

## Le Polymorphisme

### Définition

- variation individuelle de la séquence nucléotidique
- entraînant l'existence au même locus d'au moins 2 formes différentes de la séquence, appelés allèles
- dont la fréquence est  $>$  à 1% dans la population
- intra ou extra-génique
- d'effet neutre
- En dehors d'un gène
- Dans une région non codante d'un gène (intron)
- Dans une région codante mais sans altérer la signification (redundance du code génétique)

### Diversité des séquences humaines

#### - Pour les 95 % non codants

Des estimations donnent:

2 génomes humains diffèrent en moyenne par 1/100 à 1/500 nucléotides. Si on pouvait comparer entre l'ADN d'origine paternelle et l'ADN d'origine maternelle chez un individu on trouverait 10 milliard de différence correspondant à l'accumulation de mutations au cours de l'évolution humaine.

La proportion totale de bases polymorphiques «Index d'hétérozygotie d'ADN» a été estimé à 1/270 pb pour tout segment d'ADN du génome choisi au hasard.

#### - Pour les 5 % codants

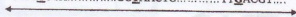
Les régions codantes sont soumises à une sélection plus importante. La présence de mutations dans ces régions est plus faible.

### Diversité nucléotidique $\pi$ :

Nombre de différences, rapporté à la longueur de la séquence examinée, entre 2 séquences prises au hasard dans la population.

AGTCCTA.....GGAAACTG.....TTTACGT....

AGTGCTA.....GGTAACTG.....TTGACGT....



7500 pb

$$\pi = 3/7500 = 1/2500$$

## I. Les Marqueurs polymorphes

Comment toutes ces variations ont été détectées?

1. RFLP bi-alléliques
2. Minisatellites
3. Microsatellites
4. SNP (single nucleotide polymorphism)

### 1. Polymorphisme de la longueur des fragments de restriction

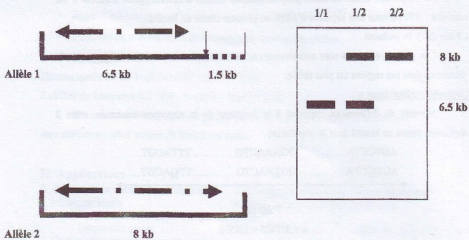
[RFLP: Restriction Fragment Length polymorphism]

- Ce sont des variations individuelles de la séquence d'ADN révélées par des modifications de la carte de restriction. Mise en évidence par le Southern qui montre des différences obtenues avec une enzyme donnée et une sonde donnée.

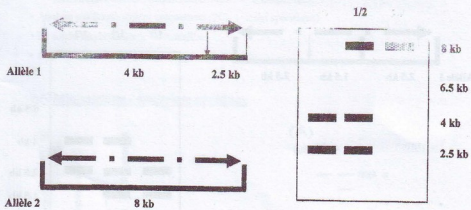
- Les enzymes de restrictions possèdent des séquences reconnues dans l'ADN, une mutation au niveau de cet ADN entraîne la formation ou la perte de site de reconnaissance, ce qui entraîne des modifications dans la taille d'un ou plusieurs fragments révélés par Southern.

- De même si un segment d'ADN situé entre 2 sites de restriction était délété ou inséré, la taille du fragment de restriction sera différente.

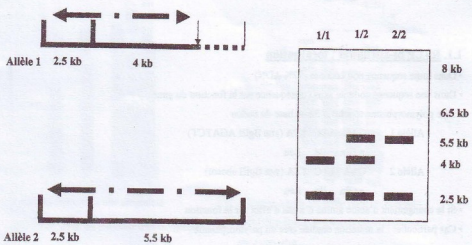
#### Les sites polymorphes



PAS DE SITE INTERNE

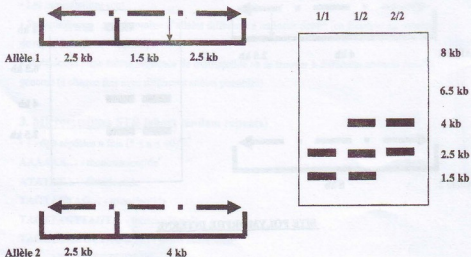


SITE POLYMORPHE INTERNE



SITE INTERNE NON POLYMORPHE

SITE INTERNE NON POLYMORPHE



**SITE INTERNE NON POLYMORPHE**

**SITE POLYMORPHE INTERNE**

### **1.1. RFLP bi-alléliques : localisation**

- Dans toute séquence non codante (97% ADN)
- Dans une séquence codante sans conséquence sur la fonction du gène
- Si le polymorphisme touche la 3ème base du codon
  - Allèle 1 CAG ATC TTA (site BglII AGATCT)
    - Gln Ile Leu
  - Allèle 2 CAA ATC TTA (site BglII absent)
    - Gln Ile Leu
- Si le changement d'acide aminé n'a pas d'effet sur la fonction
- Cas particulier : la mutation étudiée crée un polymorphisme

### EXEMPLE d'un polymorphisme multiallélique

Le polymorphisme de l'apolipoprotéine E

Protéine:

	E2	E3	E4
AA en 112	Cystéine	Cystéine	Arginine
AA en 158	Cystéine	Arginine	Arginine

ADN: Fragment de 292 pb (9 sites de coupure)

Hha I: GCG↓C

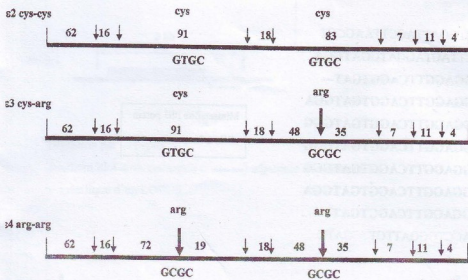
CGC = Arginine

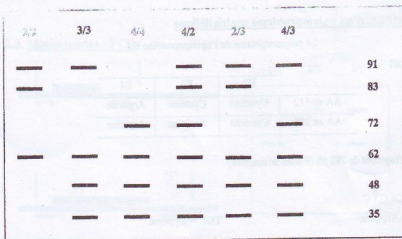
TGC = Cystéine

### Génotypage de l'apoE

■ Amplification par PCR Dénaturation : 30 s à 95°C, Hybridation : 30 s à 60°C Elongation : 1 min à 72°C

■ Digestion de l'ADN amplifié par Hha I GCG↓C (TGC cystéine CGC Arginine)





(E2/E2)=91+83+62

(E4/E2)=91+83+72+62+48+35

(E3/E3)=91+62+48+35

(E2/E3)=91+83+62+48+35

(E4/E4)=72+62+48+35

(E4/E3)=91+72+62+48+35

**2. Minisatellites: RFLP multi-alléliques VNTR (variable number tandem repeats)**

CTGGATTAGCTTAAGCGT

CTTAGTAGGATGGATGGA

GGAGGTTCAAGGTGAT--

GGAGGTTCAAGGTGATGGA

GGAGGTTCAAGGTGATGGG

GGAGGTTCAAGGTGATGGG

GGAGGTTCAAGGTGATGGG

GGAGGTTCAAGGTGATGGA

GGAGGTTCAAGGTGAT--

ACCTGGGATCTTCGATC

Minisatellite PS3 porcine

allèle à 7 répétitions

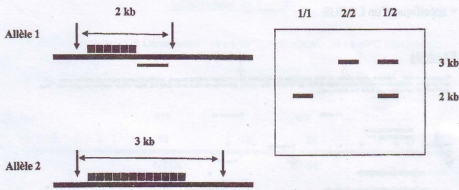
### 2.1. Répartition des minisatellites

- Répétitions en nombre variable d'une séquence de 10 à 50 pb à un locus donné
- Une même séquence de minisatellite va se retrouver à différents loci sur tout le génome
- Analyse locus par locus
- Analyse de la répartition d'un minisatellite sur le génome entier

### 2.2. Minisatellites - mise en évidence des allèles à un locus

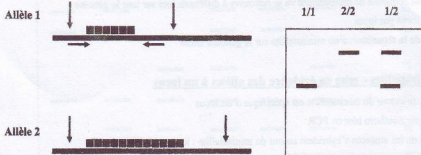
- Le contexte autour du minisatellite est spécifique d'un locus
- Analyse par Southern blot ou PCR
- La sonde ou les amorces s'hybrident autour du minisatellite : locus-spécifiques

### 2.3. Minisatellites : Southern blot



Digestion par l'enzyme ↓  
Southern blot avec une sonde ( ——— ) adjacente au minisatellite  
= spécifique d'un LOCUS

### 2.3. Minisatellites : PCR

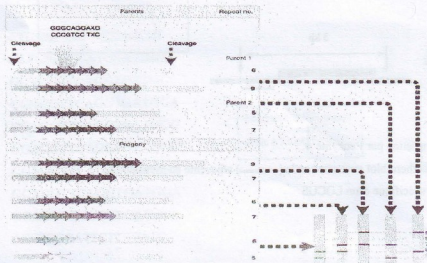


PCR avec 2 amorces encadrant le minisatellite

= spécifique d'un LOCUS

### Exemple

Figure 4.32 Alleles may differ in the number of repeats at a minisatellite locus, so that cleavage on either side generates restriction fragments that differ in length. By using a minisatellite with alleles that differ between parents, the pattern of inheritance can be followed.

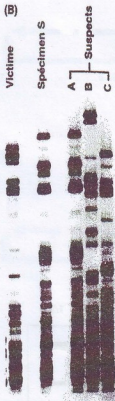




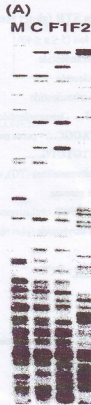
#### 2.4. Minisatellites - mise en évidence des allèles sur un gel

- Analyse par Southern blot uniquement (PCR = locus spécifique)
- La sonde s'hybride sur le minisatellite
- Applications : médecine légale, tests de paternité

#### 2.5. Minisatellites - empreinte génétique



- Chaque individu a une empreinte différente (sauf les jumeaux monozygotes) :
- Utilisation en médecine légale



- Chaque bande présente chez un individu doit l'être également soit chez sa mère soit chez son père :
  - test d'exclusion de paternité
- (M = mother, C = child, F = father)

## 2.6. Minisatellites - récapitulation

- Les minisatellites sont:
- Poly-alléliques : grand nombre d'allèles différents à un locus donné (en fonction du nombre de répétitions de la séquence)
- Multi-locus : une même séquence de minisatellite va se trouver à différents endroits dans le génome (à chaque fois avec différents allèles possibles)

## 3. Microsatellites STR (short tandem repeats)

- 1 - 6pb répétées n fois ( $5 \leq n \leq 40$ )

AAAAAA... - mononucléotide

ATATAT... - dinucléotide

TAGTAGTAG... - trinucléotide

TAGTTAGTTAGT... - tetranucléotide

TAGGCTAGGCTAGGC... - penta nucléotide

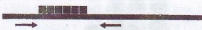
• Ex: TGTGTGTGTG.....

•  $\geq 50\ 000$  microsatellites (TG)<sub>n</sub> dans tout le génome humain

- Présents sur tout le génome
- Mise en évidence par PCR exclusivement
- Migration en gel d'acrylamide (plus résolutif que l'agarose)

## 3.1. Microsatellites : PCR

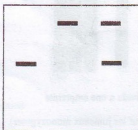
Allèle 1



Allèle 2



1/1    2/2    1/2



PCR avec 2 amorces encadrant le microsatellite = spécifique d'un LOCUS

### 3.1. Microsatellites - PCR

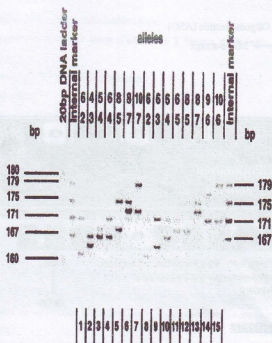


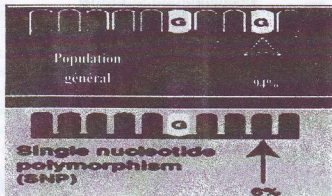
FIG. 1. Resolution of different alleles in genotype analysis by the nonradioactive method. 100 ng of DNA samples was subjected to PCR amplification using D18553-specific primer sets. The PCR products were analyzed in a 7% denaturing PAGE and bands were visualized by ethidium bromide fluorescence. A 20-bp DNA ladder (Cilco/BRL) was used for a standard size marker. Sizes of different alleles in basepairs were determined by using the size marker and the "internal marker." Each line (1- 15, as numbered at the bottom) represents analysis of PCR products from a single subject whose genotype was scored as shown at the top.

#### 4. Single Nucleotide polymorphism = SNP

- Causés par mutagenèse chimique ou erreurs de répliation
- bi-alléliques
- ratio des allèles : 99,99 : 0,01 → 50:50
- Nb loci identifiés chez l'Homme >> 5 millions
- taux de mutation  $1.10^9$ /génération
- en conséquence, peu de mutations nouvelles dans l'espèce

Mise en évidence :

- idéal : le polymorphisme correspond à un RFLP
- PCR-SSCP
- Allele-Specific Oligonucleotide (ASO)
  - Micro-arrays
- séquençage
- ...



### Marqueurs génétiques

- est caractérisé par des allèles correspondant à chaque variation possible de la séquence
- degré d'hétérozygotie d'un marqueur  $H=1 - (\sum p_i^2)$
- Fonction du nombre et de la fréquence de chacun des allèles
- **Exemple** : 3 allèles de fréquence 0,2 / 0,3 / 0,5

Hétérozygotie est de  $1 - (0,04 + 0,09 + 0,25) = 0,62$

2 allèles de fréquence 0,2 / 0,8 :  $1 - (0,04 + 0,64) = 0,32$

Un marqueur biallélique ne pourra pas avoir une hétérozygotie supérieure à 0,5 qui sera atteinte en cas d'allèles de fréquence égale.

### **II. Applications**

1. Gestion de la variabilité génétique
2. Sélection assistée par marqueur génétique
3. Médecine légale
  - Diagnostic indirect par analyse de liaison
  - Diagnostic direct
4. Diagnostic de maladies génétiques
5. Cartographie et clonage positionnel