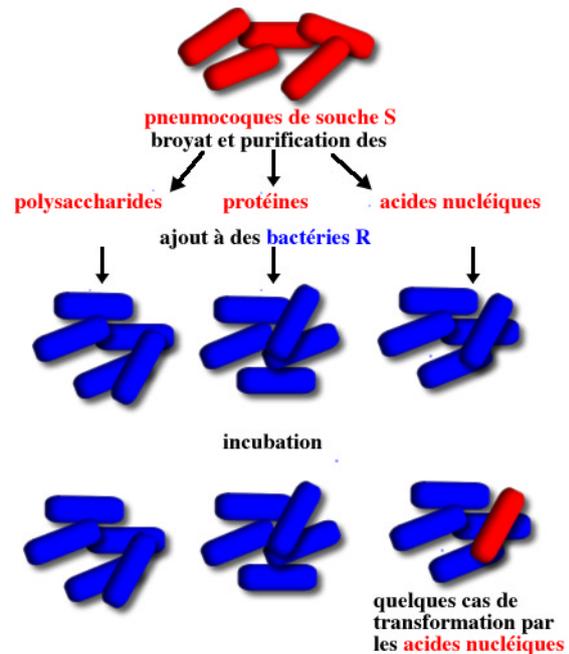


Interprétation : les bactéries vivantes au départ sont de type R, non virulent. Griffith, suppose qu'elles ont été «transformées» par un élément provenant des bactéries tuées par la chaleur. La fréquence de la transformation et des arguments d'ordre immunologique excluent un phénomène de réversion.

À l'époque, Griffith ne peut que conclure à l'existence d'un **facteur transformant** or il s'agit bien d'une transformation génétique au sens actuel du terme.

En 1944, Avery, McLeod, Mc Carthy reprennent ces travaux avec des tests un peu plus sophistiqués qui préfigurent la génétique microbienne moderne : ils mélangent, dans un tube, des bactéries S tuées par la chaleur et des bactéries R. Après un temps de culture suffisamment long, un anticorps anti R est ajouté, (plusieurs souches de ces bactéries ont été répertoriées qui diffèrent par leurs propriétés antigéniques), les bactéries R sont agglutinées et sédimentent au fond du tube, le surnageant est ensuite étalé sur un milieu nutritif et l'on s'aperçoit que des colonies de phénotype S se développent. Cette manipulation *in vitro* va servir à l'identification du principe transformant en ajoutant à des bactéries R, non plus des S tuées par la chaleur mais des extraits relativement purifiés de celles-ci. Les auteurs de ce travail se tournent

d'abord vers les polysaccharides de la paroi, sans aucun résultat puis vers les protéines sans plus de succès, ils n'obtiennent quelques cas de transformation qu'avec des préparations d'acides nucléiques. Des traitements enzymatiques montrent que la ribonucléase n'affecte pas le «pouvoir transformant» des préparations alors que la désoxyribonucléase l'abolit. C'est donc l'**ADN** qui est en cause.



Une bactérie peut donc acquérir de nouveaux caractères phénotypiques, de nouvelles fonctions métaboliques (sécrétion de polysaccharides, virulence) par l'intermédiaire d'ADN provenant d'une autre.

L'importance extraordinaire de ces travaux n'a pas été reconnue pendant longtemps pour plusieurs raisons :

- la structure chimique de l'ADN bien que déterminée d'une façon très incomplète, semblait trop simple pour pouvoir contenir une information aussi complexe que l'information génétique. Les protéines semblaient de bien meilleures candidates comme support de cette information.
- la génétique microbienne était à ses débuts et l'on n'était pas certain que le passage d'un type S à un type R soit le fait de mutations, ni que l'hérédité des organismes supérieurs soit comparable à celle des bactéries.

1.2 TRAVAUX DE HERSHEY ET CHASE :

C'est cependant l'étude d'un bactériophage (virus utilisant le métabolisme des bactéries), mais beaucoup plus tard, qui a convaincu le monde scientifique des rapports entre la molécule d'ADN et l'information.

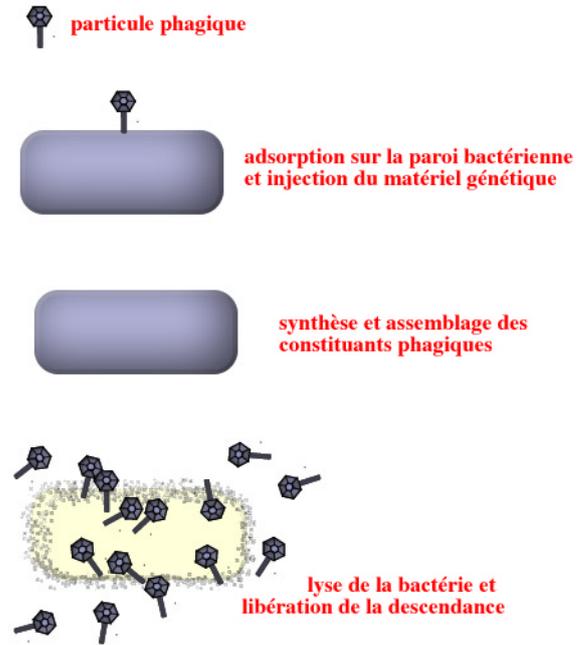
En 1952, Hershey et Chase étudient la reproduction du bactériophage T2 dans la bactérie *Escherichia coli* (qui deviendra le monstre sacré de la génétique moderne).

La figure ci-contre rappelle le mode de reproduction très particulier des bactériophages (plus simplement : phages).

Hershey et Chase vont utiliser une technique qui va désormais se développer rapidement : l'utilisation d'isotopes radioactifs comme traceurs, comme marqueurs, permettant de suivre la destinée de macromolécules. Dans un premier temps, ils cultivent les bactéries (*E. coli*) sur un milieu contenant du phosphore ^{32}P et du soufre ^{35}S . Après un certain temps de culture, les éléments constitutifs des bactéries contiennent ces marqueurs.

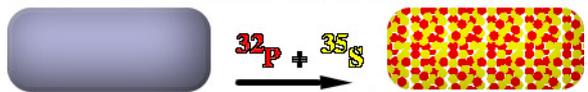
On infecte alors la culture avec une suspension de phages T2, ceux-ci vont réaliser un cycle lytique en utilisant les molécules radioactives de leurs cellules hôtes. La descendance phagique est recueillie et sert à infecter des bactéries «normales» (non radioactives).

On sait que, dans la première étape de l'infection, les phages s'adsorbent sur la bactérie et injectent à l'intérieur une molécule informative. Après cette étape d'adsorption, les auteurs agitent violemment la suspension pour décrocher ce qui reste à l'extérieur des bactéries, après centrifugation, on obtient un culot bactérien contenant l'information phagique et un surnageant contenant la capsidie. Or, le culot contient le ^{32}P et le surnageant le ^{35}S ; c'est la démonstration éclatante que l'information génétique du bactériophage, qui pénètre à l'intérieur de la bactérie est de l'ADN et que la capsidie protéique ne sert que d'emballage.

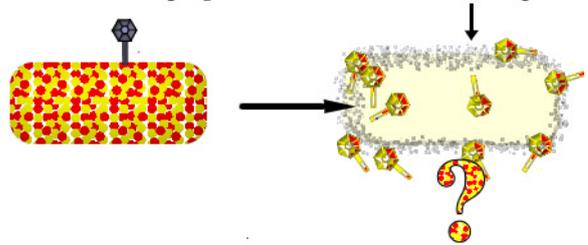


préparation des bactéries :

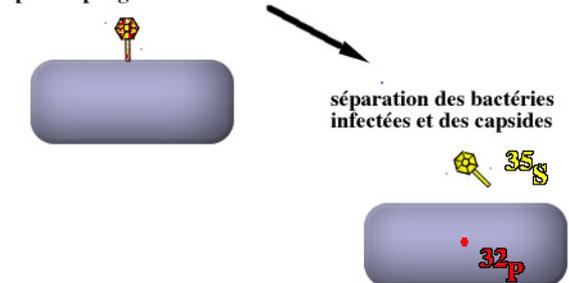
culture avec des sels radioactifs pendant plusieurs générations



infection de bactéries radioactives et récolte d'une descendance de phages dont les constituants sont marqués



infection de bactéries normales par les phages radioactifs



LE SUPPORT DE L'INFORMATION 2 - L'ADN

Watson et Crick, en 1953, ont immédiatement compris la portée des travaux de Hershey et Chase et leur proposition d'un modèle moléculaire en double hélice maintenue par des liaisons hydrogène entre des bases précises a marqué une autre étape sensationnelle et décisive de la génétique moderne. Les éléments de cette découverte sont intéressants à rappeler car ils fournissent un bel exemple de ce que permet l'intégration d'observations résultant de disciplines différentes. Watson et Crick ont progressivement élaboré leur modèle moléculaire à partir d'images, souvent difficiles à interpréter, de diffraction des rayons X par la molécule d'ADN laissant supposer une certaine régularité et une certaine répétition dans cette molécule très longue. Ils ont également tenu compte des observations d'Erwin Chargaff portant sur la composition en bases d'ADN provenant de différentes sources. A l'époque, il n'était pas question d'obtenir la séquence de ce polymère, mais il était possible, après hydrolyse complète (c'est à dire rupture de toutes les liaisons covalentes unissant les monomères entre eux), de séparer ceux-ci par chromatographie sur papier. On sépare ainsi quatre constituants («A», «T», «G» et «C») que l'on peut doser afin d'évaluer leurs proportions respectives.

Le tableau suivant, présente des caractéristiques que Chargaff a su interpréter : les purines et les pyrimidines sont également représentées (50% de chaque), quelle que soit la source de l'ADN, la proportion d'adénine est la même que celle de thymine et la proportion de guanine est la même que celle de cytosine, par contre, le rapport $A + T / G + C$ semble caractéristique de la source d'ADN.

Organisme	pourcentage de				rapport A+T/G+C
	A	T	G	C	
E. coli	26,0	23,9	24,9	25,2	1,00
D.pneumoniae	29,8	31,6	20,5	18,0	1,59
Levure	31,3	32,9	18,7	17,1	1,79
Rat	28,6	28,4	21,4	21,5	1,33
Homme	30,3	30,3	19,9	19,8	1,52

2.1 STRUCTURE DES ACIDES NUCLÉIQUES

Les acides nucléiques sont des **polymères de nucléotides** de très grande taille (macromolécules).

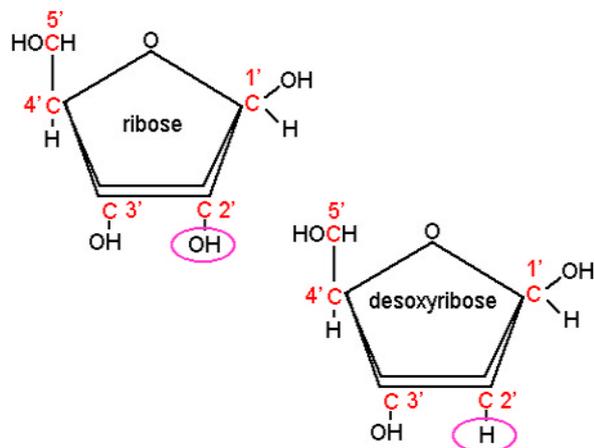
Dans ce rappel très rapide on ne fera que souligner ce que les particularités structurales entraînent comme conséquences pratiques. Ainsi la composition détaillée des différents nucléotides est supposée connue mais il est bon de garder à l'esprit que leur enchaînement conduit effectivement à des molécules de très grandes tailles.

Celles d'ADN se mesurent en centimètres pour la longueur avec un diamètre pour la double hélice de 20 Angström. Conséquence pratique : étant donné les moyens mécaniques mis en jeu dans les différentes étapes de purification de l'ADN, il est impossible d'obtenir des molécules intactes, on ne travaille qu'avec des morceaux. De plus, les cassures étant aléatoires, on travaille avec un mélange hétérogène de morceaux, c'est ce qui, pendant longtemps a bloqué la progression de l'analyse fine du gène.

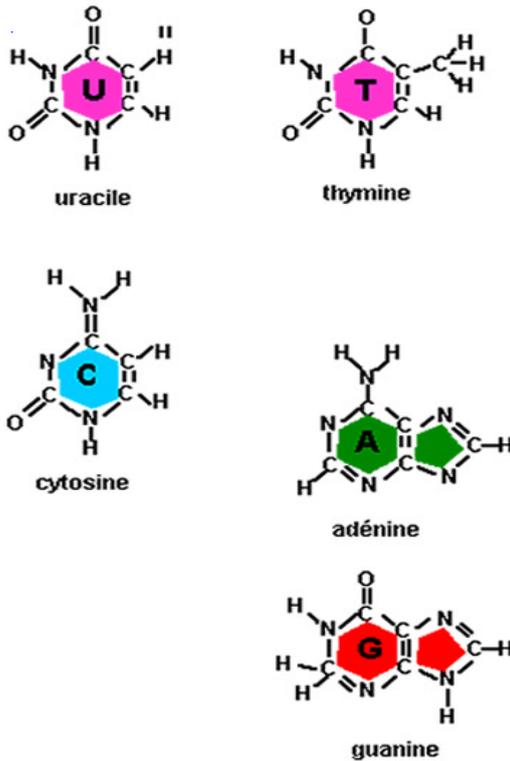
2.1.1 LES MONOMÈRES ET LEUR ENCHAÎNEMENT

Le nucléotide est lui même composé de trois molécules :

1 - un **pentose** sous forme cyclique (furane): le ribose en ce qui concerne l'acide ribonucléique (ARN) et le 2'désoxyribose pour l'acide désoxyribonucléique (ADN)



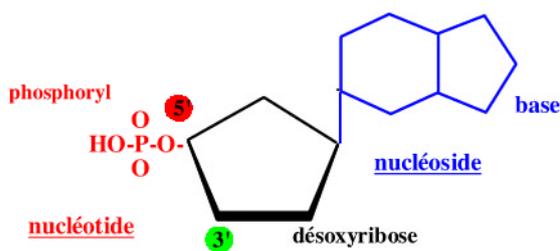
2 - une base organique est reliée au sucre en 1' :



soit une base purique adénine (A) ou guanine (G)
soit une base pyrimidique : cytosine (C), thymine (T) ou uracile (U).

L'ensemble constitue un nucléoside

3 - un acide phosphorique vient estérifier une fonction alcool du sucre en 3' pour constituer le nucléotide.



Nucléoside monophosphate

Les nucléotides précurseurs de la synthèse d'acides nucléiques sont sous forme triphosphate, deux phosphoryls (β et γ) seront éliminés au cours de la polymérisation.

La polymérisation s'effectue par l'estérification d'un alcool situé en 3' d'un nucléotide par le phosphate d'un autre nucléotide (la liaison covalente ainsi établie est dite liaison 3'-5' phosphodiester).

* Remarques pratiques :

1) En raison de leurs doubles liaisons conjuguées, les bases absorbent fortement la lumière ultraviolette. Il est donc très facile de doser les acides nucléiques en solution en mesurant la densité optique à 260 nm (maximum d'absorption).

2) En raison des groupes phosphoryls, les acides nucléiques sont chargés négativement, ils se comportent comme des polyanions et migrent vers l'anode lors d'une électrophorèse.

L'électrophorèse est actuellement l'outil de base de la génétique moléculaire car elle permet de séparer des fragments d'acides nucléiques en fonction de leur taille.

* Remarque importante : les molécules se déplacent dans un champ électrique uniquement parce qu'elles sont chargées mais la migration nécessite un support pour ces molécules. Sur le plan électrique, ce support (un gel poreux) n'intervient pas, par contre, sa porosité va déterminer, à charges égales, la vitesse de migration : des grosses molécules seront beaucoup plus freinées par le gel poreux que des petites. Cette propriété, même si l'explication reste intuitive, a un champ d'application considérable : la distance de migration est inversement proportionnelle au logarithme de la masse, ainsi, on va pouvoir, à l'aide de fragments de taille connue (échelle), « calibrer » les supports d'électrophorèse et déterminer, avec une précision qui dépend de la nature du gel utilisé, la taille des fragments d'ADN contenus dans un mélange.

Seules les bases distinguent les différents nucléotides.

Les deux acides nucléiques sont des polymères à quatre monomères possibles symbolisés par A T G C pour l'ADN et A U G C pour l'ARN.

Ces polymères ne sont pas ordonnés (de type ABCD ABCD ABCD ABCD...) mais séquencés ceci se traduit par une succession de nucléotides qui nous semble aléatoire. Cependant la séquence des monomères de tout acide nucléique est capitale pour la transmission et l'expression du matériel génétique. La séquence des nucléotides dans le polymère impose ce que l'on appelle la structure primaire des polynucléotides.

Souvent, les molécules sont linéaires et possèdent donc une extrémité 5' phosphorylée et une extrémité 3' hydroxylée d'où la représentation simplifiée suivante :



2.1.2 STRUCTURES SECONDAIRES

La structure secondaire des acides nucléiques est imposée par l'appariement des bases c'est-à-dire la formation de liaisons hydrogène entre deux bases organiques.

L'ADN est constitué de deux chaînes polynucléotidiques ainsi appariées. Etant donné que, pour des raisons d'encombrement stérique, deux paires de bases seulement sont stables A..T et G..C, la séquence d'un brin de cette molécule bicaténaire implique celle de l'autre.

Cette notion de complémentarité est fondamentale pour tout ce qui concerne le métabolisme de l'ADN *in vivo* et *in vitro*.

De plus (voir ci-contre), les deux bases d'une paire ne sont suffisamment proches l'une de l'autre et ne sont capables d'établir des ponts hydrogène que si les deux chaînes nucléotidiques sont anti-parallèles soit :



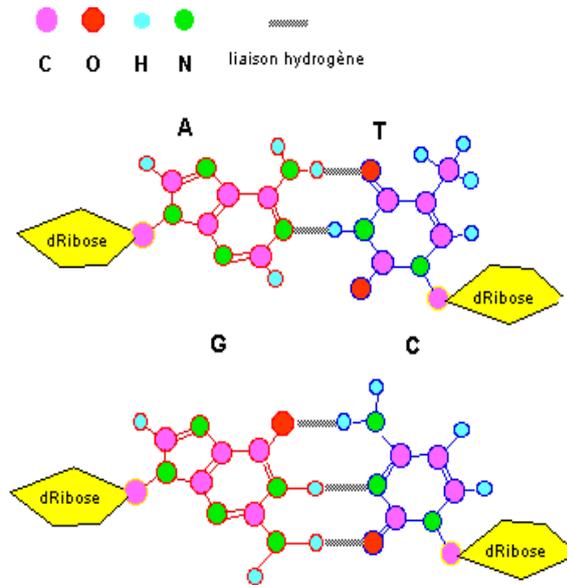
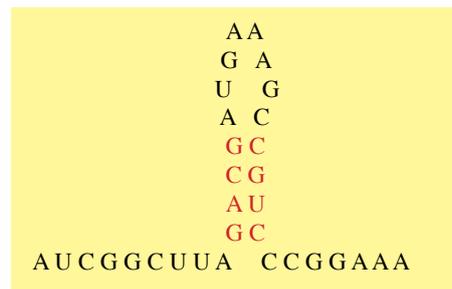
Il en résulte, dans l'espace, une double hélice dont le pas de 34 Angström correspond à l'empilement de 10 paires de bases.

**Remarque : la figure montre qu'il existe 3 liaisons hydrogène entre G et C et deux seulement entre A et T, ceci implique qu'un ADN riche en G-C sera plus stable qu'un ADN comportant plus de couples A-T.*

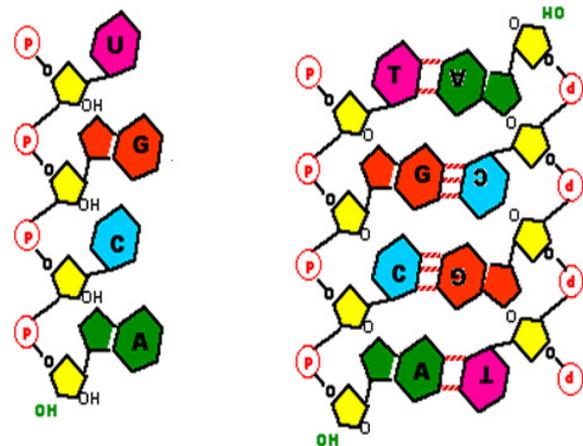
Des molécules simple brin adoptent souvent, localement, des structures secondaires, selon le même principe d'appariement des bases. C'est le cas de l'ARN : en raison des séquences (en rouge) dans la molécule ci-dessous :



elle peut prendre la structure secondaire suivante:



Ci-dessous : comparaison (très schématique) des chaînes d'ARN et d'ADN.



Enfin, parmi les structures particulières, il faut signaler celle des **molécules circulaires** : il n'y a pas d'extrémité 5' ni 3' (une liaison phosphodiester les relie). De telles structures (double ou simple brin) sont fréquentes chez les virus et des éléments génétiques particuliers tels que les plasmides, le «chromosome» bactérien est également une molécule circulaire .

2.1.2 PROPRIÉTÉS TOPOLOGIQUES DE L'ADN.

Les liaisons hydrogène, relativement fragiles, peuvent être détruites par chauffage ou par un pH élevé, dans ces conditions les structures secondaires disparaissent. Pour l'ADN, le résultat est la séparation complète des deux brins qui le composent : il y a **dénaturation** de la molécule.

En raison de la stricte complémentarité des bases, **la dénaturation est réversible**, les deux brins peuvent, dans des conditions appropriées de température et de force ionique, rétablir des liaisons hydrogènes entre leurs bases et reprendre la configuration en double hélice d'origine.

Il faut bien comprendre que ce phénomène ne dépend que de la complémentarité de deux séquences nucléotidiques et non de l'origine de chaque brin.

Dans le cas de renaturation, on a formation de molécules «homoduplex» c'est à dire de structures doubles, chaque partenaire ayant une même origine. On peut très bien mélanger des ADN dénaturés d'origines différentes, s'il y a des complémentarités de séquences, des «hétéroduplex» c'est-à-dire des molécules dont les deux brins sont d'origine différente pourront se former. Enfin, toujours par complémentarité de séquences, on peut appairer un ARN avec un brin d'ADN (il existe une possibilité de liaison A..U) on obtient alors une molécule «hybride».

Cette propriété (appariement de séquences complémentaires) représente un outil puissant dans la reconnaissance de séquences précises à l'aide de petits fragments polynucléotidiques appelés «sondes» car, dans un ensemble complexe de séquences différentes, elles sont capables de trouver celle qui leur correspond, de s'y hybrider.

La détection de ces hybrides sera facilitée par l'utilisation de sondes «marquées».

molécule native:

```
AATGCCGTCACCTTTAGCTATA
TTACGGCAGTGAAATCGATAT
```

molécule dénaturée :

```
AATGCCGTCACCTTTAGCTATA
                                TTACGGCAGTGAAATCGATAT
```

homoduplex :

```
AATGCCGTCACCTTTAGCTATA
TTACGGCAGTGAAATCGATAT
```

heteroduplex :

```
AATGCCGTCACCTTTAGCTATA
                                GAAATCGATATTGCC
```

hybride :

```
AATGCCGTCACCTTTAGCTATA
TTACGGCAGTGAAATCGATAT
```

sondes :

```
AATGCCGTCACCTTTAGCTATA
AAAT C
```

ou

```
TTTA G
TTACGGCAGTGAAATCGATAT
```