

COURS 3 : ADN eucaryote

1. Les séquences fonctionnelles non codantes

Au niveau de l'ADN eucaryote existe des séquences d'ADN non codantes mais qui jouent un rôle fondamental dans la régulation de l'expression des gènes, dans la transcription et dans la réplication. On peut citer :

- Les télomères contiennent des répétitions en tandem de séquences simples d'ADN qui ne codent aucun ARN mais présente une fonction définie indispensable à la réplication des télomères :

Exemple: la séquence TTAGGG de l'homme qui permet à la télomérase de régler le problème de la réplication des chromosomes linéaires de l'homme. Chez le cilié *Tetrahymena*, la séquence TTGGGG est répétée.

- tata box
- Séquences de début et de fin d'intron
- Les enhancers en cis qui permettent de lier les facteurs de transcription.
 - Les répresseurs, aussi appelés insulateurs diminuent l'expression génétique.

2. Les séquences sans fonction connue

Certains ADN répétés n'ont pas de fonction connue. Environ 20% du génome humain est constitué de séquences répétitives non fonctionnelles :

2.1. L'ADN hautement répétitif encadrant le centromère (Satellite) :

Cet ADN satellite est constitué de multiples répétitions en tandem de courtes séquences (10 bases) nucléotidiques qui peuvent s'étendre sur des centaines de Kb de long dans les régions d'hétérochromatine qui sont des régions très compacte encadrant les centromères (10% du génome humain). Cet ADN est appelé satellite car par centrifugation en gradient de densité de chlorure de césium, cet ADN forme une bande satellite

La plupart des gènes actifs sont situés dans l'euchromatine, région plus lâche que l'hétérochromatine.

2.2 l'ADN minisatellite

- Ce sont des séquences répétées en tandem dont la taille du motif unitaire est comprise entre 10 à 60 nucléotides et présentes chez de nombreuses espèces.

- Elles sont bien étudiées chez l'homme. On les retrouve dans environ 1000 localisations différentes dans l'ADN humain.
- En raison de la variabilité du nombre de répétitions en tandem (répétition en tandem polymorphe ou de l'anglais VNTR: variable number of tandem repeats) d'une personne à l'autre, cet ADN est utilisé comme empreinte génétique et est utilisé en médecine légale.
- L'ADN est coupé grâce aux enzymes de restriction qui ne présente pas de sites cibles au niveau des VNTR. Alors on verra apparaître sur un transfert Southern un grand nombre de fragment de taille variable, hybridés avec la sonde VNTR.
- Dans les enquêtes criminelles, on peut utiliser du sang, des cheveux de la salive, pour comparer avec les différents suspects (Fig 4).

Dans la figure suivante; la sonde VNTR contient une séquence de bases précises par exemple : la séquence GGAGGTGGGCAGG répétée un nombre variable de fois (selon l'individu c'est ce qui sert à l'identification) dans le premier intron du gène de la myoglobine. L'hybridation de la sonde marqué, avec l'ADN à des endroits précis et variable selon chaque individu

- Ces minisatellites sont utilisés dans les tests de paternité qui consiste à évaluer la taille de certains minisatellites bien spécifiques
- L'ADN de l'enfant et des parents seront isolés. L'ADN de chaque paire de chromosome (père +mère) est amplifié par PCR après séparation des chromosomes
- Les brins sont découpés par des enzymes de restriction qui isole les minisatellites en préservant leur longueur. Ces morceaux seront analysés par électrophorèse



2.3 Polymorphisme d'un seul nucléotide : SNP

- Le polymorphisme nucléotidique ou polymorphisme d'un seul nucléotide (SNP, *single-nucleotide polymorphism*) est, en génétique, la variation (polymorphisme) d'une seule paire de bases du génome, entre individus d'une même espèce. Ces variations sont très fréquentes.
- Les SNP représentent 90 % de l'ensemble des variations génétiques humaines, ils sont présents toutes les cent à trois cent paires de bases en moyenne dans le génome humain, où deux SNP sur trois substituent la cytosine avec la thymine.
- Les SNP peuvent se retrouver au sein de régions codantes de gènes, de régions non codantes de gènes, ou de régions intergéniques.
- Dans le cas où les SNP se retrouvent au sein des régions codantes, celles-ci ne vont pas obligatoirement modifier la séquence d'acide aminé de la protéine produite, et ce, grâce à la redondance du code génétique.
- Les SNP sont des outils permettant d'identifier des génotypes (reconnaître des personnes, par exemple) ou permettant de contribuer à la construction d'arbres généalogiques d'êtres vivants ou d'espèces.

2.4 L'ADN minisatellite

- C'est de l'ADN répétitif dispersé et composé de répétitions dinucléotidiques en tandem. Le nombre de répétition est variable.
- Ce type d'ADN a fourni de nombreux marqueurs moléculaires utiles pour la cartographie des génomes de grande taille de l'homme ou d'autres espèces.

2.5 Les séquences transposées

- Une grande proportion du génome eucaryote est constituée d'éléments répétitifs qui se sont propagés dans le génome, en fabriquant des copies d'eux-même, capables de gagner d'autres emplacements.
- Cet ADN mobile constitue les transposons.
- Ces transposons sont dits de classe II.
- Ils transposent sur le mode du:
 - - « couper-coller » (ex : Tn10, Tn5, Mos1, élément P) c'est-à-dire que leur transposition est couplée à leur excision de leur site d'origine
 - ou "copier-coller" (ex : IS911), soit à leur réplication (copier).
- Certains sont autonomes (codant une enzyme, la *Transposase*, permettant son propre transfert après réplication)
- D'autres non autonomes, devant utiliser la machinerie d'autres éléments autonomes.

2.6 Les rétrotransposons

Ces éléments à ARN sont de classe I. Ils fonctionnent à la façon du copier-coller

Les rétrotransposons se propagent d'eux-même sous l'action de la transcriptase inverse, (ADN→l'ARN). Ils seront rétro-transcrits avant leur insertion dans la molécule d'ADN.

Exp les éléments Ty de la levure (des éléments de 6 Kb, présents en 30 copies environ par génome).

On distingue les éléments dits « à LTR » (pour **L**ong **T**erminal **R**epeats), qui sont encadrés par de longues répétitions non inversées, et les éléments « sans LTR », qui n'en possèdent pas.

- **Les LTR**

Le cycle de transposition des éléments à LTR ressemble énormément à celui d'un rétrovirus

On a soupçonné pendant longtemps ces éléments d'être d'anciens virus ayant perdu leur capacité à sortir de la cellule, les récentes découvertes tendent à montrer qu'au contraire les rétrovirus sont d'anciens éléments transposables ayant gagné la capacité infectieuse.

Ces LTR sont autonomes

Ils codent pour la transcriptase inverse. L'intégrase (permet l'intégration du transposon dans l'ADN).

- **Les éléments LINE :**

Ils sont dispersés ou LINE (*long interspersed nuclear element*) des mammifères,

- Ce sont des éléments sans LTR, de 1 à 5 kb présents en 20000 à 40000 exemplaires environ dans le génome humain.
- Codent une protéine unique permettant la séparation, la réplication et l'intégration du transposon.

- **Séquence Alu :**

La séquence Alu dans le génome humain, contient un site cible pour l'enzyme de restriction Alu.

Elle constitue 5% du génome humain. C'est une séquence d'environ 200 nucléotides répétée des centaines de milliers de fois dans le génome humain

Elles sont désignées par le terme collectif de SINE pour courts éléments dispersés (*short interspersed elements*). Elles ne sont pas autonomes.

3. Les pseudogènes

Ils sont très dispersés dans le génome, ils ne contiennent pas les introns présents dans le gène fonctionnel d'origine.

Conclusion

L'existence d'ADN sans fonction connue constitue une énigme pour les généticiens, d'autant plus que la synthèse d'ADN consomme énormément d'énergie, peut être sert-il à séparer les éléments fonctionnels pour permettre une régulation efficace, cet ADN est appelé ADN égoïste (*junky ADN*) car il n'existerait que pour lui-même.

