

### III. Anomalies oculaires d'origine héréditaire chez le beagle

Le beagle est une race originaire d'Angleterre et très ancienne. Les premiers beagles ont été introduits en France vers 1860. Le beagle est un chien courant mais il est très polyvalent, possède un très bon odorat, est de taille plutôt petite (33 cm à 40 cm de garrot), avec un bon tempérament. Le beagle est un chien à poils ras avec une robe tricolore (blanc, noir, fauve) ou bicolore (blanc et fauve vif ou citron et feu). Il est utilisé aussi bien comme chien de compagnie que comme chien de chasse, chien de détection ou chien de laboratoire. (22)

Les affections oculaires rencontrées dans la race beagle sont multiples. (Tableau 1)

Tableau 1 : Affections oculaires héréditaires ou à prédisposition raciale chez le beagle (10)

Affection	Mode de transmission	Commentaires
Microphtalmie	Autosomique dominant (avec anomalies oculaires multiples) ou inconnu  Autosomique récessif (unilatérale)	Elle peut être bilatérale ou unilatérale, accompagnée d'anomalies oculaires complexes (cataracte, dysplasie vitréenne multifocale)
Ectropion	Autosomique dominant ( ?)	Très rare
Distichiasis	Inconnu	De fréquence variable, il est rencontré plutôt à la paupière supérieure
Tumeurs palpébrales	Prédisposition raciale	Pas de base génétique
Luxation de la glande nictitante	Inconnu	Chez des chiens de moins d'un an (se développe entre 6 mois et 1 an) et elle serait héritable

Dystrophie cornéenne stromale lipidique	Inconnu	Elle est très fréquente (17% des chiens âgés de 5 à 15 ans, dont 75% atteints de façon bilatérale). L'opacité est grisâtre à blanchâtre localisée selon les formes entre le tiers antérieur et le tiers profond du stroma ; elle est typiquement ovale, horizontale, centrale. Il existe trois types en fonction de leurs aspects : type néphéliion, type annulaire et type arqué. L'anomalie est en principe non invalidante et lentement évolutive
Kérato-conjonctivite sèche (KCS)	Inconnu	Chez des chiennes entre 6 et 8 ans avec une affection bilatérale et progressive
Persistance de la membrane pupillaire	Inconnu (peut être autosomique dominant)	Rare, uniquement des cas mineurs (fibres pupillaires en surélévation de la collerette, fibres irido-iriennes)
Kystes iriens	Inconnu	Fréquence pas très élevée, présente chez de chien d'au moins 8 ans

Glaucome primaire	Autosomique récessif	<p>Le glaucome à angle ouvert (rare chez le chien) est caractérisé par une élévation de la pression intra-oculaire chez des chiens âgés de 4 mois à 9 ans, avec une excavation pupillaire modérée et une désinsertion zonulaire entre 13 et 30 mois par fermeture progressive de la fente ciliaire au départ de structure normale. Les chiens glaucomateux (accès de glaucome aigu vers l'âge de 5 ans, buphtalmie et luxation du cristallin chez quelques chiens âgés de 7 ans ou plus) présentés pour traitement ne le sont pas avant 5 ans. Le gène myociline (MYO) est impliqué dans plus de 6% des glaucomes humains à angle ouvert, d'apparition précoce ou tardive, respectivement autosomiques dominant ou récessif. La myociline est une protéine présente sous forme glycosylée et non glycosylée dans l'humeur aqueuse ; elle est présente dans l'humeur aqueuse du beagle normal, sa concentration est environ multipliée par 4 chez les beagles porteurs et par 8 à 10 chez les beagles glaucomateux proportionnellement à la gravité de l'atteinte. Selon le laboratoire Optigen (<a href="http://www.optigen.com">www.optigen.com</a>), environ 1 % de la population de beagles serait porteuse de la mutation du gène ADAMTS10 (<i>A Disintegrin-like and A Metalloprotease with Thrombospondin Type 1 motif 10</i>). Cette mutation est considérée comme responsable du glaucome à angle ouvert chez le beagle</p>
<p>Cataracte :</p> <ul style="list-style-type: none"> <li>- Corticale postérieure</li> <li>- Autres</li> <li>- Congénitale (voir microphthalmie)</li> </ul>	Inconnu	<p>La cataracte sous capsulaire polaire postérieure triangulaire en principe bilatérale est décrite chez l'adulte, elle est rarement progressive. D'autres types de cataractes (capsulaire postérieure et antérieure, équatoriale, etc.) ont été sporadiquement observées</p>
Luxation du cristallin	Voir glaucome primaire	Elle est la conséquence du glaucome à angle ouvert

Dégénérescence des photorécepteurs (APR)	Autosomique récessif présumé	Les cas observés sont présentés pour cécité nocturne (éventuellement nocturne et diurne) entre 3 et 5 ans pour la plupart et de façon exceptionnelle entre 10 et 13 ans. L'hyperréflexion du tapis progresse et se généralise à partir d'une zone juxta-pupillaire en bande, horizontale, et de la périphérie du tapis. Les lésions focales claires de la zone sans tapis (nombreuses et fréquentes) sont ovales ou polyédriques
Dysplasie de la rétine	Autosomique récessif (?)	Il s'agit de plis rétiniens soit de la zone du tapis, soit de la zone sans tapis, linéaires ou ramifiés le plus souvent, parfois vermiformes. Ces plis persistent avec l'âge mais la fréquence (3% chez des chiens âgés de 3 à 6 mois) diminuerait avec l'âge et la localisation préférentielle est le quadrant inféro-nasal.
Hypoplasie et absence de tapis	Autosomique récessif (?)	Chez des chiens à iris clair (jaune ou fauve), le tapis est fréquemment absent (zone brun-rouge). La vascularisation choroïdienne est visible dans la zone sans tapis (dépigmentée). L'anomalie évolue dans le temps, la vascularisation choroïdienne devenant partout visible. La fréquence de l'anomalie serait de 1 cas pour 1500. Dans certains cas, le tapis est absent ou présent sous forme de simples îlots périphériques de petite taille. L'anomalie a été reconnue comme un défaut de développement d'origine autosomique récessive, sans incidence sur la vision. Le tapis est histologiquement visible mais ses éléments constitutifs sont irréguliers, remplis d'inclusions intracellulaires anormales visibles en microscopie électronique à transmission, carencés en zinc et en cystéine
Colobome du nerf optique	Inconnu	Rare, de dimension et de localisation variables
Hypoplasie du nerf optique	Inconnu	Rare (1 cas sur 1400). Il est associé au colobome. Les chiens affectés de façon bilatérale sont aveugles

Micropapille	Inconnu	Elle est liée à un défaut de myélinisation des fibres en arrière de la lame criblée et occasionnellement rencontrée dans la race
--------------	---------	--

## **IV. Recherche génétique appliquée à la couleur des yeux bleus chez le beagle**

Les gènes potentiellement responsables de la couleur des yeux ne sont pas connus chez le chien. Pour étudier un gène, il existe de nombreuses techniques dont l'approche « gène candidat » qui sera celle utilisée pour cette étude.

### **IV.1. Techniques génétiques utilisables pour l'identification de gènes d'intérêt**

Pour la recherche d'un gène potentiellement responsable d'un caractère phénotypique ou d'une maladie génétique, le généticien dispose d'outils et de cartes permettant de se repérer dans le génome, qui seront présentés ci-dessous.

#### **IV.1.1. Outils génétiques**

Pour notre étude, les principaux outils génétiques utilisés ont été des marqueurs génétiques de type SNP (*Single Nucleotide Polymorphism*) et la technique de PCR (*Polymerase Chain Reaction*). Nous allons détailler leur utilisation.

##### ***IV.1.1.1. Les Single Nucleotide Polymorphisms (SNPs)***

Un SNP est une mutation ponctuelle isolée, c'est-à-dire qu'une seule base est différente, et en un seul locus. Il s'agit donc d'un polymorphisme nucléotidique. Le SNP correspond à une variation stable de la séquence génomique portant sur une base, toutes les 100 à 300 bases environ et affectant au moins 10% de la population chez l'homme. Les SNPs sont une nouvelle génération de marqueurs polymorphes qui n'ont en général pas d'implications fonctionnelles, mais qui définissent un locus unique dans le génome. Les SNPs situés dans les régions codantes et dans les régions régulatrices des gènes sont utilisés pour effectuer la cartographie des maladies multifactorielles qui consiste en l'étude d'association des gènes candidats impliqués dans ces maladies. Actuellement, ce sont les outils privilégiés dans l'exploration du génome canin. Chez le chien, il existe des « puces » à SNPs qui permettent de génotyper un animal pour plusieurs milliers de SNP, en une seule fois. On dispose ainsi d'une couverture régulière et dense du génome du chien, qui permet d'identifier des locus impliqués dans des maladies héréditaires ou des caractères d'intérêt. (3) (30) (42) (49)

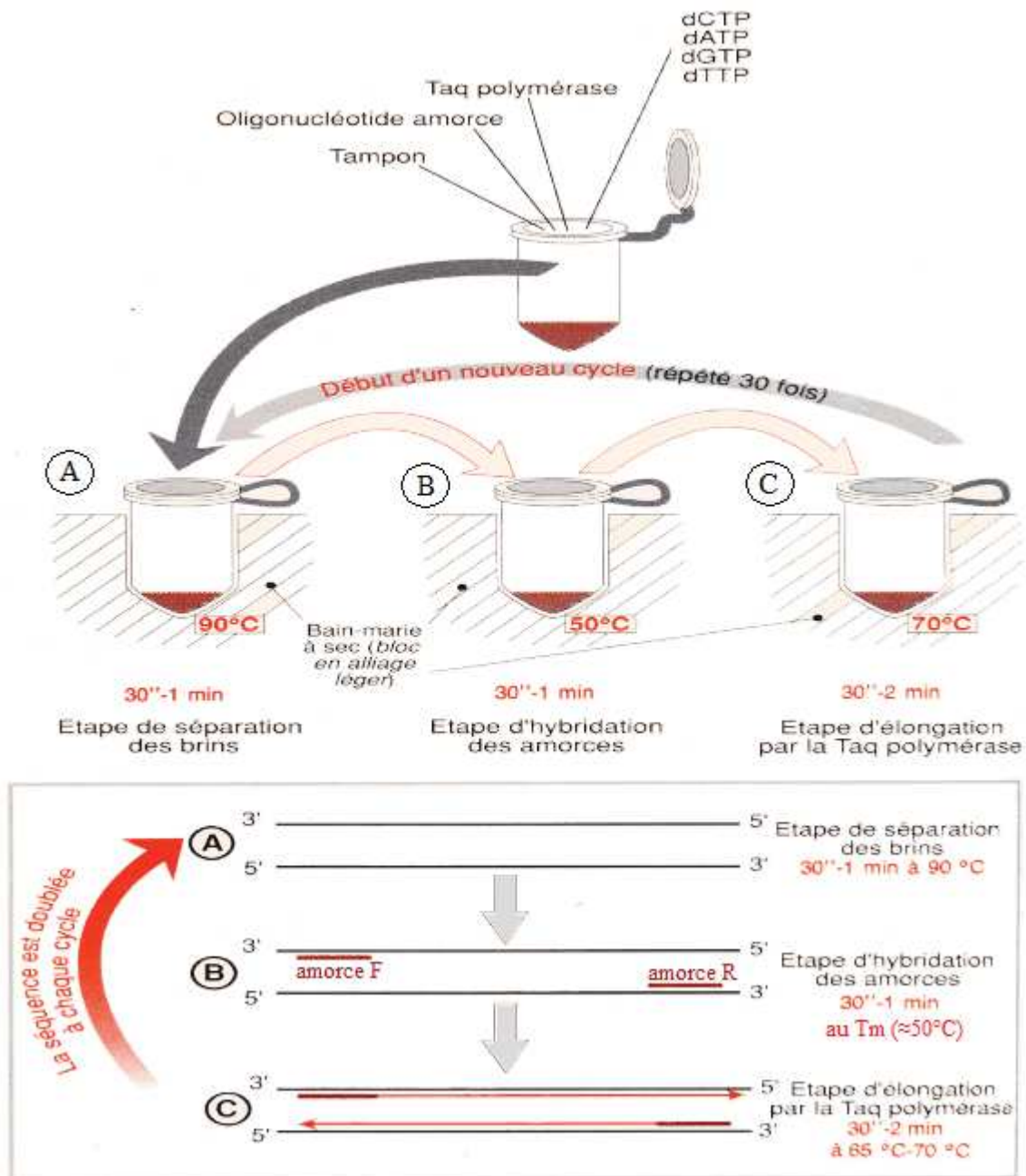
##### ***IV.1.1.2. La Polymerase Chain Reaction (PCR)***

La technique de PCR a été découverte en 1983 et mise au point en 1985 par K. Mullis. C'est une technique d'amplification *in vitro* qui utilise des ADN polymérases (appelées Taq polymérases : ce sont des enzymes thermostables) dont la particularité est de ne pouvoir synthétiser le brin complémentaire d'une séquence qu'à partir d'une amorce. L'objectif de la PCR est de multiplier de façon exponentielle une séquence d'ADN choisie. La technique de PCR consiste à enchaîner plusieurs cycles décrits ci-dessous. Le nombre de copies d'ADN double à chaque cycle. (27)

Chaque cycle est composé de trois étapes décrites sur la *Figure 16* qui sont :

- **l'étape de séparation des brins d'ADN** (appelée phase A sur la *Figure 16*). Lors de cette étape, la température du bain-marie à sec augmente jusqu'à une température supérieure à 90°C (le plus souvent 94°C). C'est l'augmentation de la température qui est à l'origine de la séparation des brins d'ADN. On parle aussi de dénaturation de l'ADN ; (27)
- **l'étape d'hybridation** (appelée phase B sur la *Figure 16*) qui correspond à l'hybridation des amorces *Forward* (F) et *Reverse* (R). Une amorce est une molécule d'ADN simple brin de 20 nucléotides environ qui encadre le morceau d'ADN que l'on souhaite amplifier. En effet, les amorces ont un double rôle : celui de délimiter la séquence d'ADN choisie et de servir d'amorçage pour l'ADN polymérase avec leur extrémité 3' OH libre. Pour cela, il est nécessaire de connaître les séquences nucléotidiques des extrémités de la région ADN amplifiée. La composition des amorces utilisées détermine la température d'hybridation des amorces ( $T_m$ ). En général, elle est comprise entre 50 et 70°C. Le choix des amorces est crucial dans la mise au point de la réaction PCR, pour cela il existe des logiciels d'analyse des séquences qui choisissent des amorces ayant des conditions particulières (des séquences qui ne soient pas complémentaires entre elles, des séquences qui ne contiennent pas de séquences répétées inversées, des  $T_m$  proches) ; (27)
- **l'étape d'élongation du brin complémentaire par la Taq polymérase** qui correspond à la phase C sur la *Figure 16*. Elle se déroule en général à 72°C et permet d'obtenir deux molécules d'ADN double brin ; (27)
- ces molécules d'ADN nouvellement formées vont subir un nouveau cycle. La PCR est composée de 30 à 35 cycles environ. Il n'est pas utile d'augmenter davantage le nombre de cycles car au-delà d'une trentaine de cycles, le nombre de copies n'augmente plus. (27)

Figure 16 : Principe de réalisation pratique d'une PCR (d'après (27))



La PCR permet la multiplication exponentielle d'une séquence d'ADN déterminée par les amorces. En phase A, l'augmentation de la température permet la séparation des brins d'ADN. Lors de la phase B, les amorces F et R s'hybrident. Au cours de la phase C a lieu l'élongation par la Taq polymérase, qui permet d'obtenir deux molécules d'ADN double brin qui vont subir un nouveau cycle. Une trentaine de cycles environ s'enchaînent.

Pour effectuer ces trois étapes qui nécessitent des transitions de températures, les microtubes contenant le mélange réactionnel sont placés dans un appareil programmable : un thermocycleur. Cet appareil permet d'exposer les tubes à des températures choisies pendant



des durées déterminées par l'expérimentateur. La réaction PCR est extrêmement rapide et ne dure que quelques heures (1 à 2 heures pour une PCR de 30 cycles).

A la fin de la PCR, pour vérifier qu'elle a bien fonctionné, le produit est visualisé en utilisant généralement une électrophorèse en gel d'agarose. Les fragments d'ADN sont séparés en fonction de leur taille par l'action d'un champ électrique. Un agent intercalant est utilisé pour visualiser les différents fragments d'ADN sous lumière ultraviolette. (27)

La PCR est utilisée pour toutes sortes d'applications : amplification d'un segment d'ADN pour divers usages (clonage cellulaire, séquençage), détection des délétions d'insertions, détection du polymorphisme des minisatellites et microsatellites, génération de sondes (radioactives, fluorescentes),... Elle est très souvent utilisée en génétique malgré l'existence de plusieurs limites : la taille de la séquence à amplifier doit être inférieure à 3kb (kilobases) avec une Taq conventionnelle, l'inefficacité de l'augmentation du nombre de cycles au-delà d'une trentaine, la contamination de l'ADN lors de manipulations (pour diminuer l'impact de ces contaminations, il est important de manipuler dans de bonnes conditions et de limiter au maximum les possibilités de transfert d'ADN en suspension dans l'air ou transmis par le manipulateur), le manque de fidélité de la Taq polymérase et des amplifications parasites (mauvaise hybridation des amorces qui sont de courte taille). (27)

Pour notre étude la PCR fut utilisée pour amplifier un segment d'ADN en vue de le séquencer.

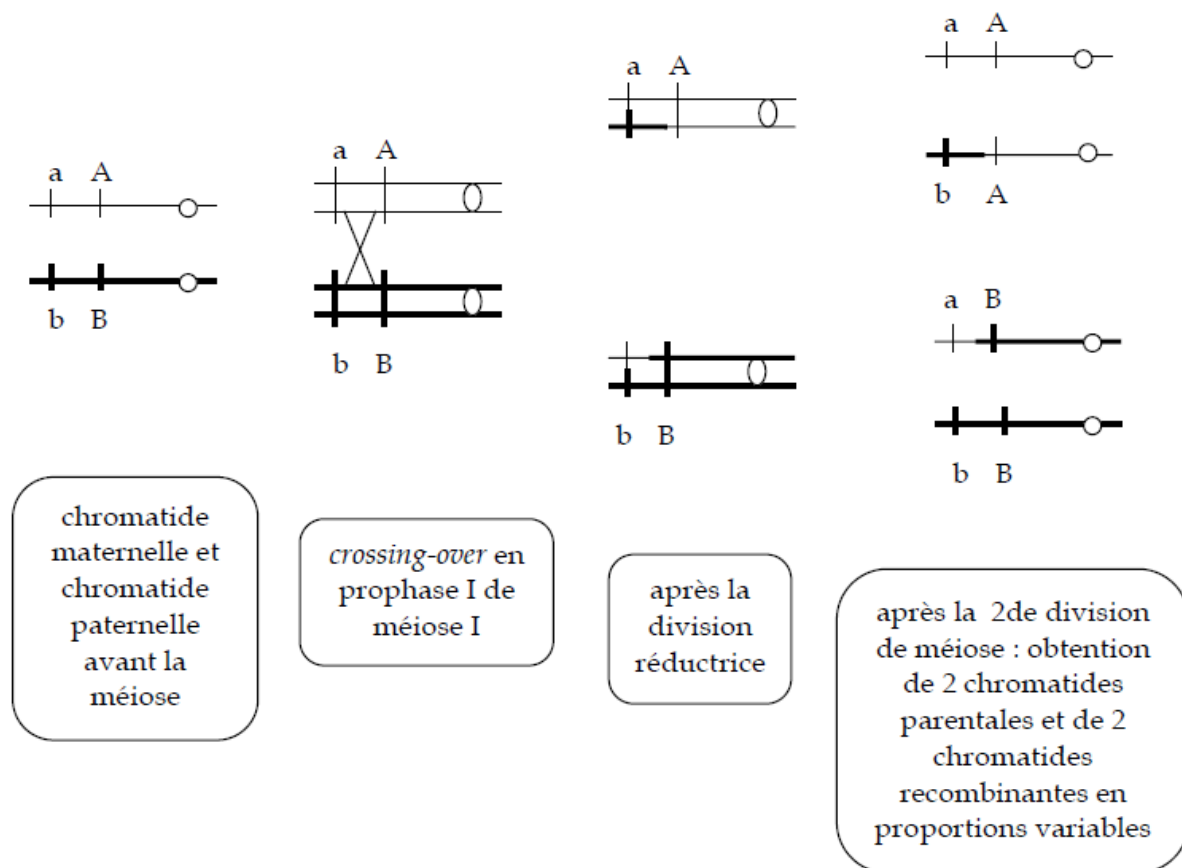
## **IV.1.2. Cartographie**

La cartographie est la localisation des gènes et marqueurs génétiques ou segments d'ADN sur les chromosomes. Elle permet de se repérer dans le génome et de pouvoir étudier les gènes s'y trouvant. Il existe de nombreuses manières de faire des cartes du génome, dont les cartes génétiques, les cartes physiques et le séquençage, décrites ci-dessous.

### ***IV.1.2.1. Cartes génétiques***

La réalisation d'une carte génétique consiste à localiser des marqueurs génétiques (SNP, microsatellites) d'une espèce les uns par rapport aux autres. Pour cela, on analyse statistiquement la capacité des marqueurs à se réassocier à la suite du brassage des chromosomes (*crossing-over*) qui se produit pendant la méiose. (*Figure 17*). Lorsque deux marqueurs ségrégent ensemble, ils sont liés et lorsque deux marqueurs ségrégent séparément, ils sont indépendants (*Figure 18*). La cartographie génétique est fondée sur le principe que deux marqueurs situés sur le même chromosome maintiennent d'autant plus fréquemment leur liaison après la méiose que la distance physique les séparant est faible. Ainsi la distance génétique sur une carte de liaison est déterminée par la fréquence des *crossing-over*. Plus les locus sont proches, plus les chances de recombinaisons méiotiques sont faibles ; et inversement, plus les locus sont éloignés et plus les chances de *crossing-over* sont élevées. Plus deux locus sont proches, moins il y a de chance qu'un *crossing-over* se produise et plus le nombre de gamètes recombinants est faible. (27)

Figure 17 : Crossing-over lors de la méiose (27)



*a/b* : allèles du gène 1

*A/B* : allèles du gène 2

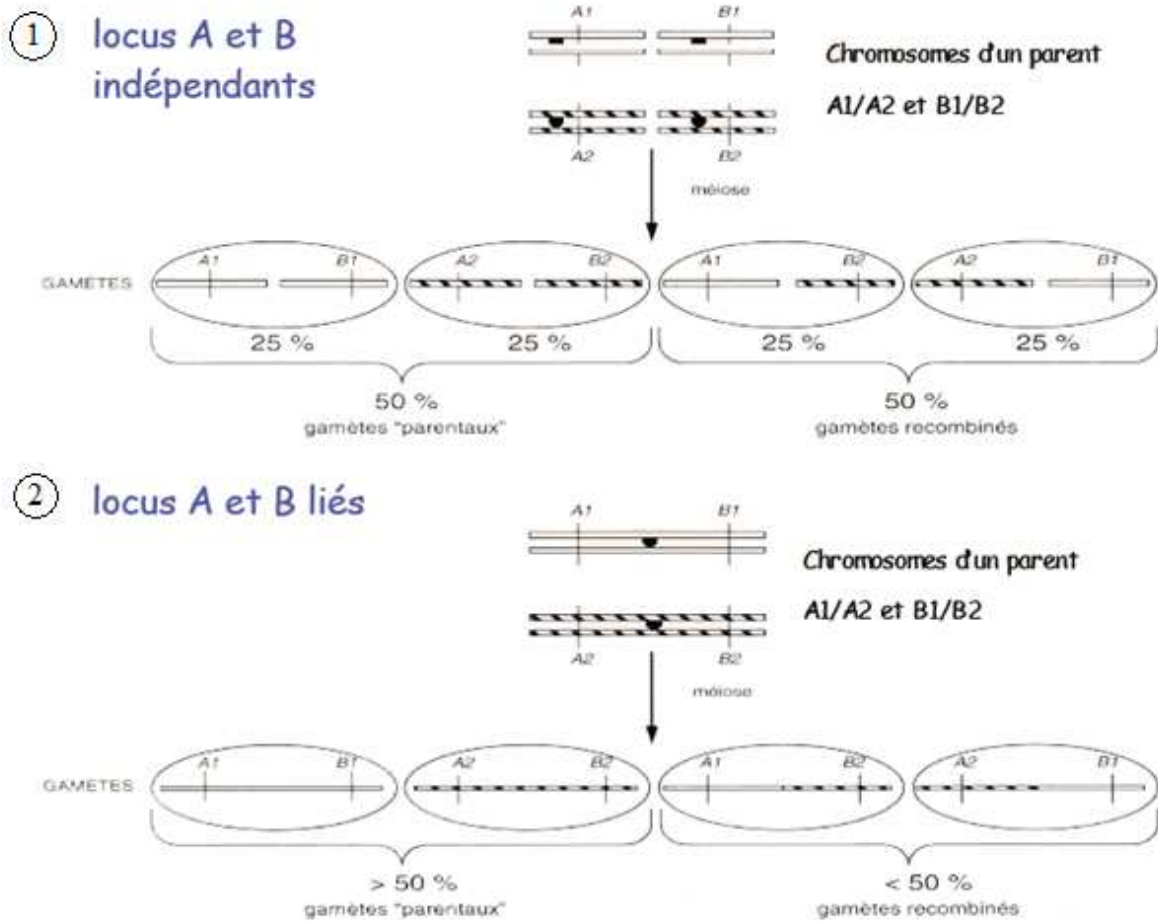
*a* et *A* appartiennent au chromosome maternel.

*b* et *B* appartiennent au chromosome paternel.

Sur cette figure, un seul chromosome est représenté. Son centromère est représenté par un cercle.

Lors de la première division de méiose, les chromosomes homologues s'apparient. À cette occasion, se produisent les crossing-over. Un échange de segments de chromatide a alors lieu. À la fin de la seconde division de méiose, on obtient 4 chromatides que l'on peut classer en 2 catégories : deux chromatides parentales, l'une paternelle portant *b* et *B*, l'autre maternelle portant *a* et *A*, et deux chromatides recombinantes portant les allèles *b* et *A* pour l'une, *a* et *B* pour la seconde.

Figure 18 : Liaison génétique et recombinaison (d'après(23) )



Soit deux locus A et B. Le locus A possède deux allèles A1 et A2. Le locus B possède deux allèles B1 et B2. Dans la partie 1, le locus A et B sont sur des chromosomes différents et sont donc indépendants. La moitié des gamètes produits est de type parental et l'autre moitié des gamètes est de type recombiné. Dans la partie 2, les locus A et B sont sur le même chromosome et sont donc liés. Plus de 50 % des gamètes est de type parental et moins de 50 % est de type recombiné.

Lors de la méiose, les individus parents donnent naissance à des individus enfants parmi lesquels on trouve des recombinants (ou recombinés) et des parentaux. Cette nouvelle génération voit l'apparition de nouvelles combinaisons d'allèles présents sur une même chromatide. La proportion d'enfants recombinants définit la fraction (ou fréquence) de recombinaison entre les deux locus appelée  $\theta$ .

$$\theta = \frac{\text{nombre de recombinants}}{(\text{nombre de recombinants} + \text{nombre de parentaux})}$$

La distance génétique exprimée en centiMorgan (cM) se calcule à partir de  $\theta$ . Un cM équivaut à un segment de chromatide sur lequel la probabilité qu'un *crossing-over* se produise est de 1%. (23) (27)

La carte génétique est établie grâce à la position relative des gènes et des marqueurs sur les chromosomes, mais elle ne permet pas de connaître la distance physique exacte séparant deux locus. En effet, malgré une fréquence de recombinaison nulle ou très proche de zéro, il est possible que plusieurs milliers de bases séparent deux locus. (23) (27)

Chez le chien, des cartes génétiques ont été établies, qui ont permis d'identifier des gènes responsables de caractères délétères et d'intérêt. (38)

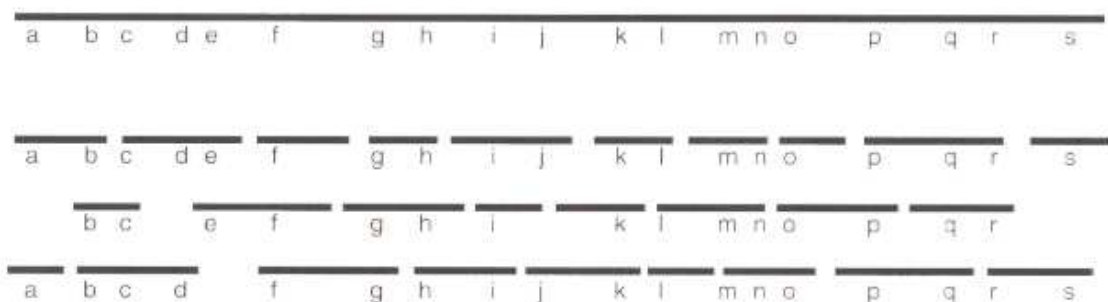
#### IV.1.2.2. Carte physique

La carte physique d'une espèce représente les chromosomes de cette espèce reconstitués artificiellement à l'aide d'une succession de clones d'ADN (morceaux d'ADN clonés c'est à dire multipliés artificiellement) et chevauchants. Les clones d'ADN sont obtenus à l'aide de l'ADN génomique de l'espèce considérée, fractionnés et inclus dans des vecteurs (de type plasmides, BAC : *Bacterial Artificial Chromosome*, ou YAC : *Yeast Artificial Chromosome*) que l'on utilise pour multiplier les clones d'ADN dans des systèmes bactériens ou en levures. Dans un second temps, il faut déterminer l'ordre des clones par l'identification des zones de chevauchement. Le positionnement des clones les uns par rapport aux autres fait appel à différentes techniques, dont le point commun est de rechercher les parties communes entre différents clones. (23) (27)

Parmi ces techniques, on peut citer :

- **l'hybridation des clones entre eux** : si deux clones peuvent s'hybrider, c'est qu'ils possèdent des séquences communes ;
- **typage de clones par PCR** pour déterminer la présence des marqueurs STS (*Sequence Tagged Sites*). Les STS sont de courtes séquences représentées de façon unique dans le génome. Si deux clones sont partiellement chevauchants c'est qu'ils contiennent un même marqueur unique. Les clones sont testés en PCR avec un marqueur donné et on regarde s'il y a présence ou non du STS ; (Figure 19)

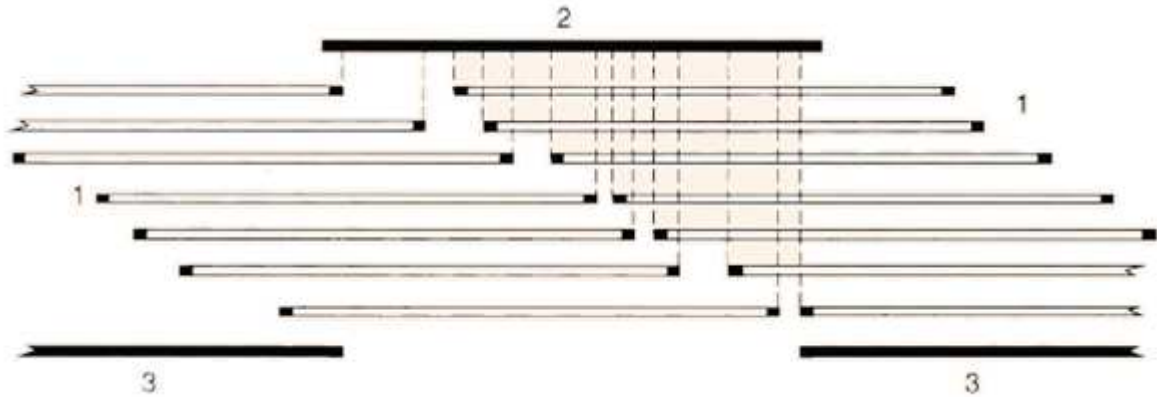
Figure 19 : Établissements d'une carte physique de fragment chevauchants à l'aide de STS (d'après (20) )



Les fragments clonés sont analysés pour la présence ou l'absence d'un ensemble de STS. Des clones ayant un ou plusieurs STS en commun sont chevauchants. On peut souvent déduire l'ordre des STS et donc celui des clones à partir du contenu en STS des clones.

- **séquençage des extrémités** : certains clones sont séquencés en entier et les autres n'ont que leurs extrémités de séquencées. Les clones séquencés partiellement sont repositionnés en fonction du clone totalement séquencé. (Figure 20) (20)

Figure 20 : Séquençage des extrémités (d'après (20) )



Les rectangles noirs représentent les segments d'ADN séquencés. Les rectangles blancs représentent les segments d'ADN non séquencés et qui sont ordonnés en fonction de leurs extrémités séquencées (carré noir).

Ainsi, les groupes de grands clones chevauchants sont positionnés le long des chromosomes ainsi reconstitués. Une fois tous les grands clones ordonnés et positionnés le long du génome, on dispose d'une carte physique de la région du génome étudiée, voire du génome complet étudié. (20) (23) (27)

#### IV.1.2.3. Les cartes d'hybrides d'irradiation

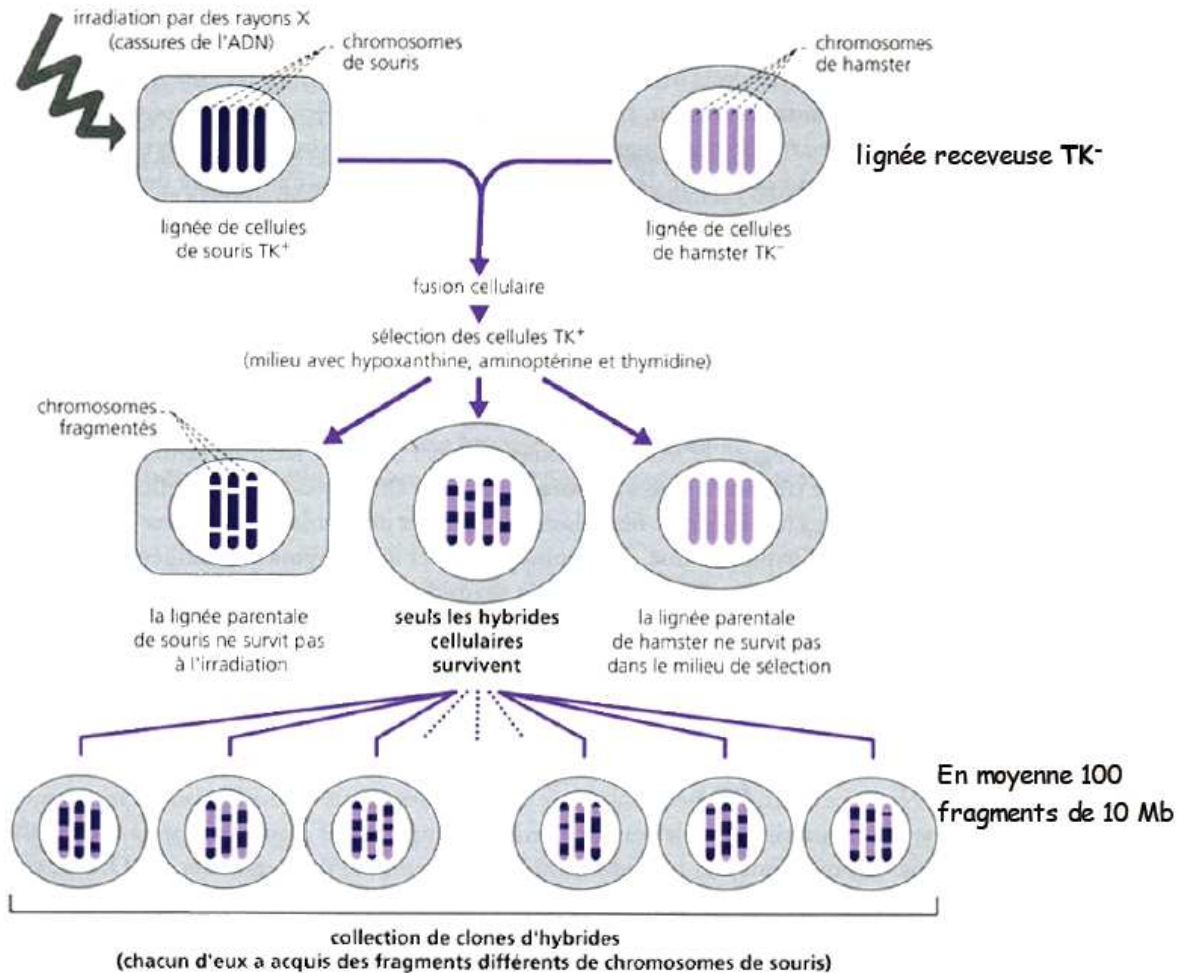
Les cartes d'hybrides d'irradiation reposent sur la fréquence d'apparition de cassures chromosomiques induites artificiellement par des rayons X, observées sur des cellules hybrides. Un hybride d'irradiation est une cellule fusionnée entre une cellule lourdement irradiées de l'animal que l'on veut cartographier et d'une cellule d'un autre animal (généralement de rongeur). Deux marqueurs initialement localisés à proximité l'un de l'autre ont plus de chance d'être simultanément retrouvés dans les mêmes hybrides que deux marqueurs éloignés. (24)

Les distances se mesurent en centiray (cR). Par définition, 1 cR<sub>(N rad)</sub> correspond à une fréquence de cassure de 1 % entre deux locus après exposition à une dose de N rad de rayons X. Une correspondance peut être établie avec les distances de cartes physiques et génétiques : 1cR pour 3000 rad équivaut environ à 200 kb, et 1cR pour 9000 rad équivaut environ à 55 kb. Ainsi, plus la dose d'irradiation (rayons X) appliquée est grande, plus la fragmentation sera élevée et plus la résolution de la carte sera fine. (24)

Les étapes de la cartographie d'hybrides d'irradiation sont :

- la **fabrication de panels d'hybrides d'irradiation** (Figure 21). Une collection de 100 à 200 hybrides d'irradiation est suffisante pour obtenir une carte au cR<sub>3000</sub> ; (24) (25) (43)

Figure 21 : Fabrication des hybrides d'irradiation (27)

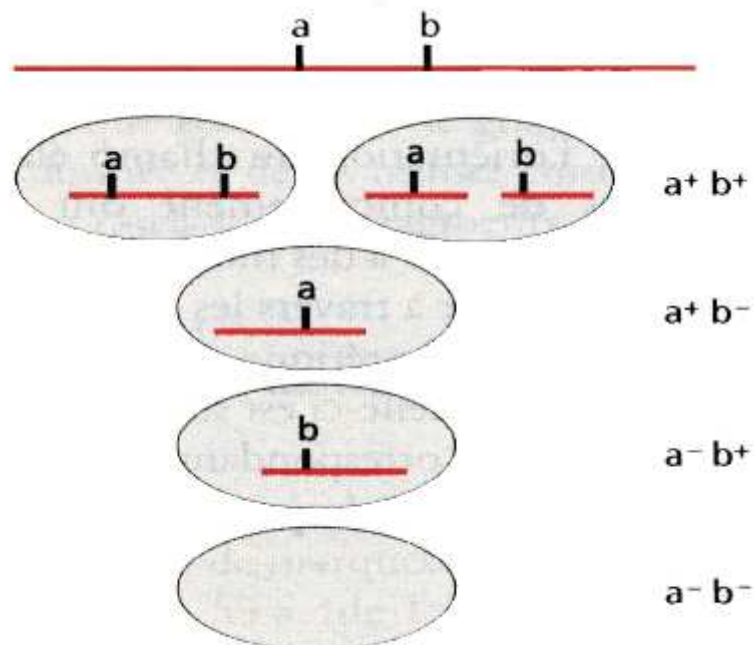


Lorsque la lignée des cellules de l'animal que l'on veut cartographier est irradiée, l'irradiation crée des cassures dans l'ADN des cellules. Ensuite, ces cellules sont fusionnées avec des cellules de hamsters déficients en thymidine kinase (TK). Puis elles sont mises dans un milieu de sélection où seuls les hybrides cellulaires (possédant le gène de la thymidine kinase) survivent. Ceci permet d'obtenir une collection d'hybrides où chacun d'entre eux possède un assortiment de fragments de chromosome de l'animal, inclus dans les chromosomes de hamster.

- chaque locus (marqueur, gène) peut être recherché en PCR à partir de l'ADN de chaque clone du panel d'hybrides. Ainsi, la fréquence de présence et d'absence de différents marqueurs moléculaires de l'animal de départ est calculée chez les hybrides. Les marqueurs n'ont pas besoin d'être polymorphes, seul la présence d'un signal positif en PCR (amplification du locus) ou négatif (pas d'amplification du locus) est analysée ; (24) (25) (43)

- l'analyse utilise une méthode statistique : la fréquence de la présence simultanée de deux locus est calculée et comparée aux fréquences de chaque locus présent individuellement. Les locus étroitement liés car physiquement proche au départ sur le chromosome de l'animal étudié sont incorporés ensemble à des fréquences élevées dans les hybrides car la probabilité qu'une irradiation par rayons X induise une cassure entre les deux locus est faible. Inversement pour des locus éloignés les uns des autres ou situés sur des chromosomes différents, ils ont peu de chances d'être présents simultanément dans un grand nombre d'hybrides. Cette technique repose donc sur l'analyse simultanée d'une collection (panel) d'hybrides contenant au minimum une centaine de clones. (Figure 22) (24) (25) (43)

Figure 22 : Différents clones d'hybrides obtenus (27)



Les différents clones d'hybrides obtenus sont  $[a^+b^+]$ ,  $[a^+b^-]$ ,  $[a^-b^+]$  et  $[a^-b^-]$ . En fonction de la fréquence de  $[a^+b^+]$  par rapport à  $[a^+b^-]$  et  $[a^-b^+]$ , les deux locus  $a$  et  $b$  sont liés ou non. Si la fréquence des clones  $[a^+b^+]$  est très supérieure à  $[a^+b^-]$  et  $[a^-b^+]$  alors les locus sont liés.

La cartographie d'hybride d'irradiation permet de cartographier des locus polymorphes et non polymorphes, de disposer de l'ADN des clones de façon illimitée (culture cellulaire) et une résolution très élevée par rapport à la cartographie génétique. (25) (43)

#### ***IV.1.2.4. Cartes cytogénétiques et cartes comparatives***

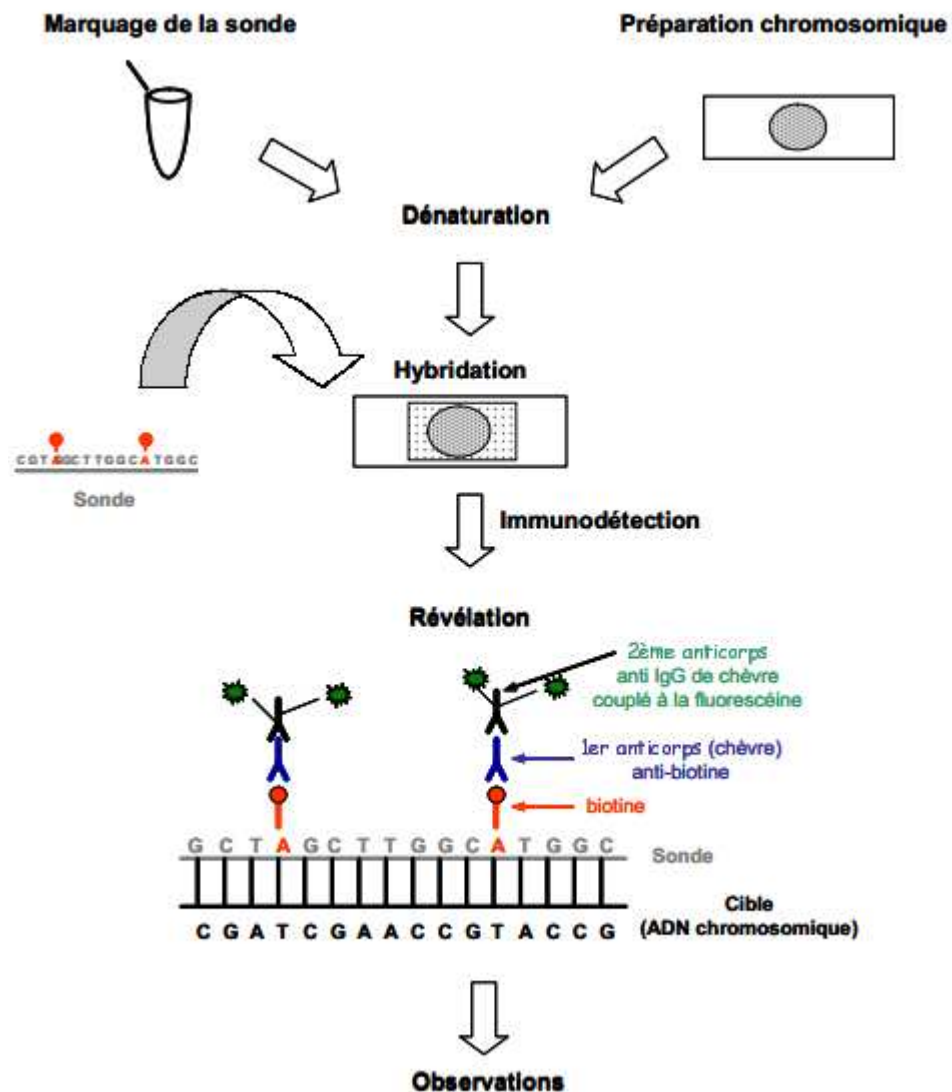
Les cartes cytogénétiques sont obtenues à partir de l'étude des chromosomes. Elles correspondent aux profils de bandes observables au microscope sur des étalements de chromosomes après coloration. Les chromosomes sont étudiés au cours du stade du cycle cellulaire où les chromosomes sont bien individualisés et observables au microscope : la métaphase mitotique. (11) (24)

Il existe différentes techniques de cytogénétiques qui sont :

- le *banding* (établissement du caryotype), il s'agit d'une coloration des chromosomes formant des bandes et permettant de les reconnaître ;
- l'hybridation *in situ* radioactive ou fluorescente sur chromosomes : elle permet d'identifier une portion chromosomique dont la séquence en bases est très homologue à celle d'un acide nucléique marqué appelé sonde. Dans la technique FISH (*Fluorescent In Situ Hybridization*), la sonde est marquée par l'incorporation de nucléotides modifiés obtenus par liaison covalente d'une molécule rapporteuse qui se liera par la suite avec une molécule fluorescente. Les chromosomes colorés sont regardés au microscope et permettent de situer la sonde dans le caryotype. L'utilisation de plusieurs sondes simultanément permet de réaliser une peinture chromosomique. (*Figure 23* et *Figure 24*) (11)



Figure 23 : Principe de l'hybridation in situ sur chromosomes (FISH) (11)



Dans cet exemple, la sonde est marquée par incorporation de biotine-14-dATP, puis hybridée sur une préparation chromosomique. Le signal d'hybridation est révélé par immunofluorescence indirecte avec deux anticorps : le premier est un anticorps IgG (immunoglobuline G) de chèvre anti-biotine, le deuxième est un anticorps anti-IgG de chèvre couplé à la fluorescéine.(11)

Figure 24 : Résultat d'une hybridation sur des chromosomiques de chien (59)



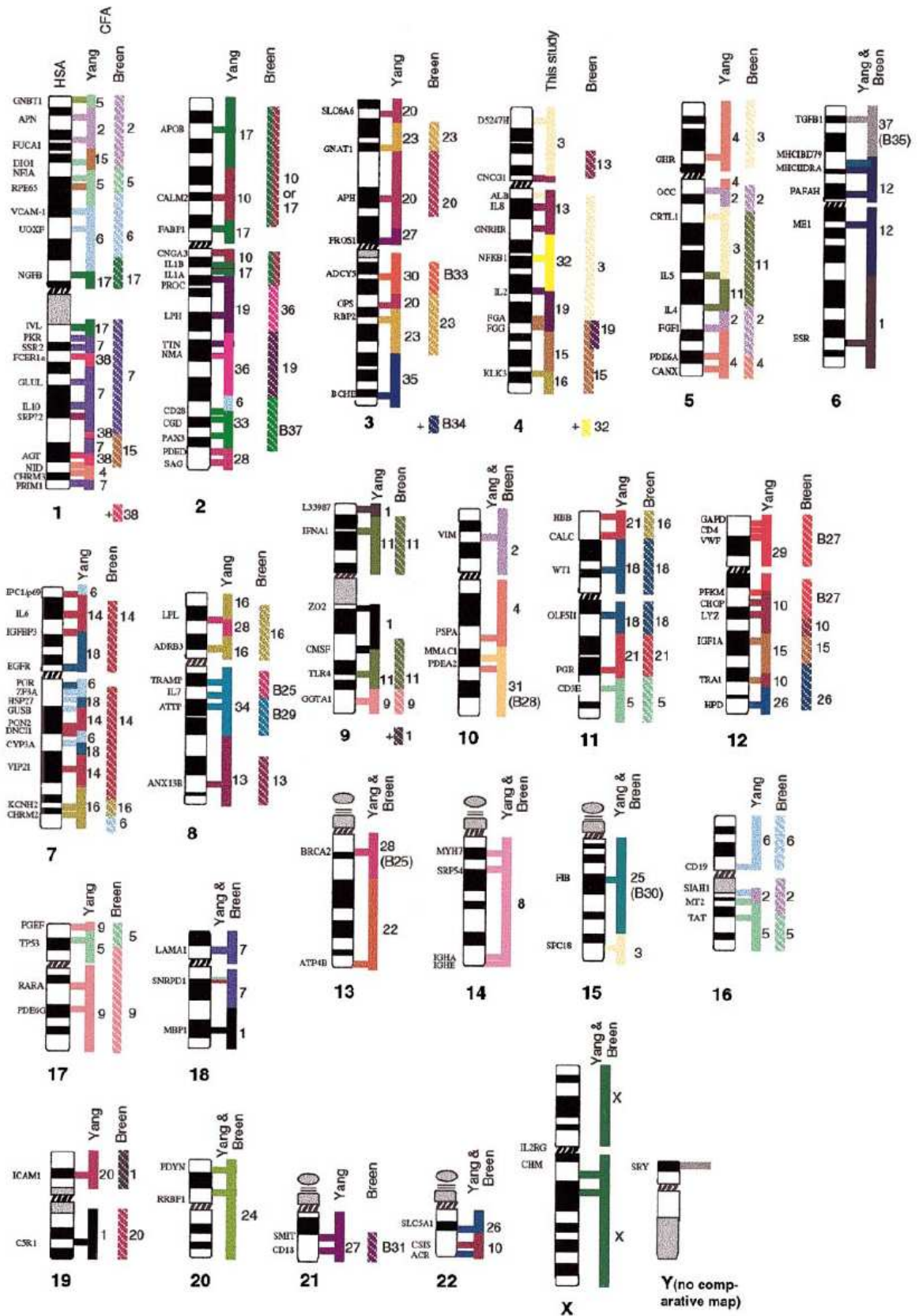
Cet exemple nous montre une photo de l'hybridation du chromosome 3 humain sur 5 paires de chromosomes canins (signal rouge).

La cytogénétique a également comme autres utilisation que de cartographier des gènes ou portions d'ADN sur les chromosomes, l'étude d'anomalies chromosomiques, des syndromes microdélétionnels, de translocations, ou la cartographie comparée entre deux espèces. (4) (46) (58) (59)

La cartographie comparée consiste à déterminer les régions chromosomiques homologues conservées dans plusieurs espèces. L'identification de ces régions permet de transposer des connaissances d'une espèce bien connue à une autre espèce moins connue ce qui facilite le l'identification de gènes impliqués dans des caractères délétères ou d'intérêt. De plus elle permet d'étudier l'évolution des génomes. La cytogénétique comparée consiste à hybrider des sondes d'ADN fluorescentes d'une première espèce, sur les chromosomes d'une autre espèce ; cette technique est appelée Zoo-FISH. (4) (7) (11) (46) (58) (59)

Chez le chien, plusieurs cartes cytogénétiques comparées avec l'homme ont été effectuées, dont celles établies par Yang et Breen (Figure 25). (4) (7) (46) (58) (59)

Figure 25 : Carte comparative entre l'homme et le chien (46)

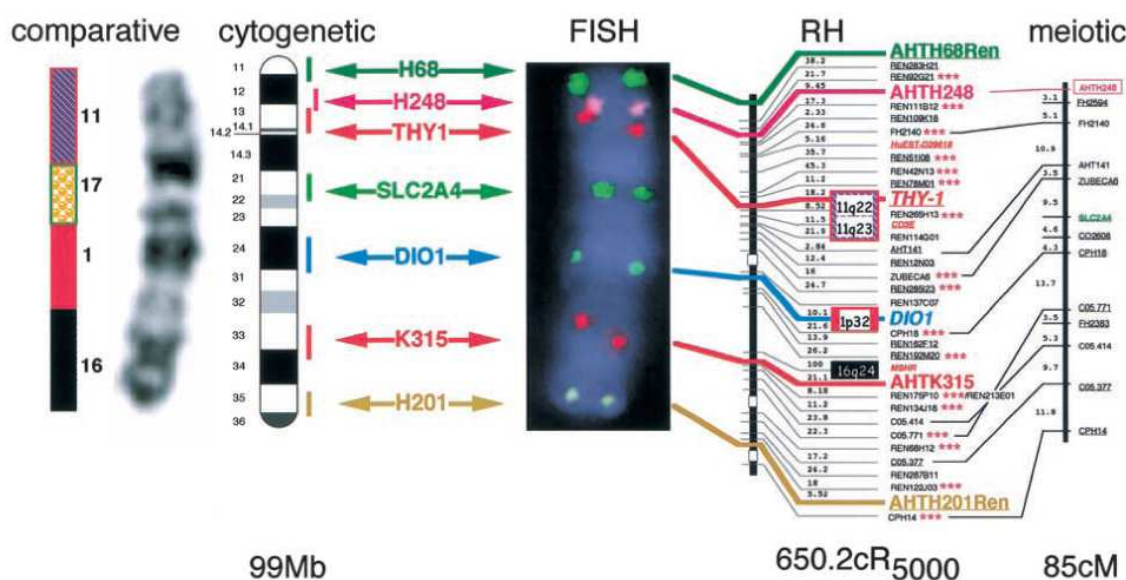


Cette carte comparative représente les chromosomes humains à gauche avec leurs bandes et à droite en couleur les chromosomes canins d'après les cartes de Yang et Breen. (46)

#### IV.1.2.5. Cartes intégrées

Les cartes intégrées sont la compilation des cartes cytogénétiques (ou chromosomiques), des cartes génétiques (indiquant les positions relatives des locus grâce aux fréquences des événements de recombinaison méiotiques), des cartes d'hybrides d'irradiation (indiquant la position relative de locus grâce aux fréquences de cassures chromosomiques dues aux rayons X) et des cartes physiques (ensembles ordonnés de clones d'ADN). Les unités de distances entre ces cartes sont différentes, ainsi une correspondance entre les cartes doit être établie, en fonction des locus présents simultanément sur les différentes cartes. Par exemple, Breen *et al.* ont établi une carte intégrée de tous les chromosomes du chien en intégrant les cartes d'hybrides d'irradiations et les cartes de liaison, avec les cartes cytogénétiques (en banding et en FISH), (Figure 26). (5) (6) (35)

Figure 26 : Carte intégrée du chromosome 5 canin (CFA 5) (6)



La carte intégrée indique la correspondance entre la carte d'hybrides d'irradiation (RH), la carte de liaison (meiotic = méiose), la carte cytogénétique en banding (cytogenetic = cytogénétique) et en FISH et la carte comparative avec l'homme. Six clones (H68, H248, THY1, DIO1, K315 et H201) ont été sélectionnés à partir de leur position le long de la carte d'hybride d'irradiation et sont indiqués en texte coloré. Le septième clone (SLC2A4) a été sélectionné sur la carte de liaison. Chacun des 7 clones ont été marqués par un fluorochrome différent et a été hybridé par la technique de FISH sur le chromosome 5 canin. Le résultat du FISH est montré au centre de la figure. Sur la gauche de la figure sont représentées les bandes du chromosome 5 (cytogenetic = cytogénétique) et à l'extrême gauche les correspondances avec les segments chromosomiques humains conservés (carte comparée = comparative). (6)

Les cartes intégrées sont utiles par exemple :

- pour établir des listes de gènes candidats pour un caractère étudié dans une espèce dont le génome est mal connu, de part leur position grâce aux synténies conservées entre deux espèces ;

- pour définir des régions candidates lors d'une approche de type clonage positionnel (voir après) ;
- pour fournir une couverture complète du génome à l'aide de marqueurs génétiques, avec des distances entre marqueurs qui soient faibles ;
- pour faciliter le clonage de gènes d'intérêt. (5) (6) (35)

#### ***IV.1.2.6. Le séquençage du génome du chien***

Le génome du chien est composé de 38 paires de chromosomes autosomes et d'une paire de chromosomes sexuels (X et Y). Le séquençage du génome du chien a été finalisé en 2005 et il a été réalisé à partir de l'ADN d'un boxer femelle. La taille totale du génome canin est de 2,41 Giga bases (Gb). Il s'est révélé plus petit que celui de l'homme (2,9 Gb). Cela s'explique par un plus faible pourcentage de séquences répétées dans le génome canin (31% dans le génome canin, 46% chez l'homme, et 38% chez la souris) (334 Mb versus 609 Mb pour l'homme). (14) (32) (42) (56)

En revanche, il existe un cas particulier d'éléments répétés appartenant à la famille des SINE (*short interspersed nuclear element*) particulièrement abondants chez le chien. Les SINE correspondent à des éléments rétrotransposables d'environ 200 pb (paire de bases) et représentent environ 7% du génome du chien (environ 11000 SINE polymorphes). Ils sont très actifs par comparaison à l'homme où le nombre d'insertions bimorphiques de SINE est inférieur à 1 000. L'insertion d'un SINE dans un exon ou dans une séquence régulatrice peut considérablement modifier le transcrit ou la régulation du gène. De la même façon, l'insertion anti-sens d'un SINE dans un intron peut engendrer une modification du mécanisme d'épissage en créant un nouveau site accepteur. Il a été montré que les SINE jouent un rôle important dans la diversification morphologique ou dans les maladies génétiques chez le chien. (14) (42)

Lors du séquençage en 2005, 19300 gènes pour le chien ont été prédits, ces gènes sont disponibles dans la base de données génomiques Ensembl (19) . Le nombre de gènes prédit chez le chien est inférieur à celui d'autres espèces telles que l'homme avec 22320 gènes ou la souris avec 23062 gènes. (14) (32) (42)

De plus, au moins 650 Mb de séquences d'ADN du génome canin coïncideraient de façon unique avec des séquences du génome humain, avec notamment près de 18500 fragments de gènes orthologues entre les gènes canins et humains, ce qui confirme l'intérêt que suscite l'espèce canine comme modèle en génétique. (14) (30) (42)

La comparaison du séquençage du boxer et du caniche et du séquençage à faible couverture de 9 autres races de chiens et de loups a permis de cataloguer plus de 2,5 millions de marqueurs génétiques de type SNP permettant une meilleure identification des locus gouvernant des phénotypes d'intérêt. (3) (14) (42) (49)

La connaissance de l'organisation du génome est essentielle pour effectuer des analyses génétiques plus poussées en vue de l'identification de gènes responsables de maladies héréditaires ou de caractères d'intérêt tels des phénotypes morphologiques. L'étude

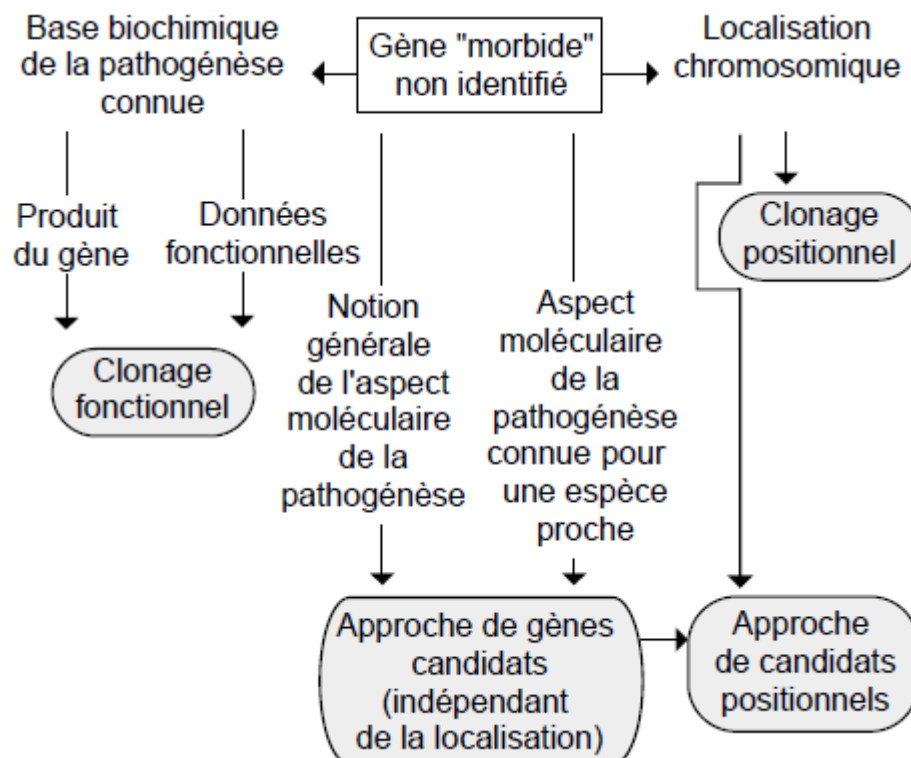
génomique comparative entre le chien et l'homme peut également permettre d'identifier des gènes d'intérêt. (49) (53)

Ainsi, les outils de cartographie permettent d'identifier des gènes impliqués dans des caractères d'intérêt. Dans cette étude, le séquençage du génome du chien nous a permis de comparer des gènes humains et leurs homologues canins afin d'identifier et d'étudier les gènes potentiellement responsables de la couleur bleue des yeux de nos chiens beagle. La technique utilisée dans cette étude pour tenter d'identifier le gène responsable du caractère yeux bleus est décrite ci-après.

### IV.1.3. Stratégies d'identification de gènes chez le chien

Il existe plusieurs méthodes pour identifier un gène dont la mutation est responsable d'une maladie génétique ou d'un caractère phénotypique (*Figure 27*). Parmi ces méthodes, la technique du clonage fonctionnel, fondée sur la biochimie, utilise des indications sur la nature ou la fonction de la protéine défectueuse et permet de remonter au gène présentant une anomalie. La cytogénétique peut également être utilisée. En effet, l'observation des chromosomes peut révéler des anomalies de type délétion ou translocation qui sont susceptibles d'orienter vers une région précise du génome. Cependant, les deux approches les plus utilisées chez le chien sont l'approche gène candidat et le clonage positionnel ; celles-ci sont décrites ci-dessous. (18) (23) (27) (28) (55)

Figure 27 : Stratégie d'identification de gène (18)



L'approche gène candidat est la technique utilisée dans la partie expérimentale de ce travail.

#### IV.1.3.1. Clonage positionnel (23) (27) (28)

Le clonage positionnel est une méthode d'identification de gènes fondée uniquement sur la localisation du gène sur les chromosomes. Il est basé sur les techniques de cartographie et de séquençage.

La première étape consiste à recenser les familles ou des cohortes dans lesquelles des individus sont atteints par la maladie ou le caractère étudié et à récolter leur ADN.

La deuxième étape est celle de la carte génétique. Elle vise à effectuer le génotypage de plusieurs centaines ou milliers de marqueurs (microsatellites ou SNP respectivement) répartis à intervalles réguliers sur l'ensemble des chromosomes de tous les individus recensés. Cette opération est nommée criblage du génome dans la littérature francophone et *genome scan* dans la littérature anglo-saxonne. Le principe consiste à utiliser des marqueurs de l'ADN suffisamment proches du gène que l'on cherche pour que la probabilité de recombinaison entre eux soit faible. On peut ainsi identifier les marqueurs dont les allèles ségrégent spécifiquement avec l'allèle responsable du caractère. Ils définissent une région sur l'un des chromosomes, appelée intervalle candidat. L'objectif est alors de multiplier le nombre (densifier) de marqueurs dans cette région afin d'affiner l'intervalle et d'approcher au plus près le gène incriminé.

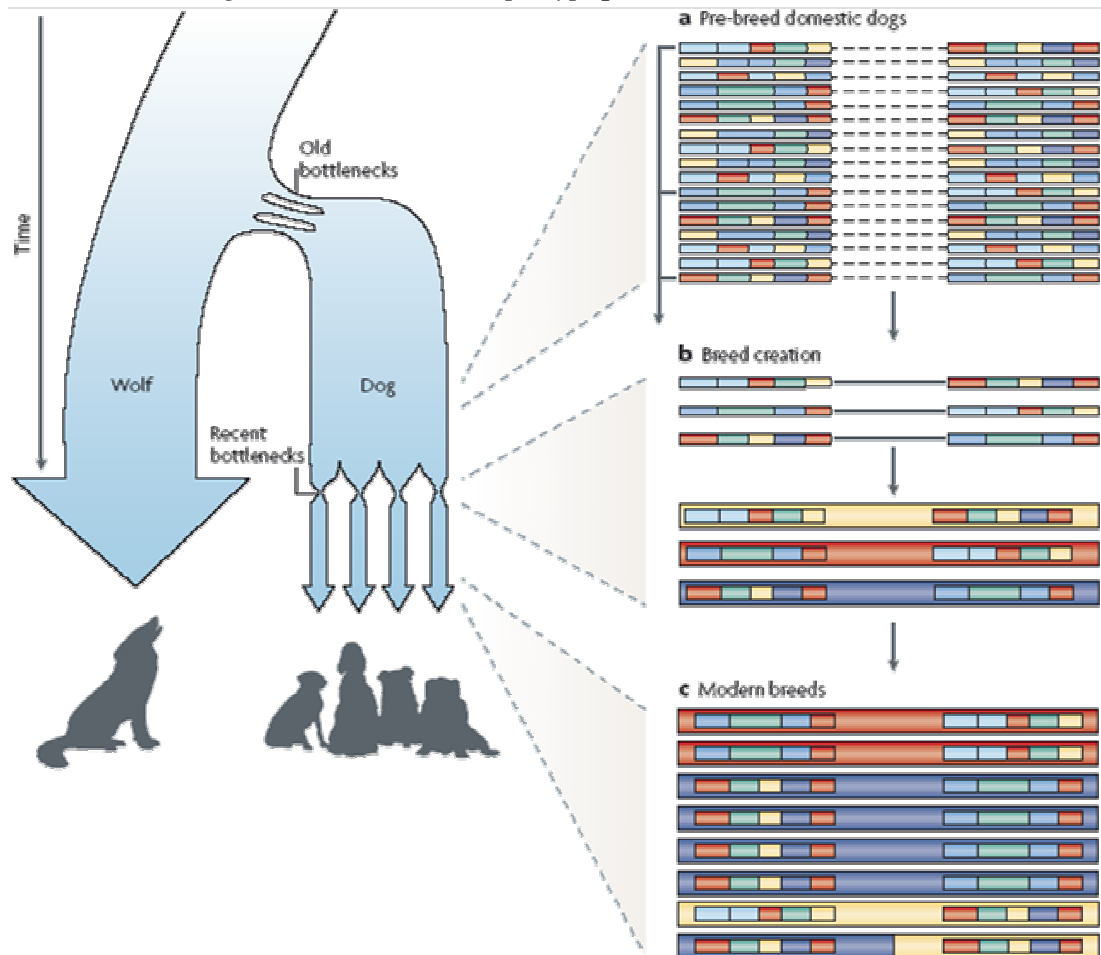
Afin de déterminer la région qui statistiquement a le plus de possibilités de contenir le gène recherché, différentes méthodes d'analyse des données de génotypage existent dont l'étude de liaison et l'étude d'association.

L'étude de liaison consiste à utiliser les liens de parentés qui existent entre les individus recrutés dans l'étude. Tous les individus sont ensuite génotypés pour les différents marqueurs. Les résultats sont alors analysés statistiquement avec, par exemple, la méthode des LOD-scores. Les LOD-scores permettent de cumuler l'information de plusieurs familles, dont l'information génétique est souvent incomplète. La méthode des LOD-scores consiste à tester statistiquement l'hypothèse de liaison génétique entre le locus recherché et le locus de chaque marqueur. Mais pour utiliser cette méthode, il est nécessaire de connaître le mode de transmission de la maladie et la pénétrance de la mutation. Il s'agit en effet d'une méthode dite paramétrique.

L'étude d'association consiste à comparer les fréquences des allèles d'un marqueur chez des sujets témoins et des sujets atteints. On ne travaille donc pas à partir des liens de parentés entre individus, mais on recrute une cohorte de patients et une cohorte de témoins. Cette méthode utilise le déséquilibre de liaison entre le marqueur et le locus du caractère étudié, qui se définit par l'existence d'une combinaison d'allèles à ces locus plus fréquente que celle attendue sous l'hypothèse d'indépendance des locus. En effet, deux locus en déséquilibre de liaison sont en général très proches sur un même chromosome. C'est-à-dire au sein d'une population, les individus portant le caractère étudié présentent la même combinaison d'allèles (un haplotype) au niveau de la région contenant le gène recherché et cet haplotype est transmis en bloc au cours des générations. On parle alors d'un effet fondateur, en supposant qu'une mutation responsable du caractère observé s'est produite chez un ancêtre commun. Les résultats des génotypages effectués dans le cas d'une étude d'association sont analysés avec le test du  $\chi^2$ . (30)

Les études d'association sont très puissantes pour identifier des gènes canins, du fait de la structure particulière du génome du chien et de l'existence de nombreuses races constituant autant d'isolats génétiques. Le clonage positionnel chez le chien utilise donc fréquemment des études d'association. Une stratégie en deux étapes a été développée pour l'espèce canine, lorsque l'on s'intéresse à un caractère partagé par plusieurs races. (*Figure 28 et Figure 29*) (28) (30) (32) (42)

*Figure 28 : Structure haplotypique du chien (28)*



*Dans l'évolution de l'espèce canine au cours du temps, il existe deux goulots d'étranglements (réduction marquée de la taille de la population suivie par la survie et l'expansion d'un petit échantillon aléatoire de la population d'origine), un ancien et un plus récent qui sont responsables de la structure haplotypique des races de chien d'aujourd'hui. Le premier correspond à la domestication du chien à partir du loup il y a environ 15000 ans. Le second est dû à la création des races dans les quelques derniers siècles. Les goulots d'étranglements ont créé la structure haplotypique et les déséquilibres de liaison (LD) particuliers des races actuelles.*

**a Pre-breed domestic dogs** / Avant la création des races actuelles, la population canine possédait un déséquilibre de liaison de courte distance (ce qui est attendu du fait de la grande taille de la population et de la longue période de temps entre les deux goulots d'étranglements).

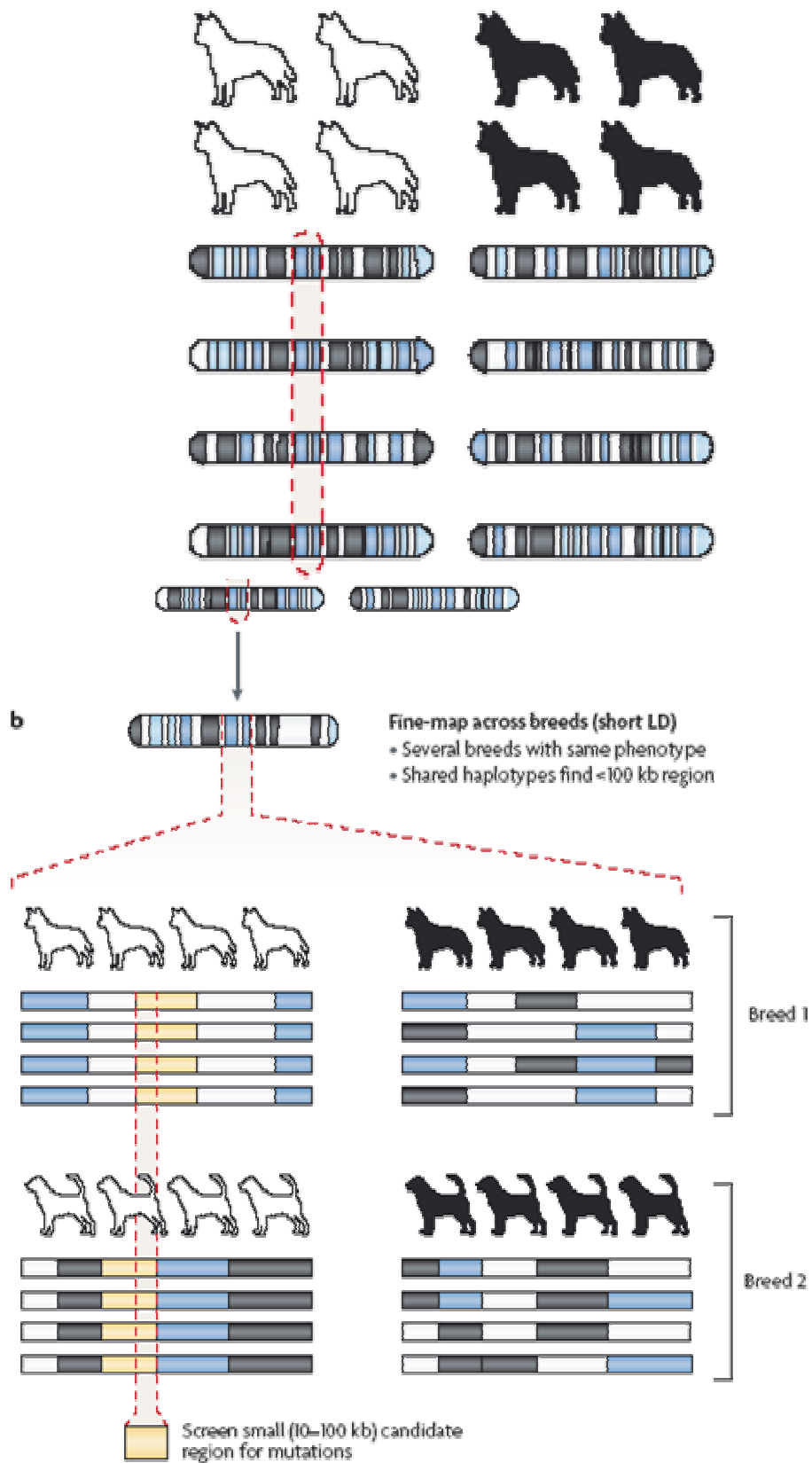
**b Breed creation** / Au moment de la création des races, des petits sous-ensembles de chromosomes ont été sélectionnés dans la population des chiens domestiques (du fait de la



*sélection d'un petit nombre d'individus fondateurs pour chaque race). Des blocs de grande taille portés par ces chromosomes sont devenus courants dans chaque race créant ainsi un déséquilibre de liaison grande taille (sur une longue distance chromosomique).*

*c Modern breeds / Depuis la création des races canines qui est récente, ces blocs de grande taille n'ont pas encore été fortement remodelés par les recombinaison. Cependant, ces blocs haplotypiques de grande taille conservent toujours les petits blocs haplotypiques de courte taille ancestraux de la population d'origine des chiens domestiques. On retrouve ces derniers lorsque l'on examine les chromosomes dans de nombreuses races.*

Figure 29 : Stratégie de cartographie en deux étapes (28)



L'approche en deux étapes tire pleinement partie du déséquilibre de liaison (LD) étendu au sein de chaque race et du déséquilibre de liaison restreint lorsque l'on compare les races

entre elles. Dans cet exemple, les chiens porteurs d'une mutation de la perte de la pigmentation (blanc) sont comparés avec les chiens normalement pigmentés (noir).

*a* | Dans la première étape, l'étude d'association à l'échelle de l'ensemble du génome, au sein d'une même race, utilise un nombre restreint d'individus et plusieurs milliers de marqueurs SNP pour identifier une ou plusieurs régions (en général de grande taille) qui sont associées au caractère étudié.

*b* | Dans la seconde étape, une cartographie fine est réalisée en utilisant beaucoup plus de SNP, dans plusieurs races qui partagent le même phénotype. Ceci permet de réduire la région candidate contenant la mutation à une courte région, du fait du court déséquilibre de liaison existant entre les races (en général une région de 10-100kb).

La troisième étape consiste à recenser, dans l'intervalle candidat, les gènes présents et sélectionner ceux pouvant être responsables du phénotype étudié, en analysant l'expression et les rôles potentiels de chaque gène. Les gènes retenus sont ensuite séquencés chez quelques cas et quelques contrôles afin d'identifier la mutation. Il reste ensuite à valider la/les mutation(s) identifiée(s) du point de vue fonctionnel : impact sur la structure de la protéine codée, fonction de la protéine et implication dans le déterminisme du phénotype étudié. Cette étape de validation fonctionnelle est parfois très complexe.

Le clonage positionnel consiste donc à compléter la carte génétique de la région, en définissant la position du locus recherché sur le chromosome, puis à identifier les gènes de cette région pour mettre en évidence un gène candidat. Il s'agit d'une méthode longue et difficile. C'est pour cela qu'elle est souvent complétée par la méthode gène candidat décrite ci-dessous.

#### **IV.1.3.2. Gène candidat**

L'approche gène candidat consiste à poser l'hypothèse qu'un gène connu peut être responsable du phénotype étudié. Lorsqu'un caractère a été décrit de façon détaillée, il devient possible d'utiliser des banques de données pour identifier les gènes candidats. L'objectif est de confirmer ou d'infirmer l'hypothèse. (18) (55)

Pour cela, il existe quatre façons de définir le gène candidat :

- gène candidat positionnel : le chromosome ou une région du chromosome responsable du phénotype a été identifié (par exemple par clonage positionnel). Dans ce cas, deux manières de rechercher le gène candidat sont possibles :
  - soit l'inventaire des gènes présents dans la région est réalisé à partir d'une base de données. Pour chaque gène de cette région, on regarde quels gènes peuvent convenir en fonction de leurs rôles et des informations connues sur le caractère étudié. Ainsi, une liste de gènes candidats est établie. Pour cela, il faut que l'espèce bénéficie d'une cartographie complète de ses gènes : tous les gènes de la région ainsi que leurs fonctions putatives doivent être connus. C'est globalement le cas des espèces dont le génome a été entièrement séquencé et annoté (la fonction putative des différents gènes a été indiquée) : homme, souris ou chien ; (18) (19) (55)

- soit pour les espèces où la cartographie est incomplète, on recherche la région homologue chez une autre espèce où la cartographie est plus complète et on en déduit les gènes homologues potentiellement présents dans la région. Grâce aux cartes comparatives entre espèces, on peut connaître les régions homologues conservées dans plusieurs espèces. (7) De la même façon que ci-dessus, on en déduit la liste des gènes candidats ; (18) (55)
- gène candidat indépendant de la position : la position du locus responsable du phénotype étudié n'est pas connue. Dans ce cas, deux manières de rechercher les gènes candidats sont possibles :
  - soit un phénotype similaire à celui étudié chez le chien a été décrit dans une autre espèce (homme, souris) pour laquelle le gène responsable est connu. On recherche alors le gène homologue chez le chien. S'il existe, il s'agit alors d'un gène candidat pour le phénotype canin ; (18) (55)
  - soit l'étude physiopathologique, cellulaire ou moléculaire du caractère analysé suggère que le gène responsable appartient à une famille de gènes connus, on en déduit alors les gènes candidats possibles pour ce caractère. (18) (55)

Ensuite, le gène candidat est séquencé chez quelques individus présentant le caractère et quelques individus témoins ne le présentant pas. En comparant les séquences pour ces deux groupes, une mutation causale peut parfois être identifiée.

Comme nous l'avons expliqué précédemment, les deux techniques de clonage positionnel et d'approche gène candidat sont complémentaires. Le clonage positionnel appartient à la génétique dite « inverse » : le gène impliqué dans le caractère héréditaire est identifié directement, sans que la protéine défectueuse soit connue. En effet, en détaillant la méthodologie du clonage positionnel suivie de l'approche gène candidat, on se rend compte qu'aucune étape ne concerne l'anomalie de la protéine codée par le gène défectueux. Seule l'anomalie génétique est étudiée, et seule la liaison (ou l'association) entre le gène incriminé et le caractère étudié est démontrée. C'est seulement lorsque la séquence nucléotidique du gène est connue que l'on peut déduire la séquence d'acides aminés qui compose la protéine correspondante. À partir de cette séquence, on peut aboutir à la structure de la protéine normale et de la protéine mutée et comprendre la physiopathologie de la maladie. (18) (55)

Dans le cas des yeux bleus chez le beagle, la recherche des gènes candidats s'est faite en déterminant les gènes responsables de la couleur des yeux bleus chez l'homme, car chez le chien aucune étude n'a été réalisée à ce jour sur le déterminisme de la couleur des yeux non liée à la couleur du pelage. De plus, chez l'homme, plusieurs gènes responsables de la couleur des yeux bleus ont été identifiés. Ils sont décrits ci-après. Nous avons donc postulé que le phénotype de l'homme et du beagle étaient comparables et nous avons analysé si les gènes canins homologues des gènes humains pouvaient être responsables de la couleur des yeux bleus chez le beagle.

## IV.2. Gènes responsables de la pigmentation de l'œil chez l'homme

Chez l'homme, la couleur des yeux a une haute héritabilité de 98%, ce qui a permis de conclure que la couleur des yeux est déterminée génétiquement (36). La transmission génétique de la couleur des yeux n'est pas strictement mendélienne bien que la couleur bleue de l'iris suive en grande partie un modèle récessif. Ainsi, plusieurs gènes seraient responsables de la transmission de la couleur des yeux : c'est un caractère polygénique. Cependant Zhu *et al.* (16) ont estimé qu'un ou plusieurs locus dans la région qui contient le gène *OCA2* du chromosome 15 expliqueraient 74% de la variance du caractère « couleur des yeux ». Ceci confirmait qu'il existe une base génétique à la variation de la couleur des yeux chez l'homme. Les gènes *OCA2* (*oculocutaneous albinism type 2*) /*HERC2* (*HECT and RLD domain containing E3 ubiquitin protein ligase 2*) expliqueraient la transmission de la couleur bleue et brune des yeux tandis que d'autres gènes tels que *SLC2A4* (*solute carrier family 2 (facilitated glucose transporter), member 4*), *TYR* (*tyrosinase*), *TYRP1* (*tyrosinase related protein 1*), *SCL45A2* (*solute carrier family 45, member 2*) et *IRF4* (*interferon regulatory factor 4*) contribueraient à la nuance de la couleur des yeux. Ces derniers gènes auraient donc des effets mineurs par rapport à *OCA2/HERC2*. (33)

### IV.2.1. Gène *OCA2* (*oculocutaneous albinism type 2*)

Des mutations dans le gène *OCA2* ont été identifiées comme responsables de l'albinisme oculo-cutané de type 2. Le gène contient 24 exons et a une longueur d'environ 3186 pb. Il code une protéine transmembranaire avec 12 domaines ressemblant à des canaux ioniques. La protéine se situe dans les mélanosomes. La fonction de la protéine n'est pas connue mais elle pourrait être impliquée dans la fabrication et la circulation de la tyrosinase, dans la régulation du pH du mélanosome ou dans la régulation du métabolisme du glutathion. Le gène *OCA2* aurait ainsi un rôle dans la régulation de la mélanogénèse, comme le montre l'hypopigmentation de la peau, des cheveux et des yeux qui survient à la suite d'une délétion de la région englobant le gène *OCA2* situé sur le chromosome 15 humain (syndromes de Prader-Willi et d'Angelman). Inversement, des insertions de copies de la région englobant le gène *OCA2* entraînent une hyperpigmentation de la peau. De plus le gène *OCA2* humain possède un homologue chez la souris qui est responsable du phénotype yeux rouges dilué dans cette espèce. (15) (16) (51) (52)

Donnelly *et al.* (15) ont réalisé une étude d'association entre les différents SNP dans la région du gène *OCA2* chez des individus de différentes populations. Trois SNP dans l'intron 1 du gène *OCA2* étaient associés aux phénotypes de la couleur des yeux. Les trois SNP étaient rs4779138, rs4778241, rs7495174 et l'haplotype associé aux yeux bleus était A-C-A.

Un autre SNP (rs1800407) a été identifié dans l'exon 13 et était associé à la couleur verte des yeux. De plus, il augmentait la pénétrance du phénotype yeux bleus chez les individus A-C-A lorsque l'allèle T de ce SNP était présent. (15)

D'autres associations plus fortes ont été trouvées à proximité du gène *OCA2*, dans le gène *HERC2*. (15)

### **IV.2.2. Gène *HERC2* (*HECT and RLD domain containing E3 ubiquitin protein ligase 2*)**

Le gène *HERC2* est situé en amont du gène *OCA2*. La fonction du gène *HERC2* est inconnue mais la protéine codée montre des similitudes avec une protéine ligase ubiquine E3. De plus, il a été montré que des régions non codantes de *HERC2* contrôlent l'expression du gène *OCA2*. L'intron 86 de *HERC2* contient une région régulatrice de l'expression du gène *OCA2* et c'est une région hautement conservée entre différentes espèces.(52)

Donnelly *et al.* (15) ont réalisé une étude d'association entre différents SNP situés dans une région du gène *HERC2* chez des individus de différentes populations. Quatre SNP ont été associés à la couleur des yeux. Le SNP rs12913832 est situé dans l'intron 86, il est dans une région fortement conservée et il est fortement associé à la couleur des yeux. Le SNP rs1129038 est situé dans la région 3'UTR (*UnTranslated Region*) de l'exon 93 qui a une séquence moins conservée entre différentes espèces. De même, il est fortement associé à la couleur des yeux et il est proche du SNP rs129113832. Les yeux bleus sont associés à l'allèle T de rs1129038 et à l'allèle G de rs12913832 ce qui forme l'haplotype T-G. Le SNP rs916977 est situé dans l'intron 12. Le SNP rs1667394 est localisé dans l'intron 4. Ces deux SNP sont associés à la couleur des yeux et l'haplotype associé à la couleur yeux bleus est C-A. Mais ces marqueurs ne sont pas universels et parfois absents dans certaines populations.

Ainsi, les marqueurs SNP rs1129038 et rs12913832 étaient les meilleurs marqueurs pour les yeux bleus. Ils sont certainement proches de l'allèle responsable de la couleur des yeux bleus (en déséquilibre de liaison) à moins que ces deux SNP ne soient causaux. (15) (36) (52)

### **IV.2.3. Autres gènes**

D'autres gènes chez l'homme, dont *SLC2A4*, *TYR*, *TYRP1* (*tyrosinase related protein 1*), *SCL45A2* (*solute carrier family 45, member 2*) et *IRF4* (*interferon regulatory factor 4*) ont été identifiés comme responsables de la grande diversité de la couleur des yeux et jouent un rôle dans les nuances de la couleur. Ils auraient donc des effets mineurs sur la couleur des yeux.

Le gène *IRF4* aurait une plus faible association avec la couleur des yeux par rapport aux autres gènes.

En revanche, le SNP rs16891982 du gène *SLC45A2* était fortement associé avec les yeux clairs versus yeux foncés. Mais cette association était plus faible qu'avec le locus *OCA2/HERC2*. (36) (51)

### **IV.2.4. Détection de trois nouveaux locus par la recherche de QTL**

#### **IV.2.4.1. Notion de QTL (24)**

Un QTL (*quantitative trait locus*) se définit comme un locus dont au moins deux formes alléliques sont associées à des valeurs phénotypiques moyennes différentes pour le caractère étudié. Ainsi lors d'une étude d'association entre le phénotype d'une population et sa composition génétique, il peut exister un ou plusieurs phénotypes associés à des différences de génotype dans certaines régions du génome.

Un caractère quantitatif présente une variation continue du phénotype et peut être sous le contrôle d'un ou plusieurs QTL.

#### **IV.2.4.2. Trois nouveaux locus (33)**

Dans l'étude de Liu *et al.* (33), la couleur des yeux a été quantifiée de façon linéaire, permettant ainsi d'améliorer le pouvoir d'indentification d'associations génétiques avec la couleur des yeux. Ainsi, trois nouveaux locus ont été identifiés.

Sur le chromosome 1 (1q42.3), deux SNP (rs3768056, rs9782955) étaient associés à la couleur des yeux. Ils se situaient dans des introns du gène *LYST* (*lysosomal trafficking regulator*) qui régule la migration des mélanosomes. Le gène *LYST* peut influencer l'intensité de la couleur des robes chez la vache. Il est susceptible d'introduire de subtiles variations dans la pigmentation.

Sur le chromosome 17 (17q25.3), trois SNP étaient associés à la couleur des yeux. Deux SNP (rs7219915, rs9894419) se situaient dans un intron et un exon du gène *NPLOC4* (*nuclear protein localization 4 homolog (S. cerevisiae)*) et l'autre SNP (rs12452184) était localisé dans un intron du gène *HGS* (*hepatocyte growth factor-regulated tyrosine kinase substrate*). Dans la région, il existait de plus de nombreux petits gènes. Le meilleur candidat avec l'association observée était le gène *PDE6G* (*phosphodiesterase 6G, cGMP-specific, rod, gamma*) qui est responsable d'un dysfonctionnement de l'épithélium pigmentaire de la rétine lorsqu'il est muté.

Sur le chromosome 21 (21q22.3), cinq SNP ont été identifiés comme étant associés à la couleur des yeux. Quatre SNP (rs1003719, rs2252893, rs2835621, rs2835630) étaient situés dans l'intron du gène *TTC3* (*tetratricopeptide repeat domain 3*) et l'autre SNP (rs7277820) était localisé dans le 5'UTR du gène *DSCR9* (*Down syndrome critical region gene 9*). Ces deux gènes appartiennent au même block en déséquilibre de liaison. Le gène *DSCR9* est par ailleurs un gène spécifique de la lignée des primates.

De plus, dans cette étude, le locus *OCA2/HERC2* était aussi très fortement associé à la couleur des yeux en particulier le SNP rs12913832 localisé dans l'intron 86 de *HERC2*.

Ainsi, chez l'homme plusieurs gènes et régions chromosomiques ont été identifiés comme impliqués dans le déterminisme de la couleur des yeux.