

## 1 - Notions de base de génétique

# ADN et chromosomes

H. HAYES

INRA, Laboratoire de Génétique Biochimique  
et de Cytogénétique,  
78352 Jouy-en-Josas cedex

e-mail : hayes@biotec.jouy.inra.fr

**Résumé.** Chaque chromosome contient une seule molécule d'ADN. L'ADN déroulé d'un noyau de cellule humaine mesurerait environ 1,8 m : chaque molécule d'ADN est enroulée et compactée en plusieurs étapes, grâce à l'association de différentes protéines, et loge dans le noyau de 6 µm de diamètre. Le degré de condensation de l'ADN est variable selon les régions chromosomiques et les régions les moins condensées sont les plus riches en gènes. L'ADN est composé d'une variété de séquences codantes ou non et répétées ou non dont l'organisation dans le chromosome est caractéristique de la métaphase. Certaines séquences peuvent être corrélées en partie avec les motifs de bandes spécifiques de chaque paire chromosomique produits par les techniques de marquage chromosomique.

L'ADN, constitué de millions de nucléotides dont l'enchaînement précis détermine l'information génétique de chaque organisme, est localisé chez les mammifères en quasi-totalité (3000 Mb pour un noyau haploïde humain) dans les chromosomes de chaque noyau cellulaire. Seule une petite molécule d'ADN circulaire (16,6 kb chez l'Homme) est située dans un autre organite cellulaire, la mitochondrie. C'est uniquement au cours de la division cellulaire qu'il est possible d'observer au microscope des chromosomes individualisés dont le nombre, la taille et la forme sont constants et caractéristiques pour toutes les cellules somatiques d'une espèce donnée et qui se regroupent en paires de chromosomes homologues (à l'exception de la paire des chromosomes sexuels, souvent morphologiquement très différents). Par exemple, l'homme a 23 paires de chromosomes, la vache et la chèvre 30, le porc 19 et le cheval 32. Tous les chromosomes métaphasiques (observés au grossissement 600) présentent une constriction primaire ou centromère qui partageant en deux bras plus ou moins longs (p et q) et dont la position détermine leur forme, métacentrique si elle est médiane, acrocentrique si elle est proche d'une extrémité et submétacentrique si elle est intermédiaire.

La propriété la plus exploitée et, paradoxalement, la moins bien comprise des chromosomes est de présenter le long de leur bras, à la suite de colorations ou de traitements particuliers, des bandes plus ou moins colorées dont la succession et la taille constituent un motif précis et reproductible, visible au microscope optique. Cette propriété, très précieuse, permet d'identifier les chromosomes (figure 1), d'assembler un caryotype, de décrire finement des remaniements chromosomiques ou structuraux, de localiser des marqueurs et de comparer les chromosomes entre eux et entre espèces proches ou

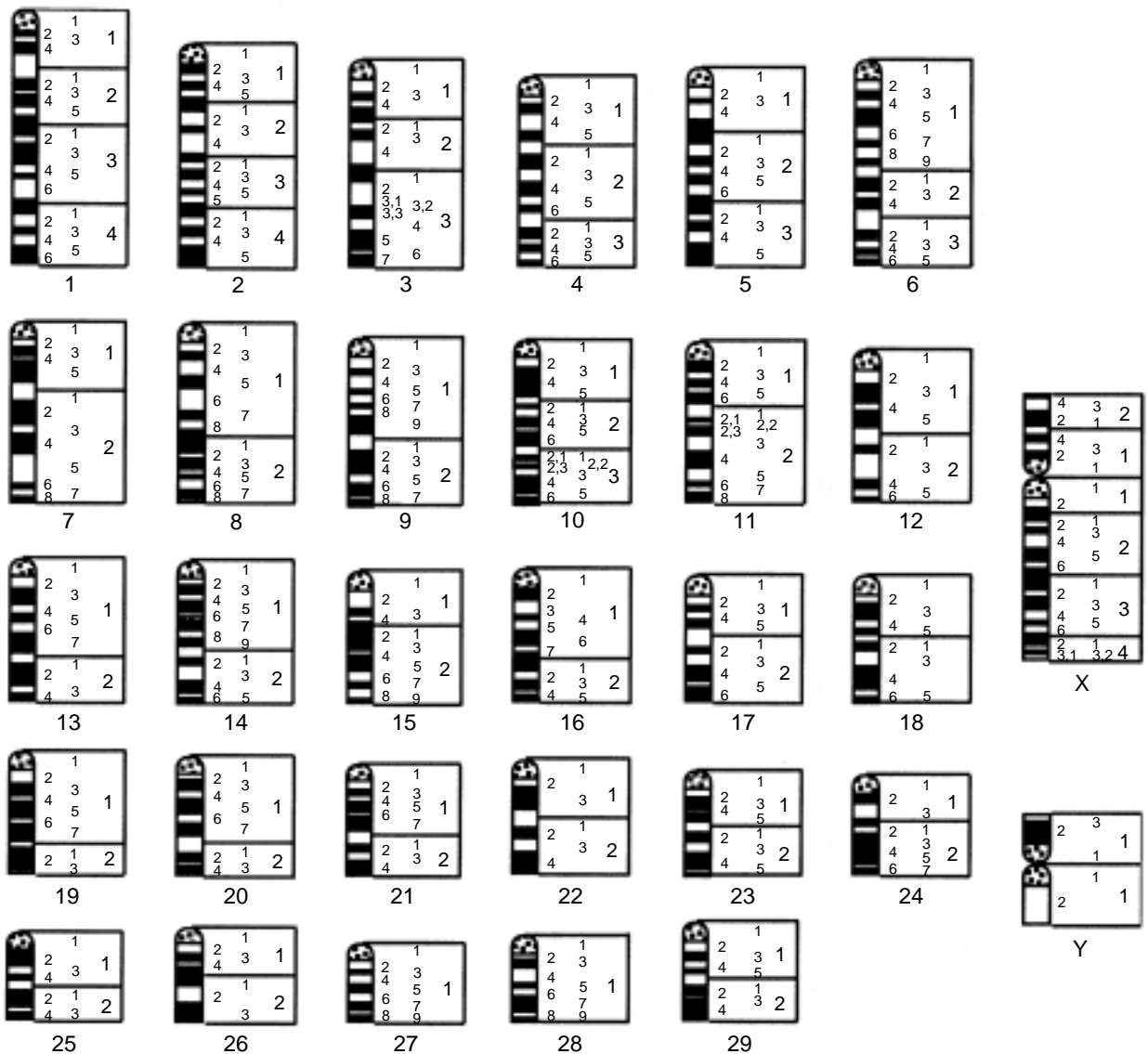
éloignées. Cependant, la compréhension des mécanismes à l'origine des bandes chromosomiques est très incomplète et les explications trouvées dans la littérature restent souvent du domaine de l'hypothèse.

Les progrès combinés des techniques de biologie moléculaire et de microscopie optique et électronique permettent de proposer une image beaucoup plus précise de l'organisation architecturale, topologique et fonctionnelle de l'ADN au sein des chromosomes eucaryotes et le lien avec le motif des bandes chromosomiques, au moins chez les mammifères.

### 1 / Repliement de l'ADN dans les chromosomes

Une cellule diploïde de mammifère contient en moyenne 6000 Mb d'ADN, soit une longueur d'environ 1,8 m contenue dans un noyau de 6 µm de diamètre grâce à un empilage ordonné et efficace de l'ADN (figure 2). Chaque chromosome contient une seule molécule d'ADN linéaire, continue, repliée de nombreuses fois sur elle-même et répartie de part et d'autre du centromère. Dans les noyaux, l'ADN n'est probablement jamais libre mais associé à d'autres molécules, principalement les histones, petites protéines basiques présentes en quantité à peu près égale à celle de l'ADN, et des protéines non histones acides représentant entre 10 et 30 % de l'ensemble. C'est ce complexe ADN-protéines, appelé aussi chromatine parce qu'il fixe les colorants, qui compose la structure des chromosomes. Les histones, dont la structure est très conservée à travers tous les organismes eucaryotes, ont un rôle essentiel dans le repliement des molécules d'ADN. La compaction (Gasser et Laemli 1987) de la molécule d'ADN chromosomique (double hélice d'environ 2 nm de diamètre) associée

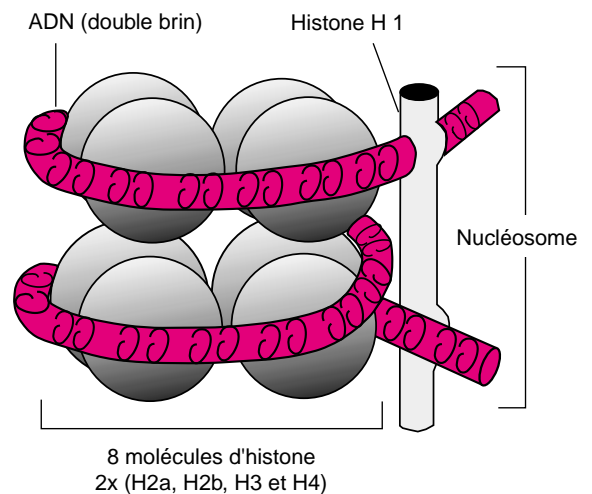
Figure 1. Idiogramme bovin en bandes RBG (nomenclature Texas 1996).



aux histones et à l'ARN passe par différents niveaux d'empaquetage (figure 3). L'unité structurale de base de la chromatine est le nucléosome (figure 4), qui est formé d'un assemblage de 8 histones (2 fois : H2a, H2b, H3 et H4) autour duquel s'enroule une portion d'ADN double brin de 146 paires de bases et qui est répété indéfiniment, donnant un aspect en "chapelet de perles" à la fibre de chromatine d'environ 11 nm d'épaisseur (cf figure 3). Grâce aux histones H1, ce chapelet de nucléosomes se comprime en formant une super hélice de 30 nm de diamètre qui s'organise elle-même en boucles d'environ 300 à 400 nm de longueur le long d'une armature constituée en grande partie par la topoisomérase II (enzyme capable de couper les deux brins de l'ADN et qui, en plus d'un rôle architectural, intervient dans le relâchement des supertours de la chromatine lors de la réplication). Les boucles de chromatine sont attachées à l'armature au niveau de régions particulières de l'ADN appelées « SAR », de l'anglais « Scaffold Attachment Regions » (cf paragraphe 2, modèle de Saitoh et Laemmli). Enfin cette organisation boucles + armature s'enroule en une spirale plus ou moins resserrée selon le stade du cycle cellulaire :

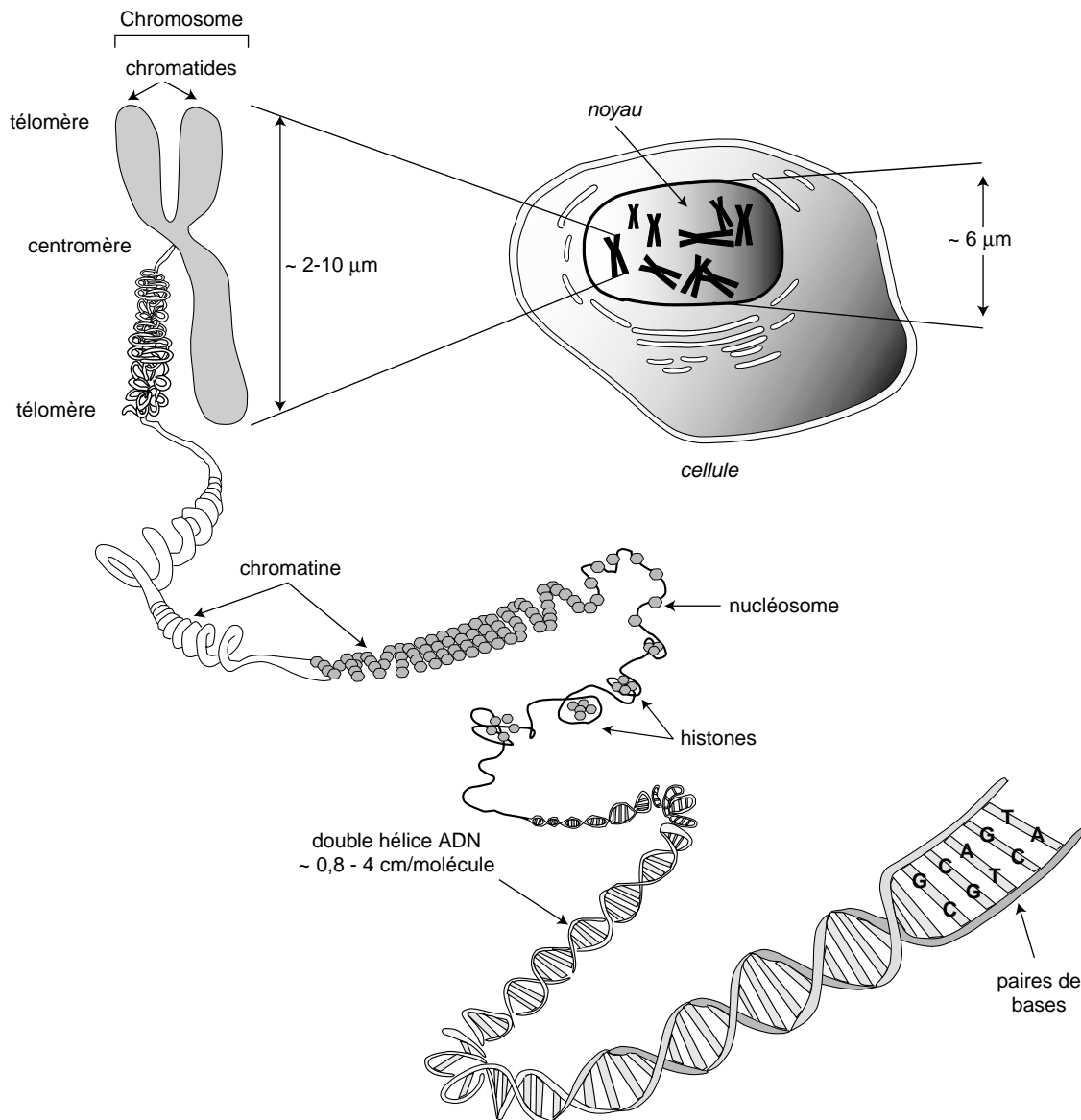
- durant l'interphase, cette spirale est relâchée, les

Figure 4. Structure d'un nucléosome.



chromosomes ne peuvent pas être distingués car ils sont très étirés et emmêlés, l'ensemble ressemblant à une pelote de laine. Néanmoins, il faut noter que

Figure 2. Structure d'un chromosome.



depuis l'utilisation d'ADN spécifique d'un chromosome comme sonde pour l'hybridation *in situ*, il est possible d'observer dans les noyaux interphasiques le domaine délimité qu'occupe un chromosome individuellement (Manuelidis 1985, Lichter *et al* 1988). L'interphase est une phase active de synthèse protéique, de synthèse des ARN et de réplication de l'ADN ;

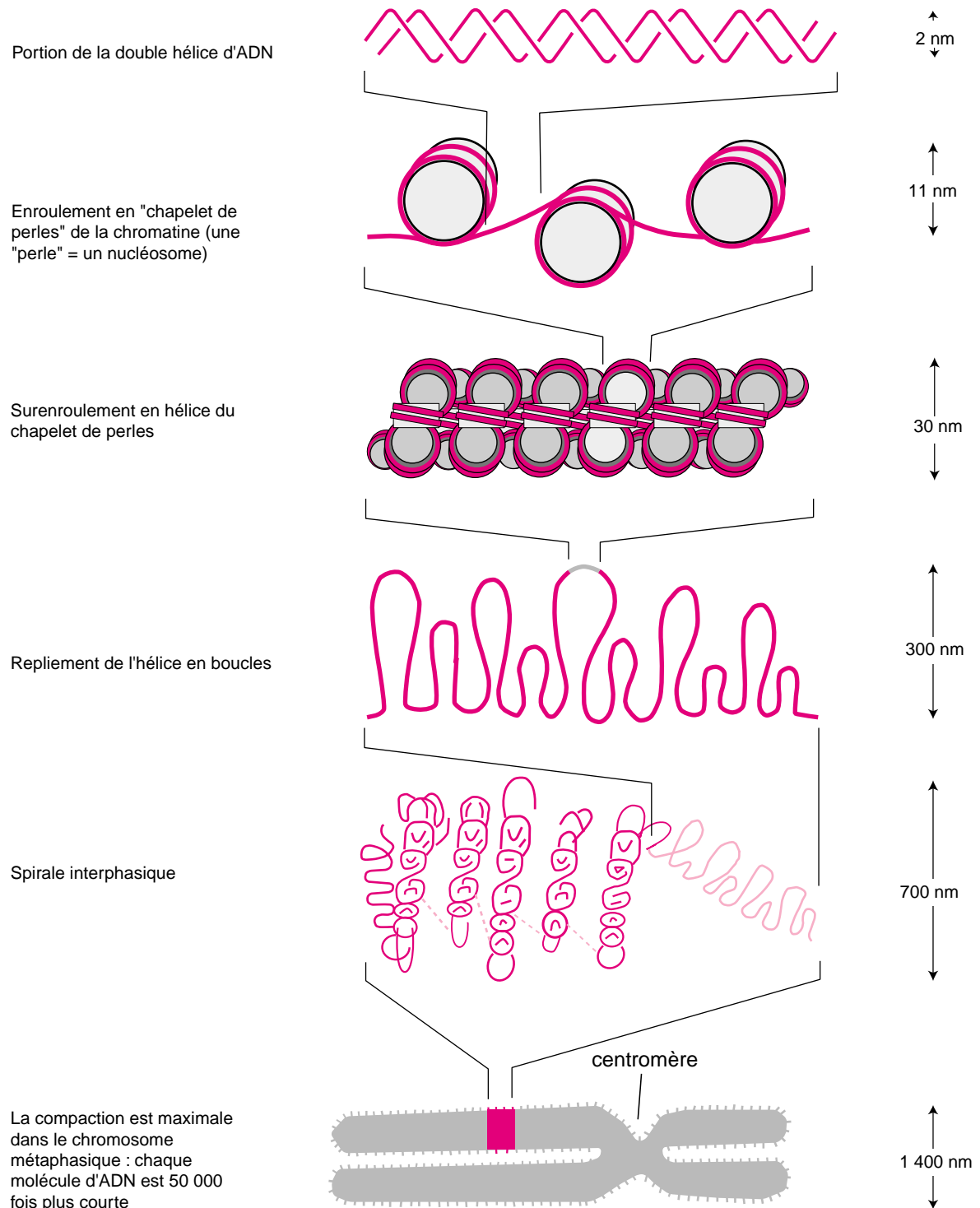
- durant la division cellulaire, la spirale de chromatine se condense encore beaucoup plus pour atteindre en métaphase un diamètre d'environ 700 nm et un degré de compaction maximal (plusieurs milliers de fois), rendant visible le chromosome (= 2 chromatides) comme une entité distincte.

## 2 / Régions ou structures révélées par les techniques de marquage chromosomique

Selon les traitements et colorations utilisés pour observer les chromosomes, on peut visualiser plusieurs types de régions ou de structures : l'euchromatine, l'hétérochromatine constitutive, les centromères et les régions organisatrices nucléolaires. L'euchromatine, l'hétérochromatine et les nucléoles sont également visibles en interphase.

### 2.1 / Euchromatine

L'euchromatine est révélée par les méthodes de marquage chromosomique produisant des motifs de bandes appelées Q/G, R et T et contenant, entre autres, presque tout l'ADN transcrit. Les dénominations des bandes Q (de quinacrine), G (de Giemsa), R (de reverse) et T (de terminale) proviennent des types de traitement et de coloration employés pour les produire. On distingue les bandes Q, G, R et T dites " structurales " parce qu'elles dépendent de la nature intrinsèque du chromosome et les bandes G et R dites " dynamiques " qui sont révélées par l'incorporation d'une base modifiée dans l'ADN de la cellule vivante et dépendent donc de la réplication. Les mêmes motifs de bandes G et R sont retrouvés. Dans tous les cas, on obtient une succession de bandes intensément colorées (positives + = foncées ou fluorescentes) ou légèrement colorées (négatives - = claires ou non fluorescentes) dont les motifs sont superposables (bandes Q+ fluorescentes = G+ foncées = R- claires et Q- non fluorescentes = G- claires = R+ foncées). Par convention, les termes bandes Q, G, R et T désignent les bandes colorées intensément ou positives et comme les motifs des bandes Q et G sont similaires, pour plus

**Figure 3.** Enroulement et compaction de l'ADN dans le chromosome.

de clarté nous ne ferons référence qu'aux bandes G et R (et parfois T, les bandes T constituant une sous-fraction des bandes R, particulièrement résistante à la dénaturation thermique, en raison de leur richesse en bases GC, et particulièrement riche en gènes).

L'euchromatine de tous les chromosomes des mammifères se subdivise en deux fractions à peu près égales de leur génome, les bandes G et les bandes R (+T). A chacune de ces fractions est associée une série de propriétés biochimiques et fonctionnelles opposées ; les principales sont regroupées dans le tableau 1 (Bickmore et Sumner 1989, Gardiner 1995).

D'après ce tableau, les bandes chromosomiques peuvent être considérées soit comme des segments d'ADN avec un contenu spécifique en paires de bases, en gènes et en séquences répétées dispersées, soit comme des unités fonctionnelles de réplication et de condensation. Ceci indique que le génome des mammifères est une structure organisée et que les caractéristiques fonctionnelles ne se distribuent pas de manière aléatoire. Jusqu'au milieu des années 80, il était admis que les bandes reflétaient une différenciation longitudinale de la structure nucléoprotéique des chromosomes, mais la nature de cette différenciation n'était pas connue. Différents modèles découpant le génome en sous-

**Tableau 1.** Principales propriétés biochimiques et fonctionnelles des bandes chromosomiques G/R.

| Bandes G   | Bandes R (+T)   |
|--|---|
| bandes G+ colorées                                     | bandes G- claires   |
| bandes Q+ fluorescentes                                | bandes Q- non fluorescentes                               |
| bandes R- et T- claires                                | bandes R+ et T+ colorées                                  |
| régions chromomériques au stade pachytène de la méiose | régions interchromomériques                               |
| réplication tardive durant phase S                     | réplication précoce                                       |
| riches en AT   | riches en GC  |
| pauvres en méthylcytosine                              | riches en méthylcytosine                                  |
| peu de gènes   | majorité des gènes de ménage et de spécificité tissulaire |
| séquences LINEs  | séquences SINEs   |

ensembles d'ADN avec des caractéristiques variables et une répartition différente selon les bandes G ou R et T ont été proposés. Ils permettent d'aller un peu plus loin dans la compréhension de l'organisation du génome et des mécanismes à l'origine des bandes chromosomiques.

**Modèle des isochores** (Bernardi *et al* 1985, Bernardi 1995)

Le génome nucléaire est divisé en 5 classes d'isochores selon leur composition en bases : L1 et L2 (riches en AT, plutôt dans les bandes G), H1, H2 et H3 (de plus en plus riches en GC, H1 et H2 plutôt dans les bandes R et H3 dans les bandes T).

**Modèle basé sur la répartition des séquences répétées dispersées SINEs et LINEs le long des chromosomes** (Korenberg et Rykowski 1988)

Les SINEs (Short Interspersed Repeated Sequences) sont des séquences inférieures à 1 kb réparties à travers le génome à raison de  $3.10^5$  à  $9.10^5$  copies et les LINEs (Long Interspersed Repeated Sequences) des séquences de 6 à 7 kb réparties à travers le génome à raison de  $10^4$  à  $10^5$  copies. Chez l'Homme, plus de 90 % de ces séquences correspondent pour les SINEs aux séquences Alu de 300 pb (ainsi nommées parce qu'elles possèdent généralement un site de restriction pour l'enzyme *AluI*) et pour les LINEs aux séquences L1 ou KpnI (ainsi nommées parce que certaines d'entre elles peuvent être isolées après digestion de l'ADN génomique avec l'enzyme de restriction *KpnI*). Il a été montré par hybridation *in situ* que ces séquences ne se répartissent pas au hasard le long des chromosomes humains, mais suivent une distribution caractéristique avec prédominance des séquences Alu (environ 18 % de leur ADN) dans les bandes R alors que les séquences L1 prédominent dans les bandes G (environ 14 % de leur ADN). Ceci suggère qu'elles sont impliquées au moins dans certaines des propriétés opposées des deux fractions de l'euchromatine.

**Modèle des « saveurs », de l'anglais « flavors »** (Holmquist 1992)

Dans ce modèle, le rapport entre composition en bases et richesse en séquences Alu est déterminé, ce qui divise la chromatine des bandes R du chromosome métaphasique en quatre classes ou

« saveurs »: riche en AT/riche en Alu, riche en AT/pauvre en Alu, riche en GC/riche en Alu et riche en GC/pauvre en Alu. La densité des gènes varie suivant ces quatre saveurs.

**Modèle de Craig et Bickmore (1994)**

L'ADN est fractionné en sous-ensembles selon la distance entre îlots CpG (îlots CpG non méthylés, donc coupés par des enzymes sensibles à la méthylation, situés surtout en 5' des gènes de ménage, d'où leur qualité d'indicateur de la densité en gènes de ménage). Quatre classes ont été définies : distance entre îlots CpG supérieure à 1000 kb, donc la moins riche en gènes, plutôt dans les bandes G, de 500 à 1000 kb plutôt dans les bandes R, et de 100 à 500 kb et de 15 à 100 kb, donc les plus riches en gènes, plutôt dans les bandes T.

**Modèle de Saitoh et Laemmli (1994)**

Dans ce modèle, le facteur déterminant serait une différence de densité de la chromatine selon les bandes, ce qui permettrait de relier l'organisation de la fibre de chromatine et l'alternance des bandes le long des chromosomes en métaphase. Comme il a été dit plus haut, la chromatine passe par différents degrés de compaction selon le stade du cycle cellulaire et, à l'un de ces degrés de compaction, la fibre de chromatine est organisée en boucles dont l'assise est liée à l'armature protéique par l'intermédiaire des régions « SAR ». Ces régions SAR contiennent des séquences nucléotidiques particulières, très riches en bases AT (plus de 65 %). Du fait de leur rôle d'ancrage des boucles de chromatine dans l'architecture chromosomique, les régions SAR seraient organisées de manière non aléatoire formant une structure centrale le long de l'axe de la fibre de chromatine appelée alignement AT. Les auteurs de ce modèle proposent que l'alternance de bandes intensément ou légèrement colorées résulte d'un arrangement et d'une compaction différents de l'alignement AT le long des chromosomes. Dans les bandes G, l'alignement AT serait plus compact, enroulé en spirale avec des boucles d'ADN courtes et orientées parallèlement à l'axe du chromosome, alors que dans les bandes R, l'alignement AT resterait sous une forme plus linéaire avec des boucles d'ADN plus longues et perpendiculaires à l'axe du chromosome. Par la méthode de marquage chromosomique des bandes G, les bandes G sont intensément colorées parce que le degré de compaction de l'alignement AT est élevé et les bandes G- (équivalentes aux bandes R) sont peu colorées parce que l'alignement AT est plus relâché.

## 2.2 / Hétérochromatine constitutive

L'hétérochromatine constitutive est colorée de manière sélective et intense par les méthodes de marquage dites de « bandes C » (de centromère) et, pour le moment, n'a pas de fonction connue. C'est une fraction de la chromatine qui ne se décondense pas durant l'interphase et qui se situe dans les régions juxtacentromériques de tous les chromosomes chez presque toutes les espèces, mais parfois dans d'autres régions aussi, par exemple sur le chromosome Y. Elle contient des séquences hautement répétées d'ADN riches en paires de bases AT ou GC selon les sites et les espèces et elle se réplique très tardivement durant la phase de synthèse S de l'ADN. Cependant, l'hétérochromatine n'est pas totalement dépourvue d'activité génétique puisque, chez la drosophile, on a pu isoler des mutations létales dans

des gènes localisés dans l'hétérochromatine (Eberl *et al* 1993). La cartographie génétique indique que l'hétérochromatine contient 100 fois moins de gènes que l'euchromatine. Par contre, on a montré que si un gène est déplacé, à la suite d'une translocation ou d'une inversion, d'une région euchromatique vers une région hétérochromatique, il peut être inactivé dans certaines cellules, probablement parce qu'il est soumis au processus de formation de l'hétérochromatine.

### 2.3 / Centromères

Les centromères sont visibles comme une constriction dite primaire sur chaque chromosome après coloration au Giemsa et sur laquelle on peut révéler, par des techniques d'immunofluorescence, les kinétochores qui sont des structures spécialisées par lesquelles les chromosomes se fixent aux fuseaux nucléaires mitotiques et méiotiques et migrent vers le plan équatorial de la cellule. Les centromères sont compris dans l'hétérochromatine constitutive.

### 2.4 / Régions organisatrices nucléolaires (NOR)

Les régions organisatrices nucléolaires colorées par la méthode Ag-NOR occupent des positions bien définies dans les chromosomes, parfois visibles comme une constriction secondaire, mais leur nombre varie d'un organisme à un autre. Ils participent à la formation et au fonctionnement des nucléoles dans les noyaux interphasiques, qui contiennent l'ARN et les protéines ribosomiques.

## 3 / Organisation des séquences transcrites et non transcrites sur le chromosome

La taille des génomes eucaryotes est beaucoup plus grande que celle des génomes procaryotes. On considère que seuls 3 à 4 % de l'ADN chromosomique humain correspondent à des séquences transcrites. Il a été montré que chez les eucaryotes, les gènes sont insérés dans un arrangement complexe de séquences répétées en tandem ou dispersées et de séquences non répétées dont la plupart n'a pas de fonction, certaines de ces séquences répétées sont aussi retrouvées dans les introns. L'ADN eucaryote peut donc se classer en plusieurs sous-ensembles : les gènes uniques, les familles de gènes et les pseudogènes, les séquences répétées fonctionnelles non codantes, les séquences répétées de fonction inconnue et les séquences non répétées non codantes. Certaines de ces séquences sont associées à des structures particulières du chromosome ou localisées dans certaines régions.

### 3.1 / Gènes uniques, familles de gènes et pseudogènes

La majorité des gènes et pseudogènes est située dans l'euchromatine de tous les chromosomes, bien que l'hétérochromatine n'en soit pas totalement dépourvue. Les études de Bernardi et collègues (Mouchiroud *et al* 1991, Saccone *et al* 1992, Bernardi 1995) ont montré que la densité des gènes n'est pas homogène le long des chromosomes (tableau 2).

**Tableau 2.** Distribution globale des isochores et des gènes dans les chromosomes humains selon les bandes T, R et G.

| Bandes           | T        | R        | G        |
|------------------|----------|----------|----------|
| Isochores        | ~H3      | ~H1 + H2 | ~L1 + L2 |
| % GC             | ~53      | ~46      | ~39      |
| % génome         | ~3       | ~31      | ~62      |
| % gènes          | 39       | 38       | 34       |
| ADN (Mb)         | 90       | 190      | 1860     |
| Densité de gènes | 1/3,2 kb | 1/25 kb  | 1/50 kb  |

Les bandes les plus riches en gènes correspondent majoritairement aux bandes T et en partie aux bandes R et elles sont peu nombreuses : environ 59/400 bandes du caryotype standard humain. La densité des gènes varie aussi entre chromosomes, ainsi certains chromosomes humains comme le chromosome 19 sont extrêmement riches en gènes et d'autres comme le chromosome 13, extrêmement pauvres en gènes (Saccone *et al* 1992).

L'étude publiée par Gardiner (1996), synthétisant les différents projets ayant pour objectif d'identifier un maximum de gènes du chromosome humain 21, présente des résultats similaires (tableau 3 et figure 5). La bande q22.3 terminale (R ou T) est la plus riche en gènes puisqu'elle contient seulement 19 % d'ADN mais 47 % des gènes identifiés et la bande q21 (G) est la plus pauvre en gènes avec 41 % de l'ADN et seulement 5 % des gènes identifiés.

Enfin, les membres d'une famille multigénique peuvent être :

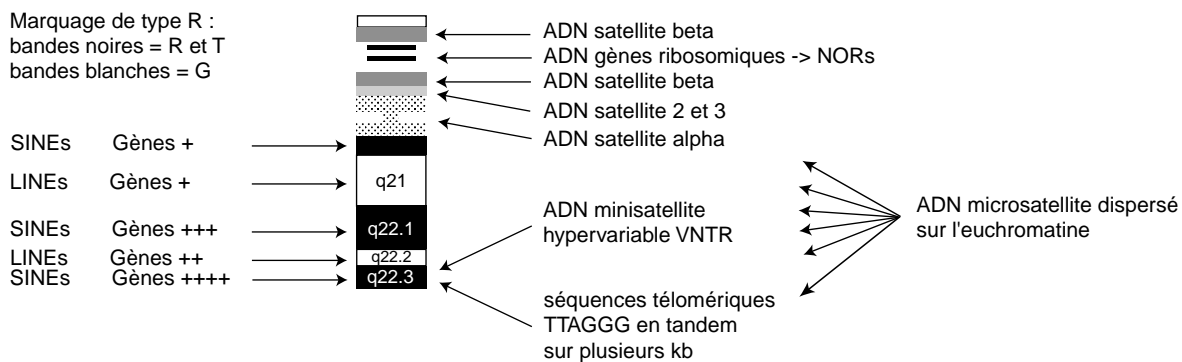
- soit regroupés à un endroit spécifique d'un chromosome, par exemple, les 7 gènes (HBA1, HBA2, HBQ1, HBZ, HBAP1, HBAP2 et HBZP) de la famille des globines alpha contenus dans un segment d'ADN d'environ 50 kb situé sur le chromosome humain 16p13.3 ;

- soit dispersés sur plusieurs chromosomes, par exemple chez l'Homme, les 3 clusters (RN5S1, RN5S2 et RN5S3) des gènes spécifiant les ARN ribosomiques 5S, localisés sur 3 sites différents des chromosomes 1 et 6, ou les 9 gènes PAX (PAX1, 2, 3, 4, 5, 6, 7, 8 et 9) localisés sur 8 chromosomes différents, chacun à un site propre.

Pour les espèces de mammifères autres que l'Homme, les cartes géniques sont à ce jour insuffisamment développées pour donner une image globale et significative mais la répartition et la densité des gènes le long des chromosomes ont très probablement les mêmes caractéristiques chez toutes les espèces. Cette hypothèse est soutenue entre autres,

**Tableau 3.** Répartition des gènes du chromosome humain 21 selon les bandes G et R.

| Bandes bras long q | % ADN | % gènes identifiés |
|--------------------|-------|--------------------|
| R q11.2            | 11    | 2                  |
| G q21              | 41    | 5                  |
| R q22.1            | 20    | 30                 |
| G q22.2            | 9     | 16                 |
| R ou T q22.3       | 19    | 47                 |

**Figure 5.** Distribution des différents types de séquence le long du chromosome humain 21.

par les expériences de peinture chromosomique hétérologue (Gellin 2000, cet ouvrage) qui révèlent l'existence de grands segments chromosomiques conservés entre génomes d'espèces différentes, la distribution de ces segments au sein des caryotypes étant spécifique d'espèce (Wienberg et Stanyon 1997).

### 3.2 / Séquences répétées fonctionnelles non codantes

Les séquences répétées fonctionnelles non codantes sont celles situées aux deux extrémités de chacun des chromosomes de toutes les espèces eucaryotes étudiées de la levure à l'Homme (Pryde *et al* 1997). Elles consistent en un alignement de répétitions hexanucléotidiques (TTAGGG chez l'Homme) en tandem qui ne spécifient aucune protéine ou ARN, mais qui ont un rôle essentiel dans la fonction des télomères. Elles protègent les extrémités des chromosomes (riches en gènes, cf paragraphe précédent) des dégradations ou pertes et elles interviennent dans le mécanisme de réplication de l'extrémité de la molécule linéaire d'ADN des chromosomes. Leur localisation peut être déterminée par hybridation *in situ*. Il a été montré que la taille de ces alignements télomériques diminue avec le vieillissement cellulaire, c'est-à-dire au fur et à mesure des divisions cellulaires, ce qui explique leur instabilité (Greider 1996).

### 3.3 / Séquences répétées de fonction inconnue

Les séquences d'ADN répétées de fonction inconnue ont généralement un nombre de répétitions beaucoup plus grand que les séquences télomériques et sont présentes en plus grande quantité dans le génome (environ 20 % du génome humain). Il en existe plusieurs classes.

#### ADN satellite

Il existe plusieurs types d'ADN satellite qui comprennent de longues séries de répétitions d'ADN plus ou moins complexes et non transcrites et qui constituent la majeure partie des régions hétérochromatiques des génomes.

Chez les primates (Willard et Waye 1987), dont l'Homme, les régions centromériques de tous les chromosomes sont riches en ADN répété de type satellite alpha ou alpoïde qui est composé de séries de monomères de 171 pb répétés en tandem sur plusieurs centaines de kb et pouvant représen-

ter jusqu'à 3 à 5 % de l'ADN de chaque chromosome. Il existe d'autres ADN satellites situés dans les régions adjacentes ou subcentromériques de certains chromosomes : ADN satellite 1, 2, 3 et beta (exemple du chromosome humain 21 dans la figure 5). Chez d'autres mammifères comme la souris, le mouton et la vache, l'ADN centromérique est également constitué de séquences répétées en tandem qui semblent spécifiques de l'espèce (Singer 1982, Burkin *et al* 1996), mais leur organisation n'est pas aussi bien connue.

#### ADN minisatellite ou VNTR

Les VNTR ou répétitions en tandem en nombre variable (Jeffreys *et al* 1985) varient beaucoup en taille mais partagent toutes un court motif d'une dizaine de paires de bases. Ce qui les caractérise particulièrement, c'est la très grande variabilité du nombre de répétitions en tandem d'un individu à l'autre (0,5 à 30 kb en moyenne), d'où leur utilisation principale pour réaliser des empreintes d'ADN spécifiques d'individu. Les VNTR se trouvent toujours dans l'euchromatine, un grand nombre est situé dans les régions proches des télomères mais ils peuvent être présents à d'autres sites chromosomiques.

#### ADN microsatellite

L'ADN microsatellite est constitué d'alignements de répétitions très courtes (2 à 5 nucléotides), généralement des dinucléotides CA, très nombreux dans le génome et souvent très polymorphes de par le nombre variable de répétitions. Ils sont uniformément dispersés dans toute la partie euchromatique des génomes. Ces qualités font qu'ils sont les marqueurs génétiques les plus utilisés pour la cartographie des génomes. L'ADN microsatellite n'a pas de fonction connue.

#### Séquences SINEs et LINEs

Les séquences répétées dispersées courtes (SINEs) et longues (LINEs) et leur répartition le long des chromosomes ont été présentées dans le paragraphe 2.1.

### 3.4 / Séquences non répétées non codantes

Ce sont des séquences non répétées situées entre les gènes et très peu caractérisées à ce jour. Elles contiennent probablement les séquences régulatrices à proximité des gènes, mais elles servent peut-être aussi à séparer plus ou moins les gènes.

En conclusion, il existe différents types de séquences codantes ou non, répétées ou non le long de la molécule d'ADN, qui ne se distribuent pas nécessairement de manière aléatoire le long des chromosomes métaphasiques. L'état actuel des connaissances

sur ce sujet est résumé sur la figure 5 en prenant pour exemple le chromosome humain 21. La majorité des études réalisées à ce jour concerne les chromosomes humains, mais il semble qu'une organisation comparable se retrouve chez beaucoup de mammifères.

## Références

- Bernardi G., 1995. The human genome: organisation and evolutionary history. *Annual Review of Genetics*, 29, 445-476.
- Bernardi G., Olofsson B., Filipski J., Zerial M., Salinas J., Cuny G., Meunier-Rotival M., Rodier F., 1985. The mosaic genome of warm-blooded vertebrates. *Science*, 228, 953-958.
- Bickmore W.A., Sumner A.T., 1989. Mammalian chromosome banding - an expression of genome organisation. *Trends in Genetics*, 5, 144-148.
- Burkin D.J., Broad T.E., Jones C., 1996. The chromosomal distribution and organization of sheep satellite I and II centromeric DNA using characterized sheep-hamster somatic cell hybrids. *Chromosome Research*, 4, 49-55.
- Craig J.M., Bickmore W.A., 1994. The distribution of CpG islands in mammalian chromosomes. *Nature Genetics*, 7, 376-381.
- Eberl D.F., Duyf B.J., Hilliker A.J., 1993. The role of heterochromatin in the expression of a heterochromatic gene, the rolled locus of *Drosophila melanogaster*. *Genetics*, 134, 277-292.
- Gardiner K., 1995. Human genome organization. *Current Opinion in Genetics and Development*, 5, 315-322.
- Gardiner K., 1996. Base composition and gene distribution: critical patterns in mammalian genome organization. *Trends in Genetics*, 12, 519-523.
- Gasser S.M., Laemmli U.K., 1987. A glimpse at chromosomal order. *Trends in Genetics*, 3, 16-22.
- Gellin J., 2000. La cartographie comparée des génomes des vertébrés. INRA Productions Animales, numéro hors série « Génétique moléculaire : principes et application aux populations animales », 95-102.
- Greider C.W., 1996. Telomere length regulation. *Annual Review of Biochemistry*, 65, 337-365.
- Holmquist G.P., 1992. Review article: chromosome bands, their chromatin flavors, and their functional features. *American Journal of Human Genetics*, 51, 17-37.
- Korenberg J.E., Rykowski J.R., 1988. Human genome organization : Alu, Lines, and the molecular structure of metaphase chromosome bands. *Cell*, 53, 391-400.
- Jeffreys A.J., Wilson V., Thein S.L., 1985. Hypervariable minisatellite regions in human DNA. *Nature*, 314, 67-73.
- Lichter P., Cremer T., Borden J., Manuelidis L., Ward D.C., 1988. Delineation of individual human chromosomes in metaphase and interphase cells by in situ suppression hybridization using recombinant DNA libraries. *Human Genetics*, 80, 224-234.
- Manuelidis L., 1985. Individual interphase chromosome domains revealed by in situ hybridization. *Human Genetics*, 71, 288-293.
- Mouchiroud D., D'Onofrio G., Aïssani B., Macaya G., Gautier C., Bernardi G., 1991. The distribution of genes in the human genome. *Gene*, 100, 181-187.
- Pryde F.E., Gorham H.C., Louis E.J., 1997. Chromosome ends: all the same under their caps. *Current Opinion in Genetics and Development*, 7, 822-828.
- Saccone S., De Sario A., Della Valle G., Bernardi G., 1992. The highest gene concentrations in the human genome are in the T-bands of metaphase chromosomes. *Proceedings of the National Academy of Sciences USA*, 89, 4913-4917.
- Saitoh Y., Laemmli U.K., 1994. Metaphase chromosome structure: Bands arise from a differential folding path of the highly AT-rich scaffold. *Cell*, 76, 609-622.
- Singer M., 1982. Highly repeated sequences in mammalian genomes. *International Review of Cytology*, 76, 67-112.
- Wienberg J., Stanyon R., 1997. Comparative painting of mammalian chromosomes. *Current Opinion in Genetics and Development*, 7, 784-791.
- Willard H.F., Wayne J.S., 1987. Hierarchical order in chromosome-specific human alpha satellite DNA. *Trends in Genetics*, 3, 192-198.