

GÉNÉTIQUE MOLÉCULAIRE HUMAINE

4^e édition



Tom STRACHAN et Andrew READ

Médecine Sciences
Publications

Lavoisier

Hybridation des acides nucléiques : principes et applications

7

CONCEPTS CLÉS

- Dans l'hybridation des acides nucléiques, des populations d'acides nucléiques ou d'oligonucléotides bien caractérisées (*sondes*) sont utilisées afin d'identifier des séquences apparentées dans un mélange complexe et souvent mal compris d'acides nucléiques à tester.
- L'hybridation des acides nucléiques repose sur la spécificité de l'appariement des bases. Les acides nucléiques de la sonde et du mélange complexe à tester sont sous forme de simples brins et mélangés pour permettre la formation d'hétéroduplex entre la séquence de la sonde et toute séquence de l'échantillon à tester qui serait en partie ou en totalité (*cible*) complémentaire.
- Les hétéroduplex sont détectés par marquage d'un acide nucléique en solution aqueuse qui pourra s'hybrider avec un mélange d'acides nucléiques non marqués, fixés sur un support solide. Après lavage, afin d'enlever les acides nucléiques marqués non hybridés, tout marquage restant sur le support solide est constitué d'hétéroduplex sonde-cible.
- La stabilité d'un hétéroduplexe dépend du nombre de bases appariées. Elle est affectée par certains paramètres comme la longueur du segment présentant des bases appariées, la température, et la force ionique de l'environnement.
- Les sondes ADN ou ARN sont constituées en général de centaines de nucléotides et utilisées pour identifier des séquences cibles présentant un haut degré de similitude avec la sonde.
- De courtes sondes oligonucléotidiques (moins de 20 nucléotides de long) peuvent être utilisées pour différencier des cibles qui ne diffèrent que par la position d'un seul nucléotide.
- Les acides nucléiques sont marqués par incorporation de nucléotides porteurs d'un isotope radioactif ou de radicaux modifiés chimiquement pouvant être détectés par dosage approprié.
- De nombreuses techniques d'hybridation impliquent la liaison des acides nucléiques contenant la cible à une surface solide, puis son exposition à une solution contenant la sonde marquée. L'ADN immobilisé peut être isolé ou être présent dans des cellules ou des chromosomes immobilisés.
- L'hybridation sur micropuces permet de faire de nombreux essais d'hybridation simultanés. On fixe à la surface d'un support solide, sous forme d'une grille à forte densité, des milliers d'ADN ou de sondes oligonucléotidiques non marqués et on les utilise pour cribler des échantillons complexes en solution contenant un mélange d'ADN ou d'ARN marqués.
- Les principales applications de l'hybridation sur micropuces sont celles qui permettent d'établir un profil d'expression des gènes et d'analyser des modifications de l'ADN.

Les molécules d'ADN sont très longues et se cassent facilement quand elles sont isolées des cellules, ce qui les rend difficiles à étudier. Dans le chapitre 6, nous avons étudié la façon par laquelle le clonage de l'ADN – soit dans des cellules, soit *in vitro* – permet à des molécules d'ADN d'être sélectivement répliquées en un nombre de copies très élevé. Quand il est amplifié par ce moyen, l'ADN est efficacement purifié et le clonage de l'ADN permet d'étudier des séquences ADN isolées de l'ADN génomique ; il est aussi possible d'étudier des molécules d'ARN isolées une fois qu'elles ont été recopiées en ADNc. Dans ce chapitre, nous considérons une approche complètement différente. Au lieu d'essayer de purifier individuellement des séquences d'acides nucléiques, l'objectif est de les détecter spécifiquement et de suivre leur devenir au sein d'une population complexe qui représente un échantillon biologique ou d'intérêt médical.

En recherche fondamentale, l'hybridation des acides nucléiques est souvent utilisée pour suivre à la trace les transcrits ARN et ainsi obtenir des informations sur la façon dont les gènes sont exprimés. Elle permet aussi d'identifier des relations entre des séquences ADN provenant de différentes origines. Par exemple, une séquence ADN de départ peut être utilisée pour identifier des séquences apparentées provenant de différents organismes ou d'individus différents de la même espèce, ou même des séquences apparentées du même génome que la séquence de départ. Souvent, l'hybridation des acides nucléiques est aussi utilisée en tant que moyen pour identifier des allèles porteurs de maladie ou des transcrits aberrants associés à une maladie.

7.1 PRINCIPES DE L'HYBRIDATION DES ACIDES NUCLÉIQUES

Dans l'hybridation des acides nucléiques, une population d'acides nucléiques connus permet d'analyser une population d'acides nucléiques mal connus

L'hybridation des acides nucléiques est un outil de base en génétique moléculaire. Elle repose sur la capacité des acides nucléiques simple brin, dont les séquences sont partiellement ou totalement complémentaires, à former une molécule bicaténaire (double brin) par appariements des bases les unes avec les autres (**hybridation**). L'Encadré 7.1 est un glossaire des termes qui s'y rapportent, pour aider les lecteurs non familiers avec certains aspects de la terminologie.

ENCADRÉ 7.1 GLOSSAIRE POUR L'HYBRIDATION DES ACIDES NUCLÉIQUES

Hybrider. Permettre la formation de liaisons hydrogène entre deux simples brins. Si deux acides nucléiques simples brins partagent suffisamment de *bases complémentaires*, ils formeront un complexe double brin d'acides nucléiques. Hybrider est le contraire de *dénaturer*.

ARN antisens. Séquence ARN qui est complémentaire d'un transcrit ARN, ce qui permet l'appariement des bases entre ces deux séquences.

Complémentarité des bases. Degré par lequel les séquences de deux acides nucléiques simple brin peuvent former un complexe double brin selon la règle de l'appariement de Watson et Crick (A s'apparie à T ou U, C s'apparie à G).

Dénaturer. Séparer les brins d'un complexe double brin par rupture des liaisons hydrogène les maintenant assemblés. Peut être obtenu par chauffage ou par exposition de l'ADN à un milieu salin ou à un solvant très polaire comme l'urée ou la formamide. Dénaturer est le contraire d'*hybrider*.

Puce à ADN. Tout arrangement très dense de rangées bien ordonnées (*micropuce*) de clones d'ADN ou d'oligonucléotides, qui est utilisé dans une analyse par hybridation.

Caractéristique (d'une micropuce d'ADN ou d'oligonucléotide). Le grand nombre de molécules d'ADN ou d'oligonucléotides identiques situées en tout point de la micropuce.

Hétéroduplex. Acide nucléique double brin formé par l'appariement des bases entre deux acides nucléiques simple brin qui ne proviennent pas du même allèle. Il peut y avoir 100 % de bases appariées entre les deux séquences d'un hétéroduplex, surtout quand une des séquences hybridée est une sonde oligonucléotidique ou quand les séquences proviennent de différents allèles du même gène ou encore qu'elles sont proches et apparentées du fait de l'évolution.

Homoduplex. ADN double brin formé à partir de deux séquences simple brin ayant pour origine le même allèle qui se sont ensuite hybridées. Un homoduplex ayant 100 % de ses bases appariées, il est donc plus stable qu'un hétéroduplex.

Dosage par hybridation. Expérience dans laquelle une même population bien caractérisée d'acides nucléiques ou d'oligonucléotide (la *sonde*) est dénaturée en simples brins et

utilisée pour rechercher des séquences complémentaires (les *cibles*) dans une population mal connue de séquences nucléotidiques qui s'hybrident pour former des hétéroduplex.

Stringence d'hybridation. Degré de tolérance de mésappariements de bases dans des hétéroduplex dans les conditions de l'hybridation. À forte stringence, seules les séquences les plus parfaitement complémentaires peuvent appairer leurs bases. Néanmoins, des hétéroduplex présentant un nombre significatif de mésappariements peuvent devenir stables quand les conditions de stringence sont abaissées, par réduction de la température d'hybridation ou par augmentation de la concentration saline.

Hybridation *in situ*. Réaction d'hybridation dans laquelle une sonde marquée est hybridée aux acides nucléiques pour détecter leur localisation morphologique dans des cellules ou des chromosomes fixés.

Température de fusion (T_m). Température correspondant au point situé à mi-chemin dans la transition entre la forme double brin et la forme simple brin des acides nucléiques.

Micropuces. Surface solide, constituée d'une grille de format bien défini et très dense, sur laquelle on peut fixer les molécules d'intérêt. Elles sont utilisées pour certains dosages. Une micropuce à ADN ou à oligonucléotides a de nombreuses molécules d'ADN ou d'oligonucléotides non marquées, fixées sur des positions précises de la micropuce, qui agissent comme des sondes dans une expérience de mesure quantitative d'hybridation. Chaque position particulière comprend des milliers de copies identiques d'un type particulier d'ADN ou d'oligonucléotide, constituant sa *caractéristique*.

Sonde. Population connue d'acides nucléiques ou d'oligonucléotides, utilisée dans une étude par hybridation afin d'analyser une population souvent complexe d'acides nucléiques pour identifier des séquences apparentées à la *cible* pouvant former des hétéroduplex.

Ribosonde. Sonde à ARN.

Similitude des séquences. Degré par lequel deux séquences sont identiques en séquence ou en complémentarité des bases.

Cible. Séquence d'acides nucléiques qui présente suffisamment de similitude dans sa séquence avec la sonde à laquelle elle peut s'apparier pour former un hétéroduplex stable.

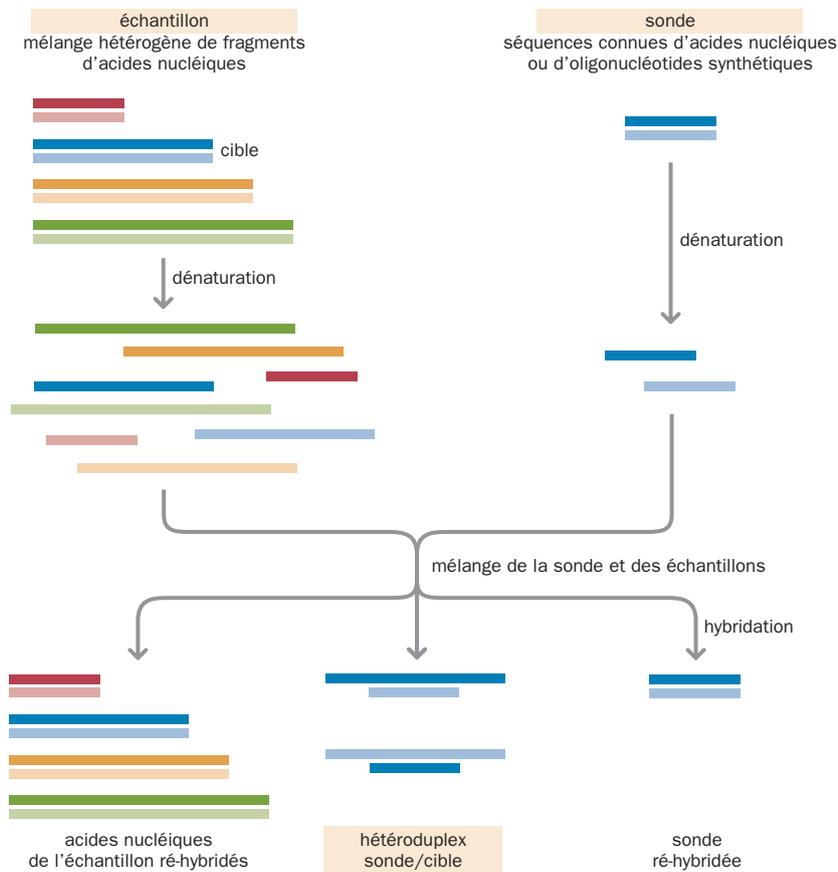


Figure 7.1 Formation d'hétéroduplex sonde/cible lors d'un test par hybridation d'acides nucléiques. Un échantillon test est composé d'un mélange complexe d'acides nucléiques et d'une population de sondes bien définies de composition connue, soit en acides nucléiques soit en séquences oligonucléotidiques. Ces deux populations sont transformées en simples brins, puis mélangées pour, à nouveau, en permettre l'hybridation. Les séquences qui avaient au préalable été appariées dans l'échantillon test et la sonde s'apparient à nouveau pour former des homoduplex (en bas à gauche et à droite). De plus, de nouveaux hétéroduplex se forment aussi entre la sonde et certaines séquences cibles ayant des séquences complémentaires ou partiellement complémentaires (en bas, au centre). Les conditions d'hybridation peuvent être ajustées pour favoriser la formation d'hétéroduplex. De cette façon, les sondes se lient de façon sélective et permette d'identifier les acides nucléiques apparentés dans une population complexe d'acides nucléiques.

L'hybridation des acides nucléiques peut être faite de nombreuses façons différentes, mais il y a un principe de base sous-jacent commun : une population *connue* et bien caractérisée de molécules d'acides nucléiques ou d'oligonucléotides synthétiques est utilisée pour analyser une population complexe (mêlée) d'acides nucléiques dans un échantillon à tester imparfaitement connu, ayant un intérêt biologique ou médical.

L'échantillon test à étudier peut contenir des molécules d'ADN comme l'ADN total des leucocytes d'un individu ou d'un type particulier de cellules tumorales, ou il peut contenir de l'ARN comme les ARN totaux ou l'ARNm exprimé par une lignée particulière de cellules ou de tissu. Dans tous les cas, si elles doivent participer à une hybridation des acides nucléiques, les molécules d'acides nucléiques de l'échantillon doivent être mises sous forme de simples brins (**dénaturation**). L'ADN cellulaire est naturellement double brin, mais l'ARN a aussi une quantité significative de régions double brin dues aux liaisons hydrogène intrachâînes. Les liaisons hydrogène intra- ou interchâînes peuvent être rompues de différentes façons. La dénaturation initiale implique le chauffage ou les traitements alcalin ; selon les besoins, les acides nucléiques peuvent aussi être exposés à des molécules fortement polaires, comme la formamide ou l'urée, afin de les conserver sous forme simple brin pendant de longues périodes.

La population connue qui est utilisée pour l'analyse comprend des séquences d'acides nucléiques très bien définies qui doivent être dénaturées, ou des oligonucléotides simples brins synthétiques. Chaque type de molécule dans cette population agira en tant que **sonde** afin de localiser les acides nucléiques complémentaires ou partiellement complémentaires (**cibles**) dans l'échantillon à tester. Pour ce faire, les simples brins de l'échantillon à tester et de la population sonde sont mélangés pour permettre l'appariement des bases entre les simples brins de la sonde et ceux de la cible complémentaire (**hybridation**). L'objectif est de former des **duplex hétérogènes** sonde-cible (Figure 7.1) ; la spécificité de l'interaction entre la sonde et les séquences cibles dépend de leur degré d'appariement.

Les hétéroduplex formés entre la sonde et la cible sont plus simples à identifier après fixation sur un support solide

Il est difficile d'identifier les duplex hétérogènes sonde-cible en milieu liquide. Pour faciliter cette identification, les échantillons d'acides nucléiques à tester ou les sondes

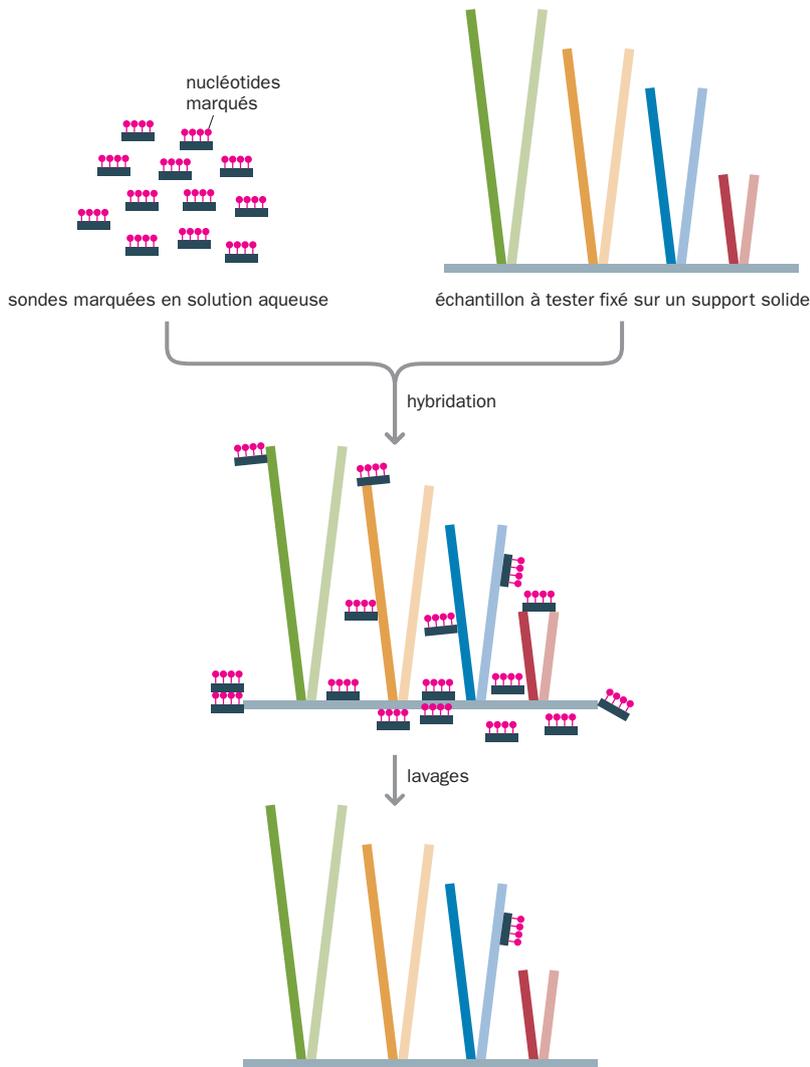


Figure 7.2 L'identification des hétéroduplex sonde-cible est bien plus efficace après capture sur un support solide. Les dosages par hybridation impliquent habituellement soit la fixation de la population d'échantillons à tester à un support solide (comme montré ici), soit la fixation de la population de sonde au support solide. La population immobilisée est liée de telle sorte que différentes séquences sont fixées en des positions bien définies sur la surface solide, mais pas nécessairement ancrées par l'une de leurs extrémités comme présenté ici. L'autre population est marquée puis ajoutée à la population immobilisée. Ici, nous avons utilisé l'exemple présenté Figure 7.1 dans lequel la sonde est homogène et l'échantillon à tester un mélange hétérogène de différentes molécules d'ADN. Pour plus de simplicité, nous ne montrons ici qu'un seul brin de sonde marquée. Au cours du passage des molécules de sondes simple brin marquées au contact du support solide, elles peuvent être captées par hybridation aux molécules cibles complémentaires fixées au support solide. Une fois que l'hybridation a eu lieu, on lave abondamment pour retirer l'excès de sonde marquée restant en solution ou liées de façon non spécifique à la surface. Tout marquage restant sur le support solide indique la présence d'hétéroduplex sonde-cible, ce qui permet d'identifier la séquence cible.

sont liés par un moyen quelconque à un support solide, souvent une membrane en plastique, en verre ou une mince couche de quartz. L'autre population – soit la sonde, soit l'échantillon à tester – est apportée en solution liquide et est marquée en y attachant une molécule qui porte un marqueur radioactif ou un groupement chimique pouvant être facilement détecté. La population marquée est mise en contact avec le support solide afin que les deux populations puissent interagir et former des hétéroduplex.

Une autre approche qui consiste à utiliser une sonde marquée en solution alors que l'échantillon à tester est immobilisé est illustrée Figure 7.2. Les molécules d'ADN cible qui se trouvent dans l'échantillon à tester captent les molécules de sonde marquée qui portent une séquence complémentaire, ou partiellement complémentaire, pour former de vrais complexes marqués sonde-cible. De plus, quelques molécules de sonde marquée peuvent se fixer de manière non spécifique soit sur le support solide, soit à des molécules non cible. Mais, après hybridation, le support solide est lavé abondamment et le seul marquage pouvant rester et être identifié sur le support solide provient des seules vraies molécules de complexe sonde-cible. Pour se donner un maximum de chance de former des complexes corrects, la cible est normalement présente en grand excès par rapport à la sonde.

L'autre alternative – utilisation de sondes immobilisées et marquage des acides nucléiques de l'échantillon à tester – sera considérée plus loin dans ce chapitre, lorsque nous verrons l'hybridation sur micropuces.

La dénaturation et l'hybridation sont modifiées par la température, l'environnement chimique et la quantité de liaisons hydrogène

Quand des séquences d'acides nucléiques complémentaires ou partiellement complémentaires s'associent pour former des duplex, le nombre de nouvelles liaisons

TABLEAU 7.1 ÉQUATIONS PERMETTANT DE CALCULER LA T_m

Hybride	T_m (°C)
ADN-ADN	$81,5 + 16,6 (\log_{10}[\text{Na}^+]^a) + 0,41 (\%GC^b) - 500/L^c$
ADN-ARN ou ARN-ARN	$79,8 + 18,5(\log_{10}[\text{Na}^+]^a) + 0,58 (\%GC^b) + 11,8 (\%GC^b)^2 - 820/L^c$
Oligo-ADN ou oligo-ARNd	
Pour < 20 nucléotides	2 (l_n)
Pour 20-35 nucléotides	$22 + 1,46 (l_n)$

^a Ou pour tout cation monovalent, mais n'est précis que dans les limites de 0,01-0,4 M. ^b Seulement précis pour un % en GC compris entre 30 et 75 %. ^c L, Longueur du duplex en paires de bases. ^d Oligo, oligonucléotide ; l_n , longueur effective de l'amorce = 2 (nombre de G + C) + (nombre de A + T). Pour chaque 1 % de formamide, la T_m est abaissé d'environ 0,6 °C, et la présence d'urée 6M abaisse le T_m d'environ 30 °C.

hydrogène formées dépend de la longueur du duplex. Plus la molécule d'acide nucléique est longue, plus il y a de liaisons hydrogène et plus il faut d'énergie pour les rompre. L'augmentation de la stabilité des duplex n'est pas directement proportionnelle à la longueur, et c'est surtout sur les molécules courtes que les modifications de longueur ont des effets particulièrement visibles. Comme nous le verrons dans la section suivante, les mésappariements diminuent la stabilité des duplex.

La composition en bases et l'environnement sont aussi des facteurs importants de la stabilité des duplex. Un haut pourcentage de paires de bases G-C implique une plus grande difficulté à séparer les deux brins du duplex, car elles ont trois liaisons hydrogène alors que les paires A-T n'en ont que deux. La présence de cations monovalents (Na^+ , par exemple) stabilise les liaisons hydrogène des molécules double brin, alors que les molécules fortement polaires (comme l'urée et la formamide) perturbent et rompent les liaisons hydrogène et agissent donc comme des dénaturants chimiques.

Une augmentation progressive de la température rend, elle aussi, les liaisons hydrogène instables et les casse. La température qui correspond au point situé à mi-parcours lors de la transition de la forme double brin à la forme simple brin est appelée **température de fusion** (T_m). Pour les génomes de mammifères dont la composition en bases est souvent de 40 % de G-C, l'ADN se dénature à une T_m d'environ 87 °C dans des tampons dont le pH et la concentration saline avoisinent les conditions physiologiques. La T_m des hybrides parfaits formés d'ADN, d'ARN ou d'oligonucléotides, peut être déterminée selon les formules données dans le [Tableau 7.1](#).

Des conditions d'hybridation rigoureuses (astringentes) augmentent la spécificité de formation des duplex

Quand on effectue une hybridation d'acides nucléiques, les conditions dans lesquelles on se place sont habituellement prévues pour que la formation des duplex soit maximale, même au prix d'un peu d'hybridation non spécifique. Par exemple, la température d'hybridation peut être choisie à 25 °C en dessous de la T_m et, ainsi, les molécules de sonde s'apparient avec des molécules d'acides nucléiques aux séquences apparentées de façon lointaine, en plus des molécules cibles aux séquences très proches. Dans ce cas, la **stringence d'hybridation** est dite faible.

Après avoir encouragé la formation de duplex forts entre la sonde et la cible, des lavages successifs sont pratiqués dans des conditions de moins en moins tolérantes vis-à-vis d'éventuels mésappariements dans les hétéroduplex. Cela est obtenu en augmentant progressivement la température, ou en réduisant progressivement la concentration en NaCl dans le tampon. L'augmentation progressive de la stringence d'hybridation révèle des séquences cibles de plus en plus proches de la sonde. Le dernier lavage correspond aux conditions optimales d'hybridation permettant d'assurer la formation d'hétéroduplex spécifiques.

Les hétéroduplex sonde-cible sont beaucoup plus stables sur le plan thermodynamique lorsque la séquence hybridée est constituée de bases parfaitement appariées. Les mésappariements entre les deux brins d'un hétéroduplex diminuent la T_m : pour des sondes ADN normales, chaque 1 % de mésappariements réduit la T_m d'environ 1 °C. Cependant, cet effet diminue avec la taille de la région appariée. Ainsi un degré considérable de mésappariements peut être toléré si la totalité de la région des bases

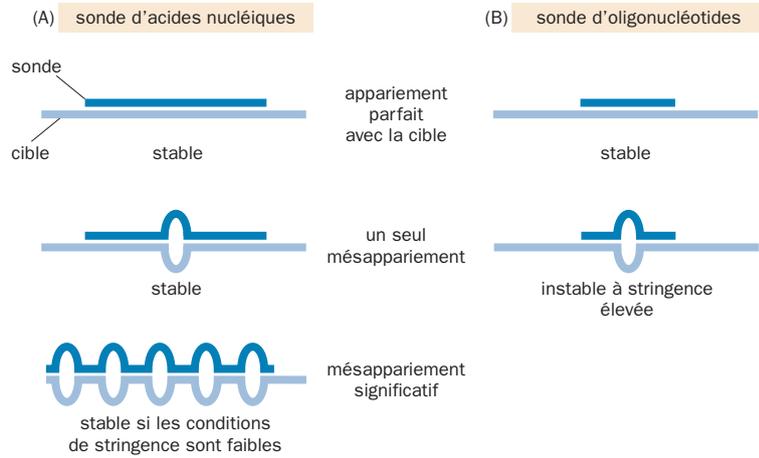


Figure 7.3 Utilisation des mésappariements sonde-cible au cours de l'hybridation des acides nucléiques. (A) Si la rigueur des conditions d'hybridation est diminuée (stringence plus faible), les sondes d'acides nucléiques plus longues (plus de 100 nucléotides) formeront des hétéroduplex stables avec des séquences cibles similaires, mais pas identiques à la sonde. Cela permet des analyses comparatives entre espèces et l'identification de membres lointainement apparentés de gènes d'une même famille. (B) Les sondes oligonucléotidiques courtes sont, en revanche, beaucoup moins tolérantes par rapport aux mésappariements de séquences. En utilisant une hybridation à stringence élevée, on peut sélectionner les seuls duplex parfaitement appariés, ce qui permet de distinguer des séquences d'allèles qui ne diffèrent que par un seul nucléotide.

complémentaires est longue (plus de 100 pb). Au contraire, si la région des bases complémentaires est courte, comme c'est le cas lorsque des oligonucléotides sont impliqués (15-20 nucléotides en général), on peut choisir les conditions d'hybridation telles qu'un seul mésappariement rende le complexe hybride instable (Figure 7.3).

La cinétique de réassociation de l'ADN dépend aussi de la concentration en ADN

La vitesse à laquelle des simples brins complémentaires se réassocient pour former un ADN double brin dépend de la concentration initiale de l'échantillon d'ADN. S'il y a une forte concentration en séquences d'ADN complémentaires, le temps nécessaire pour qu'une molécule d'ADN simple brin trouve le brin complémentaire et forme un duplex sera diminué. La cinétique de réassociation peut être calculée en utilisant la concentration de départ (C_0) de la séquence spécifique d'ADN en moles par litre et le temps de réaction (t) en secondes. Cependant, la valeur C_0t (généralement appelée valeur du Cot) varie aussi en fonction de la température de réassociation et de la concentration en cations monovalents. Par conséquent, on fixe en général des valeurs de référence : une température de réassociation de 65 °C et une concentration en NaCl de 0,3 M.

La fréquence des séquences cibles dans un échantillon test peut varier dans des proportions considérables. Quand une sonde homogène est hybridée dans un échantillon présentant une population hétérogène d'acides nucléiques, la concentration d'une séquence cible donnée peut être très basse, ce qui rend la vitesse de réassociation lente. Par exemple, si une sonde du gène de la β -globine est hybridée dans un échantillon d'ADN génomique humain, la séquence cible sera présente à une concentration très basse (le gène de la β -globine représente 0,00005 % de l'ADN génomique humain). Pour conduire la réaction d'hybridation, il faudra augmenter la quantité d'ADN cible, et donc plusieurs microgrammes d'ADN génomique humain seront nécessaires. Au contraire, des séquences hautement répétitives dans l'ADN génomique humain représenteront une plus grande proportion de l'ADN de départ et elles se réassocieront relativement plus vite.

Quand une sonde marquée est utilisée, la force du signal d'hybridation est aussi proportionnelle au nombre de copies de la séquence cible : les gènes présents en une seule copie donnent des signaux d'hybridation faibles, les séquences ADN hautement répétitives donnent au contraire des signaux intenses. Si une sonde particulière est hétérogène et contient une copie unique de la séquence d'intérêt, par exemple un gène spécifique, mélangée à de l'ADN hautement répétitif, le faible signal d'hybridation obtenu avec la séquence en faible nombre de copies sera complètement masqué par le signal d'hybridation fort venant de l'ADN répétitif. Cet effet peut cependant être surmonté par une étape de pré-hybridation, appelée *hybridation compétitive*. Cela est effectué en ajoutant à la sonde marquée une population d'ADN non marqué, enrichie en ADN répétitif. Le mélange d'acides nucléiques est alors dénaturé puis, on le laisse se réassocier ; les éléments répétitifs de la sonde marquée sont ainsi éliminés. Les éléments répétitifs marqués s'hybrident avec les séquences répétitives non marquées et seules les séquences non répétitives libres dans la sonde peuvent encore s'hybrider.

7.2 MARQUAGE DES ACIDES NUCLÉIQUES ET DES OLIGONUCLÉOTIDES

Dans l'hybridation des acides nucléiques, une population d'acides nucléiques ou d'oligonucléotides est marquée, puis hybridée aux séquences complémentaires d'une autre population. Dans cette section, nous étudierons surtout les principes et la méthodologie du marquage. Nous les illustrerons en nous référant au marquage des populations de sondes, qui sont souvent des populations homogènes d'ADN, d'ARN ou des séquences oligonucléotidiques. Les réactions d'hybridation dans lesquelles l'échantillon test d'acides nucléiques est marqué, impliquent généralement le marquage de populations hétérogènes et souvent complexes d'ADN ou d'ARN ; ces réactions seront étudiées dans la section 7.4.

Différentes classes de sondes d'hybridation peuvent être préparées avec pour substrat de l'ADN, de l'ARN ou des oligonucléotides

Lorsque les techniques d'hybridation utilisent des sondes marquées, le substrat servant de marqueur peut être de l'ADN, de l'ARN ou un oligonucléotide synthétique ; ils sont obtenus par des méthodes différentes. Selon les besoins, la sonde marquée doit être dénaturée par chauffage afin de séparer les deux brins ou pour casser les liaisons hydrogènes internes.

Les sondes d'ADN conventionnelles sont habituellement isolées par les techniques de clonage cellulaire de l'ADN ou par amplification de l'ADN par PCR. Dans les deux cas, les sondes sont habituellement double brin à l'origine. L'ADN cloné dans des cellules peut avoir des tailles variant de 0,1 kb à des centaines de kilobases ; en revanche, l'ADN obtenu par PCR fait souvent moins de quelques kilobases de long. Les sondes sont souvent marquées en y incorporant des désoxyribonucléotides triphosphates marqués (dNTP) au cours d'une réaction de synthèse d'ADN *in vitro*. Comme elles proviennent en général d'un ADN double brin, les sondes ADN simple brin représentent un mélange de séquences des deux brins et elles ne sont donc pas idéales pour suivre des transcrits ARN spécifiques (car elles identifient simultanément les transcrits des brins ADN sens et antisens).

Les sondes d'ARN sont obtenues à partir de molécules d'ARN simple brin, longues en général de quelques centaines de bases à plusieurs kilobases. L'ARN est préparé à partir d'ADN cloné dans un plasmide spécialisé, un vecteur d'expression (*voir* Chapitre 6), qui comporte une séquence promotrice de phage. À partir de ce promoteur sélectif, l'ARN polymérase du phage transcrit l'ADN inséré et en fait des copies ARN. Comme la synthèse d'ARN est faite en présence des quatre ribonucléotides triphosphate (rNTP), dont l'un au moins est marqué, des transcrits ARN marqués sont produits à partir de l'insert ADN cloné. Les sondes antisens simple brin sont utiles pour identifier des transcrits sens complémentaires particuliers.

Les sondes oligonucléotidiques sont simples brin et très courtes (typiquement 15 à 50 nucléotides). Elles sont produites par synthèse chimique plutôt que par clonage de l'ADN. Leur synthèse implique l'addition successive de nucléotides un par un à un nucléotide de départ qui a été fixé sur un support solide. Elles sont marquées par addition d'un groupement marqué à leur extrémité 5'. Les sondes oligonucléotidiques permettent de distinguer deux allèles qui ne diffèrent que par un seul nucléotide (*voir* Figure 7.3).

Les longues sondes d'acides nucléiques sont habituellement marquées au cours de la synthèse du brin en y incorporant des nucléotides marqués

Pour certains usages, les acides nucléiques sont marqués en leur attachant des groupements marqués aux extrémités de la molécule. Les oligonucléotides simple brin sont marqués en général grâce à une kinase qui ajoute un phosphate marqué à leur extrémité 5'. De longs fragments d'ADN peuvent aussi être marqués à leur extrémité en utilisant d'autres méthodes comme des amorces modifiées portant un groupement marqué à leur extrémité 5'. Au cours de la PCR, l'amorce ayant l'extrémité 5' marquée est incorporée dans le produit de la PCR.

Cependant, avec le marquage des seules extrémités, seul un très petit nombre de groupements marqués peut être inséré dans la molécule. Les groupements marqués ne représentent donc souvent qu'un très petit pourcentage de la masse totale de la sonde, surtout pour les longues séquences de nucléotides. La méthode de choix pour marquer l'ADN et l'ARN repose donc sur le marquage au cours de la synthèse du brin : un nouveau brin d'ADN ou d'ARN est synthétisé *in vitro* et marqué par incorporation, à différents

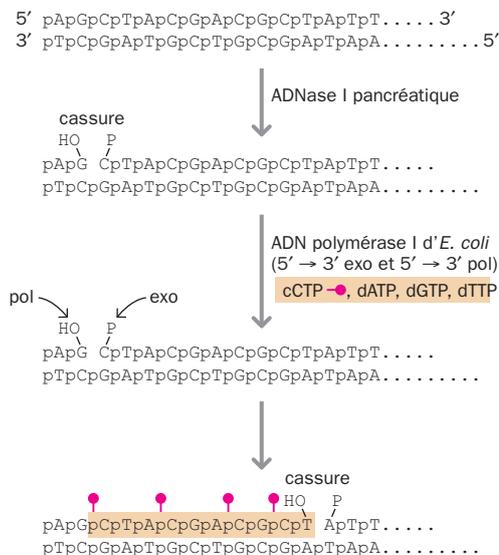


Figure 7.4 Marquage de l'ADN par nick translation. L'ADNase I pancréatique introduit des cassures dans les ADN simple brin en coupant la liaison phosphodiester (p) pour libérer un hydroxyl terminal en 3' et un groupement phosphate en 5'. Ces cassures sont des substrats pour l'ADN polymérase I d'*E. coli* qui comporte de nombreuses sous-unités. Cette protéine a deux activités enzymatiques : une activité exonucléase 5' → 3' (exo) qui attaque les extrémités 5' terminales exposées d'une cassure et enlève séquentiellement les nucléotides dans la direction 5' → 3', et une activité d'ADN polymérase (pol) qui ajoute de nouveaux nucléotides aux extrémités 3' OH libres, dans la direction 5' → 3', remplaçant ainsi les nucléotides qui ont été enlevés par l'exonucléase et causant donc un déplacement latéral de la cassure. Au moins un des quatre désoxyribonucléotides triphosphate (dNTP) est marqué (ici, cercle rouge de dCTP) et un marqueur est donc incorporé chaque fois qu'un CTP est introduit dans le nouveau brin d'ADN en cours de synthèse.

endroits tout le long du brin nouvellement synthétisé, de nucléotides marqués. Une ADN ou une ARN polymérase utilise généralement un ADN cloné approprié pour faire des copies ADN ou ARN, qui sont marquées en utilisant un mélange dans lequel au moins un des quatre dNTP ou rNTP est marqué.

Marquage de l'ADN par *nick translation*

La technique de *nick translation* (déplacement de cassure) repose sur la réparation de cassures (*nicks*) simple brin dans l'ADN, formant des extrémités 3' OH et 5' P terminales exposées. Cette cassure peut être obtenue en utilisant une endonucléase appropriée, comme la désoxyribonucléase I du pancréas (ADNase I). La cassure ainsi exposée va alors servir de point de départ pour l'introduction de nouveaux nucléotides en utilisant l'ADN polymérase d'*E. coli*, une enzyme composée de nombreuses sous-unités qui possède des activités d'ADN polymérase et d'exonucléase 5' → 3'. Alors que l'ADN polymérase ajoute de nouveaux nucléotides à l'extrémité 3'-hydroxyle de la cassure, les nucléotides sont supprimés à l'autre extrémité de la cassure par l'activité exonucléasique 5' → 3' de la même enzyme. La cassure est ainsi progressivement déplacée le long de l'ADN dans la direction 5' → 3' (Figure 7.4). Si la réaction est réalisée à une température relativement faible (environ 15 °C), elle ne se poursuit par au-delà du renouvellement complet de la séquence nucléotidique existante. Bien qu'il n'y ait pas de synthèse nette d'ADN à cette température, la réaction de synthèse permet l'incorporation de nucléotides marqués à la place de ceux qui existaient précédemment dans le nucléotide non marqué.

Marquage de l'ADN à l'aide d'amorces aléatoires (*Random primed DNA labelling*)

La technique de marquage de l'ADN à l'aide d'amorces aléatoires repose sur l'hybridation d'un mélange de nombreux hexanucléotides différents qui se lient au hasard à des séquences complémentaires de l'ADN matrice dénaturé, et initient la synthèse de nouveaux brins d'ADN (Figure 7.5). La synthèse des nouveaux brins complémentaires est catalysée par la sous-unité de Klenow de l'ADN polymérase I d'*E. coli* (qui contient l'activité de polymérase sans l'activité exonucléasique 5' → 3' associée).

Marquage au cours de la synthèse de brins par PCR

La réaction classique de PCR peut être facilement modifiée pour inclure un ou plusieurs précurseurs nucléotidiques marqués qui s'incorporeront dans les produits de PCR sur toute leur longueur.

Marquage des ARN

Des sondes à ARN (*ribosondes*) peuvent être obtenues par transcription *in vitro* d'ADN cloné dans un vecteur d'expression plasmidique approprié. Ce vecteur contient une séquence promotrice immédiatement adjacente au site d'insertion de l'ADN étranger, assurant ainsi la transcription de la séquence insérée. Des promoteurs de phages très efficaces sont utilisés communément, comme ceux des phages SP6, T3 et T7 et la polymérase correspondant au phage est fournie afin d'assurer une transcription spécifique de l'ADN cloné. L'incorporation de ribonucléotides marqués assure un marquage à l'ARN nouvellement synthétisé (Figure 7.6).

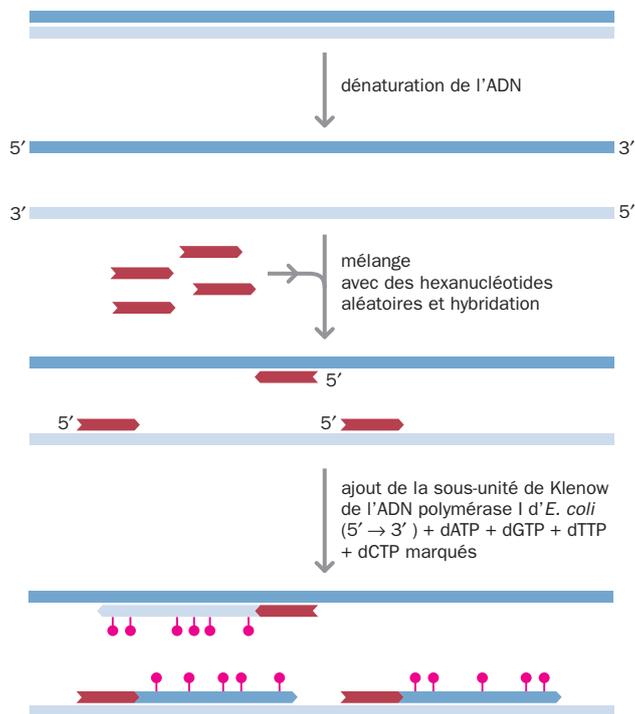


Figure 7.5 Marquage de l'ADN par amorce aléatoire. On dénature de l'ADN double brin puis on y ajoute un mélange d'hexanucléotides synthétisés de façon presque aléatoire. Les hexanucléotides se lieront à des sites complémentaires de l'ADN simple brin. Ils pourront alors servir d'amorce pour la synthèse d'un nouveau brin complémentaire marqué. L'enzyme de Klenow est une sous-unité de l'ADN polymérase I d'*E. coli* qui est utilisée dans ce type de réaction car elle possède l'activité ADN polymérase $5' \rightarrow 3'$, mais pas celle d'exonucléase.

Les isotopes radioactifs peuvent être utilisés pour marquer les acides nucléiques, mais ils sont fugaces et peuvent être dangereux

Traditionnellement, les acides nucléiques étaient marqués par incorporation de nucléotides portant un isotope radioactif qui pouvait être détecté en solution ou mieux encore dans un support solide (**autoradiographie** ; Encadré 7.2). Une autoradiographie donne une représentation bidimensionnelle de la distribution du marqueur radioactif dans

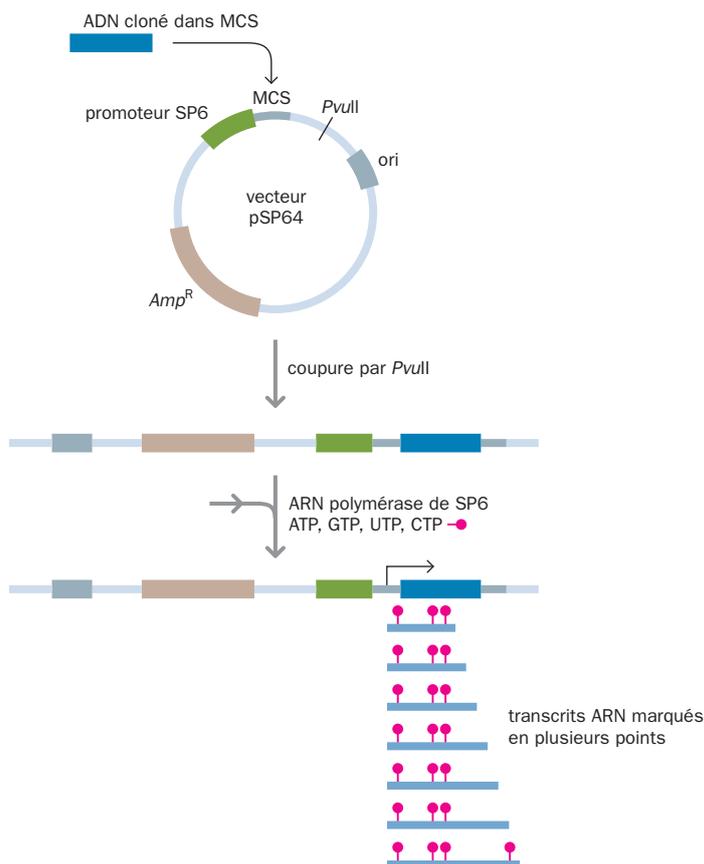


Figure 7.6 Les sondes ARN sont habituellement produites par transcription d'inserts ADN clonés grâce à une ARN polymérase de phage. Le plasmide vecteur pSP64 contient une séquence promotrice pour l'ARN polymérase du phage SP6 liée à un site de clonage multiple (MCS), à une origine de répllication (ori) et à un gène de résistance à l'ampicilline (*Amp^R*). Après clonage du fragment d'ADN approprié dans le site MCS du plasmide vecteur, l'ADN recombinant purifié est linéarisé par coupure avec l'enzyme de restriction *PvuII*. L'ARN polymérase de SP6 et un mélange de NTP, dont au moins un est marqué, sont alors ajoutés, ce qui permet l'initiation de la transcription à partir d'un site particulier du promoteur de SP6 ; elle se poursuit en englobant aussi l'ADN inséré. Des transcrits ARN de l'insert ADN, marqués plusieurs fois, sont alors produits. De nombreux vecteurs d'expression similaires possédant deux types différents de promoteurs de phages situés de chaque côté du système à clonage multiple (MCS), comme ceux des phages T3 et T7, et les polymérases des mêmes phages, sont aussi utilisés pour créer des ribosondes à partir de l'un ou l'autre des deux brins d'ADN.

ENCADRÉ 7.2 AUTORADIOGRAPHIE

L'autoradiographie est une technique permettant de localiser et de visualiser un composé marqué par un isotope radioactif dans un échantillon solide par production d'une image dans une émulsion photographique. Pour les applications de génétique moléculaire, les composés marqués sont très souvent des molécules d'ADN ou de protéines et l'échantillon solide peut être de la chromatine fixée ou un échantillon de tissu monté sur une lame de verre. Les échantillons d'ADN ou de protéines peuvent aussi être inclus dans un gel d'électrophorèse séché ou fixé à la surface d'une membrane de nylon ou de nitrocellulose.

L'échantillon solide est placé au contact immédiat d'un film à rayons X, feuille de plastique recouverte d'une émulsion photographique. L'émulsion photographique est composée de cristaux d'halogénure d'argent en suspension dans une phase gélatineuse transparente. L'émission radioactive provenant de l'échantillon traverse l'émulsion et transforme les ions Ag^+ en atomes d'Ag. La position des cristaux d'halogénure d'argent altérés est révélée lors du développement, une technique d'amplification dans laquelle le reste des ions Ag^+ provenant des cristaux altérés

est réduit pour donner de l'argent métallique. Le procédé de fixation permet d'enlever tous les cristaux d'halogénure d'argent non exposés. Les régions sombres sur le film photographique donnent une représentation en deux dimensions du marqueur radioactif dans l'échantillon original.

L'autoradiographie directe est mieux adaptée à la détection des radio-isotopes à émission β faible ou moyenne (par exemple, ^3H et ^{35}S). Cependant, les particules β à haute énergie (par exemple, ^{32}P) passeront à travers le film en perdant la majorité de leur énergie (Tableau 1). Pour les échantillons qui émettent des radiations de forte énergie, une modification est nécessaire. Elle consiste en la conversion en lumière de l'énergie émise grâce à un agent chimique adéquat (scintillateur ou fluor). L'autoradiographie indirecte utilise des plaques composées d'un scintillateur inorganique solide comme écran intensifiant qui sont placées derrière le film photographique. Les émissions qui passent à travers l'émulsion photographique sont absorbées par l'écran et converties en lumière qui réduit aussi les ions Ag^+ et donc intensifie l'image autoradiographique directe.

TABLEAU 1 CARACTÉRISTIQUES DES RADIO-ISOTOPES UTILISÉS EN GÉNÉRAL POUR MARQUER LES SONDAS ADN ET ARN

Isotope	Demi-vie	Type d'émission	Énergie émise (MeV)	Temps d'exposition	Adéquation pour des études à haute résolution
^3H	12,4 années	β^-	0,019	Très long	Excellente
^{32}P	14,3 jours	β^-	1,710	Court	Mauvaise
^{33}P	25,5 jours	β^-	0,248	Moyen	Intermédiaire
^{35}S	87,4 jours	β^-	0,167	Moyen	Intermédiaire

l'échantillon original. L'intensité du signal en autoradiographie dépend de l'énergie de radiation émise et du temps d'exposition.

L'isotope radioactif ^{32}P a été très utilisé dans les analyses par hybridation des acides nucléiques car il émet des particules β à haute énergie pouvant être détectées très facilement. Cependant, ces particules sont si pénétrantes que le signal diffuse, ce qui est un inconvénient lorsqu'une résolution physique fine est nécessaire. D'autres isotopes, comme le ^{35}S , sont plus utiles pour l'étude de l'expression des gènes dans les cellules et tissus pour laquelle une résolution morphologique est requise.

Les isotopes radioactifs sont facilement détectés, mais ils constituent aussi un danger pour la santé et la radioactivité décroît avec le temps, rendant nécessaire la synthèse de nouvelles sondes avant toute expérience. Ainsi, les marqueurs non radioactifs contenant des groupements chimiques distinctifs, qui sont à la fois stables et détectés efficacement, sont maintenant employés en routine.

Des fluorophores sont souvent utilisés pour le marquage non isotopique des acides nucléiques

Le marquage non radioactif des acides nucléiques implique l'incorporation de nucléotides portant un groupement chimique ou une molécule pouvant être détecté facilement et spécifiquement. Le groupement incorporé peut servir lui-même de marqueur et être détecté par une mesure directe dans le dosage. On utilise souvent un **fluorophore**, groupement chimique qui peut facilement être détecté car il absorbe l'énergie d'une certaine longueur d'onde qui lui est spécifique (**longueur d'onde d'excitation**) et la réémet à une longueur d'onde plus longue (**longueur d'onde d'émission**), mais également spécifique (Encadré 7.3).

Alternativement, une mesure indirecte peut être utilisée. Un groupement chimique incorporé sert de **rapporteur** spécifiquement reconnu et lié par une molécule d'affinité, comme un anticorps dédié. La molécule d'affinité a un groupement **marqueur** qui lui est lié, un groupement chimique ou une molécule, qui peut alors être facilement dosé (Figure 7.7).

ENCADRÉ 7.3 MARQUAGE FLUORESCENT DES ACIDES NUCLÉIQUES

Le marquage fluorescent des acides nucléiques a été développé dans les années 1980 et s'est montré extrêmement précieux pour de nombreuses applications, y compris l'hybridation chromosomique *in situ*, l'hybridation tissulaire *in situ* et le séquençage automatique de l'ADN.

Un **fluorophore** est un groupement chimique qui absorbe l'énergie lumineuse lorsqu'il est exposé à une lumière d'une certaine longueur d'onde (excitation) et la réémet à une longueur spécifique mais plus longue (Tableau 1 et Figure 1). Le marquage direct des acides nucléiques par des fluorophores est réalisé par incorporation d'un nucléotide modifié (souvent la 2' désoxyuridine 5' triphosphate) contenant un fluorophore approprié.

Des systèmes de marquage indirect sont aussi utilisés. Dans cette technique, le fluorophore est utilisé en tant que marqueur et attaché à une molécule d'affinité (comme un anticorps spécifique de la streptavidine ou de la digoxigénine) qui se lie de façon spécifique à un nucléotide modifié contenant un radical rapporteur (comme la biotine ou la digoxigénine) (voir Figure 7.7).

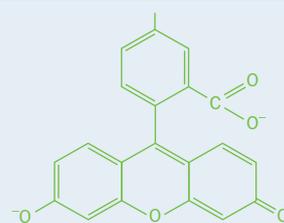
TABEAU 1 FLUOROPHORES POUR LE MARQUAGE DES ACIDES NUCLÉIQUES

Fluorophore	Longueurs d'onde maximales (nm)	
	Excitation	Émission
Bleu		
AMCA	350	450
DAPI	358	461
Vert		
FITC	492	520
Fluorescéine (voir Figure 1)	494	523
Rouge		
CY3	550	570
TRITC	554	575
Rhodamine (Voir Figure 1)	570	590
Rouge Texas	596	620
CY5	650	670

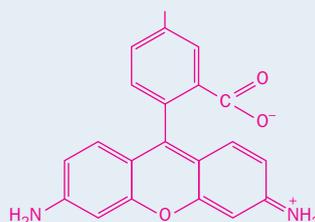
AMCA, aminométhylcoumarine ; DAPI, 4', 6-diamino-2-phénylindole ; FITC, fluorescéine isothiocyanate ; CY3, indocarbocyanine ; TRIT, tétraméthylrhodamine isothiocyanate ; CY5, indodicarbocyanine.

Détection des acides nucléiques marqués par un fluorophore

Les acides nucléiques marqués par un fluorophore peuvent être détectés par scanner laser ou microscopie de fluorescence. Les fluorophores sont détectés en faisant passer un rayon de lumière provenant d'une source appropriée (un laser à l'argon est utilisé dans le séquençage automatique de l'ADN, une lampe à vapeurs de mercure est utilisée en microscopie de fluorescence) à travers un filtre de couleur appropriée. Le filtre étant conçu pour transmettre la lumière à la longueur d'onde d'excitation désirée. Dans les systèmes de microscopie de fluorescence, la lumière est réfléchi sur l'échantillon marqué par fluorescence, posé sur le porte-objet du microscope, en utilisant un miroir dichroïque qui réfléchit la lumière de certaines longueurs d'onde, alors que la lumière des autres longueurs d'onde passe directement à travers (Figure 2). La lumière excite alors le fluorophore qui devient fluorescent ; il émet alors de la lumière à une longueur d'onde légèrement plus longue,



fluorescéine



rhodamine

Figure 1 Structure de deux fluorophores communs. TRITC et divers autres fluorophores dérivent de la rhodamine.

la longueur d'onde d'émission. La lumière émise par le fluorophore est renvoyée directement à travers un miroir dichroïque, puis un filtre approprié, dans l'oculaire du microscope. Un second dispositif séparateur de faisceau peut aussi permettre d'enregistrer la lumière à l'aide d'une caméra couplée à l'appareil.

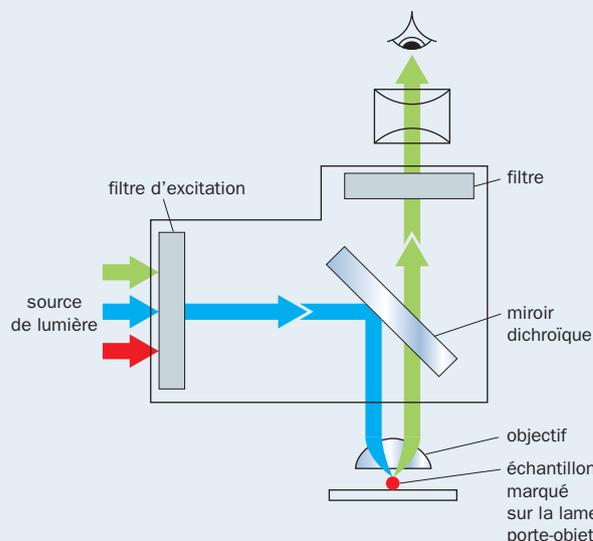


Figure 2 Microscopie à fluorescence. Le filtre d'excitation permet à la seule lumière de longueur d'onde appropriée de passer (par exemple ici, la bleue). La lumière bleue transmise est réfléchi par le miroir dichroïque (qui sépare les rayons lumineux) sur l'échantillon marqué, qui devient alors fluorescent et émet une lumière à une longueur d'onde légèrement plus grande (lumière verte, sur la figure). La lumière verte émise passe alors directement à travers le miroir dichroïque, puis à travers un autre filtre qui bloque les signaux fluorescents non désirés, laissant la seule émission de fluorescence verte désirée passer à travers l'oculaire du microscope.

Deux techniques sont largement utilisées pour détecter un marquage indirect. Le système à biotine-streptavidine repose sur la très grande affinité entre deux ligands naturels. La biotine (une vitamine) agit en tant que rapporteur car elle est spécifiquement liée par la streptavidine, une protéine bactérienne, avec une constante d'affinité (aussi appelé constante de dissociation) de 10^{-14} , une des plus élevées qui soit connue en biologie. Les sondes biotinylées peuvent être facilement

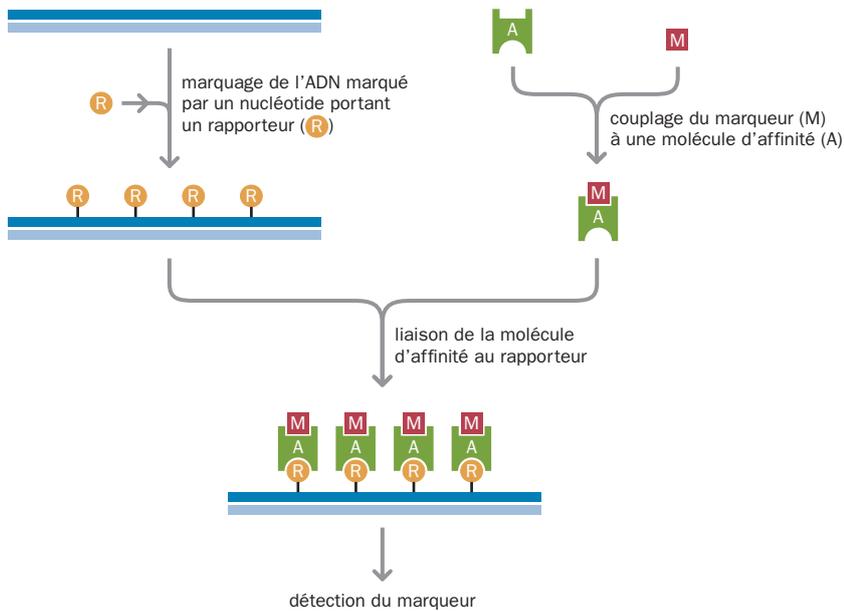


Figure 7.7 Détection indirecte de groupements marqués dans les acides nucléiques. Les acides nucléiques peuvent être marqués par des groupements chimiques qui ne sont pas détectés directement. Les groupements incorporés servent plutôt de groupements rapporteurs car ils sont liés avec une grande spécificité par une molécule d'affinité qui, elle, porte un marqueur détectable. Ce marqueur peut être détecté de différentes façons. Si c'est un colorant fluorescent, il pourra être détecté par microscopie de fluorescence. Fréquemment, l'alternative est d'utiliser une enzyme, comme la phosphatase alcaline, qui convertit un substrat pour donner un produit coloré qui peut être dosé par colorimétrie.

obtenues en incluant un nucléotide adéquat biotinylé au cours de la réaction de marquage (Figure 7.8). La streptavidine sert alors de molécule d'affinité révélatrice. Un autre rapporteur largement utilisé est la digoxigénine, un stéroïde obtenu à partir de la plante *Digitalis* (voir Figure 7.8). Un anticorps spécifique contre la digoxigénine sert de marqueur d'affinité.

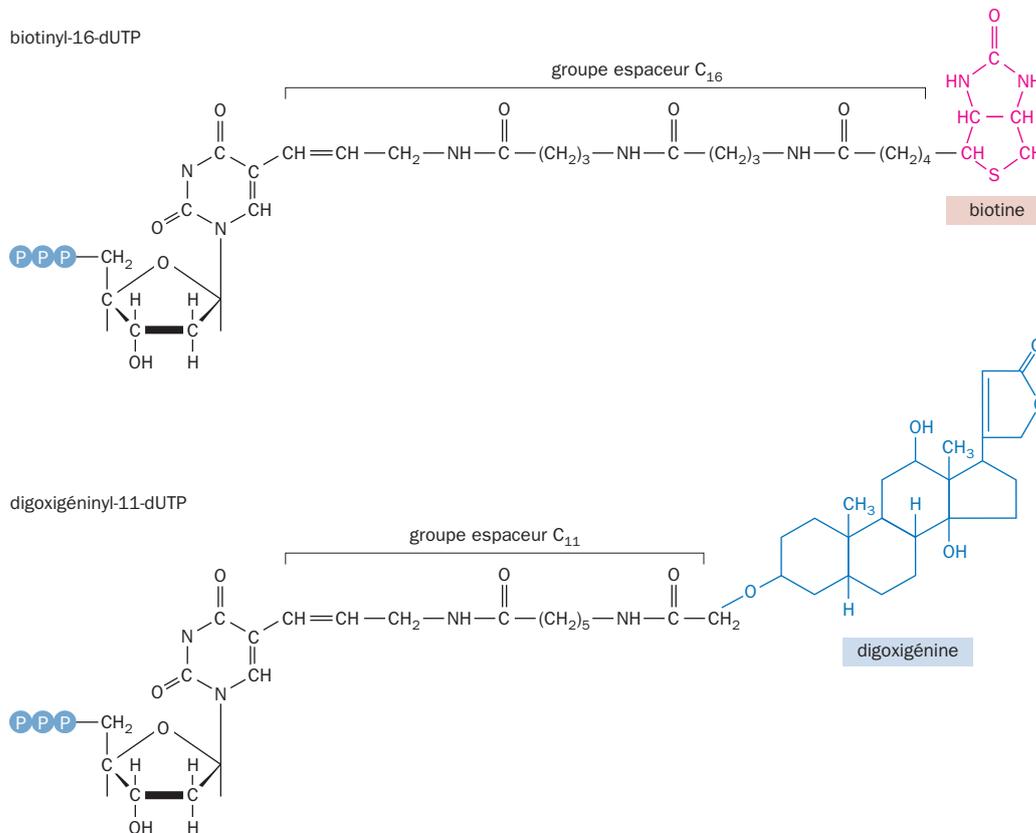


Figure 7.8 Structure des nucléotides modifiés par la biotine et la digoxigénine. La biotine et la digoxigénine sont des groupements rapporteurs qui sont liés ici à l'atome de carbone 5' de l'uridine du dUTP par des groupes espaceurs formés respectivement, d'un total de 16 atomes de carbone (biotinyl-16-dUTP) ou de 11 atomes de carbone (digoxigéninyl-11-dUTP). Ces groupes d'espaceur sont nécessaires pour assurer une séparation physique du groupe rapporteur par rapport au squelette des acides nucléiques afin que le groupement rapporteur soit exposé assez loin du squelette d'acide nucléique pour permettre aux molécules d'affinité de s'y lier.

GÉNÉTIQUE MOLÉCULAIRE HUMAINE

Tom STRACHAN et Andrew READ

4^e édition



Depuis la publication, en 2004, de la séquence complète du génome humain grâce au Projet Génome Humain, de grandes banques de séquences d'ADN sont produites chaque année. De puissants programmes bioinformatiques sont capables de comparer les génomes d'organismes différents. La génétique comparative permet ainsi de comprendre l'évolution du génome humain et celui de nombreux autres organismes modèles, ouvrant la voie à la recherche biomédicale.

De grands projets internationaux ont été mis en place afin de comprendre le fonctionnement de nos gènes : le projet ENCODE vise à créer une encyclopédie des éléments fonctionnels de notre ADN ; le projet HapMap étudie les variations génétiques dans la population mondiale, permettant aux chercheurs d'identifier les gènes impliqués dans les maladies. Enfin le développement constant de nouvelles technologies telles que le génotypage à haut débit des SNP ou les micropuces à ADN offre des outils puissants et hautement performants.

La 4^e édition de **Génétique moléculaire humaine** intègre ces avancées scientifiques et technologiques. Véritable pont entre les manuels de base et la littérature issue de la recherche, l'ouvrage fournit en 21 chapitres le cadre permettant de comprendre ce domaine, depuis les aspects fondamentaux jusqu'aux notions les plus pointues de biologie moléculaire.

Après des chapitres généraux sur la structure et la fonction des acides nucléiques, des chromosomes, la répartition des gènes dans la population, il présente les techniques d'étude et d'analyse des gènes.

Des chapitres très développés et à la pointe des connaissances actuelles portent sur l'épigénétique, les ARN non codants, les cellules souches.

Les principaux organismes modèles utilisés dans les études génétiques et l'étude des maladies humaines sont décrits en détail.

Les traitements des maladies humaines – et les questions d'éthique qui en découlent – sont abordés à travers les tests et analyses génétiques, l'utilisation des cellules souches, la thérapie cellulaire, génique, oligonucléotidique et la médecine personnalisée.

Enfin, un glossaire de plus de 600 définitions et un index de près de 10 000 entrées complètent ce livre exhaustif, actuel et richement illustré.

L'ouvrage intéresse les étudiants en médecine, en biologie, en sciences et leurs enseignants ; les généticiens, les biologistes, les pédiatres ainsi que les personnes désireuses d'appréhender de manière scientifique la génétique moléculaire humaine et ses applications médicales.

www.medecine.lavoisier.fr



9 782257 204196