

Cour 06 : carte génétique, carte physique et logiciel de cartographie

Introduction

- **La cartographie génomique : pourquoi faire, comment ?**

Les génomes ont été longtemps considérés comme un ensemble de gènes n'entretenant entre eux que des relations fonctionnelles. Le séquençage complet des génomes a relancé l'étude des relations spatiales entre gènes (et autres éléments génétiques) le long des génomes. La prise en compte de l'arrangement linéaire des gènes a montré à la fois son intérêt dans la compréhension des mécanismes évolutifs et son efficacité dans l'exploration de mécanismes biologiques aussi divers que la recombinaison, la recherche de gènes orthologues ou encore la fluidité des génomes. Plus récemment l'accès à l'arrangement spatial des chromosomes au sein d'une cellule a mis en avant l'intérêt d'une cartographie en 3D. Il apparaît en effet de plus en plus clairement que la proximité spatiale de gènes distants dans le génome, à un sens biologique.

Avant de faire un bilan des différents types de cartes utilisées sur les génomes, les concepts communs peuvent être rapidement cités.

Chromosome : Le génome des organismes est (dans la plupart des cas chez les eucaryotes) constitué de plusieurs molécules d'ADN. Chacune est intégrée au sein d'une structure complexe impliquant de multiples protéines. Deux structures jouent un rôle important à la fois fonctionnel et dans la description des chromosomes ce sont le centromère et les télomères. Les chromosomes ont deux bras de part et d'autre du centromère un bras court (p) et un bras long (q). Si les deux bras ont des longueurs voisines le chromosome est dit métacentrique, si au contraire le bras p est réduit à une zone très courte principalement hétérochromatique le chromosome est acrocentrique. Enfin si le centromère est à l'extrémité du chromosome celui-ci est dit télocentrique. Il faut enfin signaler le cas d'organismes n'ayant pas de centromère clairement identifié, on parle alors de chromosomes holocentriques (c'est entre autres le cas des lépidoptères).

Marqueur : n'importe quelle structure pouvant être localisée sur un génome (un gène, une séquence anonyme, une région colorée par un produit, ...). En pratique deux classes de marqueurs peuvent être définies en fonction de l'importance ou pas qu'il soit polymorphe. Les marqueurs peu polymorphes ont une localisation qui peut être facilement déterminée (par hybridation expérimentale ou *in silico*), ils peuvent ainsi servir de points d'ancrage pour la cartographie. Les marqueurs fortement polymorphes sont nécessaires à la fois pour les cartes génétiques et pour les études populationnelles (détermination de paternité, flux de gènes, ...).

méiose : Division cellulaire conduisant à la formation des gamètes ...

Carte : c'est une succession de marqueurs ou d'ensembles de marqueurs. Certaines cartes sont associées à une distance entre marqueurs.

Segment conservé : la notion la plus floue ! En théorie un ensemble de gènes orthologues synténiques et dont l'ordre dans les deux chromosomes est le même. Les difficultés sont de deux sortes :

- un relachement plus ou moins important de la contrainte d'ordre
- la prise en compte des gènes sans orthologues connus ou dupliqués différemment dans les deux génomes.

En pratique il est impossible actuellement d'utiliser ce terme sans donner la définition que l'on utilise.

Réarrangements : les cartes génomiques évoluent sous l'effet de réarrangements du génome. Ces réarrangement génomiques sont de plusieurs natures et surviennent au cours de l'évolution avec des probabilité différentes dépendant à la fois de la nature du réarrangement et du taxon en cause. Les données permettant de mesurer l'impact évolutif des différents réarrangements sont à la fois dispersées et de natures diverses. Les données médicales fournissent une vue relativement exhaustive des réarrangements associées à des dysfonctionnements (maladies génétiques) non léthaux mais donnant lieu à des symptômes.

Le polymorphisme 'caryotipique' est largement étudié dans de nombreux taxons en particulier celui des rongeurs. Dans ce contexte les réarrangements Robertsioniens sont clairement les plus fréquents, même se ce ne sont pas les seuls.

EST (Expressed Sequence Tags) :

Marqueurs de séquences exprimées. Les séquences EST, issues du séquençage systématique des extrémités d'ADNc partiels, sont des séquences STS codantes. Une séquence EST est donc une séquence partielle d'ADNc de quelques centaines de nucléotides, correspondant à une des extrémités d'un ARNm et qui lui est spécifique. Elle comporte le plus souvent de nombreuses erreurs.

Polymorphisme

Hétérogénéité allélique au sein d'une population. Elle peut se détecter au niveau du caractère phénotypique, au niveau chromosomique ou moléculaire. Polymorphisme nucléotidique : voir SNP.

QTL (Quantitative Trait Locus)

Locus dont l'unité de fonction contribue, pour une part plus ou moins importante, à l'élaboration d'un caractère quantitatif (locus responsable de la variabilité de ce caractère). Plusieurs loci intervenant dans la réalisation d'un même caractère sont souvent groupés dans une même région chromosomique. Les QTL sont étudiés tout particulièrement chez les plantes : les marqueurs génétiquement liés à de tels QTL permettent de sélectionner, parmi un grand nombre de plantes, les individus les plus performants. Ils peuvent ainsi être utilisés pour construire un génotype idéal par croisements successifs ou pour améliorer l'évaluation de la valeur des individus.

Pedigree

Représentation arborescente du phénotype des membres d'une famille (arbre généalogique). L'étude de ces pedigrees peut seule permettre l'analyse génétique lorsque des croisements expérimentaux sont irréalisables.

Synténie

Groupes de liaison conservés entre les génomes de deux espèces. Conservation de l'ordre de gènes homologues dans certaines portions de l'ADN d'espèces différentes.

- **Les différents types de cartes**

La diversité des cartes génomiques résulte à la fois de la diversité des approches expérimentales mais aussi de l'existence de plusieurs concepts biologiques de distances entre marqueurs génomiques. Si la distance physique comptée en nombre de bases séparant les deux marqueurs semble la plus naturelle, il faut être conscient que cette distance n'a pas de réel sens biologique. Le nombre moyen de crossing over est beaucoup plus significatif à la fois en termes évolutifs que génétique.

Dans tous les organismes étudiés, les généticiens ont pour démarche de localiser les gènes pour identifier leur nature et leur rôle dans le contrôle des différents caractères phénotypiques.

La cartographie des gènes est la détermination de la position d'un locus (gène ou marqueur génétique) sur un chromosome. Il y a deux approches fondamentalement différentes pour

localiser les gènes sur les chromosomes humains : la cartographie physique et la cartographie génétiques.

La carte génétique est une représentation graphique de la position des gènes d'un chromosome. Elle se différencie essentiellement de la carte physique, du fait qu'elle saisit l'ordre de localisation des gènes. Les distances exactes entre ces derniers sont, en effet, saisies sur la carte physique.

La cartographie génétique :

Permettant de représenter l'emplacement des marqueurs ou des gènes sur un chromosome, la cartographie génétique est une véritable révolution dans le domaine de la recherche médicale. Se présentant sous forme d'un graphique, la carte génétique démontre quelques différences par rapport à la carte physique.

La cartographie génétique est la construction d'une carte soit localisée autour d'un gène, soit à base large portant sur le génome entier. Plus généralement, c'est la détermination de la position d'un locus (gène ou marqueur génétique) sur un chromosome en fonction du taux de recombinaison génétique. Cette fréquence de recombinaisons méiotiques est utilisée pour estimer les distances entre les marqueurs. La distance génétique est une mesure statistique estimée en centi-Morgan (cM).

La première carte génétique est réalisée en 1913 par Thomas Hunt Morgan et Alfred Sturtevant sur le chromosome X de la drosophile. La première plante cartographiée est le maïs en 1935 grâce à la technique de marqueurs moléculaires. Ces cartes se basaient essentiellement sur la ségrégation de caractères phénotypiques disposés linéairement sur un chromosome et de l'existence de crossing-over se produisant a priori aléatoirement le long de celui-ci. Ainsi, plus la distance séparant deux gènes portés par le même chromosome est importante, **plus la probabilité de recombinaison entre ces deux gènes est élevée**. La mesure du taux de recombinaison reflète donc la distance linéaire entre les gènes. Le centimorgan (cM), correspondant à 1 % de recombinaison, fut adopté comme unité de distance génétique.

L'analyse de liaison, en suivant la transmission de caractères polymorphes, permet ainsi d'établir la position relative des gènes sur le chromosome et de mesurer la distance qui les sépare.

Il y'a 2 aspects fondamentaux en cartographie génétique :

- La détermination de l'ordre dans lequel les unités génétiques sont disposées les unes par rapport aux autres. On peut aussi représenter les marqueurs génétiques. Exemple : Les

marqueurs RFLP utilisés dans la cartographie de la pomme de terre sont répartis sur les 12 chromosomes de la pomme de terre.

- La détermination de la distance relative entre 2 unités génétiques par la carte physique.

La cartographie génétique est aussi utilisée pour étudier l'assortiment de caractères par analyse génétique ou l'hérédité.

1-Les recombinants et les non recombinants

Les paires de chromosomes homologues s'apparient au cours de la méiose I et subissent un certain nombre de recombinaisons ; échangeant des sections homologues par crossing-over (CO) et créant de nouvelles combinaisons d'allèles dans les produits de la méiose. Le crossing-over a lieu au moment de la méiose où il existe quatre chromatides (tétrade). Cependant, chaque crossing-over implique seulement 2 des 4 chromatides. Pour chaque recombinaison, il existe quatre produits : deux qui sont recombinants (R) et deux non recombinants (NR).

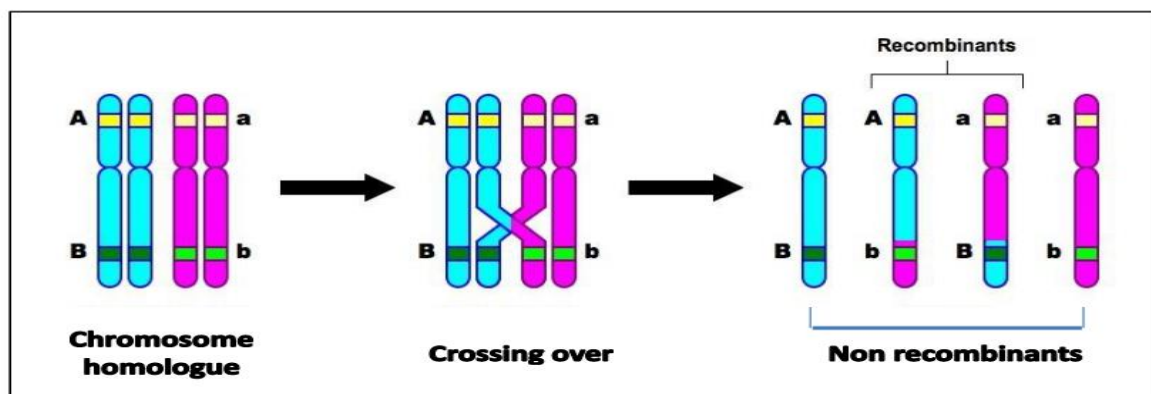


Figure : les recombinants et les non recombinants

2-Ségrégation des gènes

Il y a trois façons possibles pour une paire d'allèle de ségréger en méiose:

2-1- Premier cas de ségrégation (gènes indépendants)

Les allèles situés sur différents chromosomes se combinent de façon indépendante en méiose. Pour deux loci A et B avec les allèles A/a, et B/b, nous observons quatre types de gamètes en proportions égales (25%), les deux gènes seront considérés comme non liés et la ségrégation devient de type indépendant. On dit que deux gènes sont indépendants quand leur distance génétique est supérieure de 50 cM ($D \geq 50 \text{ cM}$, $P=R=50\%$). Ceci peut signifier que les deux gènes sont sur des chromosomes différents ou très éloignés sur le même chromosome.

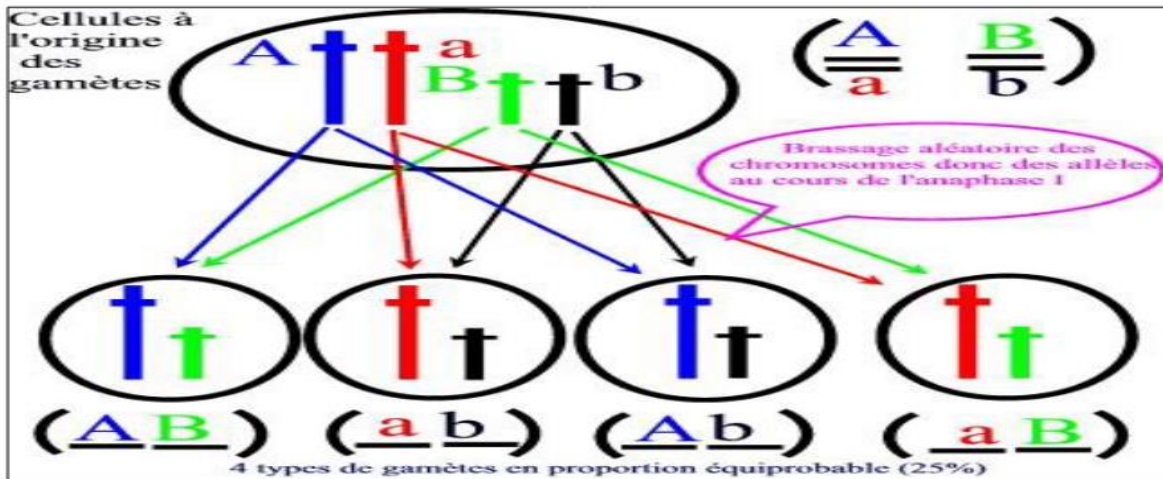


Figure : les gènes indépendants

2-2- Deuxième cas de ségrégation (liaison génétique)

Les gènes situés côte à côte sur le même chromosome sont transmis pratiquement toujours ensemble. En conséquence, seuls deux des combinaisons alléliques possibles vont être observées dans la descendance ; lesquelles dépendront de la combinaison particulière des chromosomes parentaux AB et ab. L'assortissement ne sera pas indépendant et aucune recombinaison ne sera observée. La probabilité d'avoir un crossing-over entre les deux gènes tend à s'annuler lorsqu'ils sont proches ($D < 10$ cM), il s'agit d'une liaison totale (pas d'échange).

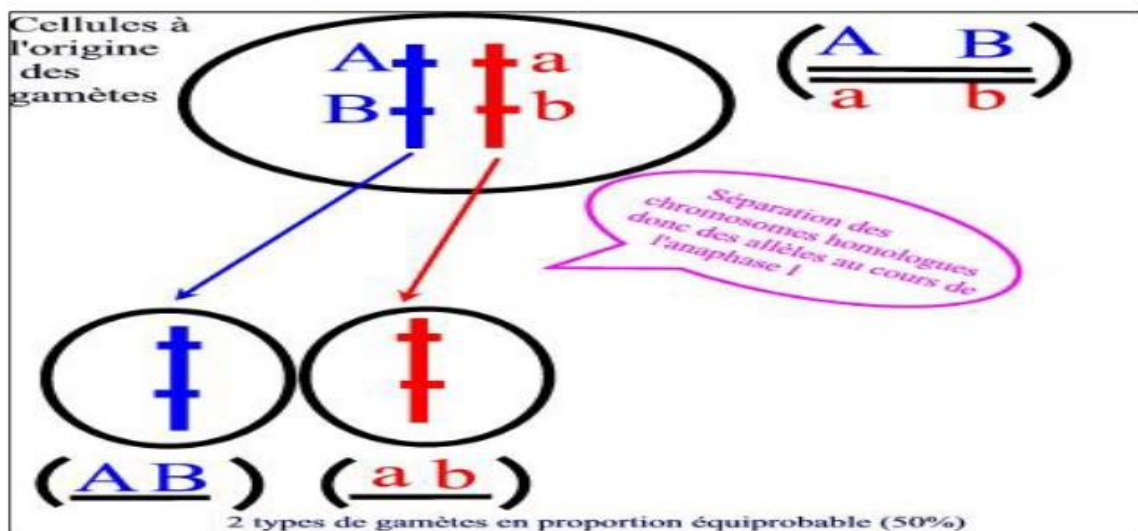


Figure : liaison absolue des gènes

2-3- Troisième cas de ségrégation (gènes liés)

Entre ces deux extrêmes, les allèles de deux loci localisés à une certaine distance l'un de l'autre sur le même chromosome tendent à être transmis ensemble, au moins qu'une recombinaison au cours de la méiose ne crée une nouvelle combinaison. Dans ce cas, nous observons quatre combinaisons d'allèles dans la descendance. Les proportions relatives de ces combinaisons dépendront de fréquence de recombinaison entre les deux loci. On dit que deux gènes sont liés quand leur distance est inférieure à 50cM ($P > 50\%$ et $R < 50\%$). Ceci signifie que les gènes sont proches sur le même chromosome.

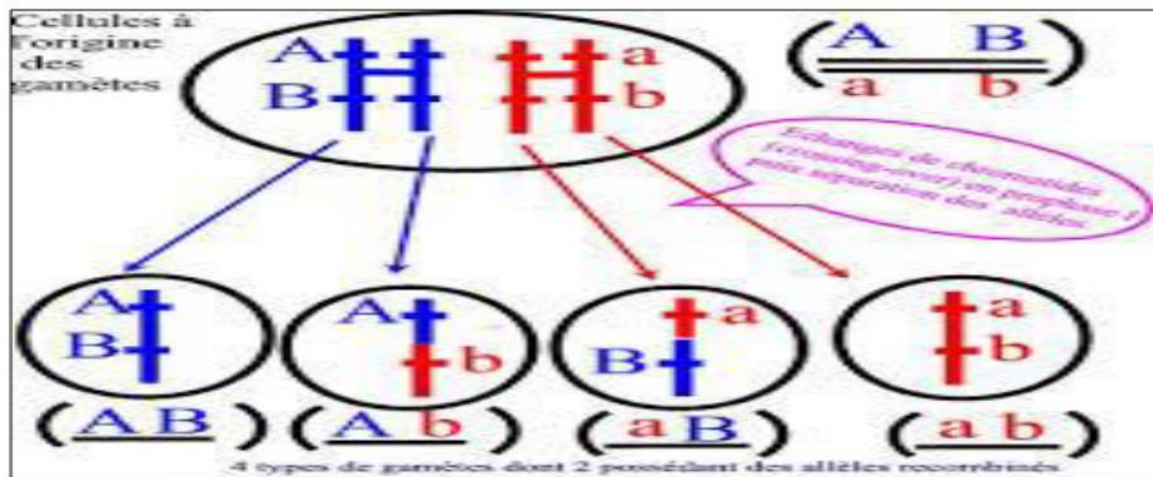


Figure : les gènes liés

3-La liaison génétique

La liaison génétique peut être définie comme la tendance pour deux allèles situés proches l'un de l'autre sur le même chromosome d'être transmis ensemble comme une seule unité au cours de la méiose. La force de liaison peut alors être utilisée comme unité de mesure permettant d'estimer la proximité de deux loci distincts. Cette unité de mesure est un reflet de la distance physique.

4-La distance génétique

La distance génétique (D) est la distance qui sépare deux loci sur un chromosome.

La distance entre deux loci peut être déduite en estimant la fréquence de recombinaisons survenant dans les familles. 1cM est approximativement égal à une fréquence de recombinaison de 1%. La distance génétique est proportionnelle à la fréquence de recombinaison entre deux gènes liés. L'unité de la distance génétique est le centi Morgan (cM).

Pour calculer les distances entre deux régions, il faut donc calculer le pourcentage de chromatides recombinés parmi la totalité des chromatides.

Pour calculer la distance entre deux gènes, il faut donc déterminer le nombre de chromatides ayant une des deux combinaisons parentales (P1 et P2) et le nombre de celles ayant une des deux combinaisons recombinées (R1 et R2).

La distance génétique en cM sera alors égale à $100 \times (R1+R2)/(P1+P2+R1+R2)$.

$$D (A-B) = \text{fréquence des gamètes recombinants pour les gènes A et B} \times 100$$

$$D = \frac{\text{Nombre de recombinants}}{\text{Nombre total de descendants}} \times 100$$

Exemple : Supposons que l'on réalise un test-cross sur un individu double hétérozygote AC/ac et que l'on trouve dans les phénotypes de la descendance

37% AC (phénotype dominant sur les deux loci)

37% ac (phénotype récessif sur les deux loci)

13% Ac (phénotype dominant sur le premier locus et récessif sur le deuxième locus)

13% aC (phénotype récessif sur le premier locus et dominant sur le deuxième locus)

D'une manière évidente les deux derniers phénotypes dont le génotype correspondant est Ac/ac et aC/ac sont produits par des gamètes recombinés des doubles hétérozygotes. Ainsi 26% de tous les gamètes (13%+13%) sont de types recombinés et la distance entre les loci A et C est estimée à 26 centimorgan ou unités de carte génétique.

5- L'équivalence entre la distance génétique et la distance physique

1 cM est une unité de longueur génétique sur laquelle on observe une recombinaison une fois sur 100. La longueur totale du génome a été estimée à 3000 cM sur la base du nombre de chiasmats observés au cours de la méiose I dans la spermatogénèse. Puisque le génome haploïde (gamète) correspond à une longueur physique d'environ 3×10^9 pb donc 1cM est équivalent à environ 1 million de pb.

$$1\text{cM} = 10^6 \text{ pb}$$

Selon cette estimation, il est suggéré qu'un chromosome moyen est long de 100 à 300 cM. On s'attend par conséquent à observer en moyenne 1 à 3 recombinaisons par chromosome, par méiose. Cependant, il ne s'agit que d'une relation très approximative.

Plusieurs facteurs sont susceptibles d'influencer les taux de CO:

- les CO sont environ 1,5 fois plus fréquents au cours de l'ovogenèse que pendant la spermatogenèse
- les CO tendent à être plus fréquents à proximité des télomères des chromosomes que dans les régions des centromères
- certaines régions chromosomiques montrent des taux de CO particulièrement plus élevés, ces régions sont qualifiées de points chauds de recombinaison (hot spot); les séquences Alu semblent particulièrement sujettes à des CO.

La cartographie physique

La cartographie physique fait appel à une variété de méthodes cytogénétiques et moléculaires réalisées sur des cellules somatiques dans le laboratoire et capables de localiser les gènes le long des chromosomes. La position des gènes sur la carte est décrite en unité de mesure physique et réelle. La distance physique se mesure en paires de bases (pb) et kilo, mega ou giga pb.

Selon les techniques de la cartographie physique, le niveau de résolution peut varier du chromosome entier à la paire de base. Nous distinguerons deux classes de cartographie physique selon le niveau de résolution :

➤ **La cartographie physique à faible résolution** dont la plus petite unité de carte pouvant être résolue varie habituellement entre une et quelques méga bases d'ADN.

Les techniques de la cartographie physique à faible résolution sont :

- les hybrides cellulaires somatiques
- les hybrides mono-chromosomiques

- hybrides de translocation et de délétion
- cartographie par hybrides d'irradiation
- anomalies chromosomiques de structure
- cartographie par dosage génique
- cartographie des gènes par hybridation in situ
- l'hybridation in situ en fluorescence standard (FISH)

➤ **La cartographie physique à haute résolution** : la résolution est habituellement très grande, de quelques centaines de kilobases jusqu'à un seul nucléotide. Les techniques de haute résolution sont :

- La FISH à haute résolution
- la cartographie de restriction
- le séquençage.

1-La cartographie physique à faible résolution

1-1-Les hybrides cellulaires somatiques

C'est une méthode qui a joué un rôle important dans la localisation de nombreux gènes pathologiques. Elle exploite la propriété des cellules somatiques provenant de différentes espèces, de parfois fusionner pour former des cellules hybrides lorsqu'elles sont cultivées en présence de substance comme le polyéthylène glycol (PEG) ou le virus Sendai.

Pour la cartographie humaine, la technique consiste à construire, en général des hybrides cellulaire en fusionnant des cellules humaines (46 chromosomes) et des cellules de rongeurs (habituellement la souris ou le hamster, 40 chromosomes). Les produits initiaux de fusion sont appelés hétérocaryons, car les cellules contiennent à la fois un noyau humain et de rongeur. Les hétérocaryons peuvent parfois entrer en mitose, ce qui entraîne la dissolution des deux enveloppes nucléaires. Les chromosomes humains et de rongeurs sont alors mélangés en un seul noyau. La plupart des chromosomes humains ne se répliquent pas lors des cycles de division suivants et sont perdus. Cela peut donner naissance à différentes lignées cellulaires hybrides, contenant la totalité des chromosomes de rongeurs plus un petit nombre de

chromosomes humains. La base de la perte ou de la rétention chromosomique est encore inconnue.

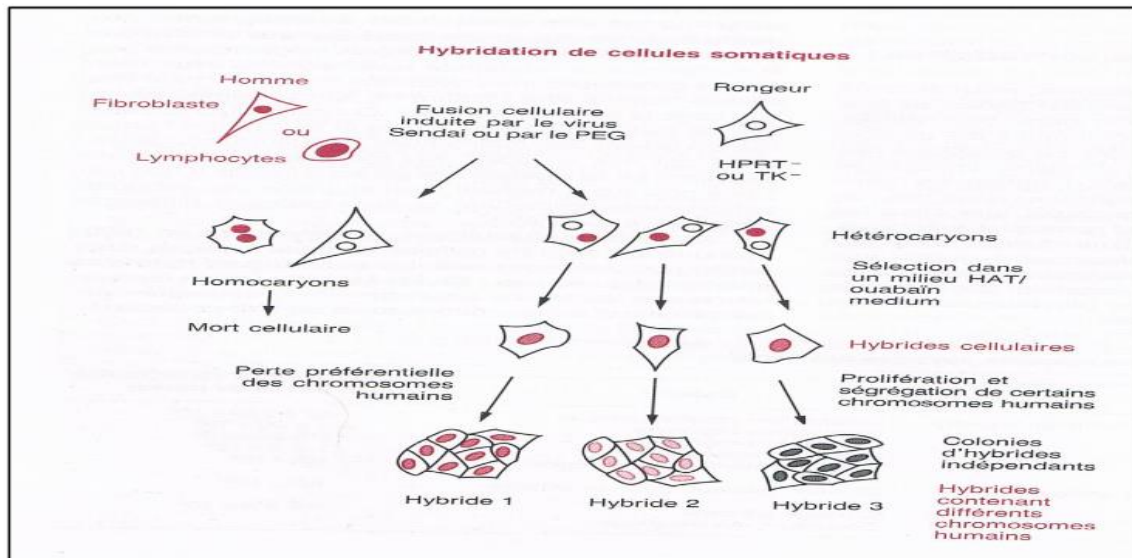


Figure : technique d'hybridation interspécifique de cellules somatiques

Les cellules hybrides retiennent un nombre différent et des combinaisons différentes des chromosomes humains. La distinction des chromosomes humains des chromosomes de rongeurs peut être déterminée par les techniques de cytogénétique : le caryotype.

Des panels de telles cellules sont utilisés pour établir une corrélation entre la présence d'un gène et la présence constante d'un chromosome. La présence du gène dans la cellule hybride peut être détectée de différentes manières selon que le produit du gène soit connu ou inconnu.

- **le produit du gène est connu**

Si le gène code pour une enzyme qui est produite par la cellule, alors un dosage enzymatique peut être utilisé, une électrophorèse de protéines peut également être utilisée pour détecter un produit du gène humain, et pour le distinguer de la protéine de la souris.

- **le produit du gène est inconnu**

Le plus souvent aujourd'hui, les investigateurs effectuent une hybridation des lignées cellulaires fusionnées avec une sonde marquée contenant l'ADN auquel s'intéresse, en utilisant les techniques de Southern blot, ou de PCR. La présence de ce gène humain peut être corrélée avec la présence ou l'absence d'un chromosome spécifique.

1-2-Les hybrides monochromosomiques

Le désavantage des hybrides cellulaires somatiques traditionnels est que les cellules hybrides contiennent habituellement plusieurs chromosomes humains et non un seul. Afin de limiter la quantité du matériel génétique humain transféré dans la cellule receveuse de rongeur, on peut appliquer la technique de fusion microcellulaire :

- la première étape consiste à soumettre les cellules donneuses à un arrêt mitotique prolongé par exposition continue à un inhibiteur de la formation du fuseau mitotique comme la colcémide.
- ce traitement entraîne le fractionnement du contenu cellulaire chromosomique en paquet subnucléaire distinct formant des micronoyaux ou des micronuclei.
- les micronoyaux peuvent être isolés physiquement des cellules par centrifugation en présence de cytochalasine B (un inhibiteur du fuseau mitotique), entraînant la formation de microcellules, particules constituées d'un seul micronoyaux et d'une fine bande de cytoplasme entourée par une membrane cytoplasmique intacte.
- comme pour les cellules donneuses normales, les microcellules peuvent être fusionnées avec les cellules receveuses. Les hybrides obtenus sont appelés hybrides microcellulaires.
- Certains contiennent un petit nombre de chromosomes donneurs, mais les plus simples contiennent un seul chromosome donneur, il s'agit d'hybride monochromosomique.

Le gros avantage des hybrides monochromosomiques est qu'ils peuvent mettre en évidence, de façon non ambiguë, la présence ou l'absence d'une séquence définie d'ADN humain sur un chromosome spécifique.

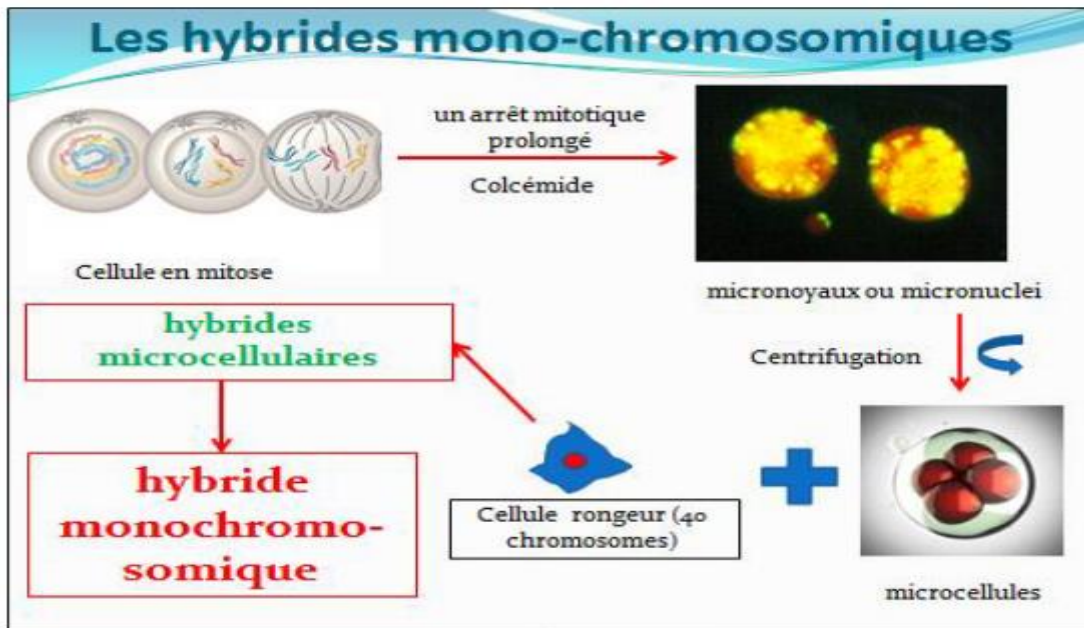
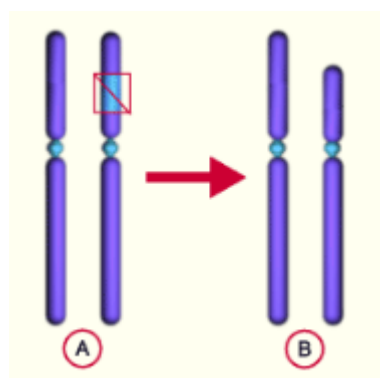


Figure : technique des hybrides mono-chromosomiques

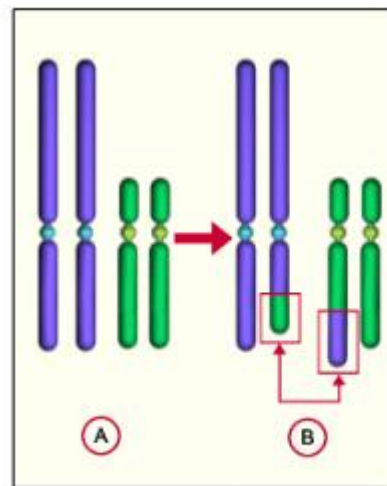
1-3-Les hybrides de translocation et de délétion

Pour obtenir une localisation sub-chromosomique, il est nécessaire d'utiliser des hybrides spécialisés contenant seulement une partie des chromosomes humains. Les fragments sub-chromosomiques peuvent être obtenus par la cassure spontanée des chromosomes résultant de translocation ou de délétion ou être induite artificiellement.

- **La délétion** est une anomalie de structure des chromosomes caractérisée par la perte de matériel génétique sur un chromosome. La taille des délétions varie d'une paire de base à toute une région chromosomique et les délétions peuvent survenir n'importe où sur le chromosome. Il s'agit toujours d'une anomalie déséquilibrée. La délétion est dite interstitielle quand il y a perte d'un fragment intermédiaire (deux points de cassure comme dans l'inversion), terminale quand l'extrémité d'un bras chromosomique est concernée (un seul point de cassure).



- Une **translocation réciproque** est un échange de fragments chromosomiques entre 2 chromosomes non homologues après cassure sur chacun des chromosomes impliqués. Si cet échange s'accompagne d'une perte de matériel génétique, il est déséquilibré, sinon la translocation est dite équilibrée, car la totalité de l'information génétique est présente. Les problèmes surgissent lors de la formation des gamètes.



On peut réaliser des hybrides cellulaires somatiques en fusionnant des cellules de rongeur à des cellules humaines contenant des translocations chromosomiques ou des chromosomes possédant une délétion visible cytogénétiquement. Les cellules hybrides résultantes peuvent être testées pour la présence du chromosome anormal. Des lignées cellulaires peuvent être ensuite réalisées à partir d'hybrides contenant le chromosome anormal, mais ne contenant pas l'homologue normal du chromosome d'intérêt.

Bien que l'approche d'hybride de délétion et de translocation améliore de manière importante le niveau de résolution de la localisation des gènes, les régions chromosomiques définies sont encore très larges en comparaison de la taille moyenne d'un gène.

1-4-Cartographie par hybrides d'irradiation

C'est une technique qui commence généralement avec un hybride de cellules somatiques monochromosomique. Les cellules sont irradiées avec les rayons X ou les rayons γ afin d'induire de multiples cassures doubles brins dans les chromosomes humains. Certaines cellules, dénommées hybrides d'irradiation, contiennent de petits fragments de chromosomes humains fusionnés à des chromosomes de rongeur.

La présence de fragments de chromosomes humains peut être détectée par la recherche des séquences Alu, ce sont des séquences répétées, réparties régulièrement dans les chromosomes humains, espacées de quelques kilobases, mais qui ne sont pas retrouvées dans les chromosomes de rongeur.

Dans la mesure où les radiations cassent les chromosomes humains selon des intervalles aléatoires, les loci qui sont situés à proximité les uns des autres se retrouveront plus fréquemment sur le même fragment chromosomique. Des techniques comme la PCR peuvent alors être utilisées pour déterminer quels hybrides contiennent des combinaisons spécifiques de locus humains.

Des amorces spécifiques de chaque locus sont utilisées pour déterminer si les segments d'ADN de chacun d'eux peuvent être amplifiés, indiquant leur présence dans l'hybride. Ensuite, des méthodes statistiques sont appliquées pour évaluer à quelle fréquence chaque paire de loci est retrouvée sur le même fragment chromosomique, fournissant ainsi une estimation des distances relatives entre les loci.

Il faut habituellement environ 100 à 200 hybrides pour construire une carte chromosomique humaine. Ce type d'approche n'est pas réalisable en pratique pour cartographier tout le génome du fait du nombre énorme de cellules hybrides qui seraient nécessaires.

1-5-Les anomalies chromosomiques de structure

Des anomalies de type délétion et translocation peuvent n'avoir aucune conséquence clinique par elles-mêmes, servant ainsi de marqueurs ; ou peuvent être responsables de maladies et ainsi peuvent être utilisées en cartographie:

- lorsque le réarrangement est non équilibré, il y a perte ou gain de matériel
- lorsque l'un des points de cassures chromosomiques est à l'origine d'une maladie.

Les délétions de matériel génétique constituent la cause directe de la maladie génétique. Les translocations chromosomiques équilibrées n'ont souvent aucun effet sur le porteur de la translocation, dans la mesure où l'individu garde une copie complète de son matériel génétique. Cependant lorsqu'une translocation a pour conséquence l'interruption d'un gène, elle peut être à l'origine d'une maladie génétique.

1-6-Cartographie par dosage génique

C'est une approche qui tire aussi profit de chromosomes dont la structure est anormale pour localiser des gènes, mais ne repose pas sur la nécessité de voir les chromosomes anormaux ségréger dans des cellules somatiques hybrides. Cette approche repose sur la détection d'un effet de dose soit des produits géniques soit des séquences géniques elles-mêmes observées dans les lignées cellulaires provenant de patients et contenant un nombre anormal de copies d'un gène particulier. Bien que cette méthode nécessite une analyse et une interprétation très rigoureuses, elle a été utilisée pour localiser ou exclure des gènes au niveau d'une région impliquée dans une duplication ou une délétion.

1-7-Cartographie des gènes par hybridation in situ

Les méthodes de cartographie exposées jusqu'ici sont indirectes, en ce sens qu'elles fournissent une information sur la localisation physique d'un gène sur un chromosome particulier, sans préciser la position du gène sur la carte.

Une approche plus directe est celle de l'hybridation in situ qui implique l'hybridation de sonde de l'ADN sur des préparations de chromosome en métaphase sur une lame dont l'ADN a été dénaturé sur place (in situ). La sonde s'hybride, par appariement des bases complémentaires, avec l'ADN dénaturé d'un segment de chromosome spécifique. Le marqueur indique sur le chromosome la position de la sonde hybridée. Les marqueurs radioactifs sont visualisés en plaçant un film radiographique sur la lame (autoradiographie), les marqueurs fluorescents sont visualisés avec un microscope à fluorescence.

Cette méthode était à l'origine très laborieuse, mais très lente, nécessitant un très long temps d'exposition des lames à l'émulsion photographique pour détecter la localisation des sondes hybridées marquées à l'aide d'isotopes de faible intensité tel que le tritium. La localisation nécessitait l'analyse d'un grand nombre de préparations en métaphase afin de distinguer le signal d'hybridation du bruit de fond de la radioactivité.

1-8-L'hybridation in situ en fluorescence standard

La FISH comprend les étapes suivantes :

- marquage de sonde
- dénaturation
- préhybridation (hybridation compétitive)

- hybridation sélective
- révélation
- observation au microscope épifluorescent

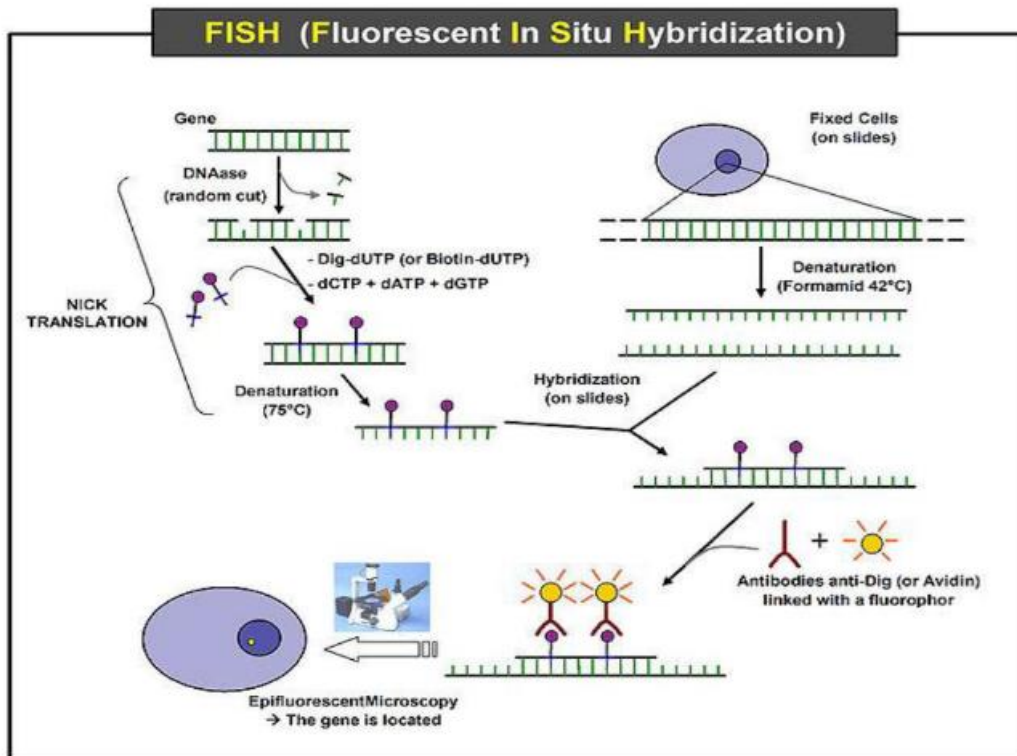


Figure : hybridation in situ en fluorescence standard

2-Cartographie physique à haute résolution

Les techniques décrites ci-dessus peuvent être complétées par des techniques de cartographie moléculaire qui assurent une résolution de 1 pb à plusieurs mégabases.

2-1- La FISH à haute résolution

La longueur linéaire de l'ADN d'un chromosome humain de taille moyenne est environ 5 cm, mais il est compacté par différents niveaux de repliement dans les chromosomes métaphasiques. Pour obtenir une résolution supérieure, des sondes d'ADN sont hybridées à des chromosomes interphasiques naturellement étalés, à de la chromatine artificiellement étalée ou à des fibres d'ADN préparées par différentes méthodes.

2-1-1-La FISH sur chromosome prophasique

Les chromosomes des noyaux prophasiques sont beaucoup plus étalés que les chromosomes métaphasiques ou pro-métaphasiques. Cette technique permet donc une résolution de cartographie supérieure et peut aider à déterminer l'ordre physique de certains clones d'ADN synténique. Les chromosomes étalés pouvant se replier en boucles dans certaines régions, une détermination précise de l'ordre linéaire de clones d'ADN synténiques impose une analyse statistique des résultats de la cartographie à partir de nombreux noyaux pro-phasiques différents. Même ainsi, la méthode reste surtout adaptée pour ordonner des séquences séparées par des intervalles compris entre 50 et 500 Kb.

2-1-2-La FISH sur chromatine artificiellement étalée

La FISH à haute résolution peut également être réalisée en utilisant des techniques qui permettent d'étaler l'ADN des chromosomes sur une lame de microscope avant l'hybridation. Une de ces méthodes est appelée DIRVISH (direct visual hybridization) : les cellules sont lysées avec un détergent à une extrémité de la lame ; la lame est ensuite penchée et l'ADN en solution glisse vers le bas. De telles préparations permettent des résolutions de cartographie extrêmement élevées, de 700 Kb jusqu'à 5 Kb.

2-1-3-La FISH sur fibre d'ADN

Plus récemment, le principe de l'étalement artificiel des fibres de chromatine a été étendu à l'ADN sans protéine, l'ADN cible est préparé à partir de cellules incluses dans des blocs de gel d'électrophorèse en champs pulsé. Cette méthode utilise des fibres d'ADN linéarisées non fixées sur une lame de microscope, et sa résolution est comprise entre 500 Kb et quelques kb.

2-2- La cartographie de restriction

La cartographie de restriction est l'utilisation des enzymes de restriction pour analyser et produire une carte physique de restriction des génomes, des gènes, ou d'autres segments d'ADN. Une carte de restriction est une suite linéaire de sites séparés par une distance réelle (en nombre de paires de bases). Ces sites correspondent à des coupures de la séquence par des enzymes de restriction. Les tailles des fragments qui résultent de ces coupures sont ensuite mesurées.

La cartographie de restriction a pour objectif de reconstituer la séquence d'ADN en réordonnant les fragments issus des digestions. La cartographie de restriction à grande échelle nécessite des enzymes qui coupent rarement l'ADN et le fractionnement des grands fragments de restriction par électrophorèse sur gel en champ pulsé.

La plupart des enzymes de restriction qui reconnaissent une séquence de 4 à 6 pb coupent habituellement l'ADN. Elles engendrent donc des fragments dont la taille est habituellement de quelques centaines de Kb. Les petits fragments de restriction peuvent être fractionnés selon la taille sur un simple gel d'électrophorèse en agarose. Cependant la capacité des gels d'agarose à fractionner de grands fragments d'ADN est très limitée.

L'ADN génomique préparé de façon conventionnelle ne se prête pas à la cartographie de restriction à grande échelle, car les techniques de lyse cellulaire et de purification de l'ADN utilisées entraînent une fragmentation considérable de cet ADN.

L'ADN doit être isolé de façon à minimiser la cassure artificielle des grandes molécules avant la digestion avec les enzymes de restriction. Pour préparer de l'ADN de haut poids moléculaire, les échantillons de cellules, par exemple les leucocytes, sont mélangés avec de l'agarose fondu, puis transférés dans les puits d'un bloc calibré et refroidi. Les cellules sont ainsi emprisonnées dans les blocs d'agarose solide. Les blocs sont incubés avec des enzymes hydrolytiques qui diffusent à travers les pores de l'agarose et digèrent les composants cellulaires, mais laissent l'ADN chromosomique de haut poids moléculaire pratiquement intact. Ces blocs peuvent ensuite être incubés dans un tampon contenant l'enzyme de restriction coupant rarement.

Afin de séparer les grands fragments de restriction, on utilise l'électrophorèse sur gel en champ pulsé PFGE (pulsed field gel electrophoresis). Au cours de la PFGE, l'orientation relative du gel et du champ électrique est modifiée périodiquement, en délivrant de brèves impulsions électriques qui activent alternativement deux champs orientés différemment. Le principe du champ électrique discontinu se fait de sorte que les molécules d'ADN sont forcées de changer leur conformation et leur direction de migration périodiquement au cours de leur passage à travers le gel. Ce changement dépend strictement de la taille, des fragments d'ADN allant jusqu'à plusieurs mégabases peuvent ainsi être fractionnés efficacement.

Logiciels de cartographie

Végétaux : MapMaker, CarthaGene, JoinMap

Animaux : CRIMAP,

Homme : MapMaker

Hybrides irradiés : RHMAP, RHO, CarthaGene

Voir <http://linkage.rockefeller.edu/soft/>