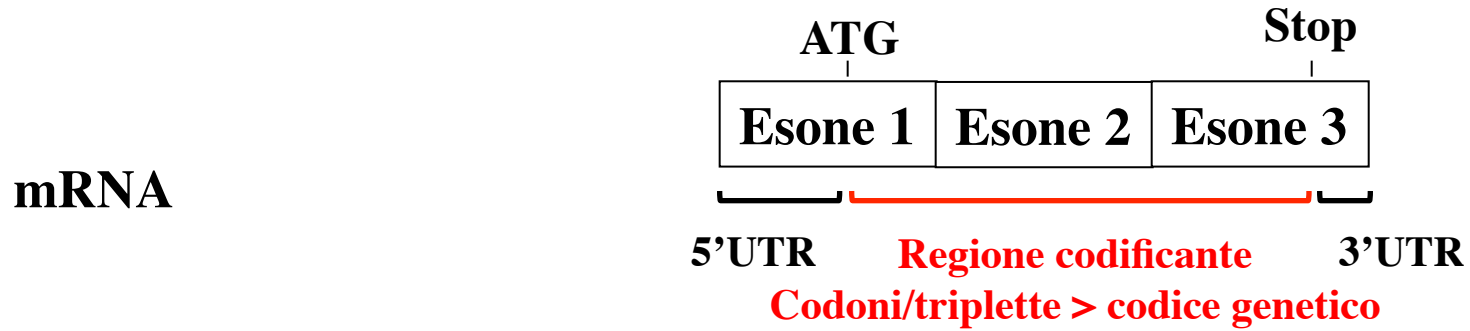
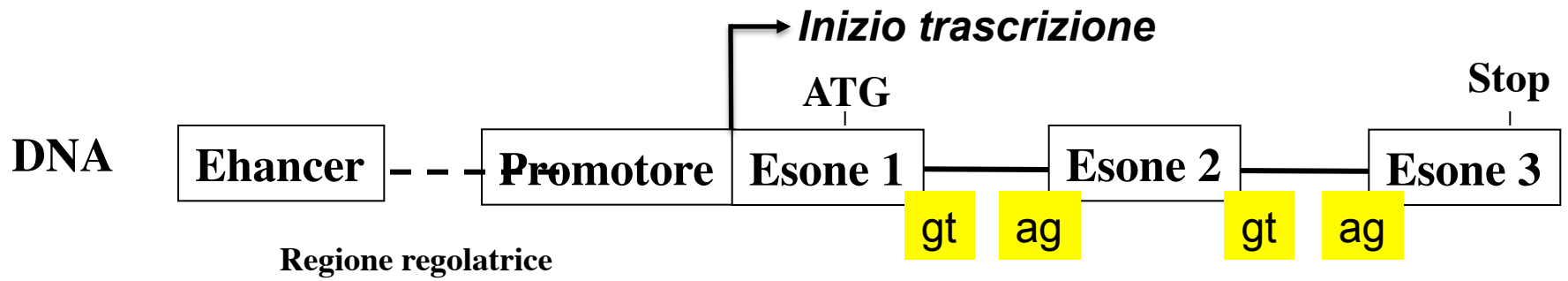


STRUTTURA GENOMICA DEL GENE

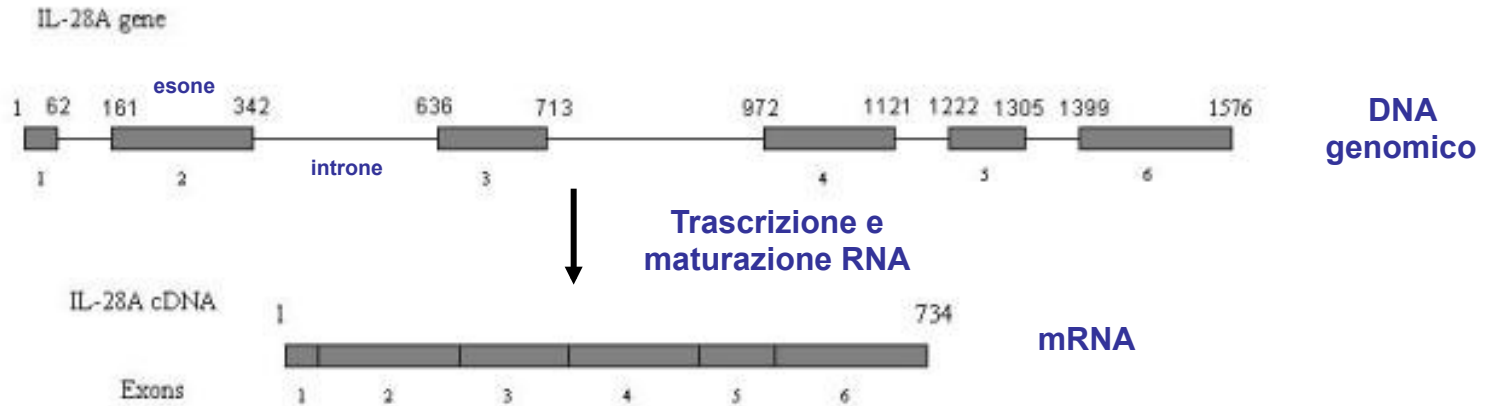
MUTAZIONI GENICHE



		Seconda base					
		U	C	A	G		
Prima base	U	UUU Phe	UCU	UAU Tyr	UGU Cys	U	Terza base
		UUC	UCC Ser	UAC	UGC	C	
		UUA Leu	UCA	UAA Stop	UGA Stop	A	
		UUG	UCG	UAG Stop	UGG Trp	G	
	C	CUU	CCU	CAU His	CGU	U	
		CUC	CCC Pro	CAC	CGC	C	
		CUA	CCA	CAA Gln	CGA	A	
		CUG	CCG	CAG	CGG	G	
	A	AUU	ACU	AAU Asn	AGU Ser	U	
		AUC Ile	ACC Thr	AAC	AGC	C	
		AUA	ACA	AAA Lys	AGA	A	
		AUG Met	ACG	AAG	AGG	G	
	G	GUU	GCU	GAU Asp	GGU	U	
		GUC Val	GCC Ala	GAC	GGC	C	
		GUA	GCA	GAA Glu	GGA	A	
		GUG	GCG	GAG	GGG	G	

Struttura del gene

A)



B)

```

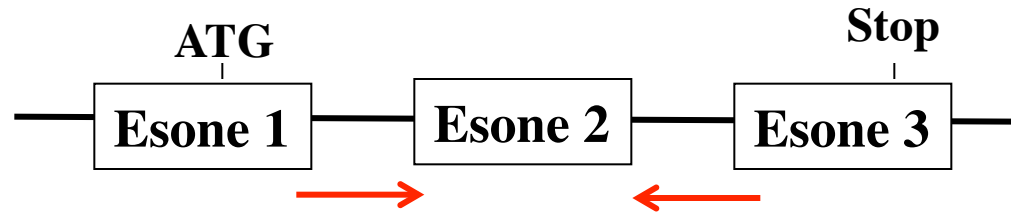
1  tgggtgacag cctcagagtg ttttctctgc tgacaaagac cagagatcag gataaact
61  aggtgagtc cactctctg tccgtgctca gctcctgcag cccctgccct cagtgggcag
121 cctctccatc cctcagatc cctttctctc tgtgacacag acatgactgg ggaactgcag
181 ccagtgtctg tctgtatggc cgcagtgtct accgtgactg gacagcttcc tctctccagg
241 ctccacgggg ctctcccgga tgcaaggggc tgccacatag ccagttccaa gtccctgtct
301 ccacaggagg tgcaggcctt taagaggggc aaagatgctt tagtgagtct cccctgtccc
361 tcttgccatg gactagcttc caccctccct ccaaggctca ccatgcttcc cactccctag
421 cttctctcac tgggtatgac tccacctctc ctgcagtggg ctatctcatg ctctactgtt
481 agggactgac tcatgttttc ctgtagaaga gggctctctc ccatctctcc agcagttaac
541 ctccctctac ctgtttgtcag ccctctctca atccctccag gatggtctaa cctccacctc
601 tctgtctggg gctaacctgt gcccttctct tctaggaaga gtctctctct ctgaaggact
661 gcagggtgca ctcccgcttc ttccctcagg cctgggacct gaggcagctg cagggtgagag
721 ggggagtcag gccacctctc gctctctcag cccctctcac ctggctctgt agtggccctt
781 tcacgtcttc tttctctctt gctctctctc attctctctc cactgtctct ccttctcttc
841 cgtctctcac tgacctctac ggtgtgtctc tctctctctg gctgtctaac ttcattgttt
901 cctctctctc ctgtctctca acctgtctcc ctctctctcc cctcactctg ctcttctctc
961 ctctctctca ggtgagggag cgcctctctg ctttggaggg tgagctgtgc ctgacgtgca
1021 aggttcttga gccacctctc gactctgacc cagcctctgt ggaactcttg gaccagctcc
1081 ttcacacctc gccacctctc ctctctcagt tccgggctct tgtgagctgt tggggctctg
1141 gacacctctc ctgtgagctc tgagcagctg cctctctctc gccaggctcc cggctctctc
1201 accgctctcc tctgctctca gatcagctct cagcctctct cagggtctct gaccgtctct
1261 cgcctctctc attggtctga cgggtctctc gaggctctca aaaaggtgag tgacctctga
1321 agagagggag tgaggtcttg ggagctctct ggagctctca accagctctc cctctctctc
1381 atccctctct cctctctctc gtctctcttg tctctctctg cctctctctc ctctctctc
1441 ttctctctcc tctctctctc cctgaattgt gttgctctct gggctctctg tgtctctctc
1501 tctctctctc cctctctctc gagctctctc ttataaatta gccctctctc ttaattctct
1561 gccacctctc cgtctctc
    
```

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)

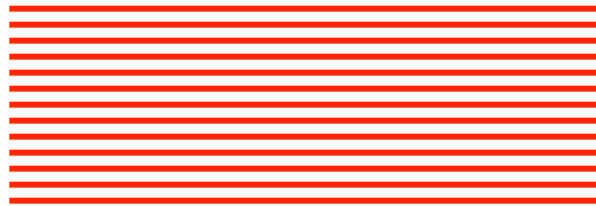
Screening di mutazioni



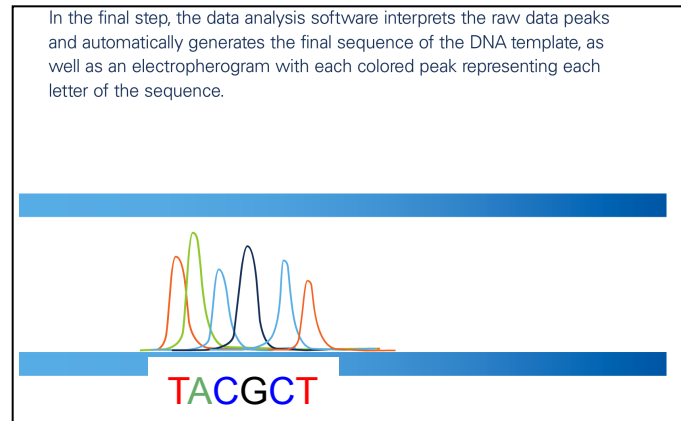
PCR

Amplificazione di DNA

(sintesi in vitro di molecole del frammento prescelto)



Sequenziamento
Sanger



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?

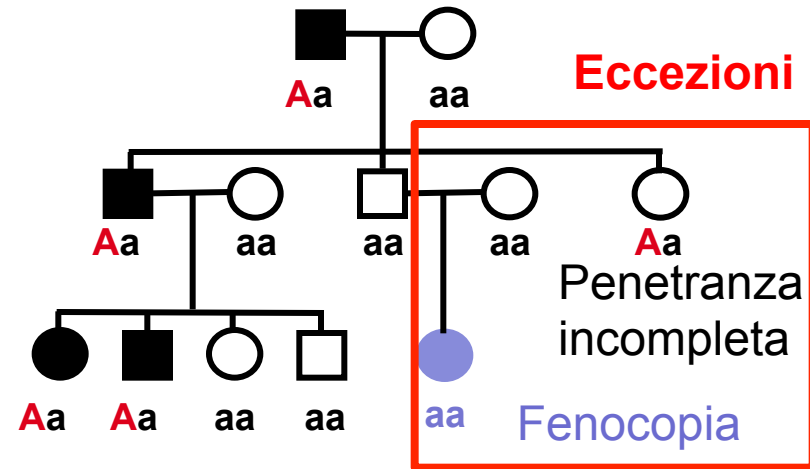
	Beta-globin DNA and amino acid sequence	Beta-globin protein																		
•normal DNA sequence •normal amino acid sequence •normal protein	Normal Sequence <table><tr><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td><td>8</td></tr><tr><td>CTG</td><td>ACT</td><td>CCT</td><td>GAG</td><td>GAG</td><td>AAG</td></tr><tr><td>Leu</td><td>Thr</td><td>Pro</td><td>Glu</td><td>Glu</td><td>Lys</td></tr></table>	3	4	5	6	7	8	CTG	ACT	CCT	GAG	GAG	AAG	Leu	Thr	Pro	Glu	Glu	Lys	
3	4	5	6	7	8															
CTG	ACT	CCT	GAG	GAG	AAG															
Leu	Thr	Pro	Glu	Glu	Lys															
•single base change in DNA sequence •altered amino acid sequence •abnormal protein causing sickle cell anemia	Missense Mutation <table><tr><td>CTG</td><td>ACT</td><td>CCT</td><td>GTG</td><td>GAG</td><td>AAG</td></tr><tr><td>Leu</td><td>Thr</td><td>Pro</td><td>Val</td><td>Glu</td><td>Lys</td></tr></table>	CTG	ACT	CCT	GTG	GAG	AAG	Leu	Thr	Pro	Val	Glu	Lys							
CTG	ACT	CCT	GTG	GAG	AAG															
Leu	Thr	Pro	Val	Glu	Lys															
•single base change in DNA sequence •no change in amino acid sequence •normal protein	Silent Mutation sinonima <table><tr><td>CTG</td><td>ACT</td><td>CCT</td><td>GAA</td><td>GAG</td><td>AAG</td></tr><tr><td>Leu</td><td>Thr</td><td>Pro</td><td>Glu</td><td>Glu</td><td>Lys</td></tr></table>	CTG	ACT	CCT	GAA	GAG	AAG	Leu	Thr	Pro	Glu	Glu	Lys							
CTG	ACT	CCT	GAA	GAG	AAG															
Leu	Thr	Pro	Glu	Glu	Lys															

Effetto?

Effetto?

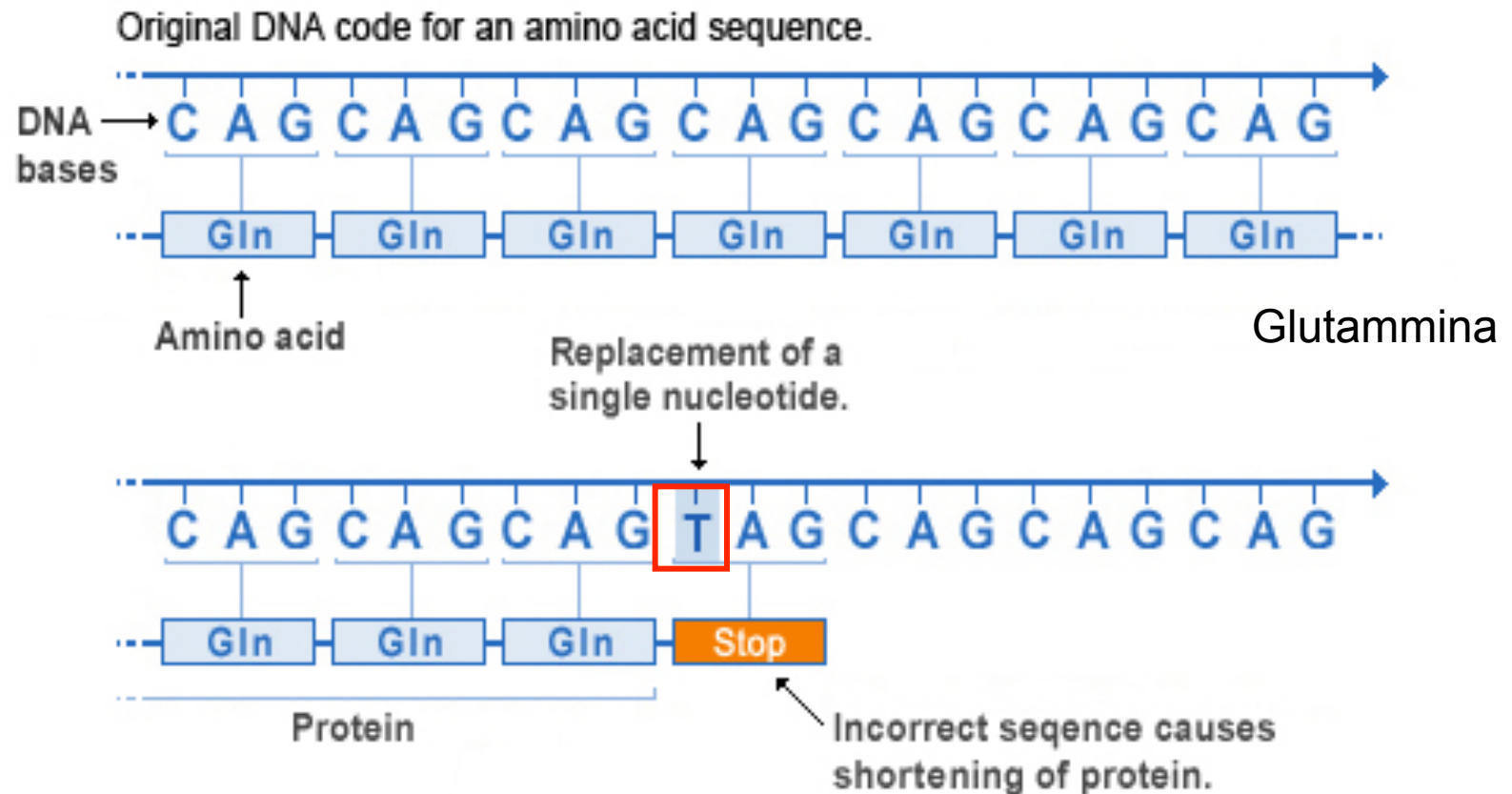
Effetto patogenetico delle varianti missense? (problematica frequente in NGS)

- 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	GLFGRKTGQAPGYS	YTAANKNKGI
Mm	GLFGRKTGQAAGFS	YTDANKNKGI
Gg	GLFGRKTGQAEGFS	YTDANKNKGI
Dr	GLFGRKTGQAEGFS	YTDANKSKGI
Dm	GLIGRKTGQAAGFAY	YTDANKAKGI
At	GLFGRQSGTTAGYS	YSAANKNKAV
Sc	GIFGRHSGQAEGYS	YTDANIKKNV
Bm	GLFGRHSGSVEGYK	YSDANKNSGI

Nonsense mutation



Sostituzioni
nucleotidiche



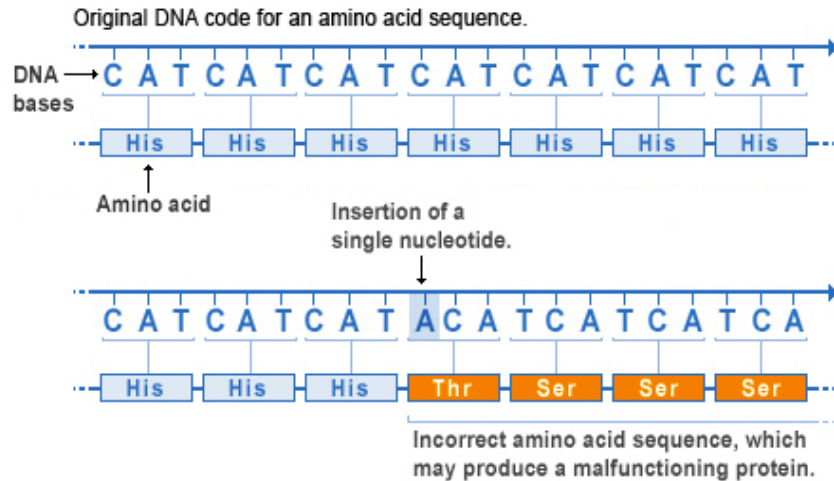
nonsenso



**Effetto
Patogenetico
Generalmente
deleterio**

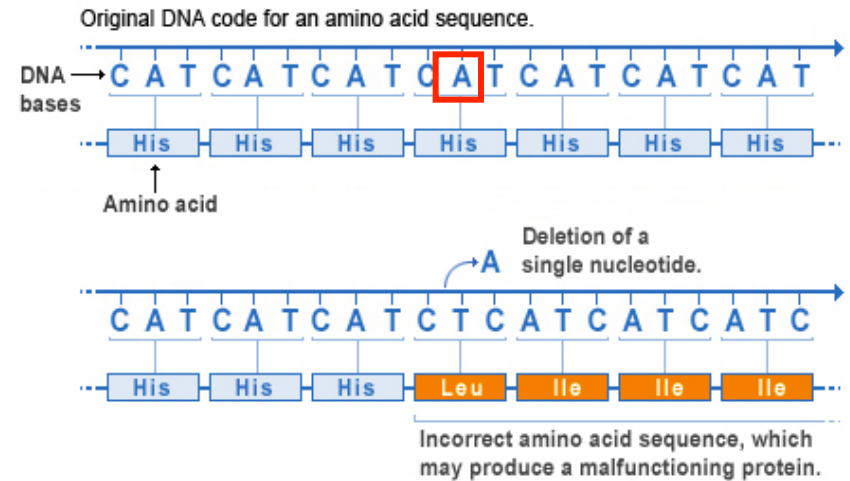
Frameshift mutations

Insertion mutation



U.S. National Library of Medicine

Deletion mutation

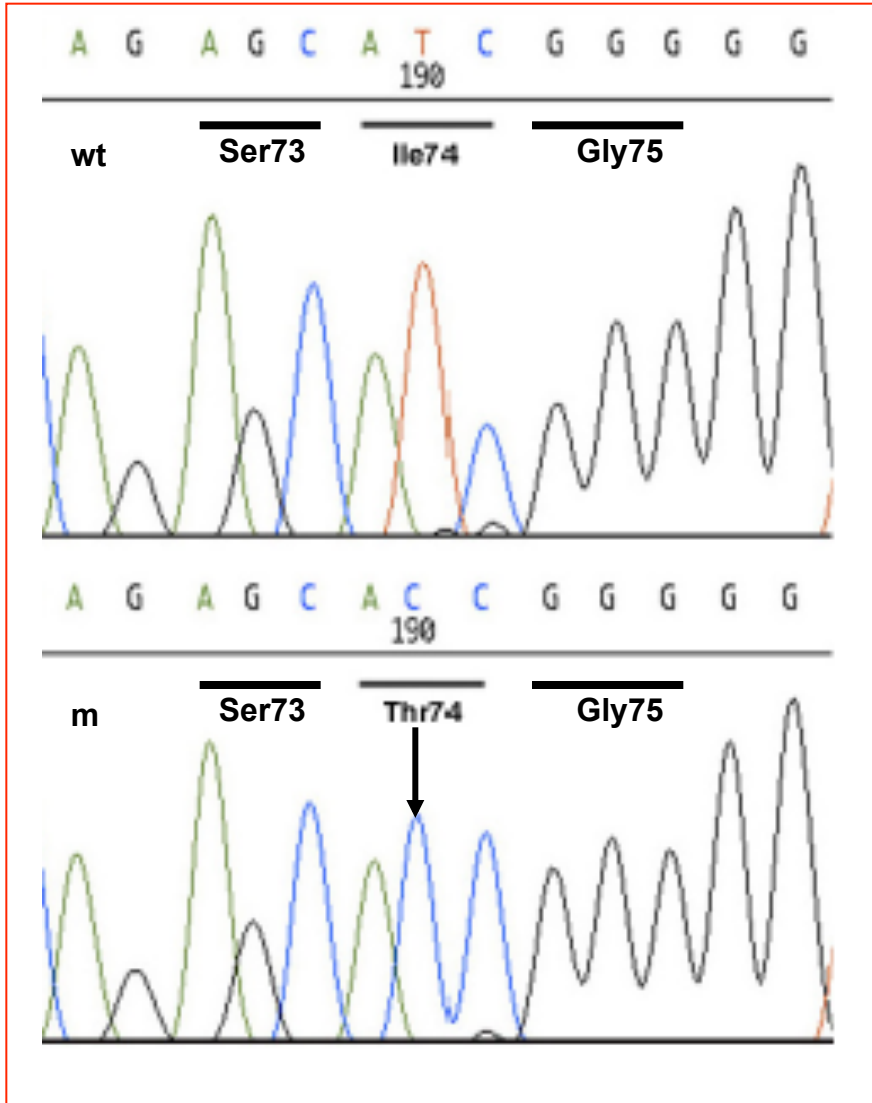


U.S. National Library of Medicine

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

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

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



Sequenza normale

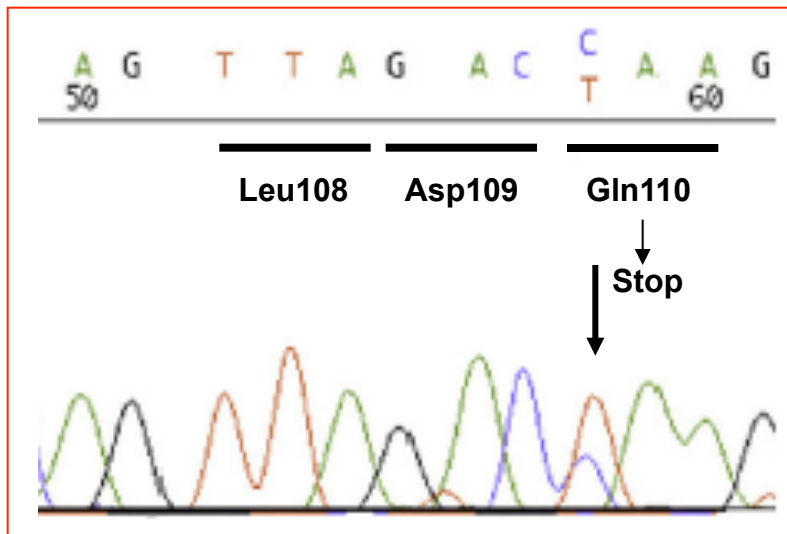
AGAGCA**T**CGGGGG wt allele 1

AGAGCA**T**CGGGGG wt allele 2

Sequenza mutata in omozigosi

AGAGCA**C**CGGGGG m allele 1

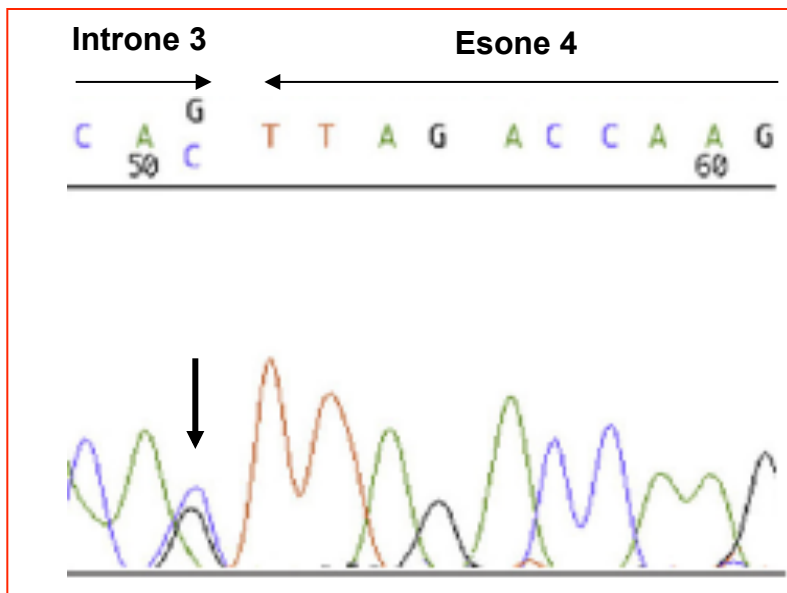
AGAGCA**C**CGGGGG m allele 2



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

AGTTAGAC**C**AAG wt allele 1

AGTTAGAC**T**AAG m allele 2



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

cag TTAGACCAAG wt allele 1

ca**C** TTAGACCAAG m allele 2

Mutazione di splicing

Muzioni geniche: nomenclatura

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

AGTTAGAC**C**AAG wt allele 1

AGTTAGAC**T**AAG m allele 2

c.328C>T

5'



”

A: primo nucleotide +1

p.Gln110*

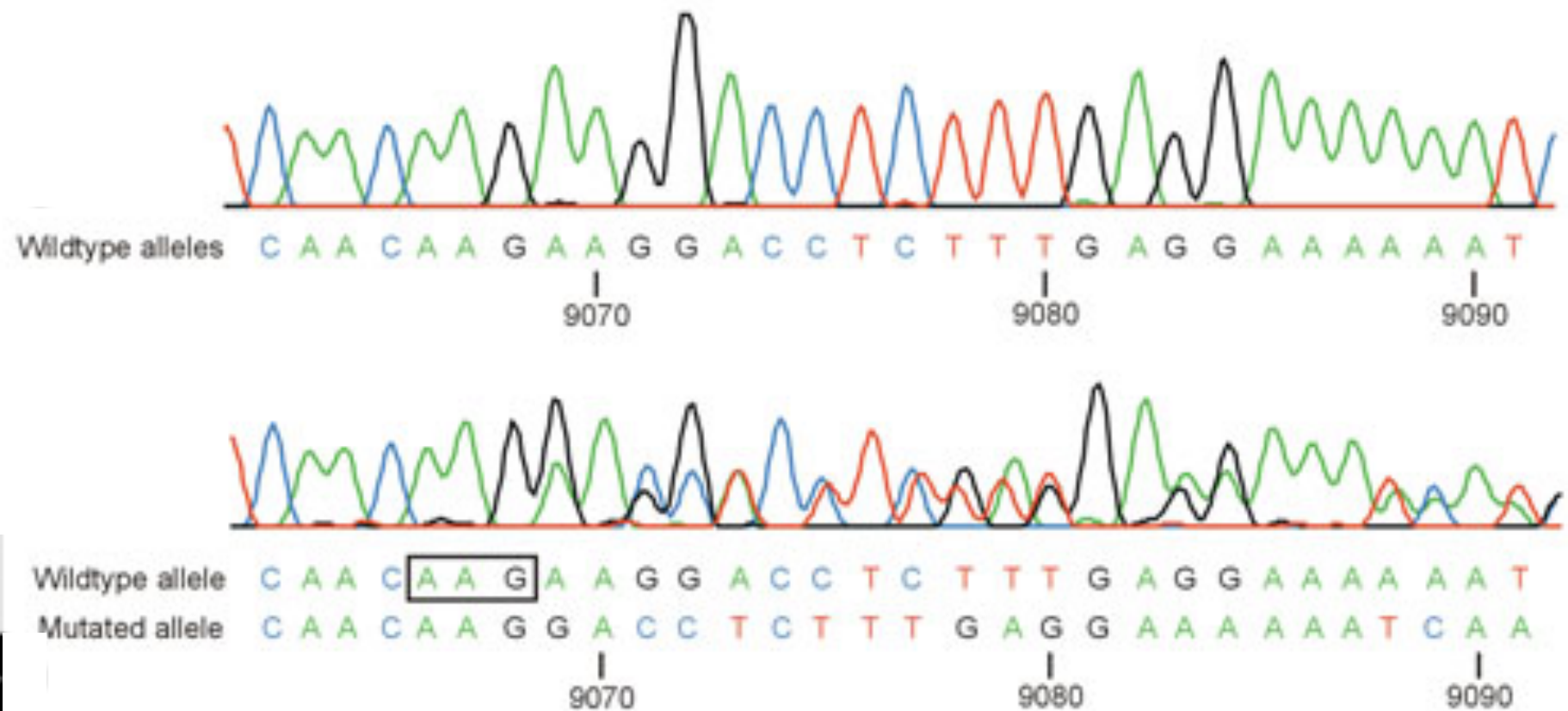
NH2

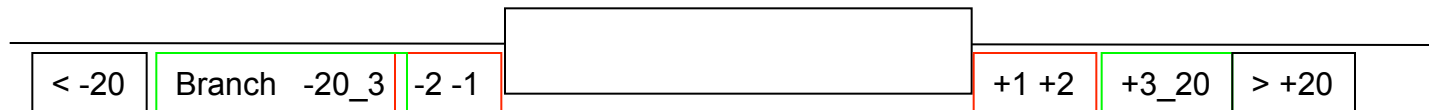
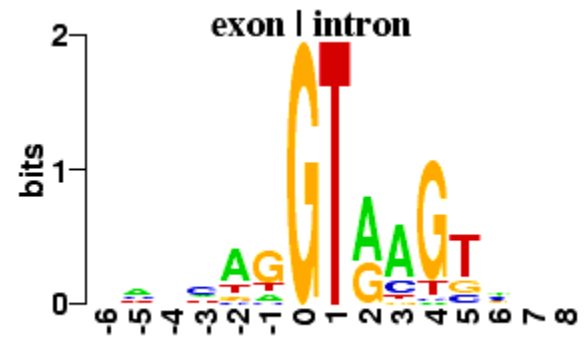
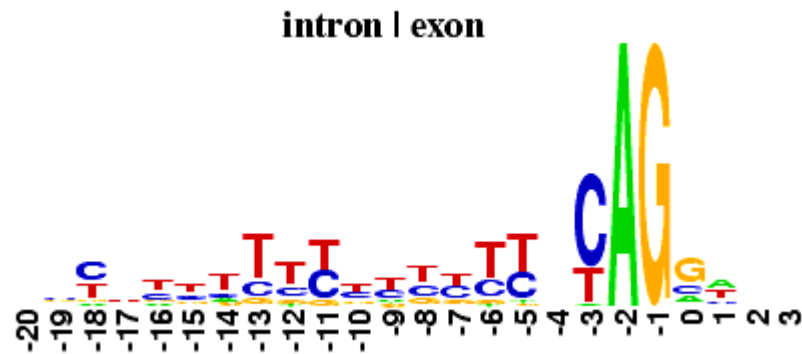
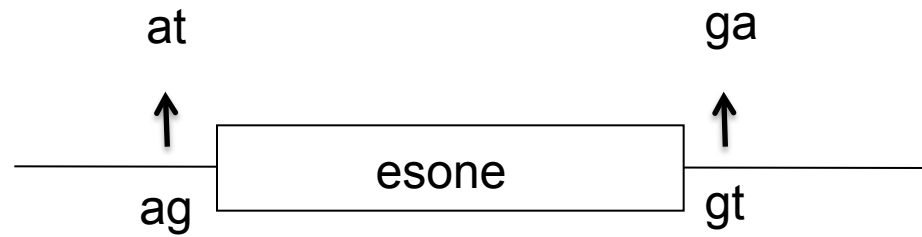


COOH

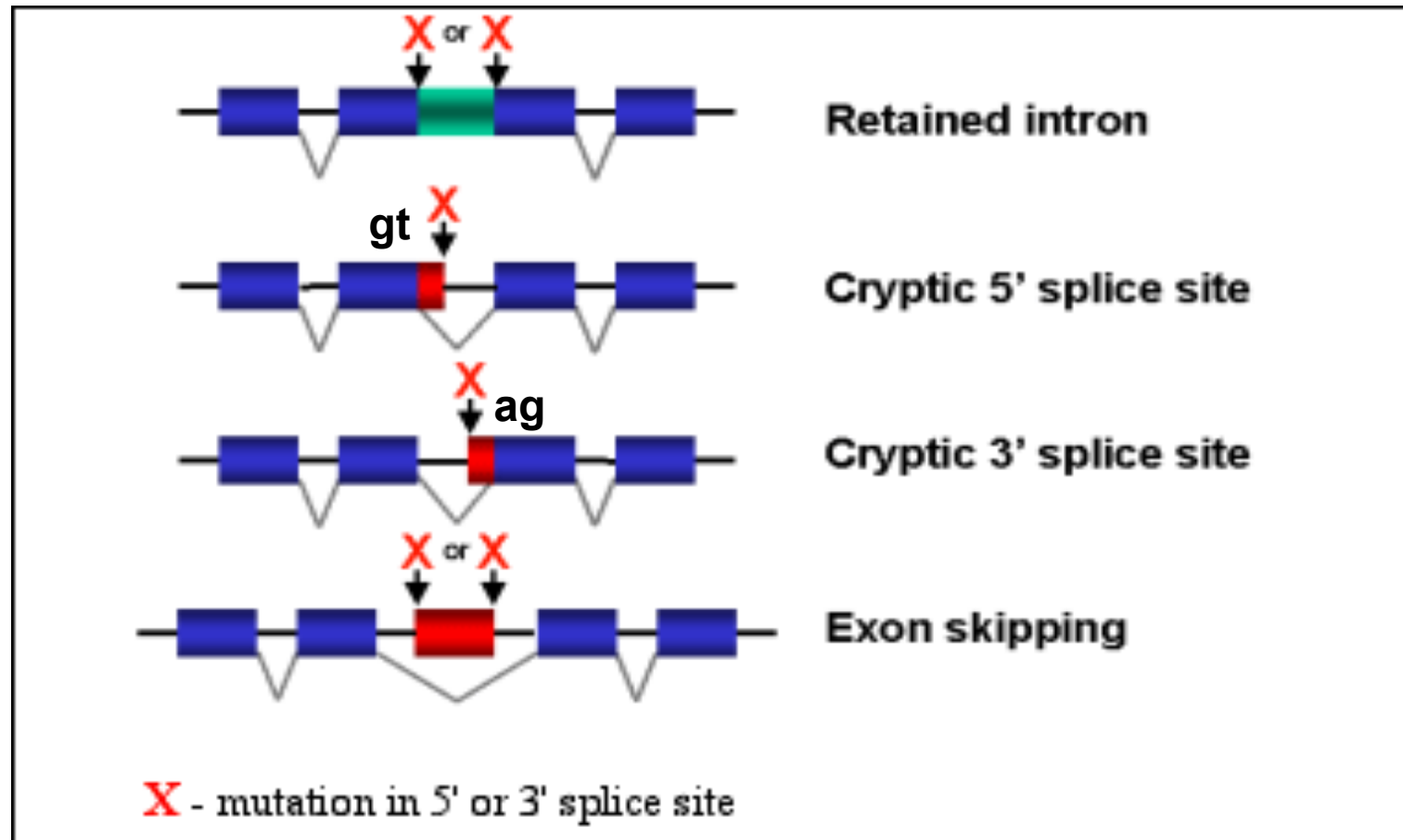
Met: primo amminoacido +1

DELEZIONE ETEROZIGOTE di tre nucleotidi - AAG





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

Allele 1

Allele 2

Affetto

NH₂

X

COOH

NH₂

X

COOH

Sano

NH₂

X

COOH

NH₂

COOH

Sano

NH₂

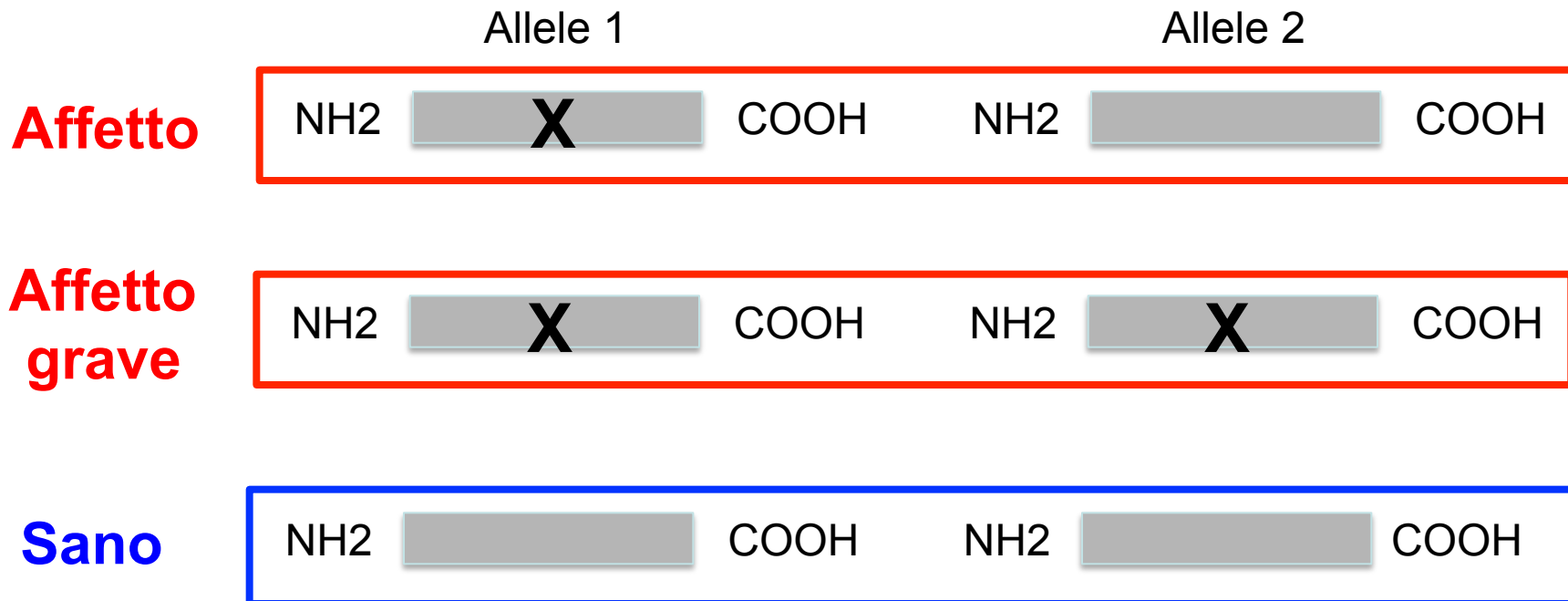
COOH

NH₂

COOH

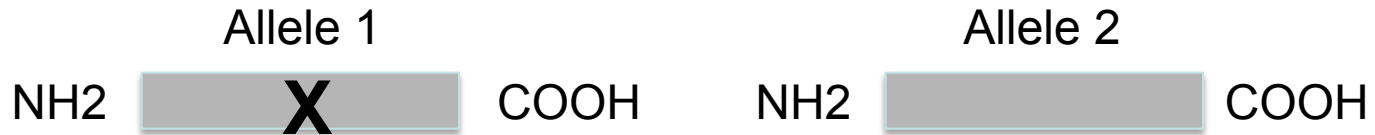
Malattie dominanti

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

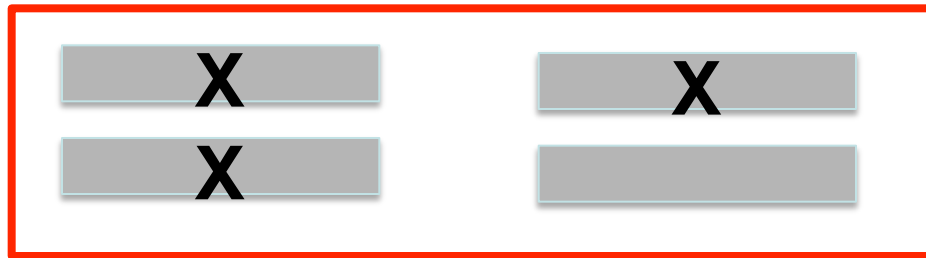


Classificazione delle mutazioni in base all'effetto sulla funzione

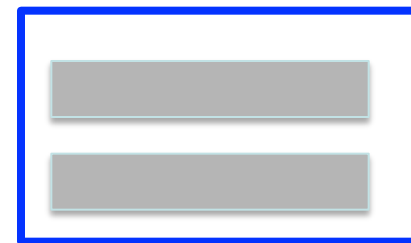
3) Effetto dominante negativo: proteina anomala interferisce con la funzione dell'allele normale



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

Allele 1

NH₂



COOH

Allele 2

NH₂



COOH

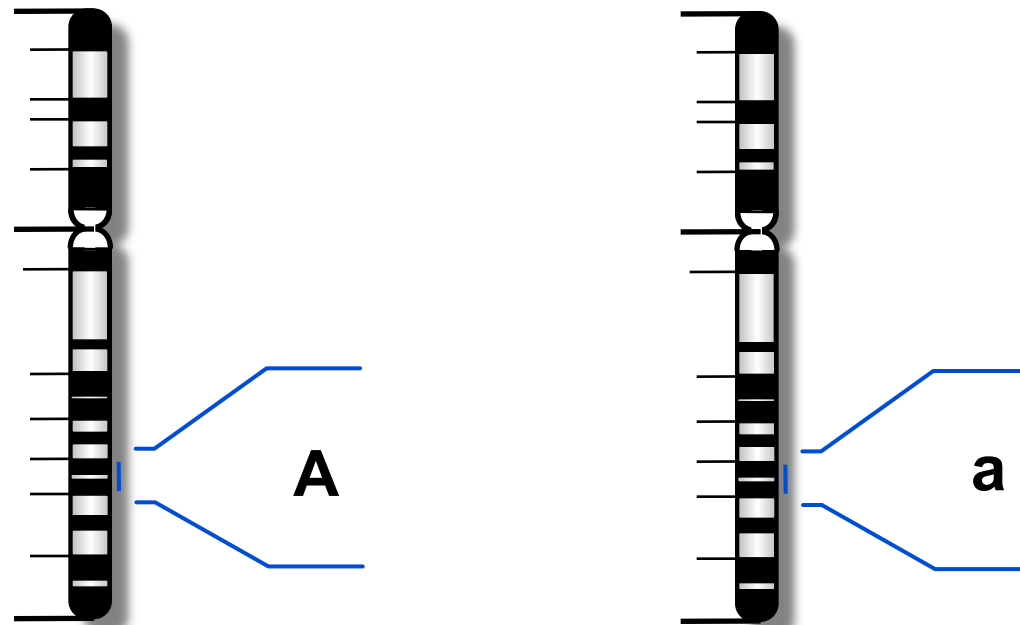
Non solo mutazioni ...

- 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**

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

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

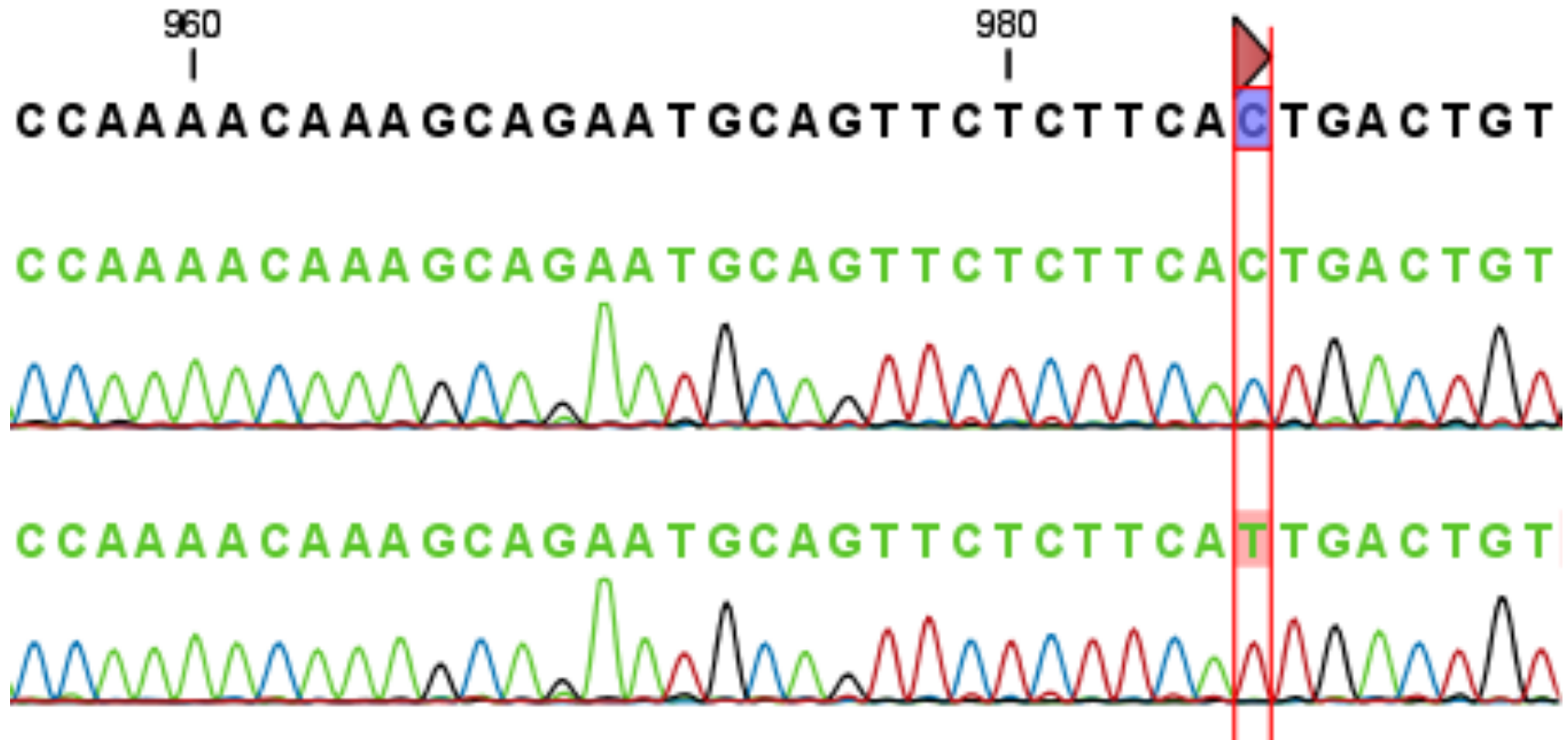
$$F(A3) = r = 1000/10000 = 0,10$$

$$p + q + r = 1$$

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

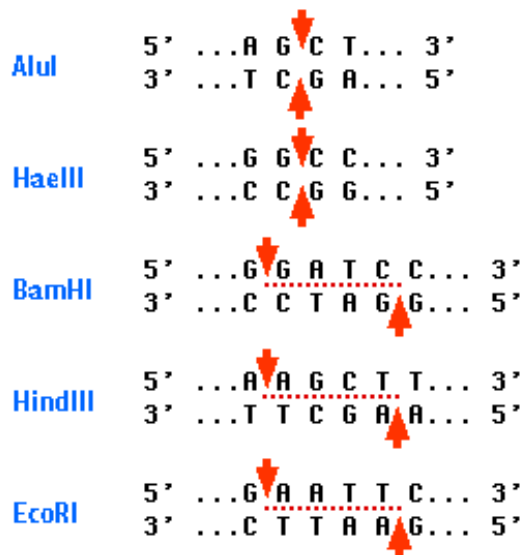
ACTGGGTACG**G**ATCCATTTC

450 bp

400bp

ACTGGGTACG**C**ATCCATTTC

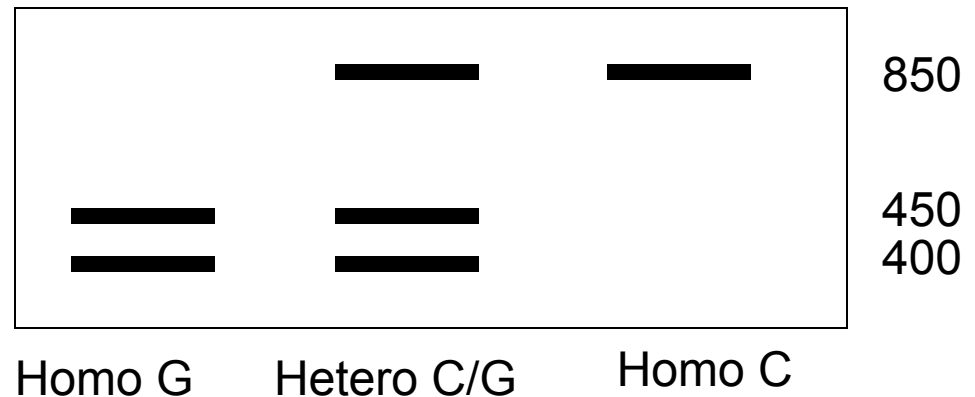
850 bp



AluI and **HaeIII** produce blunt ends

BamHI **HindIII** and **EcoRI** produce "sticky" ends

Corsa elettroforetica



[illegible]

13 12 10 8 6 N di ripetizioni

– +

Elettroforesi
(DNA carico – migra al polo + in base alla grandezza)

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**

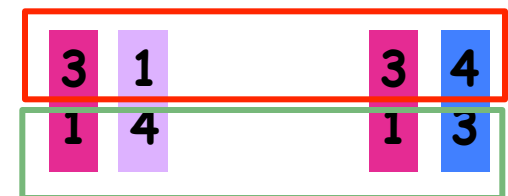
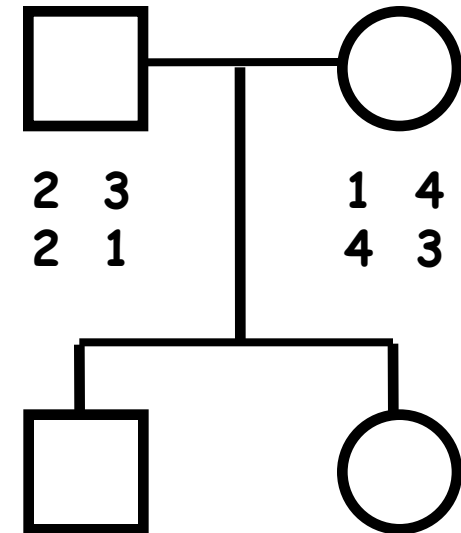
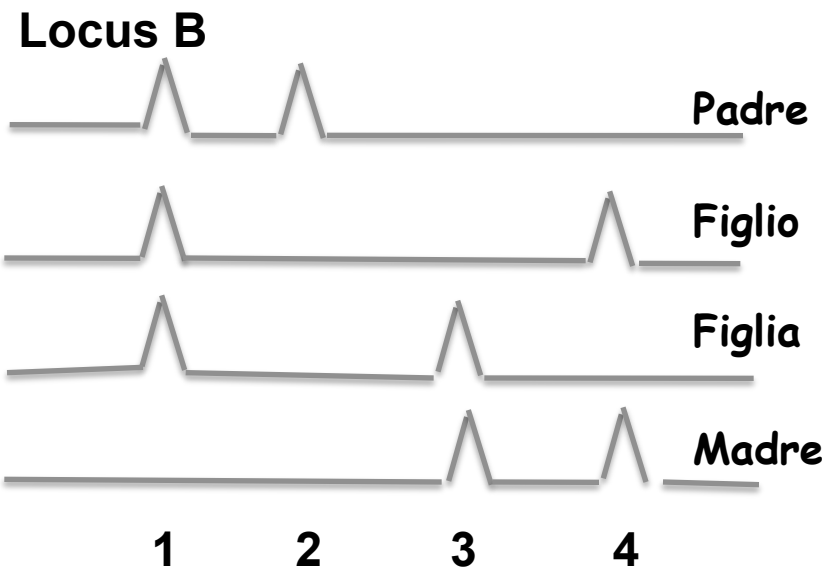
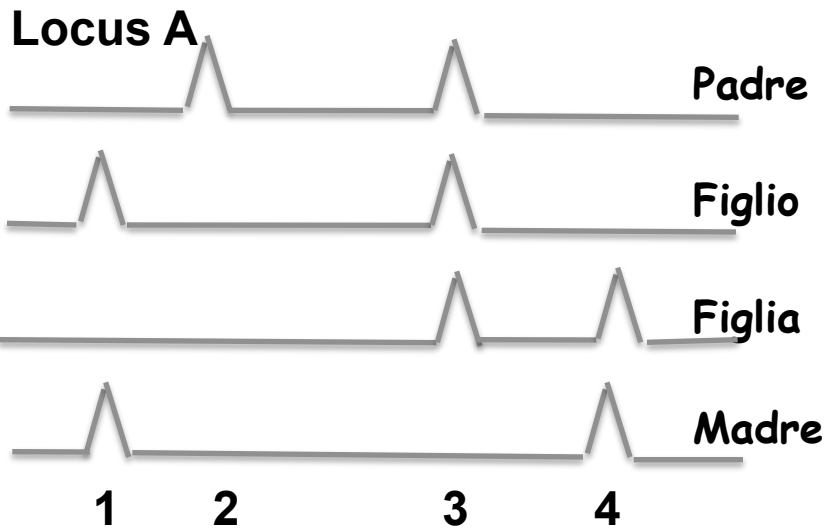
- **Genoma**
- **Esoma**



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

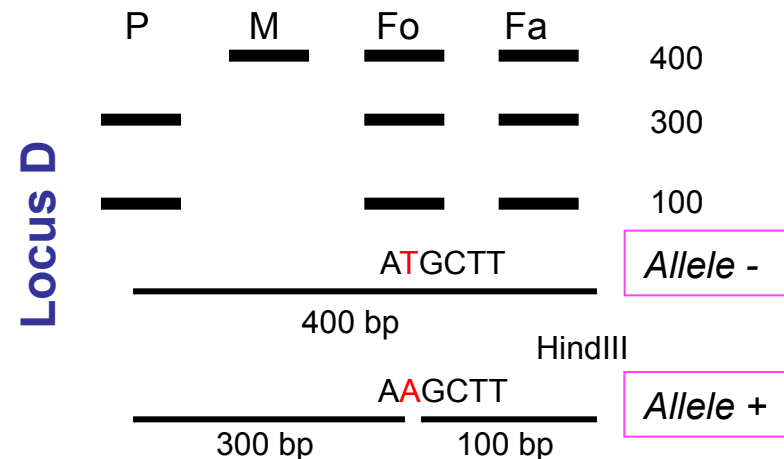
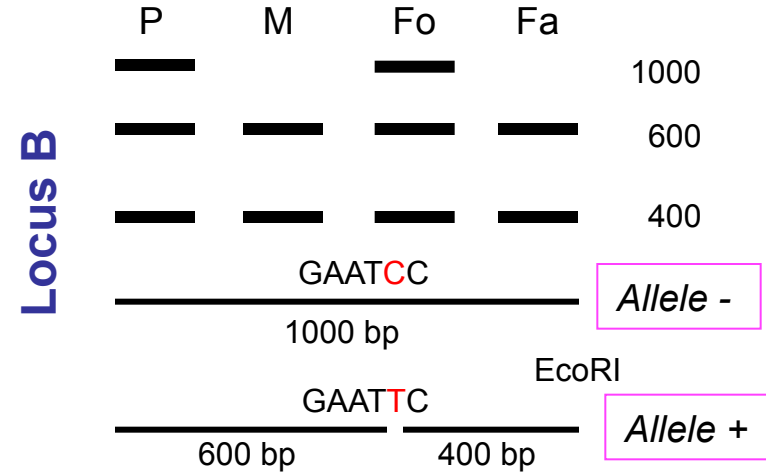
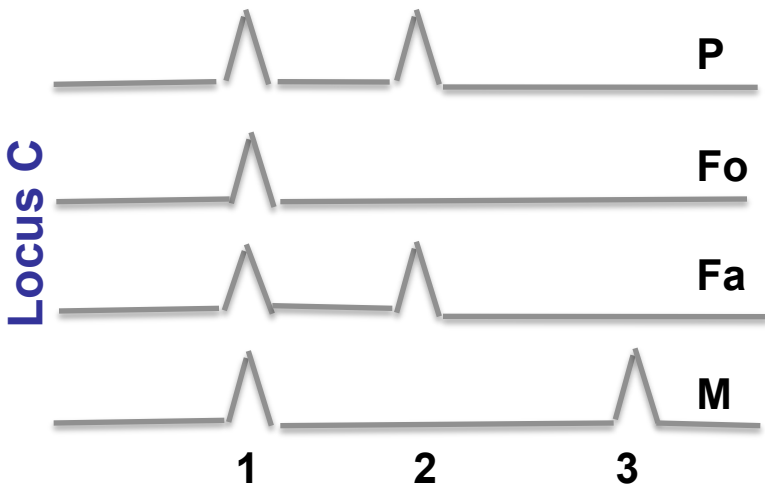
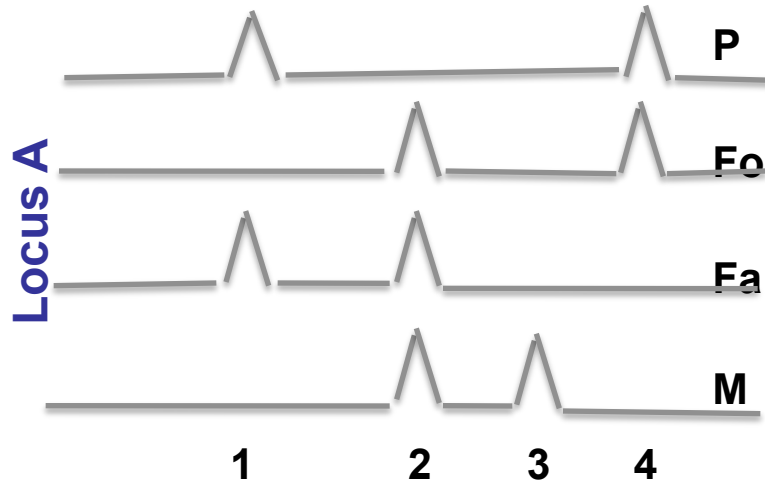
Dati i seguenti 4 loci, costruire l'aplotipo (vedi soluzione diapositiva successiva)

A

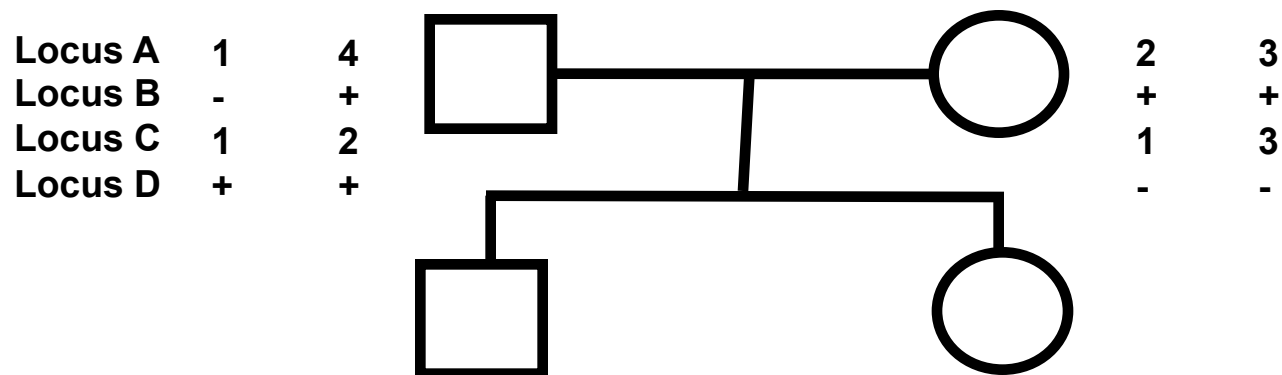
B

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	+ +	- -	+ -	+ -



Locus A
 Locus B
 Locus C
 Locus D

4
 -
 1
 +

2
 +
 1
 -

1
 +
 2
 +

2
 +
 1
 -

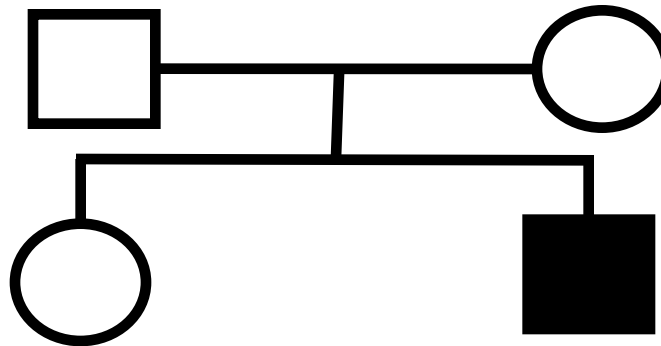
Aplotipo trasmesso dal padre

Aplotipo trasmesso dalla madre

Utilizzo loci polimorfici: identificazioni eventuali delezioni

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

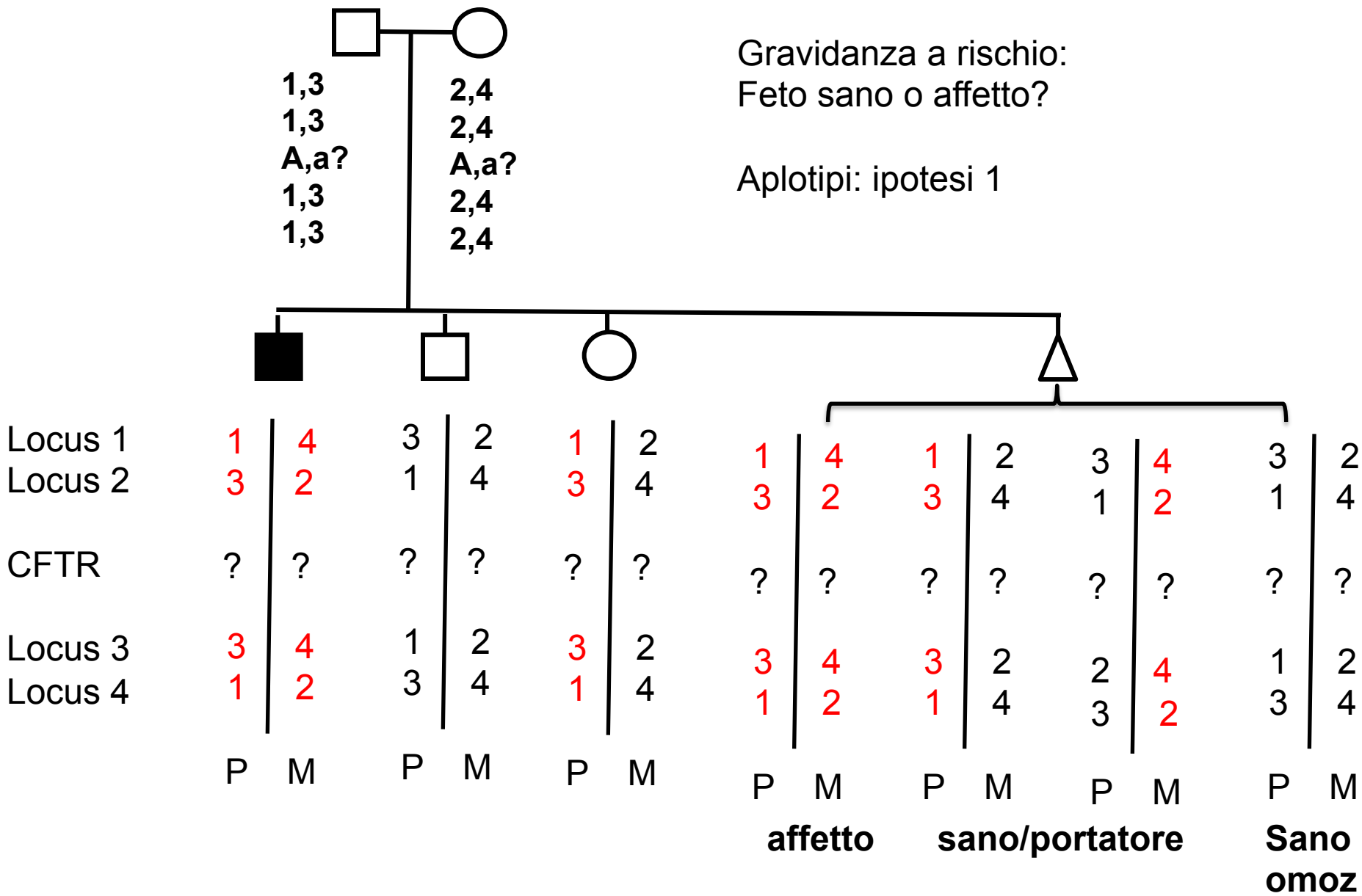
135	131
150	154
250	252
179	185
234	224
120	120
157	167

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

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

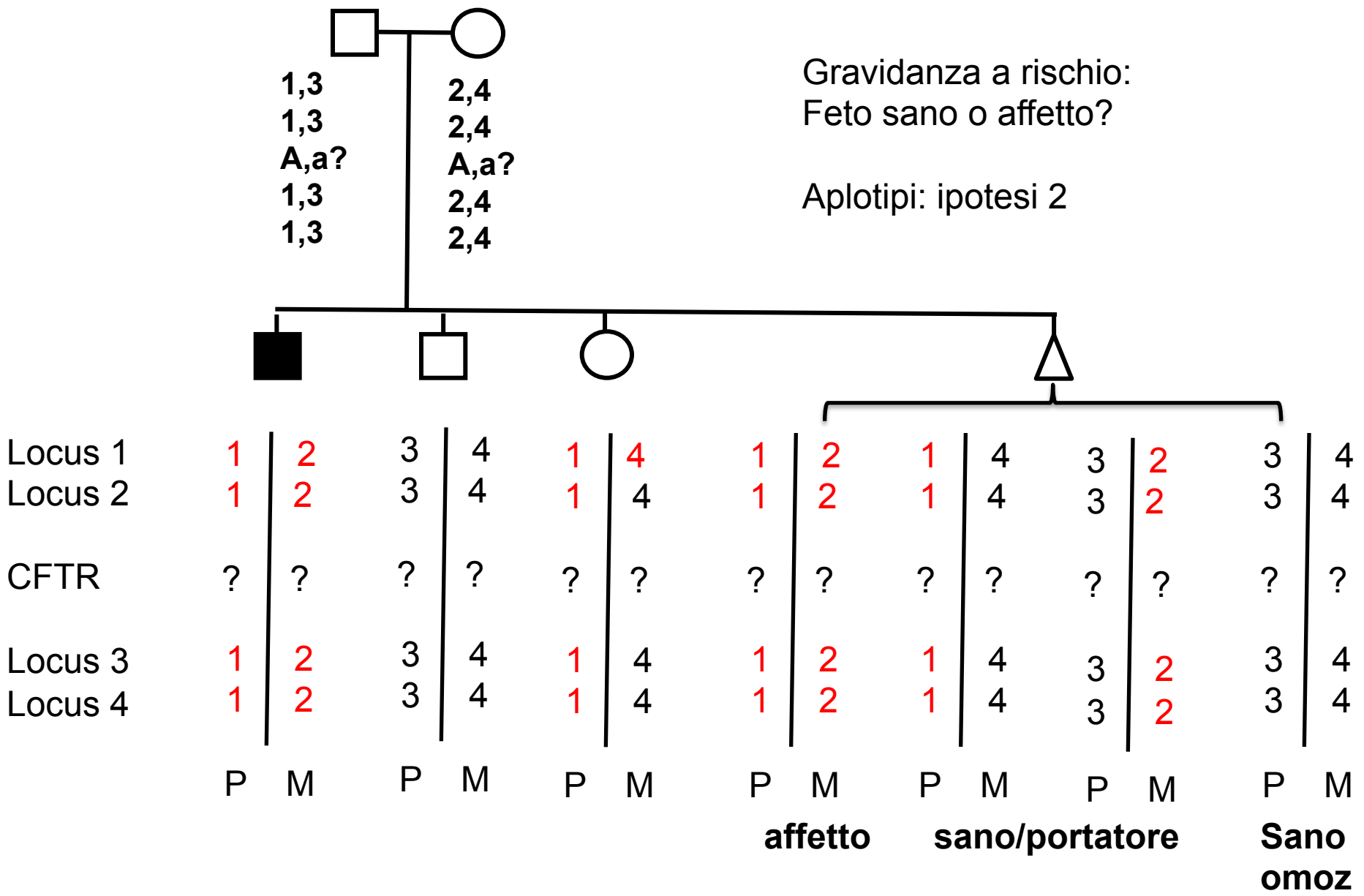
Utilizzo loci polimorfici:

Diagnosi indiretta (CF)/mutazioni non note

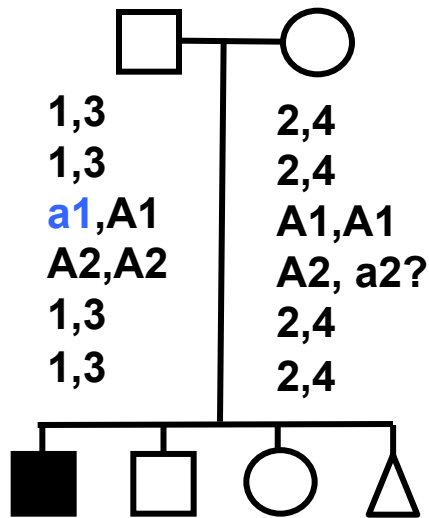


Utilizzo loci polimorfici:

Diagnosi indiretta (CF)/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 dei genitori
 - a) **4242 > affetto a_1a_2**
 - b) **2424 > portatore Aa_2**

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

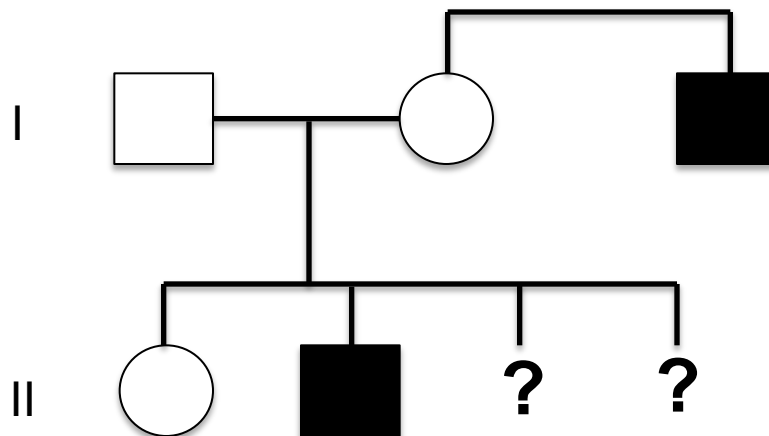
	1	4
	3	2
	a1	A1
	A2	a2?
	3	4
	1	2
	P	M
	affetto	

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

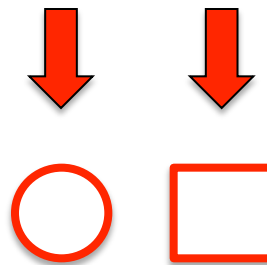
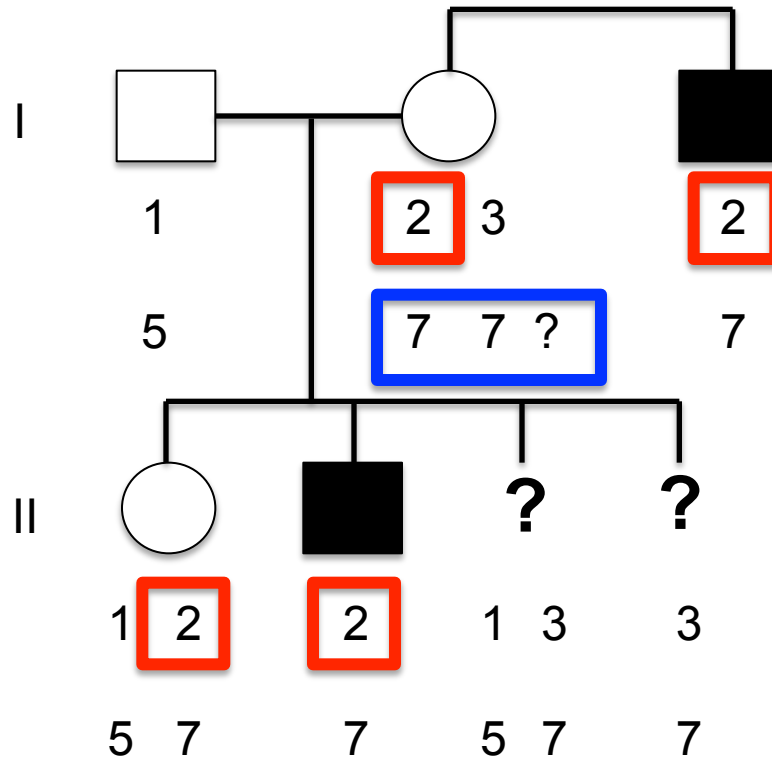
Data la seguente famiglia con malattia X-linked recessiva e i relativi genotipi di due loci molto vicini al gene che causa la malattia, stabilire:

- 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
LOCUS A	1	2,3	2	1,2	2	1,3	3
LOCUS B	5	7	7	5,7	7	5,7	7



	I-1	I-2	I-3	II-1	II-2	II-3	II-4
LOCUS A	1	2,3	2	1,2	2	1,3	3
LOCUS 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)	1 3	4 7	1 7	5 7
B (3q)	1 2	3 4	2 3	3 9
C (15p)	1 2	4 5	1 5	4 6

Fa non figlia di P