

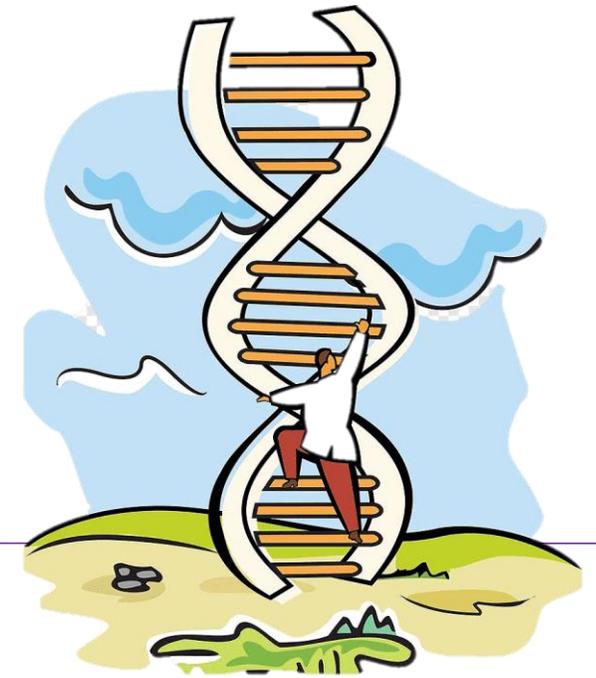
Analisi molecolari per il fingerprinting delle varietà vegetali

STATO DELL'ARTE



Alberto Acquadro

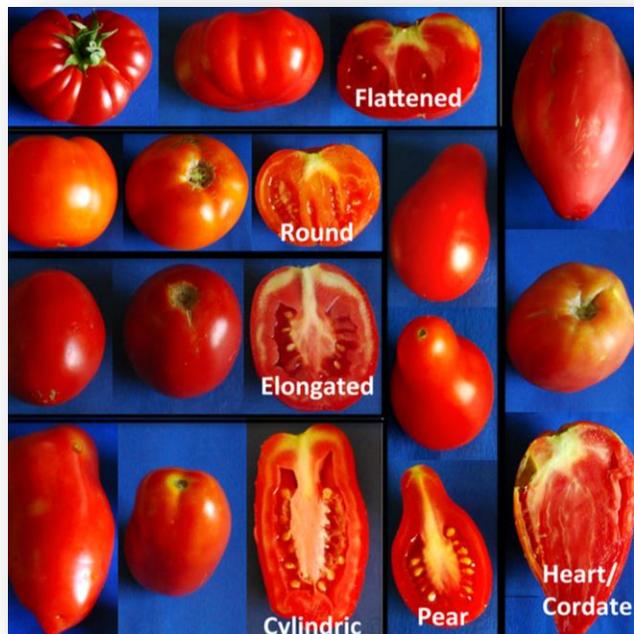
DISAFA – Genetica Vegetale
Università degli studi di Torino



MARCATORE GENETICO

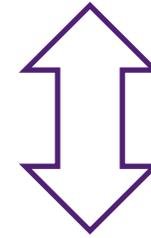
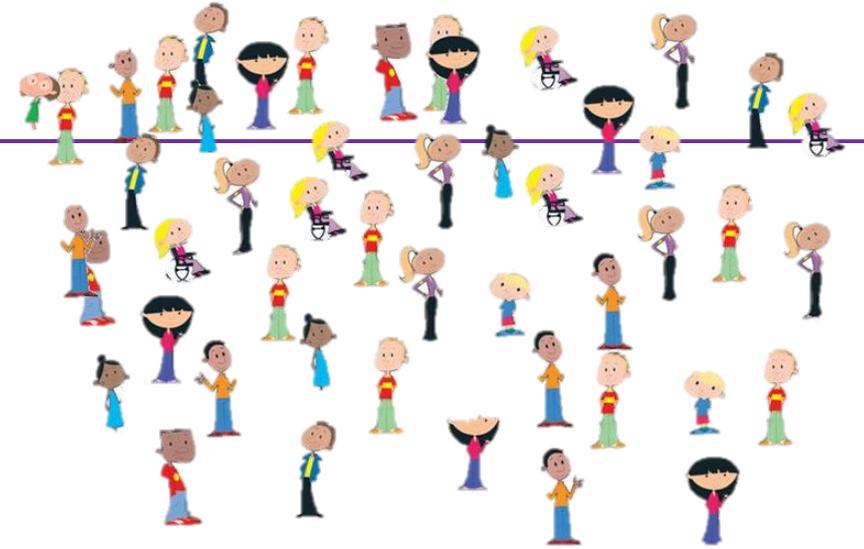
MARCATORE GENETICO

carattere mendeliano che può essere utilizzato per seguire la segregazione di una particolare regione cromosomica



MARCATORI GENETICI: IMPIEGHI

- Analisi di paternità
- Diagnosi di anomalie genetiche
- **Analisi forensi**
- Identificazione di QTL e geni utili
- Miglioramento delle specie vegetali ed animali (MAS)
- Studio della struttura, evoluzione e biodiversità delle popolazioni
- Studi di epidemiologia
- **Tracciabilità dei prodotti animali e/o dei GMO**
- Conservazione della natura
- Analisi di campioni da erbari e collezioni
- Studi di filogenesi ed evoluzione



MARCATORI GENETICI: TIPOLOGIE

Marcatori morfologici - caratteri mendeliani

- colore della cariosside
- forma della foglia
- presenza assenza di corna

Marcatori biochimici

- gruppi sanguigni
- isoenzimi
- proteine (dell'endosperma, del plasma, del latte ecc.)

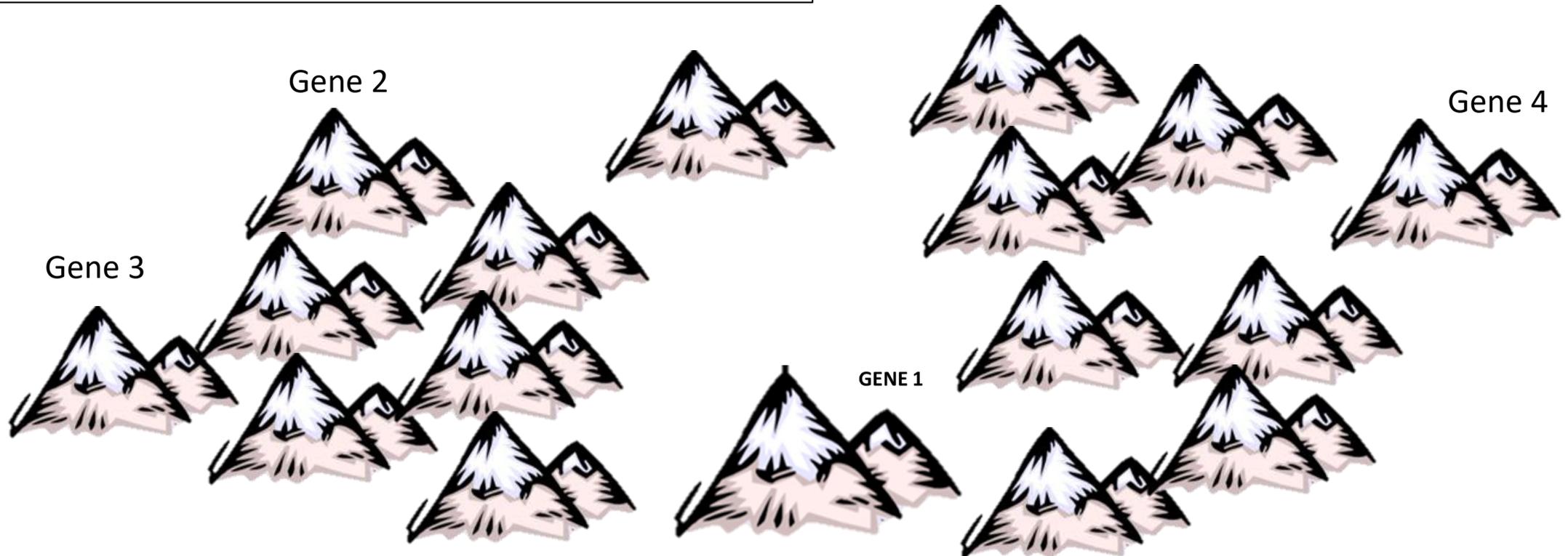
Marcatori molecolari

una sequenza DNA o proteica facilmente identificabile e la cui ereditabilità sia monitorabile



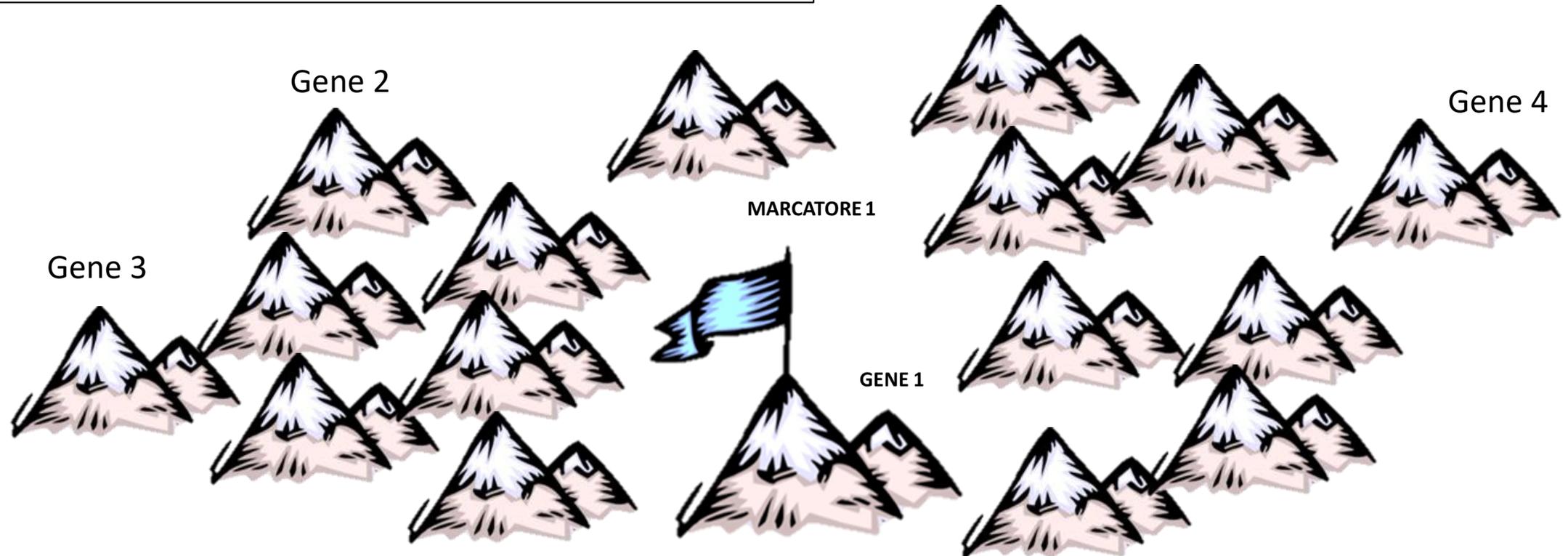
MARCATORI MOLECOLARI: CHE COSA SONO

Un marcatore è un locus che definisce in modo caratteristico e inequivocabile un tratto genomico



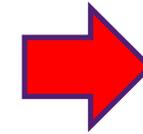
MARCATORI MOLECOLARI: CHE COSA SONO

Un marcatore è un locus che definisce in modo caratteristico e inequivocabile un tratto genomico

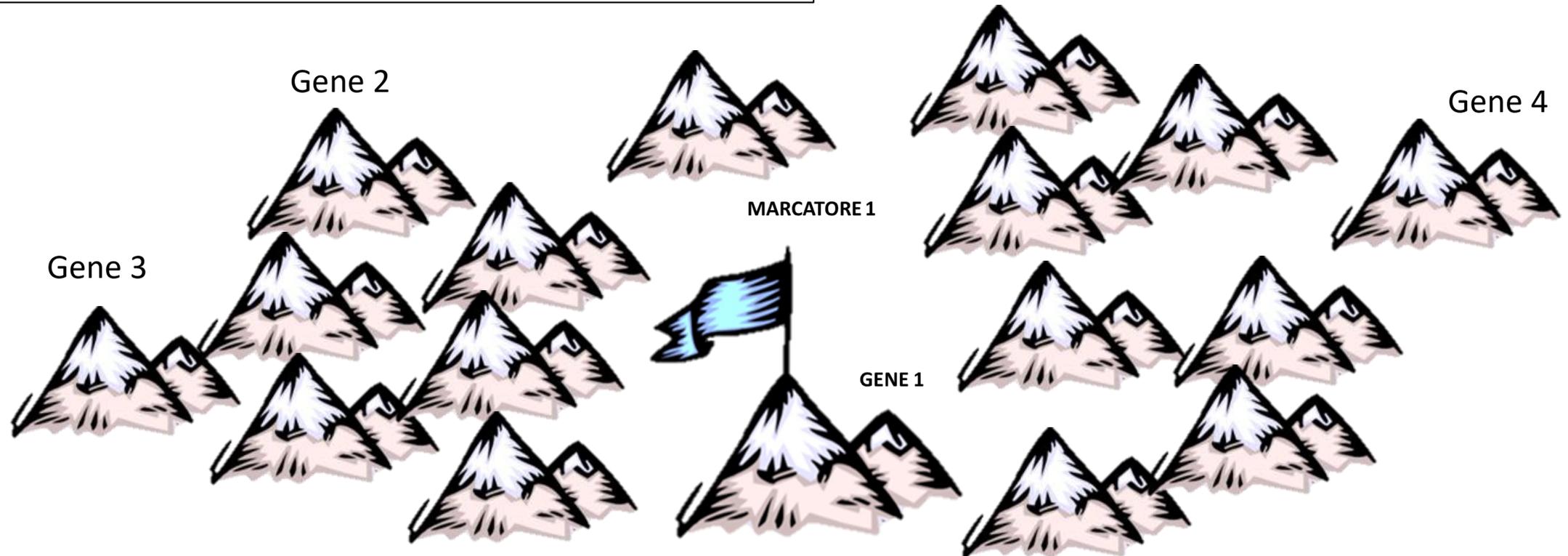


MARCATORI MOLECOLARI: CHE COSA SONO

Un marcatore è un locus che definisce in modo caratteristico e inequivocabile un tratto genomico



Rilevabile con metodologie molecolari



IL MARCATORE IDEALE

- Comportamento mendeliano
- Non influenzato dall'ambiente
- Stabile
- Facile da monitorare
- Numeroso
- Codominante
- Polimorfico (incluso anche sequenze non codificanti)
- Presente in qualsiasi tessuto
- Indipendente da sesso ed età
- Analisi automatizzabile
- Analisi veloce ed economica



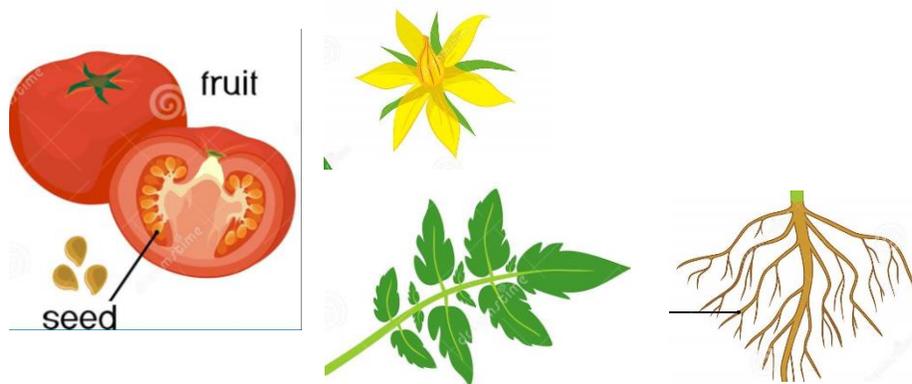
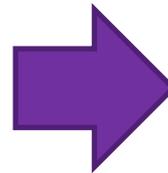
MARCATORI MOLECOLARI: VANTAGGI

Le caratteristiche morfologiche e biochimiche possono variare in relazione all'ambiente di coltivazione



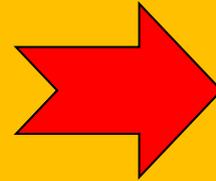
INFLUENZATI
DALL'AMBIENTE

TECNICHE DI ANALISI DEL DNA
IDENTIFICAZIONE DI UNA VARIETÀ
IN QUALSIASI STADIO DI SVILUPPO

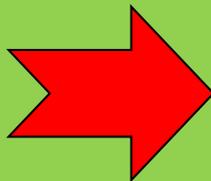


MARCATORI MOLECOLARI: VANTAGGI

DNA



- Marcatori molecolari**
- Potenzialmente infiniti
 - Non influenza ambiente

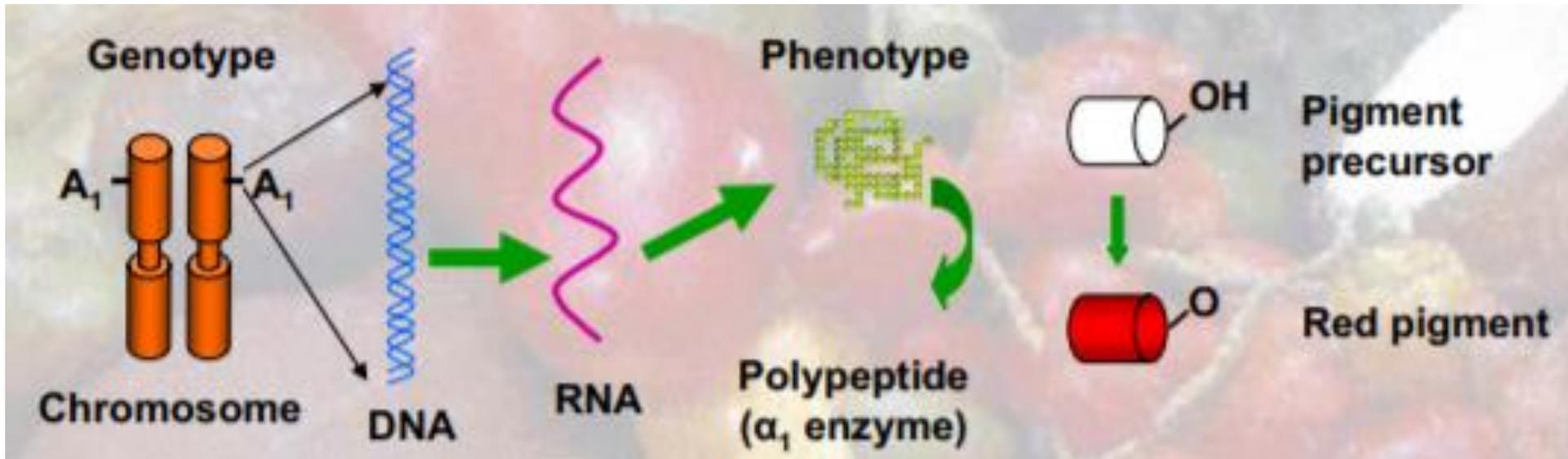


- Caratteri morfologici
e fisiologici**
- Pochi caratteri utilizzabili
 - influenzati dall'ambiente

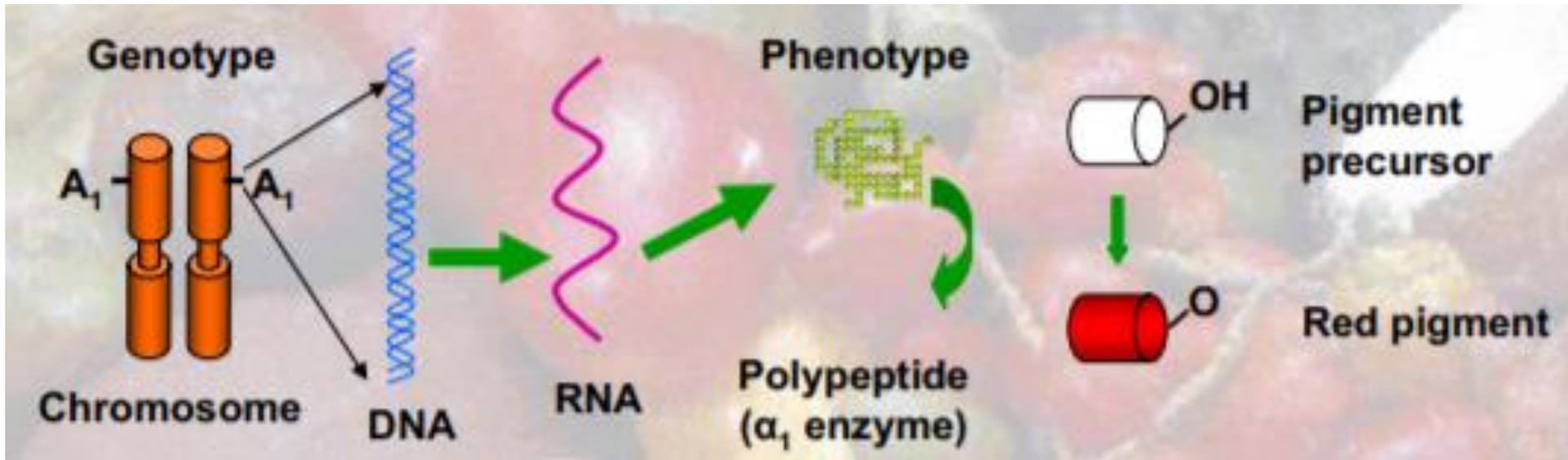


Aumenta il livello di precisione

IL FENOTIPO (CARATTERE MORFOLOGICO) HA SEMPRE UNA BASE NEL DNA

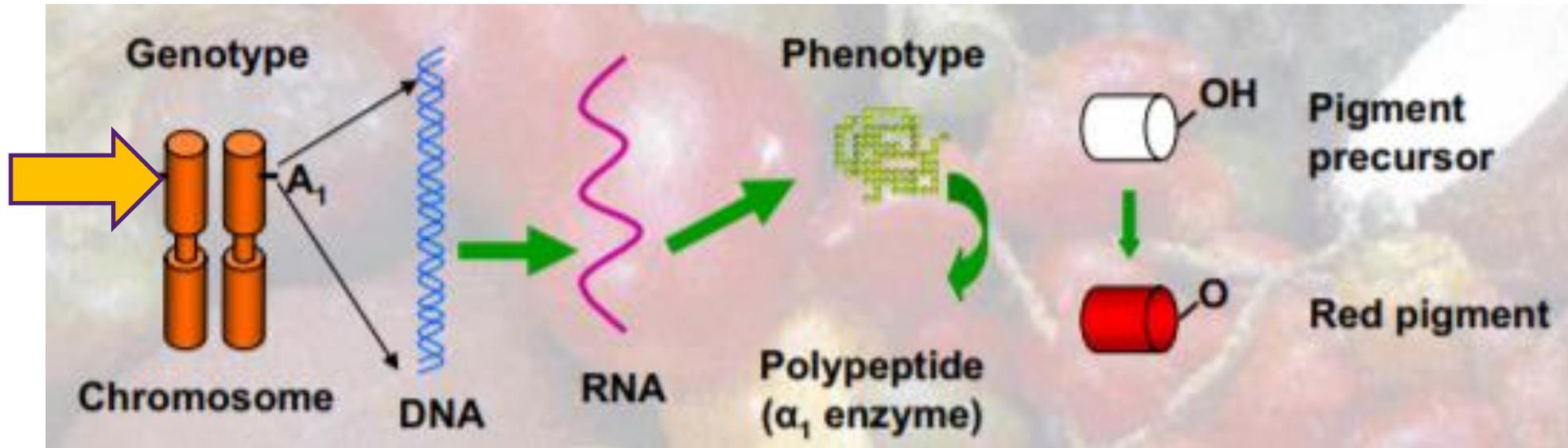


IL FENOTIPO (CARATTERE MORFOLOGICO) HA SEMPRE UNA BASE NEL DNA



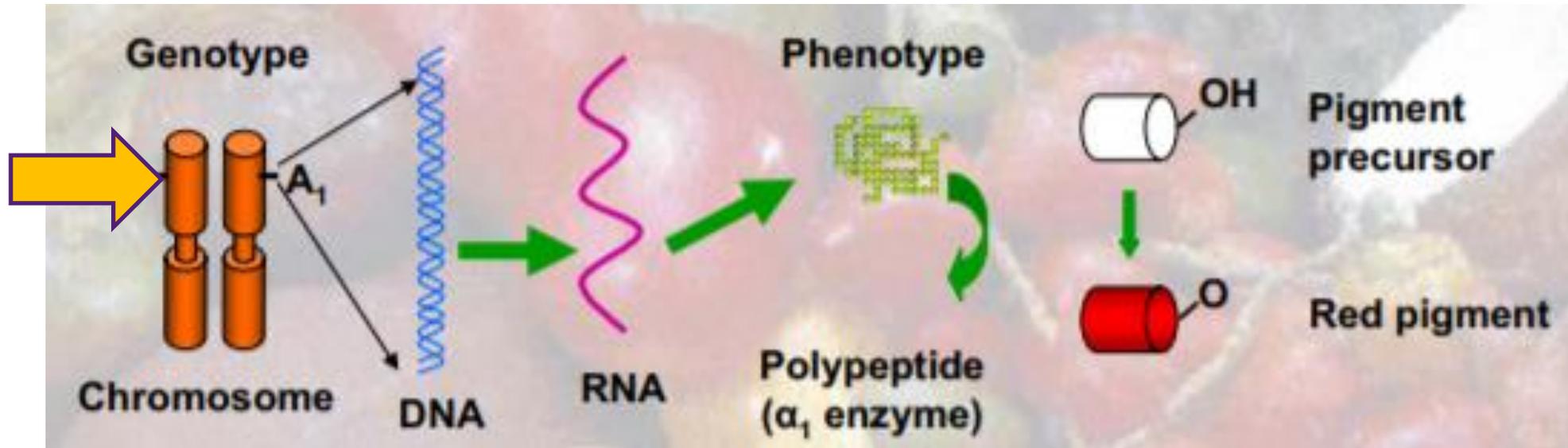
IL FENOTIPO (CARATTERE MORFOLOGICO) HA SEMPRE UNA BASE NEL DNA

MUTAZIONE in gene chiave



IL FENOTIPO (CARATTERE MORFOLOGICO) HA SEMPRE UNA BASE NEL DNA

MUTAZIONE in gene chiave



1 sola mutazione nel DNA può portare a differenze macroscopiche!



**Nel corso della domesticazione
(circa 10mila anni di storia dell'agricoltura)**

**Singole mutazioni chiave hanno modificato
uno o pochi geni determinando cambiamenti
Fenotipici incredibili**

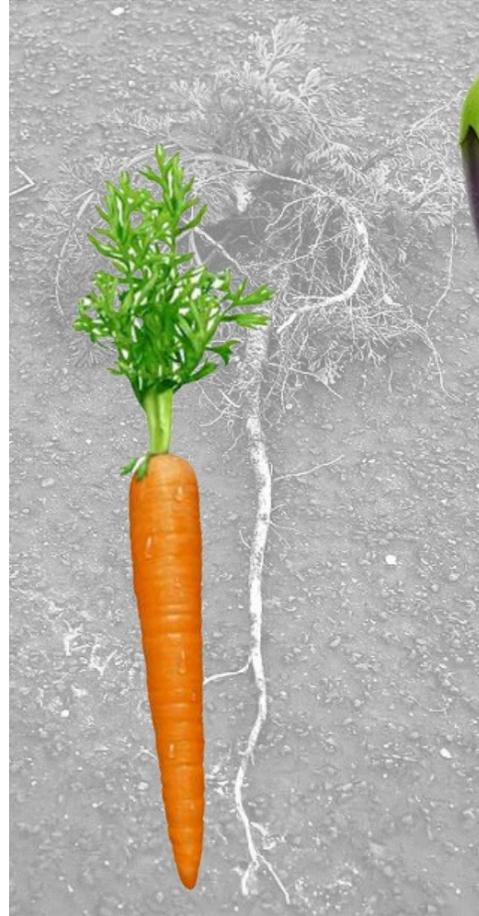
1 sola mutazione nel DNA può portare a differenze macroscopiche!

Guardate questi 5 esempi di piante selvatiche



È molto difficile collegarle con le «crops» di oggi

1 sola mutazione nel DNA può portare a differenze macroscopiche!

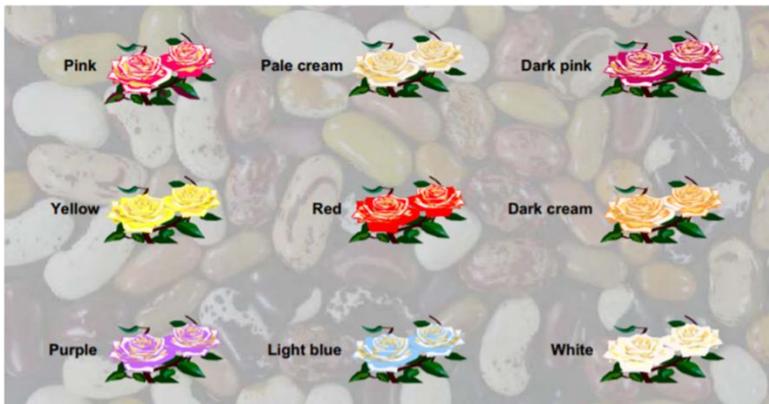


POLIMORFISMO

Il termine **polimorfismo** (dal greco, letteralmente, «avere molte forme») assume significati specifici in diversi contesti

In biologia/genetica si parla di polimorfismo genetico rispetto ad una variazione genetica e in particolare quando questa abbia una frequenza maggiore dell'1% nella popolazione.

Polimorfismo



Marcatori morfologico: colore

Monomorfismo



Marcatori morfologico: colore

POLIMORFISMO (A LIVELLO MOLECOLARE)

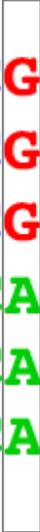
I polimorfismi a livello molecolare si possono classificare in:

- **polimorfismi di sequenza**
(conseguenza di differenze nelle sequenze di nucleotidi)
- **polimorfismi di inserzione e delezione**
(causati da mutazioni puntiformi)
- **polimorfismi nel numero di ripetizioni**

POLIMORFISMO DI SEQUENZA: **SNP**

SNP = Single Nucleotide Polymorphism

| | |
|-----------|---------------------|
| seq_1 (A) | ATGCGGC G AT |
| seq_2 (A) | ATGCGGC G AT |
| seq_3 (A) | ATGCGGC G AT |
| seq_1 (B) | ATGCGGC A AT |
| seq_2 (B) | ATGCGGC A AT |
| seq_3 (B) | ATGCGGC A AT |



Definizione

Un **polimorfismo a singolo nucleotide** è una variazione, del materiale genico a carico di un unico nucleotide tale per cui l'allele polimorfico risulta presente nella popolazione in una proporzione superiore all'1%. Al di sotto di tale soglia si è soliti parlare di variante rara (SNV = single nucleotide variant). Vengono rilevati mediante sequenziamento.

POLIMORFISMO DI INSERZIONE/DELEZIONE: **INDEL**

indel = insertion/deletion

| | |
|----------------|--------------------------|
| Indiv 1 | ATGGATATAGGCTGA-TAGGCTA |
| Indiv 2 | ATGGATATAGGCTGA-TAGGCTA |
| Indiv 3 | ATGGATATAGGCTGAC TAGGCTA |
| Indiv 4 | ATGGATATAGGCTGAC TAGGCTA |

Definizione

il termine **Indel** si usa come abbreviazione di un evento di mutazione/ricombinazione che può far parte di due classi: una inserzione (insertion) o una cancellazione (deletion). Vengono rilevate mediante sequenziamento (in genere insieme agli SNP)

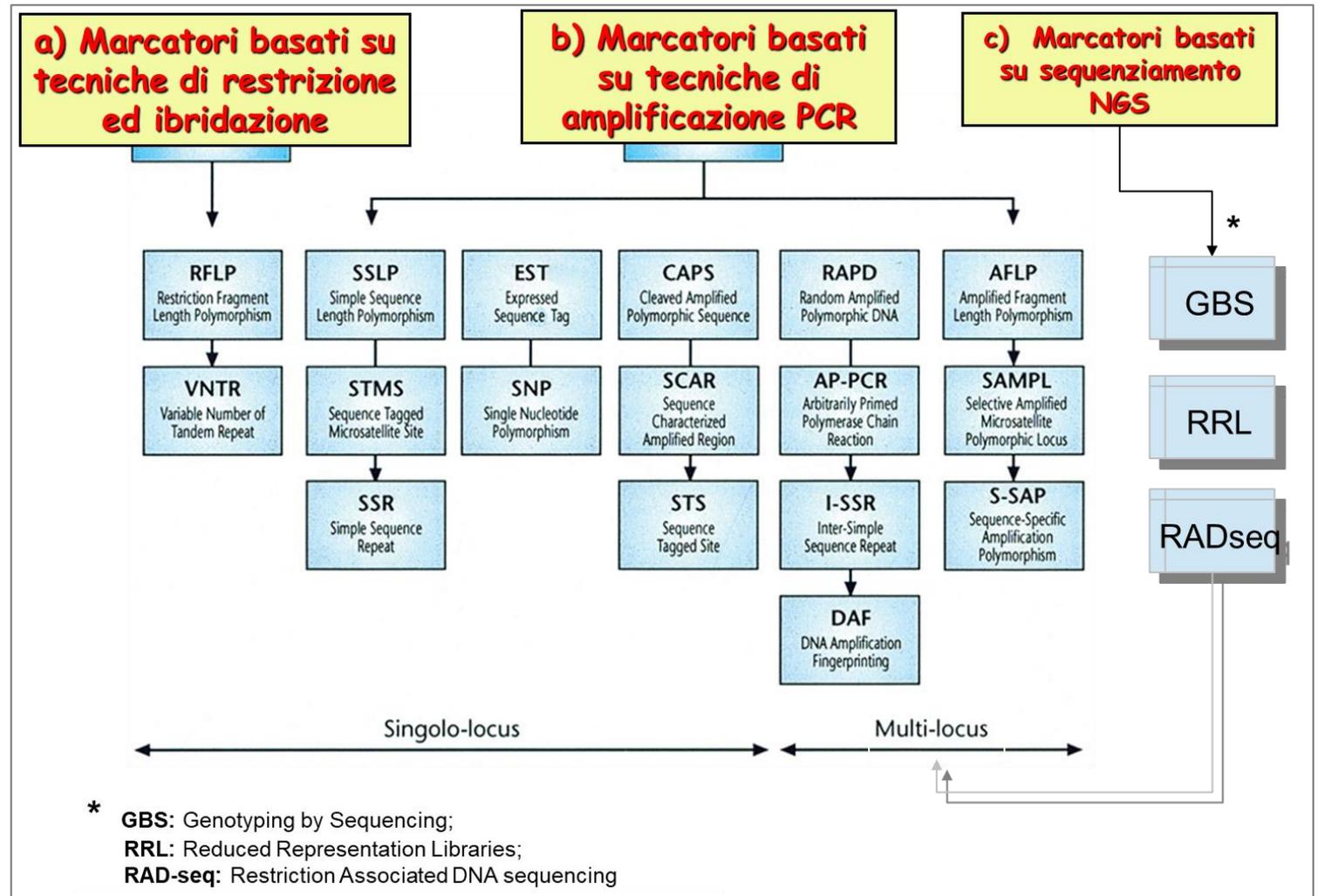
POLIMORFISMO NEL N° RIPETIZIONI: MICRO- MINI- MIDI-SATELLITI

Indiv 1 ATGGA**TATATATATA**GGCTGATAGGCTA
Indiv 2 ATGGA**TATA**GGCTGATAGGCTA
Indiv 3 ATGGA**TATA**GGCTGACTAGGCTA
Indiv 4 ATGGA**TATA**GGCTGACTAGGCTA

Definizione

Il polimorfismi legati al numero di ripetizioni ricadono, ad esempio, nei marcatori SSR (Simple Sequence Repeats) e sono formati da sequenze di DNA che si ripete in tandem. Alla base del loro grande successo c'è l'elevata variabilità intra-locus tra individui della stessa specie.

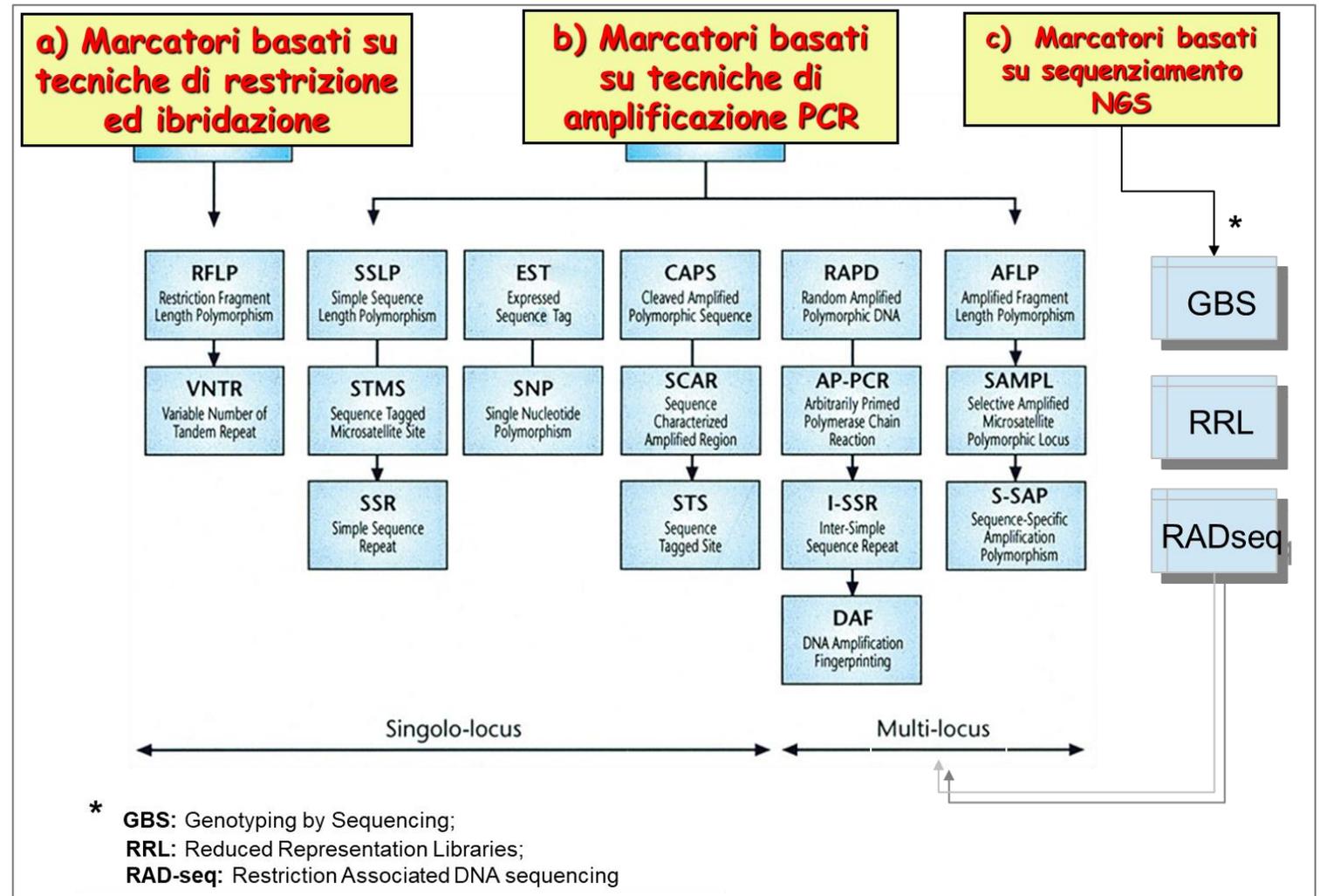
MARCATORI MOLECOLARI: METODOLOGIE



MARCATORI MOLECOLARI: METODOLOGIE

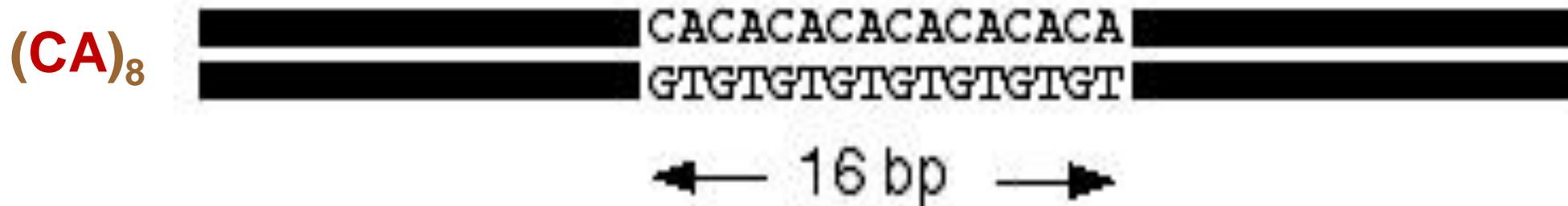


Molte tecniche



MICROSATELLITI: CHE COSA SONO

- DNA ripetuto in tandem
- Anche chiamati SSR (Simple Sequence Repeats)



MICROSATELLITI: CHE COSA SONO

Tipologie di SSR

AAAAAAAAAAAAA = (A)₁₃

MONO-nucleotidici SSR

ATATATATATAT = (AT)₆

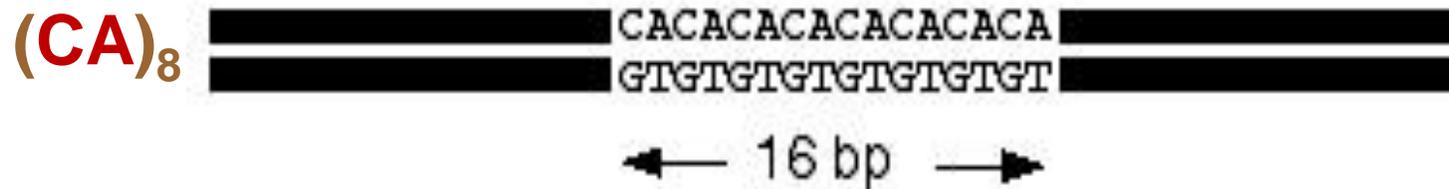
DI-nucleotidici SSR

CGACGACGACGA = (CGA)₄

TRI-nucleotidici SSR

ATCGATCGATCG = (ATCG)₃

TETRA-nucleotidici SSR



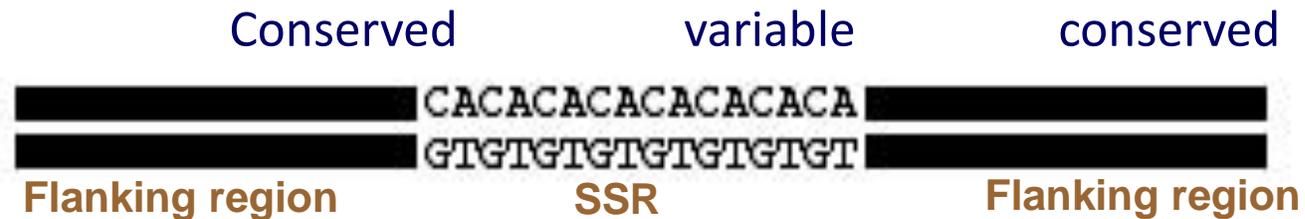
MICROSATELLITI: CHE COSA SONO

- Ripetizioni ubiquitarie in natura
- Presenti in tutti gli eucarioti:
 - Primati
 - Roditori
 - Altri mammiferi
 - non mammiferi (vertebrati)
 - insetti
 - Piante
- Meno riconosciuti nei procarioti

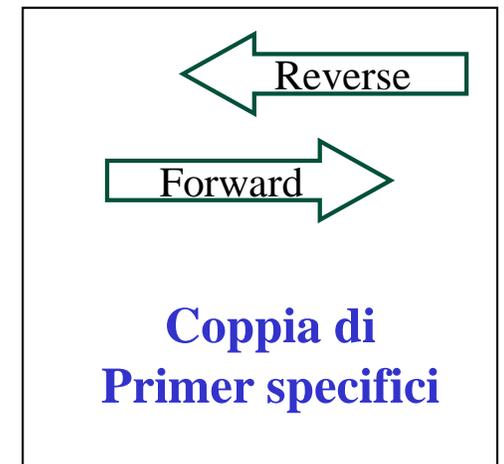
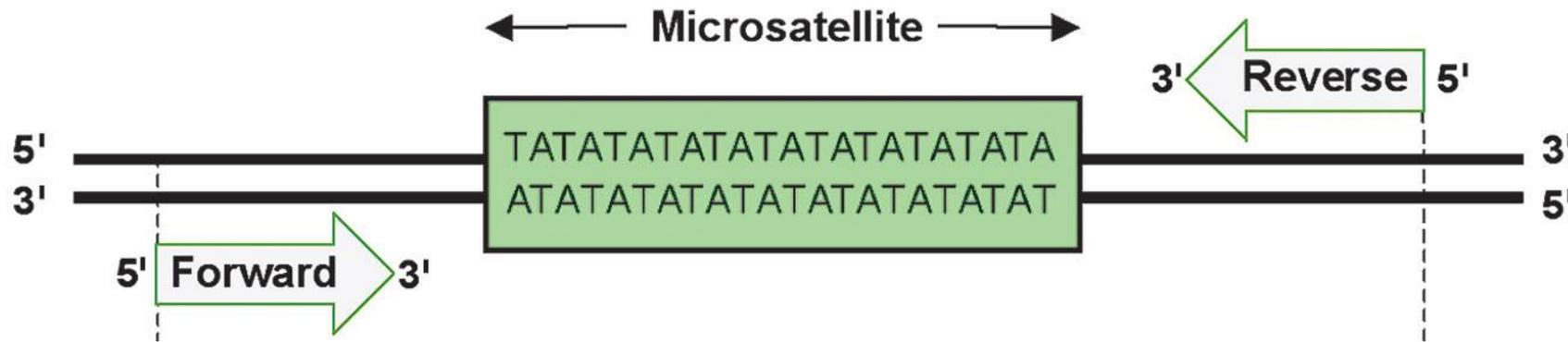
MICROSATELLITI: CARATTERISTICHE PRINCIPALI

Altamente variabili (elevato tasso mutazionale) a livello di:

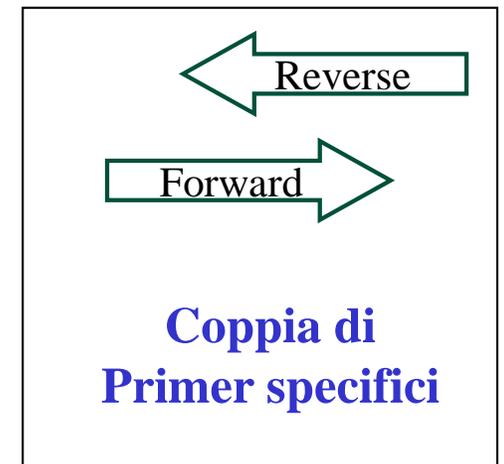
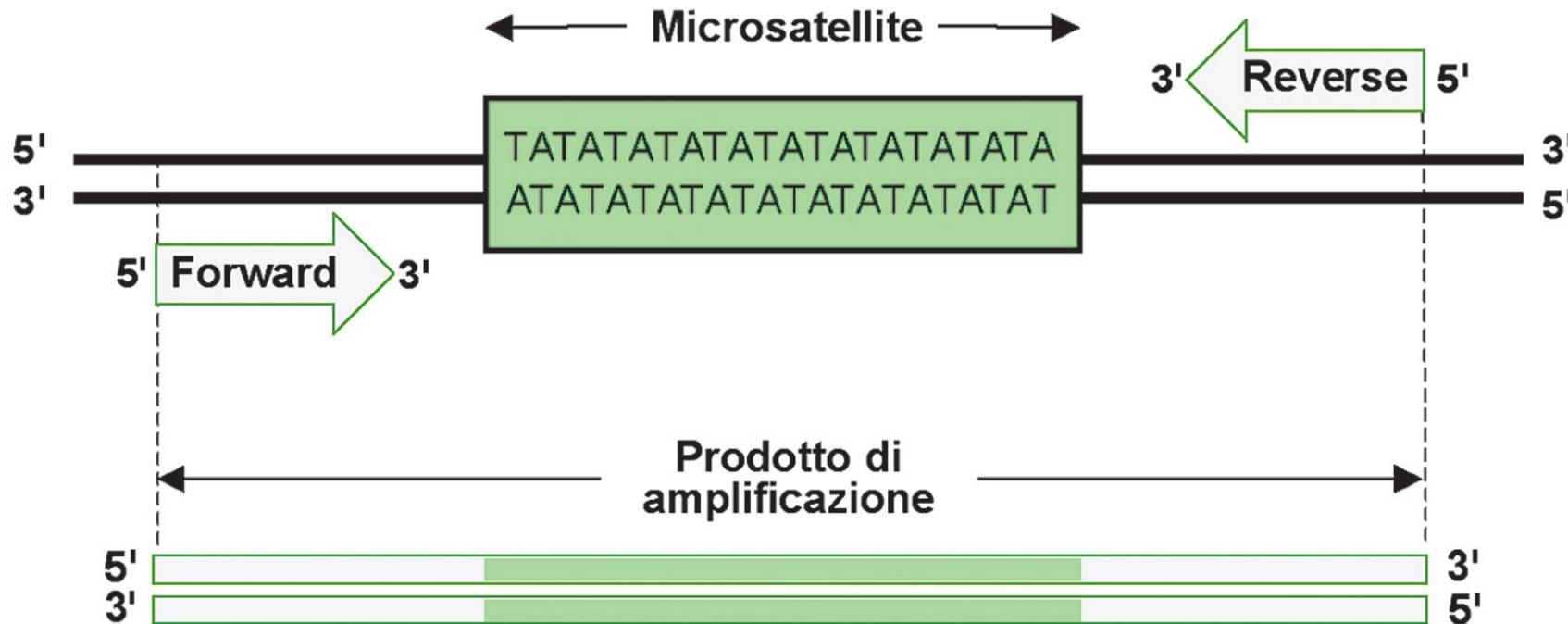
- **Ogni locus >>> molti alleli** (2-100 alleli).
- **Le regioni fiancheggianti sono conservate** entro specie
- **Distribuiti casualmente** nel genoma



CARATTERISTICHE DEI MICROSATELLITI: RILEVABILI CON PCR



CARATTERISTICHE DEI MICROSATELLITI: RILEVABILI CON PCR



APPLICAZIONE

Multiplex: 14 loci in contemporanea

Avere la sequenza di DNA

```
ATGAGCCTAGCACAGCTGATCAATTGCCTGTGTA
GTGTGTGTGTGTGTGTGTGGTGTGCAAGGGTG
GGTGGTGGGATAGAAAAGAGATC
```



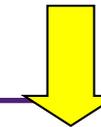
Forward primer



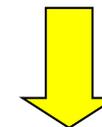
Reverse primer



**DISEGNARE 2 primer
FIANCHEGGIANTI il motivo SSR**



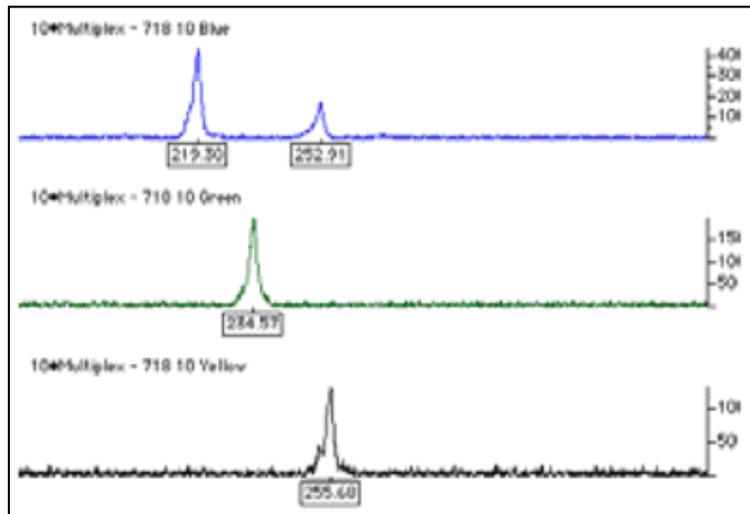
**AMPLIFICAZIONE
PCR**



VISUALIZZAZIONE

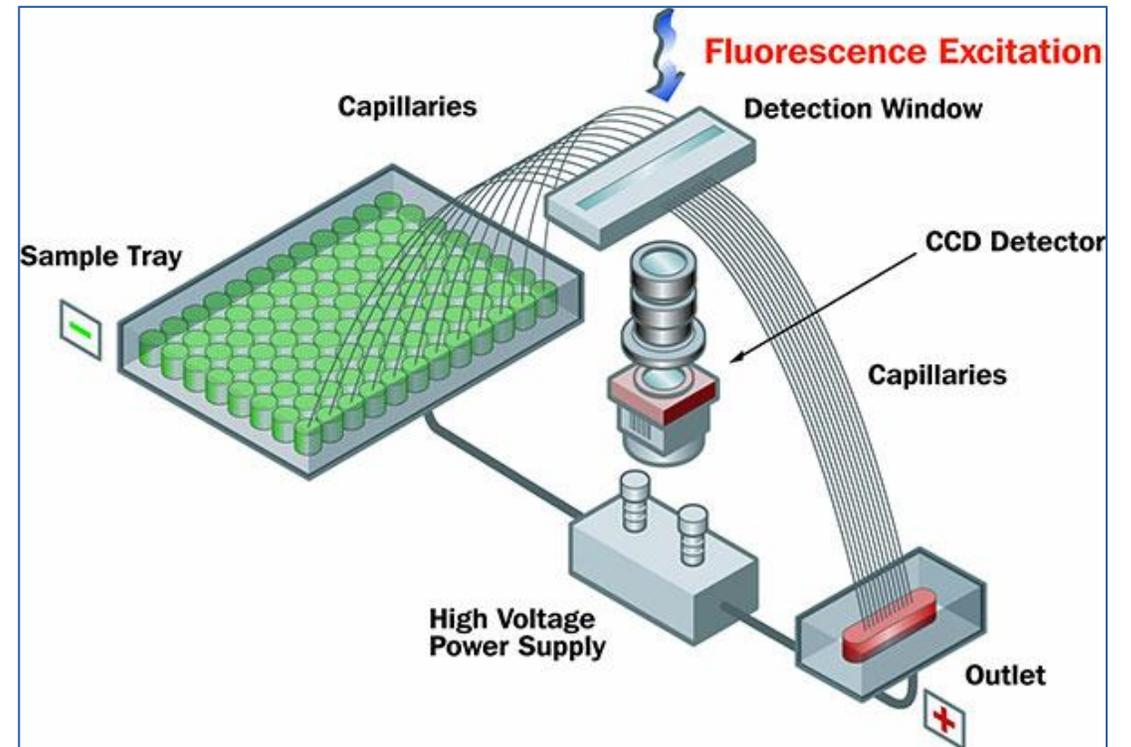


Elettroforesi capillare

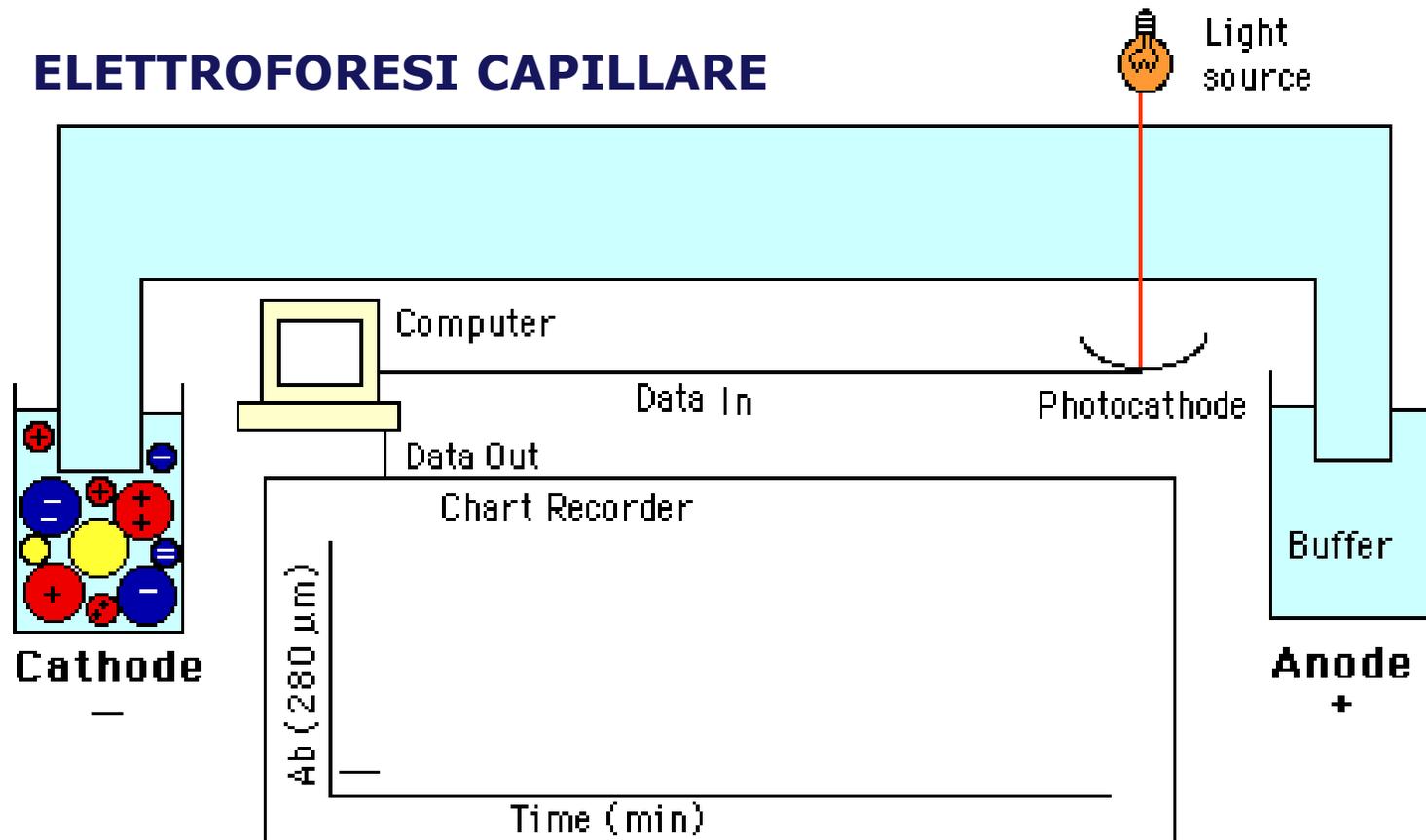


REQUISITI: **VISUALIZZAZIONE**

ELETTROFORESI CAPILLARE

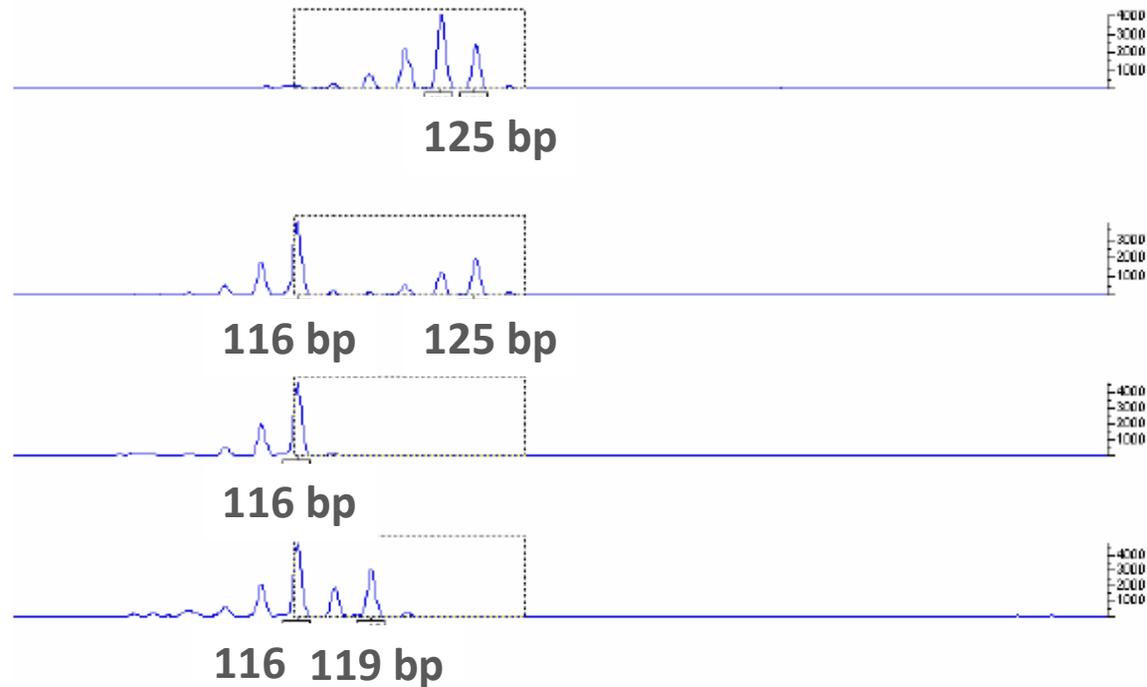


REQUISITI: CONOSCENZA SPECIE, DNA, PCR



REQUISITI: CONOSCENZA SPECIE, DNA, PCR

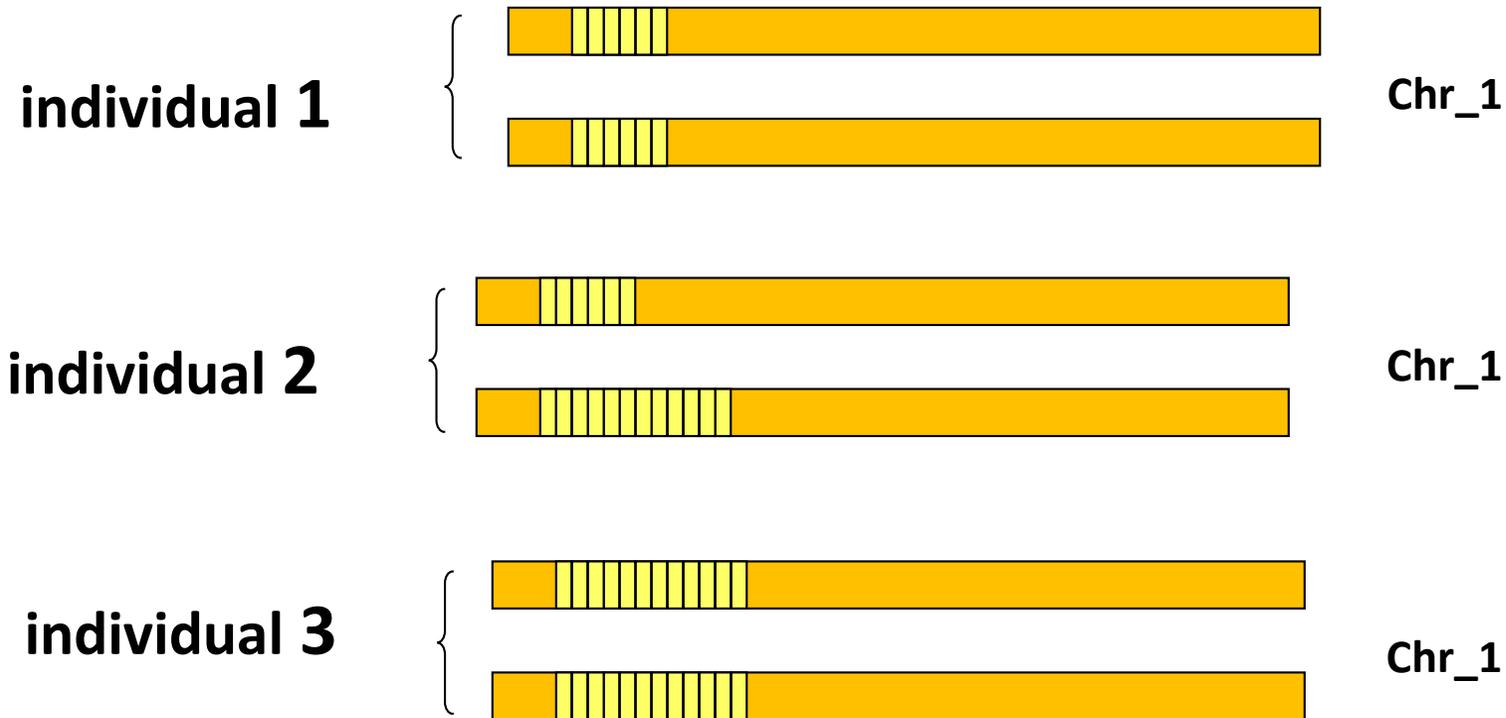
Elettroferogramma



L'intensità del segnale dipende dalla quantità di frammento e determina l'area e l'altezza dei picchi

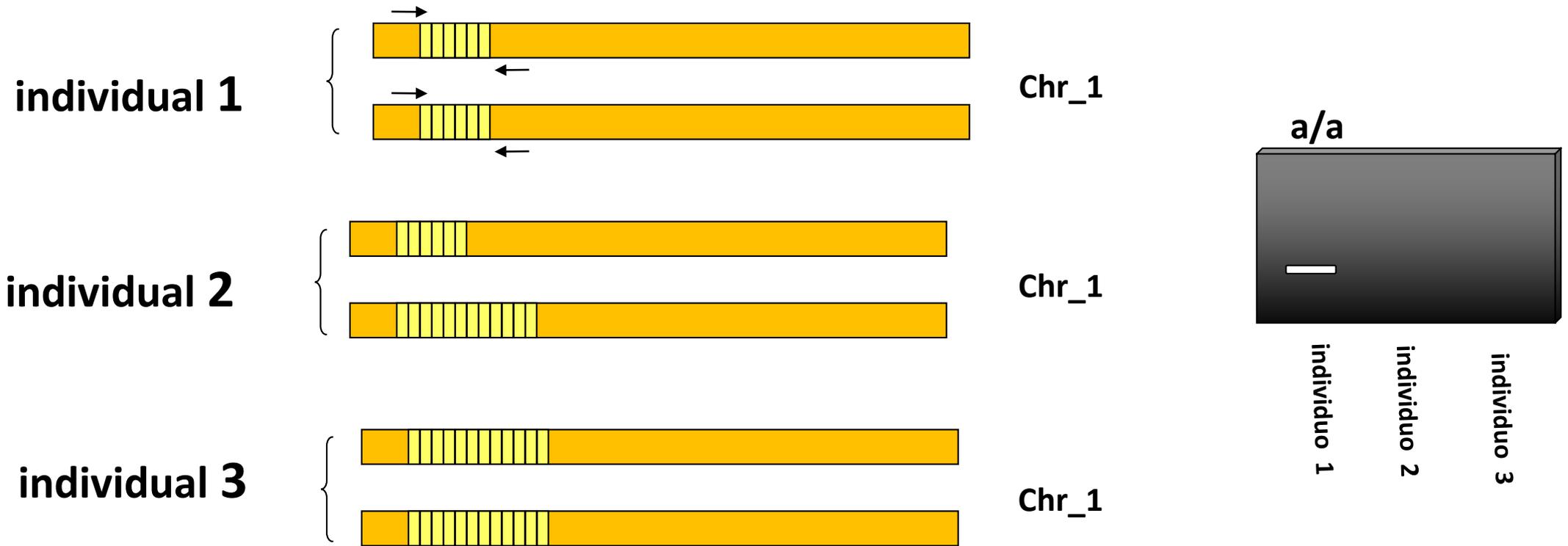
CARATTERISTICHE DEI MICROSATELLITI: CO-DOMINANTI

È POSSIBILE DISTINGUERE L'OMOZIGOTE DALL'ETEROZIGOTE



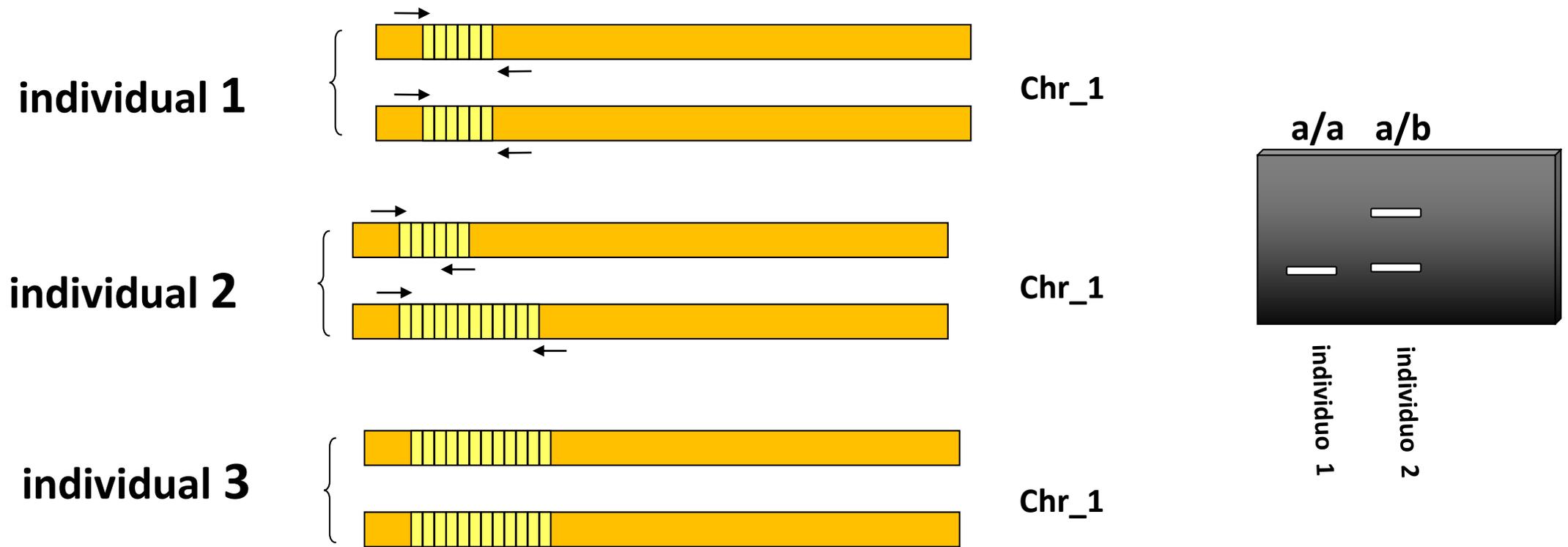
CARATTERISTICHE DEI MICROSATELLITI: CO-DOMINANTI

È POSSIBILE DISTINGUERE L'OMOZIGOTE DALL'ETEROZIGOTE



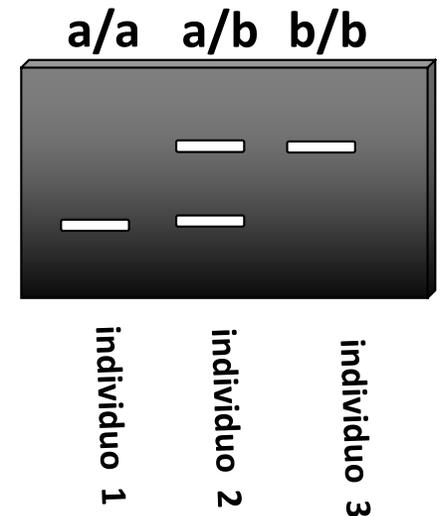
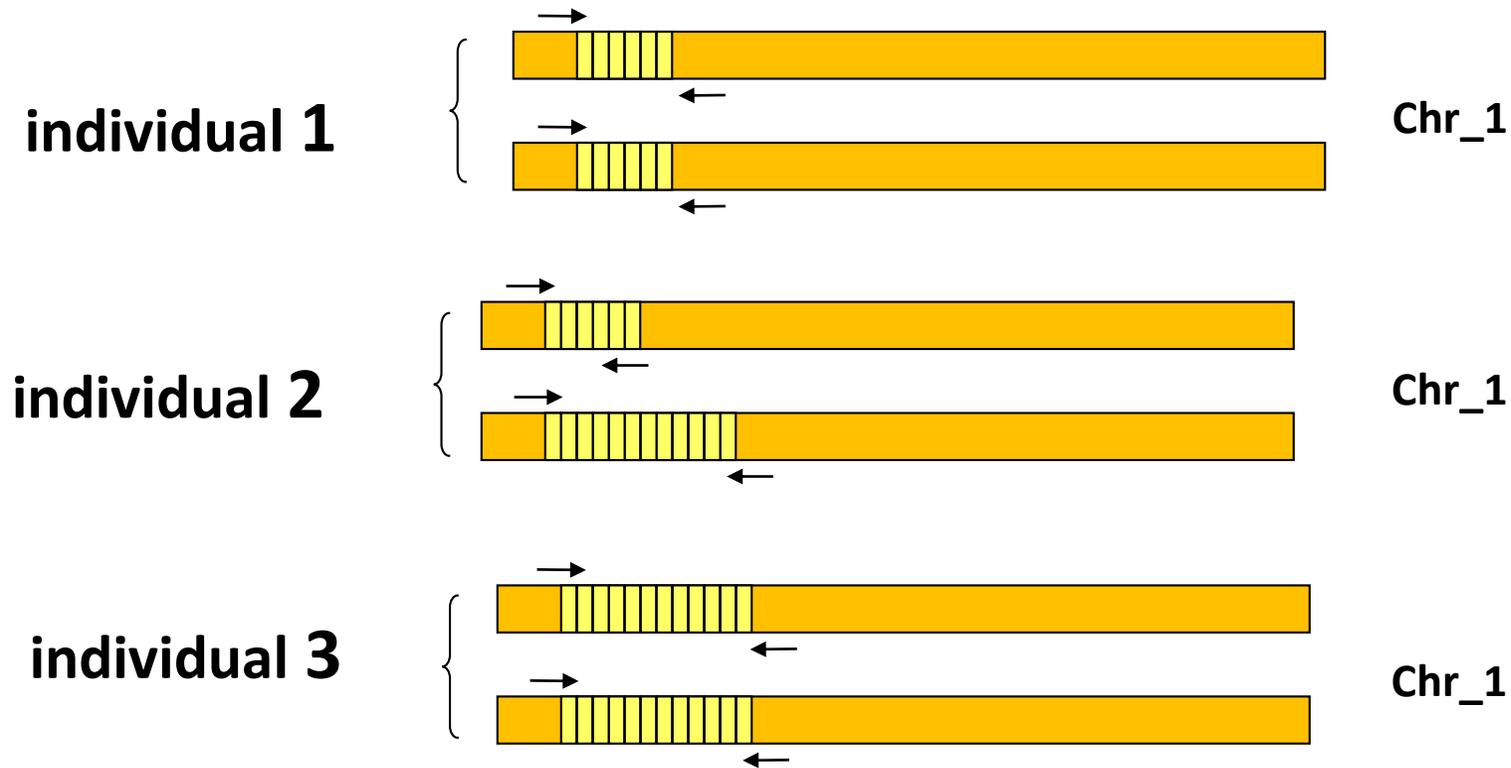
CARATTERISTICHE DEI MICROSATELLITI: CO-DOMINANTI

È POSSIBILE DISTINGUERE L'OMOZIGOTE DALL'ETEROZIGOTE



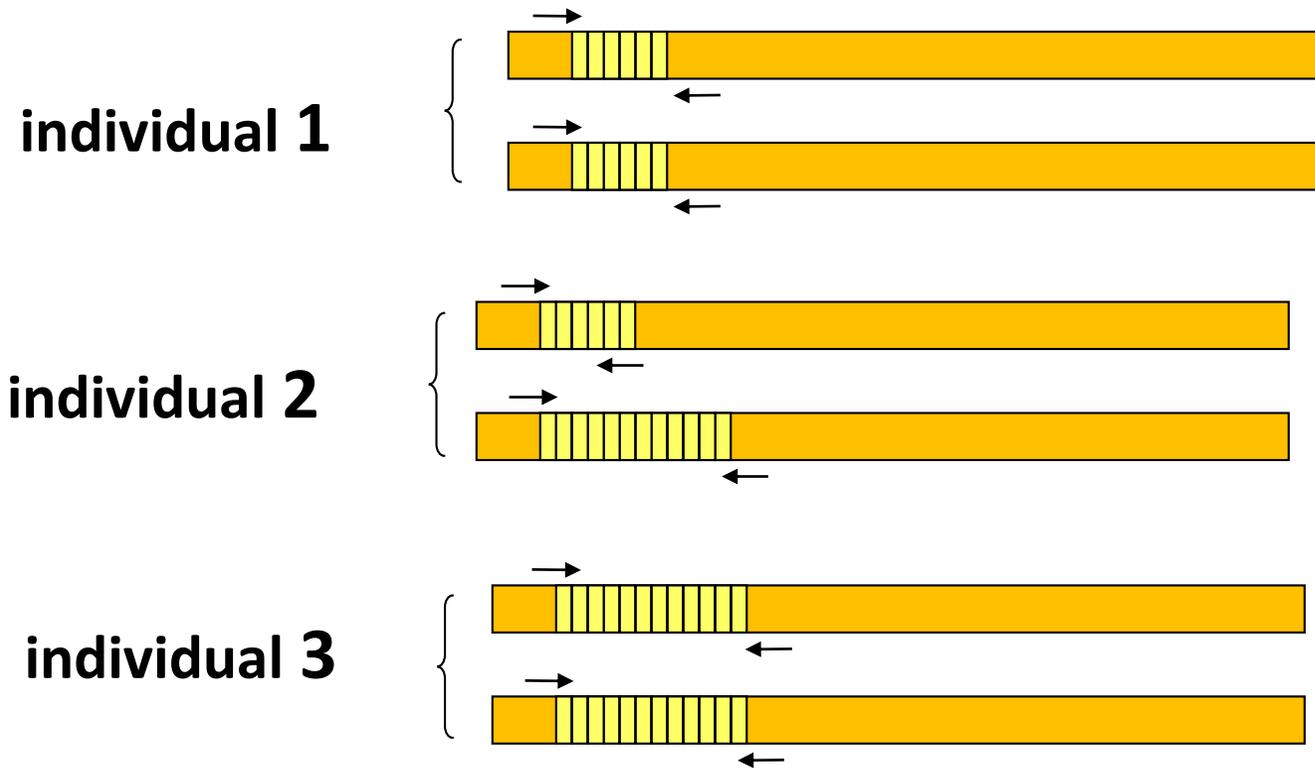
CARATTERISTICHE DEI MICROSATELLITI: CO-DOMINANTI

È POSSIBILE DISTINGUERE L'OMOZIGOTE DALL'ETEROZIGOTE



CARATTERISTICHE DEI MICROSATELLITI: CO-DOMINANTI

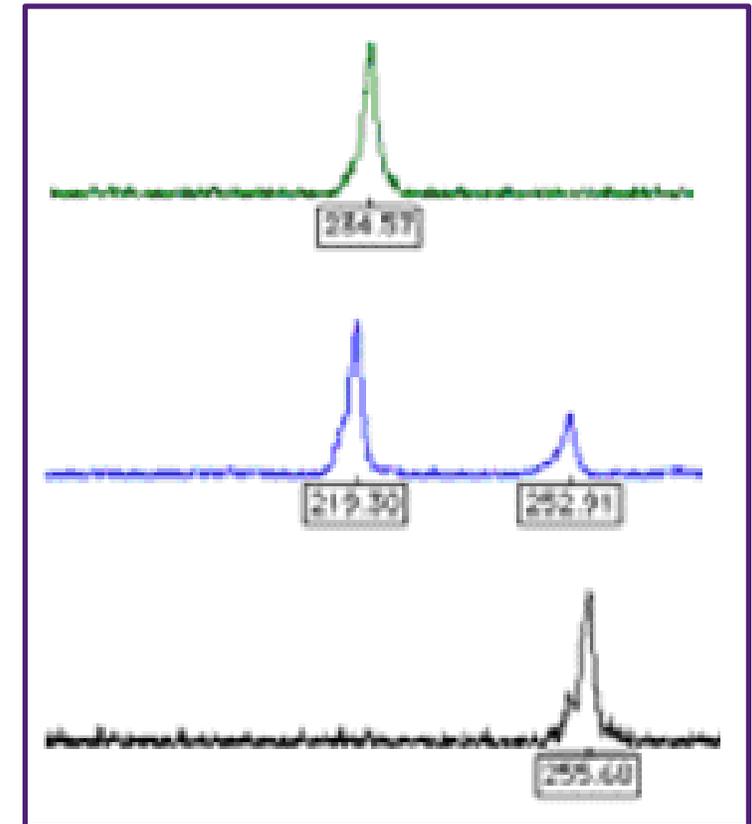
È POSSIBILE DISTINGUERE L'OMOZIGOTE DALL'ETEROZIGOTE



Chr_1

Chr_1

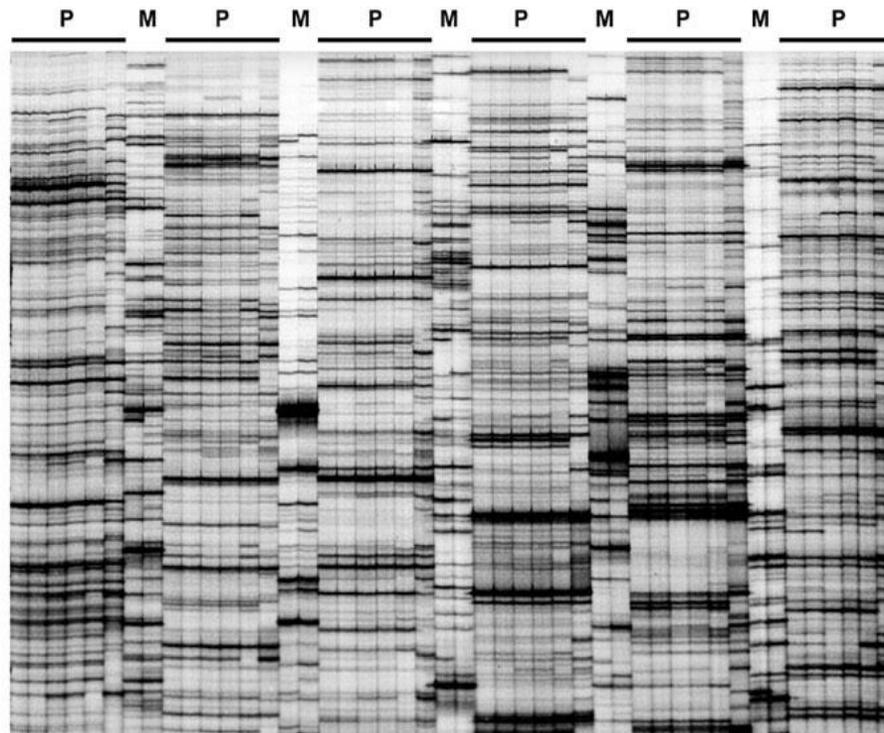
Chr_1



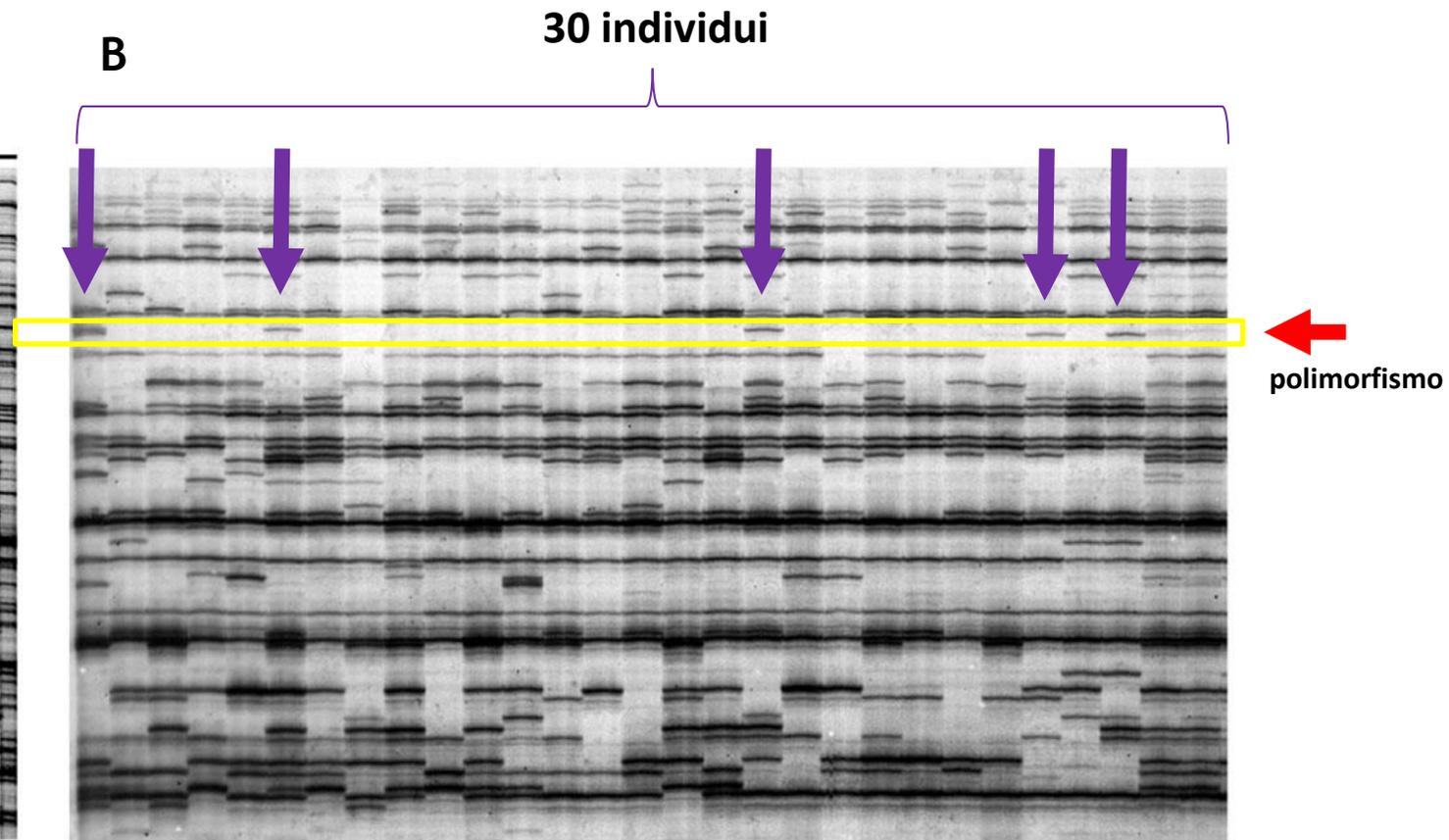
MARCATORI AFLP

(AMPLIFIED FRAGMENT LENGTH POLYMORPHISM)

A



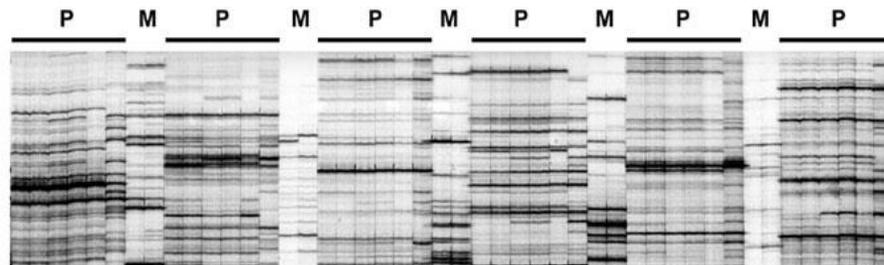
B



MARCATORI AFLP

(AMPLIFIED FRAGMENT LENGTH POLYMORPHISM)

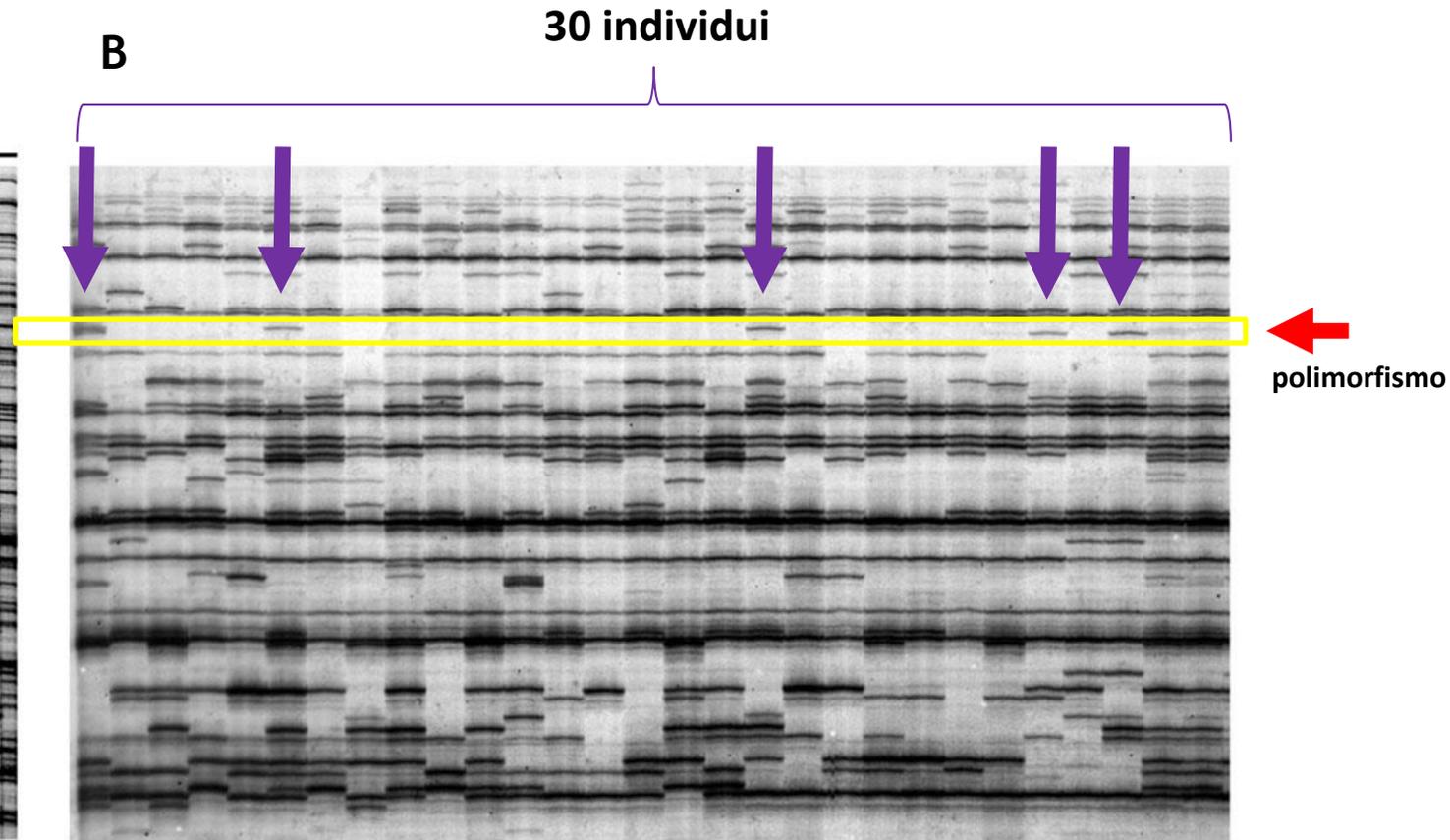
A



AFLP: peculiarità

- MULTI-LOCUS (50-100 per esperimento)
- NO CONOSCENZE apriori del GENOMA
- COSTI RIDOTTI

B



MARCATORI SNP (SINGLE NUCLEOTIDE POLYMORPHISM)

| | | | |
|-----------|---------|---|----|
| seq_1 (A) | ATGCGGC | G | AT |
| seq_2 (A) | ATGCGGC | G | AT |
| seq_3 (A) | ATGCGGC | G | AT |
| seq_1 (B) | ATGCGGC | A | AT |
| seq_2 (B) | ATGCGGC | A | AT |
| seq_3 (B) | ATGCGGC | A | AT |



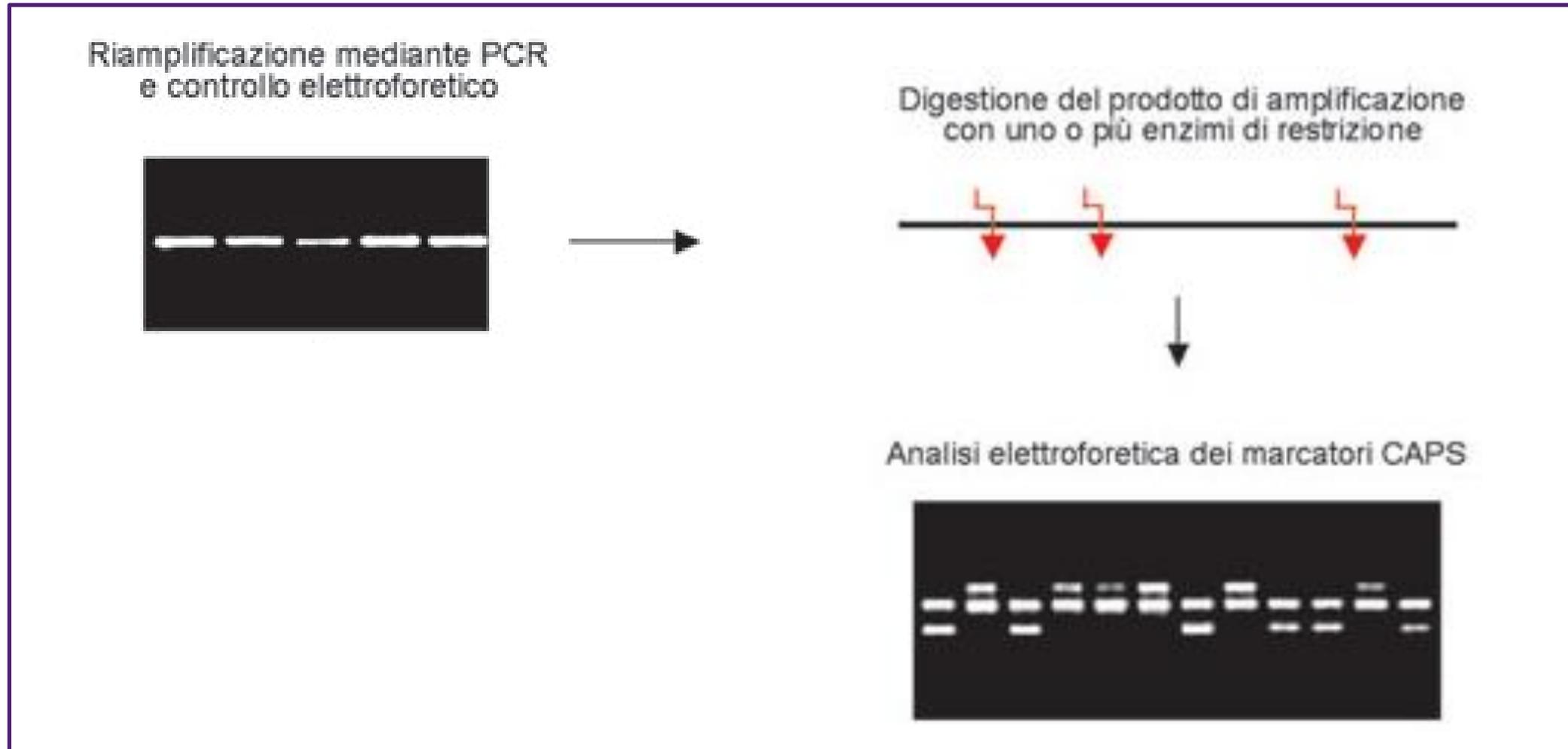
Definizione

Un polimorfismo a singolo nucleotide è una variazione, del materiale genico a carico di un unico nucleotide tale per cui l'allele polimorfico risulta presente nella popolazione in una proporzione superiore all'1%. Al di sotto di tale soglia si è soliti parlare di variante rara (SNV = single nucleotide variant). Vengono rilevati mediante sequenziamento.

Identificabile mediante allineamento multiplo della stessa sequenza derivante da individui diversi

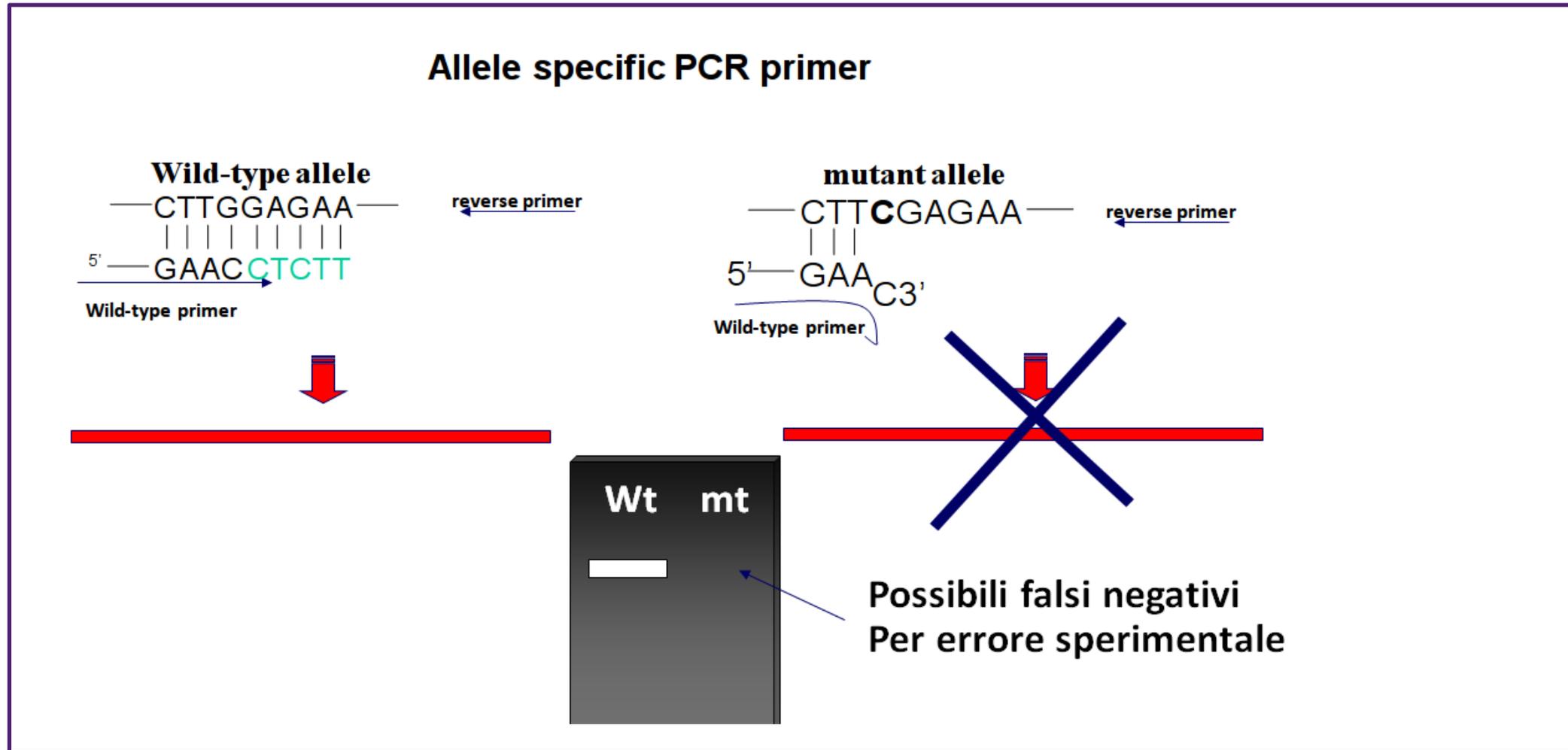
MARCATORI SNP: RILEVAZIONE

CAPS (cleaved amplified polymorphic sequence)



MARCATORI SNP: RILEVAZIONE

ARMS (Allele Refractory Mutation System)



MARCATORI SNP: RILEVAZIONE

Tetra-primer ARMS

Mwg2062

Consensus sequence

```
.....TTGTG TCAAGCATAT CGGTTGCTCT TGCTATCGTT
TTGGTGCACA A TCTATGGGG G TAT GCATAC AGCAA CGACA AGGAAAGT GGT GGAGT ACATT
TCAAGAATTA TGCCAATTA TGGCGTGGC R TTC TTGT TTG ATGACATGCA GTGTGTTCTT
TCAAGGTATTA TA AGGAGGAT CCTA TTGTTT TCGAACGTGC TG GAAATA TG CGACTTTA TT
AATA AGCATA TTTTTCAGG TA TTGTTA GG GGCTGCGGCT TTCAAAAGAT TGGCTCCTA T
GTCAAT CTTA GTGCGTACTA CCT.....
```

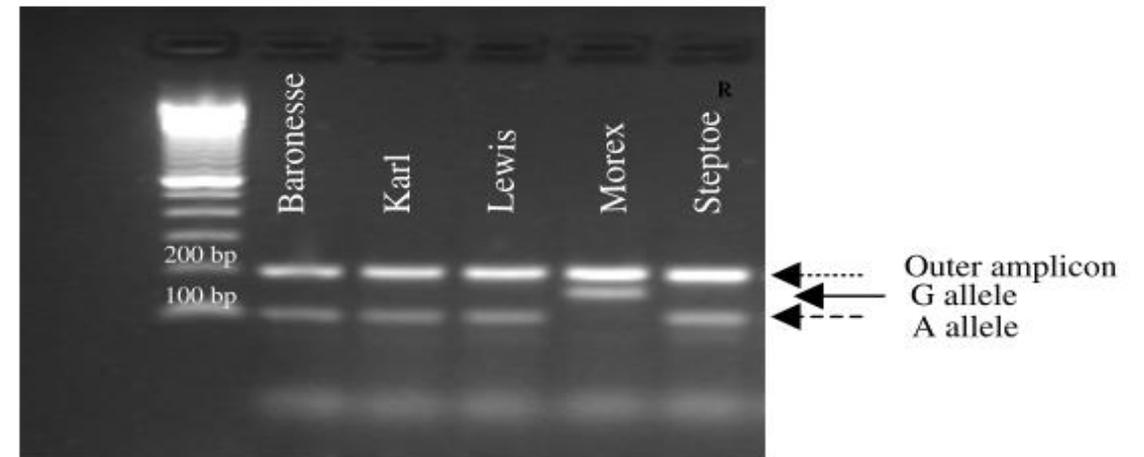
Morex: allele G

Baronesse, Karl, Lewis: allele A

@325gccaattattggcgtggc(**g/a**)ttcttgttgat....

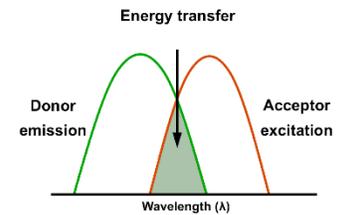
PCR and gel electrophoresis

B



MARCATORI SNP: RILEVAZIONE

SAGGI SNP MULTIPLEXED



Più loci
in contemporanea
(fluorescenza)

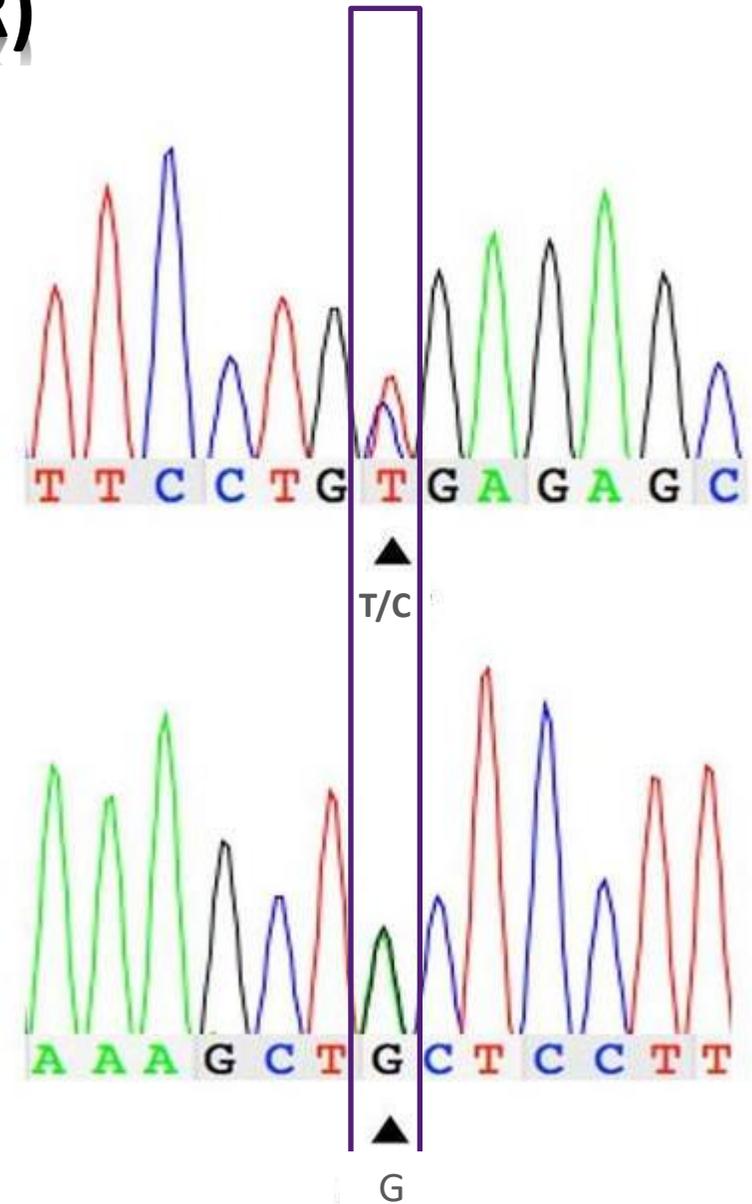
- Allele Specific Hybridisation
- Taqman
- Primer Extension
- Oligonucleotide Ligation assay (OLA)



MARCATORI SNP: RILEVAZIONE (SANGER)

Rilevazione dei marcatori SNPs

Sono visualizzati attraverso sequenziamento di prodotti di PCR di campioni di DNA isolati da un certo numero di individui geneticamente differenziati

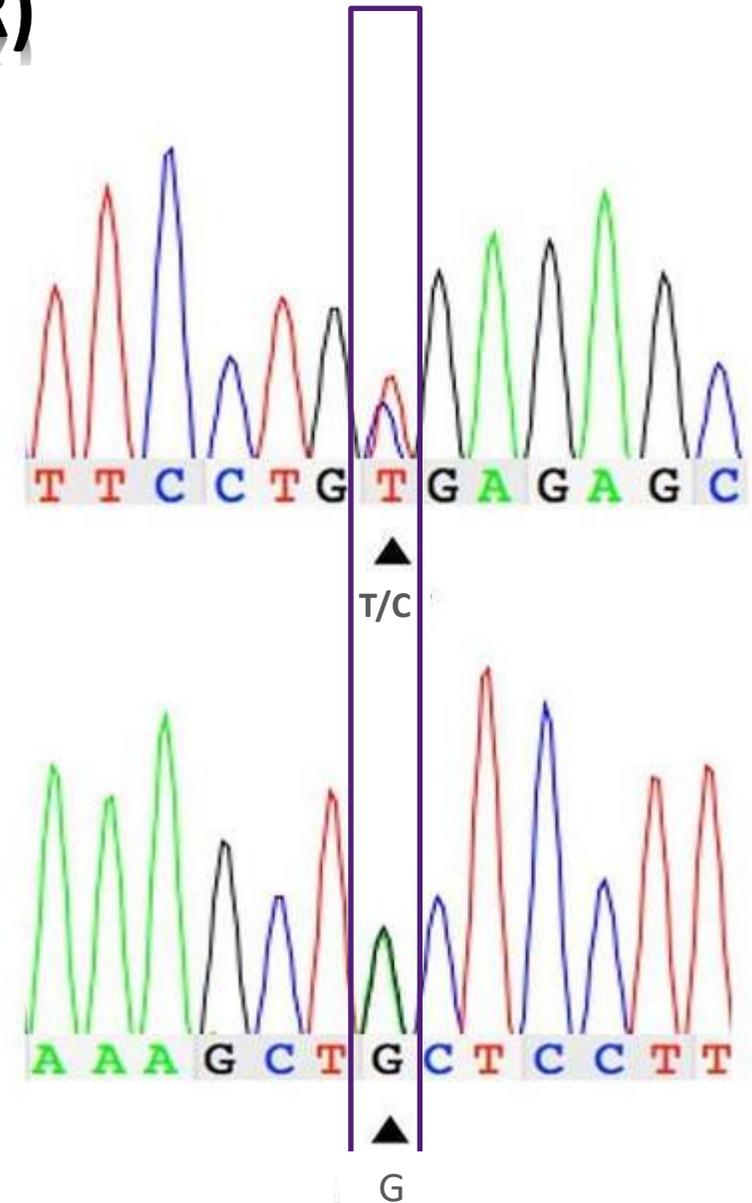


MARCATORI SNP: RILEVAZIONE (SANGER)

Rilevazione dei marcatori SNPs

Sono visualizzati attraverso sequenziamento di prodotti di PCR di campioni di DNA isolati da un certo numero di individui geneticamente differenziati

5 euro per frammento di 500 bp



MARCATORI SNP: RILEVAZIONE (ILLUMINA)

Oggi è possibile sequenziare direttamente i campioni da analizzare (anche se essi sono in numero di > 100) e a **prezzi contenuti**



15 euro per un miliardo di basi



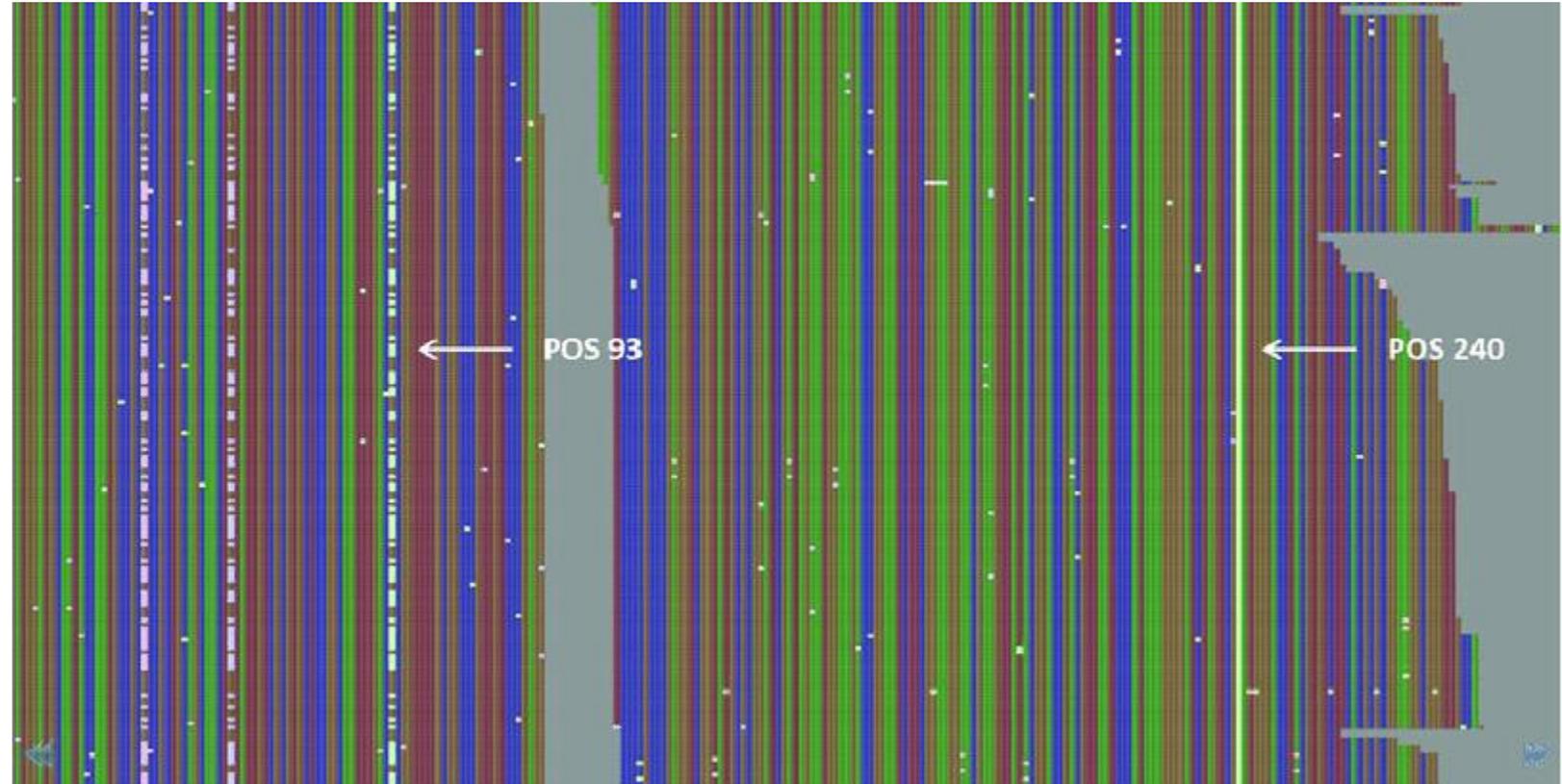
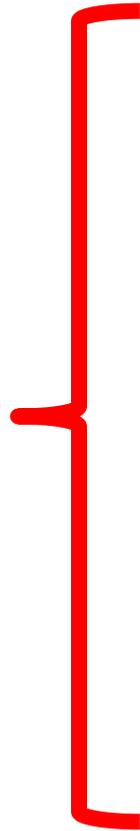
**450 euro per sequenziare
un individuo
(con genoma di 1Gbp, 30X)**

SNP: SEQUENZIAMENTO «WHOLE GENOME»

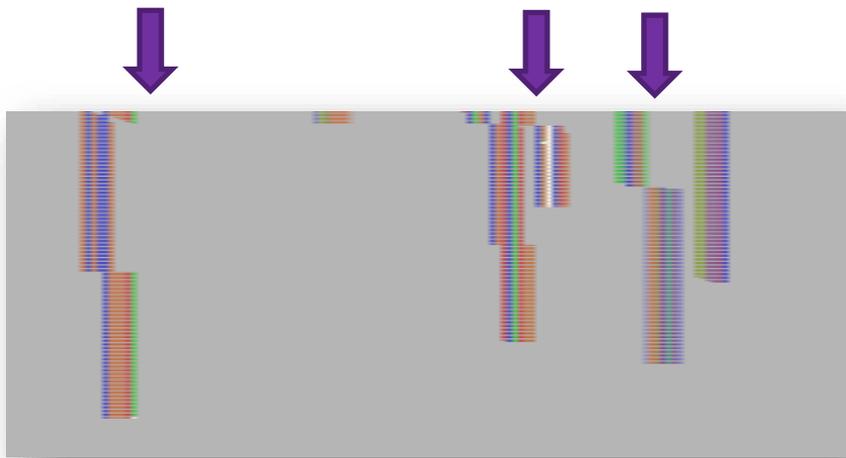
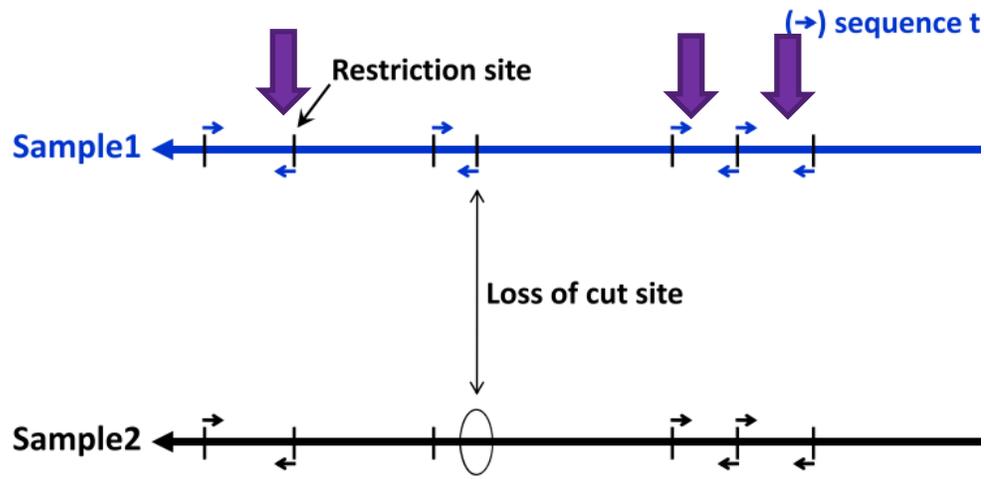
Serie di 3 SNP consecutive



Pila di 50 reads



METODOLOGIE “REDUCED COMPLEXITY”

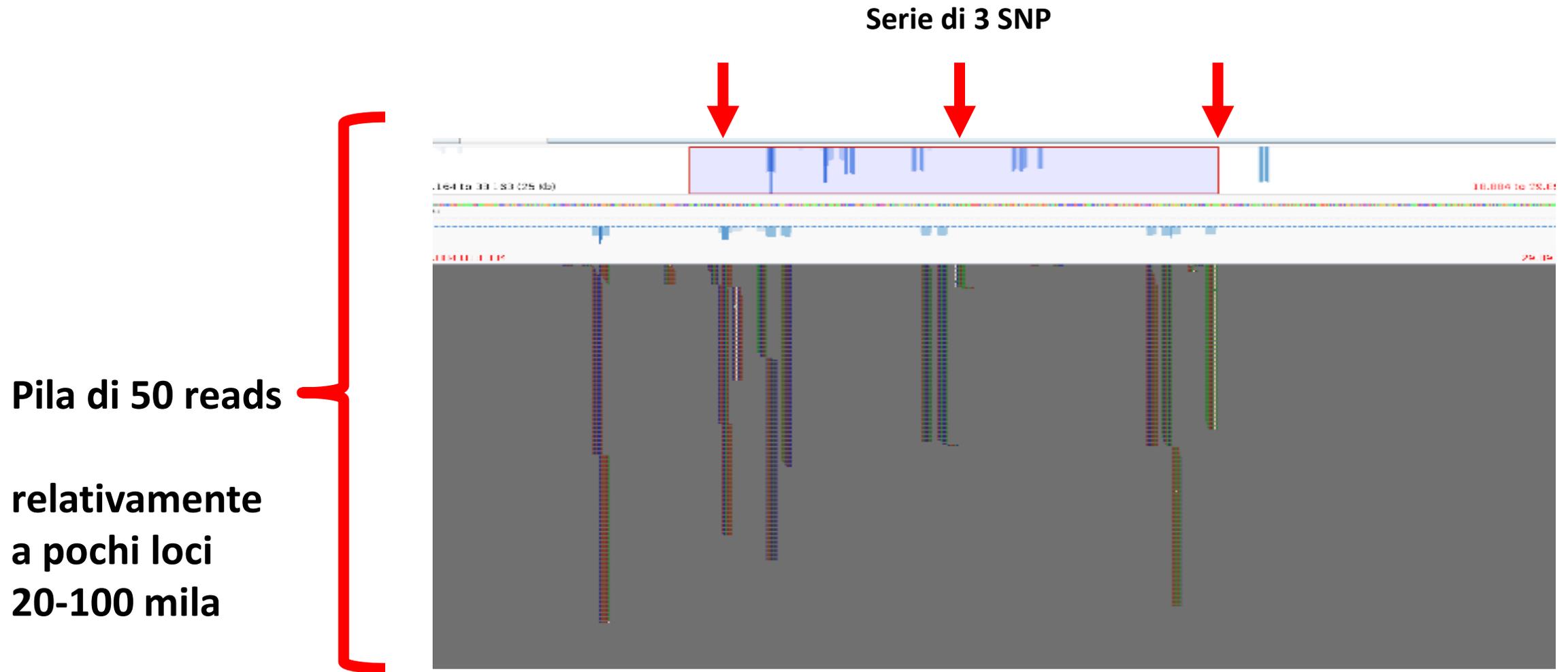


Tramite metodologie “reduced complexity” possono accadere due cose:

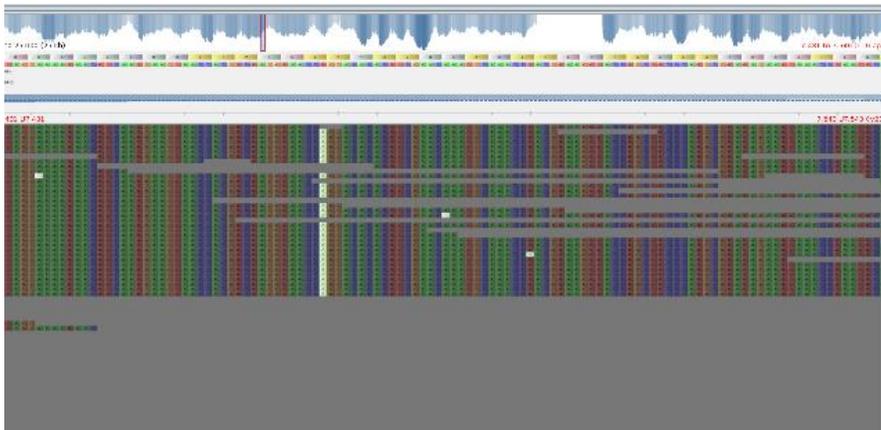
1) Si può perdere un sito di restrizione (e pertanto il frammento non viene sequenziato). E' è un polimorfismo che non viene visto direttamente ma che si può contare

2) Si osservano SNP nei frammenti sequenziati

SNP: SEQUENZIAMENTO A RIDOTTA COMPLESSITÀ



«WHOLE GENOME» VS RIDOTTA COMPLESSITÀ

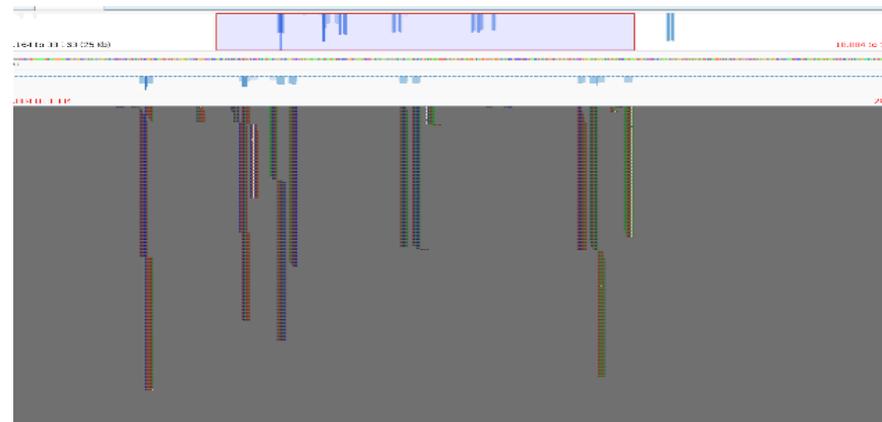


whole genome seq



15 euro / 1 Gbp

450 euro per sequenziare
un individuo
(con genoma di 1Gbp, 30X)



Reduced complexity seq



15 euro / 1 Gbp

30 euro per sequenziare
un individuo
(con genoma di 1Gbp, 30X)

METODOLOGIE “REDUCED COMPLEXITY”

Le più importanti “campionano” la situazione nell’intorno di un tipo di siti di restrizione

Le tecniche oggi disponibili per questo sono:

- Radseq (a uno o a due enzimi)
- CROPS (Complexity reduction of polymorphic sequences)
- GBS (genotyping by sequencing)

OPEN ACCESS Freely available online



A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species

Robert J. Elshire¹, Jeffrey C. Glaubitz¹, Qi Sun², Jesse A. Poland³, Ken Kawamoto¹, Edward S. Buckler^{1,4}, Sharon E. Mitchell^{1*}

METODOLOGIE “REDUCED COMPLEXITY”

Altre “campionano” la situazione nell’intorno di regioni selezionate dall’utente
(sequenze esoni/introniche)

Le tecniche oggi disponibili per questo sono:

- SPET (Single Primer Enrichment Technology)
- ARRAY (infinium, Illumina)

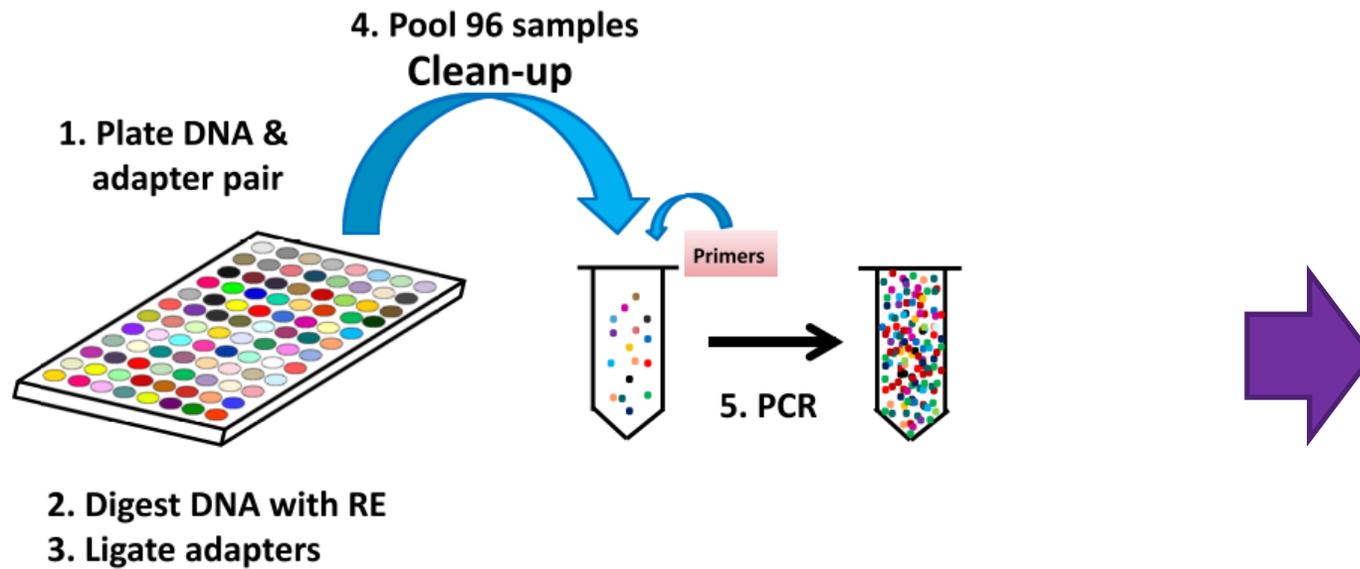


Single Primer Enrichment Technology (SPET) for High-Throughput Genotyping in Tomato and Eggplant Germplasm

Lorenzo Barchi¹, Alberto Acquadro¹, David Alonso², Giuseppe Aprea³, Laura Bassolino⁴, Olivia Demurtas³, Paola Ferrante³, Pietro Gramazio², Paola Mini³, Ezio Portis¹, Davide Scaglione⁵, Laura Toppino⁴, Santiago Vilanova², María José Díez², Giuseppe Leonardo Rotino⁴, Sergio Lanteri^{1}, Jaime Prohens^{2*} and Giovanni Giuliano³*

OPEN ACCESS

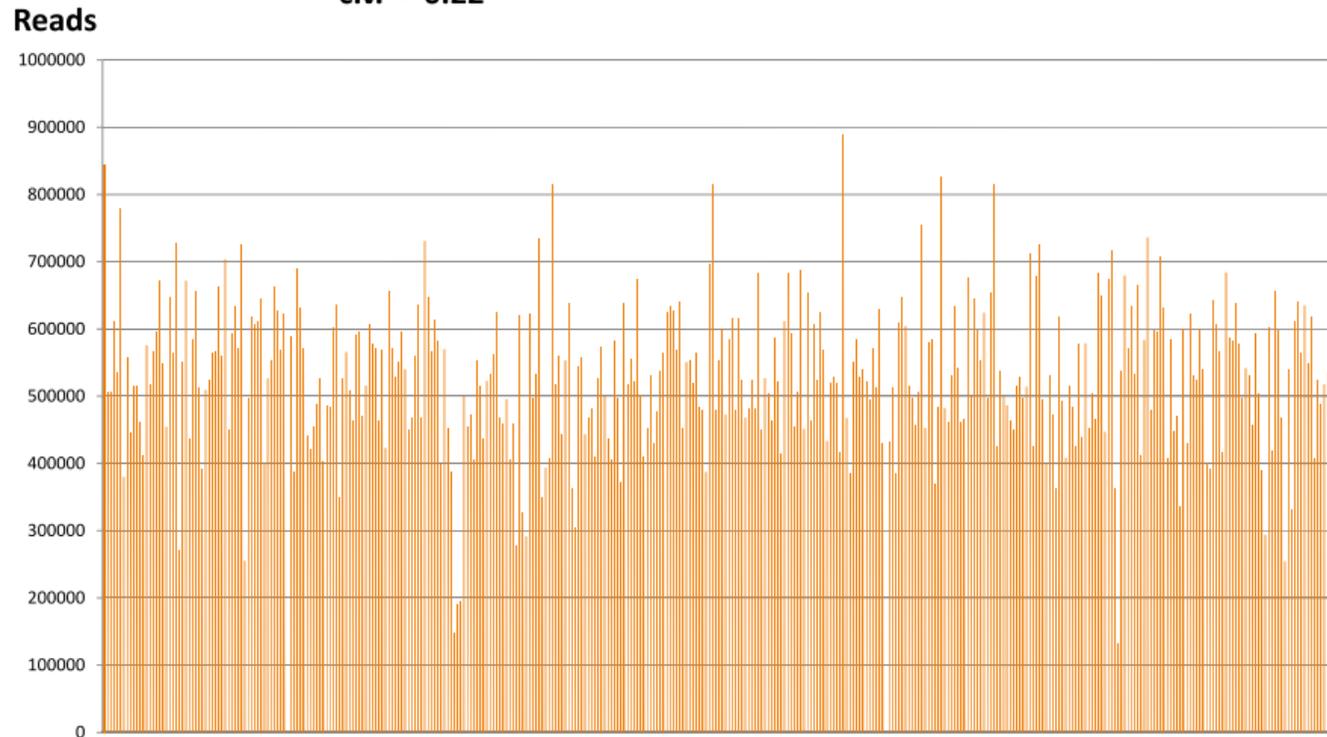
METODOLOGIE “REDUCED COMPLEXITY”



METODOLOGIE “REDUCED COMPLEXITY”

384-plex GBS Results for Maize

Mean read count per line = 528,000
c.v. = 0.22



**Fino a 384 piante
sequenziate insieme**

EDV E SEQUENZIAMENTO

Da un punto di vista legislativo una nuova varietà deve essere

DUS

Distinta – Uniforme - Stabile



Quali soglie per definire una varietà **DISTINTA**
in caso di sequenziamento Illumina?

EDV E SEQUENZIAMENTO

Da un punto di vista legislativo una nuova varietà deve essere

DUS

Distinta – Uniforme - Stabile



Quali soglie per definire una varietà **DISTINTA**
in caso di sequenziamento Illumina?



DIPENDE

EDV E SEQUENZIAMENTO

Le varietà hanno caratteristiche diverse in relazione al sistema riproduttivo

autogamia



linea pura

allogamia



popolazioni di piante selezionate che si inter-incrociano

prop.vegetativa



clone



Un ulteriore categoria di varietà è rappresentata dagli ibridi F1 (linee inbred – eterosi)

EDV E SEQUENZIAMENTO

| GAMIA | Struttura varietà | SOGLIA EDV | 1gb genome |
|------------------|---|------------|-------------------|
| Allogama | popolazioni di piante selezionate che si inter-incrociano | ? | < 5 su 5.000 LOCI |
| Autogama | Linea pura | ? | ? |
| Prop. Vegetativ. | Clone | ? | ? |
| Ibrido F1 | Piante uniformi | ? | ? |

- Occorre portare avanti un lavoro preliminare (es. g2pSOL)
- SPECIE per SPECIE
- Occorre sequenziare più individui della stessa varietà per verificarne la variabilità
- Non ancora in vigore in UPOV



GRAZIE PER L'ATTENZIONE

