

Séquençage automatisé de l'ADN

Le séquençage de l'ADN à grande échelle nécessite des procédures automatisées basées sur le marquage par fluorescence de l'ADN et des systèmes de détection appropriés. En général, un marqueur fluorescent peut être utilisé directement ou indirectement. Les marqueurs fluorescents directs, tels qu'utilisés dans le séquençage automatisé, sont des fluorophores. Ce sont des molécules qui émettent une couleur fluorescente distincte lorsqu'elles sont exposées à la lumière UV d'une longueur d'onde spécifique. Des exemples de fluorophores utilisés dans le séquençage sont la fluorescéine, qui émet une fluorescence verte pâle lorsqu'elle est exposée à une longueur d'onde de 494 nm; la rhodamine, qui émet une fluorescence rouge à 555 nm; et l'acide aminométhyl cumarin acétique, qui émet une fluorescence bleue à 399 nm. De plus, une combinaison de différents fluorophores peut être utilisée pour produire une quatrième couleur. Ainsi, chacune des quatre bases peut être marquée distinctement.

Une autre approche consiste à utiliser des produits amplifiés par PCR (séquençage thermique). Ceci présente l'avantage que l'on peut utiliser comme matériau de départ de l'ADN double brin plutôt que simple brin. Et puisque de petites quantités d'ADN matrice sont suffisantes, l'ADN à séquencer n'a pas besoin d'être cloné au préalable.

Séquençage du cycle thermique

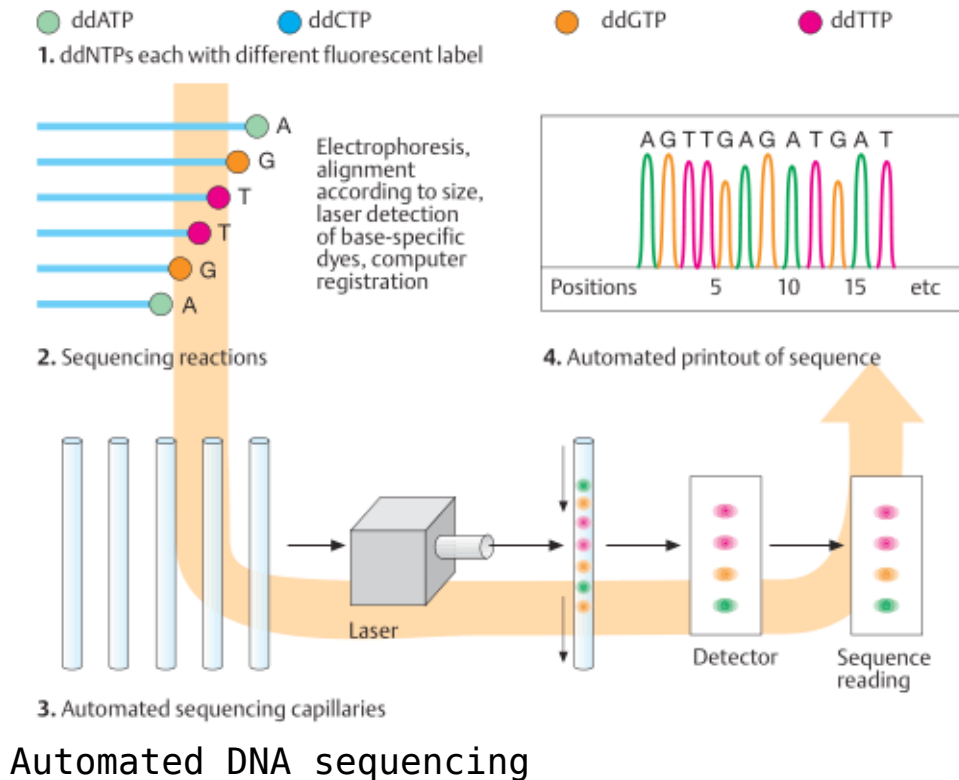
L'ADN à séquencer est contenu dans l'ADN vecteur <fn> Brown, T.A. Genomes. Bios Scientific Publ., Oxford, 1999. </fn>. L'amorce, un oligonucléotide court avec une séquence complémentaire du site de fixation sur l'ADN monocaténaire, est utilisée comme point de départ. Pour le séquençage de

brins courts d'ADN, une amorce universelle est suffisante. Il s'agit d'un oligonucléotide qui se liera à l'ADN du vecteur adjacent à l'ADN à séquencer. Cependant, si ce dernier est plus long qu'environ 750 pb, seule une partie de celui-ci sera séquencée. Par conséquent, des amorces internes supplémentaires sont requises. Ceux-ci s'hybrident à des sites différents et amplifient l'ADN dans une série d'expériences contiguës de chevauchement de chaînes qui se chevauchent. <fn> Rosenthal, N .: Structure fine d'un séquençage gène-ADN. Nouveau Eng. J. Med. 332: 589-591, 1995 </fn>. Ici, chaque amorce détermine quelle région de l'ADN matrice est en cours de séquençage. Dans le séquençage du cycle thermique <fn> Strachan, T., Read, A.P .: Human Molecular Genetics. 2 e éd. Bios Scientific Publishers, Oxford, 1999. </fn>, une seule amorce est utilisée pour effectuer des réactions de PCR, chacune avec un didésoxynucléotide (ddA, ddT, ddG ou ddC) dans le mélange réactionnel. Cela génère une série de différents brins terminés par une chaîne, chacun dépendant de la position de la base nucléotidique particulière où la chaîne est terminée <fn> Wilson, R.K., et al.: Développement d'une procédure automatisée pour le séquençage d'ADN fluorescent. Genomics 6: 626-636, 1990. </fn>. Après de nombreux cycles et avec l'électrophorèse, la séquence peut être lue comme indiqué dans la plaque précédente. Un avantage du séquençage du cycle thermique est que l'ADN double brin peut être utilisé comme matière de départ.

Séquençage automatisé de l'ADN (principe)

Le séquençage automatisé de l'ADN implique quatre fluorophores, un pour chacune des quatre bases nucléotidiques. Le signal fluorescent résultant est enregistré à un point fixe lorsque l'ADN traverse un capillaire contenant un gel électrophorétique. Les marqueurs fluorescents spécifiques de la base sont attachés aux didésoxynucléotides triphosphates

appropriés (ddNTP). Chaque ddNTP est marqué avec une couleur différente, par exemple, vert ddATP, bleu ddCTP, jaune ddGTP et ddTTP rouge <fn> Brown, T.A. Genomes. Bios Scientific Publ., Oxford, 1999. </fn>. (Les couleurs réelles pour chaque nucléotide peuvent être différentes.) Toutes les chaînes terminées à l'adénine (A) donneront un signal vert; toutes les chaînes terminées à une cytosine (C) donneront un signal bleu, et ainsi de suite. Les réactions de séquençage basées sur ce type de terminaison de la chaîne à nucléotides marqués <n Rosenthal, N .: Structure fine d'un séquençage d'ADN-gène. Nouveau Eng. J. Med. 332: 589-591, 1995 </fn> sont réalisées automatiquement dans des capillaires de séquençage <fn> Strachan, T., Read, A.P .: Human Molecular Genetics. 2 e éd. Bios Scientific Publishers, Oxford, 1999. </fn>. La migration électrophorétique des chaînes marquées ddNTP dans le gel dans le capillaire passe devant un faisceau laser focalisé sur une position fixe. Le laser induit un signal fluorescent qui dépend du marqueur spécifique représentant l'un des quatre nucléotides. La séquence est lue électroniquement et enregistrée et est visualisée sous forme de pics alternés dans l'une des quatre couleurs, représentant les nucléotides alternés dans leurs positions de séquence. En pratique, les pics ne montrent pas nécessairement la même intensité maximale que dans le diagramme schématique montré ici. (Illustration basée sur celle de Brown, 1999, et Strachan et Read, 1999).



Séquençage du Génome

La connaissance de la séquence nucléotidique d'un gène fournit des informations importantes sur sa structure, sa fonction et sa relation évolutive avec d'autres gènes similaires dans le même organisme ou dans des organismes différents. Ainsi, le développement dans les années 1970 de méthodes relativement simples de séquençage de l'ADN a eu un grand impact sur la génétique. Deux méthodes de base pour le séquençage de l'ADN ont été développées: une méthode de clivage chimique (A. M. Maxam et W. Gilbert, 1977) et une méthode enzymatique (F. Sanger, 1981). Un bref aperçu des principes sous-jacents suit.

Séquençage par dégradation chimique

Cette méthode utilise le clivage de l'ADN spécifique de la

base par certains produits chimiques. Quatre produits chimiques différents sont utilisés dans quatre réactions, une pour chaque base. Chaque réaction produit un ensemble de fragments d'ADN de différentes tailles. Les tailles des fragments dans un mélange réactionnel sont déterminées par des positions dans l'ADN du nucléotide qui a été clivé. Un fragment d'ADN à séquencer à double brin ou à simple brin est traité pour obtenir un seul brin marqué avec un isotope radioactif à l'extrémité 5' <fn>Brown, T.A. Genomes. Bios Scientific Publ., Oxford, 1999. </fn>. Ce brin d'ADN est traité avec l'un des quatre produits chimiques pour l'une des quatre réactions. Ici, la réaction sur les sites de guanine (G) par du diméthylsulfate (DMS) est montrée. Le sulfate de diméthyle attache un groupe méthyle au cycle purine des nucléotides G. La quantité de DMS utilisée est limitée de sorte qu'en moyenne, un seul nucléotide G par brin soit méthylé, pas les autres (représentés ici dans quatre positions différentes de G). Lorsqu'un second produit chimique, la pipéridine, est ajouté, le cycle purine nucléotidique est éliminé et la molécule d'ADN est clivée au niveau de la liaison phosphodiester juste en amont du site sans la base. La procédure globale résulte en un ensemble de fragments marqués de tailles définies en fonction des positions de G dans l'échantillon d'ADN en cours de séquençage. Des réactions similaires sont effectuées pour les trois autres bases (A, T et C non représentées). Les quatre mélanges réactionnels, un pour chacune des bases, sont placés dans des voies séparées d'une électrophorèse sur gel de Polyacrylamide. Chacune des quatre voies représente l'une des quatre bases G, A, T ou C. Le plus petit fragment migre le plus vers le bas, le suivant un peu moins loin, etc. On peut alors lire la séquence dans le sens opposé à la migration pour obtenir la séquence dans la direction 5' vers 3' (ici TAGTCGCAGTACCGTA).

Séquençage par terminaison de chaîne (séquençage Sanger)

Cette méthode, maintenant beaucoup plus utilisée que la méthode de clivage chimique, repose sur le principe que la synthèse de l'ADN est terminée quand au lieu d'un désoxynucléotide normal (dATP, dTTP, dGTP, dCTP), un didésoxynucléotide (ddATP, ddTTP, ddGTP, ddCTP) est utilisé. Un didésoxynucléotide (ddNTP) est un analogue du dNTP normal. Il diffère par l'absence d'un groupe hydroxyle à la position 3 'du carbone. Lorsqu'un didésoxynucléotide est incorporé pendant la synthèse de l'ADN, aucune liaison entre sa position 3 'et le nucléotide suivant n'est possible parce que le ddNTP ne possède pas le groupe hydroxyle en 3'. Ainsi, la synthèse de la nouvelle chaîne est terminée sur ce site. Le fragment d'ADN à séquencer doit être simple brin <fn> Brown, T.A. Genomes. Bios Scientific Publ., Oxford, 1999. </fn>. La synthèse d'ADN est initiée en utilisant une amorce et l'un des quatre ddNTP marqués avec ^{32}P dans les groupes phosphate ou, pour un séquençage automatisé, avec un fluorophore (voir plaque suivante). Ici, un exemple de terminaison de chaîne utilisant ddATP est montré <fn> Strachan, T., Read, A.P. : Human Molecular Genetics. 2e éd. Bios Scientific Publishers, </fn>. Chaque fois qu'une adénine (A) apparaît dans la séquence, le didésoxyadénine triphosphate provoque la terminaison de la nouvelle chaîne d'ADN synthétisée. Ceci produira un ensemble de fragments d'ADN différents dont les tailles sont déterminées par les positions des résidus d'adénine apparaissant dans le fragment à séquencer. Des réactions similaires sont faites pour les trois autres nucléotides. Les quatre réactions parallèles donneront un ensemble de fragments avec des tailles définies en fonction des positions des nucléotides où la nouvelle synthèse d'ADN a été terminée. Les fragments sont séparés selon la taille par électrophorèse sur gel comme dans la méthode chimique. Le gel de séquence est lu dans la direction allant de petits

fragments à de grands fragments pour dériver la séquence nucléotidique dans la direction 5 'vers 3'. Un exemple d'un gel de séquençage réel est montré entre les panneaux A et B.