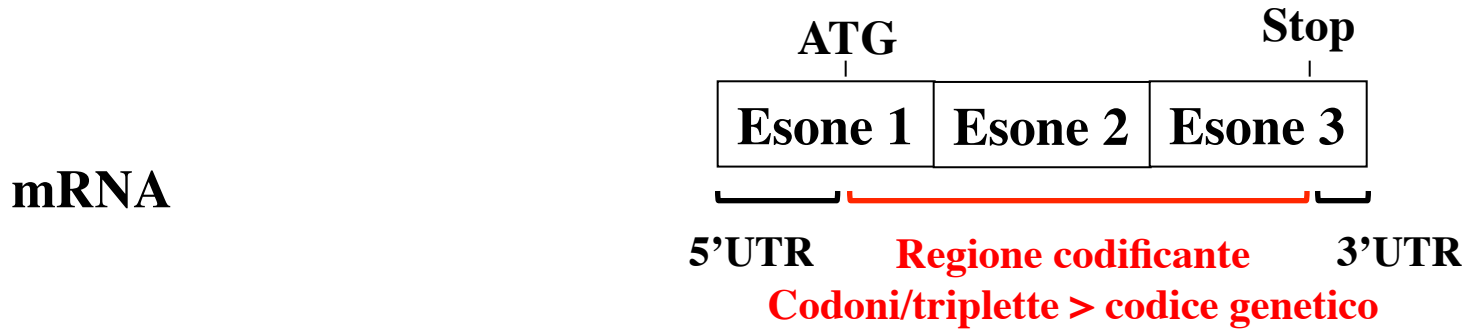
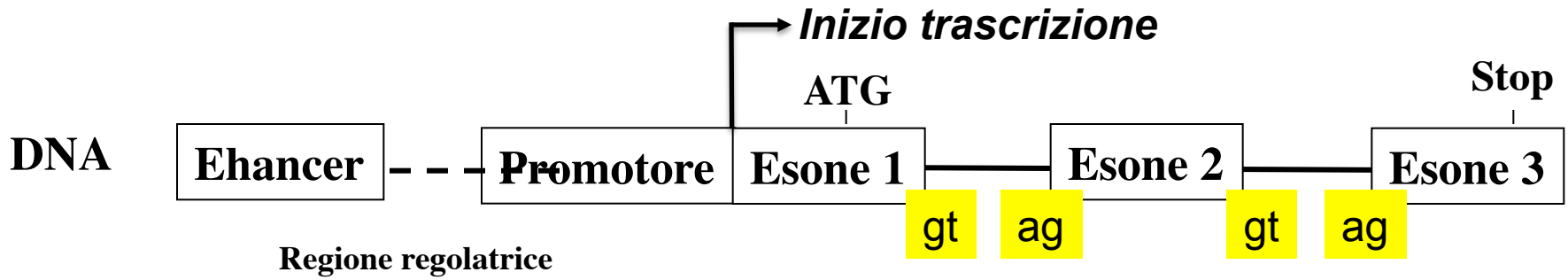


# **STRUTTURA GENOMICA DEL GENE**

**MUTAZIONI GENICHE**

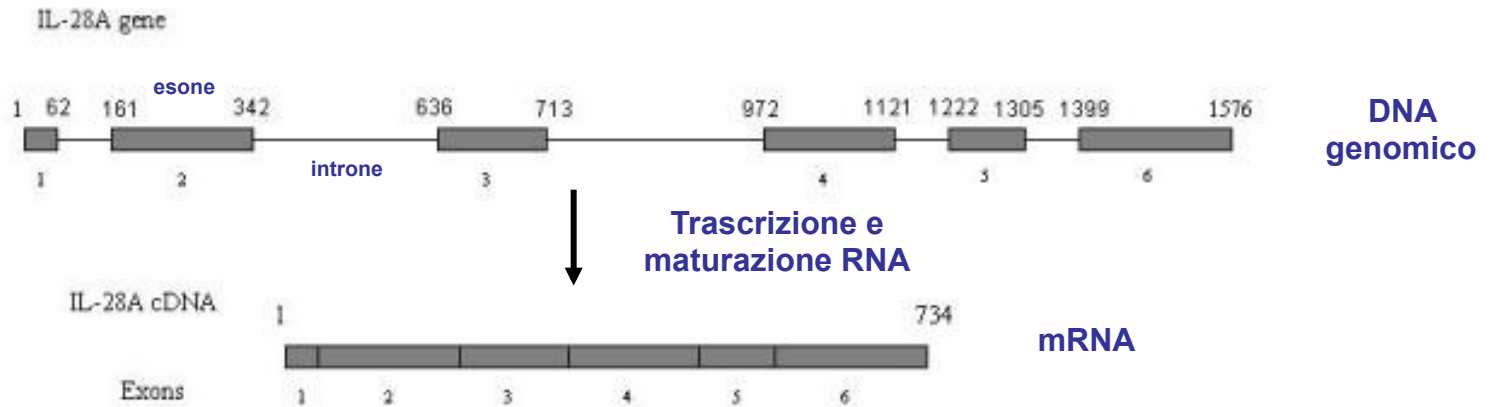


## Seconda base

		Seconda base					
		U	C	A	G		
Prima base	U	UUU	UCU	UAU	UGU	Terza base	U
		UUC	UCC	UAC	UGC		C
		UUA	UCA	UAA	UGA		A
		UUG	UCG	UAG	UGG		G
	C	CUU	CCU	CAU	CGU	U	
		CUC	CCC	CAC	CGC	C	
		CUA	CCA	CAA	CGA	A	
		CUG	CCG	CAG	CGG	G	
	A	AUU	ACU	AAU	AGU	U	
		AUC	ACC	AAC	AGC	C	
		AUA	ACA	AAA	AGA	A	
		AUG	ACG	AAG	AGG	G	
	G	GUU	GCU	GAU	GGU	U	
		GUC	GCC	GAC	GGC	C	
		GUA	GCA	GAA	GGA	A	
		GUG	GCG	GAG	GGG	G	

# Struttura del gene

A)



B)

```

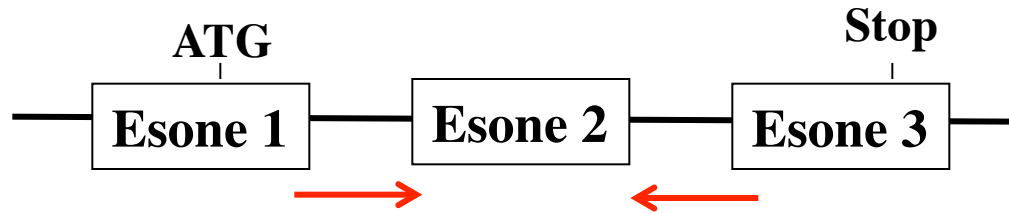
1  tgggtgacag cctcagagtg tttcttctgc tgacaaagac cagagatcag gataaaact
61  aggtgagtc cacatctctg tccgtgctca gctcctgcag cccctgccct cagtgggcag
121 cctctccatc cctcagatc cctttctctc tgtgacacag acatgactgg ggaactgcag
181 ccagtgctgg tctctgatgg cgcagtgctg accgtgactg gaggcagttc tctctccagg
241 ctccacgggg ctctcccgga tgcagggggc tgcacatag cccagttcaa gtccctgtct
301 ccaacaggagc tgcaggcctt taagaggggc aaagatgect tagtgagtct ccccttgcct
361 tcttgcactg gactagcctc caccctccact ccaagcgtca ceatgettct ccactcccag
421 ctctctccac tgggtagcc tccacctctc ctgcagtggt ctatctcctg ctctactctg
481 agggactgac tcatgttttc ctgtagaaga gggctctcta ccatctctcc agcagttaac
541 ctccctctatc ctgtttgctag ccctctctca atcccaccag gatggtctaa cctccacccc
601 tcttctctgg gctaacctgt gcccttctct tctaggaaga gtctctctct ctgaaggact
661 gcaggtgcca ctccctctct ttcctccagg cctgggactt gaggcagctg caggtgagag
721 ggggagtcag gccaccctct gctctctccag cccactctac ctggctctgt agtggccctt
781 tcaactctct tttctctctt gctctctctc ctctctctca cactctctct cctctctctc
841 cgtctctctc tgaccacact ggtgtgctcc tctctctctg cctctctctc tctctctctt
901 cctctctctc ctgctctctc acctgtctcc ctctctctcc cctctctctg ctctctctca
961 ctctctctca ggtgagggag cgcctctctg ctttggagge tgagctgctc ctgactctga
1021 aggttctgga gcccactctg gactctgacc cagcctctgt ggaactcttg gaccagctcc
1081 ttcacacctt gcaccatctc ctctctcagt tccgggctct tgtgagctgt tggggctctg
1141 gcaacctctg ctgtgagctc tgagcagctg cctctctctt gccaggctcc cggctctctc
1201 accctctctc tctgctctca gatccagctc cagctctctg cagggctctc gaccctctct
1261 cgcctctctc attggtctga cggctctctc gaggctctca aaaaggtctg tgaccctctg
1321 agagagggac tgaggtctgt ggagcctctg ggagctctca acccagctct cctctctctc
1381 atccctctct cctctctctg gctctctctg tctctctctg cctctctctc ctctctctca
1441 tctctctctc tctctctctg cctctctctg gttctctctg gggctctctg tctctctctc
1501 tctctctctc cctctctctg gatctctctt ttataaatta gccctctctc tctctctctt
1561 gccactctct cgtctctct
    
```

# Mutazioni geniche

(riferite ad alterazioni di pochi nucleotidi)

- Alterazioni che possono colpire un gene ovunque (regioni regolatrici, esoni, introni, ecc.)
- Classificazione:
  - a) Sostituzioni nucleotidiche
  - b) Piccole delezioni, duplicazioni, inserzioni (**indel**)
- Terminologia:
  - a) **Variante**: alterazione con effetto sconosciuto (neutro o patogenetico?)
  - b) **Mutazione**: variante con effetto patogenetico
- Valutazione dell'effetto (meccanismi di patogenicità)

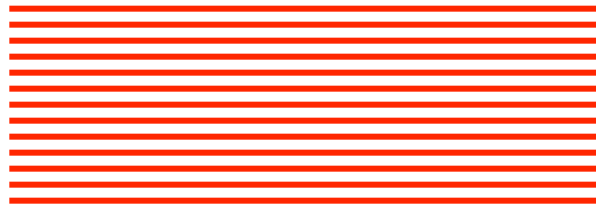
# Screening di mutazioni



## PCR

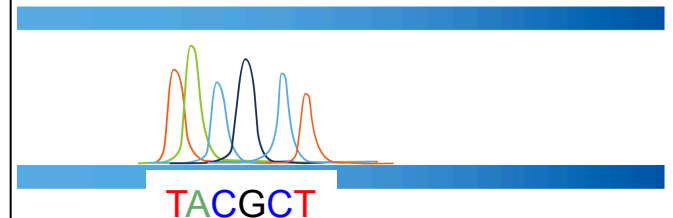
Amplificazione di DNA

(sintesi in vitro di molecole del frammento prescelto)



Sequenziamento  
Sanger

In the final step, the data analysis software interprets the raw data peaks and automatically generates the final sequence of the DNA template, as well as an electropherogram with each colored peak representing each letter of the sequence.



# Sanger sequencing vs Next Generation Sequencing (NGS)

500-1000 bp > entire genome



**Sanger sequencing:** consente di sequenziare un frammento alla volta

**NGS** (anche high-throughput sequencing, sequenziamento ad alta resa): consente di sequenziare moltissimi frammenti in parallelo con riduzione tempi e i costi della ricerca di mutazioni.

## ***NGS applications***

- **Whole Genome Sequencing (WGS):** entire genome (3000 Mb)
- **Whole Exome Sequencing (WES):** all exons (180,000-200,000 exons; 1% of genome; 30 Mb)
- **Transcriptome analysis (RNA-seq):** the quantification of transcript levels and the sequence information
- **Epigenomics:** study of changes in the regulation of gene activity and expression that are not dependent on the DNA sequence
- **Metagenomics:** access of the genetic content of entire communities of organisms



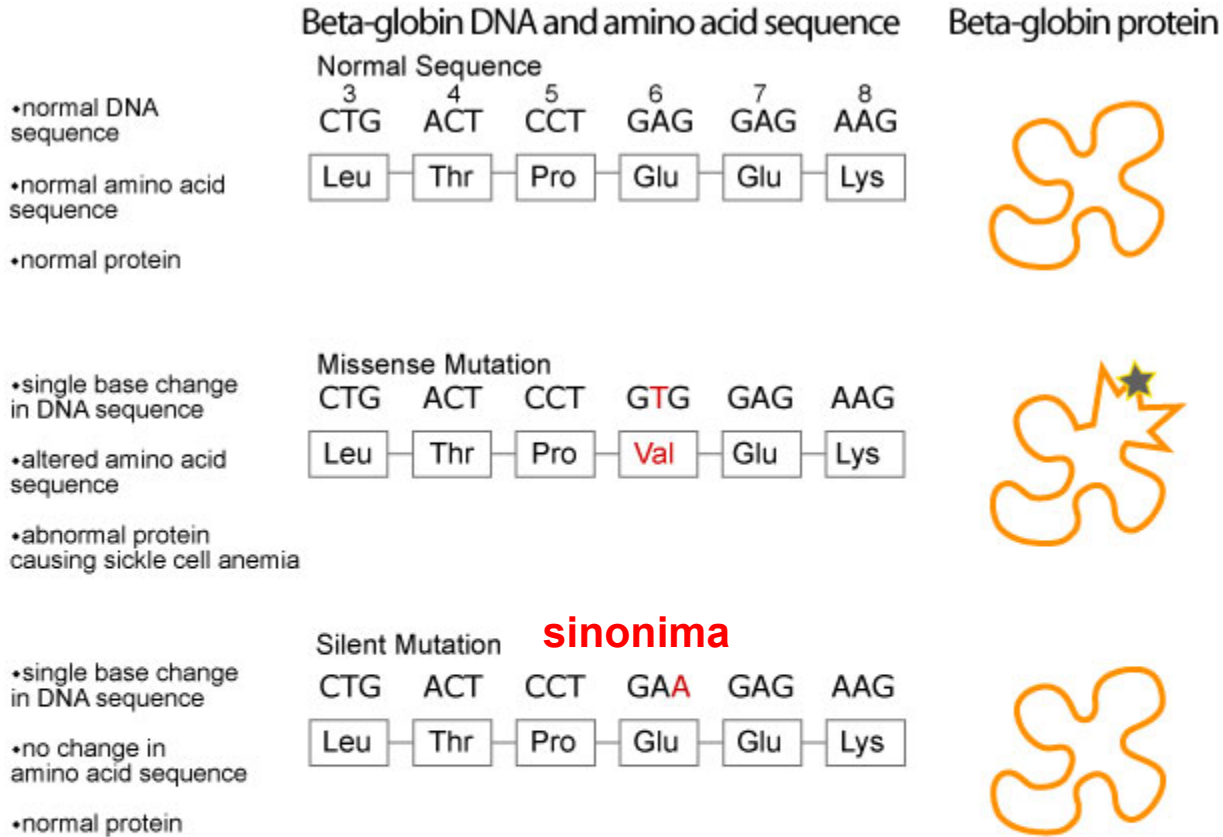
# Sostituzioni nucleotidiche



# Effetto sulla proteina



# Effetto Patogenetico?

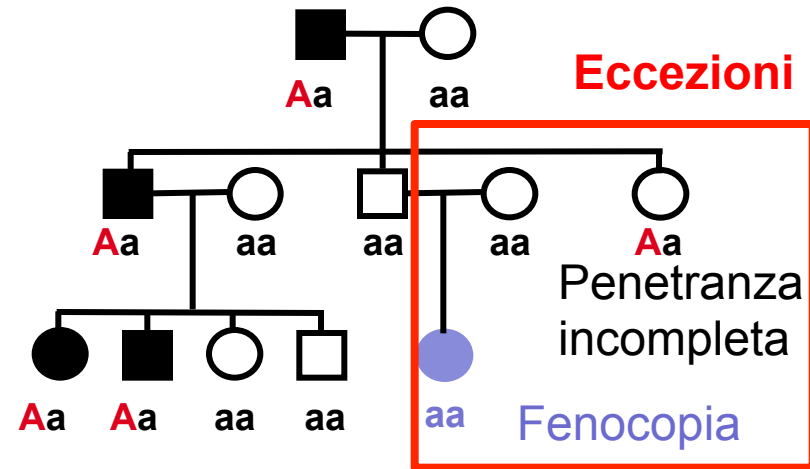


Effetto?

Effetto?

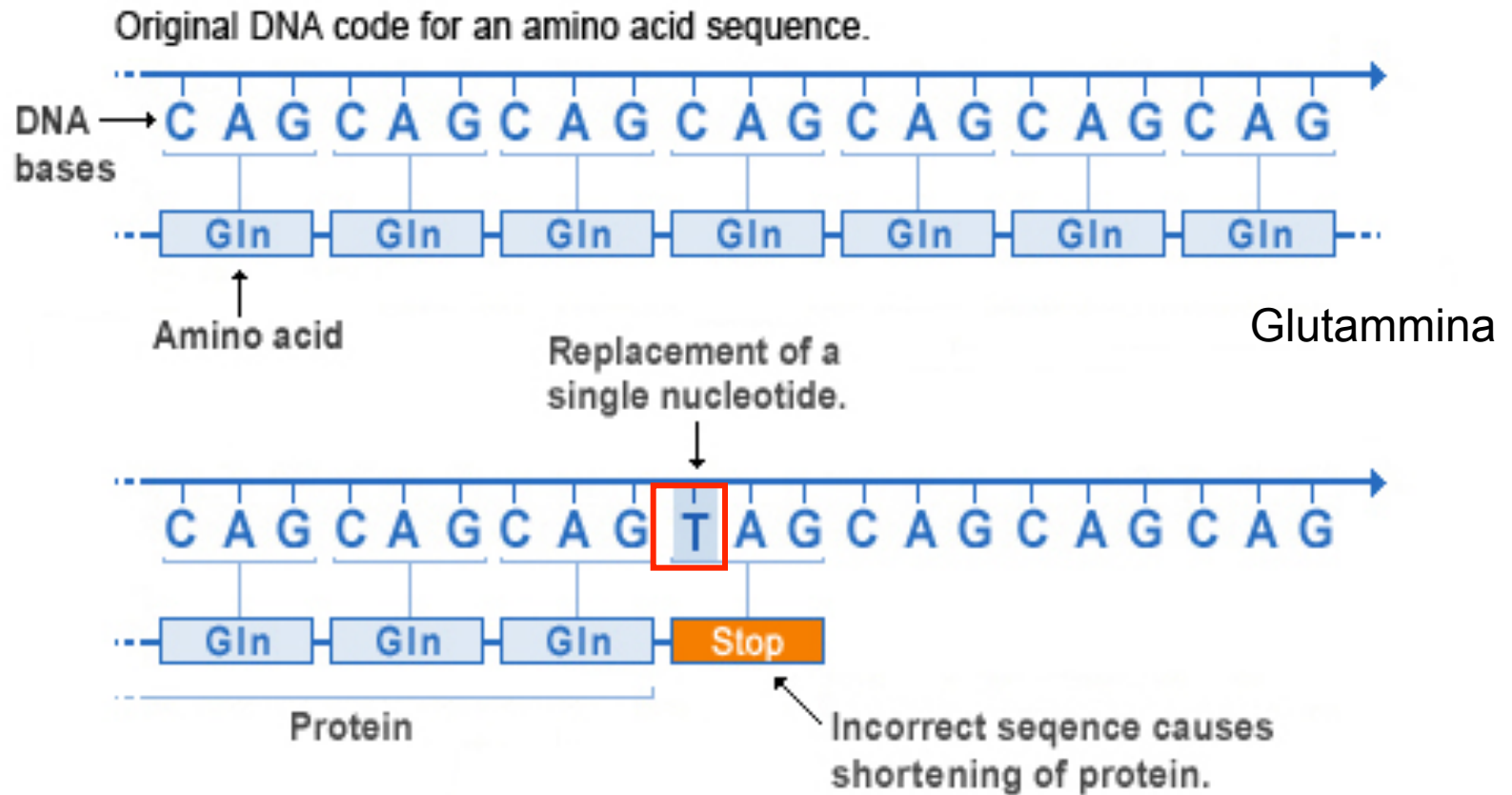
# Effetto patogenetico delle varianti missense? (problematica frequente in NGS)

- Variante rara nella popolazione
- **Segregazione nella famiglia**
- **Conservazione dell'amminoacido durante l'evoluzione**
- Programmi predittivi
- Studi funzionali in vitro o in modelli animali



	G42S	Y48H
Hs	GLFGRKT <b>G</b> QAPGYS <b>Y</b> TAANKNKGI	
Mm	GLFGRKT <b>G</b> QAAGFS <b>Y</b> TDANKNKGI	
Gg	GLFGRKT <b>G</b> QAEGFS <b>Y</b> TDANKNKGI	
Dr	GLFGRKT <b>G</b> QAEGFS <b>Y</b> TDANKSKGI	
Dm	GLIGRKT <b>G</b> QAAGF <b>A</b> YTDANKAKGI	
At	GLFGRQ <b>S</b> GTTAGYS <b>Y</b> SAANKNKAV	
Sc	GIFGRHS <b>G</b> QAEGYS <b>Y</b> TDANIKKNV	
Bm	GLFGRHS <b>G</b> SVEGY <b>K</b> YSDANKNSGI	

# Nonsense mutation



**Sostituzioni  
nucleotidiche**

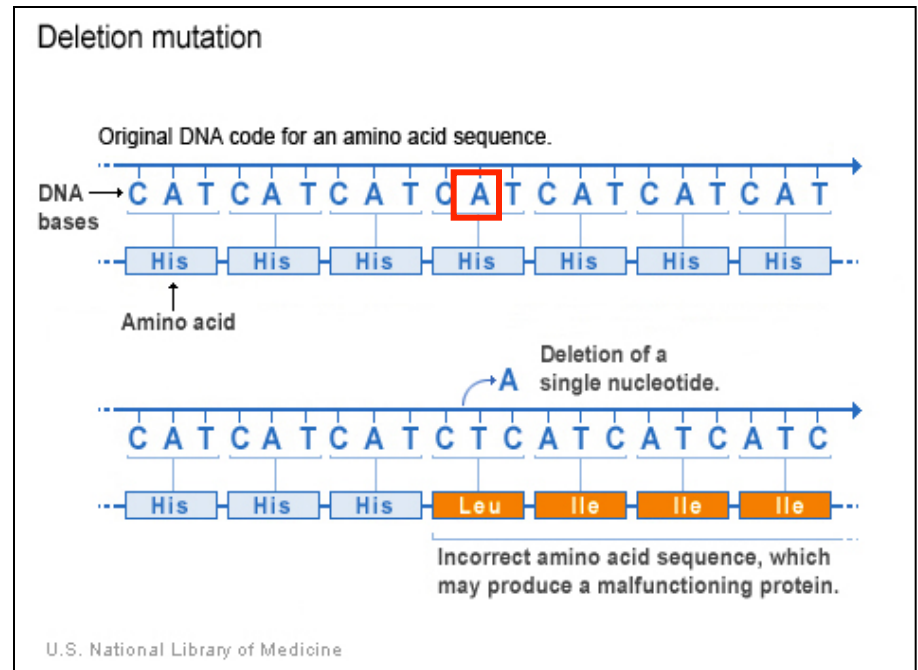
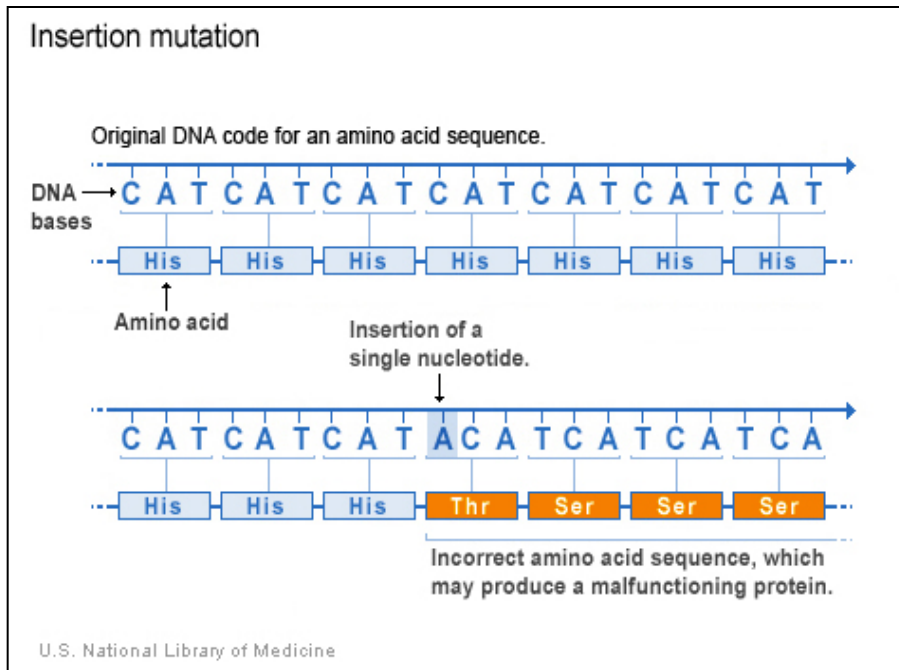


**nonsense**



**Effetto  
Patogenetico  
Generalmente  
deleterio**

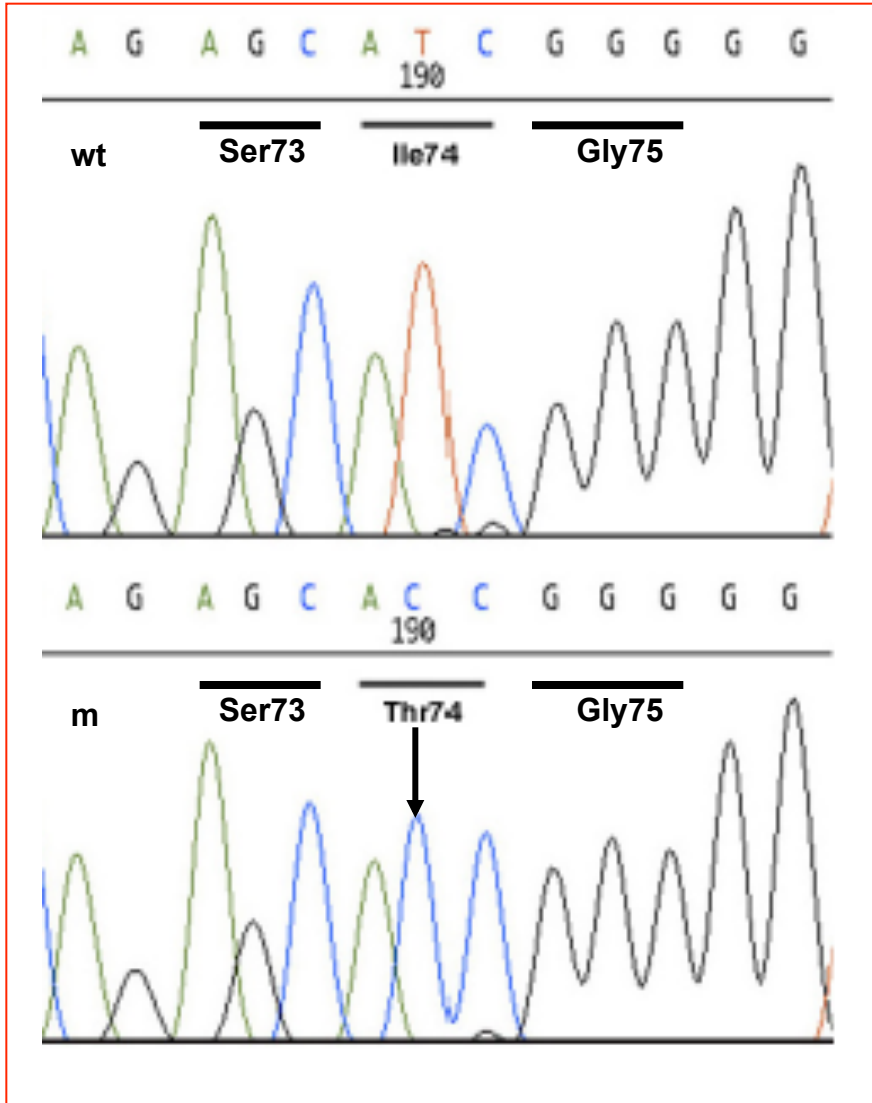
# Frameshift mutations



Indel → Frameshift > Nonsense → **Effetto patogenetico**

Indel (3n) → In-frame → **Effetto patogenetico  
Generalmente deleterio**

## c.221T>C/p.Ile74Thr in omozigosi



Sequenza normale

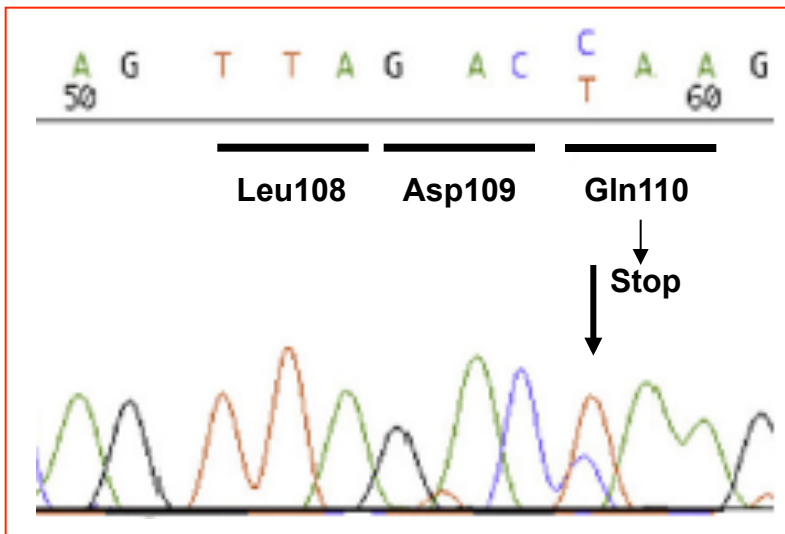
AGAGCA**T**CGGGGGG wt allele 1

AGAGCA**T**CGGGGGG wt allele 2

Sequenza mutata in omozigosi

AGAGCA**C**CGGGGGG m allele 1

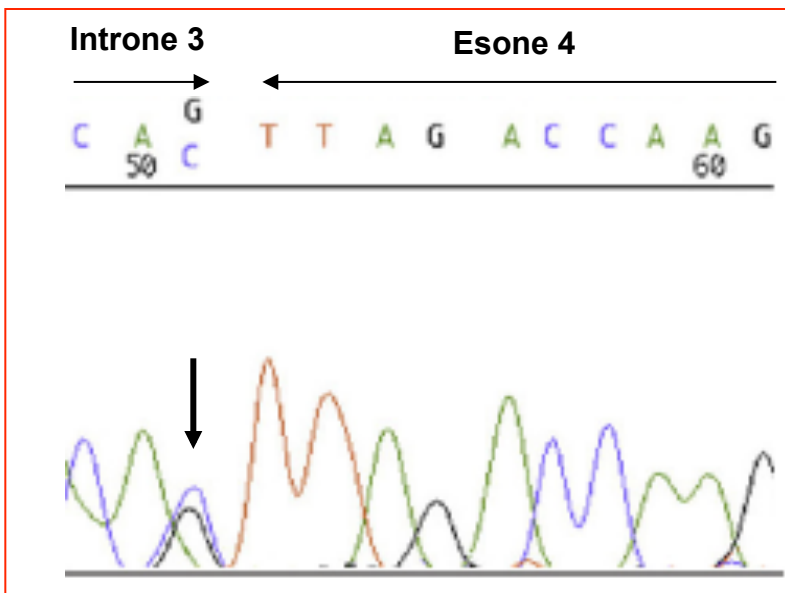
AGAGCA**C**CGGGGGG m allele 2



## c.328C>T/p.Gln110\* in eterozigosi

AGTTAGACCAAG wt allele 1

AGTTAGACTAAG m allele 2



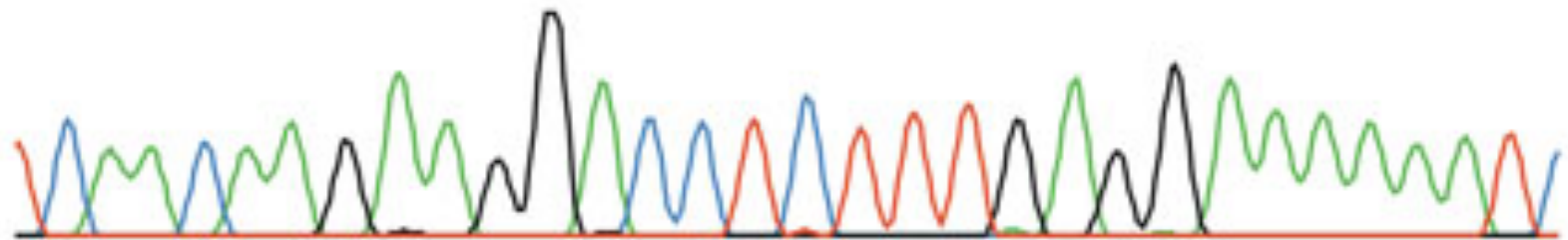
## c.322-1G>C/p.? in eterozigosi

cag TTAGACCAAG wt allele 1

caC TTAGACCAAG m allele 2

**Mutazione di splicing**

# DELEZIONE ETEROZIGOTE di tre nucleotidi - AAG



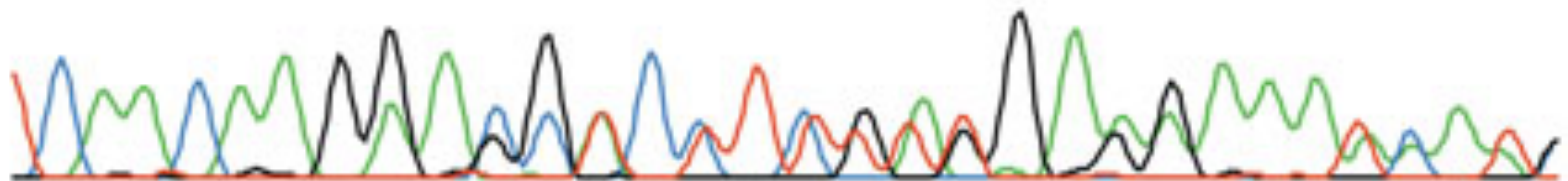
Wildtype alleles

C A A C A A G A A G G A C C T C T T T G A G G A A A A A A T

9070

9080

9090



Wildtype allele

C A A C A A G A A G G A C C T C T T T G A G G A A A A A A T

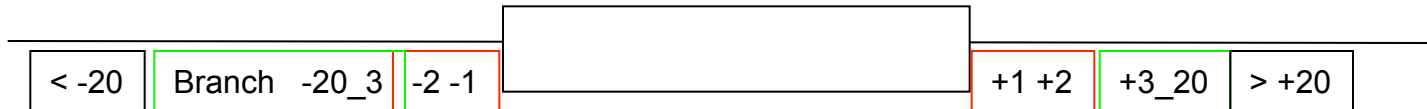
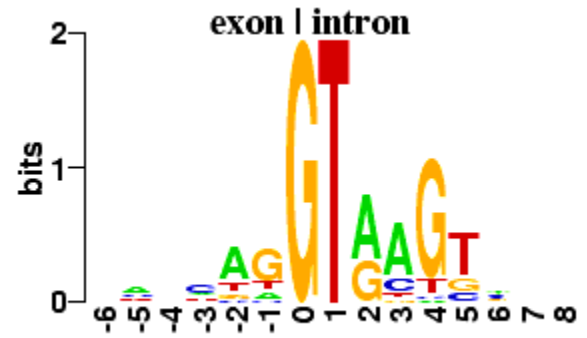
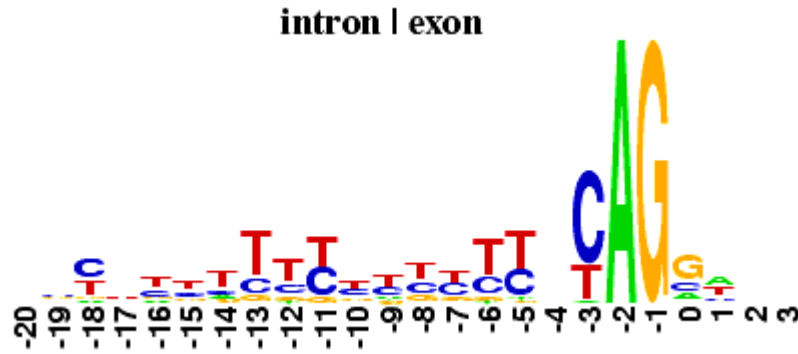
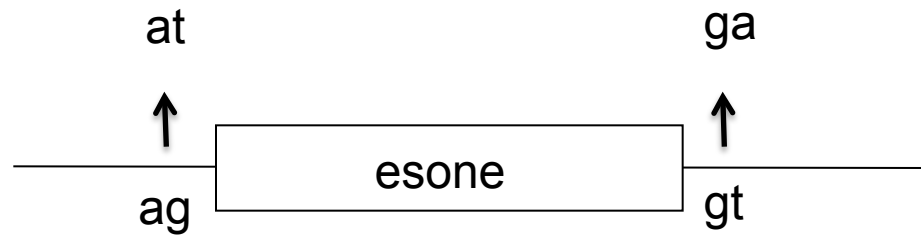
Mutated allele

C A A C A A G G A C C T C T T T G A G G A A A A A A T C A A

9070

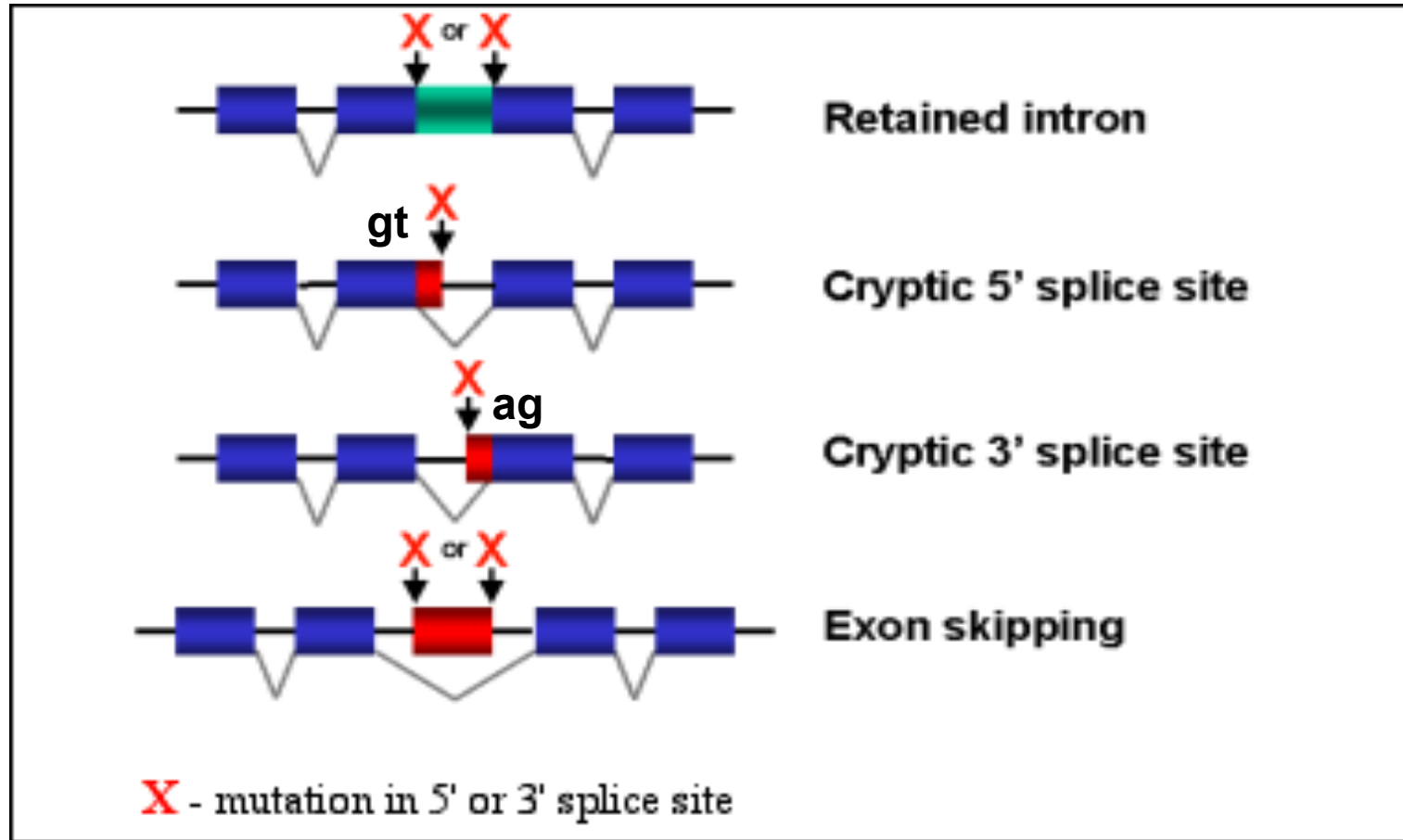
9080

9090





## Conseguenze mutazione nei siti di splicing



Splicing  
mutation



Effetto  
mRNA



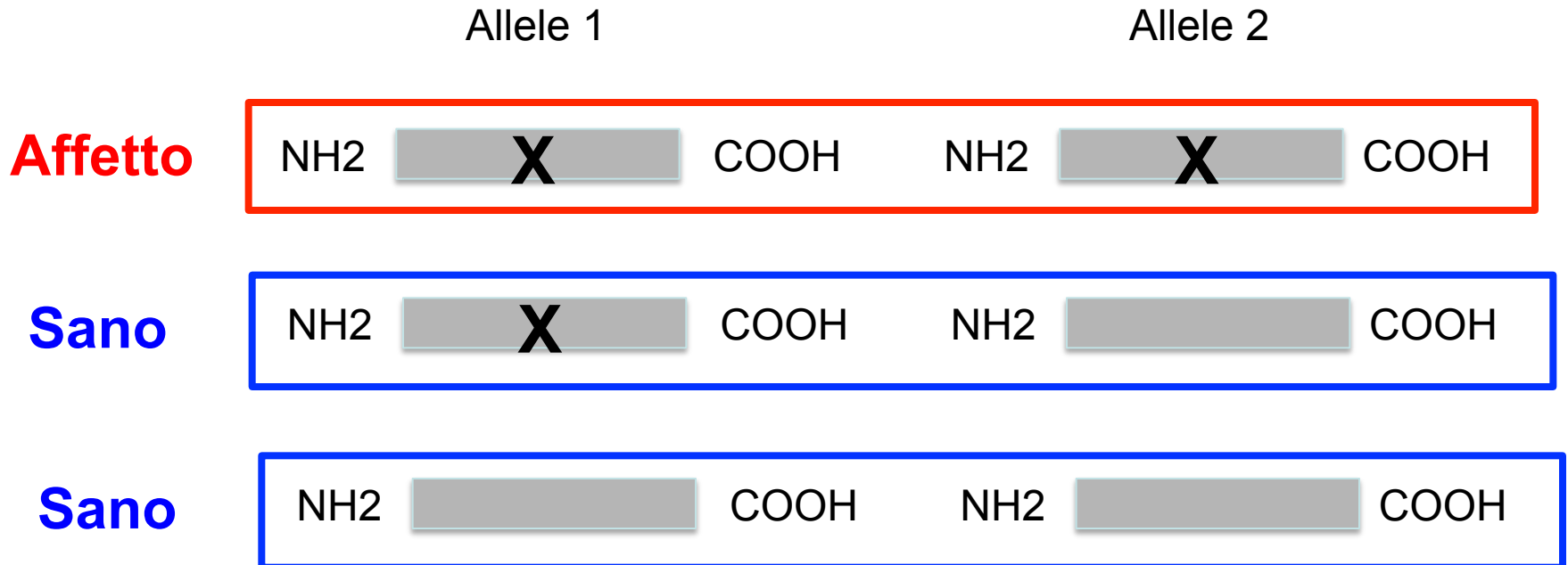
Frameshit  
In-frame

## Classificazione mutazioni

Livello DNA		Livello Proteina
Regione codificante	Sostituzioni nucleotidiche	sinonima
		missense
nonsense		
	Delezioni, inserzioni, duplicazioni	Delezioni/Inserzioni in frame (N. nucleotidi: $3n$ ) Frameshift (N. nucleotidi: $\neq 3n$ )
Regioni non codificanti	Siti di splicing	Effetti diversi sulla maturazione dell'mRNA
	Regioni regolatrici/5' o 3' UTR	Effetti diversi sui livelli di espressione dell'mRNA e della proteina

## Classificazione delle mutazioni in base all'effetto sulla funzione

**Malattie recessive** { **1) Perdita di funzione (loss of function):** riduzione o perdita di funzione; negli eterozigoti si mantiene un margine di attività che permette una normale funzione



**Malattie dominanti**



**2) Aploinsufficienza:** contributo di un allele normale non è sufficiente per prevenire un difetto; necessario più del 50% di proteina per la normale funzione

Allele 1

Allele 2

**Affetto**



**Affetto grave**

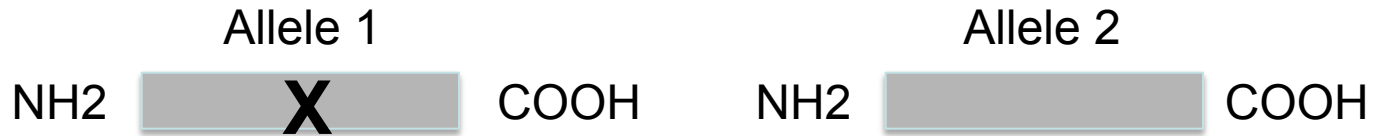


**Sano**

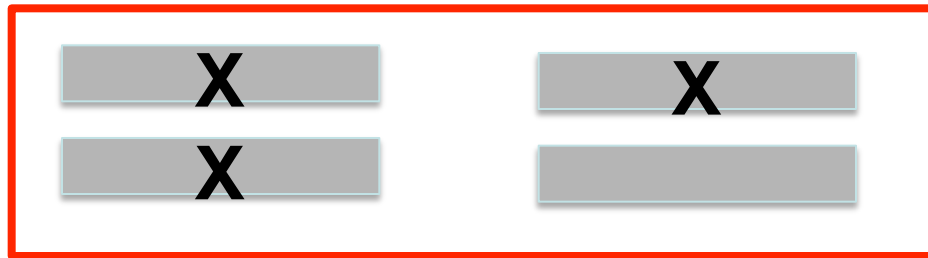


# Classificazione delle mutazioni in base all'effetto sulla funzione

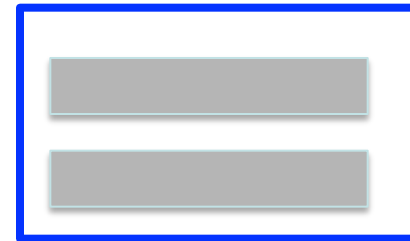
**3) Effetto dominante negativo:** proteina anomala interferisce con la funzione dell'allele normale



Molecola matura formata da dimeri nelle seguenti combinazioni



**Prodotto non attivo**



**Prodotto attivo  
Insufficiente per garantire funzione**

1 : 2 : 1

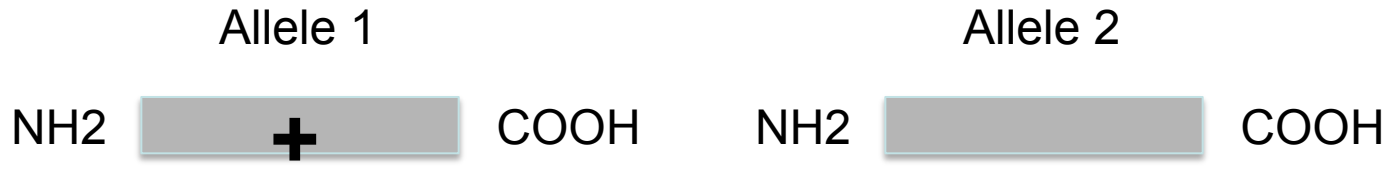
Malattie dominanti

# Classificazione delle mutazioni in base all'effetto sulla funzione

Malattie dominanti

## 4) Gain of function

- Aumenta l'attività funzionale
- Nuova funzione della proteina



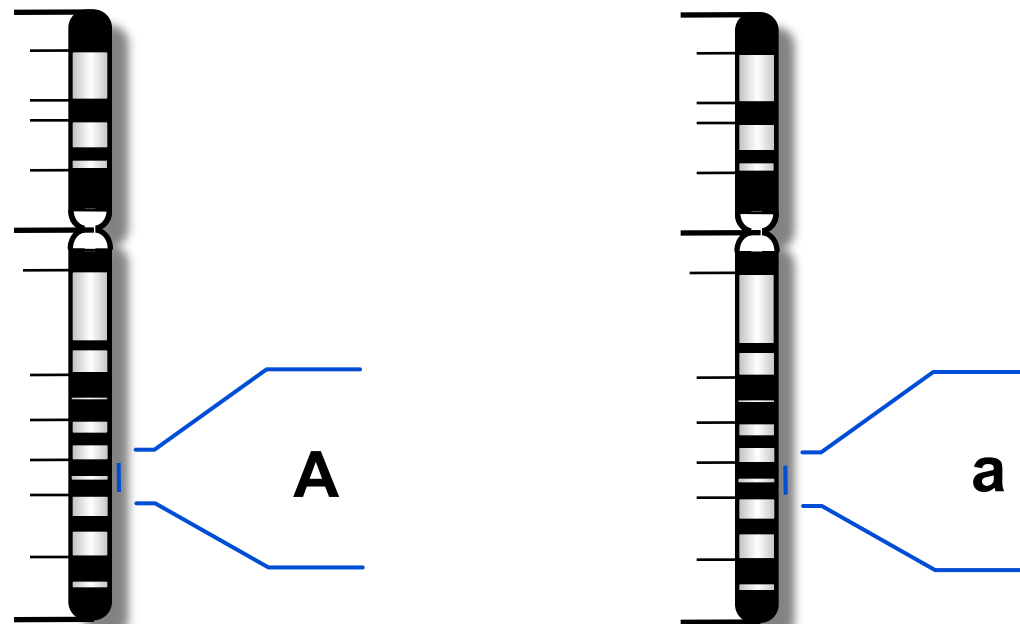
## Non solo mutazioni ...

- Allele molto raro nella popolazione
- Effetto patogenetico sulla funzione del prodotto genico

## ... ma anche polimorfismi ...

- Allele più raro con frequenza  $\geq 1\%$
- Generalmente considerato una variante “neutra” senza effetti sul fenotipo

# Stima delle frequenze genotipiche e alleliche



**Genotipo: Aa**



# Calcolo delle frequenze alleliche

	Genotipo			
	AA	Aa	aa	Totale
N° individui	40	47	13	100
Frequenza genotipo	0,40	0,47	0,13	1
N° alleli "A"	80	47	0	127
N° alleli "a"	0	47	26	73
Totale N° alleli				200

**Frequenza  
allelica**

$$F(A) = p = 127/200 = 0.635$$

$$F(a) = q = 73/200 = 0.365$$

$$p + q = 1$$

## Calcolo delle frequenze alleliche

	Genotipi/fenotipi						
	A1A1	A1A2	A1A3	A2A2	A2A3	A3A3	Totale
N° individui	2450	1400	700	200	200	50	5000
Frequenza genotipi	0,49	0,28	0,14	0,04	0,04	0,01	
N° alleli "A1"	4900	1400	700				7000
N° alleli "A2"		1400		400	200		2000
N° alleli "A3"			700		200	100	1000
Totale N° alleli							10000

**Frequenza  
allelica**

$$F(A1) = p = 7000/10000 = 0,70$$

$$F(A2) = q = 2000/10000 = 0,20$$

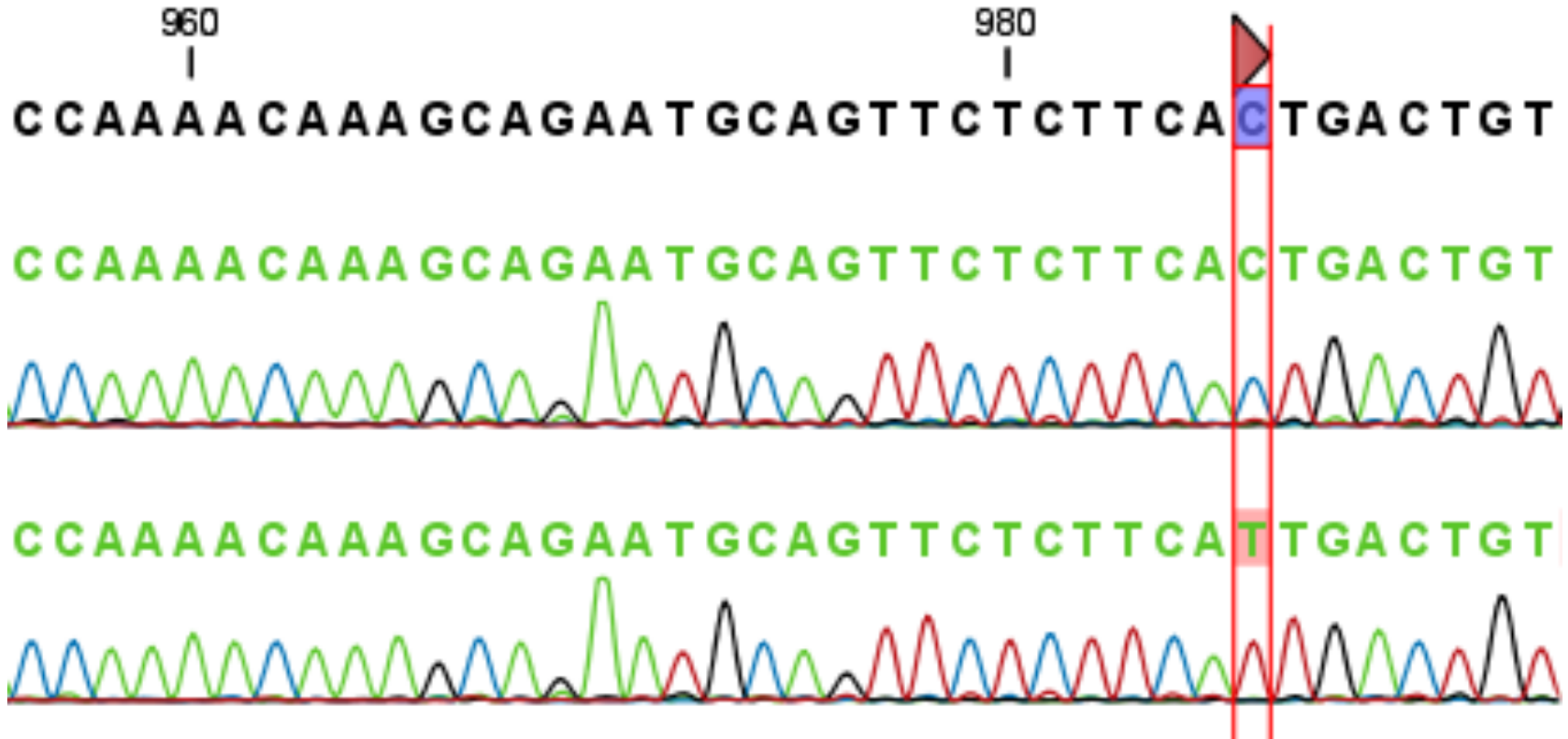
$$F(A3) = r = 1000/10000 = 0,10$$

$$p + q + r = 1$$

# Diversi tipi di polimorfismi

- Polimorfismi del DNA
  - **SNP (single nucleotide polymorphism)**: sostituzioni singoli nucleotidi;
  - **RFLP (restriction full length polymorphism)**: SNP che crea o distrugge un sito riconosciuto da un enzima di restrizione;
  - Loci ipervariabili: **microsatelliti**, minisatelliti, VNTR
  - **CNV (copy number variations)**
- Polimorfismi sierologici e immunologici (HLA, sistema ABO, etc.)
- Polimorfismi dei cromosomi

# SNP



# SNP/RFLP

BamHI

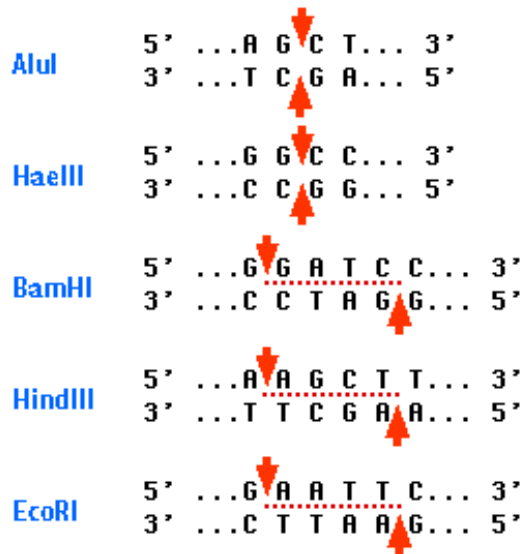
ACTGGGTACG**G**ATCCATTCA

450 bp

400bp

ACTGGGTACG**C**ATCCATTCA

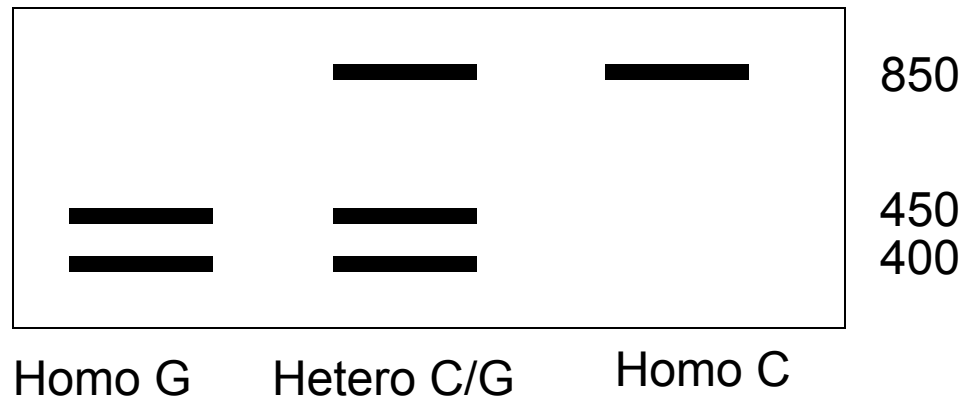
850 bp



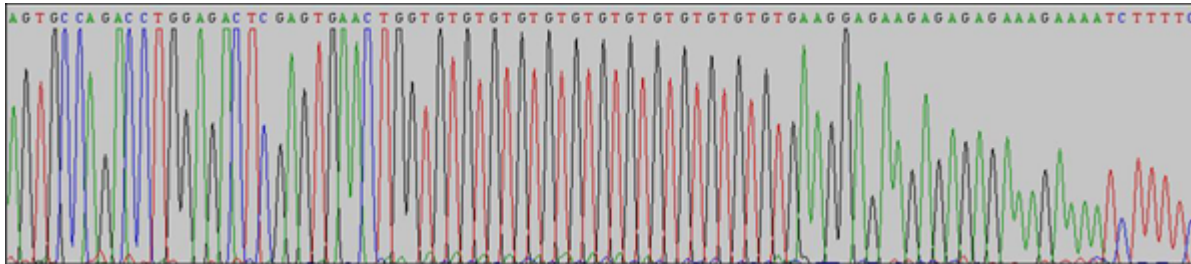
AluI and HaeIII produce blunt ends

BamHI HindIII and EcoRI produce "sticky" ends

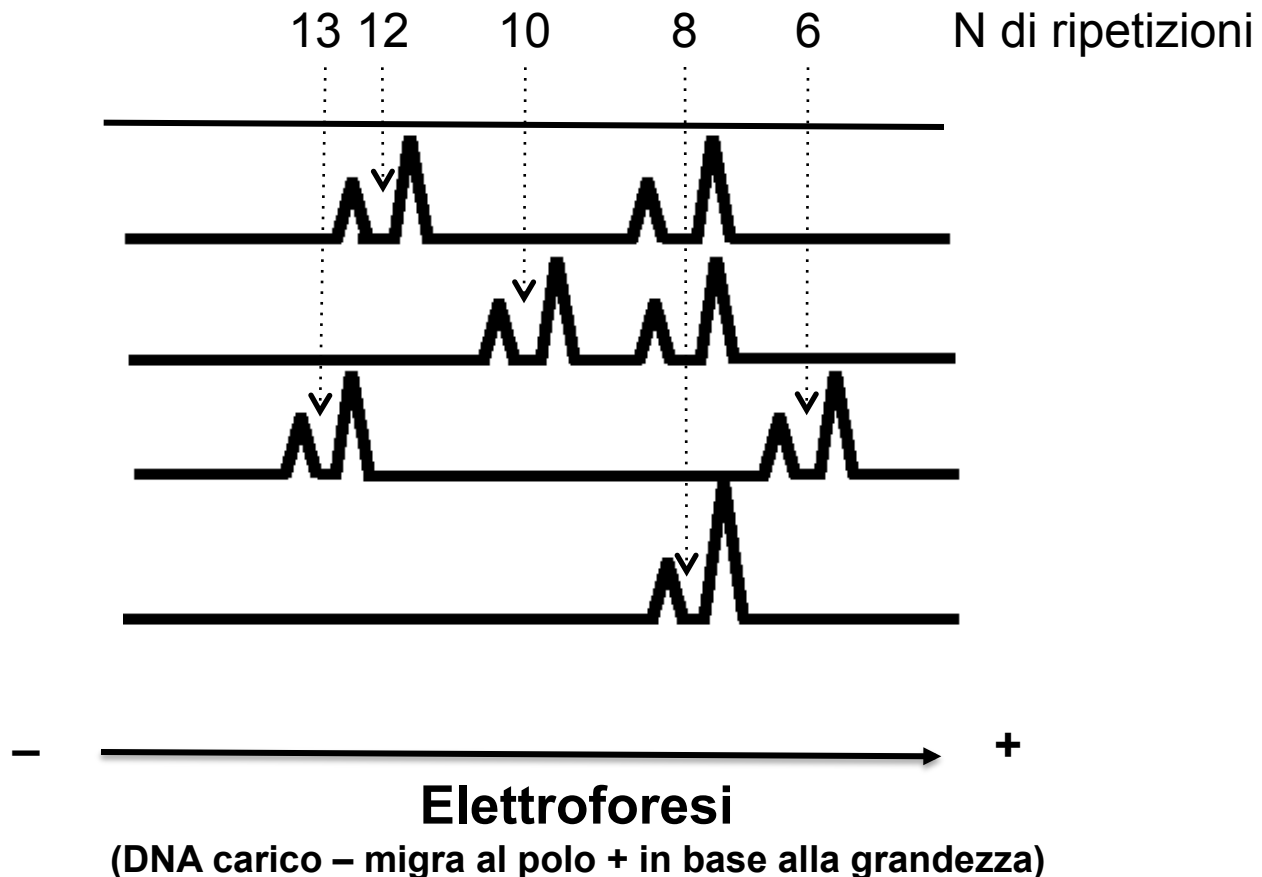
Corsa elettroforetica



# Microsatelliti (ripetizioni di dinucleotidi)



(GT)<sub>n</sub>



# Genotipizzazione di loci polimorfici

**Un locus**



**PCR**

Amplificazione di DNA

(sintesi in vitro di molecole del frammento prescelto)

**Più loci**

1) **SNP array**

2) **NGS**

- **Genoma**
- **Esoma**



Analisi di  
Sequenza  
(altre metodiche)

**SNP**

Digestione con  
Enzimi di restrizione  
E elettroforesi

**RFLP**

Elettroforesi  
capillare

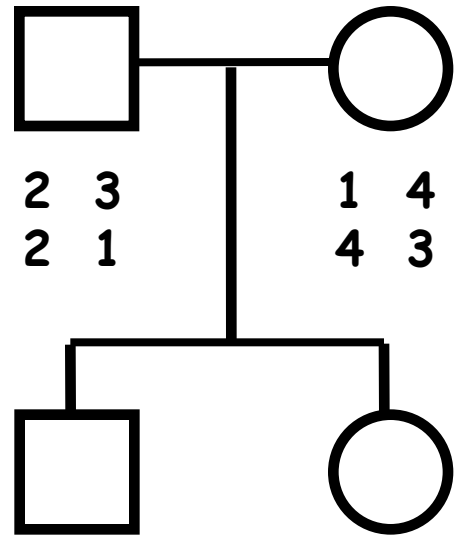
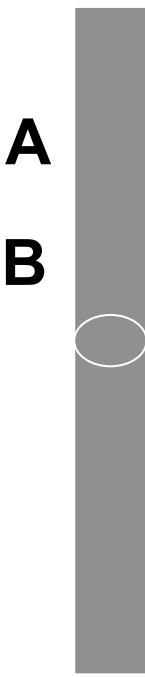
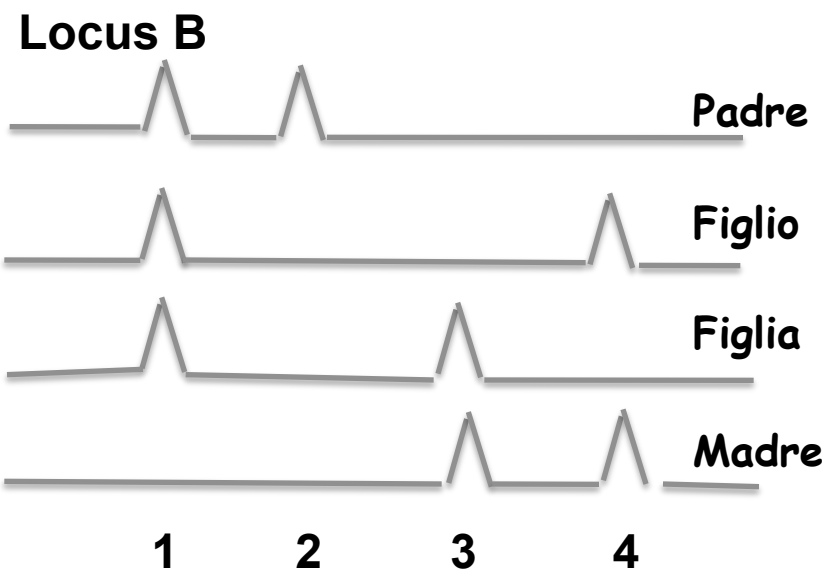
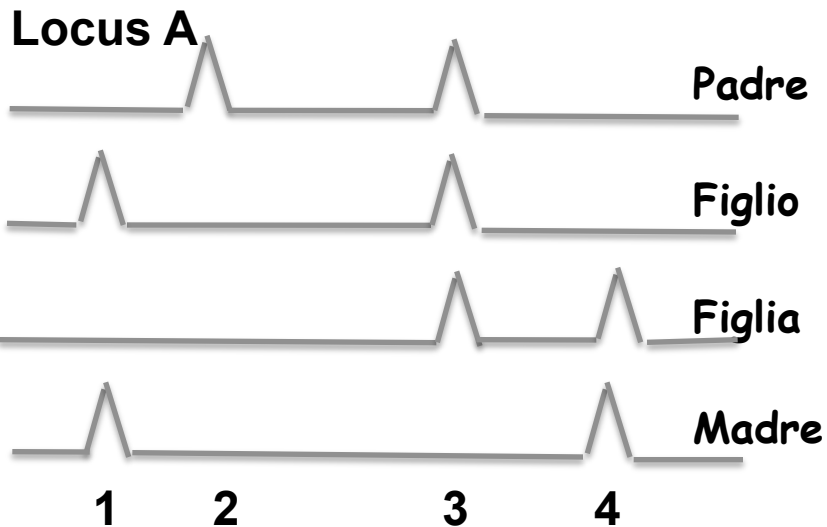
**Microsatelliti**

# Applicazione dei polimorfismi

- Costruzione aplotipo (set di alleli che caratterizzano i cromosomi o porzioni cromosomiche)
- Analisi di linkage per localizzare geni-malattia (identificazione di geni in malattie mendeliane)
- Studi di associazione (malattie multifattoriali)
- Numerose applicazioni in diverse problematiche (esempi nelle diapositive seguenti)



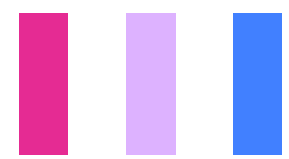
# Analisi di microsatelliti e costruzione dell'aplotipo



Locus A



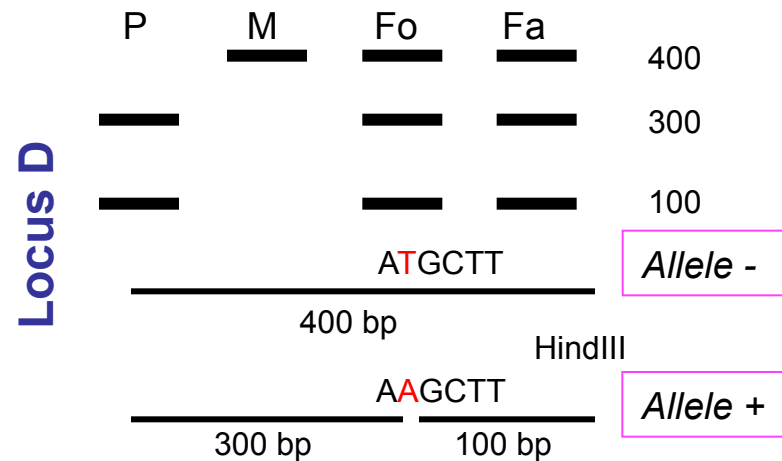
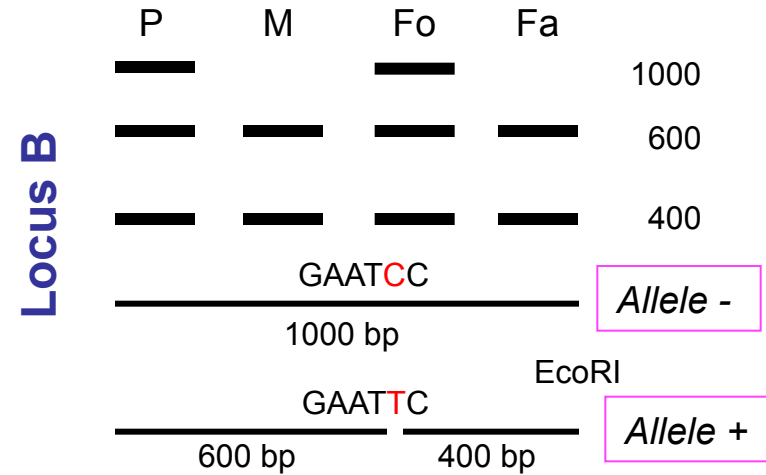
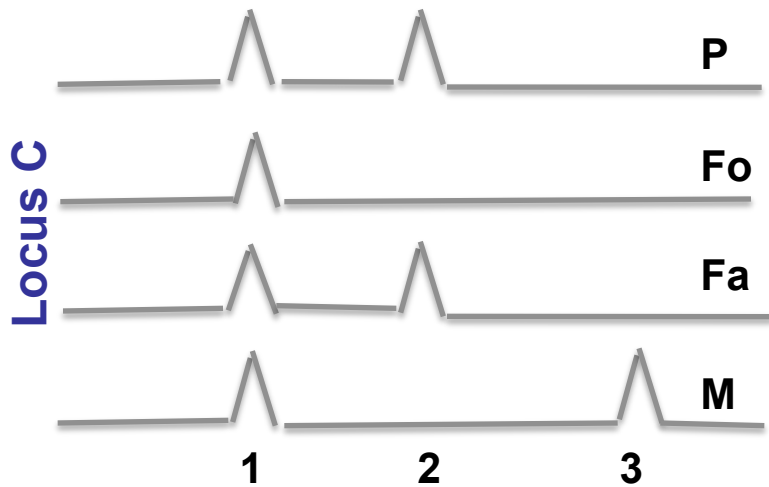
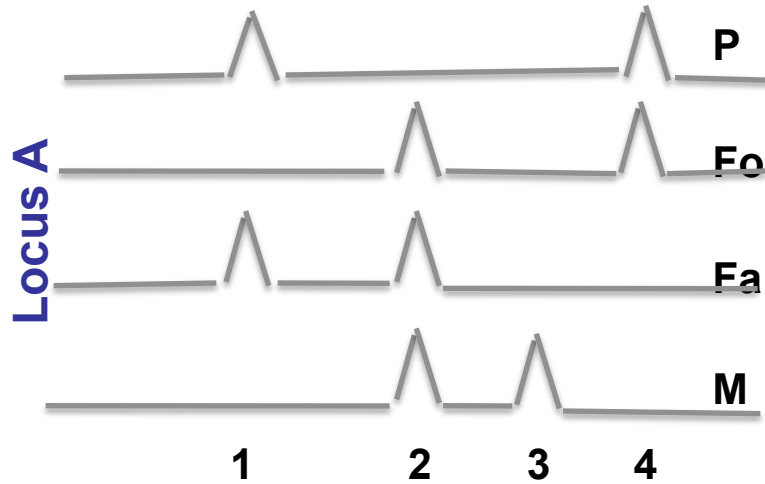
Locus B



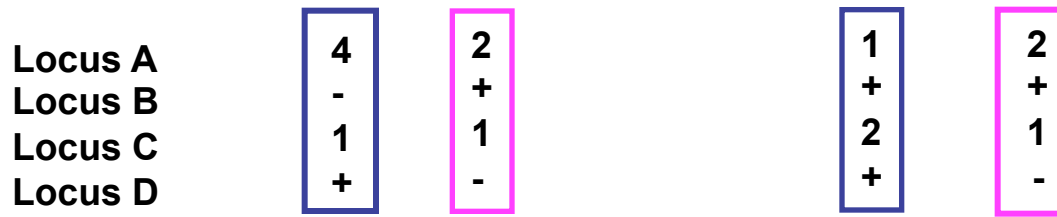
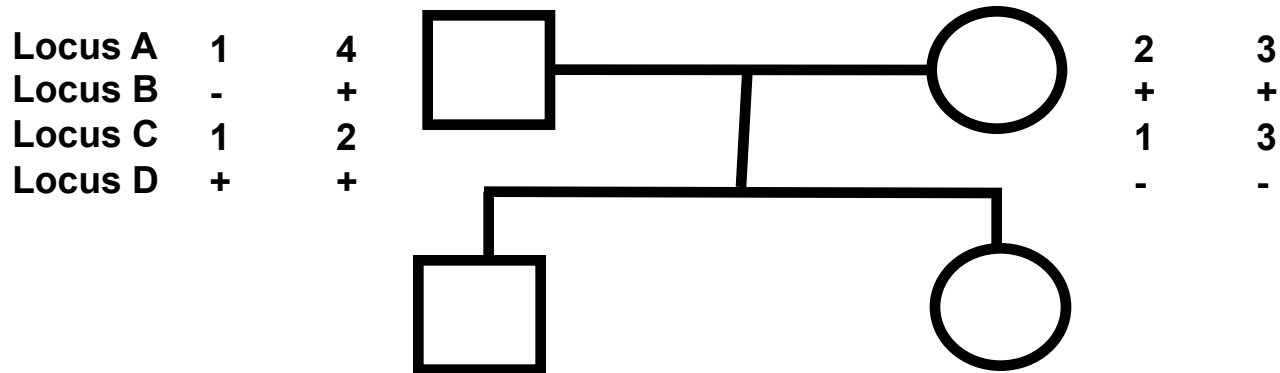
Aplotipo:  
assetto allelico su cromosomi

# Dati i seguenti 4 loci, costruire l'aplotipo (vedi soluzione diapo successiva)

A  
B  
C  
D



Locus	P	M	Fo	Fa
A	1 4	2 3	2 4	1 2
B	+ -	+ +	+ -	+ +
C	1 2	1 3	1 1	1 2
D	+ +	- -	+ -	+ -



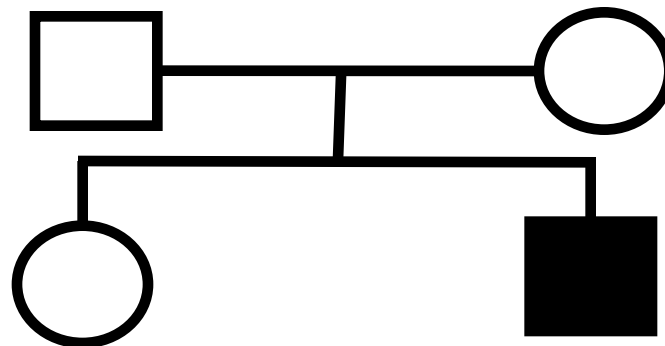
Aplotipo trasmesso dal padre

Aplotipo trasmesso dalla madre

# Utilizzo loci polimorfici: identificazioni eventuali delezioni

Locus	P	M	Fa	Fo
A	133 135	<b>131 133</b>	131 135	131 133
B	150 160	<b>154 156</b>	150 154	160 160
C	244 250	<b>246 252</b>	250 252	244 244
D	179 183	<b>185 195</b>	179 185	183 183
E	228 234	<b>224 234</b>	224 234	224 228
F	120 127	<b>114 120</b>	120 120	120 127
G	157 169	<b>153 167</b>	157 167	167 169

**Apparente omozigosità**



Sospetto  
Sindrome di Williams (7q)

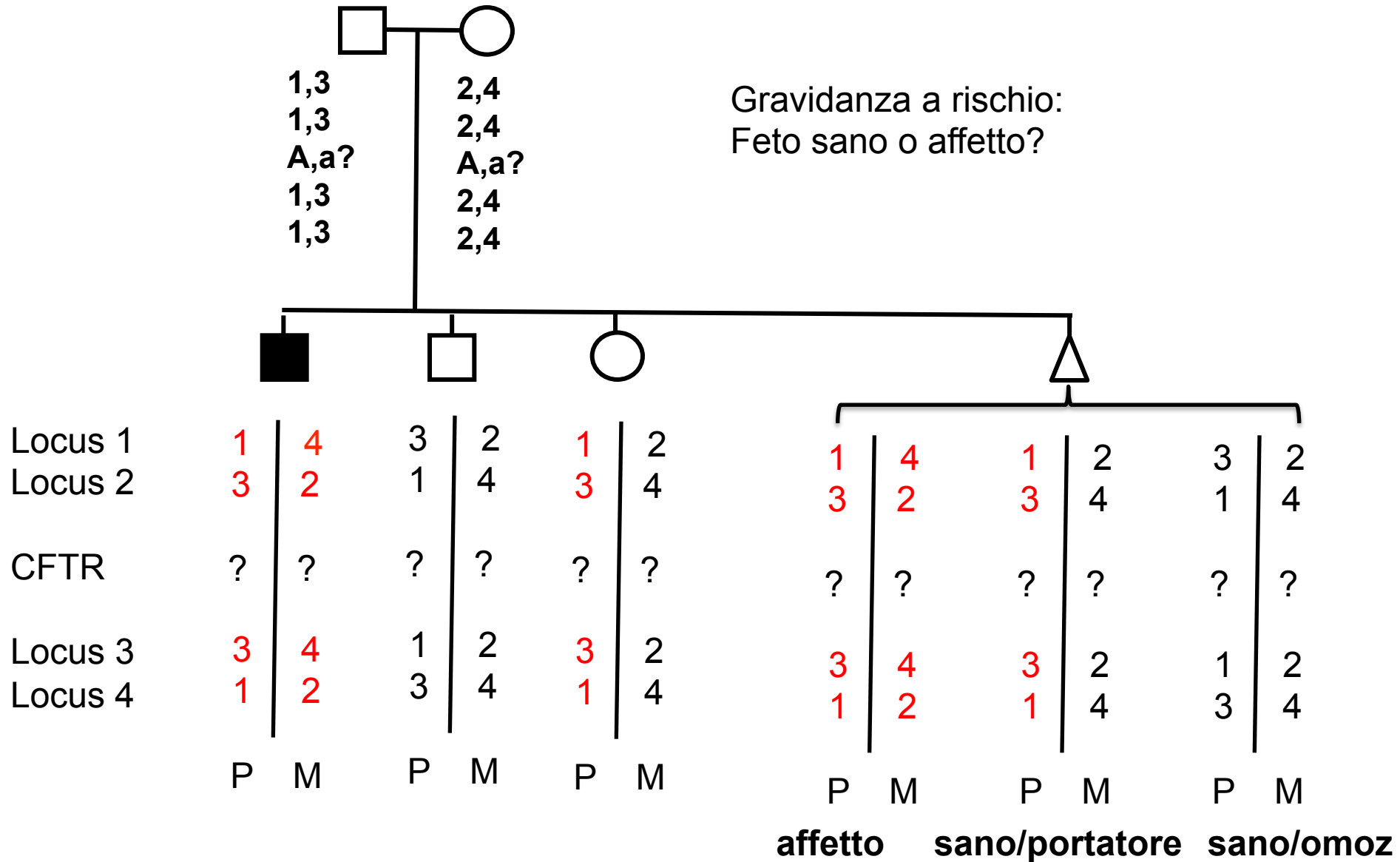
Marker A  
Marker B  
Marker C  
Marker D  
Marker E  
Marker F  
Marker G

135	131
150	154
250	252
179	185
234	224
120	120
157	167

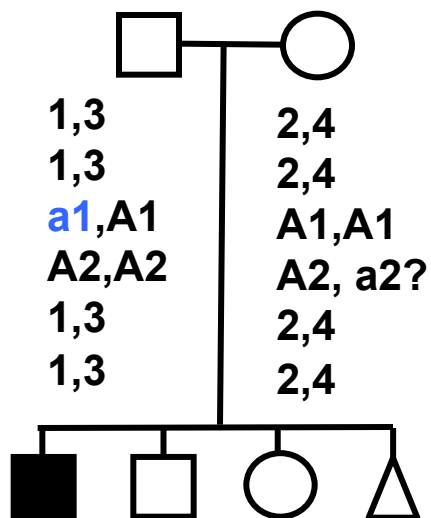
133	131
160	-
244	-
183	-
228	224
127	120
169	167

**Delezione  
Emizigosità  
ai loci B, C e D**

# Utilizzo loci polimorfici: Diagnosi indiretta (CF)/mutazioni non note



# Utilizzo loci polimorfici: Diagnosi indiretta (CF) / una mutazione nota



Gravidanza a rischio:  
Feto sano o affetto?

- 1) Analisi mutazione nota  $a_1$ 
  - a) Se assente > genotipo  $Aa_2$  oppure  $AA$
  - b) Se presente > genotipo  $Aa_1$  oppure  $a_1a_2$
- 2) Se presente, procedere con l'analisi dell'aplotipo paterno
  - a) 4242 > affetto  $a_1a_2$
  - b) 2424 > portatore  $Aa_2$

Locus 1	1	4	
Locus 2	3	2	
CFTR	a1	A1	nota
	A2	a2?	non nota
Locus 3	3	4	
Locus 4	1	2	
	P	M	

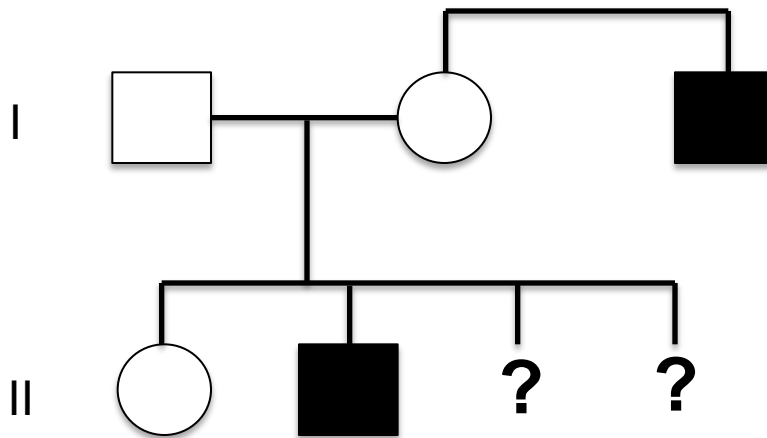
	1	4
	3	2
	a1	A1
	A2	a2?
	3	4
	1	2
P	M	
<b>affetto</b>		

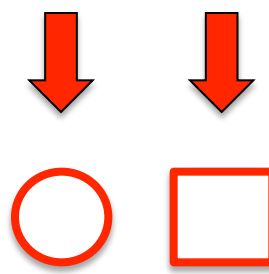
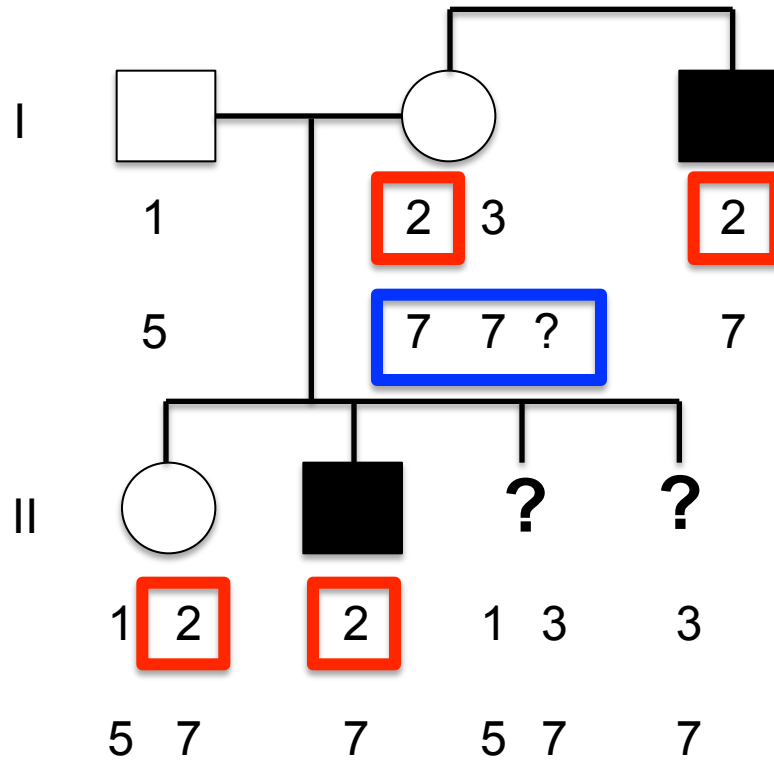
	1	2
	3	4
	a1	A1
	A2	A2
	3	2
	1	4
P	M	
<b>sano/portatore</b>		

Data la seguente famiglia con malattia X-linked recessiva e i relativi genotipi di due loci molto vicini al gene che causa la malattia, stabilire:

- A) Il locus informativo ai fini di una diagnosi prenatale
- B) Genere e condizione (sano/affetto/portatore) di II-3 e II-4

	I-1	I-2	I-3	II-1	II-2	II-3	II-4
<b>LOCUS A</b>	1	2,3	2	1,2	2	1,3	3
<b>LOCUS B</b>	5	7	7	5,7	7	5,7	7





**Sani**

**Locus A è informativo**



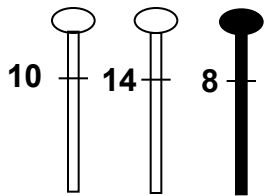
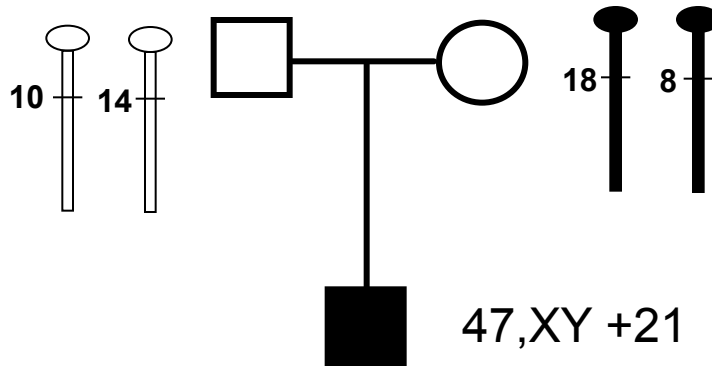
## Utilizzo loci polimorfici: paternità (analisi di almeno 20 loci)

Locus	P	M	Fo	Fa
A (1p)	<b>1</b> 3	4 <b>7</b>	<b>1</b> <b>7</b>	<b>5</b> <b>7</b>
B (3q)	1 <b>2</b>	<b>3</b> 4	<b>2</b> <b>3</b>	<b>3</b> <b>9</b>
C (15p)	<b>1</b> 2	4 <b>5</b>	<b>1</b> <b>5</b>	4 <b>6</b>

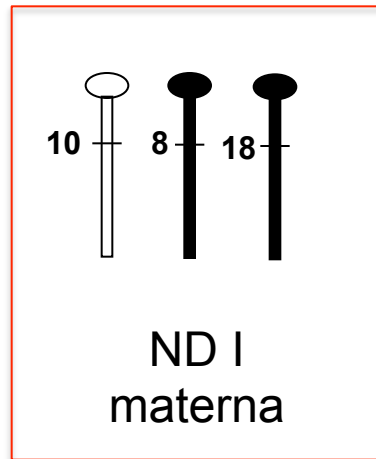
**Fa non figlia di P**

# Utilizzo loci polimorfici: Come determinare l'origine parentale della non-disgiunzione (ND) in I o II divisione meiotica (sindrome di Down)

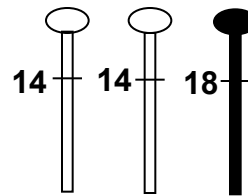
Esempio con un marcatore (microsatellite)



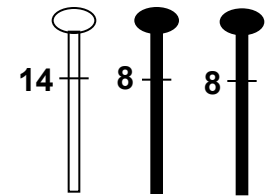
ND I  
paterna



ND I  
materna



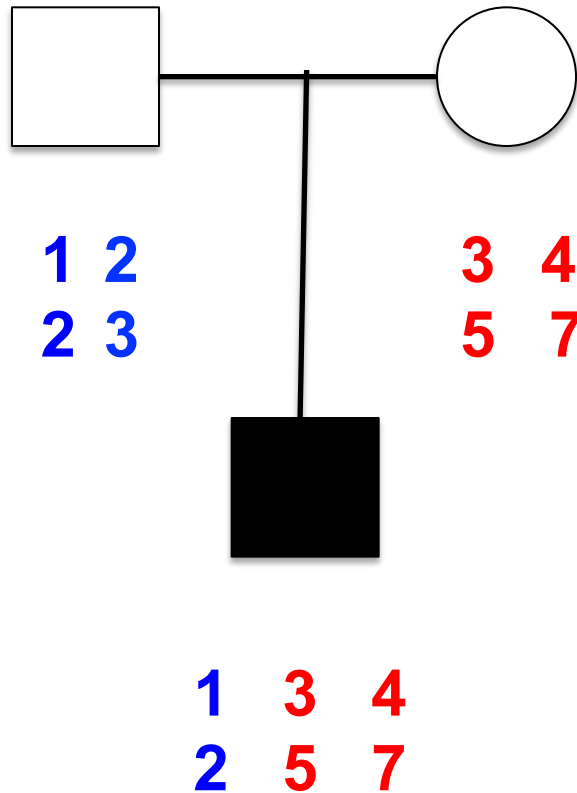
ND II  
paterna



ND II  
materna

In una famiglia con un figlio affetto dalla sindrome di Down, è possibile, dall'analisi dei genotipi di due loci, A e B, localizzati sul cromosoma 21, stabilire la causa della malattia?

	Locus A Chr 21	Locus B Chr 21
Padre	1,2	2, 3
Madre	3,4	5,7
Figlio Down	1, 3, 4	2, 5, 7



**Non disgiunzione  
I divisione meiotica materna**