

# TEMA 4

## Identitat genètica

1. Ús de marcadors moleculars per a l'obtenció d'empremtes genètiques
2. Assignació de probabilitats en genètica forense
3. Assignació de probabilitats en proves de paternitat
4. Resolució de problemes

# 1. Ús de marcadors moleculars per a l'obtenció d'empremtes genètiques

L'empremta genètica (també anomenada Prova d'ADN) ens permet obtenir un patró característic i específic per a cada individu mitjançant la detecció de polimorfismes a l'ADN.

Va ser descrita per primera vegada per Sir Alec Jeffreys, de la Universitat de Leicester (Regne Unit) al 1985. (Jeffreys, A.J., Wilson, V. and Thein, S.L. 1985. Nature 316, 76-79)

- La seua anàlisi es basava en l'alta variabilitat entre individus de zones no codificants del DNA, en les quals es repetia una seqüència en tàndem.
- Va observar que s'obtenia un patró específic per a cada individu, excepte en bessons univitel·lins.
- El marcador polimòrfic va rebre el nom de VNTR (Variable Number of Tandem Repeats) o minisatèl·lit.

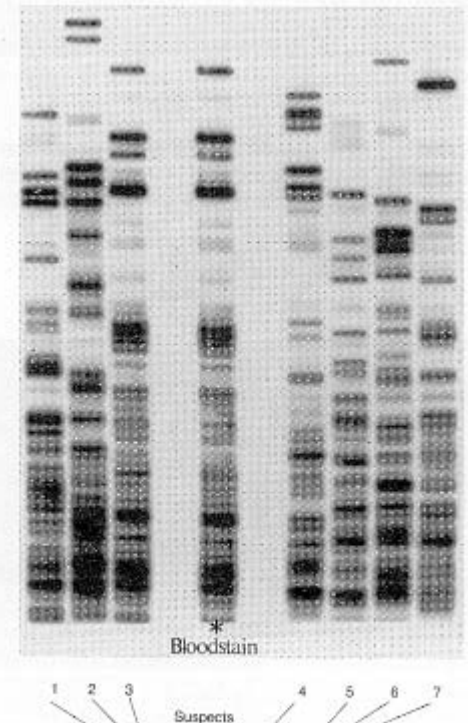
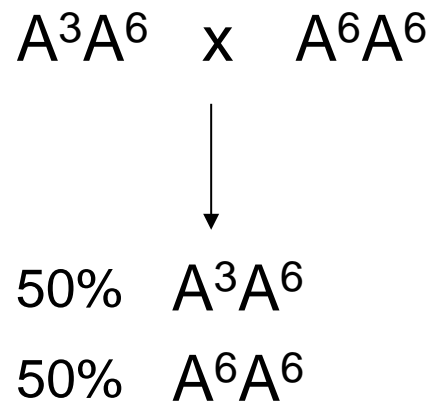
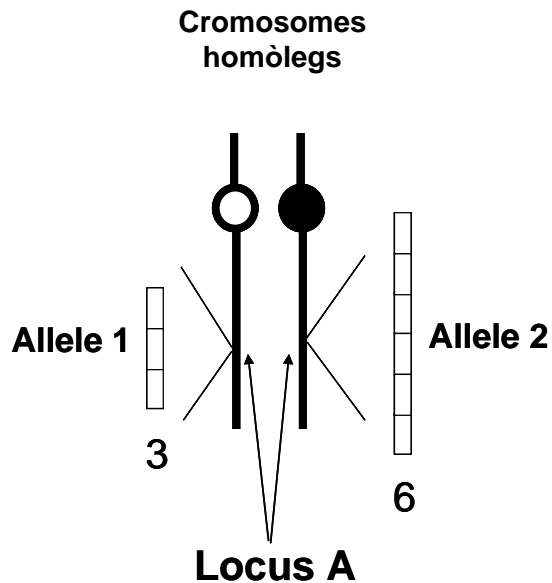


Figure 12-10 DNA fingerprints from a bloodstain at the scene of a crime and from the blood of seven suspects. (Cellmark Diagnostics, Germantown, MD.)

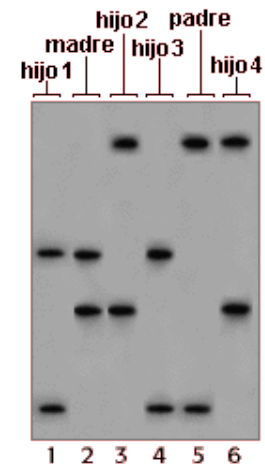
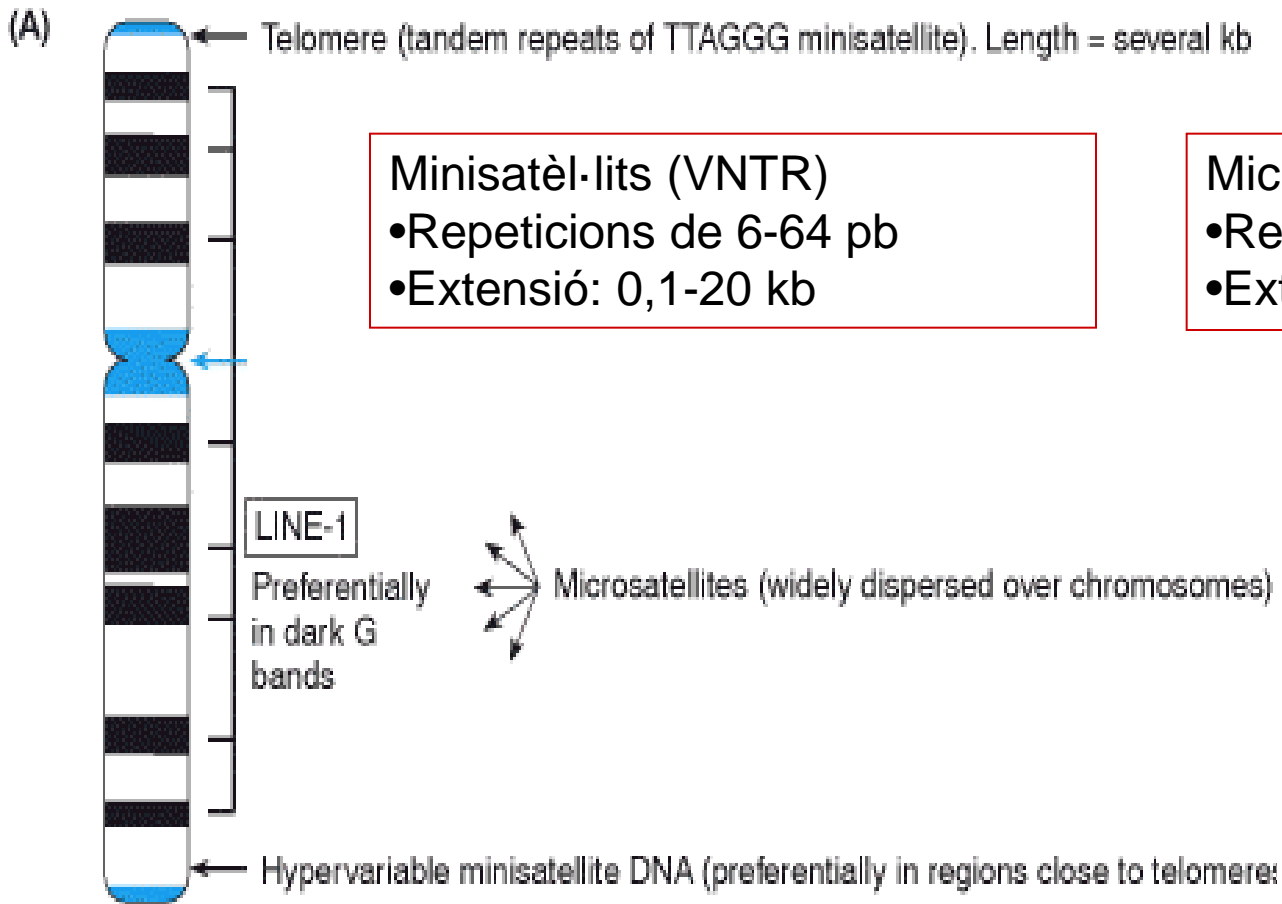
Al 1989 es van descriure altres marcadors moleculars útils per a l'obtenció d'empremtes genètiques: els STR (també anomenats microsatèl·lits).

Encara que es poden utilitzar altres marcadors (RAPDs, AFLPs), els VNTR i els STR continuen sent els marcadors d'elecció per a l'obtenció d'empremtes genètiques. **L'herència d'aquests marcadors és codominant:**



En alguns casos es fan servir també els SNPs, que en anys recents estan adquirint un pes important. **També són codominants.**

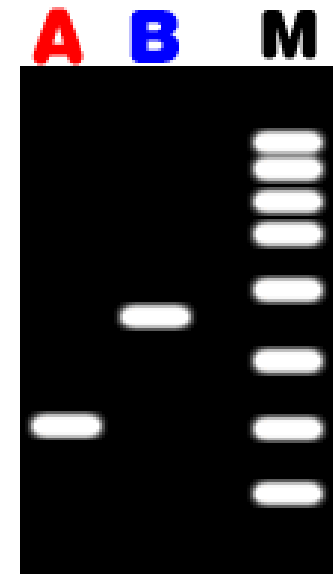
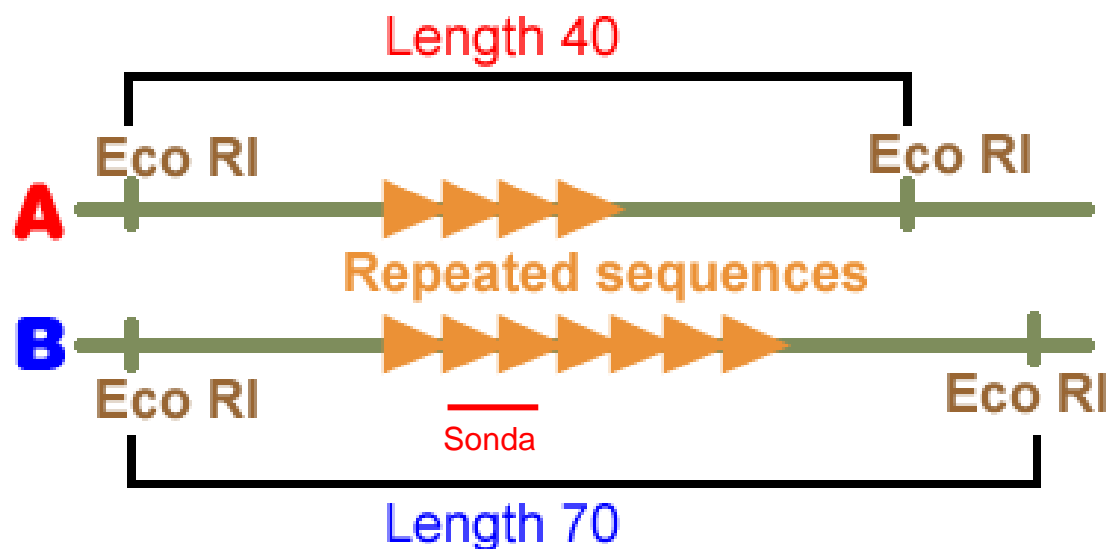
# Minisatèl·lits i microsatèl·lits:



ÚS: Identitat genètica. Usos forenses, paternitat.

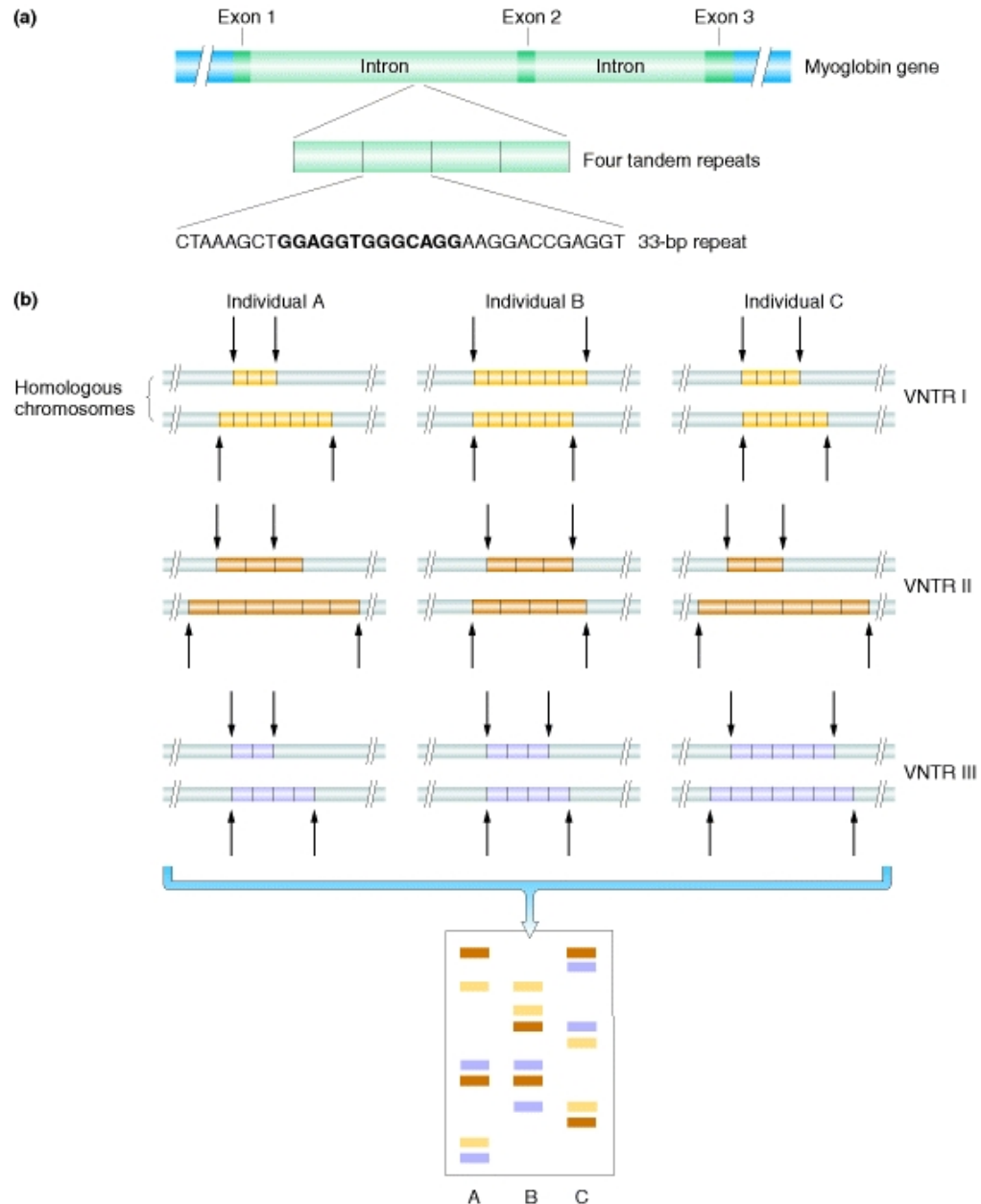
## Detecció de minisatèl·lits:

- L'augment o disminució en la llargària d'un minisatèl·lit resulta de canvis en el nombre de repeticions.
- El perfil de VNTR apareix quan un enzim de restricció talla fora de la repetició.
- Es detecten fent servir sondes.



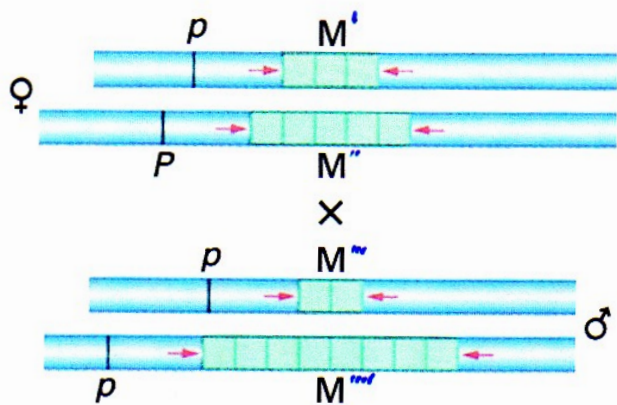
## Detecció de minisatèl·lits:

Normalment, minisatèl·lits derivats de la mateixa unitat de repetició es troben dispersos pel genoma, de manera que amb una sola sonda es pot detectar més d'un locus ("sonda multilocus").

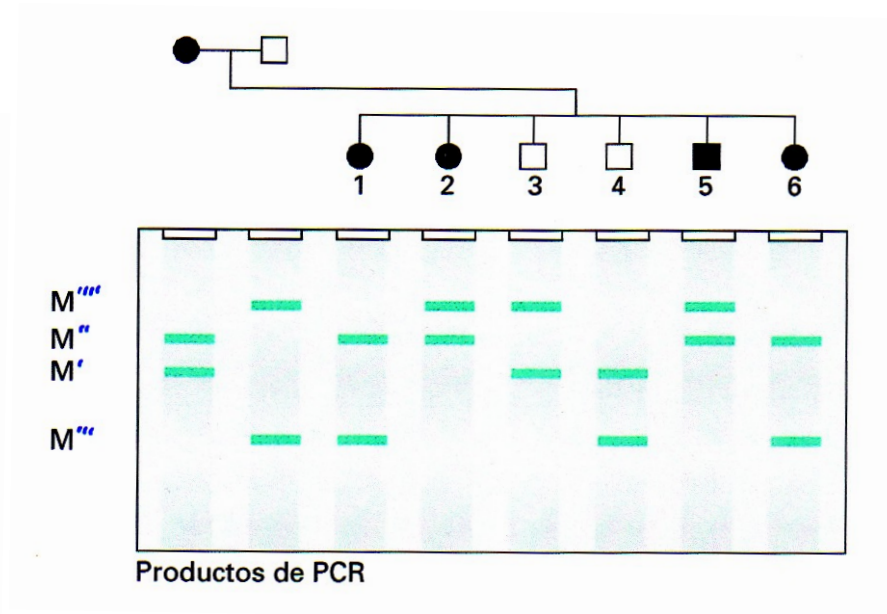


## Detecció de microsatèl·lits (STR):

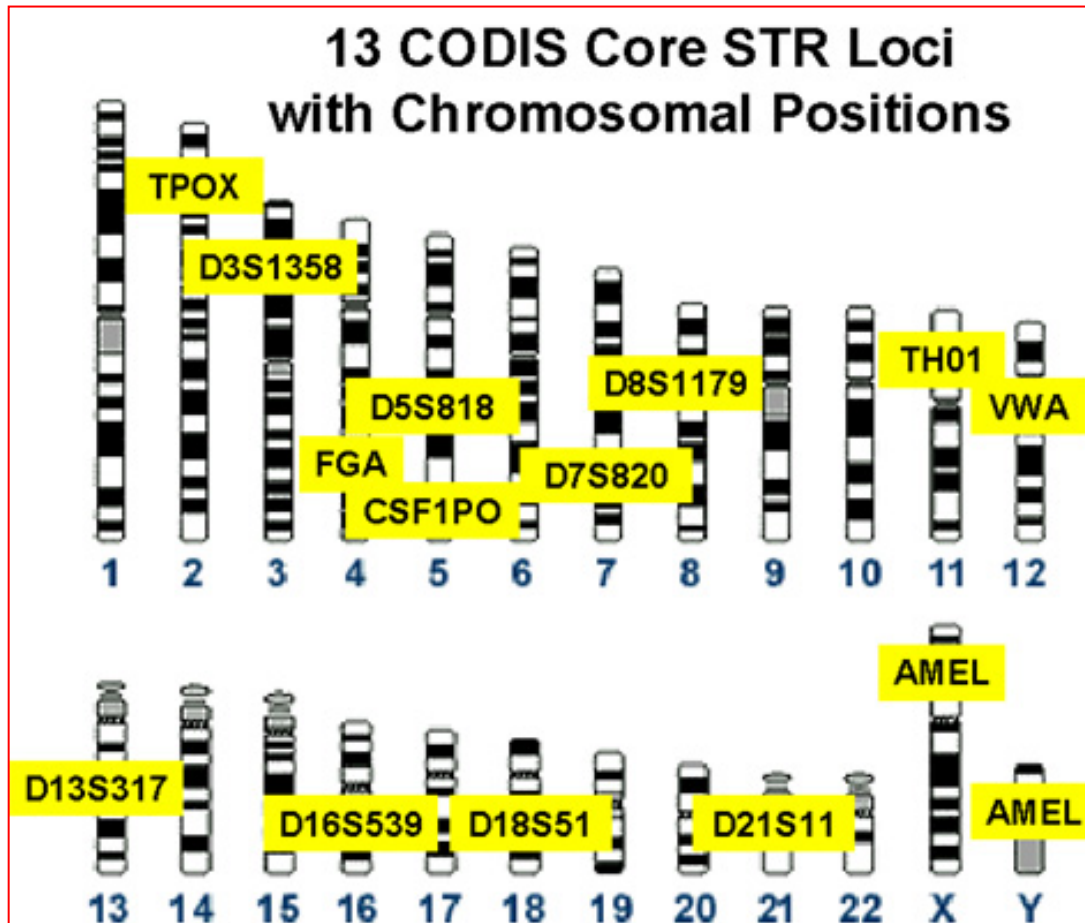
Es detecten per PCR, per amplificació de la seqüència repetida utilitzant un parell d'encebadors que hibriden amb les regions flanquejants.



- Clave
- Cebadores de PCR
  - Repeticiones del microsatélite
  - P* Alelo dominante de la enfermedad
  - $M^1 - M^4$  Marcadores moleculares



Al 1997, l'FBI va anunciar la selecció de 13 loci STR a utilitzar per a la identitat genètica humana, que formarien el nucli central de la base de dades nacional dels EUA, anomenada CODIS. Tots els STR del CODIS són seqüències repetitives tetramèriques.



THE FBI's  
**COMBINED DNA  
INDEX SYSTEM  
PROGRAM**

**CODIS**

**CODIS**

CODIS

A Federal/State Partnership Fighting Violent Crime

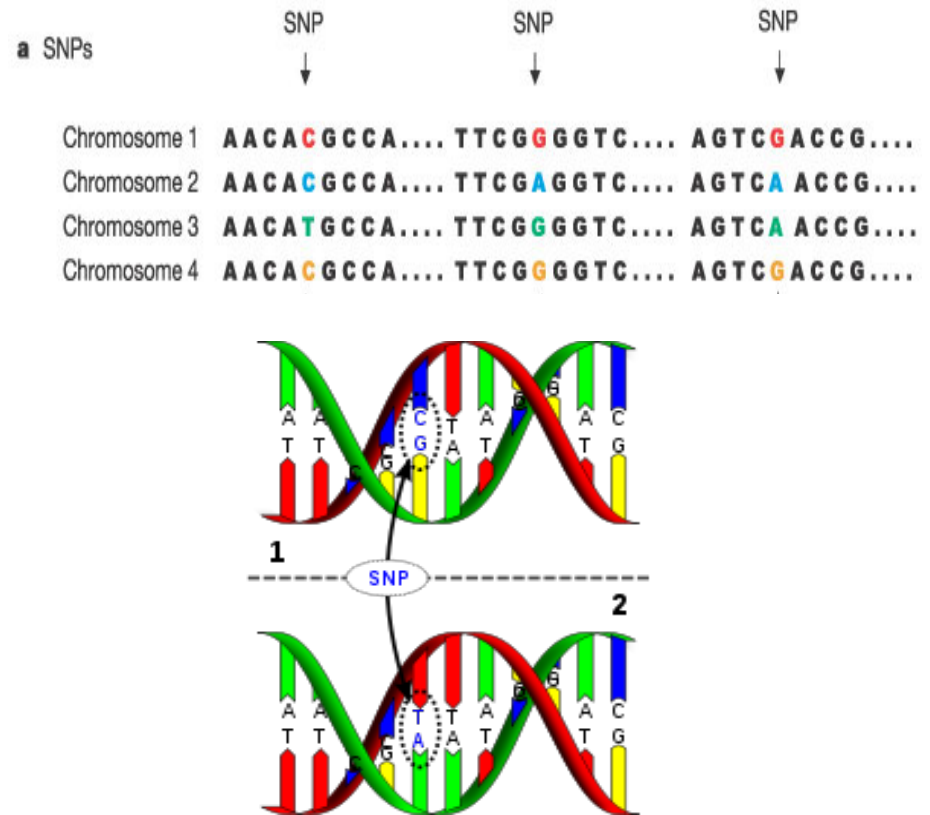


# SNPs: *Single Nucleotide Polymorphisms* (polimorfismes d'un únic nucleòtid)

L'SNP apareix per un canvi en el nucleòtid present en una posició específica d'un cromosoma.

Els SNP es consideren una forma de mutació puntual que ha tingut prou èxit evolutiu per fixar-se en una part significativa de la població d'una espècie.

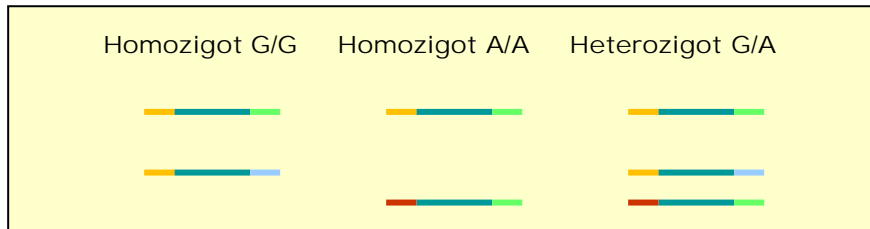
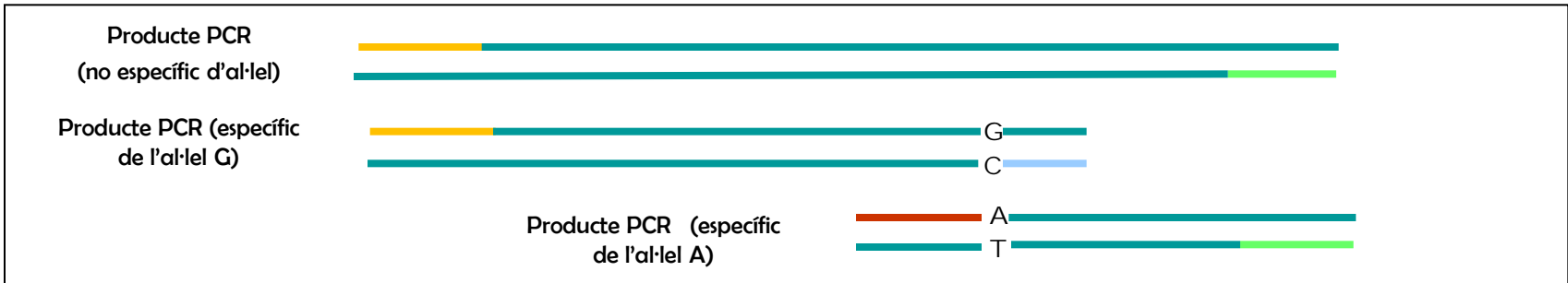
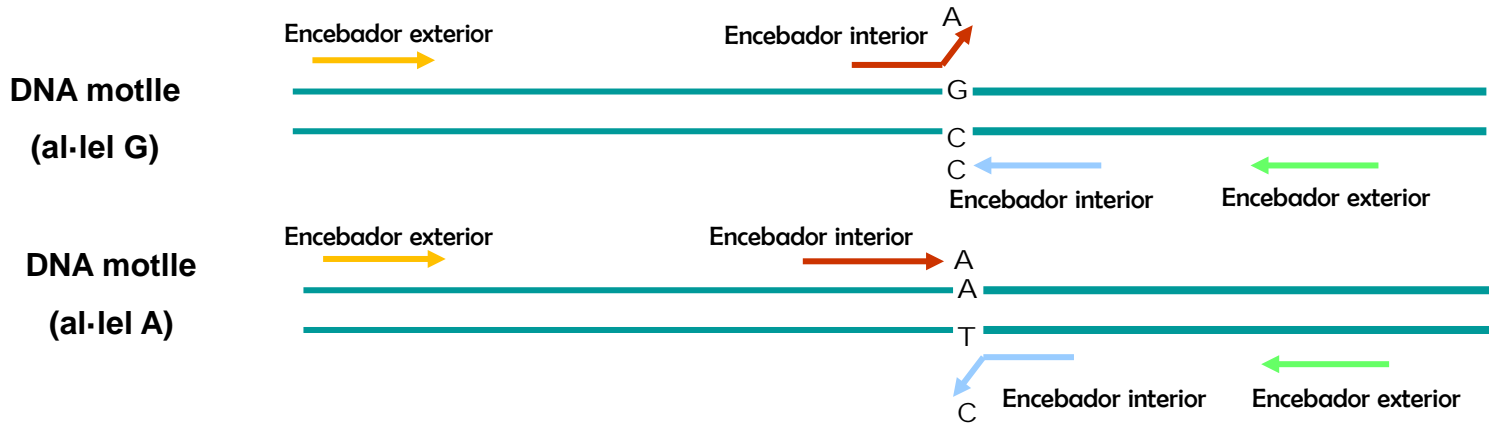
Els marcadors SNP es basen en la capacitat de detecció del canvi d'un únic nucleòtid.



ÚS: Mapeig genòmic, lligament a marcadors fenotípics (malalties), identitat genètica.

# Genotipat d'SNPs:

Actualment els SNPs se solen identificar mitjançant PCR. Un dels mètodes és el que fa servir dos encebadors externs i dos d'interns, fent coincidir l'extrem 3' amb l'SNP:



Ye *et al* (2001), Nucl. Acids Res., 29, (17) e88

# SNPs utilitzats en genètica forense:

*L.A. Dixon et al./Forensic Science International 154 (2005) 62–77*

Table 1

Allele frequencies for each of the 20 SNP loci used in the multiplex for each race group studied and combined likelihood ratios for each group.

| SNP locus | Allele 1 (green)/allele 2 (blue) | White Caucasian |          | British Afro-Caribbean |          | Indian sub-continent |          |
|-----------|----------------------------------|-----------------|----------|------------------------|----------|----------------------|----------|
|           |                                  | Allele 1        | Allele 2 | Allele 1               | Allele 2 | Allele 1             | Allele 2 |
| D         | T/C                              | 0.52            | 0.48     | 0.27                   | 0.73     | 0.49                 | 0.51     |
| U6        | A/T                              | 0.37            | 0.63     | 0.10                   | 0.90     | 0.32                 | 0.68     |
| B6        | A/T                              | 0.64            | 0.36     | 0.77                   | 0.23     | 0.57                 | 0.43     |
| N4        | A/T                              | 0.57            | 0.43     | 0.51                   | 0.49     | 0.53                 | 0.47     |
| Y3        | G/C                              | 0.92            | 0.08     | 0.96                   | 0.04     | 0.94                 | 0.06     |
| P5        | T/A                              | 0.72            | 0.28     | 0.59                   | 0.41     | 0.80                 | 0.20     |
| A4        | C/G                              | 0.71            | 0.29     | 0.51                   | 0.49     | 0.66                 | 0.34     |
| O6        | A/T                              | 0.75            | 0.25     | 0.82                   | 0.18     | 0.77                 | 0.23     |
| Z2        | C/T                              | 0.56            | 0.44     | 0.45                   | 0.55     | 0.45                 | 0.55     |
| K3        | G/C                              | 0.31            | 0.69     | 0.25                   | 0.75     | 0.39                 | 0.61     |
| J2        | C/T                              | 0.92            | 0.08     | 0.94                   | 0.06     | 0.94                 | 0.06     |
| Y6        | T/A                              | 0.63            | 0.37     | 0.57                   | 0.43     | 0.53                 | 0.47     |
| P7        | T/A                              | 0.62            | 0.38     | 0.73                   | 0.27     | 0.79                 | 0.21     |
| J8        | A/T                              | 0.77            | 0.23     | 0.86                   | 0.14     | 0.72                 | 0.28     |
| X         | C/A                              | 0.79            | 0.21     | 0.89                   | 0.11     | 0.78                 | 0.22     |
| F         | C/A                              | 0.78            | 0.22     | 0.85                   | 0.15     | 0.80                 | 0.20     |
| G         | T/C                              | 0.75            | 0.25     | 0.64                   | 0.36     | 0.59                 | 0.41     |
| L2        | C/T                              | 0.79            | 0.21     | 0.94                   | 0.06     | 0.90                 | 0.10     |
| W3        | C/G                              | 0.77            | 0.23     | 0.87                   | 0.13     | 0.77                 | 0.23     |
| H8        | A/T                              | 0.11            | 0.89     | 0.10                   | 0.90     | 0.15                 | 0.85     |

Les freqüències al·lèliques dels marcadors varien segons la població de referència.

## Avantatges/beneficis dels SNPs davant els STRs:

- La mesura de l'ampliació és xicoteta → millor per a l'anàlisi de mostres degradades.
- Abundant al genoma humà (mitjana: 1 SNP/1000 bases).
- Taxa de mutació més baixa que per a STRs ( $10^{-8}$  vs.  $10^{-3}$ ).
- S'estan desenvolupant mètodes de tipat múltiple, que permetrien l'automatització.

## ¿Reemplaçaran els STRs per obtenir empremtes genètiques?

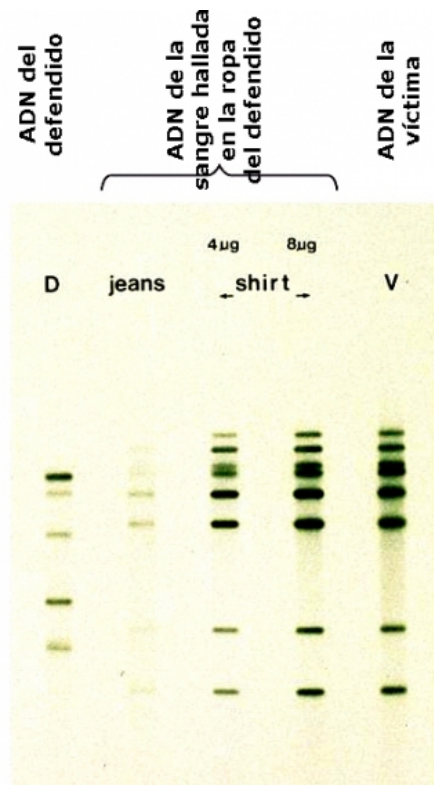
Malgrat els seus avantatges, es considera improbable que els SNPs reemplacen els STR com el mètode més utilitzat en mostres forenses, almenys a curt o mitjà termini. No obstant això, s'utilitzen per a usos especialitzats :

- Anàlisi de mtDNA.
- Anàlisi de Y-SNPs: llinatge, estudis de població, discriminació de mostres.
- SNPs autosòmics: per a mostres molt degradades, o escasses (pèls), determinació ètnica/geogràfica...

Les **principals aplicacions** de l'empremta genètica en l'assignació d'identitats són en l'anàlisi criminal (o forense) i en l'establiment de relacions de parentiu (proves de paternitat):

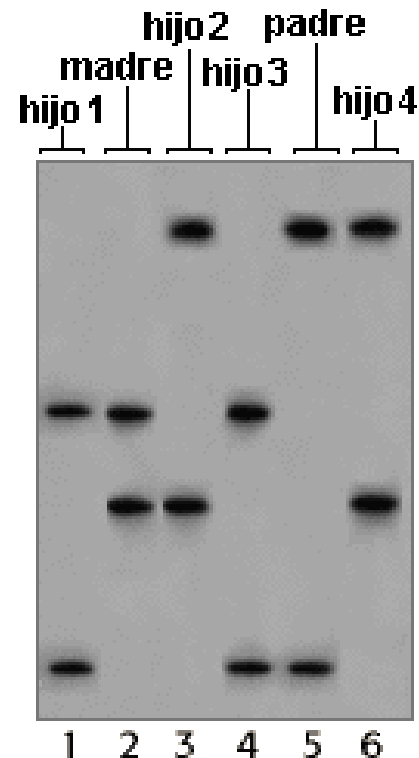
Aplicacions forenses:

Inculpació/exculpació de sospitosos.



Proves de paternitat:

Exclusió o paternitat probable



# Altres aplicacions de l'empremta genètica:

- **En genètica humana:**
  - Indispensable per identificar restes d'accidents massius (accidents aeris, naufragis).
  - Permet rastrejar l'evolució geogràfica de poblacions humanes, així com el possible origen i les relacions genètiques entre poblacions separades.
  - Permet deduir com s'han produït els fluxos migratoris en una escala temporal relativament extensa, al llarg de la història i prehistòria humanes.
- **En microbiologia:** és una eina molt útil capaç d'aconseguir una precisió a nivell de subespècie, on altres tècniques no són capaces d'arribar.
- **En veterinària:** estudis d'evolució d'espècies salvatges, per determinar el grau de distanciament entre animals domèstics o per definir el possible aïllament de poblacions animals causades per l'impacte humà, entre altres molts exemples.
- **En agroalimentació:** traçabilitat d'animals i plantes.

## 2. Assignació de probabilitats en genètica forense

L'anàlisi genètica forense es basa en aportar evidències (proves d'ADN) que el sospitós ha estat o no en l'escena del crim.

L'anàlisi consisteix a comparar el patró de DNA del vestigi biològic identificat en l'escena del crim i el del possible sospitós.

Els resultats poden exonerar o culpabilitzar el sospitós.

Perquè la prova d'ADN pugui tenir valor jurídic s'han d'acomplir una sèrie de requisits, que es corresponen amb 4 etapes bàsiques :

| Etapa                                | Requisits  |
|--------------------------------------|--|
| 1. <b>Obtenció de les mostres</b>    | <b>Custòdia apropiada fins a la realització de l'anàlisi</b>             |
| 2. <b>Realització de l'anàlisi</b>   | <b>Laboratoris validats i protocols estandarditzats</b>                  |
| 3. <b>Comparació dels resultats</b>  | <b>Valoració probabilística de la prova</b>                              |
| 4. <b>Comunicació dels resultats</b> | <b>Asèpticament i de forma intel·ligible per a jutge, fiscal i jurat</b> |

En comparar els resultats de l'empremta genètica de la mostra (restes de sang, semen, etc.) amb l'empremta genètica de l'inculpat (o de la víctima, si el vestigi s'ha trobat sobre l'inculpat) podem trobar-nos amb dues possibilitats :

- **Ambdues empremtes difereixen** (encara que siga en un sol locus):



Si els patrons comparats són diferents significa que no pertanyen al mateix individu.

- **Ambdues empremtes coincideixen exactament** →



Si els patrons comparats són idèntics, podria tractar-se de la mateixa persona, però també podria tractar-se d'un altre individu de la població que compartira aquest mateix genotip.

→ S'ha de calcular la probabilitat que la mateixa empremta genètica es poguera donar en altres individus de la població.

Necessitem fer una anàlisi probabilística.



## Vejam-ne un exemple:

Es troba una taca de sang a l'escena del crim,  
que no pertany a la víctima.

S'analitza i es compara amb la de l'acusat del crim:



Anàlisi del grup sanguini: Ambdues són del grup B

Un 15% de la població és del grup B → valor probatori baix

Anàlisi d'ADN, sistema *D1S80*: Ambues mostres tenen el fenotip 19-29

Un 6,7% de la població té aquest fenotip → valor probatori baix

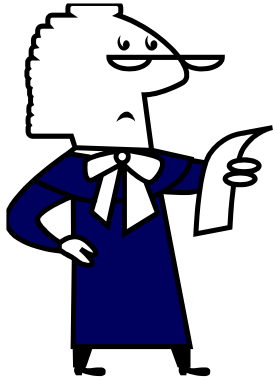
Anàlisi conjunta:

Únicament un 1% de la població posseeix el fenotip ( $0,15 \times 0,067 = 0,01$ )

↙ valor probatori més alt

Com més loci analitzem, més resolutive serà l'anàlisi.  
Les probabilitats són independents (si els loci no estan lligats).

# Però hi ha moltes formes de presentar la prova:



**FISCAL:** “L'anàlisi forense té molta importància en aquest cas. Només un 1% de la població té aquest fenotip. Per tant, només hi ha un 1% de probabilitats que siga un altre individu el culpable”.

→ 99% que siga culpable.

**DEFENSA:** “La prova forense té molt poca importància. Només l'1% de la població té aquest fenotip, però en una ciutat com aquesta (per exemple Madrid), de 500.000 persones en edat de cometre el crim, podríem trobar el fenotip en 5.000 persones”.

→ L'acusat és només una de les 5.000 persones de la ciutat que van poder cometre el crim. La probabilitat que siga culpable és 1/5.000.

Els dos arguments són falsos ("fal·làcia del fiscal" i "fal·làcia de la defensa"), però només presentant el percentatge de la població que pot tenir aquest fenotip, la majoria d'individus cauria en una de les dues fal·làcies. La interpretació no és intuïtiva.

La manera correcta de valorar dita prova és recorrent a l'anàlisi bayesiana:

### Teorema de Bayes:

El teorema de Bayes permet calcular la probabilitat final d'un succés a partir d'una probabilitat inicial tenint en compte informacions addicionals que s'obtenen *a posteriori* (la prova genètica):

$$P_{\text{final}} = P_{\text{inicial}} \times LR$$

- La probabilitat inicial és el grau de convicció prèvia que posseeix el jutge sobre la culpabilitat de l'acusat (p.e., 10 a 1 a favor de la culpabilitat).
- La “LR” és la raó de versemblança (de l'anglès *likelihood ratio*) o raó bayesiana de probabilitat, i es calcula en base a las proves d'ADN.

La raó de versemblança es calcula de la següent manera:

$$LR = \frac{\text{Probabilitat de la troballa científica donada la culpabilitat}}{\text{Probabilitat de la troballa científica donada la innocència}} = \frac{P(E/C)}{P(E/I)}$$

La probabilitat de la troballa científica donada la innocència fa referència a la probabilitat de trobar aquest mateix perfil en un individu de la població seleccionat aleatòriament.

Aquesta probabilitat dependrà dels marcadors emprats i de la **població de referència** utilitzada.

Tornem a l'exemple anterior:

Normalment del que es parteix no és de les freqüències fenotípiques, sinó de les freqüències al·lèliques en la població de referència.

Suposem que  $f(B) = 0,15$  i  $f(i) = 0,43$ , i per al locus  $D1S80$ ,  $f(19) = 0,10$  i  $f(29) = 0,33$

Si ell és el culpable, la probabilitat que els fenotips coincidiren és 1.

Si ell és innocent, la probabilitat de trobar aquests fenotips en la població és:

Per al grup sanguini:  $P_B = (0,15)^2 + 2(0,15 \times 0,43) = 0,1515$  ( $\equiv p^2 + 2pq$ )

Per al locus  $D1S80$ :  $P_{19/29} = 2(0,10 \times 0,33) = 0,066$  ( $\equiv 2pq$ )

Per tant, la raó de versemblança serà:

$$LR = \frac{P(E/C)}{P(E/I)} = \frac{1}{0,1515 \times 0,066} = \frac{1}{0,01} = 100$$

Per als perits genètics, els percentatges de probabilitat es transformen en predicats verbals per tal de comunicar al jutge. En el nostre exemple :

“L'acusat té 100 vegades més probabilitat de ser el culpable que qualsevol altre individu de la població”.

El jutge haurà de multiplicar l'LR per la probabilitat de culpabilitat que es tenia *a priori* sobre l'individu. La **probabilitat *a posteriori*** serà:

$$\frac{\Pr(C | E)}{\Pr(I | E)} = \frac{\Pr(E | C)}{\Pr(E | I)} \times \frac{\Pr(C)}{\Pr(I)}$$

Probabilitat final que siga culpable (davant que no ho siga), coneguda l'evidència científica.

LR

Probabilitat a priori que siga culpable (davant la seua innocència), segons proves del judici.

**Probabilitat *a priori* de culpabilitat**

(basada en altres proves)

1.000 a 1 en contra (1/1000)

100 a 1 en contra (1/100)

10 a 1 en contra (1/10)

Igual (1 a 1)

10 a 1 a favor



**Probabilitat *a posteriori* de culpabilitat**

(després de la prova d'ADN)

10 a 1 en contra

Igual a favor i en contra

10 a 1 a favor

100 a 1 a favor

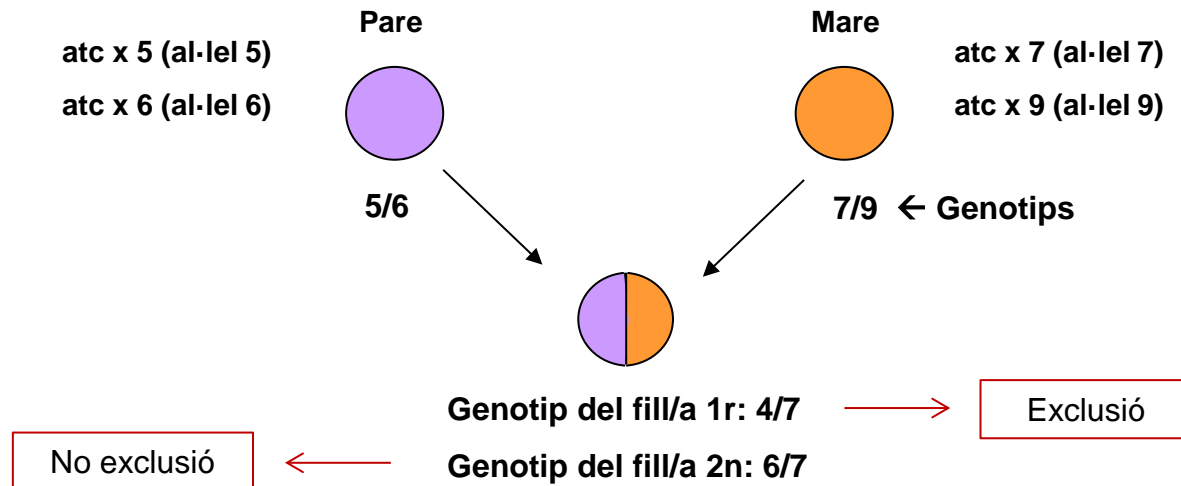
1.000 a 1 a favor

### 3. Assignació de probabilitats en proves de paternitat

En les proves de paternitat, la investigació es fonamenta en el fet que el material genètic del fill/a prové a parts iguals del pare i de la mare.

Per establir paternitat, s'obté el perfil genètic del fill/a, del suposat pare i de la mare, per a diversos marcadors (generalment, STR).

Els resultats poden excloure una persona com a pare d'un xiquet/a. En cas de no exclusió s'ha de fer una anàlisi probabilística que siga el pare biològic.



Per al segon fill cal fer una **anàlisi probabilística de la prova**, que consisteix en contrastar les dues hipòtesis contràries:

- P: El suposat pare és el pare biològic del fill/a.
- $\bar{P}$ : El suposat pare NO és el pare biològic i, per tant, ho és un altre home.

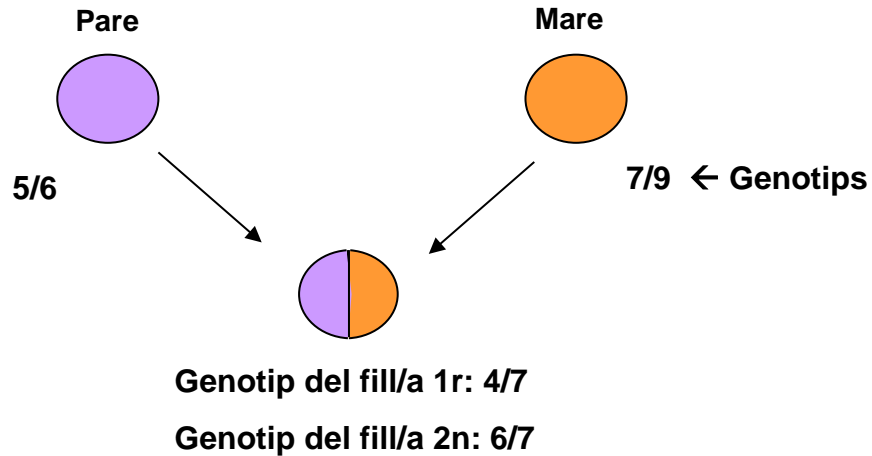
Si anomenem E l'evidència genètica (genotip del fill/a), l'aplicació del Teorema de Bayes implica el càlcul d'un factor, basat en l'evidència, que multiplique aquest grau de creença inicial.

Aquest factor és la raó bayesiana (LR), que en els casos de paternitat rep el nom d'Índex de Paternitat (IP):

$$IP = \frac{\text{Probabilitat de tenir un fill/a amb aquest genotip si SP és el pare}}{\text{Probabilitat de tenir un fill/a amb aquest genotip si el pare és qualsevol altre home}} = \frac{P(E/P)}{P(E/\bar{P})}$$



A l'exemple inicial:



Fill/a 1r:

L'IP seria 0, ja que l'SP no pot ser el pare d'aquest/a fill/a.

Fill/a 2n:

Per poder calcular el denominador, hem de conèixer les freqüències d'aquests al·lels en la població. Suposem que són:

$$f(A4) = 0.33, \quad f(A5) = 0.12, \quad f(A6) = 0.21, \quad f(A7) = 0.09, \quad f(A9) = 0.28$$


$$IP = \frac{P(E/P)}{P(E/\bar{P})} = \frac{\frac{1}{2} \times \frac{1}{2}}{(\frac{1}{2} \times 0.21)} = \frac{0.25}{0.105} = 2,38$$

És 2,38 vegades més probable que siga ell el pare, que no que ho siga un altre individu de la població.

(Si s'analitzaren més loci, l'IP total s'obtidria multiplicant els IP per a cada locus i el resultat seria més resolutiu.)

Abans de recollir l'evidència genètica, s'haurien de tenir unes creences a priori sobre la probabilitat que el suposat pare sigui el pare veritable.

Un cop informat el jutge sobre el valor d'IP, aquest aplicarà el teorema de Bayes per calcular les probabilitats finals que el SP siga el pare, davant que no ho siga:

$$\frac{\Pr(P | E)}{\Pr(\bar{P} | E)} = \frac{\Pr(E | P)}{\Pr(E | \bar{P})} \times \frac{\Pr(P)}{\Pr(\bar{P})}$$


Probabilitat final que siga el pare (davant que no ho siga), coneguda l'evidència.

IP

Probabilitat a priori que siga el pare (davant que no ho siga).

Una altra manera d'interpretar l'IP és convertint-lo a probabilitat de paternitat, representada com a  $W$ , que es defineix com:

$$W = \frac{P(E/P)}{P(E/P) + P(E/\bar{P})} \times 100$$

El valor de  $W$  representa la probabilitat que l'home en qüestió siga el veritable pare del xiquet/a.

L'expressió anterior es pot convertir en una de més senzilla, dividint numerador i denominador per  $P(E/\bar{P})$ :

$$W = \frac{IP}{1+IP} \times 100$$

A l'exemple anterior ( $IP = 2,38$ ), la  $W = 70,4\%$

El perit genètic també pot transmetre al jutge el valor de  $W$  en forma de predicat verbal:

| <b>Probabilitat de paternitat<br/>(<math>W</math>)</b> | <b>Predicat verbal</b>            |
|--|-----------------------------------|
| Major del 99,7 %                                       | Paternitat pràcticament provada   |
| De 99,0 al 99,7 %                                      | Paternitat altament probable      |
| De 95,0 al 99,0 %                                      | Paternitat molt probable          |
| De 90,0 al 95,0 %                                      | Paternitat probable               |
| De 80,0 al 89,9 %                                      | Certa insinuació de paternitat    |
| Menor del 80 %   | Paternitat despreciable o no útil |