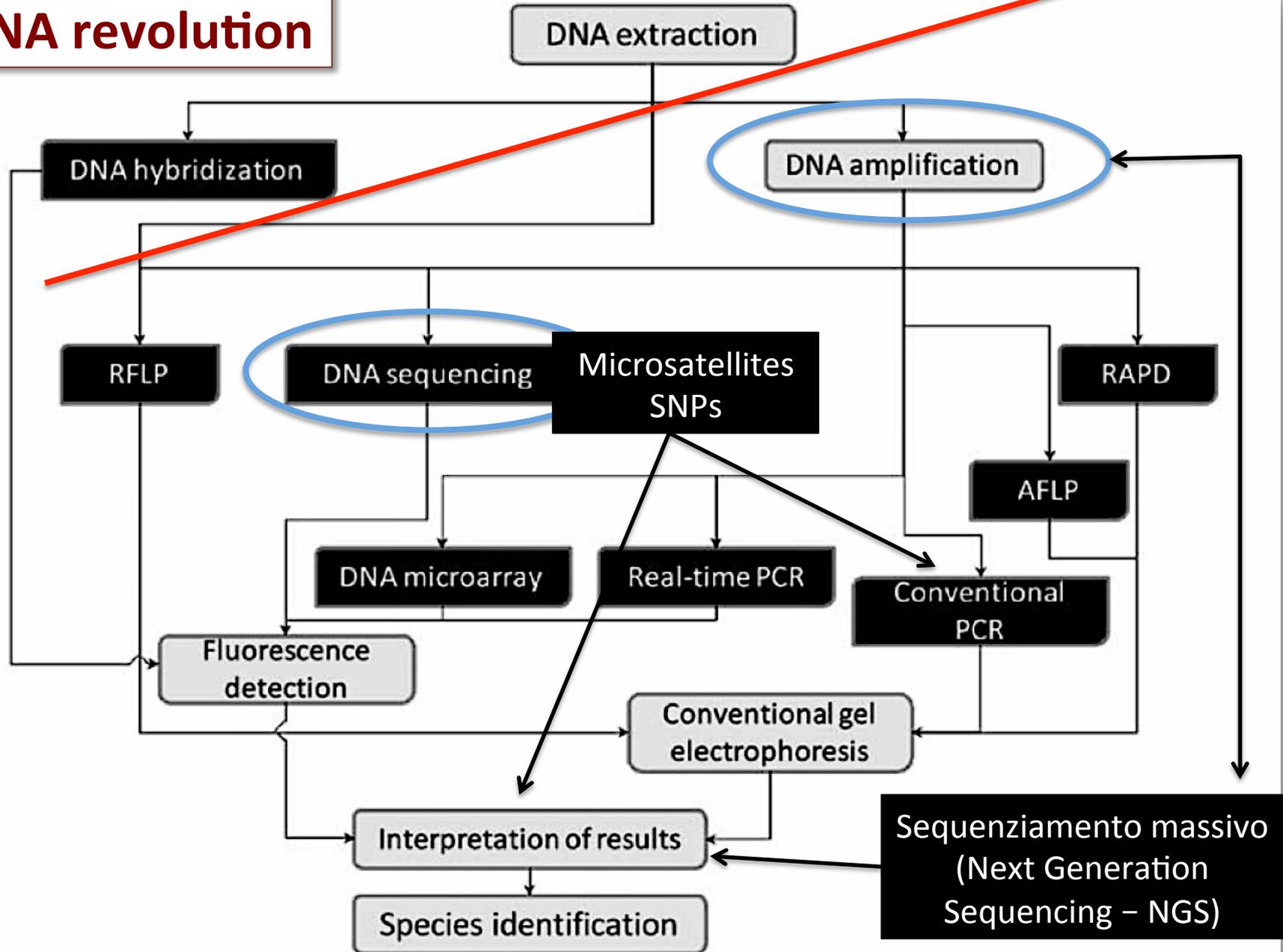


# DNA revolution

*I marcatori molecolari e il loro contributo alle scienze forensi (II parte)*



# DNA revolution



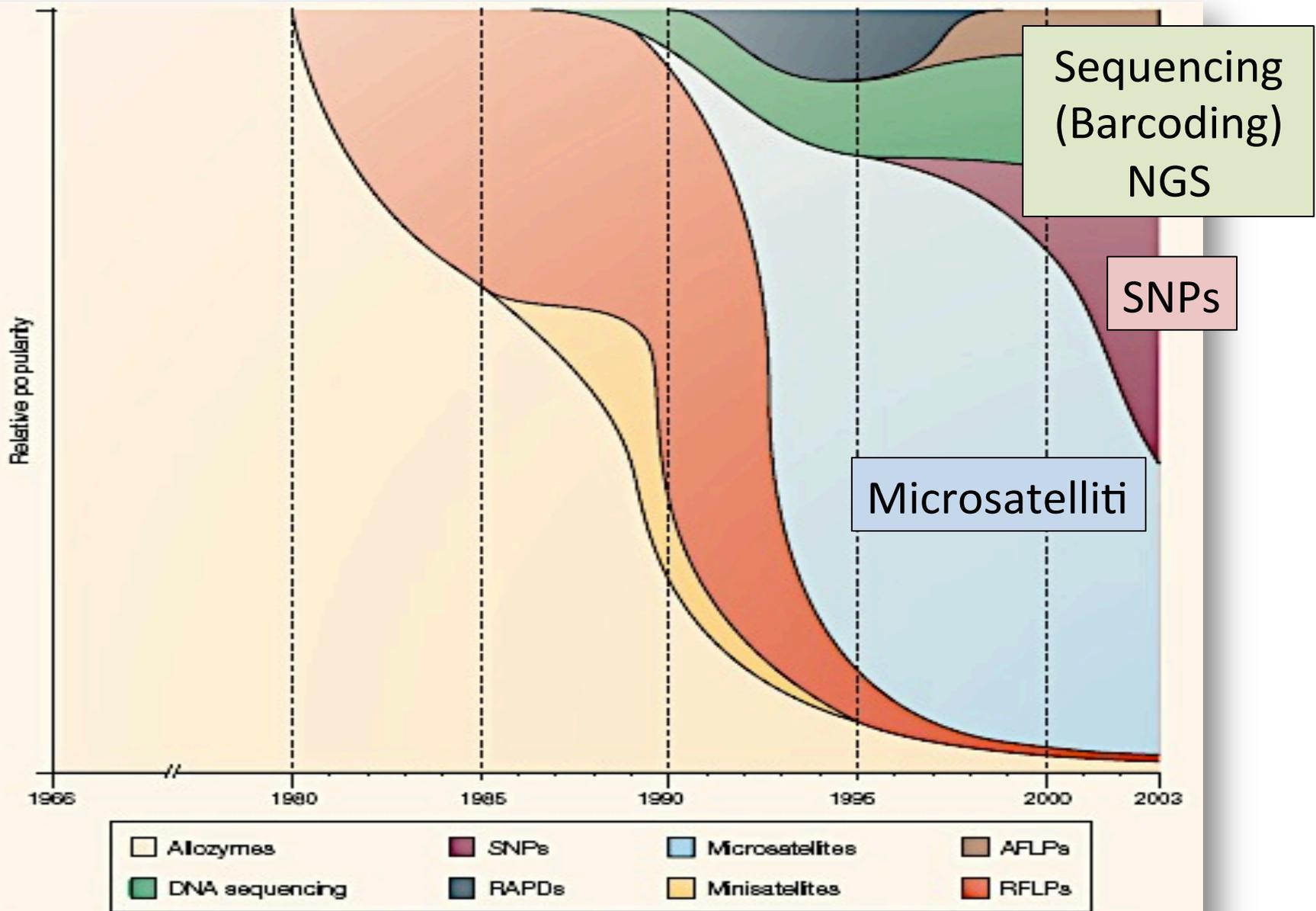


Figure 1 | Subjective view of the the changing relative importance of different molecular markers.

# Sequenziamento

Principio base: Determinazione della sequenza nucleotidica di DNA

Anni '70

Metodo di  
degradazione chimica  
di Maxam e Gilbert

Metodo Enzimatico  
**SANGER**  
di Fred Sanger

Generazione di **popolazioni di frammenti di DNA marcati radioattivamente che iniziano da un punto fisso e terminano in punti diversi (coprendo l'intera lunghezza del frammento)**

**Lettura per autoradiografia dopo corsa elettroforetica**

# Sequenziamento – Metodo Sanger

## Metodo di terminazione della catena con dideossi

DNA polimerasi

Miscela di deossi-nucleotidi (dNTPs = dATP, dCTP, dGTP, dTTP)

Miscela di dideossi-bucleotidi (ddNPs) in **bassa concentrazione**

Primer marcato



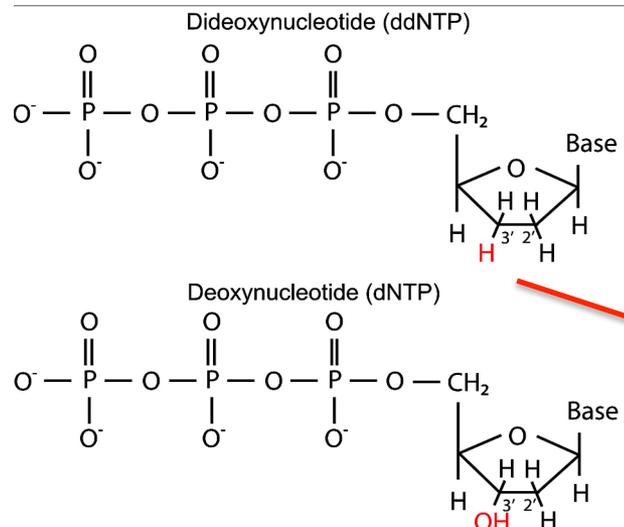
Filamento copia



DNA stampo a singolo filamento (stDNA)

**ddNTPs =**

Nucleotidi **privi del gruppo ossidrilico** in 3' necessario per l'elongazione della catena



Non permette il legame fosfodiesterico con il successivo deossinucleotide, interrompendo la reazione di sintesi

# Sequenziamento – Metodo Sanger

4 reazioni di sequenza

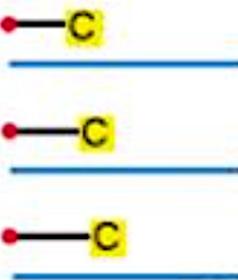
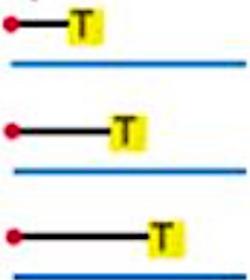
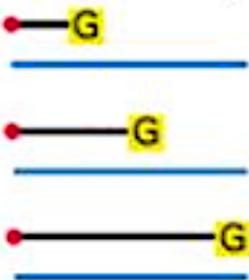
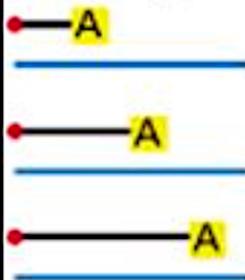
5' —  
3' ————— 5'  
DNA polymerase  
+ dNTPs (100 μM)

+ Dideoxy A  
(1 μM)

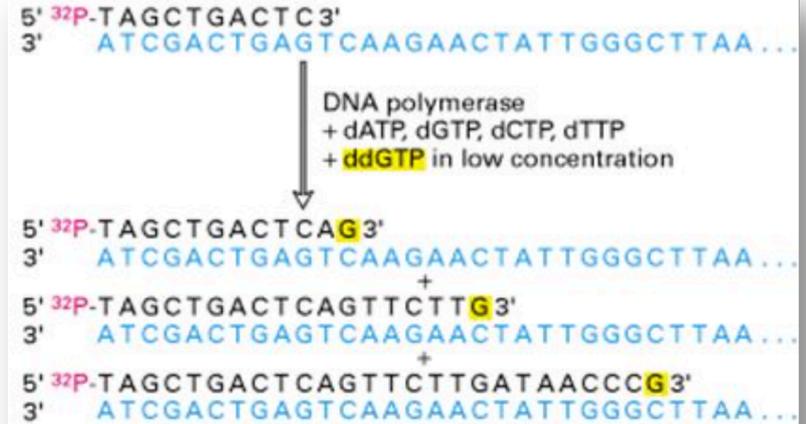
+ Dideoxy G  
(1 μM)

+ Dideoxy T  
(1 μM)

+ Dideoxy C  
(1 μM)



← Denature and separate by electrophoresis →

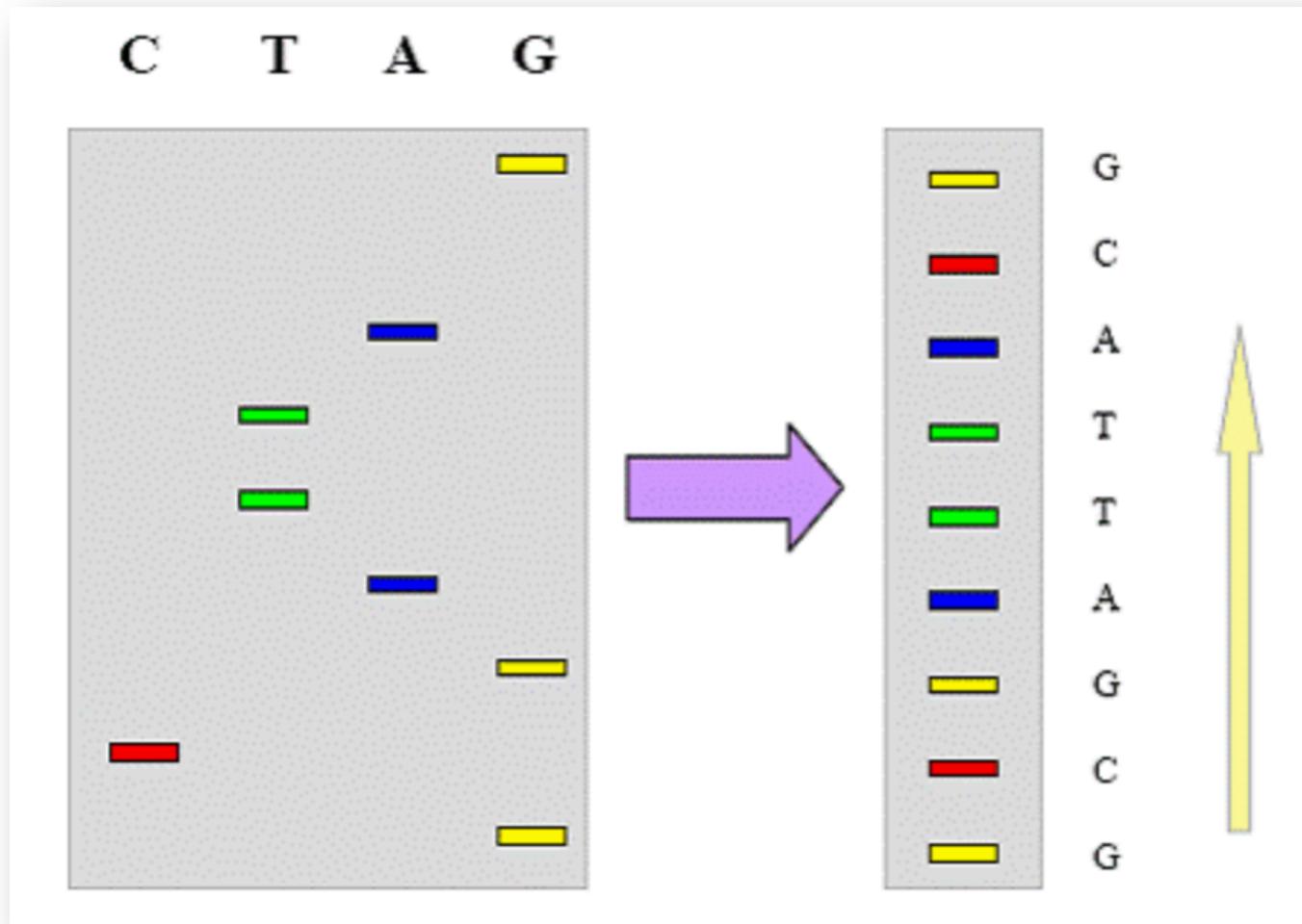


Per ogni reazione si formeranno popolazioni di frammenti terminanti in tutte le posizioni dove è presente quel nucleotide

## Sequenziamento – Metodo Sanger

Le molecole ottenute da ogni reazione vengono caricate su gel in 4 corsie parallele

La corsa viene letta per autoradiografia grazie alla presenza del primer marcato in ogni frammento



## Sequenziamento – Metodo Sanger

### Automatizzazione del sequenziamento Sanger

Uso di 4 fluorocromi per marcare i 4 ddNTPs

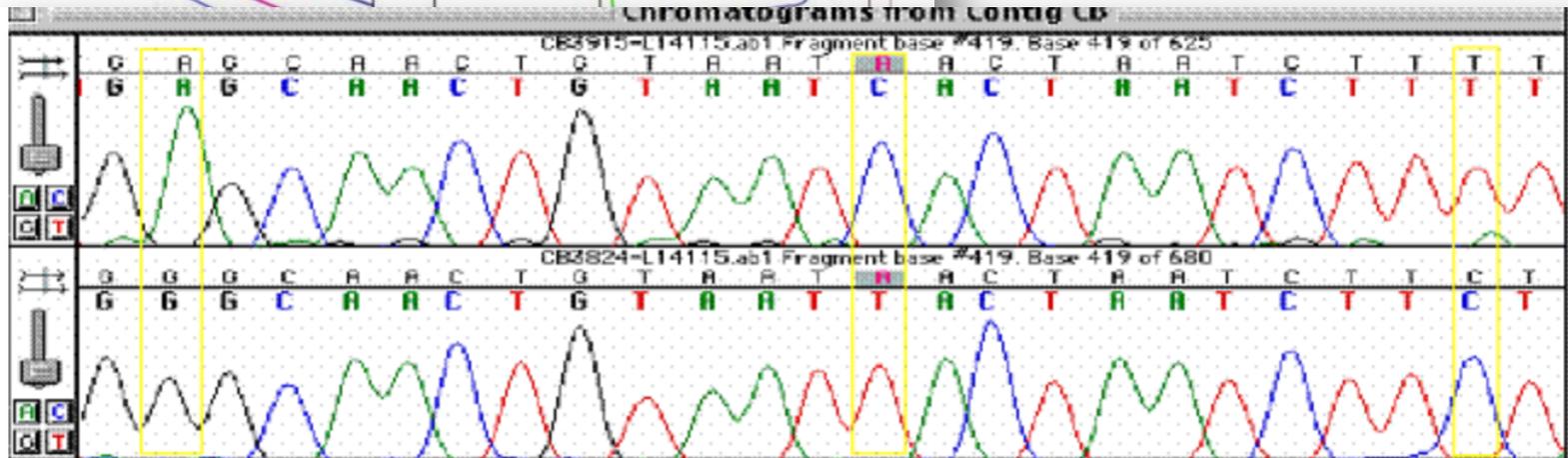
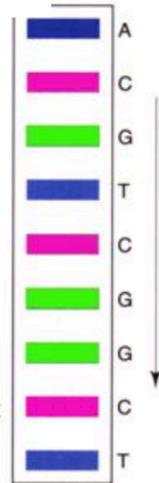
Sequenziamento in un'UNICA REAZIONE

- ✓ Minor quantità di DNA stampo
- ✓ Ridotta introduzione di impurità nella miscela di reazione

I prodotti fluorescenti della reazione di terminazione migrano lungo una singola corsia su cui passa il rivelatore

Unità laser di eccitamento

Griglia di diffrazione



# Sequenziamento



## Vantaggi:

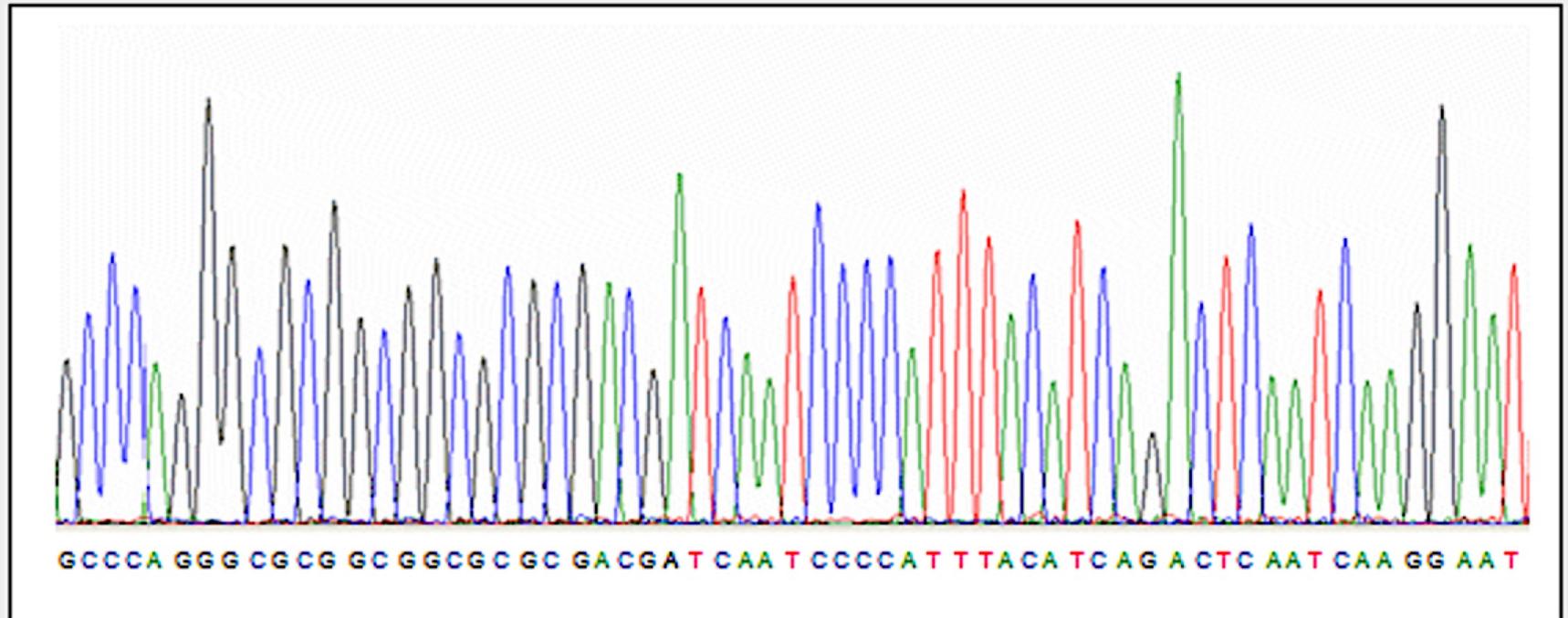
- Alta riproducibilità e affidabilità
- **Utile negli screening genomici e per la caratterizzazione di microsatelliti e SNPs**
- **Strumento base per il DNA barcoding**
- **Possibilità di cross-amplificazione con primer universali**



## Svantaggi:

- Time-consuming and expensive
- Richiede informazioni a priori (primer)
- Low throughput (possibilità di sequenziare una sequenza per volta)
- Difficoltà nell'interpretazione di alcuni risultati ambigui

# Problemi interpretativi del sequenziamento



In un elettroferogramma ogni picco corrisponde ad una base azotata e il colore del picco individua la singola base (A, G, C o T)

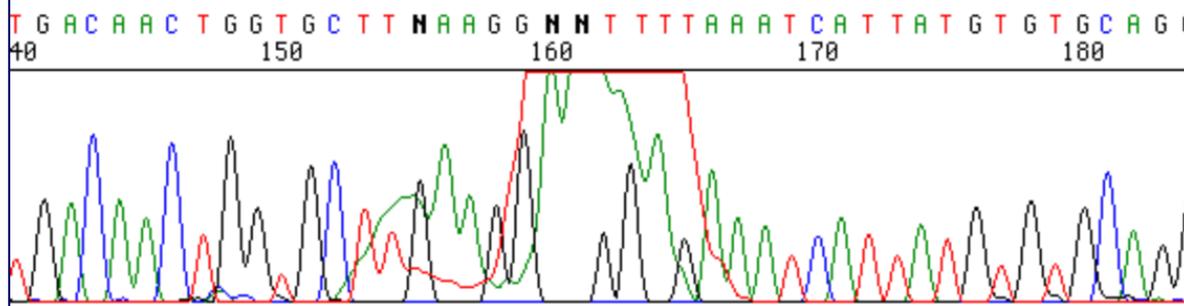
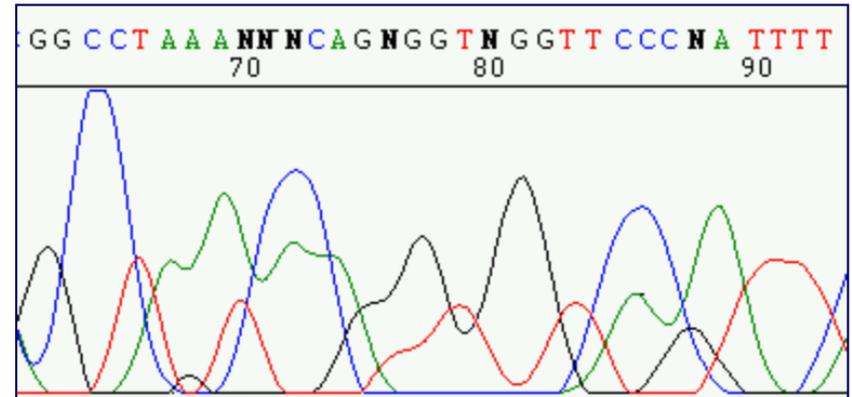
No rumore di fondo

Tutti i picchi sono stati correttamente letti dal software

# Problemi interpretativi del sequenziamento

## *Pill-up peaks*

Presenza di picchi di altri colori sotto il picco principale dovuti al sequenziamento di un quantitativo eccessivo di DNA



## *Dye blobs*

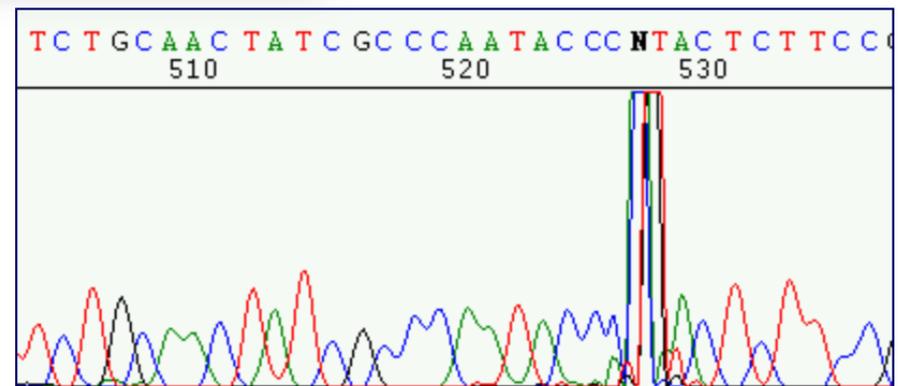
Artefatto dato da picchi ampi di un colore sovrastanti il picco corretto.

Dovuti a una non corretta purificazione (residui di dNTPs)

## *Spikes*

Artefatti dati da picchi multicolore alti e stretti che mascherano il nucleotide sottostante.

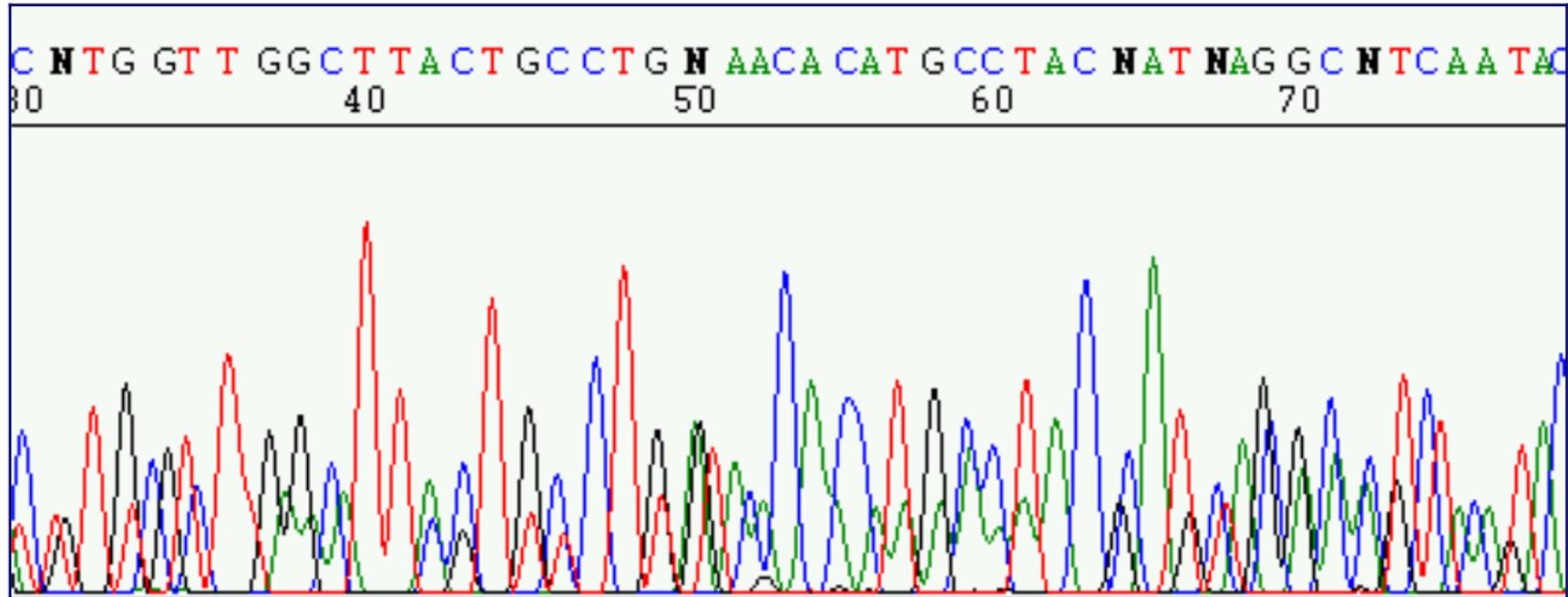
Dovuti a presenza di bolle d'aria nel capillare di sequenziamento.





# Problemi interpretativi del sequenziamento

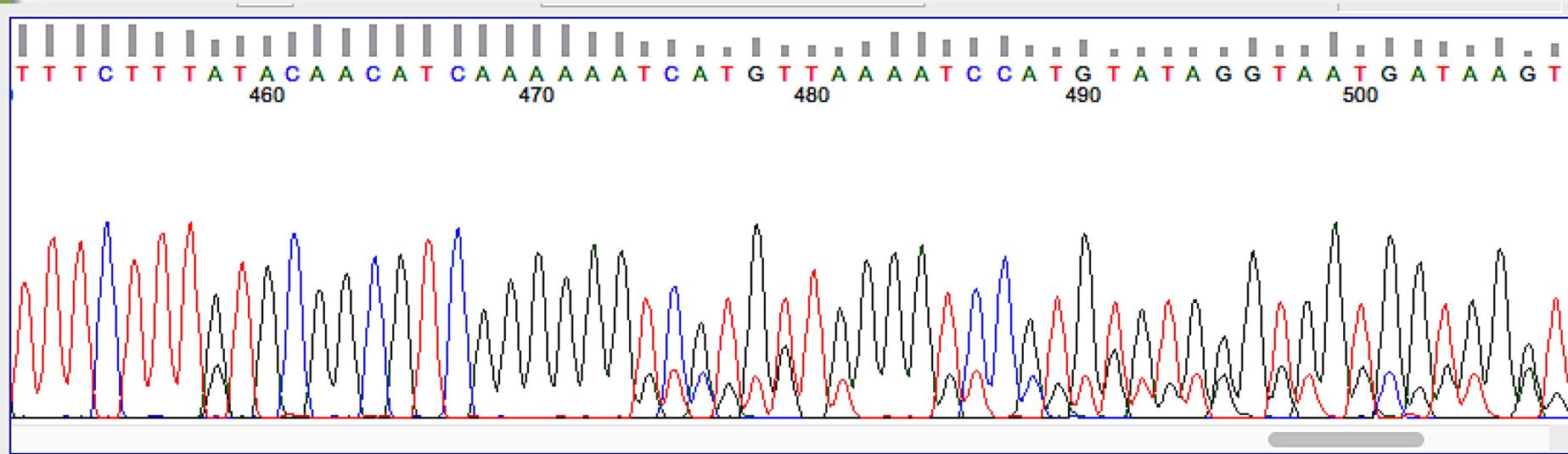
## *Contaminazione da DNA esogeno*



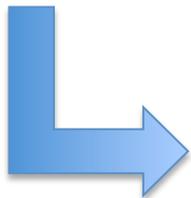
Presenza di picchi sovrapposti e disordinati  
Sequenza non interpretabile

# Problemi interpretativi del sequenziamento

## *Presenza di varianti alleliche*



Presenza di più prodotti di PCR (reali) dovuta alla simultanea amplificazione di alleli diversi allo stesso locus (es. individuo eterozigote a quel locus, campione contaminato da materiale biologico di individui diversi)



**CLONAGGIO** dei prodotti di PCR

# CLONAGGIO

Inserzione del frammento di interesse in un DNA circolare (**DNA plasmidico**)

Inserzione del DNA plasmidico in una **cellula batterica**

- Il plasmide viene così duplicato dal batterio
- Tutti i batteri discendenti da una singola cellula presenteranno copia del DNA esogeno

In studi di **genetica forense** e genetica della conservazione il clonaggio di frammenti di DNA è una tecnica cruciale quando frammenti diversi coesistono nello stesso campione, consentendone la separazione e la caratterizzazione

## A COSA SERVE il CLONAGGIO del DNA??

→ a creare 'copie identiche' (= CLONI) di un frammento di DNA

### DIFFERENZE con la PCR:

- è una reazione in vivo

→ sfrutta l'apparato di replicazione del DNA delle cellule batteriche

→ si fanno duplicare le cellule

- non necessita nessuna conoscenza a priori della sequenza del frammento di DNA da clonare

→ non si usano primers

- è un processo più accurato

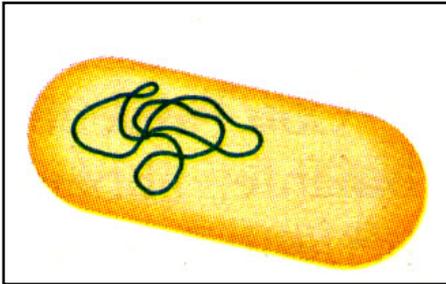
→ le copie di DNA ottenute sono esattamente identiche l'una all'altra

→ sistemi di 'correzione degli errori' cellulari

# COME FUNZIONA il CLONAGGIO del DNA?

## PREMESSE:

### 1) I BATTERI



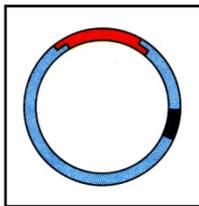
→ CELLULE PROCARIOTICHE:

assenza nucleo

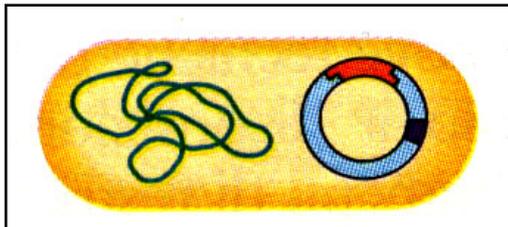
singola molecola circolare di DNA (3-4 M bp) nel citoplasma

→ possono acquisire facilmente piccole molecole di DNA esogeno

### 2) I PLASMIDI



→ piccole molecole circolari di DNA (3000-4000 bp) che portano pochi geni



→ possono facilmente essere acquisiti dalle cellule batteriche

# COME FUNZIONA il CLONAGGIO del DNA?

## PREMESSE:

### 3) GLI ENZIMI DI RESTRIZIONE

ENDONUCLEASI che tagliano il DNA in corrispondenza di specifiche sequenze di basi

#### Estremità piatte o blunt



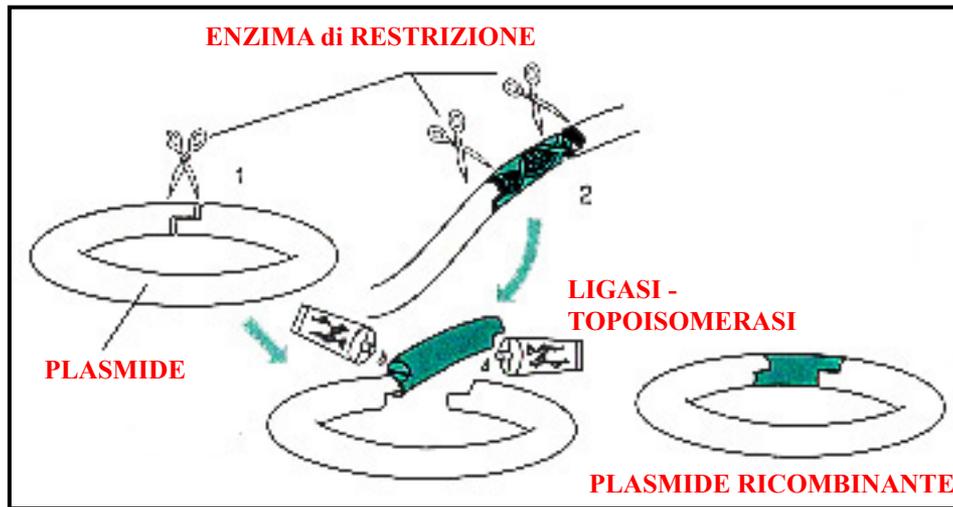
#### Estremità coesive o sticky



Caratteristiche di alcune endonucleasi di restrizione.

	Nome dell'enzima	Organismo in cui è stato trovato	Sequenze di riconoscimento e posizione del taglio *	λ	Numero di siti di taglio in DNA da pBR322
Enzimi con sequenze di riconoscimento di 6 bp	<i>Bam</i> HI	<i>Bacillus amyloliquefaciens</i> H	5' GGATCC 3' 3' CCTAGG 5'	5	1
	<i>Bgl</i> II	<i>Bacillus globigi</i>	AGATCT TCTAGA	5	0
	<i>Eco</i> RI	<i>E. coli</i> RY13	GAATTC CTTAAG	5	1
Enzimi con sequenze di riconoscimento di 4 bp	<i>Hae</i> II	<i>Haemophilus aegyptius</i>	RGCGCY YCGCGR	>30	11
	<i>Hind</i> III	<i>Haemophilus influenzae</i> R <sub>s</sub>	AAGCTT TTCGAA	6	1
	<i>Pst</i> I	<i>Providencia stuartii</i>	CTGCAG GACGTC	18	1
	<i>Sal</i> I	<i>Streptomyces albus</i> G	GTCGAC CAGCTG	2	1
	<i>Sma</i> I	<i>Serratia marcescens</i>	CCCGGG GGGCCC	3	0
	<i>Hae</i> III	<i>Haemophilus aegyptius</i>	GGCC CCGG	>50	22
	<i>Hha</i> I	<i>Haemophilus hemolyticus</i>	GCGC CGCG	>50	31
Enzimi con sequenze di riconoscimento di 8 bp	<i>Hpa</i> II	<i>Haemophilus parainfluenzae</i>	CCGG GGCC	>50	26
	<i>Sau</i> 3A	<i>Staphylococcus aureus</i> 3A	GATC CTAG	116	22
	<i>Not</i> I	<i>Nocardia otitidis-caviarum</i>	GCGGCCGC CGCCGGCG	0	0

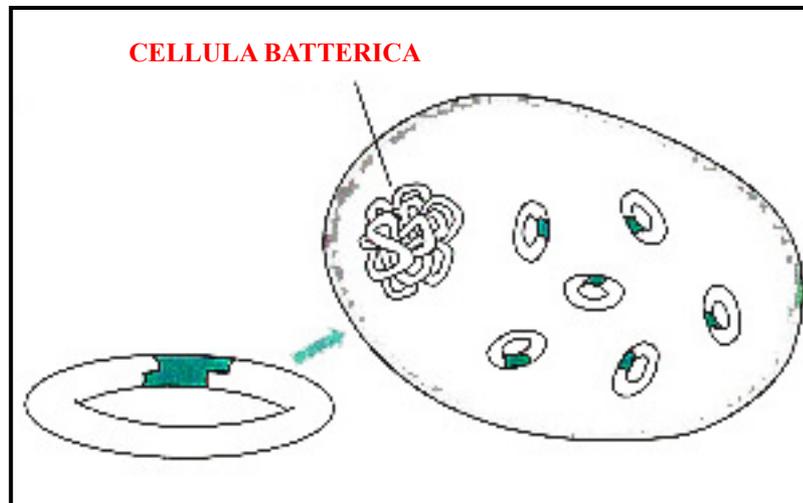
# COME FUNZIONA il CLONAGGIO del DNA?



1. Il DNA da clonare viene tagliato con un **enzima di restrizione**

2. inserito in un **plasmide** tagliato con lo stesso enzima

3. ad opera della LIGASI o della TOPOISOMERASI



4. Il PLASMIDE RICOMBINANTE viene inserito in una **cellula batterica** e al suo interno si duplica

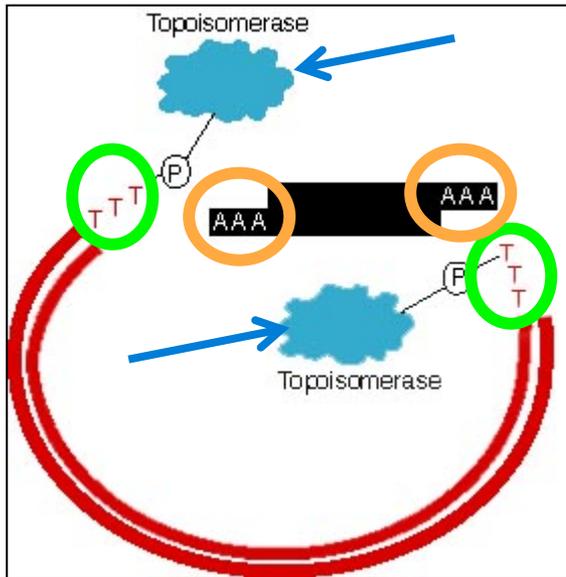
## PUO' SUCCEEDERE CHE...

- 1) La cellula non acquisisca il plasmide e quindi neanche il DNA da clonare
- 2) Il plasmide si richiuda su se stesso senza acquisire il DNA da clonare  
→ la cellula acquisisce il plasmide ma NON il DNA da clonare

## STRATEGIE per VERIFICARE che il CLONAGGIO è AVVENUTO:

- 1) Il plasmide contiene un gene che conferisce alla cellula la resistenza ad un antibiotico  
→ le cellule vengono fatte crescere (e duplicare) in un terreno selettivo contenente quell'antibiotico: **SOLO le CELLULE CONTENENTI il PLASMIDE POSSONO DUPLICARSI** nel terreno contenete l'antibiotico
- 2) Il DNA da clonare viene inserito all'interno di un gene che quando non è interrotto codifica per un enzima (che compie una determinata reazione metabolica) l'interruzione del gene impedisce la sintesi di tale enzima  
→ **CELLULE CHE MOSTRANO il PRODOTTO della REAZIONE CATALIZZATA dall'ENZIMA CONTENGONO IL PLASMIDE MA NON CONTENGONO il DNA CLONATO**

## UN ESEMPIO: il TOPO-TA Cloning

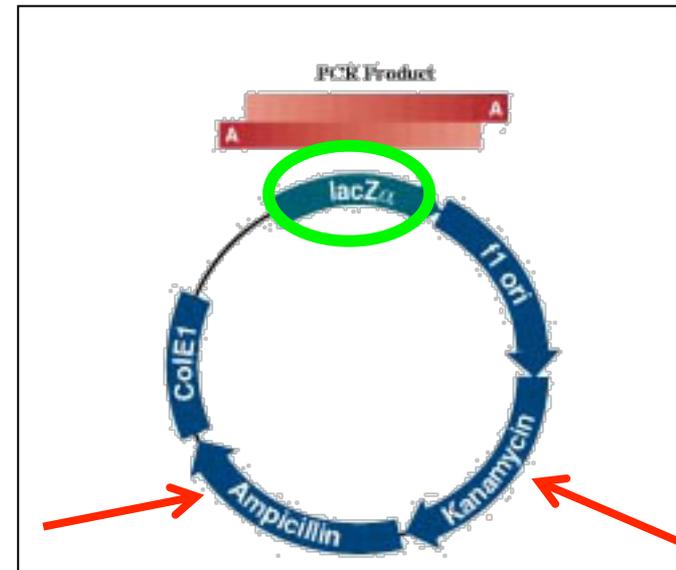


- il plasmide è tagliato in un sito che lascia delle estremità coesive di Timina mentre al frammento da clonare sono aggiunte delle estremità coesive di Adenina (aggiunte in genere dalla Taq polimerasi)

- il legame tra plasmide e DNA da clonare è fatto dall'enzima **Topoisomerasi**

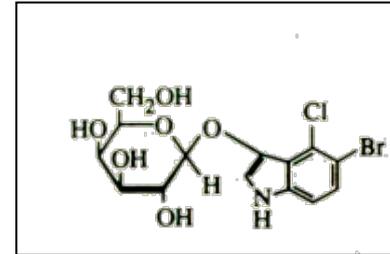
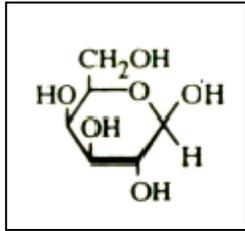
- il plasmide porta i geni per la resistenza agli antibiotici ampicillina e kanamicina

- il sito di inserzione del DNA da clonare spezza la sequenza del gene lacZ ( $\beta$ -galattosidasi)

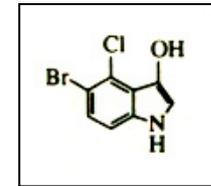


# La $\beta$ -GALATTOSIDASI e l'X-gal

La  $\beta$ -galattosidasi è un enzima coinvolto nella prima tappa del metabolismo del galattosio

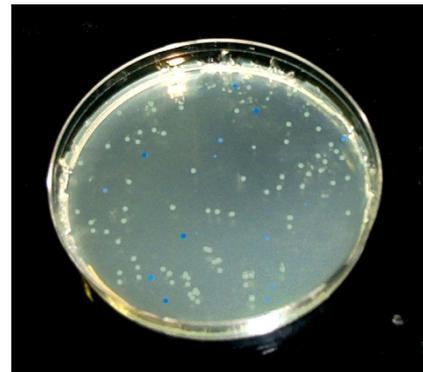
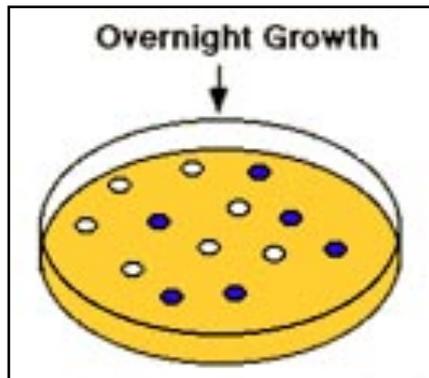
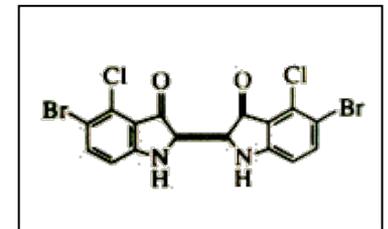


L' X-gal è un omologo del galattosio che viene scisso dalla  $\beta$ -galattosidasi in due prodotti uno dei quali è di colore blu intenso



le cellule in cui viene sintetizzata la  $\beta$ -galattosidasi (il frammento da clonare non si è inserito) appaiono di colore BLU

le cellule senza  $\beta$ -galattosidasi (il frammento da clonare si è inserito) appaiono di colore BIANCO



## IN PRATICA:

### Le FASI del CLONAGGIO

1. Preparazione del terreno di coltura e delle piastre di crescita nelle quali va messo sia l'antibiotico che l'*x-gal*
2. Preparazione del DNA da clonare
  - purificazione
  - creazione delle estremità coesive o "blunt" (nel caso di PCR ci pensa la polimerasi ad aggiungere delle A al 3' e il plasmide è già fornito linearizzato con T terminali)
3. Inserzione del DNA da clonare nel plasmide e ligazione
4. Inserzione del plasmide nelle cellule batteriche → trasformazione mediante shock termico o elettroporazione
5. Piastramento delle cellule e crescita
6. Recupero degli inserti o ampliconi se si usa PCR

# Sequenziamento



## Vantaggi:

- Alta riproducibilità e affidabilità
- **Utile negli screening genomici e per la caratterizzazione di microsatelliti e SNPs**
- **Strumento base per il DNA barcoding**
- **Possibilità di cross-amplificazione con primer universali**

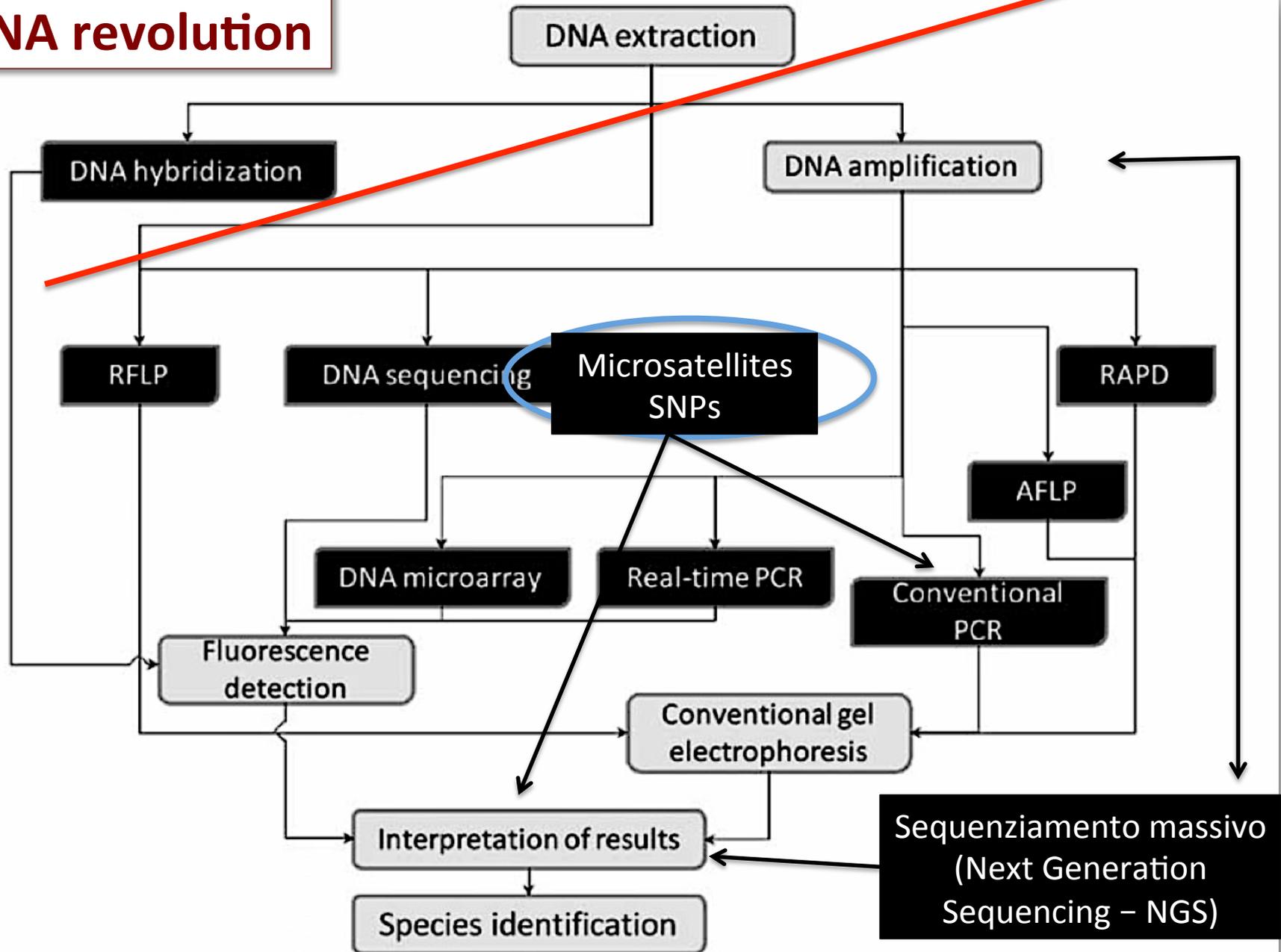


## Svantaggi:

- Time-consuming and expensive
- Richiede informazioni a priori (primer)
- Low throughput (possibilità di sequenziare una sequenza per volta)
- Difficoltà nell'interpretazione di alcuni risultati ambigui

Richiede approfondimenti con tecniche di clonaggio

# DNA revolution



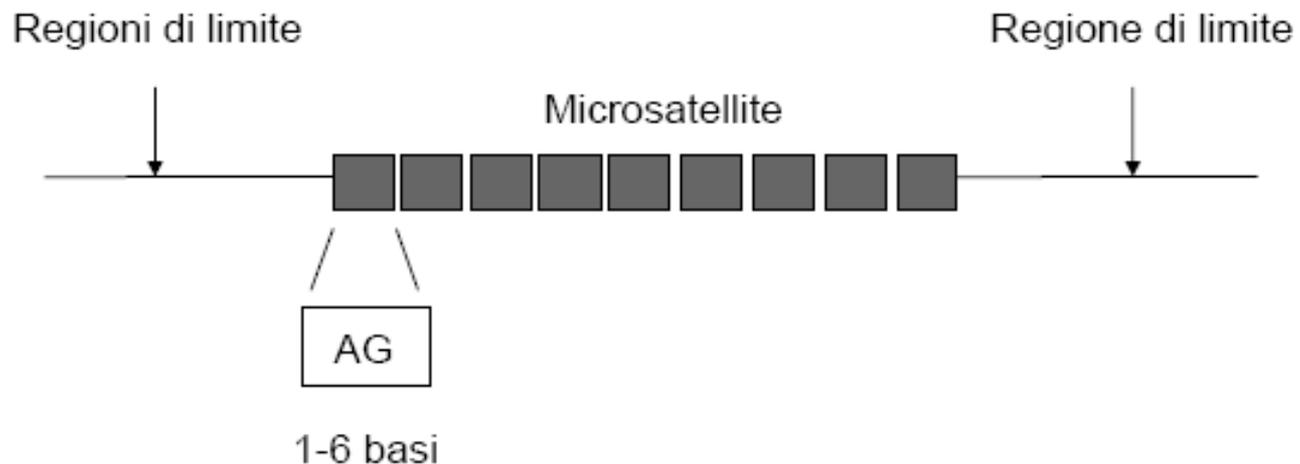
# Microsatelliti

MICROSATELLITI (10-40 sequenze ripetute)

Simple Sequences Repeats (SSRs)

Short Tandem Repeats (STRs)

Simple Sequence Length Polymorphisms (SSLPs)

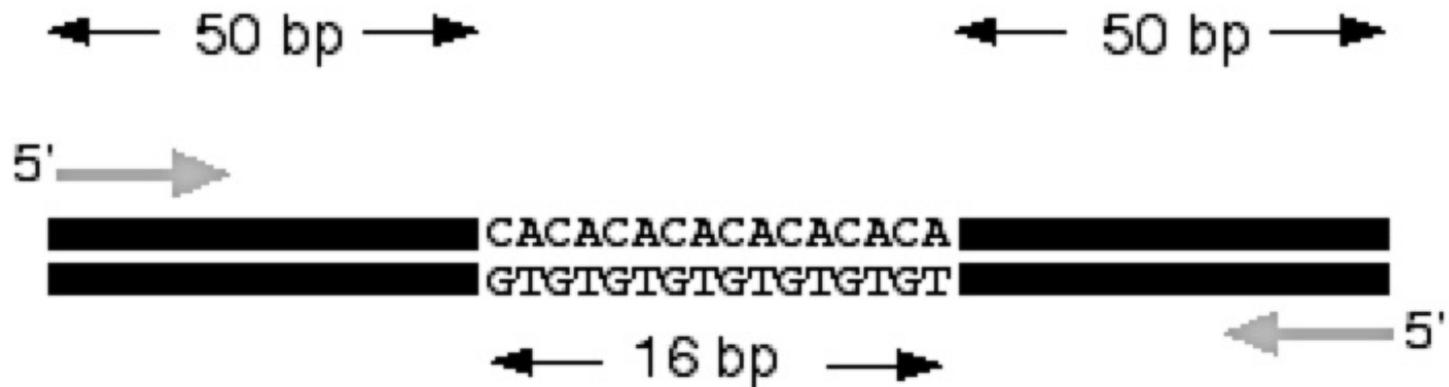


In media si trovano ogni 10,000 basi nel genoma eucariote



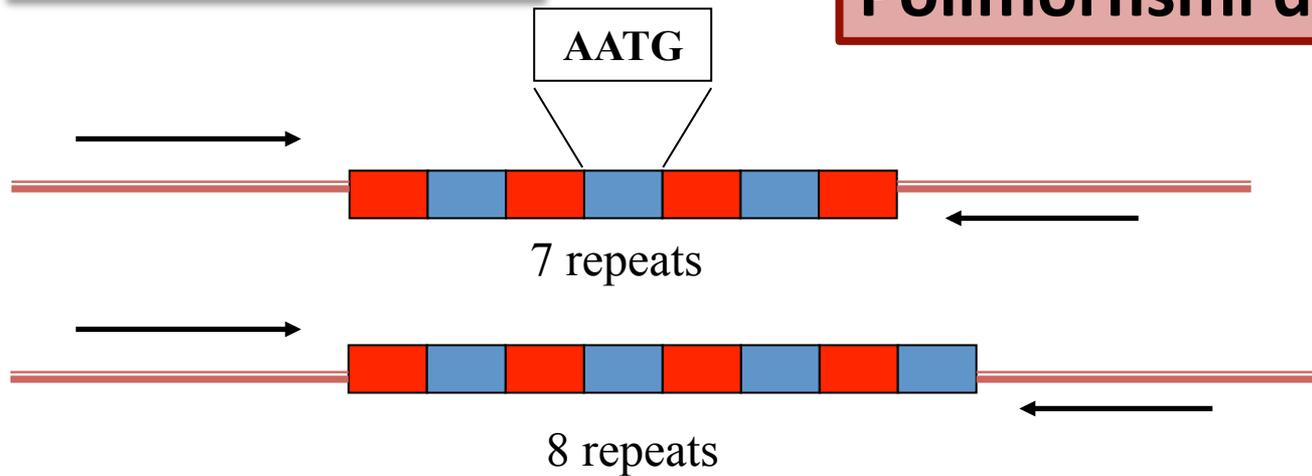
# Microsatelliti

- La variabilità microsatellite viene analizzata tramite PCR
- I primer per la PCR vengono costruiti sulle regioni fiancheggianti la regione ripetuta
- I frammenti vengono separati tramite elettroforesi su gel



# Microsatelliti

## Polimorfismi di LUNGHEZZA



*the repeat region  
is variable  
between samples  
while the flanking  
regions where  
PCR primers bind  
are constant*

Polymorphism is given by the different number of repeats (different length)

**Homozygote** = both alleles are the same length

**Heterozygote** = alleles differ and can be resolved from one another

### COME FACCIAMO A CARATTERIZZARE IL LOCUS??

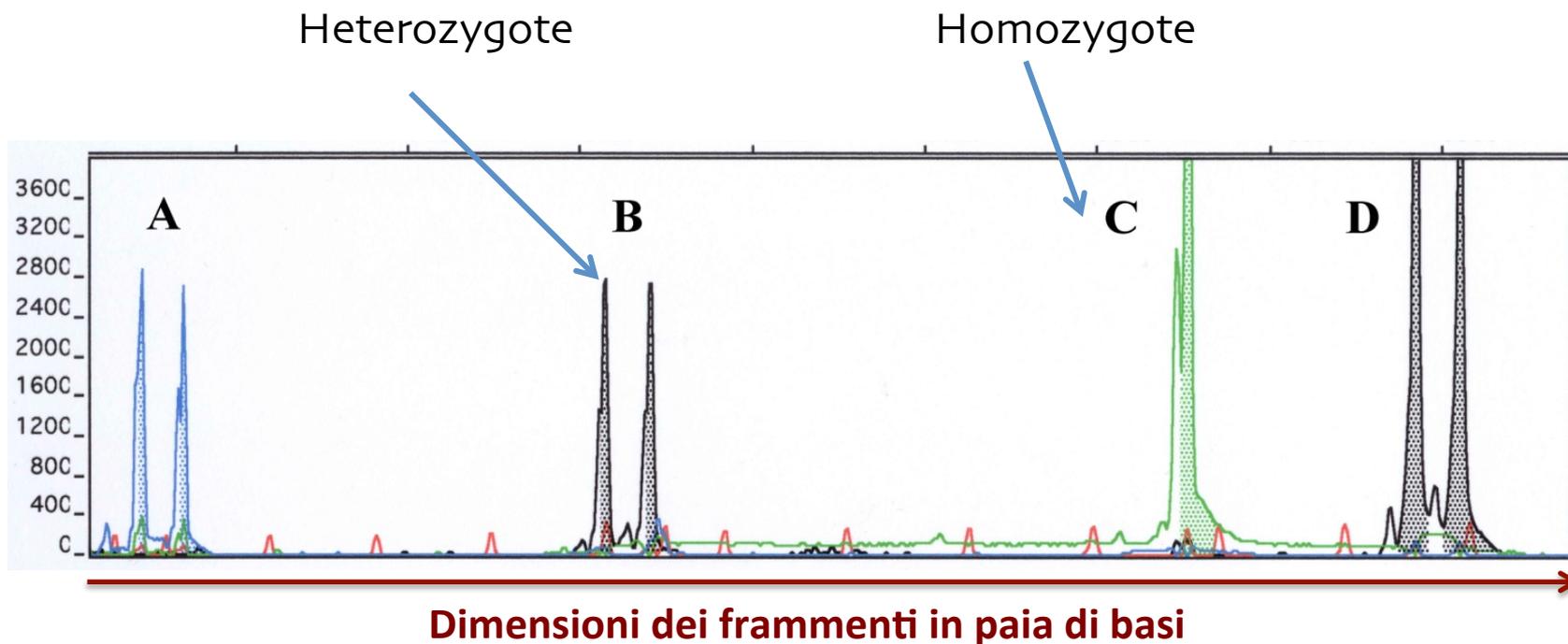
Sequenziamento  
+  
Clonaggio

**GENOTYPING**

# Microsatelliti

- Grazie all'utilizzo di primers marcati con fluorofori la tecnica del **GENOTYPING** (elettroforesi capillare e lettura della fluorescenza) permette la genotipizzazione e l'analisi di loci microsatellite

Conversione dei picchi in GENOTIPI  
(utilizzo di SOFTWARE specifici)

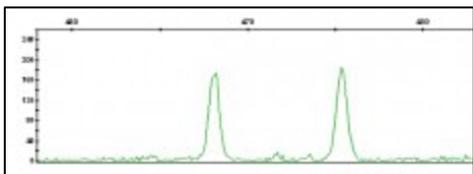


# Microsatelliti

Acquisizione dei dati  
(PCR)

Elettroforesi capillare  
(utilizzo di GeneScan software)

Elettroferogramma



Assegnazione della taglia  
(utilizzo di software)

operatore

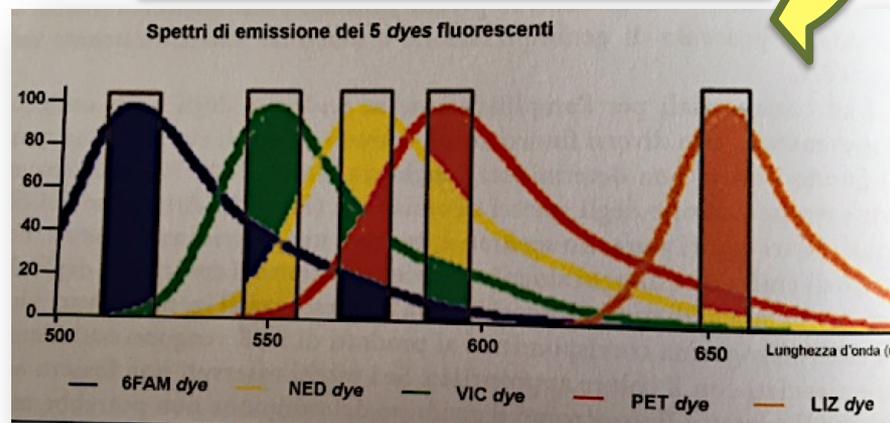
## PROFILO GENETICO

## In pratica . . .

Kit commerciali per  
amplificazione in  
**MULTIPLEX**

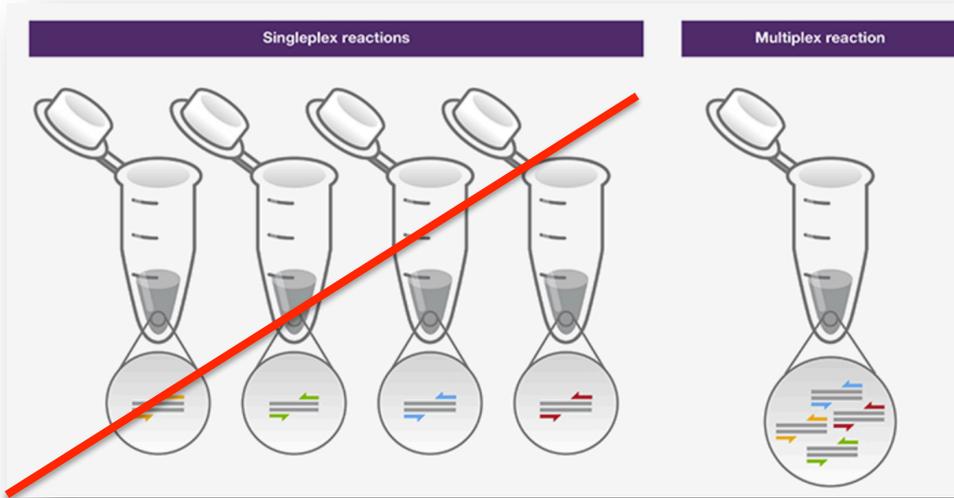
- Identificazione dei picchi (lettura della **fluorescenza**)
- Separazione dei colori

Set di 4 Fluorofori + standard

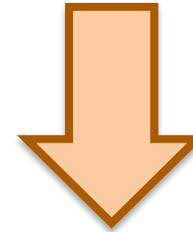


# Multiplexing

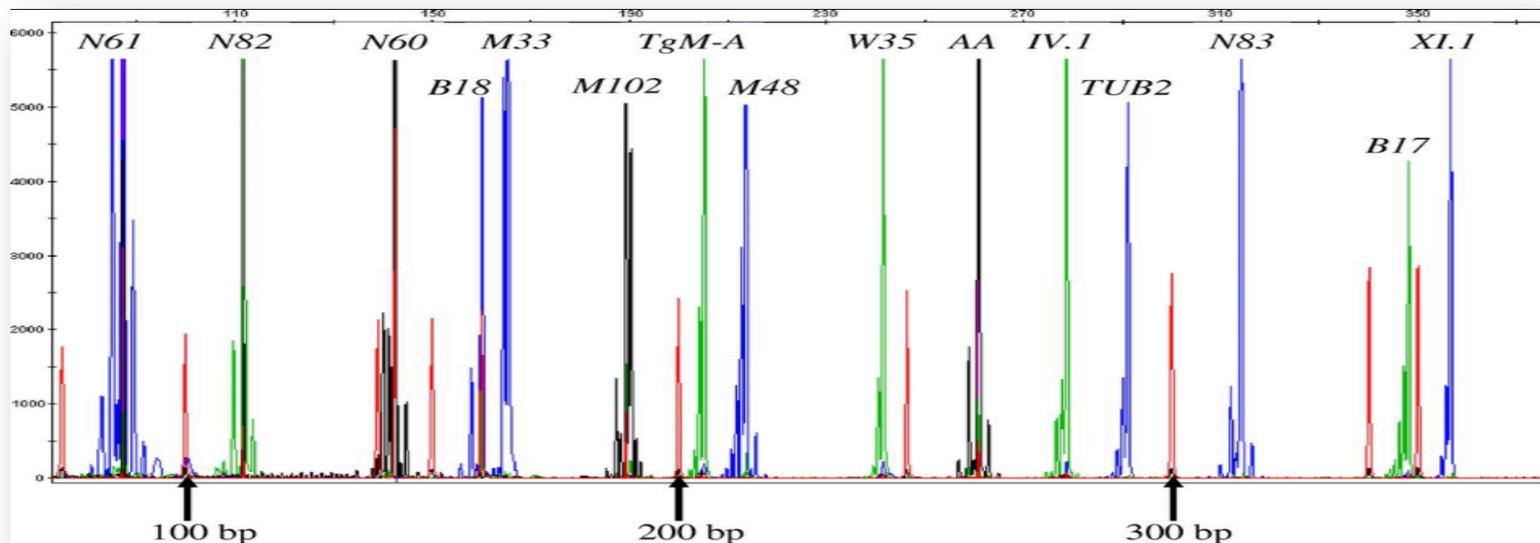
PCR con più coppie di primer con fluorofori differenti



Gel con un pattern di bande non più riconducibili al singolo locus

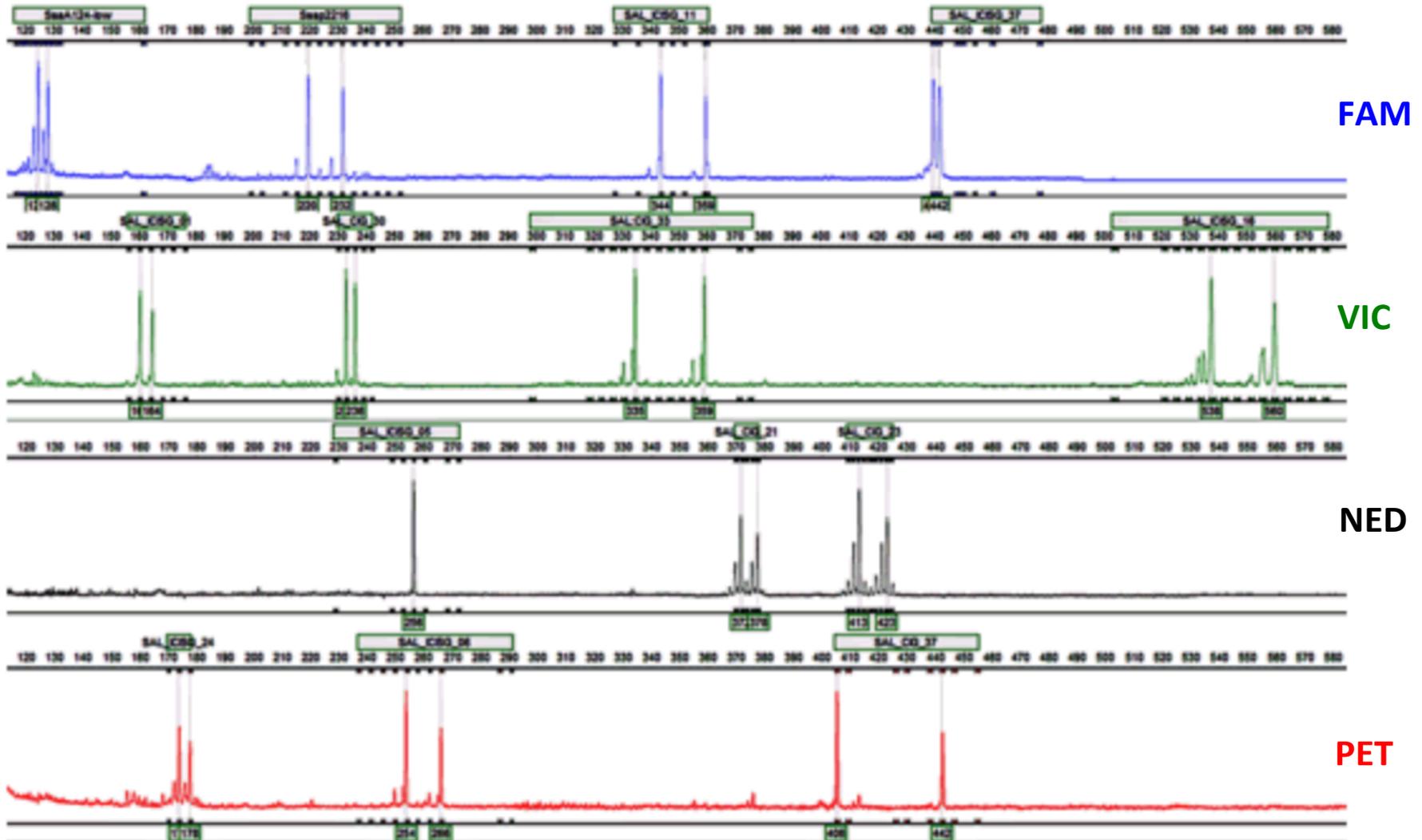


## Elettroferogramma



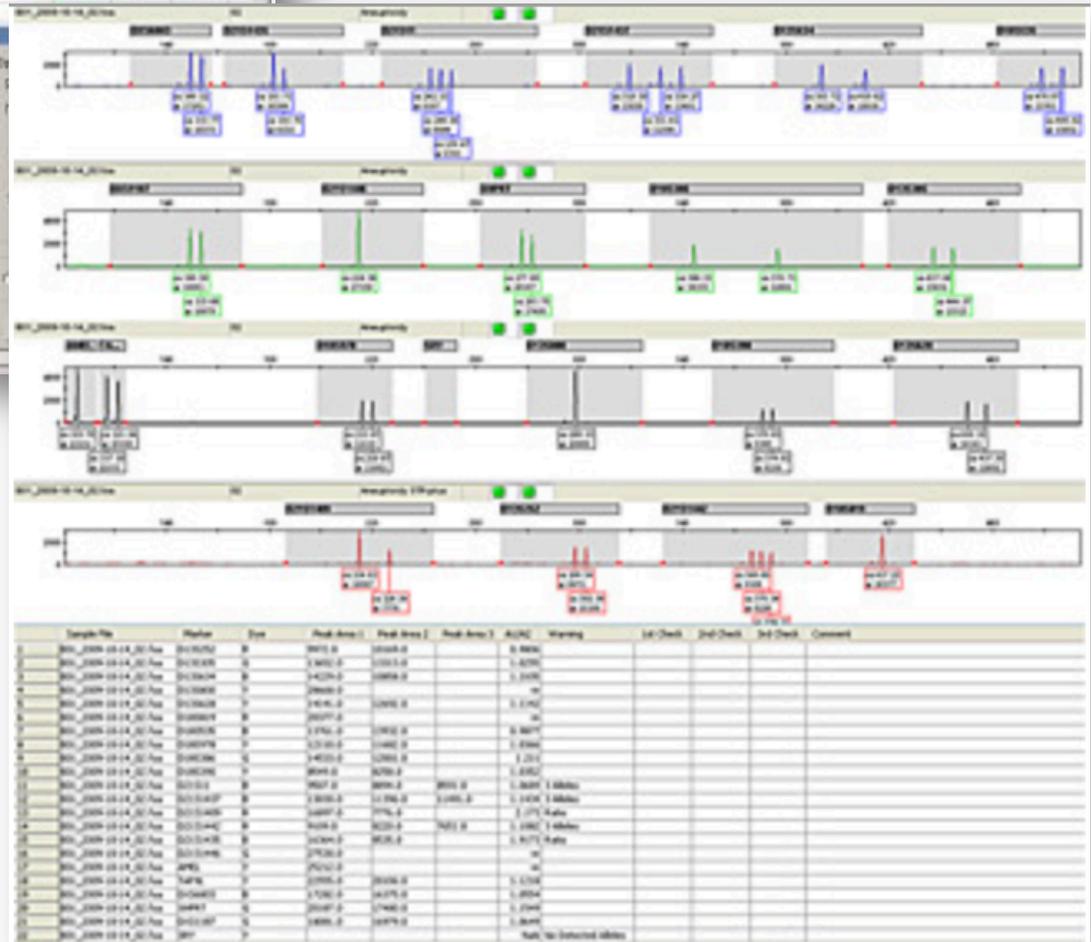
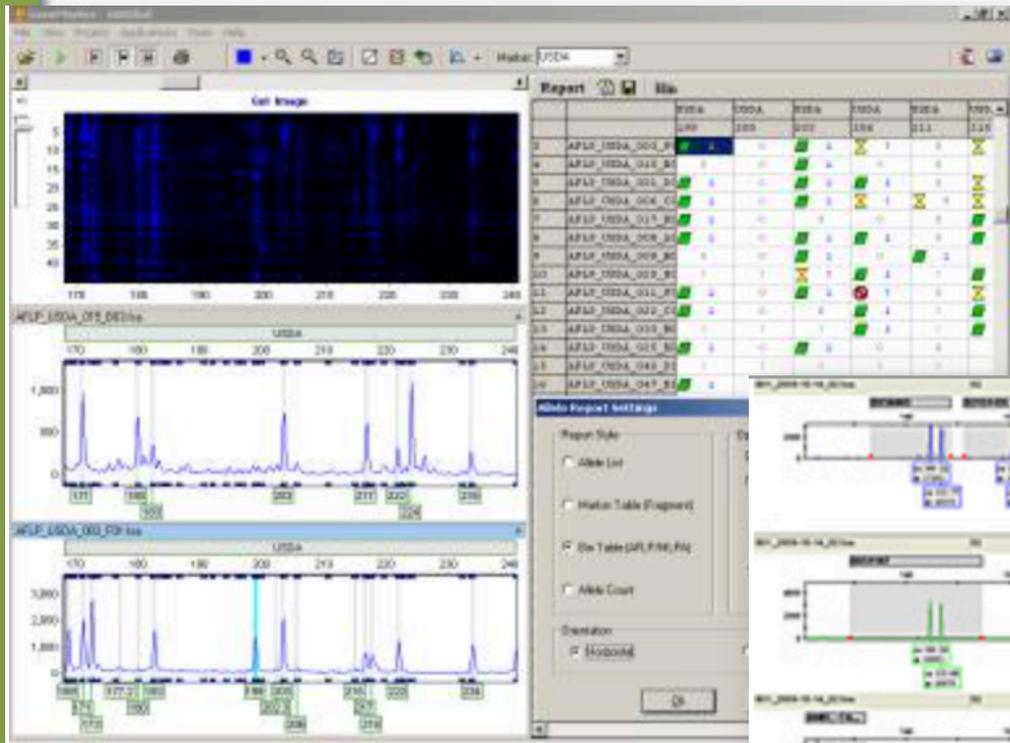
# Multiplexing

Es. Multiplexing di 14 loci microsatellite



GeneMapper, Genotyper

Esempi di software utilizzati per la genotipizzazione di loci microsatellite nelle pratiche forensi



NB: Tutti questi software sono progettati per organismi DIPLOIDI

## Come leggere l'elettroferogramma??

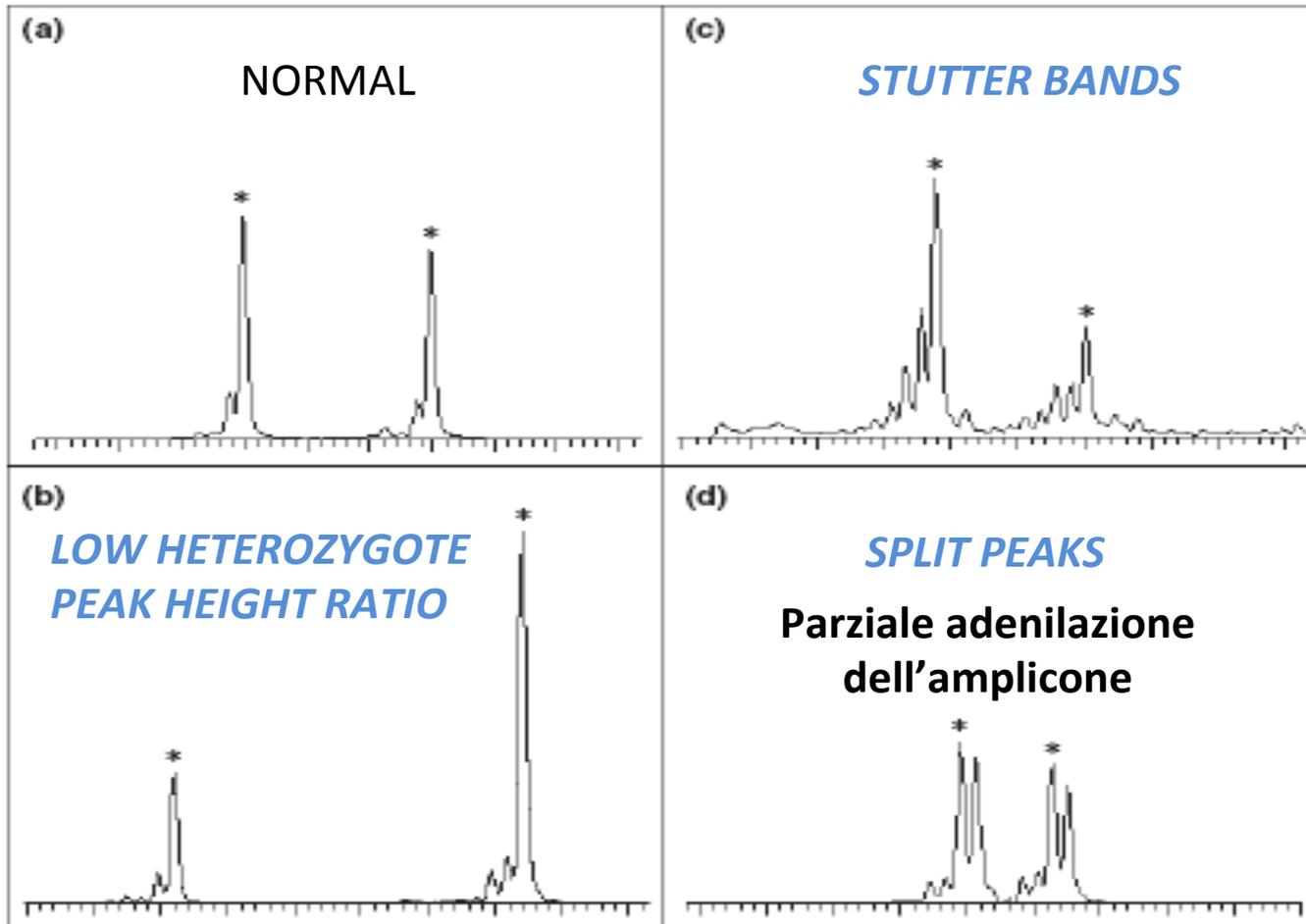
- Conversione dell'elettroferogramma in profilo genetico grazie all'utilizzo di software (*GeneMapper*, *GeneScan*, *Genotyper* ecc).
- Interpretazione esclusiva di personale con esperienza
- Linee guida:
  - Assicurarsi che il locus non crei "rientri" nei colori diversi dal fluoroforo con cui è marcato
  - **N. massimo di picchi per locus = 2 (IN INDIVIDUI DIPLOIDI)**
  - Stabilire un valore minimo come altezza dei picchi da considerare allele
  - Gli alleli del campione non devono differire più di 0.5 bp dal corrispettivo allele dello standard altrimenti sono definiti **OFF-LADDER**
  - L'altezza dei picchi ad un locus deve essere bilanciata (= comparabile)
  - Attenzione alla presenza di **Stutter** (frammenti più corti di un repeat prodotti in fase di PCR in quanto non completati che determinano un picco anomalo di basse dimensioni)

# Come leggere l'elettroferogramma??

Molecular Ecology Resources (2011) 11, 591–611

## Current trends in microsatellite genotyping

E. GUICHOUX,\*†‡ L. LAGACHE,\*† S. WAGNER,\*†§ P. CHAUMEIL,\*† P. LÉGER,\*† O. LEPAIS,\*†¶  
C. LEPOITTEVIN,\*† T. MALAUSA,\*\* E. REVARDEL,\*† F. SALIN\*† and R.J. PETIT\*†



# Problemi nell'interpretazione di profili microsatellite

## Problemi «biologici»

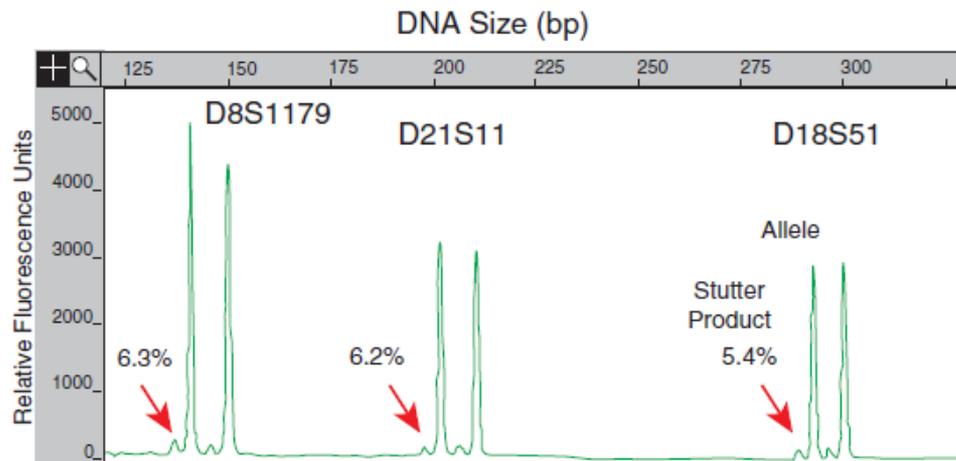
### ➤ **Prodotti *stutter***

Scivolamento della polimerasi durante l'amplificazione di ripetizioni in tandem (fenomeno analogo a quello che porta alla mutabilità in natura dei microsatelliti).



**Alcuni prodotti di PCR più corti di un'unità rispetto al numero di ripetizioni dell'allele amplificato.**

In genere, il prodotto dello stuttering ha un'altezza inferiore al 15% del picco «reale» (almeno per i loci validati in ambito forense).



Altri punti da considerare:

- Differenze tra loci e tra alleli dello stesso locus
- Differenze per lo stesso locus in esperimenti diversi e kit diversi
- Negative and positive stutter
- «Soglia stutter» nell'interpretazione
- Possibili problemi nell'interpretazione di tracce miste.

# Problemi nell'interpretazione di profili microsatellite

## Problemi «biologici»

### ➤ **pattern multiallelici**

Conseguenza della duplicazione della regione che contiene uno dei microsatelliti.

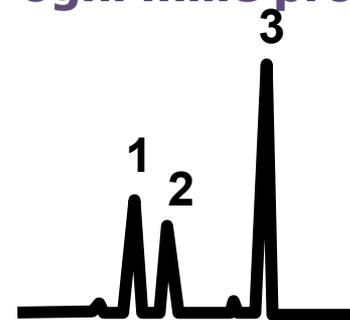
Le duplicazioni possono essere intra- o inter-cromosomiche. In alcuni casi si può trattare di aneuploidie (es. sindrome di Down). Sono stati descritti casi di mosaicismo somatico e germinale.

Per quanto si tratti di un fenomeno **relativamente raro (almeno nei microsatelliti forensi standard)** sono state riportate in letteratura decine di differenti pattern triallelici.

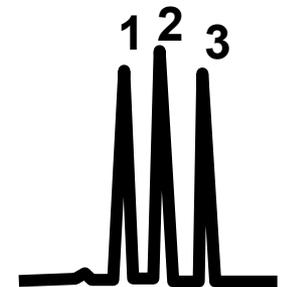
**Si osserva mediamente un triallelico ogni mille profili 15-STR.**

Punti da considerare:

- Differenza rispetto alle misture di campioni
- Problemi per il calcolo delle probabilità
- alleli off ladder estremi
- **Ploidia dell'organismo in esame**



**Type 1  
sbilanciata**



**Type 2  
bilanciata**

# Problemi nell'interpretazione di profili microsatellite

## Problemi «biologici»

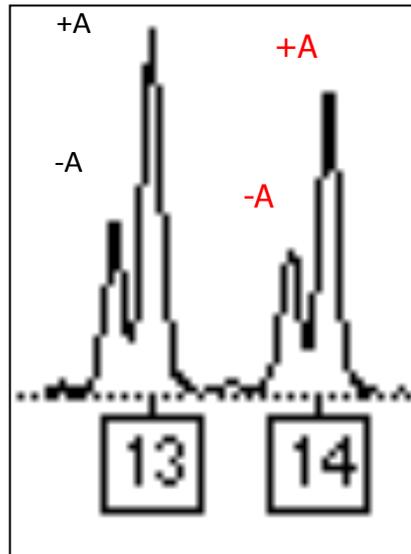
### ➤ Split peaks

Adenilazione incompleta al 3' dei prodotti PCR.

Incomplete adenylation



Per ciascun allele ottengo due picchi distanziati di una sola base nell'elettroferogramma.



D8S1179

Punti da considerare:

-Un picco più corto di una base rispetto al normale potrebbe essere interpretato come una «microvariante»

-STR diversi possono comportarsi in modo diverso

-Soluzioni: lunga estensione finale nella PCR promuove +A; oppure G al 5' del primer non marcato promuove +A

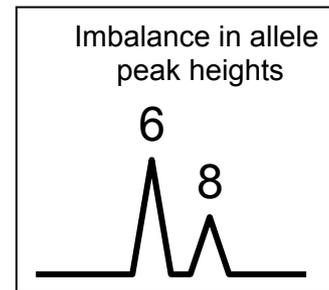
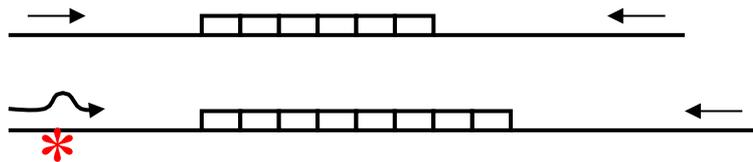
# Problemi nell'interpretazione di profili microsatellite

## Problemi «biologici»

### ➤ Alleli nulli (*Allele drop-out*)

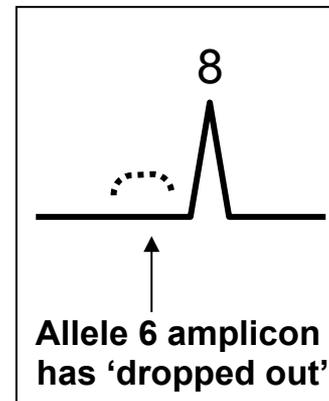
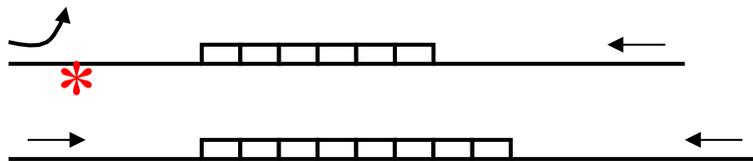
Mutazione nel sito di annealing di un primer (3') che determina il suo mancato attacco e la non amplificazione di uno dei due alleli in un eterozigote

(a)



**Mutation in middle of primer binding site**

(b)



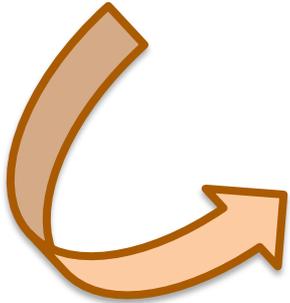
**Mutation at 3'-end of primer binding site**  
**(allele dropout)**

# Problemi nell'interpretazione di profili microsatellite

## Problemi «biologici»

### ➤ **Alleli nulli (*Allele drop-out*)**

Mutazione nel sito di annealing di un primer (3') che determina il suo mancato attacco e la non amplificazione di uno dei due alleli in un eterozigote



Grazie a studi di genetica di popolazione è possibile stabilire la presenza di alleli nulli ad un certo locus grazie all'elevata percentuale di omozigoti

### **Problema frequente in Genetica Forense lavorando con materiale degradato**

Punti da considerare:

- Kit differenti per STR multiplex con diversi set di primers
- Problemi per il confronto con databases: stringenza della ricerca di match
- Problemi nei test di paternità

# Problemi nell'interpretazione di profili microsatellite

## Problemi «tecnici»

### ➤ **“pull up”:**

Incapacità di risolvere appropriatamente nei giusti colori la luce emessa dai fluorocromi utilizzati per marcare i primers (rientri).  
(Vedi slide successiva)

### ➤ **“Dye blobs”:**

Dovuti al distaccamento dei fluorocromi dai primers con conseguente loro migrazione indipendente. (Vedi slide successiva)

### ➤ **Cristalli di urea e bolle d'aria:**

Rilevabili come falsi picchi sottili nell'elettroferogramma.

### ➤ **Spikes:**

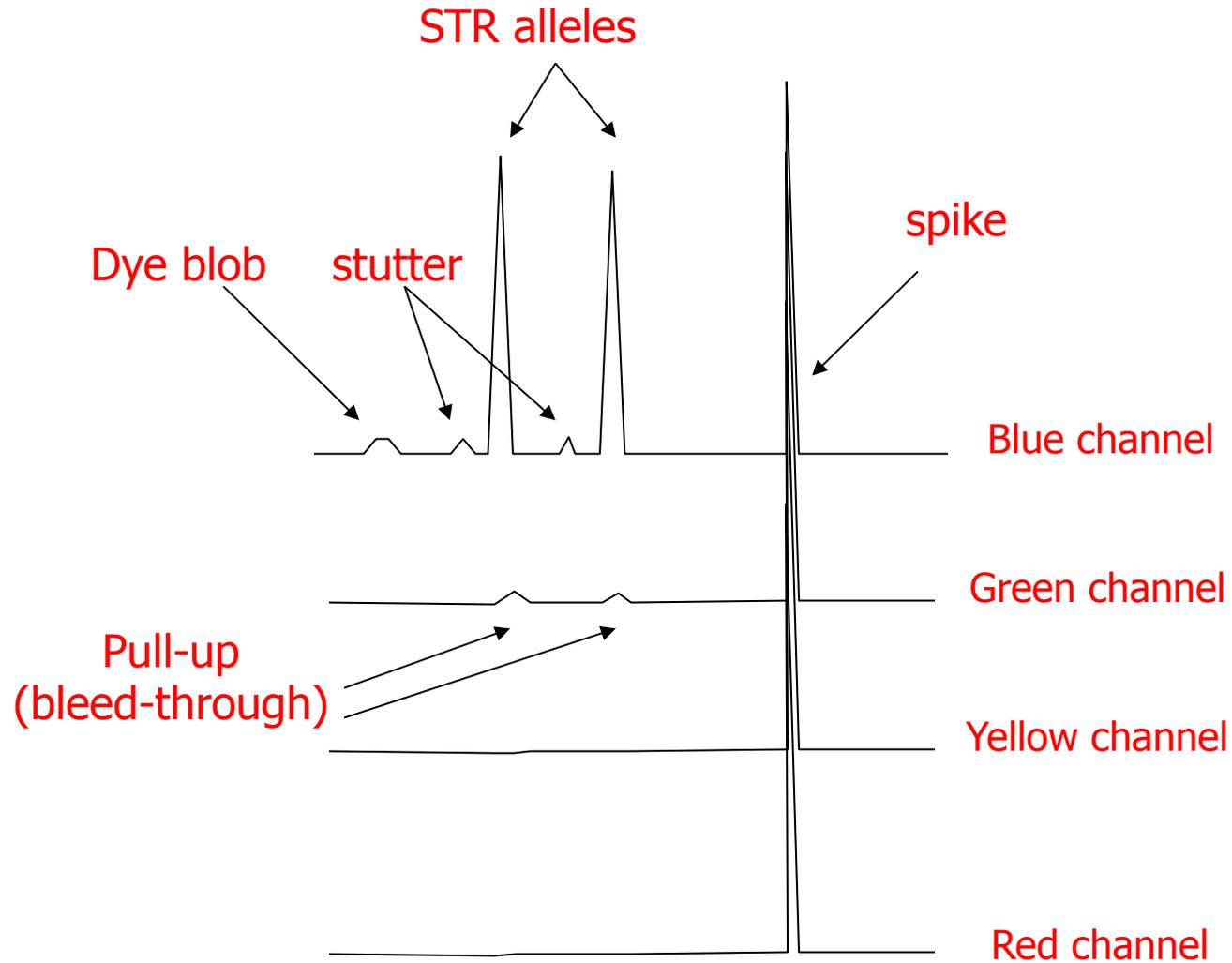
sbalzi di voltaggio, rilevabili come picchi attraverso tutti i colori (Vedi slide successiva)

### ➤ **Contaminanti**

Sostanze fluorescenti nella parte visibile dello spettro

# Problemi nell'interpretazione di profili microsatellite

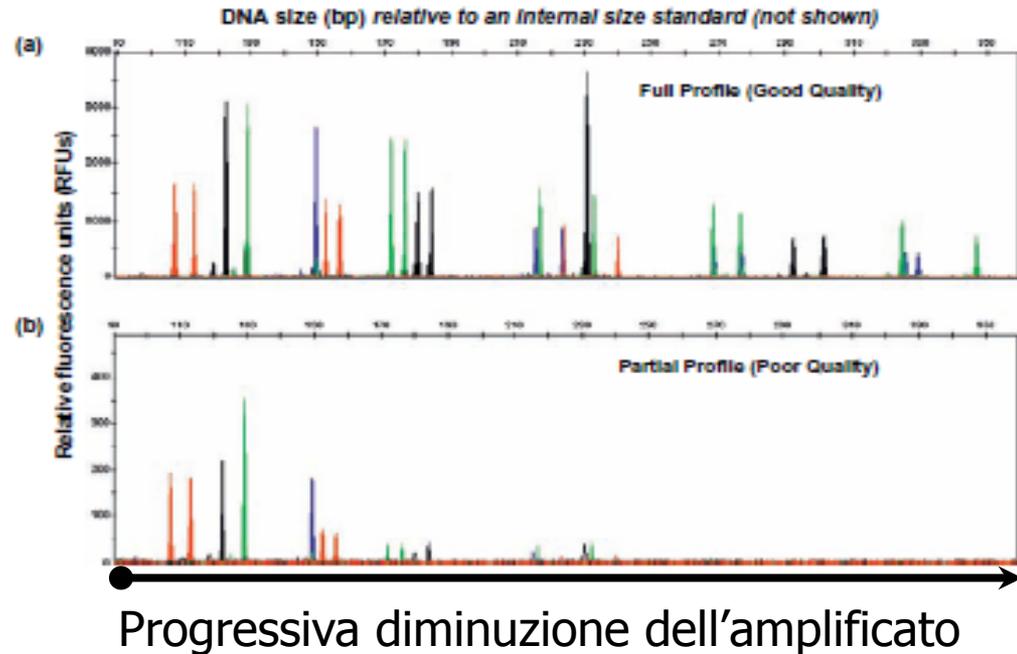
## Exaples



# Problemi nell'interpretazione di profili microsatellite

## Problemi dovuti a "campioni difficili"

### A) DNA degradato



### B) DNA scarso (low concentration DNA)

Alternativa basarsi anche su mtDNA

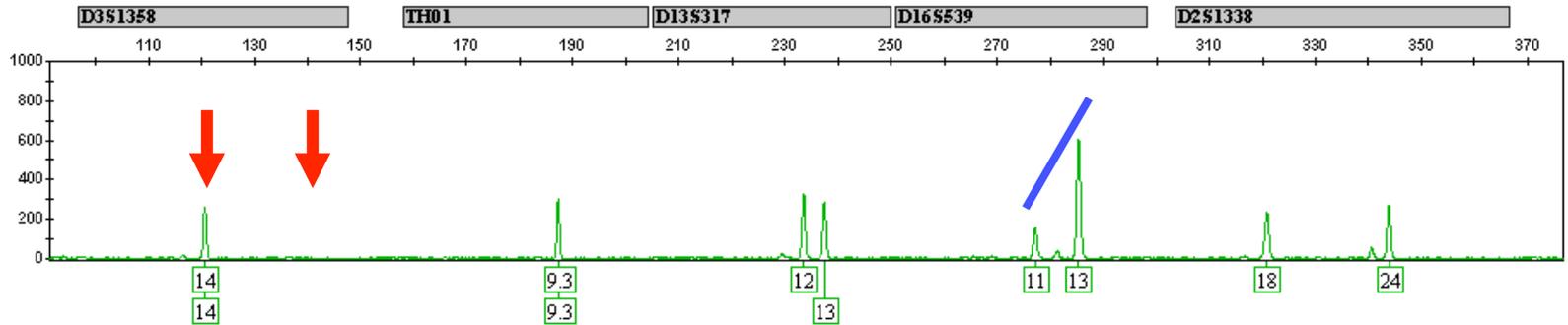
### C) Campioni misti

Possibilità di n. di alleli per individuo > 2

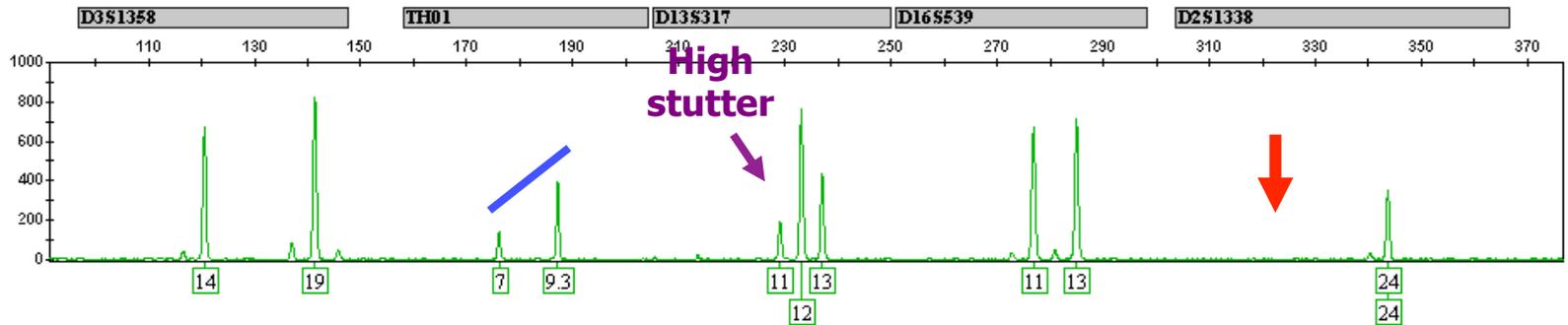
Campioni di riferimento  
Replicati

# Esempio di analisi di replicati

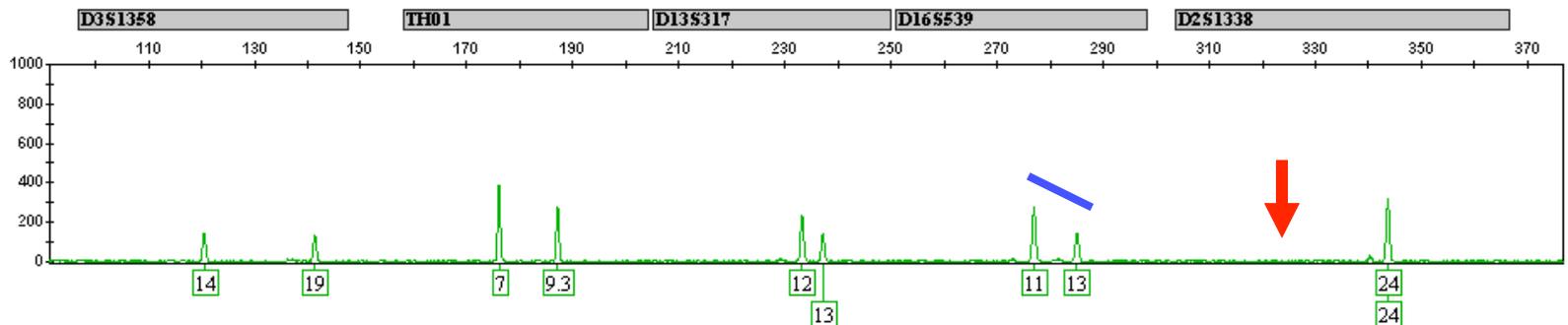
Replicate  
#1



Replicate  
#2



Replicate  
#3



**Consensus**

**Profile:**

**14,19**

**7,9.3**

**12,13**

**11,13**

**24**

**Correct Profile:**

**14,19**

**7,9.3**

**12,13**

**11,13**

**18,24**

# Microsatelliti

## Svantaggi:



- Richiede informazioni a priori (primer)
- Cross-amplificazione tra specie spesso limitata (necessità di isolare marcatori specie-specifici)
- Necessità di campioni di controllo (reference)
- Difficoltà nell'interpretazione (es. alleli nulli, allele dropout)

## Vantaggi:

- Elevata riproducibilità
- Marcatori frequenti nel genoma
- **Loci molto polimorfici** (elevato numero di alleli)
- **CO-DOMINANTI**
- **EREDITA' MENDELIANA**
- Possibilità di multiplexare i loci riducendo i costi di genotyping (**PCR in MULTIPLEX - multilocus**)



# Microsatelliti

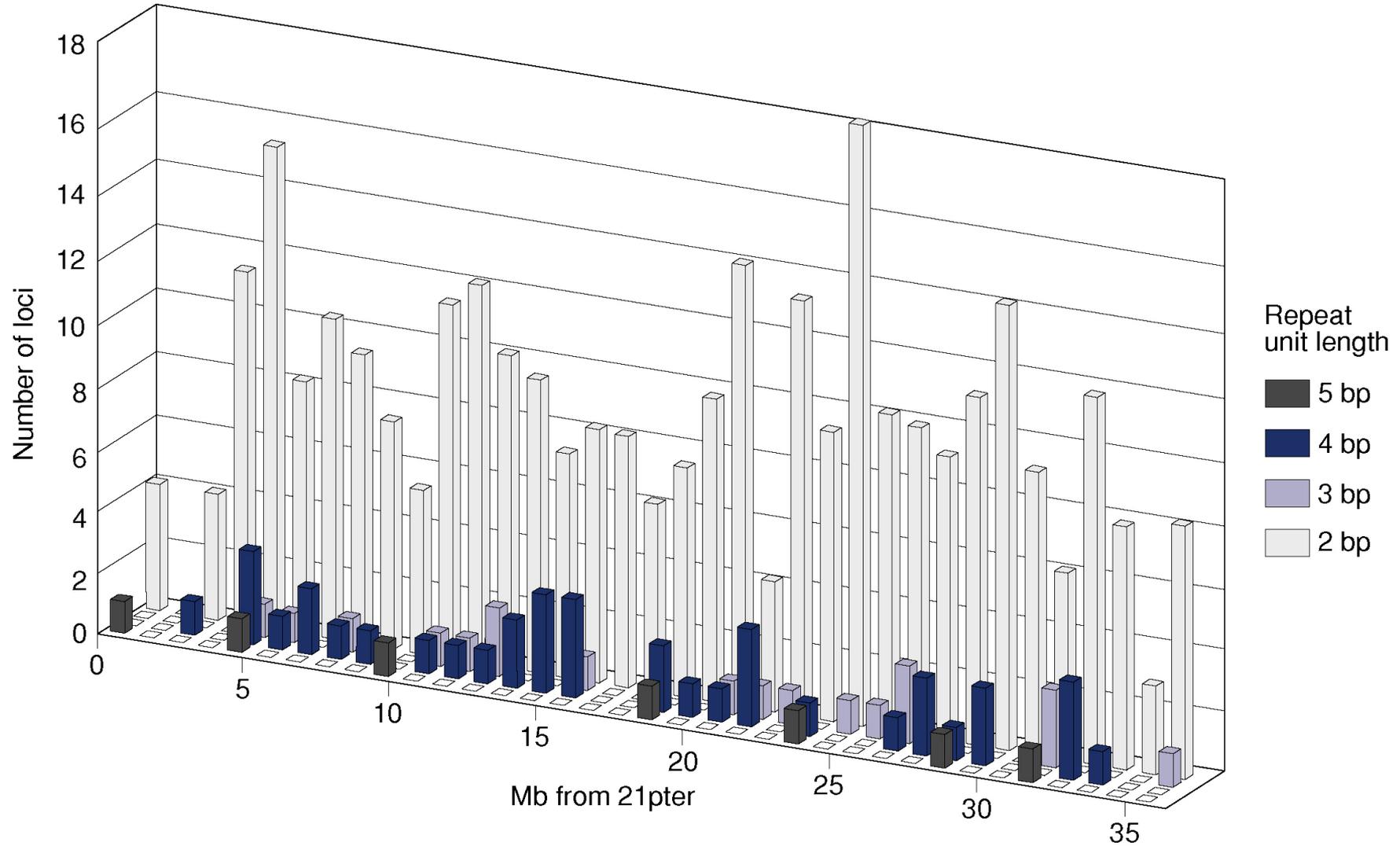
Ancor oggi i marcatori molecolari più utilizzati nelle scienze forensi!

## Vantaggi:

- Elevata riproducibilità
- Marcatori frequenti nel genoma
- **Loci molto polimorfici** (elevato numero di alleli)
- **CO-DOMINANTI**
- **EREDITA' MENDELIANA**
- Possibilità di multiplexare i loci riducendo i costi di genotyping (**PCR in MULTIPLEX - multilocus**)



# Distribuzione dei microsatelliti nel genoma umano



## «Il microsatellite ideale» (nella genetica forense)

- 1) Microsatellite con elevata eterozigosità (maggiore **discriminazione**)
- 2) Microsatellite con range allelico contenuto (**facilità multiplexing**)
- 3) Microsatellite amplificabile in corti prodotti PCR (**x DNA degradato**)
- 4) Microsatellite i cui alleli siano ben separabili su gel (**accuratezza**)
- 5) Microsatellite che presenti un ridotto fenomeno di stuttering
- 6) Microsatellite con tasso di mutazione contenuto (**x paternità**)
- 7) Microsatellite situato su autosomi ("**indipendenza**" **statistica dei loci**)

Molte di queste caratteristiche si escludono vicendevolmente:

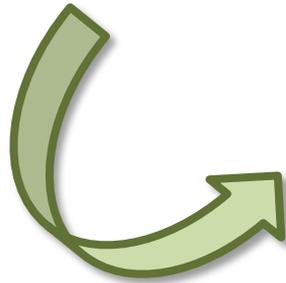
Qual è un ragionevole compromesso?

La comunità forense si è orientata principalmente su  
microsatelliti del  
tipo tetranucleotide repeats autosomici

# Microsatelliti

## CODIS (Combined Dna Index System)

Nel 1990, l'FBI decide di utilizzare loci caratteristici per indagini forensi



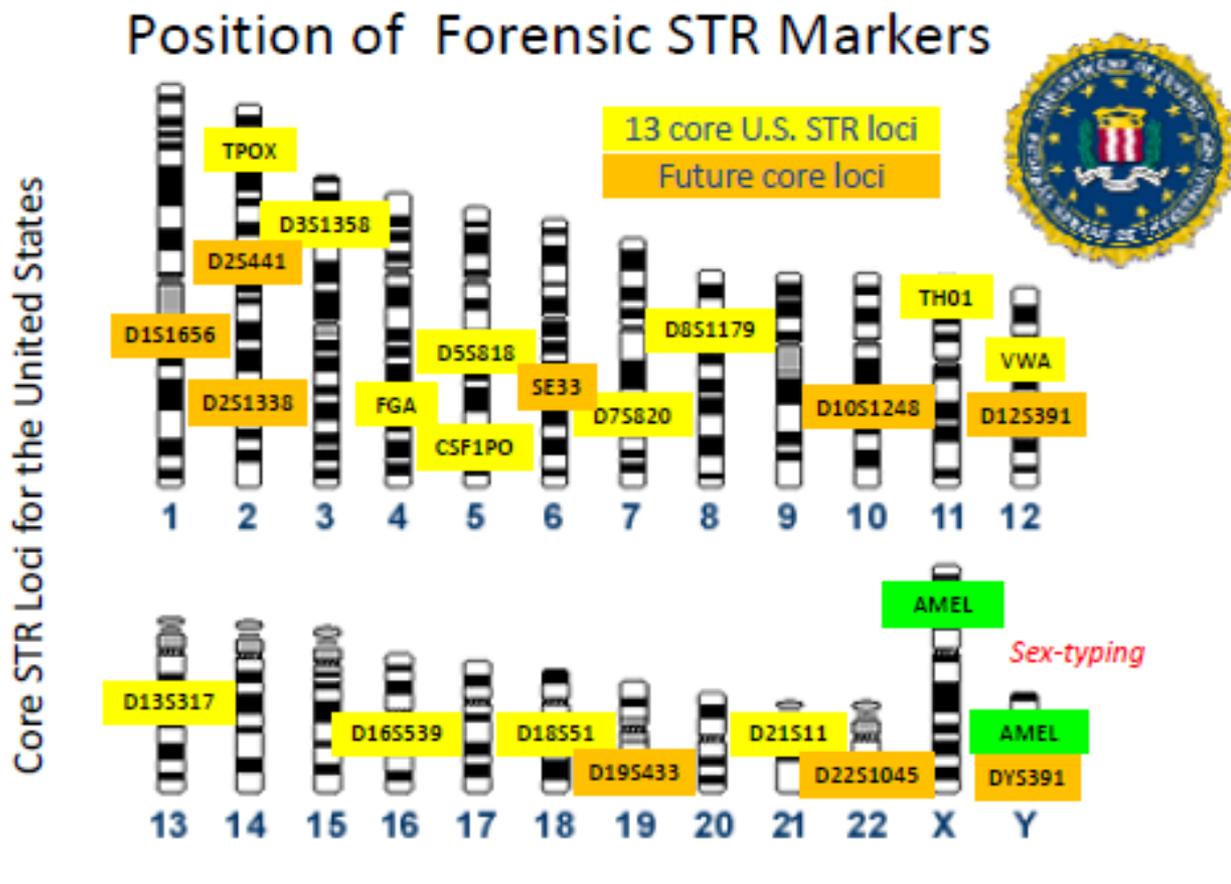
Necessità di identificare un set unico (core) di microsatelliti che sia condiviso da tutta la comunità genetica forense.

- Durante gli anni il numero di loci selezionati è aumentato
- Attualmente il **CODIS USA** ha in uso 13 microsatelliti selezionati a partire da 17 loci candidati sulla base della loro variabilità e facilità di interpretazione.
- Usato in tutti i laboratori forensi americani
- In **UK e alcuni altri stati Europei** si utilizza un core di 10 STR, 8 dei quali in comune con il CODIS

# Microsatelliti

# CODIS (Combined Dna Index System)

Pannello di 13 micro  
«core» (in giallo)  
secondo il CODIS, USA



I 13 loci sono:

- Tutti «**tetranuceotide repeats**», sia semplici che composti
- Tutti localizzati in regioni **non codificanti**
- 11 si trovano su cromosomi diversi, solo 2 sullo stesso cromosoma

- L'analisi dei 13 loci del CODIS porta a profili genetici la cui probabilità di random match è minore di 1 su un miliardo di milioni ( $10^{-15}$ ).

# Microsatelliti

## CODIS (Combined Dna Index System)

L'utilizzo dei medesimi loci nelle varie popolazioni permette di avere informazioni sulla frequenza di un determinato allele nella popolazione in esame.

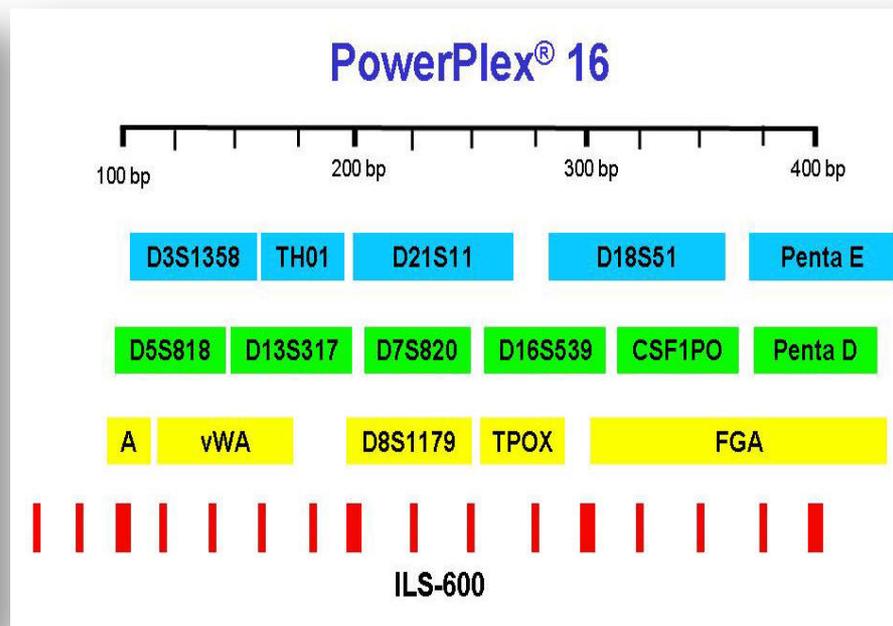
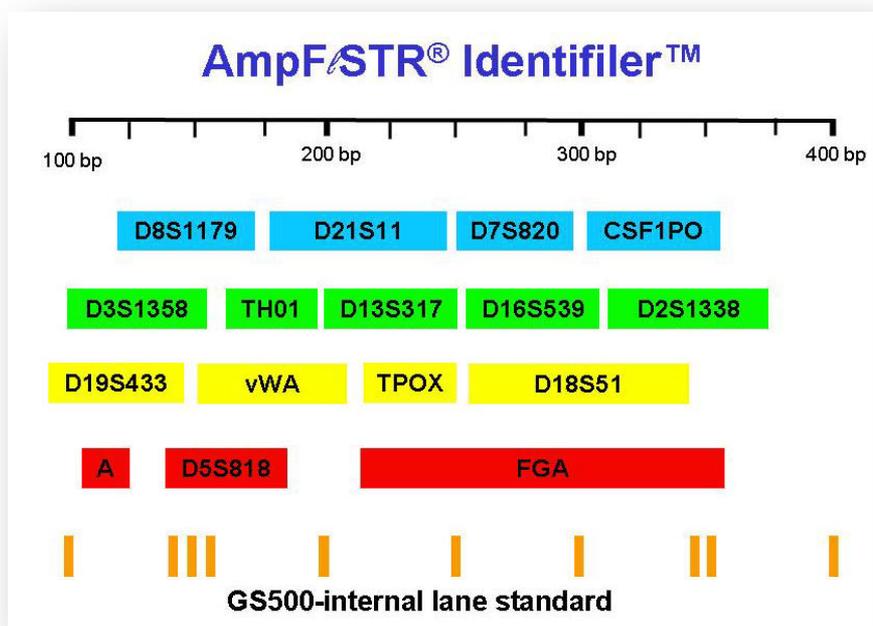


Programma dell'FBI in supporto alla giustizia che si basa sullo sviluppo e l'utilizzazione di DATABASE DI PROFILI GENETICI.

# Microsatelliti

## CODIS (Combined Dna Index System)

I kit commerciali sviluppati da Applied Biosystems (a sinistra) e Promega (a destra) permettono di amplificare i 13 STR «core» del CODIS più alcuni altri STR.



I kit commerciali sono in continua «evoluzione» inseguendo le richieste della comunità forense (maggiore numero di loci coamplificabili, mini STR, Y-STR, STR ipervariabili ecc.)

# Generare un profilo STR

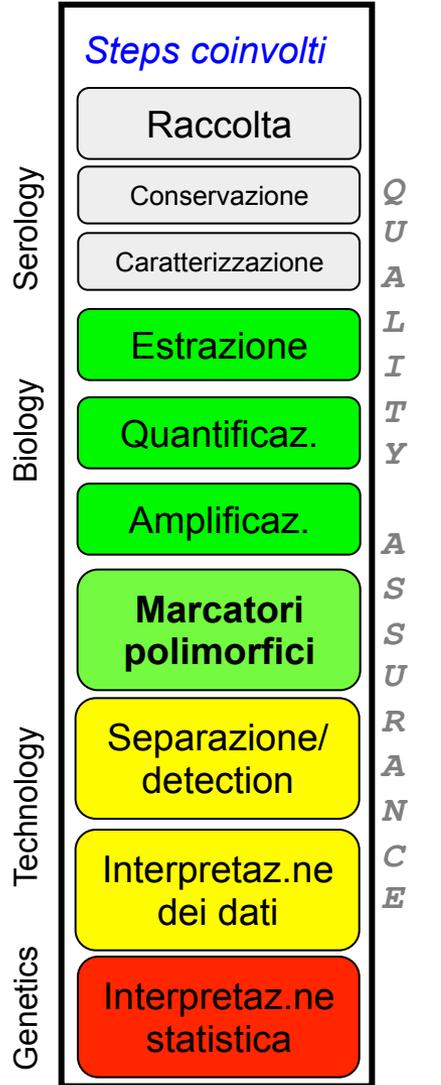
## 1. Estrazione e quantificazione del DNA

(attenzione: questi steps in genetica forense sono di fondamentale importanza e sono successivi ad uno step altrettanto importante di caratterizzazione del campione)

# Crimine

Trasferimento del materiale biologico

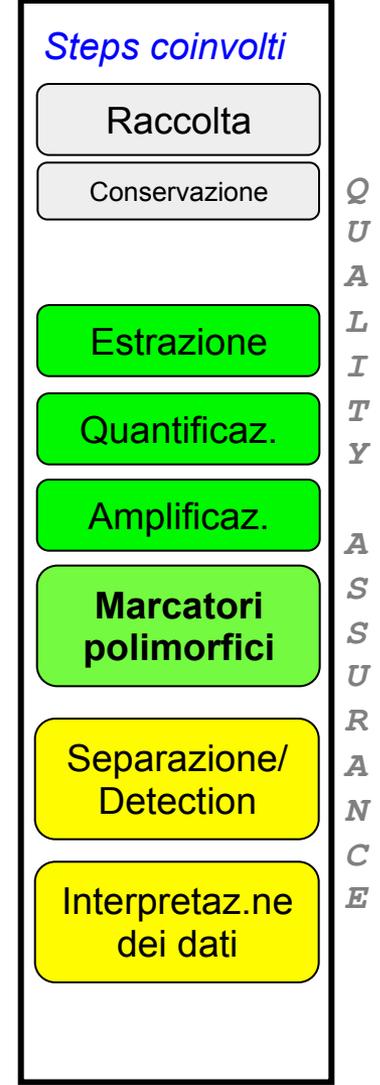
**Campione**  
"Q" (Question)



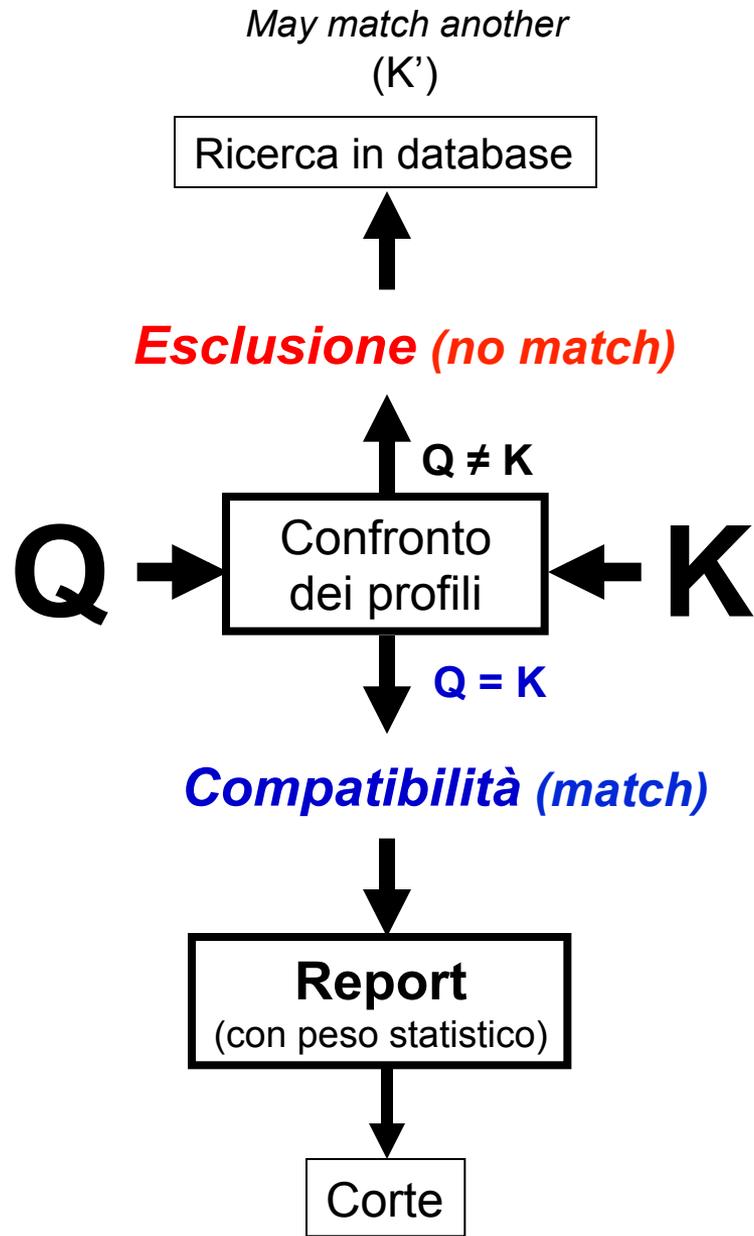
Profilo depositato in database

# Sospettato individuato

**Reference**  
campione "K" (Known)



Profilo depositato in database



# Generare un profilo STR

## 1. Estrazione e quantificazione del DNA

(attenzione: questi steps in genetica forense sono di fondamentale importanza e sono successivi ad uno step altrettanto importante di caratterizzazione del campione)

Contenuto in DNA nucleare dei campioni biologici:

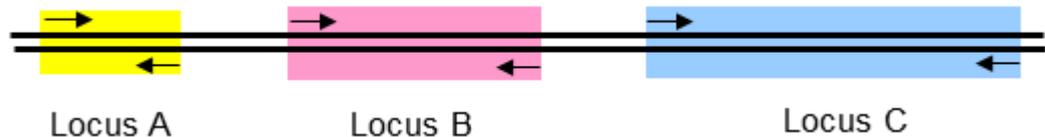
<u>Tipo di campione</u>	<u>Quantità di DNA</u>
<b>Sangue</b>	30000 ng/mL
macchia 1 cm <sup>2</sup>	200 ng
macchia 1 mm <sup>2</sup>	2 ng
<b>Liquido seminale</b>	250000 ng/mL
tampone vaginale postcoitale	0 - 3000 ng/tampone
<b>Capelli</b>	
capello strappato	1 - 750 ng/capello
capello perso	1 - 12 ng/capello
<b>Saliva</b>	10000 ng/mL
<b>Urine</b>	1 - 20 ng/mL
<b>Cellula umana diploide</b>	~7 pg

# Generare un profilo STR

## 2. Amplificazione mediante PCR multiplex



- Amplificare tramite PCR contemporaneamente più microsatelliti (**multiplex**), utilizzando primers fiancheggianti ciascun locus.



Per ottenere buoni risultati i primers devono essere scelti con attenzione:

- Le temperature di annealing dei vari primers devono essere simili
- Devono essere valutate e minimizzate le possibili interazioni intra- e inter- primers, anche appartenenti a coppie diverse (esempio complementarità al 3')

Software online per primer selection: "Primer3Plus" <http://primer3plus.com/>

- Una valutazione sperimentale della concentrazione ottimale dei singoli primers è sempre necessaria quando si mette a punto una nuova multiplex, così come una titolazione dei vari componenti della PCR.

# Generare un profilo STR

## 2. Amplificazione mediante PCR multiplex

### ■ PRESENZA DI INIBITORI:

I reperti biologici possono essere mescolati con inibitori della PCR di tipo organico ed inorganico che non sempre vengono eliminati efficacemente negli step di purificazione del DNA.

- Ulteriore purificazione,
- BSA,
- diluzione della soluzione di DNA

**Table 1. Known PCR Inhibitors.**

<b>Inhibitor</b>	<b>Source of Inhibitor</b>
bile salts	feces
complex polysaccharides	feces, plant material
collagen	tissues
heme	blood
humic acid	soil, plant material
melanin and eumelanin	hair, skin
myoglobin	muscle tissue
polysaccharides	plants
proteinases	milk
calcium ions	milk, bone
urea	urine
hemoglobin, lactoferrin	blood
immunoglobulin G (IgG)	blood
indigo dye	denim

# Generare un profilo STR

## 2. Amplificazione mediante PCR multiplex

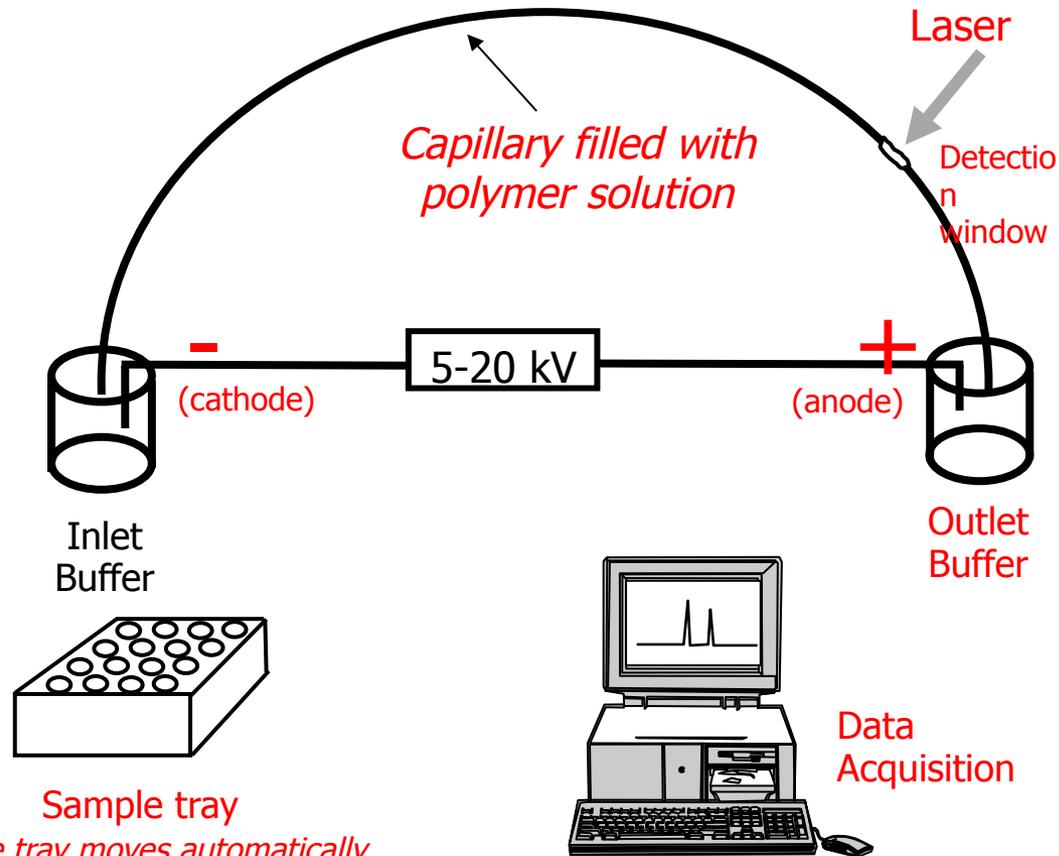
### ■ CONTAMINAZIONE IN LABORATORIO O IN FASE DI REPERTAZIONE:

a causa della sua elevata sensibilità, la PCR presenta un forte problema legato alla contaminazione, in misura tanto maggiore quanto più il DNA in esame è scarso e/o degradato.

- Utilizzazione di guanti monouso e mascherine (in lab e in repertazione)
- Strumentazione e spazi pre- e post- PCR separati;
- Utilizzazione di puntali monouso con filtro
- Irradiazione con UV dell'area dedicata alla preparazione della PCR
- Creazione di un database di profili genetici del personale del laboratorio

# Generare un profilo STR

## 3. Separazione dei prodotti di PCR mediante elettroforesi capillare

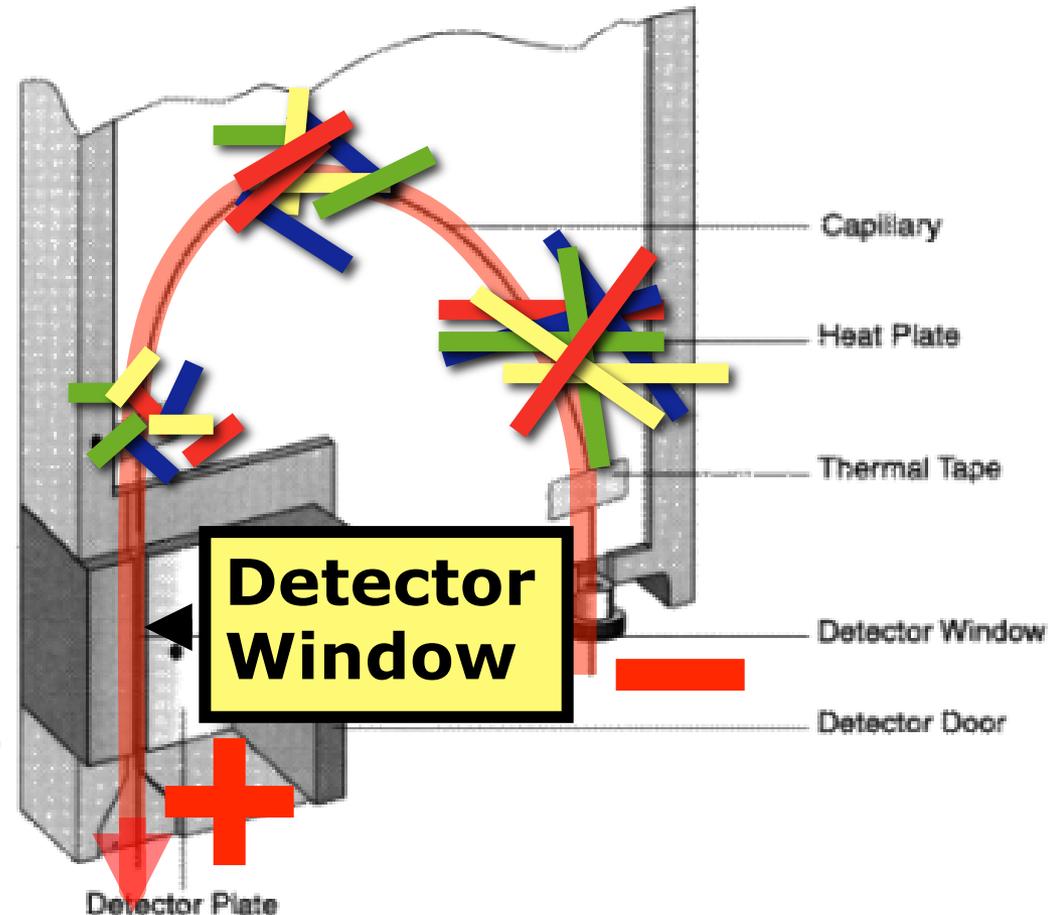


*Sample tray moves automatically beneath the cathode end of the capillary to deliver each sample in succession*

# Generare un profilo STR

## 3. Separazione dei prodotti di PCR mediante elettroforesi capillare

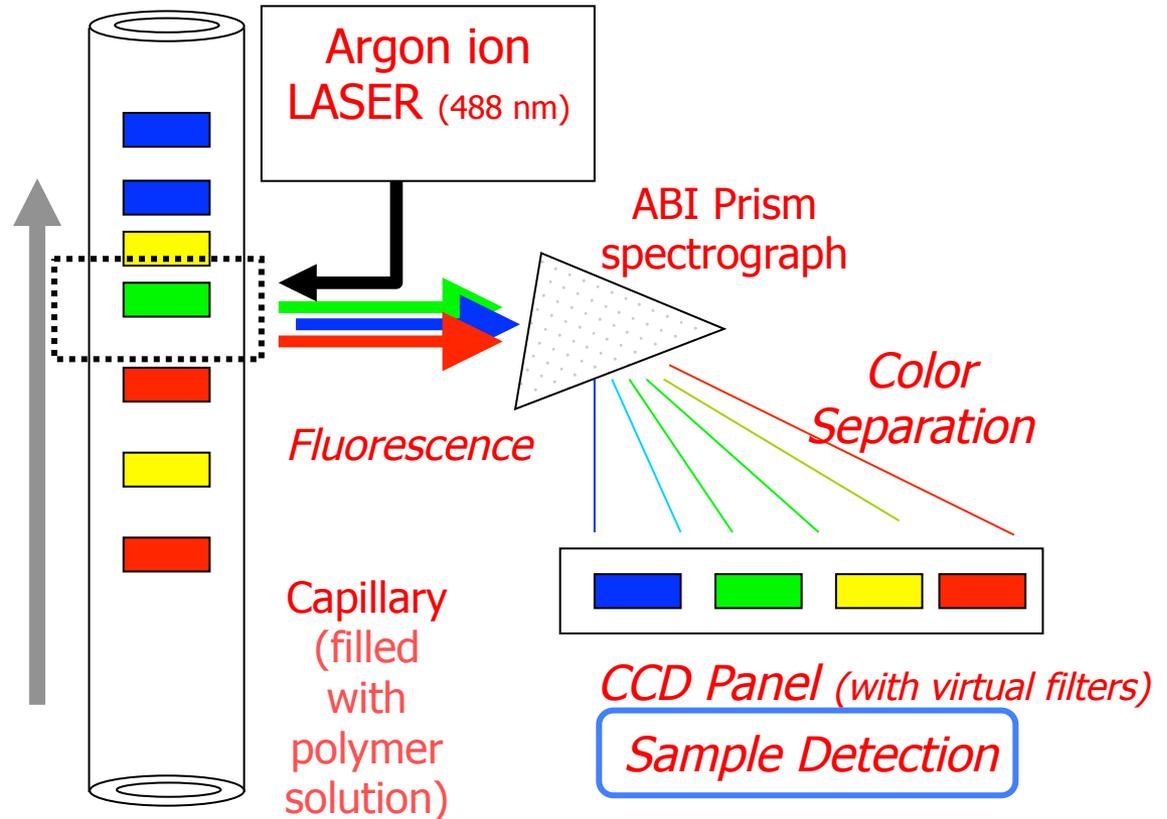
- Prodotto PCR fluorescente viene denaturato con formammide e posto nel campionatore
- Si applica voltaggio elettrico Il prodotto PCR viene elettroiniettato nel capillare
- Il DNA migra nel polimero contenuto nel capillare tanto più velocemente quanto minore è la grandezza del frammento
- Quando il frammento arriva in corrispondenza di una finestra nel capillare, una sorgente laser eccita il fluorocromo legato al prodotto PCR



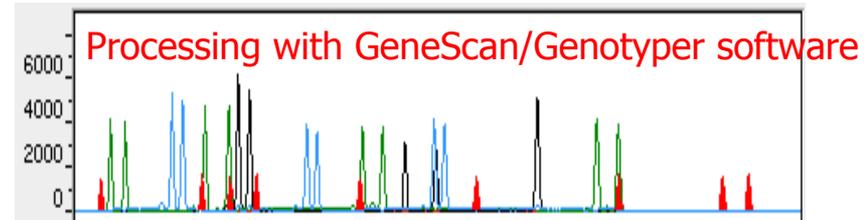
# Generare un profilo STR

## 3. Separazione dei prodotti di PCR mediante elettroforesi capillare

- Il fluorocromo eccitato dal laser emette luce nel visibile, in un determinato intervallo di lunghezze d'onda e con una intensità che è **proporzionale alla quantità di molecole di fluorocromo eccitate**.
- La luce emessa viene filtrata per determinati range di lunghezza d'onda e convertita in un **segnale elettrico che viene misurato in RFUs (Relative Fluorescence Units)**.



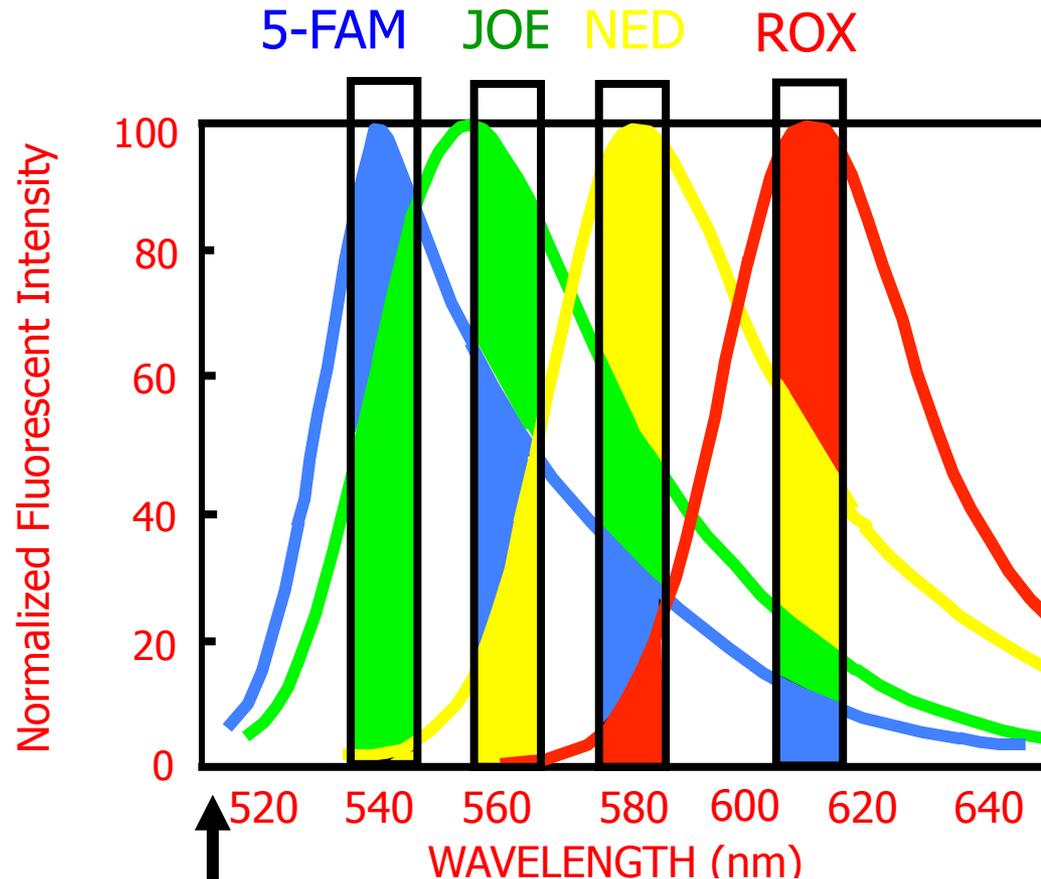
tradotto nei  
tipici picchi  
colorati



Sample Interpretation

# Generare un profilo STR

## 3. Separazione dei prodotti di PCR mediante elettroforesi capillare



Laser excitation  
(488 nm, 514.5 nm)

310 Filter Set F  
with color contributions

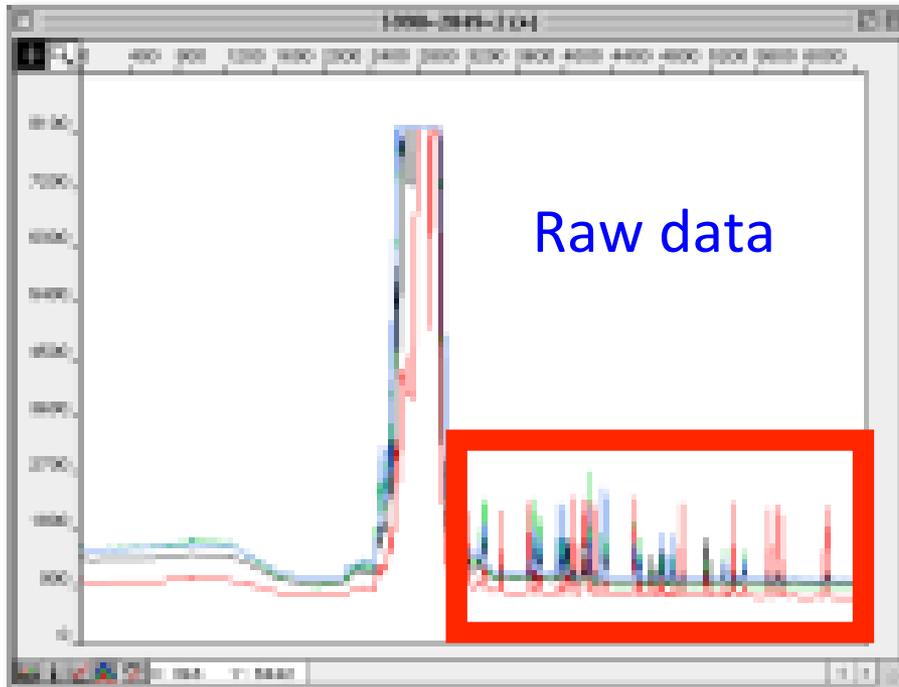
Spettri di emissione dei fluorocromi Applied B. in uso nel kit AmpFISTR.

I filtri virtuali sono rappresentati dalle quattro colonne nere centrate sul picco dello spettro di ciascuno dei 4 fluorocromi.

A causa della sovrapposizione degli spettri, una matrice di deconvoluzione (4 x 4) deve essere utilizzata per una corretta interpretazione dei dati.

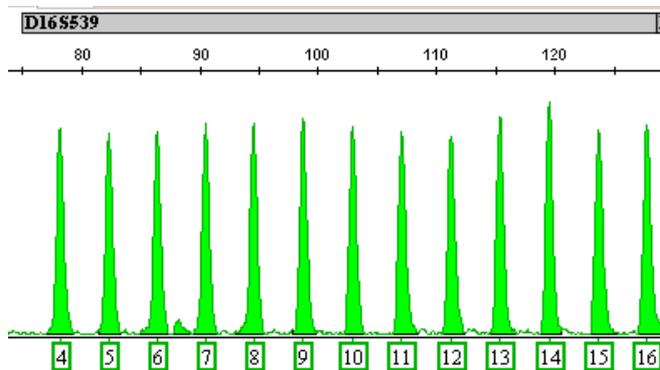
# Generare un profilo STR

## 4. Interpretazione del dato grezzo



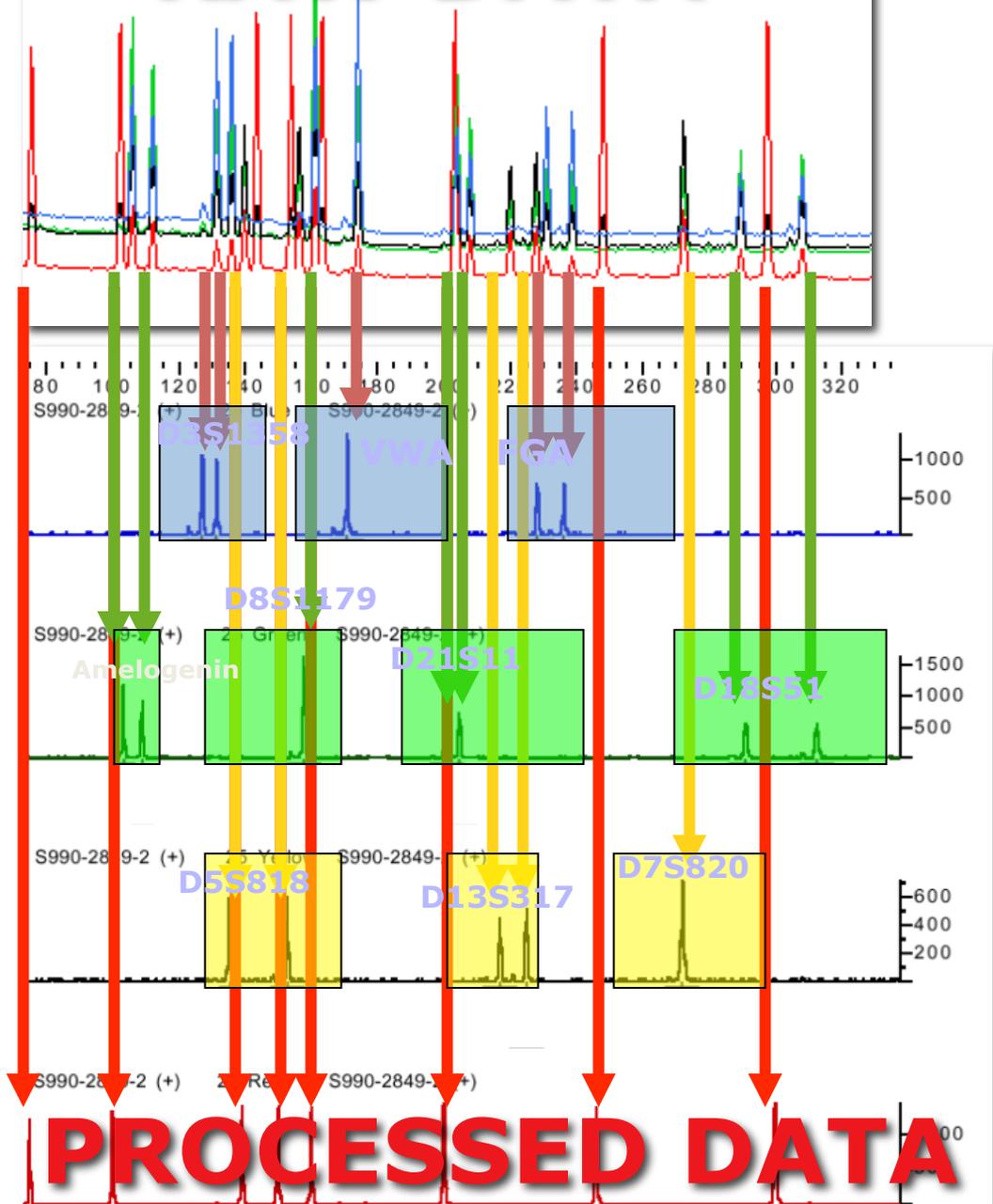
- Il software GENESCAN divide i dati grezzi in un elettroferogramma separato per ciascun colore

- Blue
- Green
- Yellow
- Red



- Il software GENOTYPER identifica i differenti loci e per ciascuno di essi riconosce gli alleli facendo riferimento ad uno standard di peso molecolare e ad un **ladder allelico**

# RAW DATA



Data Collection

Color Separation

Matrix file  
(spectral  
calibration)  
  
User-defined thresholds

Peak Identification

Peak Sizing

Internal size  
standard

Comparison to  
Allelic Ladder

Allelic ladder  
sample

Genotype Assign.  
to Alleles

Peak Editing to Remove  
Artifact Calls

Data Review by  
Analyst/Examiner

Confirmation of  
Results by Second  
Analyst/Examiner

*GeneScan*  
software

*Genotyper*  
software

*GeneMapperID*  
software

*Expert Systems*  
(e.g., *FSS-i3*, *TrueAllele*)

SACRAMENTO — Arnold Schwarzenegger said he realized he was the father of his housekeeper's child when the boy reached age 7 or 8 and the resemblance became apparent.

## Test di parentela e paternità



**SOCIETÀ**

04/07/2012 - LA STORIA

### "Voglio le prove che sia mio" Balotelli invoca il test di paternità

L'ex Raffaella Fico annuncia la maternità. Lui: «Mi assumerò tutte le responsabilità soltanto quando avrò la conferma»

**EGLE SANTOLINI**

MILANO

Lo chiamano il test «peace of mind» o della tranquillità d'animo ed è sempre più diffuso anche grazie alla pubblicità involontaria dei ricchi & famosi. Ultimo caso quello del bomber Mario Balotelli, deciso a farlo dopo che Raffaella Fico ha annunciato di aspettare un bambino da



# Microsatelliti – test di parentela e paternità

L'analisi di loci microsatellite può essere usata per studi sulla parentela di individui diversi.

La maggior parte di questi studi riguarda:

1. Test di paternità
2. Studi per determinare se determinati individui (viventi o deceduti) siano imparentati.
3. Analisi volte a risolvere casi di persone scomparse.
4. Analisi volte ad identificare cadaveri nei disastri di massa
5. Analisi di supporto ad investigazioni criminali (...il parente dell'assassino...)
6. Predizione dell'origine etnica di un DNA (estendendo il concetto di parentela a quello di «coancestry»)

# Microsatelliti – test di parentela e paternità

In tutti i casi, i test di parentela si basano su un concetto semplice:

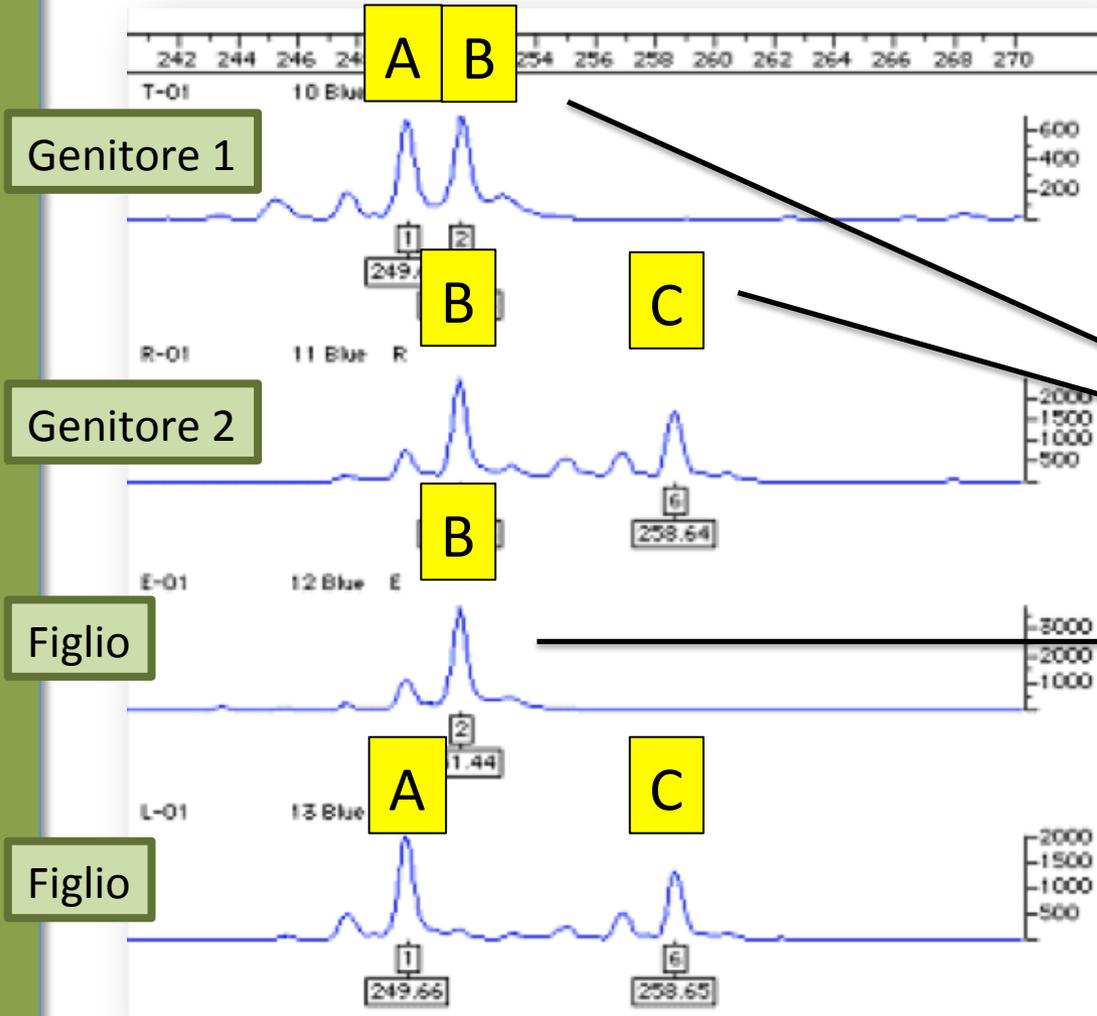
Individui imparentati condividono alleli in una maggiore proporzione (spesso misurabile) rispetto ad individui presi a caso dalla popolazione.

Sono, in altre parole, più simili geneticamente

Differenze tra autosomi, mtDNA, Cromosoma Y  
Consanguineità e alleli uguali per discendenza

I microsatelliti sono lo strumento principale per le analisi di parentela in genetica forense grazie alla loro eredità mendeliana

# Microsatelliti



Es. Esempio eredità in un locus diploide  
I genitori sono entrambi eterozigoti e condividono un solo allele

Heterozygotes

Homozygotes

# Test di paternità

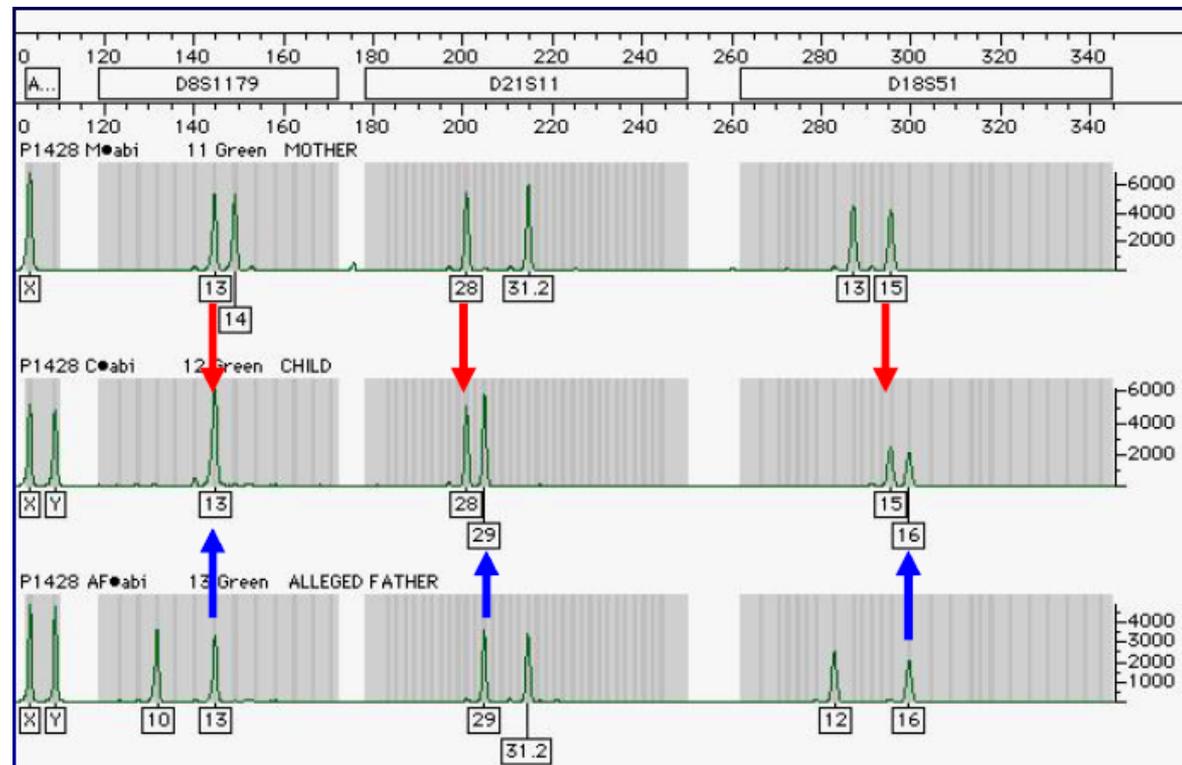
Due possibili risultati:

## (1) Compatibilità:

Tutti\* gli “**alleli obbligati paterni**” del figlio sono presenti nel padre putativo

## (2) Esclusione:

Più\* “alleli obbligati paterni” del **figlio non** sono presenti nel genotipo del padre putativo



\* Attenzione ad alleli nulli e mutazioni!!!

# Test di paternità

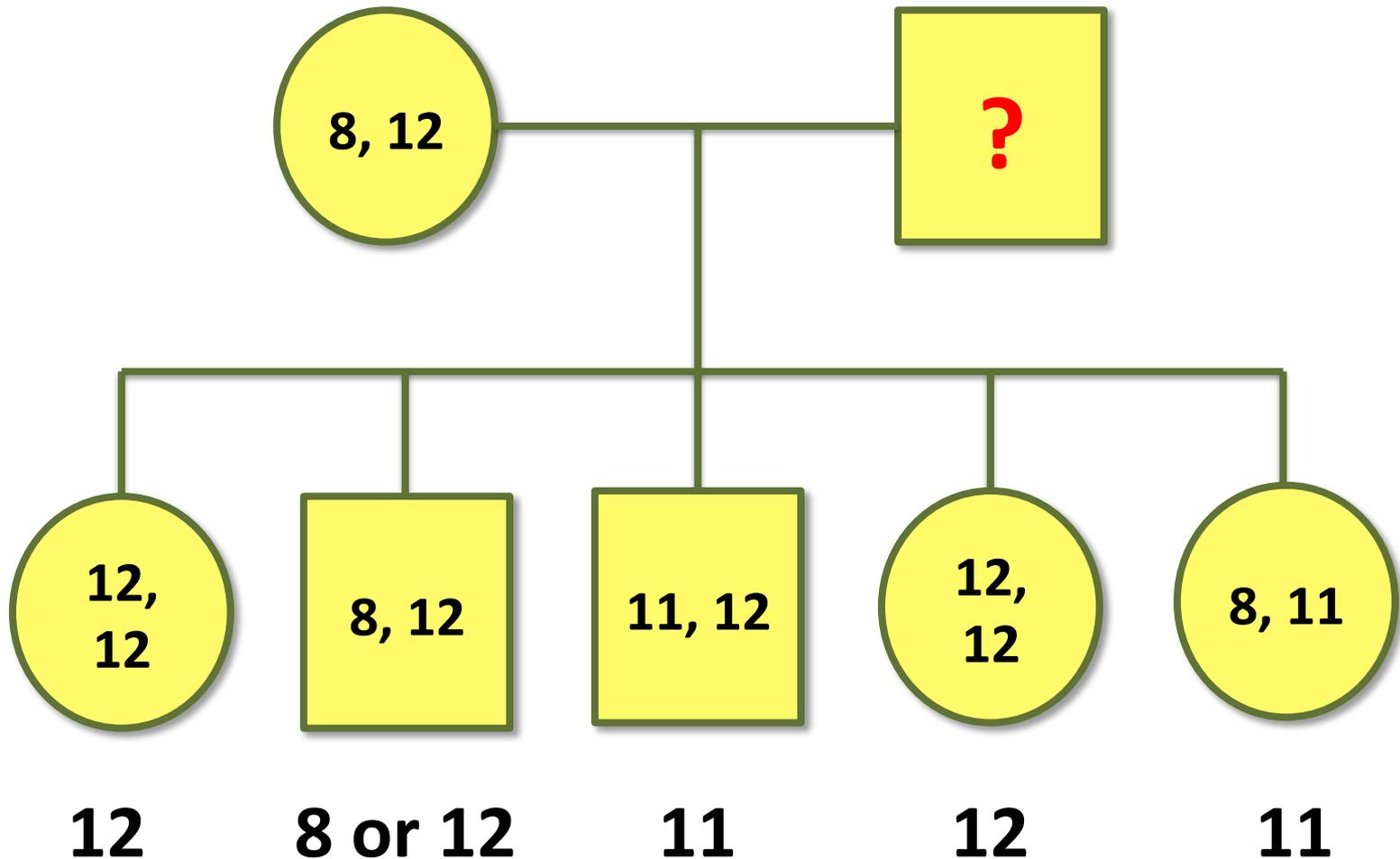
**“alleli obbligati paterni” = L’allele (i) che il padre DEVE avere**

In base alle leggi dell’ereditarietà (**1° legge Mendel**), considerando un locus autosomico e assumendo che “la madre sia certa”, se si conosce il genotipo di madre e figlio, è possibile dedurre (**nella maggior parte dei casi**) l’allele di origine paterna. Oppure (**nei casi meno fortunati**) due alleli di cui uno sia di origine paterna.

Questo allele (o alleli) deve essere presente nel padre putativo (**allele obbligato**) pena esclusione della paternità.

# Test di paternità

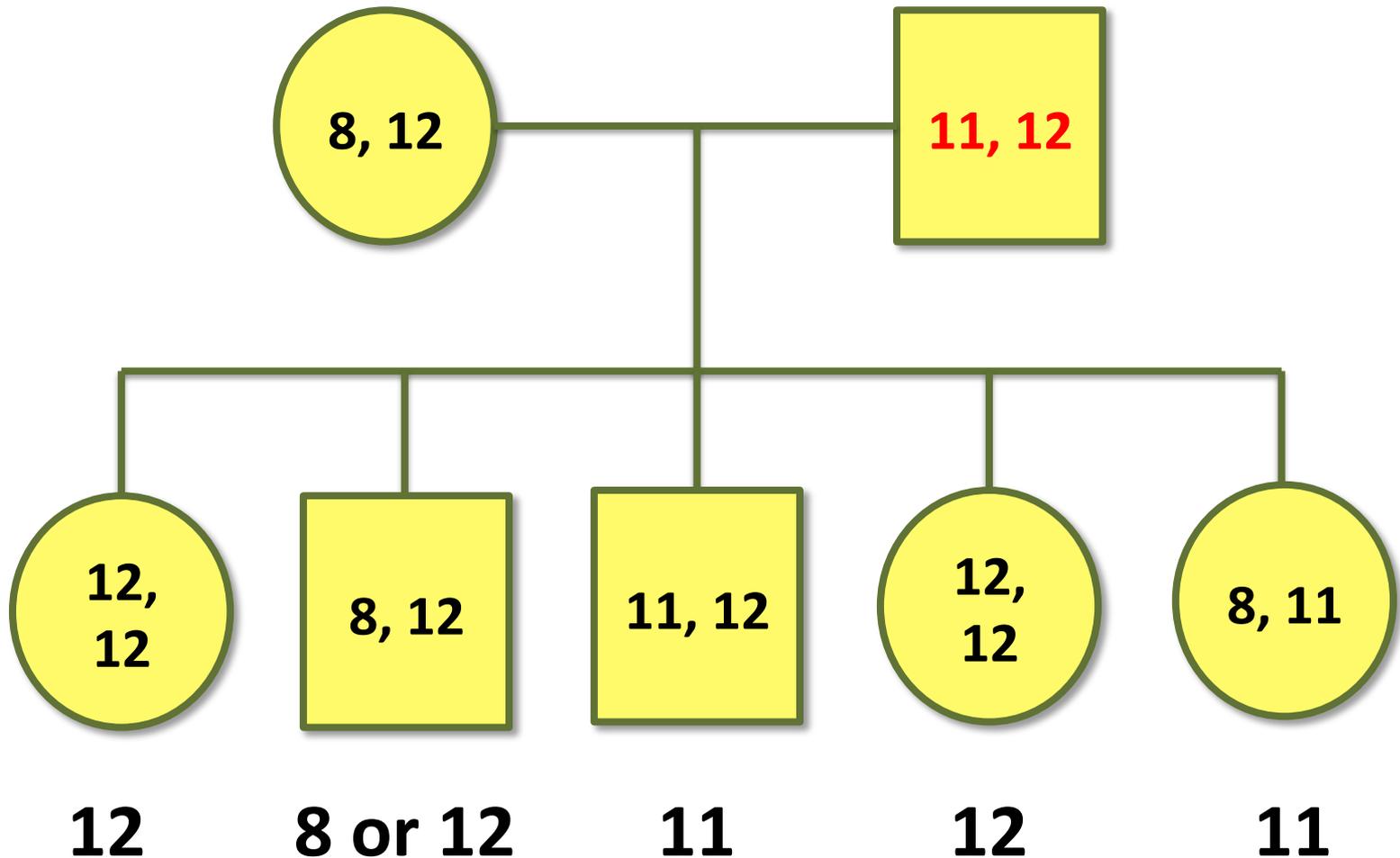
“alleli obbligati paterni” = L'allele (i) che il padre DEVE avere



Obbligate paternal allele

# Test di paternità

“alleli obbligati paterni” = L'allele (i) che il padre DEVE avere

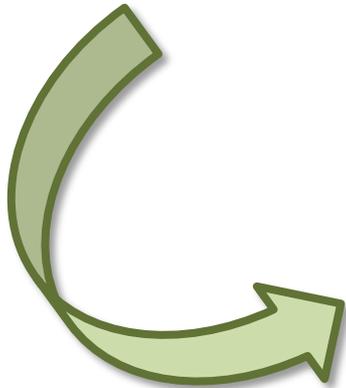


Obbligate paternal allele

# Test di paternità

In caso di **esclusione**, il test può considerarsi concluso

In caso di **compatibilità**, è necessario quantificare il “peso” dell’osservazione.

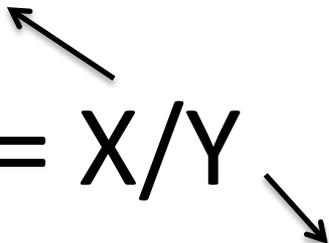


**P.I. = Indice di paternità**  
(per un locus)

**C.P.I = Indice di paternità combinato**  
(considerando più loci).

## Indice di paternità (PI)

Numeratore (X): probabilità che il padre putativo trasmetta l'allele(i) obbligato essendo il vero padre. Il trio in analisi è un vero trio

$$PI = X/Y$$


Denominatore (Y): probabilità che un uomo qualsiasi della popolazione trasmetta l'allele(i) obbligato essendo il vero padre. Il trio in analisi è un falso trio

Queste due probabilità discendono rispettivamente dalle leggi di Mendel e dalle leggi della genetica di popolazioni, il loro rapporto (PI, paternity index) esprime una verosimiglianza (Likelihood Ratio, LR).

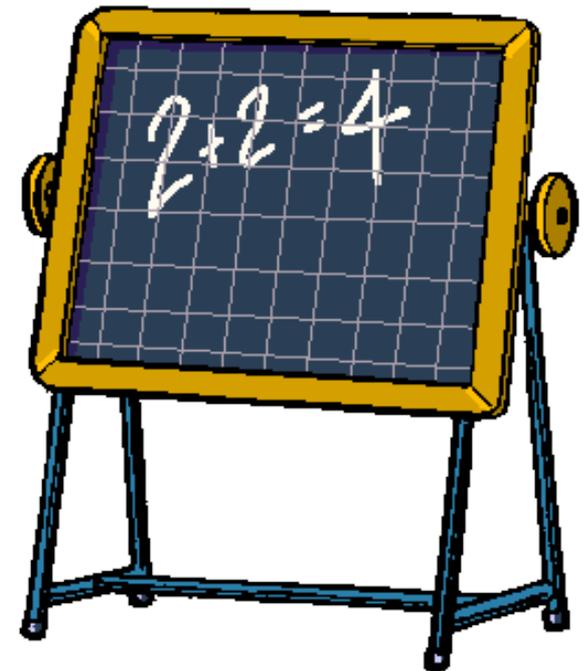
## Indice di paternità (PI)

Esempio (semplice):

Madre AB, Figlio AC, Padre AC

- Individuare l'allele/i paterno obbligato/i (allele C in questo esempio)
- Numeratore: Probabilità che il padre trasmetta l'allele obbligato =  $\frac{1}{2}$  (Mendel!!)
- Denominatore: Probabilità che un individuo a caso della popolazione trasmetta l'allele obbligato = Frequenza dell'allele nella popolazione (Hardy-Weinberg!!)
- → necessità di utilizzare database di frequenze alleliche. Se l'allele C ha frequenza nella popolazione (ad esempio)  $p_C = 0,1$ , allora (in questo esempio):

$$P.I. = 0,5/0,1 = 5$$



## Indice di paternità (PI)

Un solo locus non è generalmente sufficiente per ottenere una informazione che abbia un peso rilevante circa la compatibilità osservata.

Si considerano più loci

## C.P.I = Indice di paternità combinato = prodotto dei singoli PI

Un valore di CPI  $> 100$  viene in genere considerato sufficientemente alto da far ritenere il padre putativo il vero padre.

CPI = 100 significa: « E' 100 volte più probabile ottenere i risultati osservati (i profili genetici dei tre soggetti del trio) se il padre putativo (piuttosto che un uomo preso a caso dalla popolazione) è il vero padre

Con i loci microsatelliti attualmente in uso in genetica forense, in caso di compatibilità si ottengono valori di CPI  $\gg 100$

## Indice di paternità o probabilità di paternità???

In alternativa all'indice di paternità CPI e può essere usata una probabilità di paternità

La formula più utilizzata è la «**probabilità di Essen-Möller**» espressa come

$$W = X/(X+Y)$$

X: probabilità che il padre putativo trasmetta l'allele(i) obbligato essendo il vero padre.

Y: probabilità che un uomo qualsiasi della popolazione trasmetta l'allele(i) obbligato.

Esiste una relazione semplice tra CPI e W

**CPI può essere trasformato in  
«probabilità di paternità» come**

$$**W = CPI/(CPI+1)**$$

## Casi di dubbia interpretazione???

In caso di esclusione si osserva generalmente non compatibilità per più loci.

**Ma come considerare una situazione in cui solo uno (o "pochi") dei loci analizzati non sono compatibili?**

Bisogna tener conto delle **possibili mutazioni**

Tasso di mutazione?

Come si tiene conto delle mutazioni?

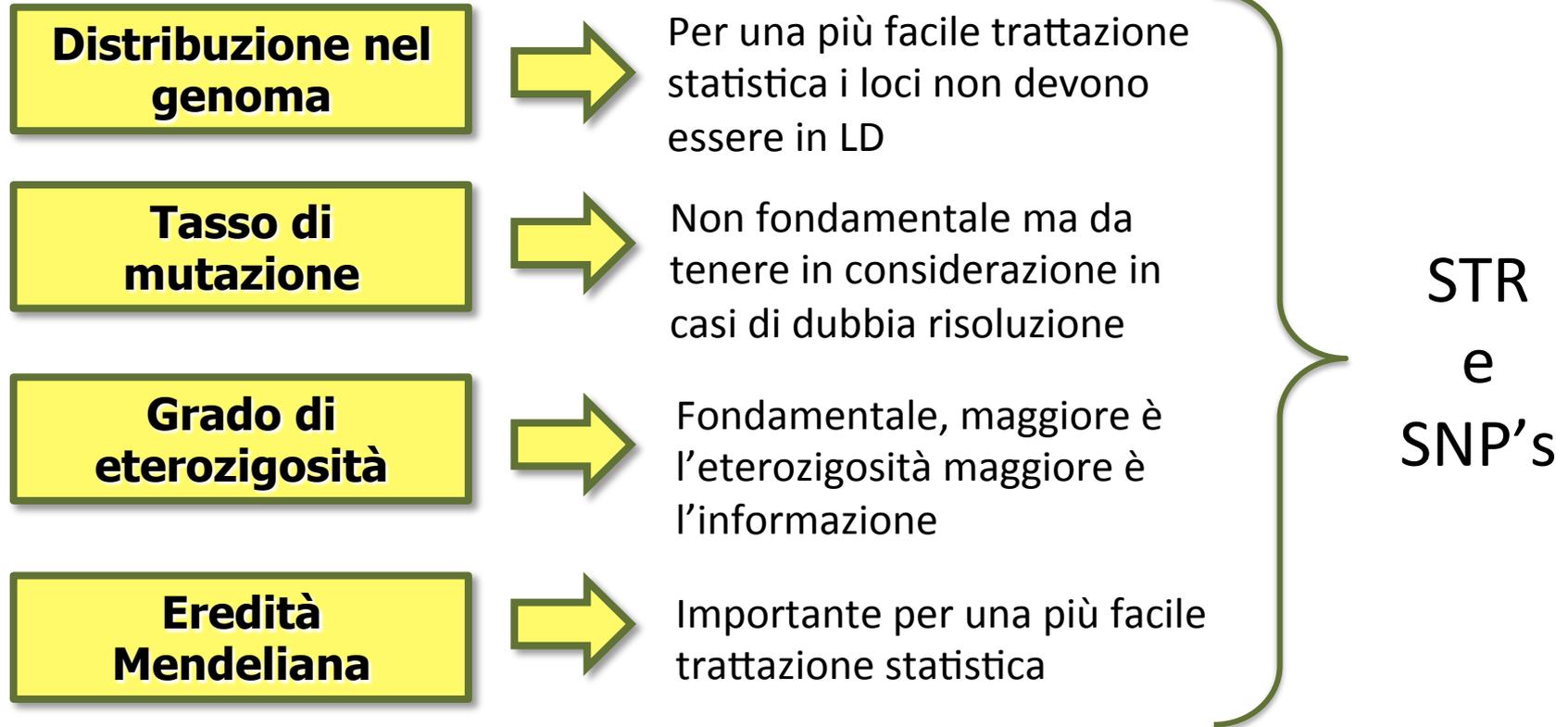
**"two exclusion rule"**

**e**

**CPI in presenza di mutazioni**

# Test di paternità

In conclusione quale marcatore scegliere per un test di paternità?



Nella pratica, per i test di paternità si utilizzano quasi esclusivamente **STRs autosomici**, con la possibilità di fare ricorso a **STR -Y, mtDNA e STR-X** in casi particolari

Cercano informazioni che interessano altri tipi di parentela.

# Test di parentela

**TABLE 14.1** Degrees of Relatedness and Probability of Sharing an Allele that Is Identical by Descent (IBD) at a Genetic Marker with a Common Ancestor

Relationship	Level	P	0 alleles ( $k_0$ )	1 allele ( $k_1$ )	2 alleles ( $k_2$ )
Self or identical twin	Same DNA	1,0	0	0	1
Parent–child	1 <sup>st</sup> degree	0,5	0	1	0
Full siblings	1 <sup>st</sup> degree	0,5	1/4	1/2	1/4
Half siblings*	2 <sup>nd</sup> degree	0,25	1/2	1/2	0
Uncle–nephew* or Aunt–niece*	2 <sup>nd</sup> degree	0,25	1/2	1/2	0
Grandparent–grandchild*	2 <sup>nd</sup> degree	0,25	1/2	1/2	0
First cousins	3 <sup>rd</sup> degree	0,125	3/4	1/4	0
Double first cousins	3 <sup>rd</sup> degree	0,125	9/16	3/8	1/16
Second cousins	4 <sup>th</sup> degree	0,062	15/16	1/16	0
Unrelated <sup>†</sup>	–	–	1	0	0

\* Half siblings, uncle- or aunt- nephew or niece, and grandparent–grandchild relationships are genetically identical with autosomal loci.

<sup>†</sup> Unrelated individuals can share alleles that are identical by coincidence (IBS) rather than IBD.

$k$  is the kinship coefficient where  $k_0$  is the probability that neither allele of one relative is inherited by the second relative,  $k_1$  is the probability that one allele of one relative is inherited by the second relative, and  $k_2$  is the probability that both alleles of one relative are inherited by the second relative. The level of confidence in a calculated relationship decreases with the degree of separation between them. Adapted from Wenk et al. (1996) and Weir (2007).

In questi casi, è di fondamentale importanza interrogarsi su quale sia il grado di «somiglianza» genetica atteso tra diversi tipi di parenti, che può essere ricostruito tenendo in considerazione semplicemente le leggi di Mendel.

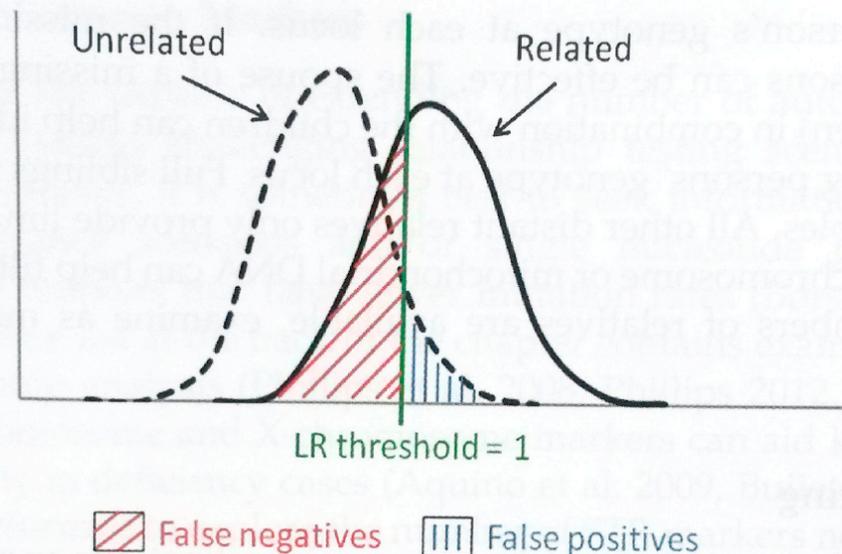
Il coefficiente di parentela (P) indica la proporzione di alleli condivisi per discesa da due individui.

PSI (Proportion of shared alleles) come misura grezza del grado di parentela...

# Test di parentela

Analogamente a quanto avviene per il test di paternità, i test di parentela vengono effettuati confrontando i due profili di STR ed utilizzando un **rapporto di due probabilità (LR, Likelihood Ratio)**

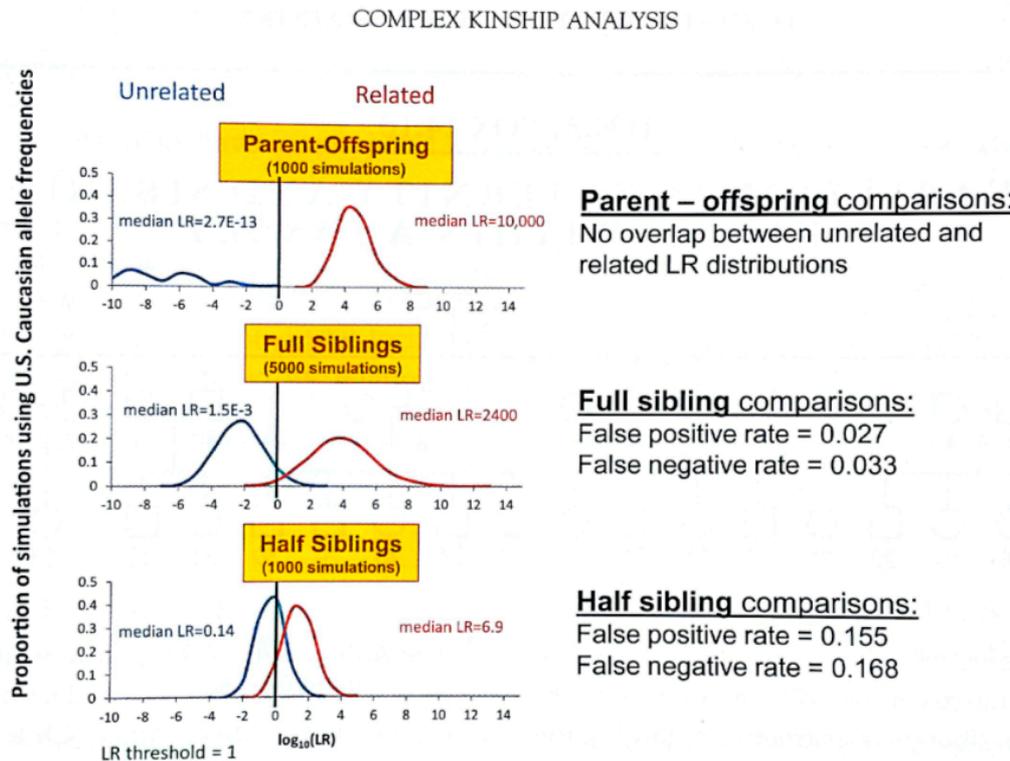
- Al numeratore la probabilità di osservare quei profili sotto l'ipotesi che due individui abbiano un determinato tipo di parentela
- Al denominatore la probabilità di osservare quegli stessi profili sotto l'ipotesi che i due individui siano tra loro non imparentati



**FIGURE 14.5** Illustration of the overlap in likelihood ratio (LR) distributions between related and unrelated individuals showing where false positives (unrelated individuals that have an  $LR > 1$ ) and false negatives (related individuals that have a  $LR < 1$ ) can arise. Figure courtesy of Kristen O'Connor, former NIST postdoc. Adapted from O'Connor (2010).

# Test di parentela

Si noti come il calcolo di LR nei test di parentela tra parenti (con l'esclusione delle relazioni genitori figli) **può facilmente condurre a falsi positivi e falsi negativi**



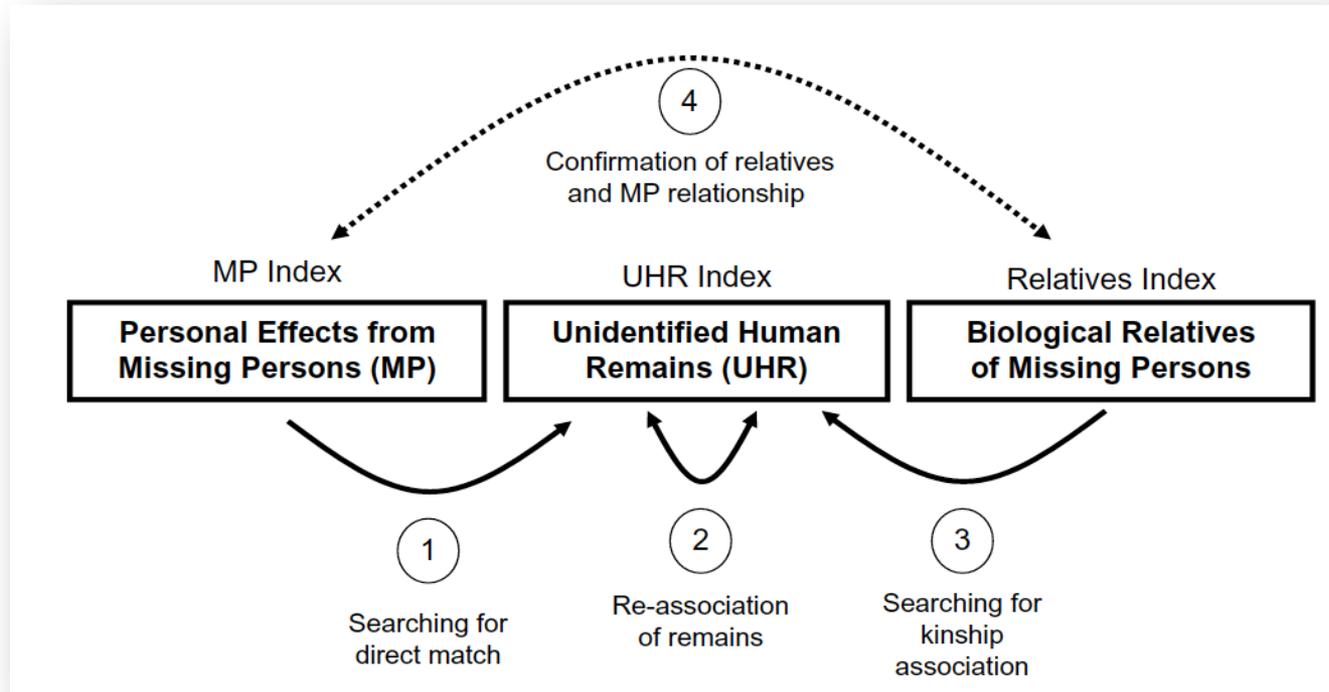
Il possibile errore è tanto **MEGGIORE** quanto

✓ è **più basso** il **coefficiente di parentela**

✓ È **minore** il **numero di loci utilizzato**

**FIGURE 14.6** Likelihood ratio (LR) distributions calculated through simulating pairwise comparisons of related and unrelated individuals using 13 STR loci and U.S. Caucasian allele frequencies. The degree of overlap in the LR distributions corresponds with possible false positive or false negative results (see Figure 14.5). Full siblings and half siblings may have true relatives with LR < 1 because these individuals may not share any alleles at a locus. Adding more STR loci can shift the related distribution to the right. Figure courtesy of Kristen O'Connor, former NIST postdoc. Adapted from O'Connor (2010).

## Identificazione di persone scomparse



In questo tipo di indagini, si utilizzano frequentemente:

- (a) **database di profili di DNA di persone scomparse (MP);**
- (b) **database di profili di resti umani non identificati (UHR);**
- (c) **database di profili di parenti di persone scomparse.**

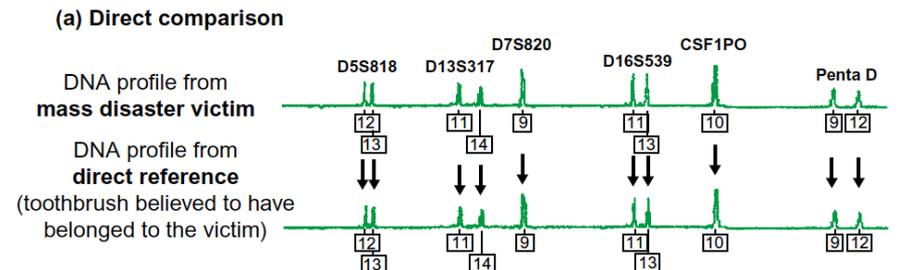
## Identificazione delle vittime dei disastri di massa (DVI)

Problema in parte analogo al precedente. In questo caso i resti da «identificare» appartengono spesso ad un numero *definito* (relativamente piccolo) di persone per le quali spesso si conosce una lista di nominativi. Anche in questo caso il profilo di riferimento può essere ottenuto in diversi modi:

- 1) da campioni provenienti dall'individuo scomparso
- 2) da parenti delle persone

### Problemi:

- Nei casi di frammentazione e dispersione dei reperti: step di «riassociazione» basato sui profili genetici dei singoli reperti
- Spesso DNA fortemente degradato a causa di incendi o avverse condizioni ambientali
- Rispetto all'identificazione di persone scomparse, essendo limitato il numero di persone cui possono appartenere i resti, è possibile utilizzare sistemi polimorfici meno informativi



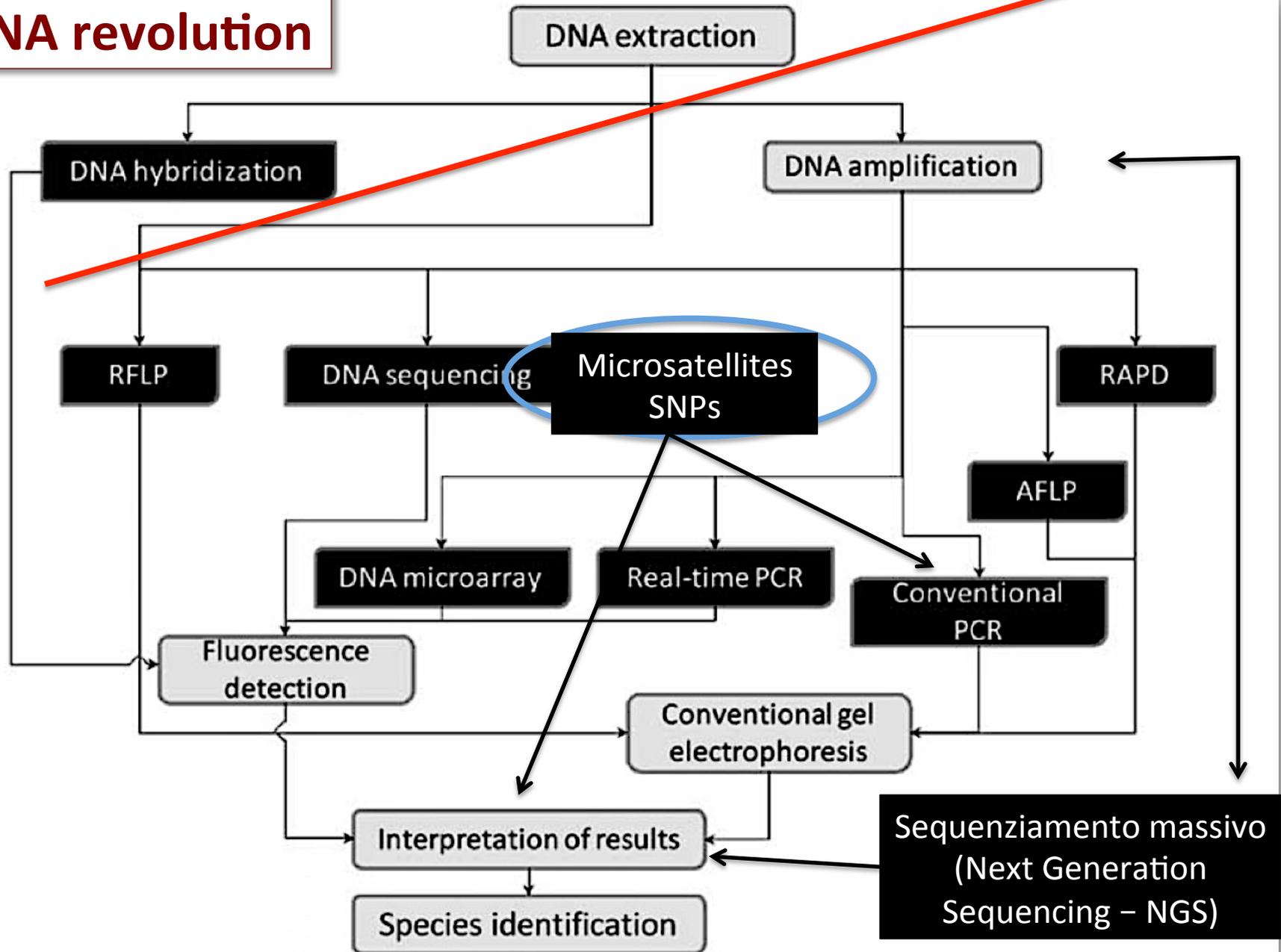
### (b) Kinship analysis

	<u>D5S818</u>	<u>D13S317</u>	<u>D7S820</u>	<u>D16S539</u>	<u>CSF1PO</u>	<u>Penta D</u>	
victim							
?							wife
wife							
?							son
son							
Predicted victim profile	11, ? or ? , 13	? , 14	9, ?	? , 13	? , 10	9, ?	victim (father)
Mass disaster victim profile	12, 13	11, 14	9, 9	11, 13	10, 10	9, 12	actual profile

- Il ricongiungimento di familiari di un immigrato può essere in alcuni stati subordinato alla effettuazione di test genetici di parentela.
- Spesso i test di parentela possono essere d'aiuto nelle indagini criminali, permettendo, attraverso il tracciamento di relazioni di parentela, l'identificazione di individui che hanno commesso dei reati.

**Un caso da manuale, che ha coinvolto test di parentela multipli e di diverso tipo, è quello relativo al caso Gambirasio (caso di studio alla fine del corso)**

# DNA revolution



# SNPs = Single Nucleotide Polymorphisms

GATTAGATCGCGATAGAG

GATTAGATCTCGATAGAG



- Differenze puntiformi tra sequenze omologhe (transizioni, transversioni o indel)
- Originano per **MUTAGENESI** o **ERRORE DI REPLICAZIONE**



Più frequenti a causa di agenti  
mutageni chimici e fisici

Sistemi di riparazione DNA:

- *Mismatch repair*
- *Nucleotide excision repair*
- Ricombinazione omologa
- *End-joining* non omologo



Frequenza di  $10^{-9}$ - $10^{-11}$   
per nucleotide

Attività esonucleasica  
della DNA polimerasi

# SNPs = Single Nucleotide Polymorphisms

GATTAGATCGCGATAGAG

GATTAGATCTCGATAGAG



- Differenze puntiformi tra sequenze omologhe (transizioni, transversioni o indel)
- Originano per MUTAGENESI o ERRORE DI REPLICAZIONE
- **Silenti o Sinonime** = Sostituzione che NON altera la codifica di un aa
- **Non-sinonime o missenso** = Sostituzione che altera la codifica di un aa
- **Non-senso** = Sostituzione che trasforma un codone per un aa in uno di STOP
- **Frameshift** = Indel nella regione codificante che causa lo slittamento della lettura del codice genetico (le più dannose)

# SNPs = Single Nucleotide Polymorphisms

GATTAGATCGCGATAGAG  
GATTAGATCTCGATAGAG

- Polimorfismo più abbondante nel genoma.

Diversità nucleotidica media stimata è  $7,51 \times 10^{-4}$  (1SNP ogni 1331 bp circa)

Genoma umano  $3,3 \times 10^9$  bp gli SNPs sono stimati nell'ordine di 3 milioni.

1,42 milioni già identificati

Recenti stime sulle frequenze dell'allele raro si spingono a oltre 11 milioni di SNPs

- **Basso tasso di mutazione:** prevalentemente loci biallelici (massimo 4).
- Sono meno variabili se comparati ai microsatelliti ma questo limite è compensato dalla loro frequenza e dalla facilità di isolamento.

## SNPs

- within coding region,
- in non-coding regions of genes,
- in the intergenic regions.

**This should be taken into account (if possible) when the panel of SNPs to be used is established**

Ciclo vitale di uno SNP:

- Comparsa
- Sopravvivenza
- Aumento di frequenza
- **FISSAZIONE**

284 mila anni



Marcatore utile per gli scopi delle scienze forensi

## Genotyping Technologies

1. RFLP
2. PCR (allele specific primers)
3. Oligonucleotide ligation
4. Primer extension (incorporate labeled nucleotides)
5. Hybridization (microarray)
6. Taqman (RT-PCR)
7. Sequencing (whole genome or targeted)
8. Genotyping by sequencing

.....and many other ones !!!

## SNPs

### Vantaggi:



- Elevata riproducibilità
- **CO-DOMINANTI**
- **EREDITA' MENDELIANA**
- Molto frequente nel genoma e presenti in tutte le regioni
- **Informativi sul fenotipo**
- **Elevate potenzialità con NGS (Next Generation Sequencing)**

### Svantaggi

- Richiede informazioni a priori (?????)
- Cross-amplificazione tra specie spesso limitata
- Poco variabili
- Attenzione a non confonderli con errori di sequenziamento



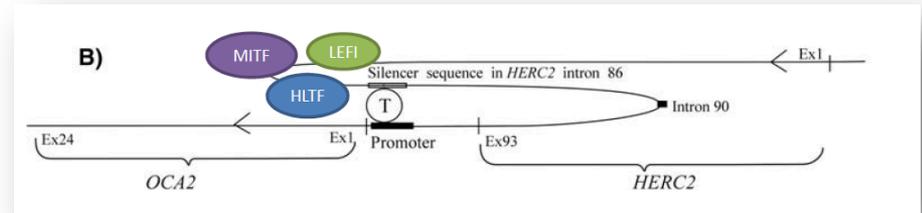
# SNPs – marcatori informativi sul fenotipo

## Colore degli occhi:

Il colore dell'occhio dipende principalmente dalla quantità totale di melanina nello strato superficiale dell'iride e dal rapporto quantitativo tra due diversi tipi di melanina (eumelanina e pheomelanina).

Da un punto di vista genetico, **il colore è un carattere multifattoriale poligenico, con un elevato grado di ereditabilità.**

Dei numerosi geni che contribuiscono al fenotipo, riveste particolare importanza OCA2, che codifica per un enzima coinvolto nella maturazione dei melanociti. **L'allele A dello SNP rs12913832**, localizzato in un enhancer del gene OCA2 (che si trova all'interno di un introne del gene adiacente HERC2), è associato ad un maggiore livello di espressione del gene, determinando frequentemente la **comparsa di occhi marroni piuttosto che azzurri (allele G).**



SNP rs12913832,  
Relazione tra genotipo e fenotipo

AA	In Europeans, 85% chance of brown eyes; 14% chance of green eyes; 1% chance of blue eyes.
AG	In Europeans, 56% chance of brown eyes; 37% chance of green eyes; 7% chance of blue eyes.
GG	In Europeans, 72% chance of blue eyes; 27% chance of green eyes; 1% chance of brown eyes.

# SNPs – marcatori informativi sul fenotipo

Colore degli occhi:

Sistema Irisplex, una multiplex basata su sei SNPs in altrettanti geni (HERC2, OCA2, SLC24A4, SLC45A2, TYR e IRF4), in grado di individuare, con un buon grado di confidenza, il colore degli occhi di un individuo a partire dal suo DNA.

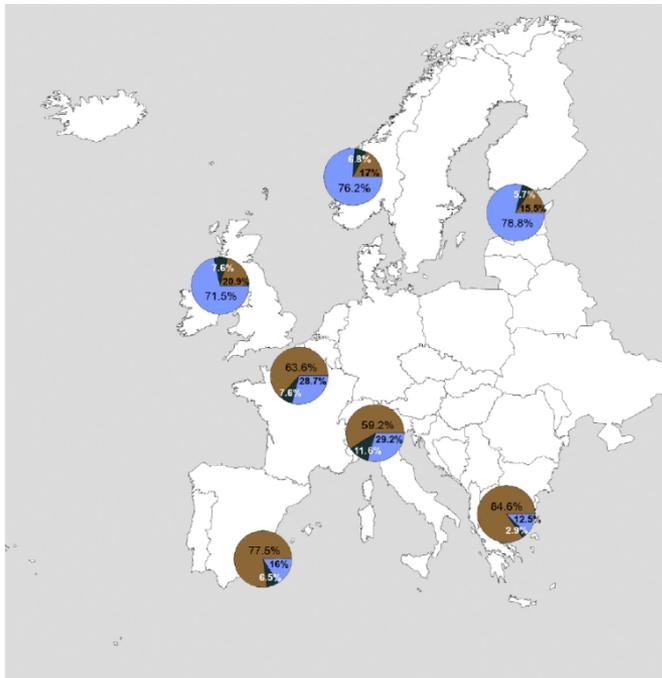


Fig. 1. Eye colour phenotype frequencies across Europe representing Bergen (Norway) (N = 547), Tallinn (Estonia) (N = 579), Belfast (UK) (N = 498), Paris (France) (N = 616), Verona (Italy) (N = 542), Thessaloniki (Greece) (N = 547) and Alicante (Spain) (N = 511). Numbers indicate the total number of individuals observed. Within each pie chart, the blue colour represents the frequency of blue eyed individuals, brown represents the frequency of brown eyed individuals, and the black colour represents the frequency of individuals whose eye colour could not be defined clearly into either the blue or the brown category from the image analysis. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of the article.)

# Test di parentela

Identificazione delle persone disperse nel World trade center (NYC) in seguito all'attacco terroristico del 11/9

- Sono stati generati
  - 52,000 profili STR
  - 44,000 profili mtDNA
  - 17,000 profili SNPs
- 850 delle 1594 vittime identificate grazie all'analisi del DNA



Photo © 2001 Caolina Salguero/ Sips