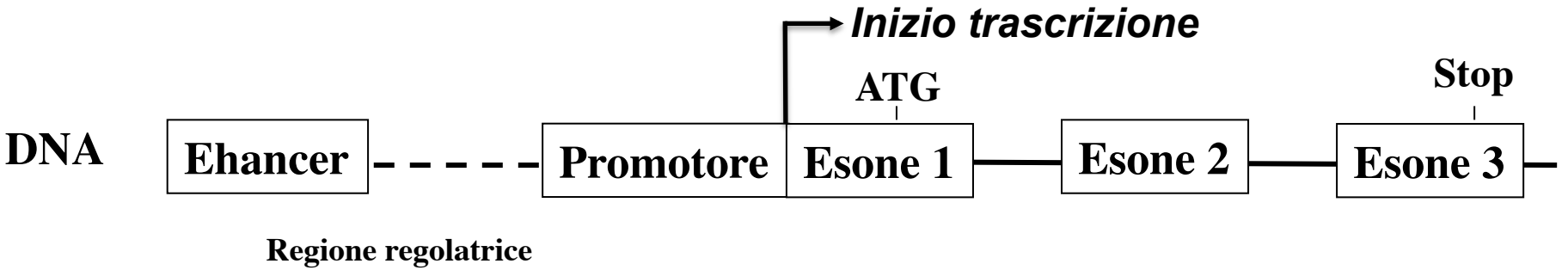
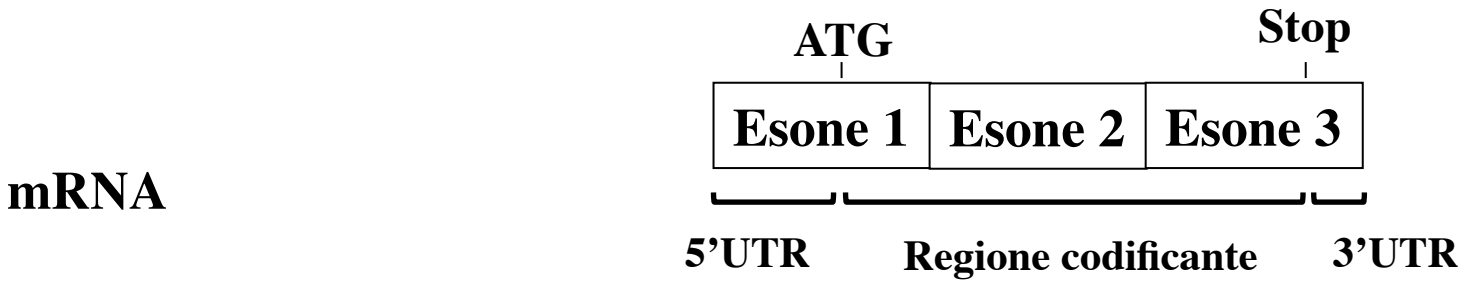
STRUTTURA GENOMICA DEL GENE

MUTAZIONI GENICHE


Inizio trascrizione
ATG
Stop
DNA
Ehancer
Promotore
Esone 1
Esone 2
Esone 3
Regione regolatrice


ATG
Stop
mRNA
Esone 1
Esone 2
Esone 3
5'UTR
Regione codificante
3'UTR

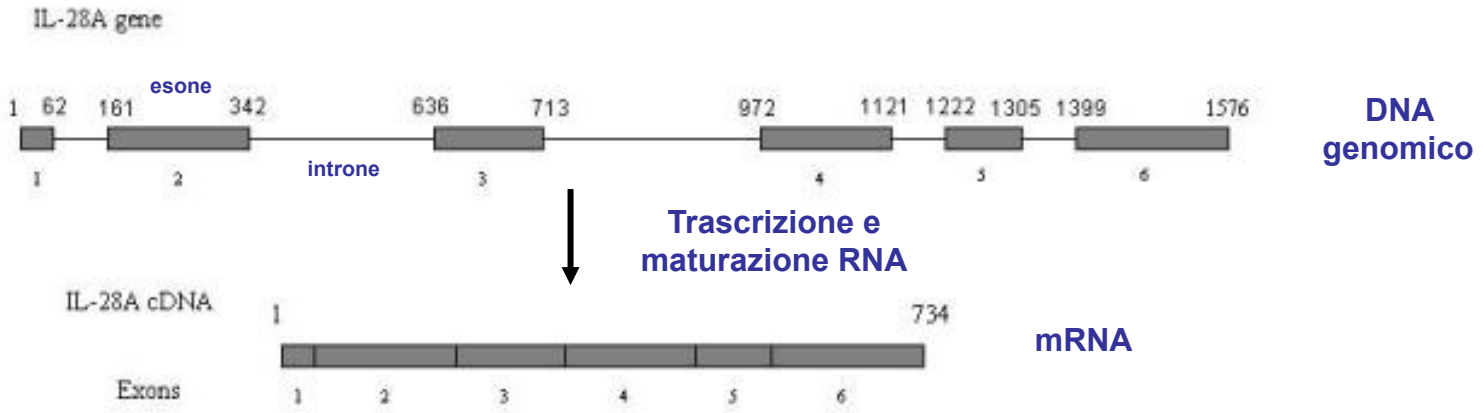

PROTEINA
NH2
Proteina
COOH


PROTEINA
Modificazioni posttraduzionali)
NH2
Proteina
COOH


FENOTIPO
Patologia

Struttura del gene

A)



B)

```
   1 tgggtgacag cctcagagtg tttcttctgc tgacaaagac cagagatcag gaatgaaact
  61 aggtgagtcc cacatctctg tccgtgctca gctcctgcag ccctgccct cagtgggcag
 121 cctctccatc ccctcagctc cctttctctc tgtgacacag acatgactgg ggactgcacg
 181 ccagtgctgg tgctgatggc cgcagtgctg accgtgactg gagcagttcc tgtcgccagg
 241 ctccacgggg ctctcccgga tgcaaggggc tgccacatag cccagttcaa gtccctgtct
 301 ccacaggagc tgcaggcctt taagagggcc aaagatgcct tagtgagtct ccccctgccc
 361 tcctgccatg gactagcctc caccccact ccaagcgtca ccatgctttc ccactcccag
 421 cttccttcac tgggctagcc tccaccctcc ctgcagtggg ctatctcatg ctcctactgt
 481 agggactgac tcatgttttc ctgtagaaga gggtcctcta ccatcctccc agcagttaac
 541 ctcccctatc ctgttgtcag ccatcctcca atcccaccag gatggtctaa cctccacccc
 601 tcctgctggg gctaacctgt gcctttgctg tctaggaaga gtcgcttctg ctgaaggact
 661 gcaggtgcca ctcccgcctc ttcccagga cctgggacct gaggcagctg caggtgagag
 721 ggggagtcag gcccacccct gctctcccag ccccactcac ctggctctgt agtggccct
 781 tcaccgtctc tttctccctt gtctctctct cttctcctca cacctgctct ccctccctc
 841 cgctcccacc tgaccacact ggctgtgccc tctcccctgt gcctgtcacc ttcacttgtt
 901 cctctctatc ctgctcccca acctgttccc ctcacctccc ccctcacctg ctctttctca
 961 cctctcctca ggtgagggag cgcccatgg ctttggaggc tgagctggcc ctgacgctga
1021 aggttctgga ggccaccgct gacactgacc cagccctggt ggacgtcttg gaccagcccc
1081 ttcacaccct gcaccatatc ctctcccagt tccgggcctg tgtgagtcgt tggggcctgg
1141 gcacccaggt ctgtgagctc tgagcagcgt ccttcccctt gccaaggccc cggctcacac
1201 accgccctcc tctgcccaca gatccagcct cagcccacgg cagggcccag gacccggggc
1261 cgcctccacc attggctgta ccggctccag gaggccccaa aaaaggtgag tgacccggga
1321 agagagggac tgaggtctgg ggagccactg ggagcccaga acccagacag cccctgaccc
1381 atccctctct ccctacagga gtcccctggc tgcctcgagg cctctgtcac cttcaacctc
1441 ttccgcctcc tcacgcgaga cctgaattgt gttgccagtg gggacctgtg tgttttgccc
1501 tcccaccagt catgcaacct gagattttat ttataaatta gccacttgtc ttaatttatt
1561 gccacccagt cgctat
```

Mutazioni geniche

(riferite ad alterazioni di pochi nucleotidi)

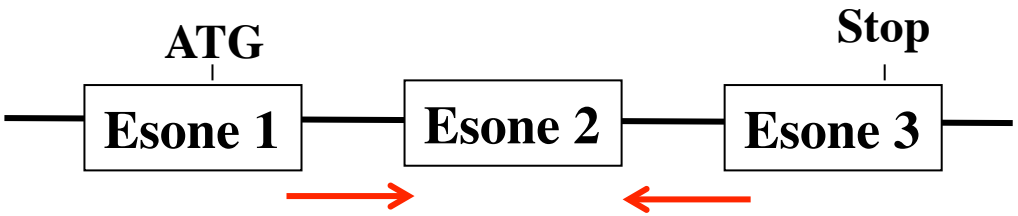
- Alterazioni possono colpire ovunque (regioni regolatrici, esoni, introni, ecc.) un gene

- 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

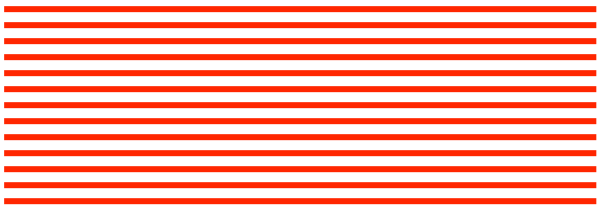
- Valutazione dell'effetto (meccanismi di patogenicità)

Screening di mutazioni

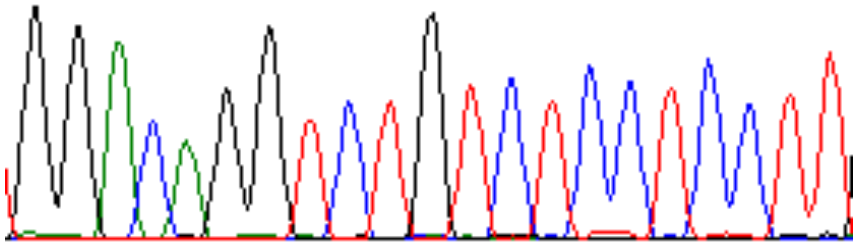


PCR

Amplificazione di DNA
(sintesi in vitro di molecole del frammento prescelto)



Sequenziamento Sanger



Sanger sequencing vs Next Generation Sequencing (NGS)

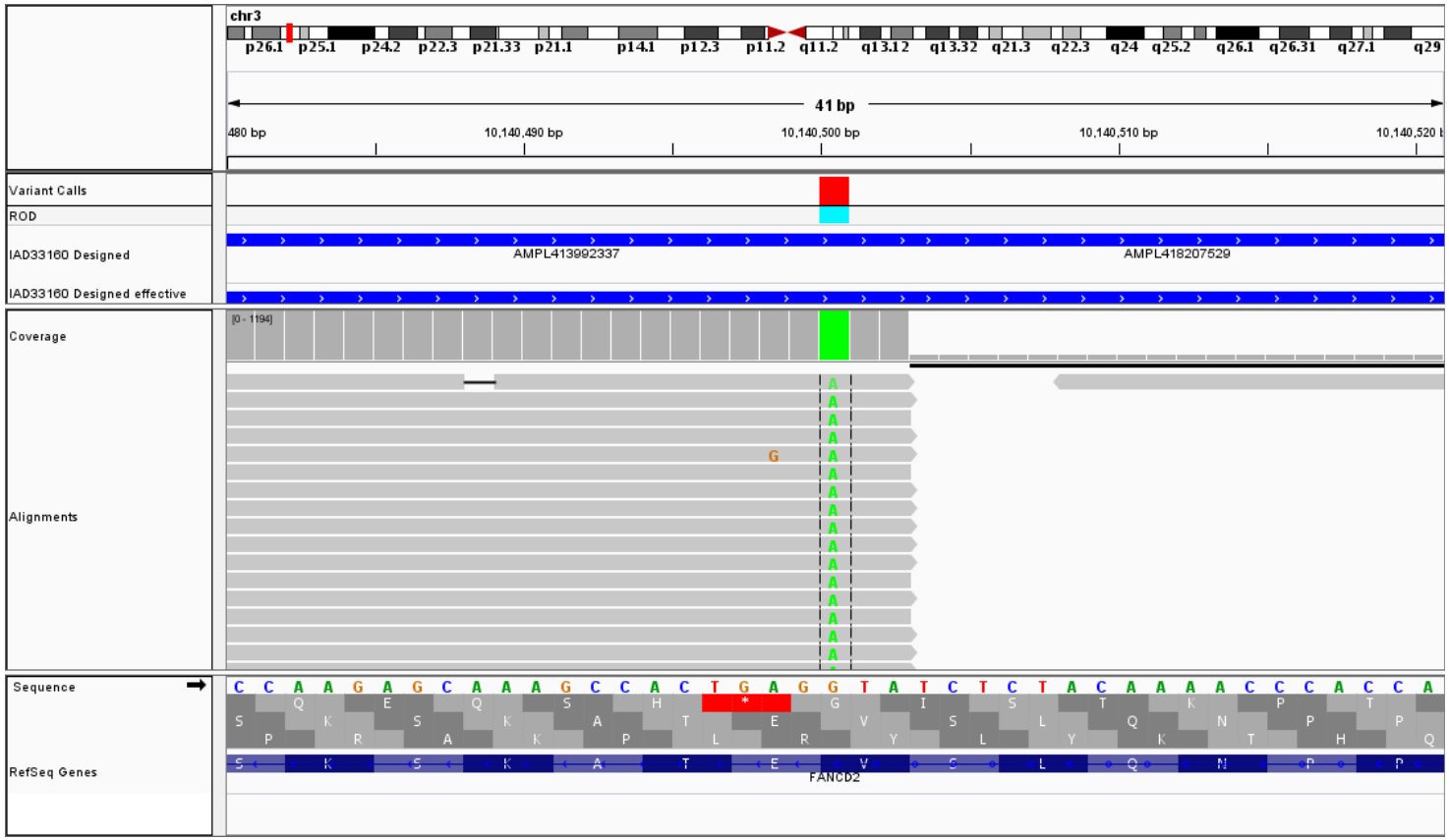
Sanger	NGS (high-throughput sequencing)*
One fragment	Up to entire genome
500-1000 bp	$3x10^9$ bp

**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

NGS: example 1

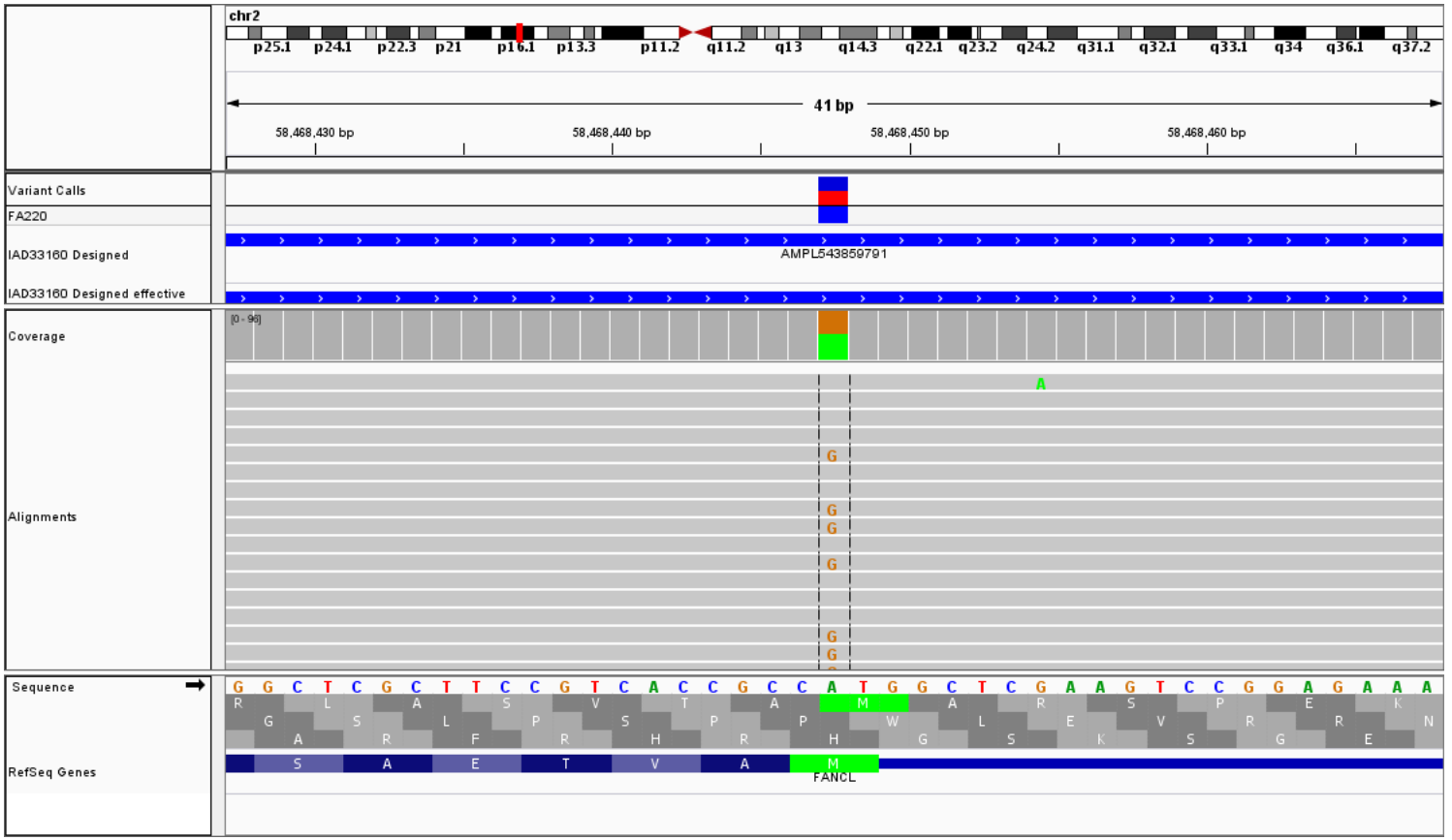
Position	Ref	Variant	Allele Call	Frequenc...
chr3:10140500	G	A	Homozygous	100.0 %



GENE: FANCD2 (NM_001018115.1)
INTRON: 43
VARIANT: c.4281+1G>A
STATUS: Homozygous

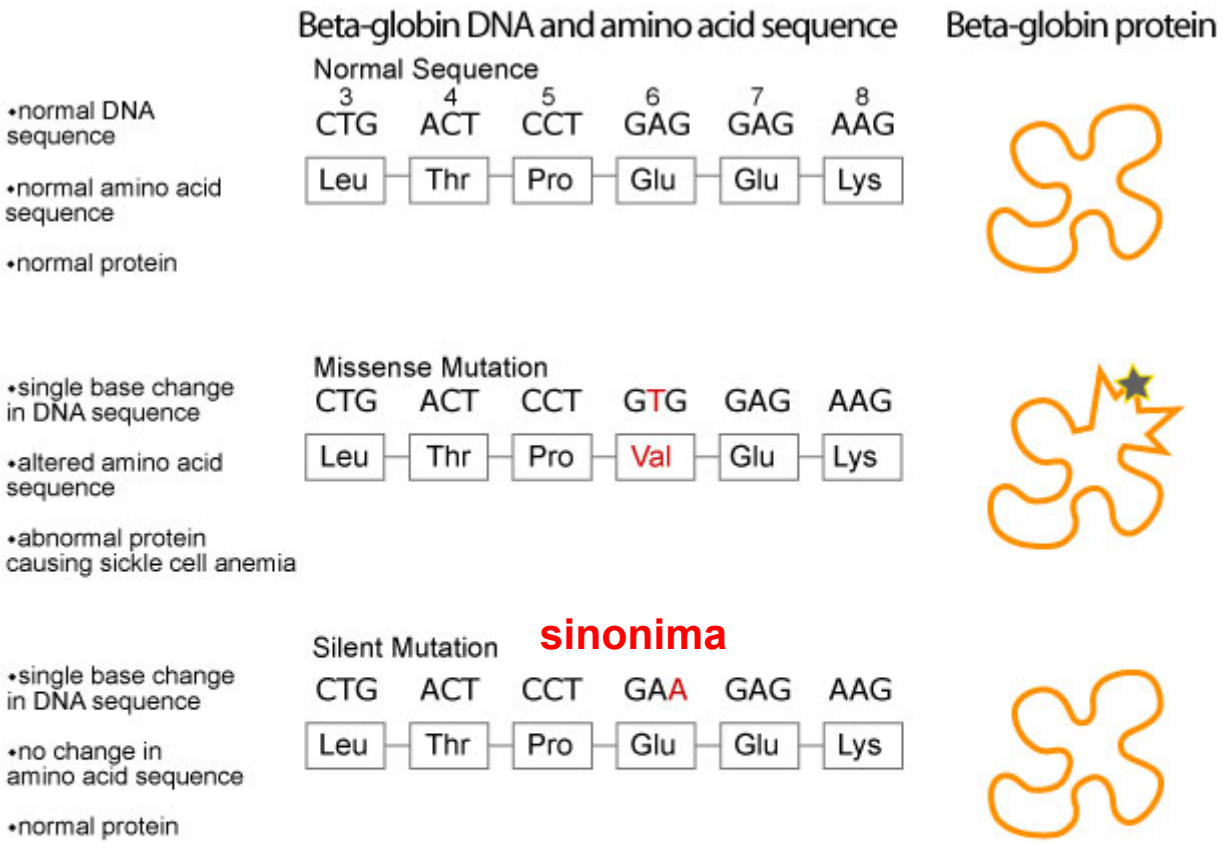
NGS: example 2

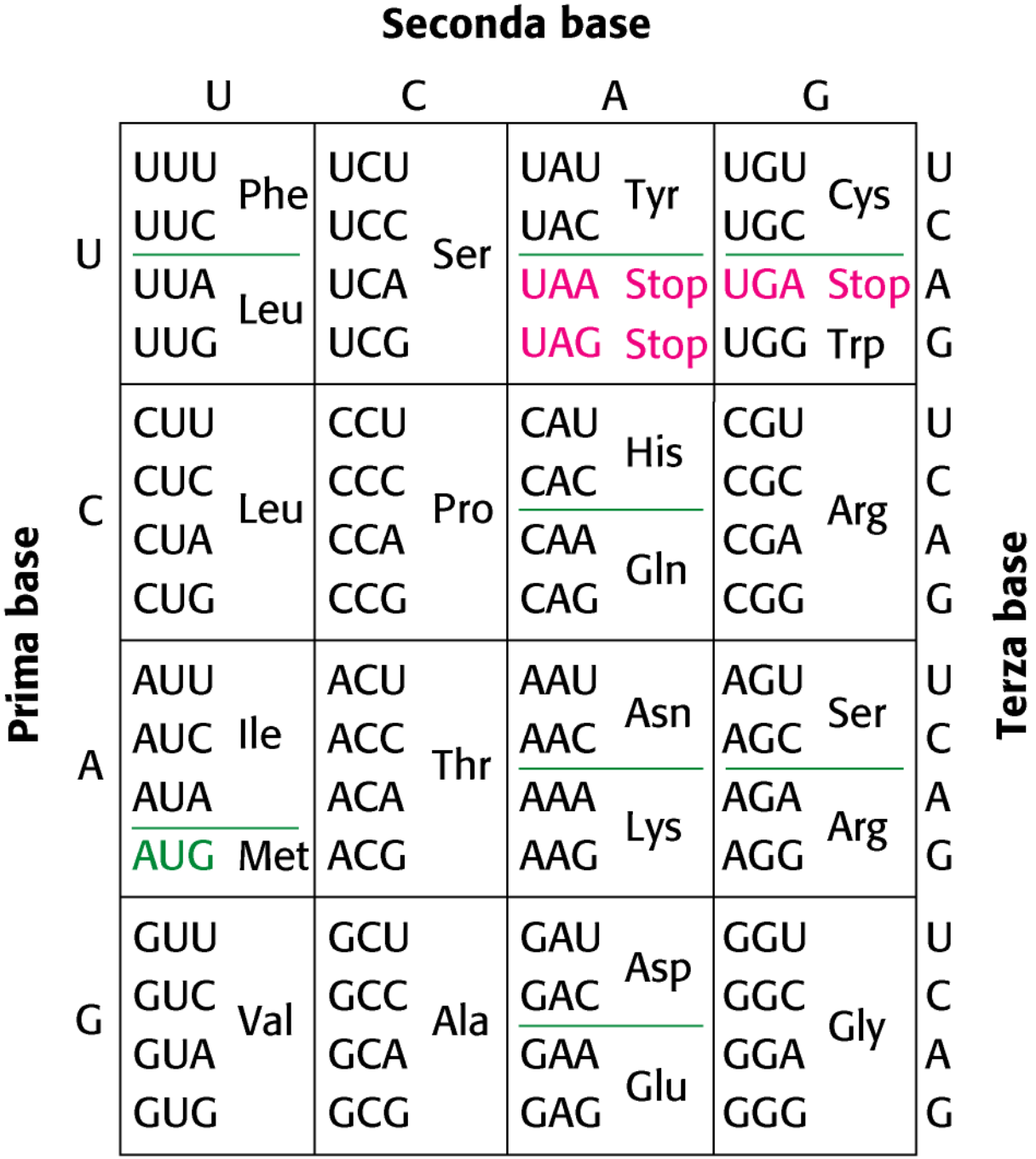
Position	Ref	Variant	Allele Call	Frequenc...
chr2:58468447	A	G	Heterozygous	47.9 %



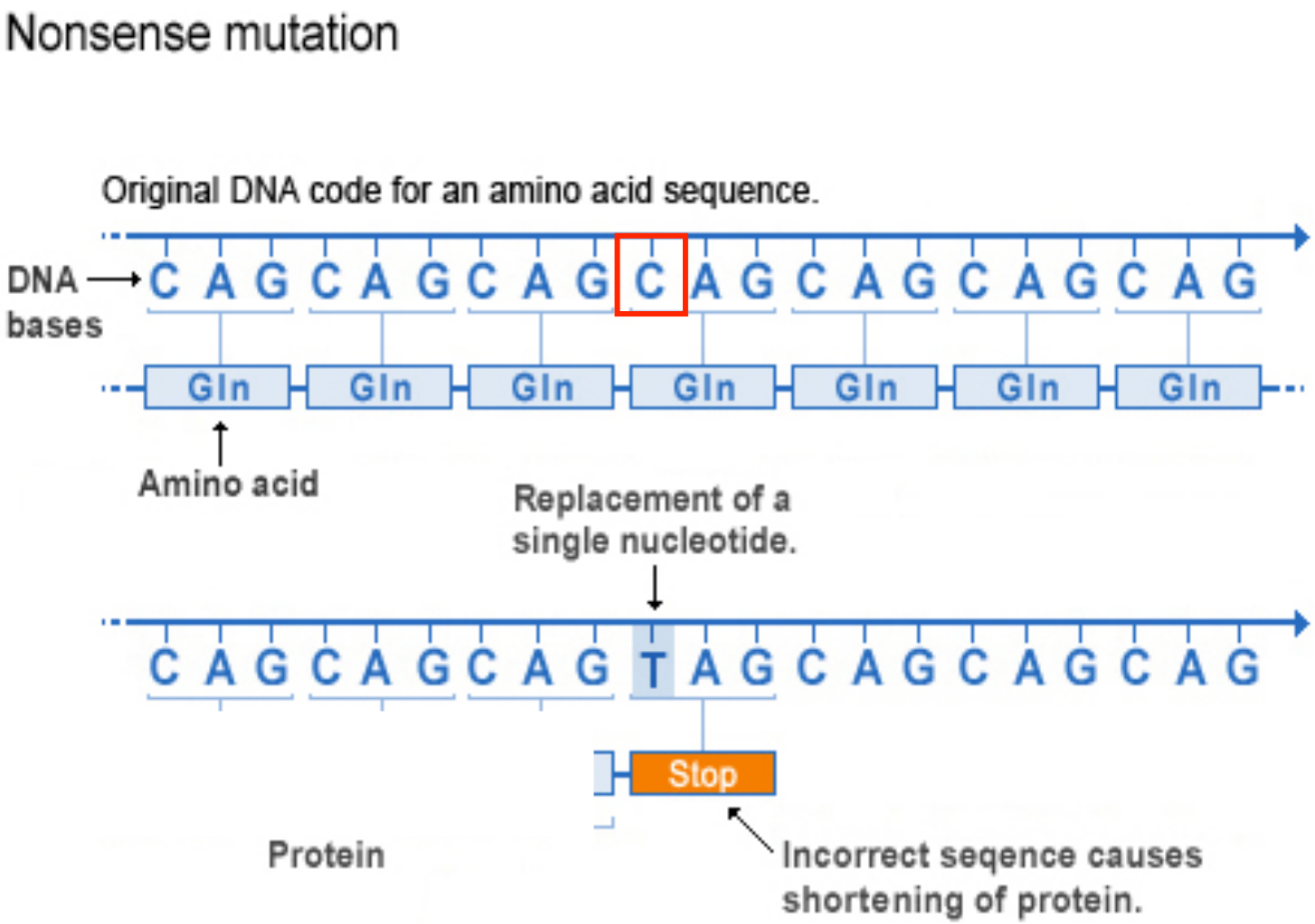
GENE: FANCL (NM_018062.3)
EXON: 1
VARIANT: c.2T>C
STATUS: Heterozygous

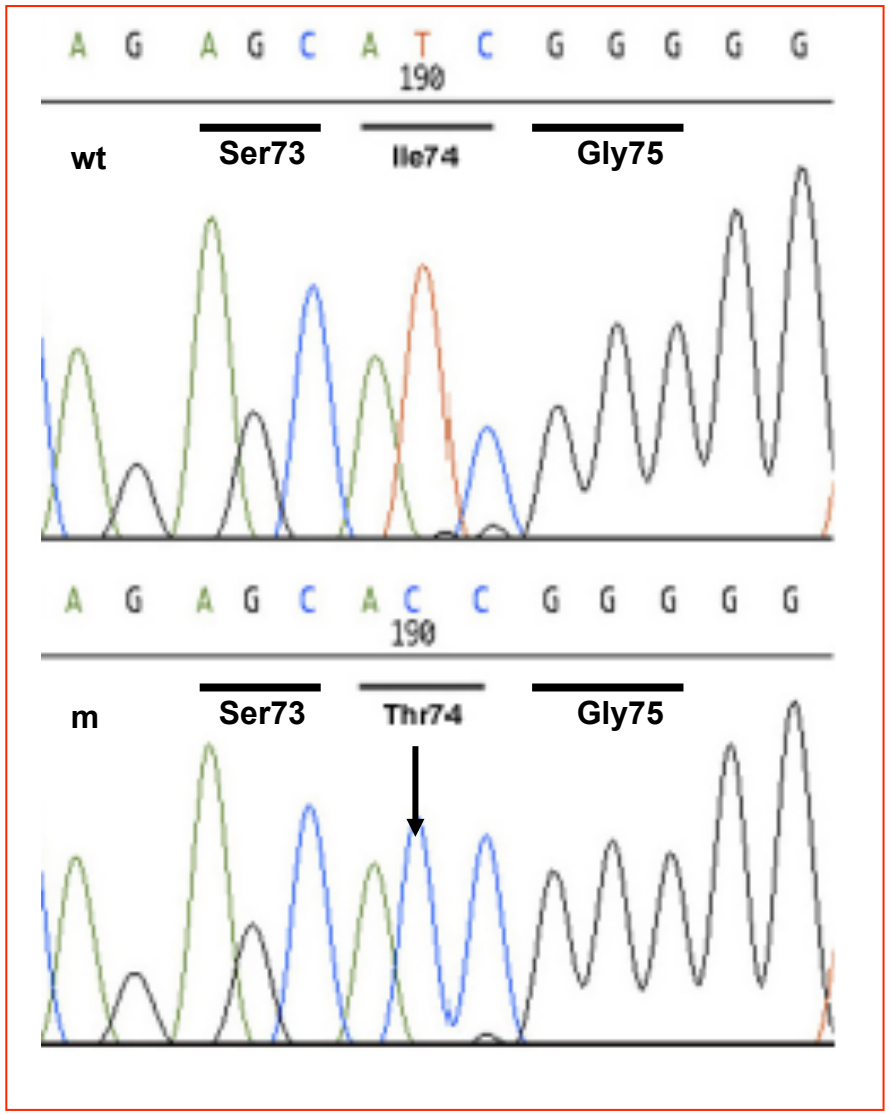
Nucleotide substitutions



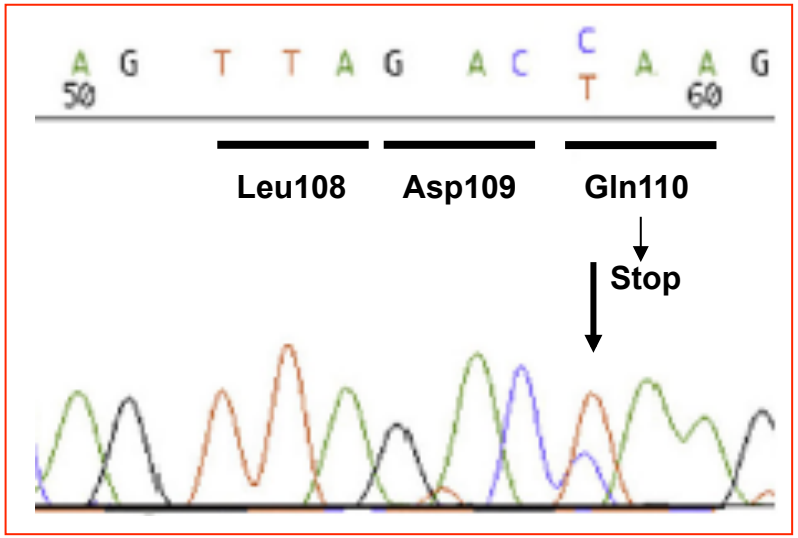

Seconda base
U
C
A
G
Prima base
Terza base
U
C
A
G
UUU Phe
UUC
UUA Leu
UUG
UCU
UCC
UCA
UCG Ser
UAU Tyr
UAC
UAA Stop
UAG Stop
UGU Cys
UGC
UGA Stop
UGG Trp
CUU
CUC
CUA
CUG Leu
CCU
CCC
CCA
CCG Pro
CAU His
CAC
CAA
CAG Gln
CGU
CGC
CGA
CGG Arg
AUU
AUC
AUA Ile
AUG Met
ACU
ACC
ACA
ACG Thr
AAU Asn
AAC
AAA
AAG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

Nucleotide substitutions





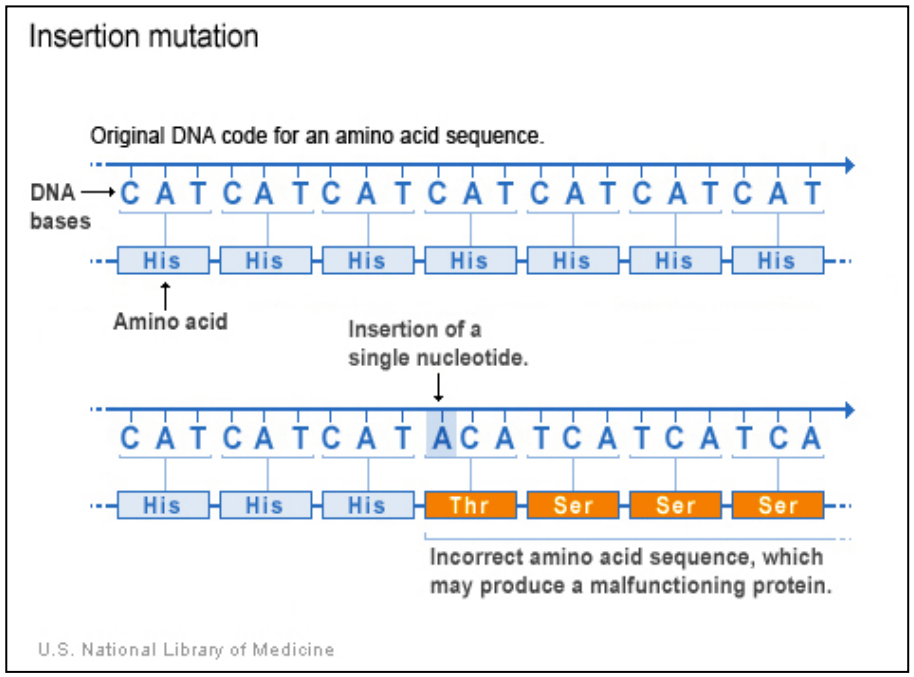
p.Ile74Thr

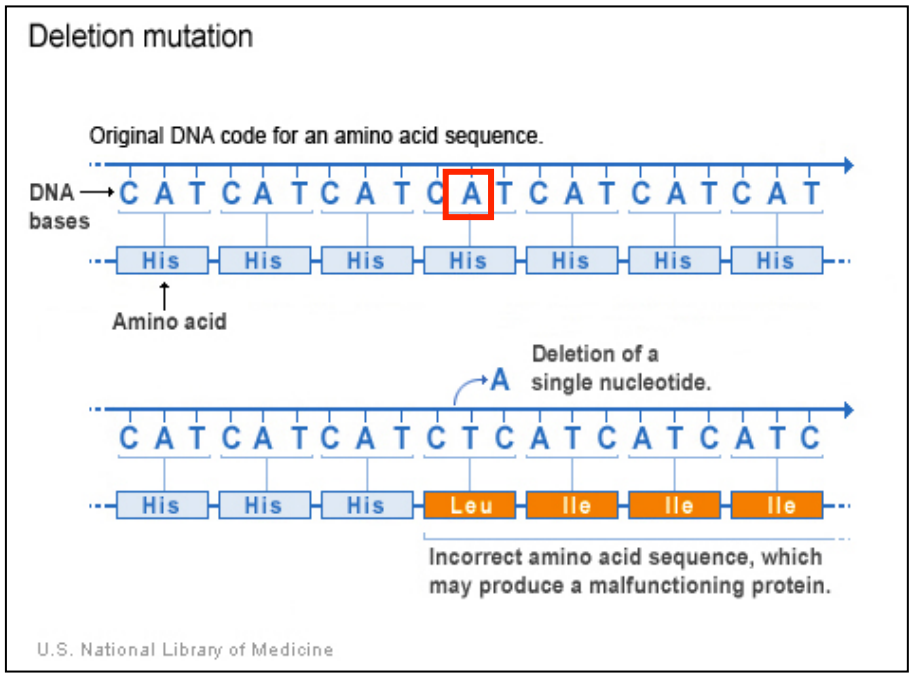


p.Q110*

Nucleotide deletion/duplication/insertion

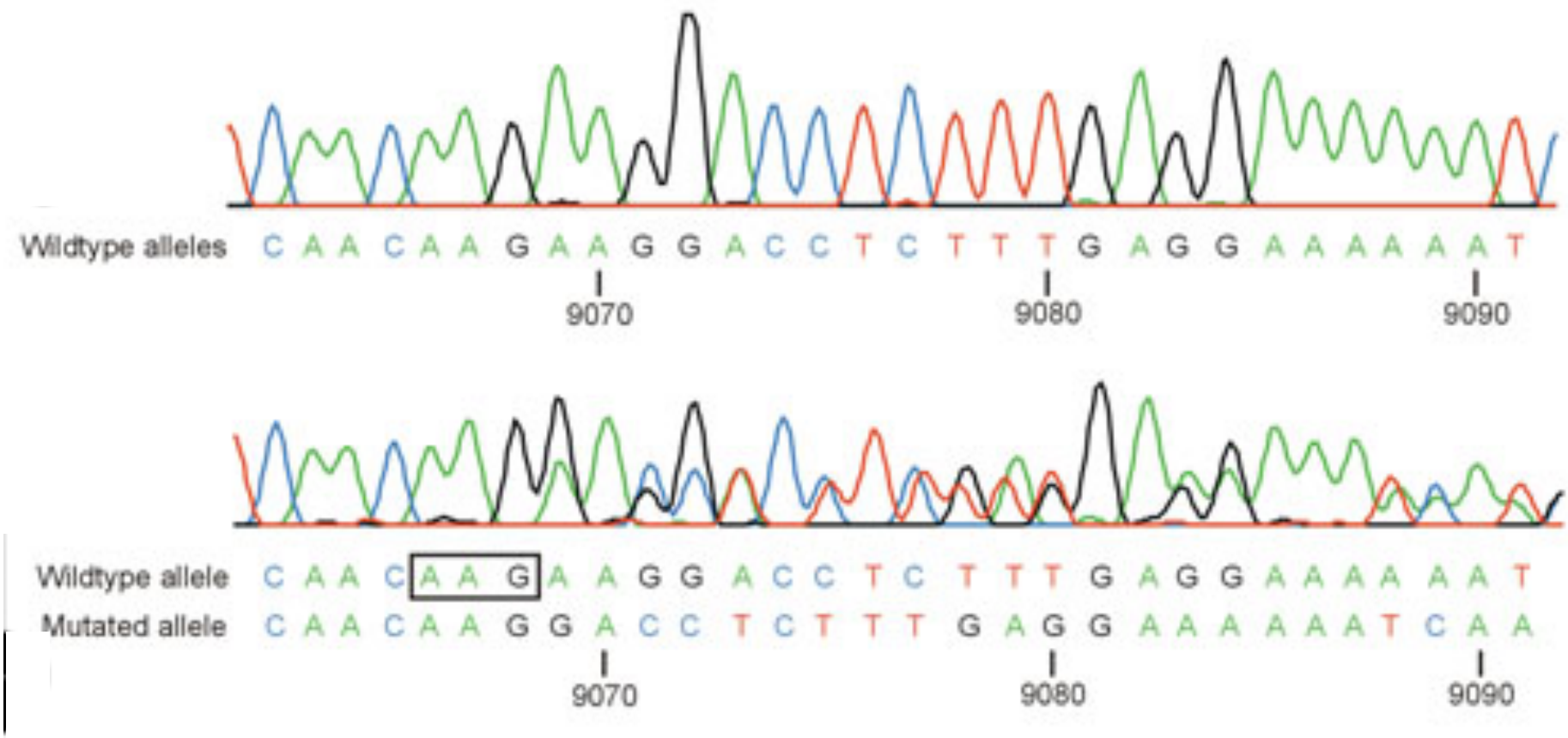
1) Frameshift mutations



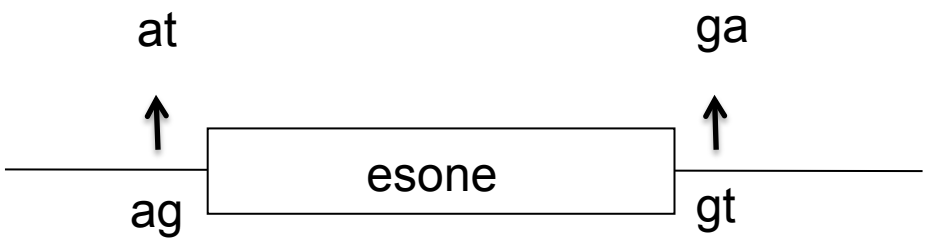


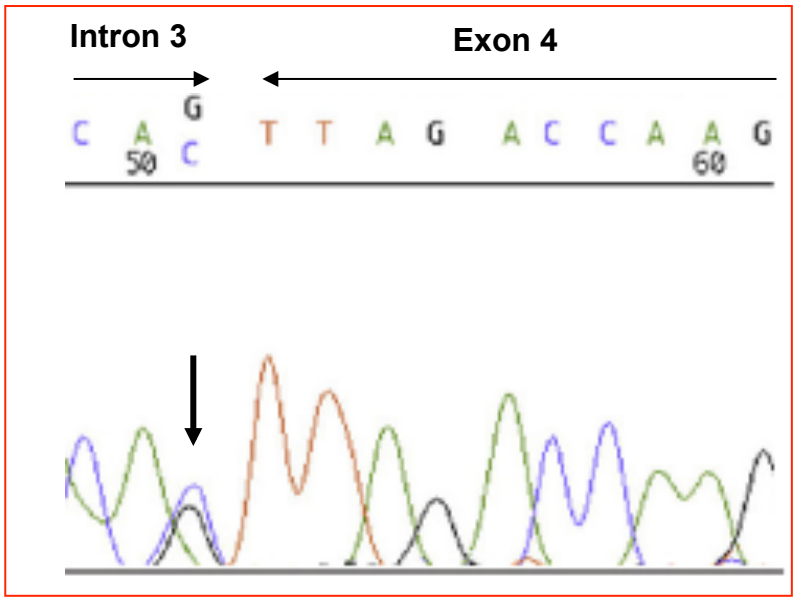
2) In-frame mutations

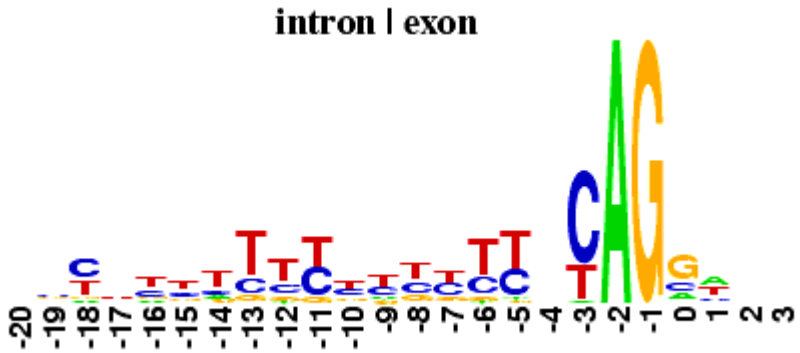
DELEZIONE ETEROZIGOTE di tre nucleotidi - AAG

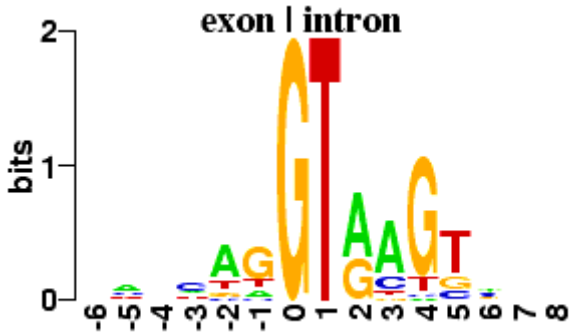


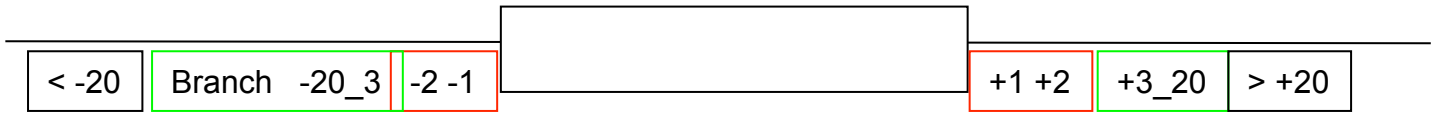
Splicing mutations



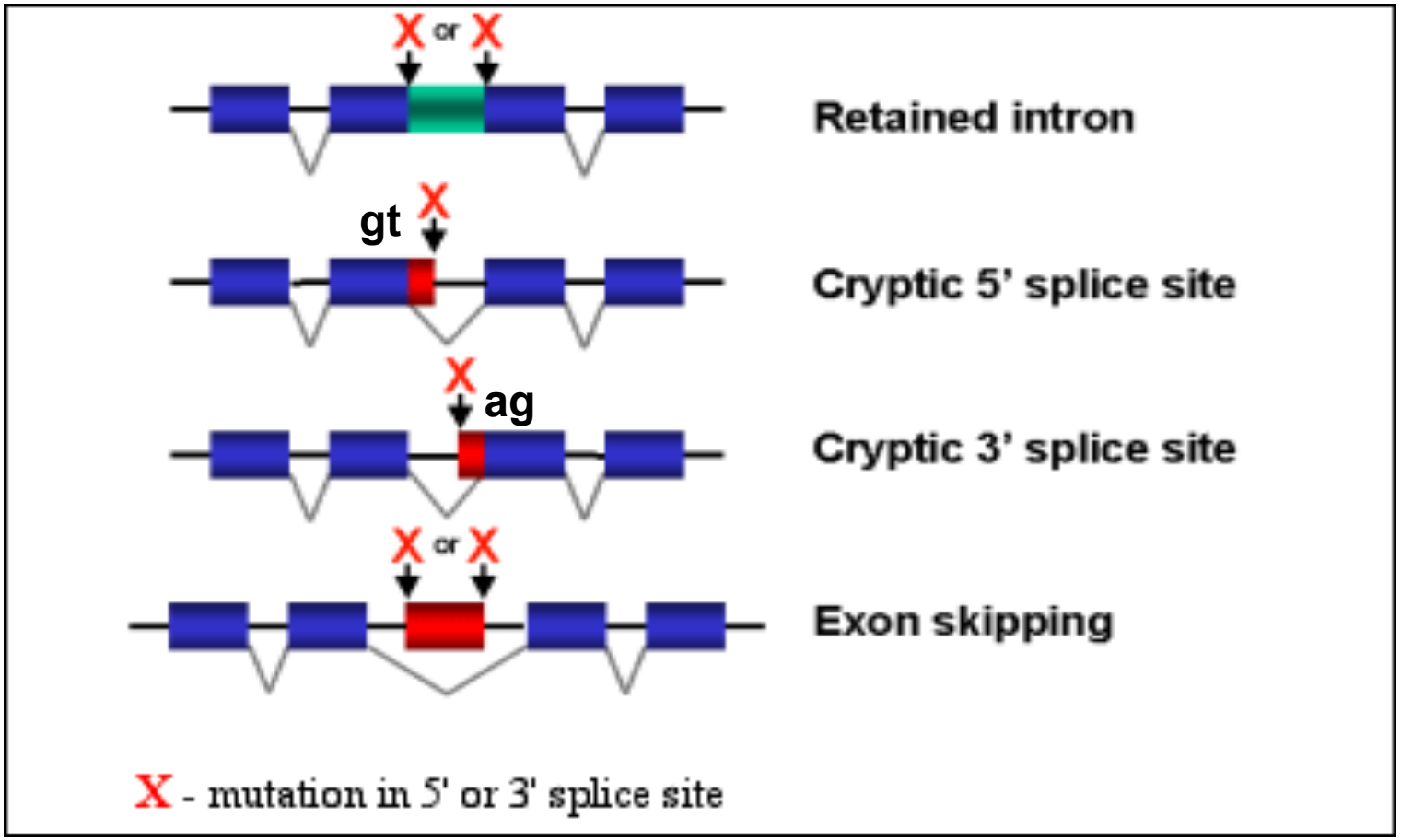








Conseguenze mutazione nei siti di splicing

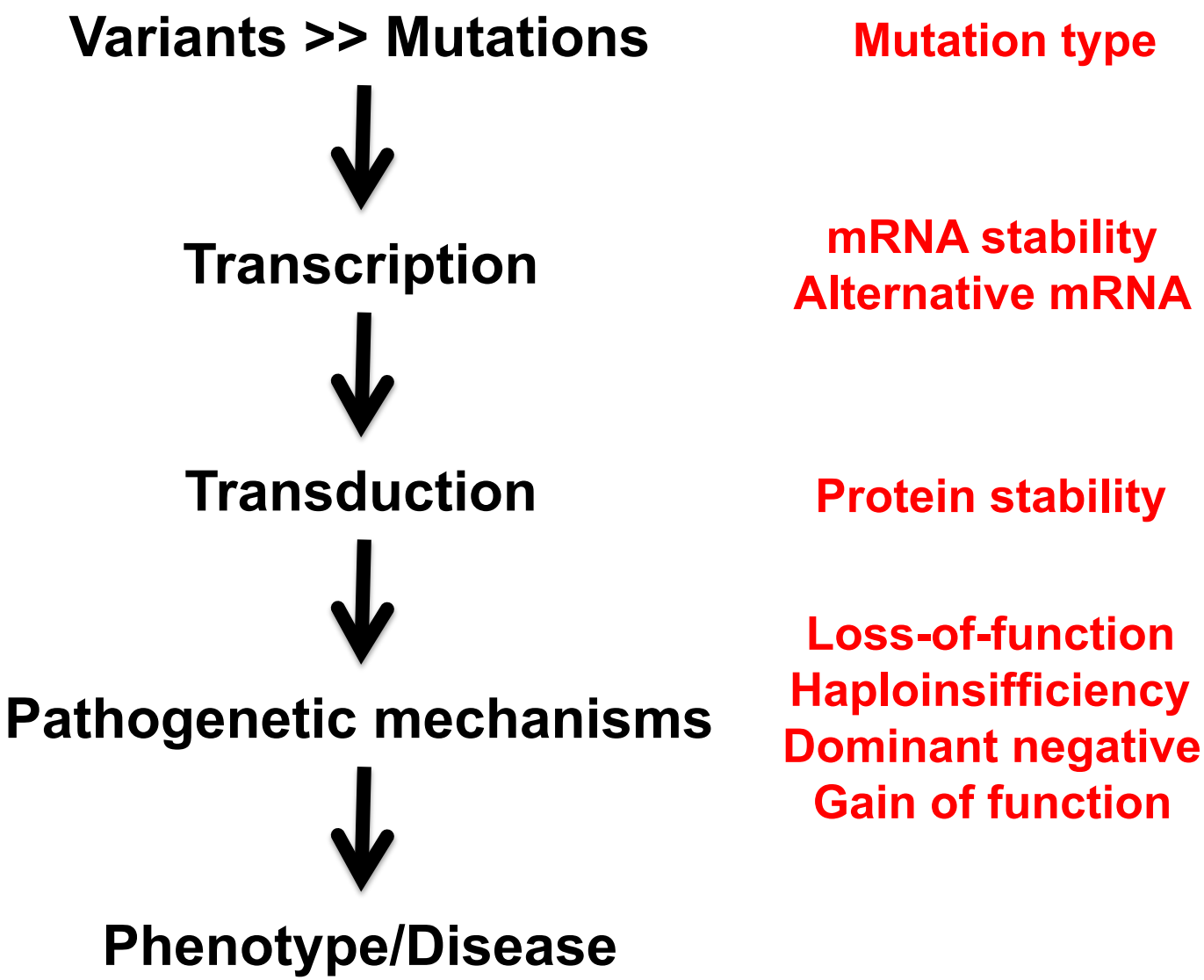


Splicing mutation → **Effetto mRNA** → **Frameshit In-frame**

Classifiction of variants (point)

DNA		Protein	
		Prediction	Effect on protein function
Coding region	Nucleotide substitution	Synonim	VUS (var of unknown significance)
		missense	VUS
		nonsense	Deleterius
	Deletion Insertion duplication	In-frame (N=3n)	Likely pathogenetic
		Frameshift (N≠3n)	Deleterius
Uncoding (coding) region	As above (Splicing sites)	unknown	unknown
	As above (Regulatory 5'/3'-UTR)	unknown	unknown

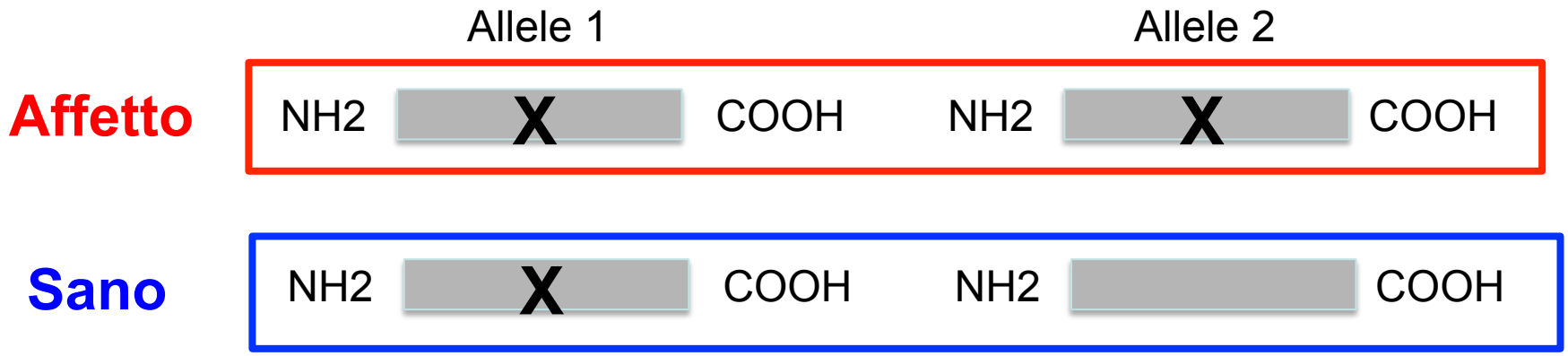
Consequences of mutations



Classificazione delle mutazioni in base all'effetto sulla funzione

Malattie ricessive

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



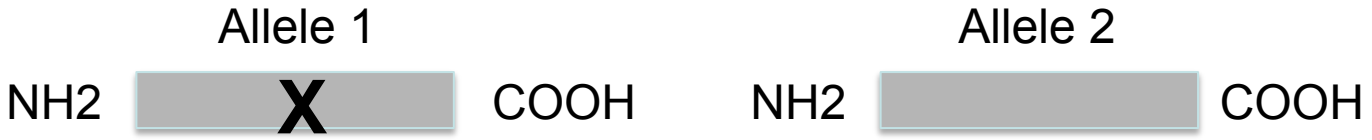
Malattie dominanti

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

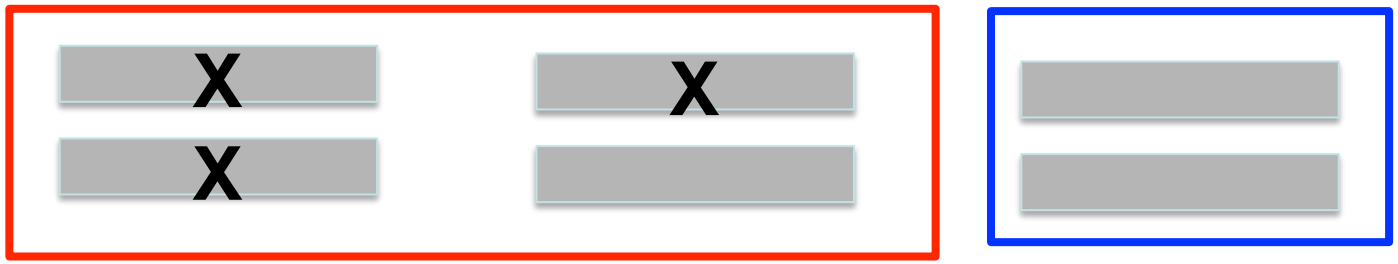
Classificazione delle mutazioni in base all'effetto sulla funzione

Malattie dominanti

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



Molecola matura formata da dimeri nelle seguenti combinazioni

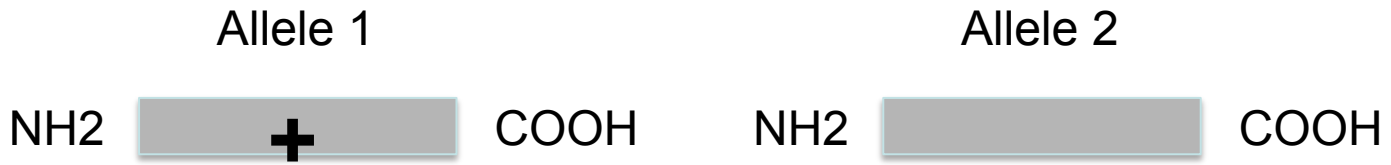


Prodotto non attivo

Prodotto attivo
Insufficiente per garantire funzione

4) Gain of function

- Aumenta l'attività funzionale
- Nuova funzione della proteina



Types of mutations (large)
DNA level

- Deletions of the entire gene
- Gene disruption for chromosomal rearrangement
- Intragenic, large (one o ore exons) deletions or duplications
- Nonsense, frameshift

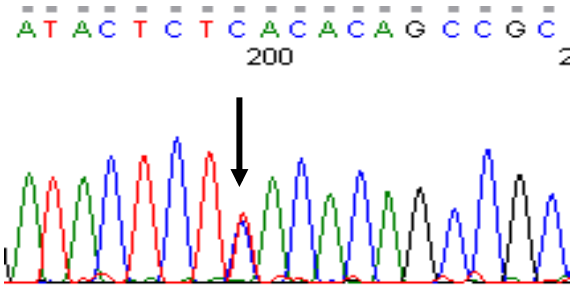
Deleterious
Pathogenic

- Duplications of the entire gene

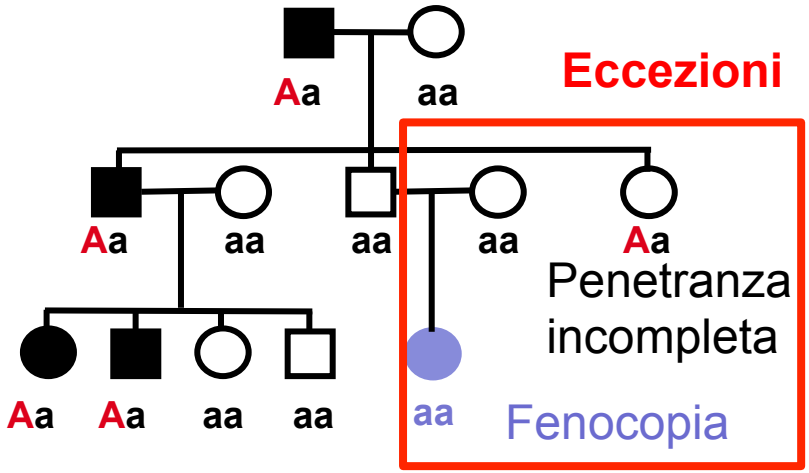
Quantitative
Gain of function?

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

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



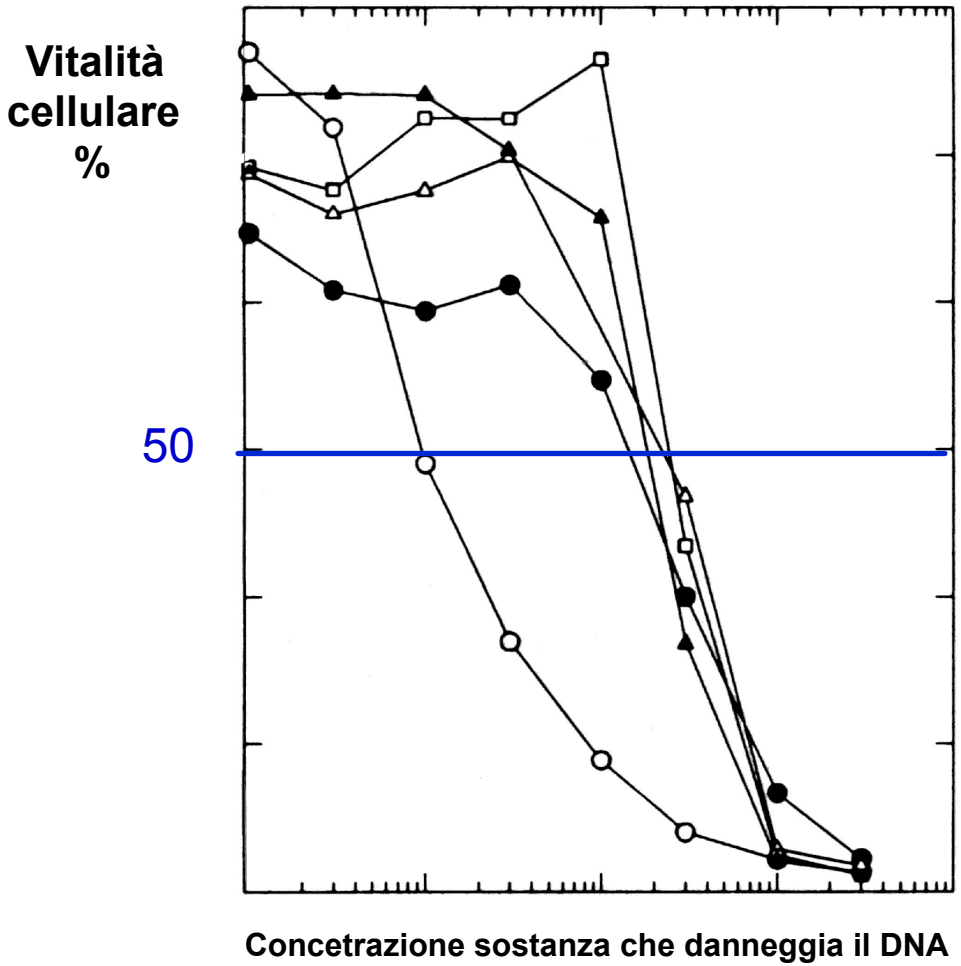
Segregazione nella famiglia

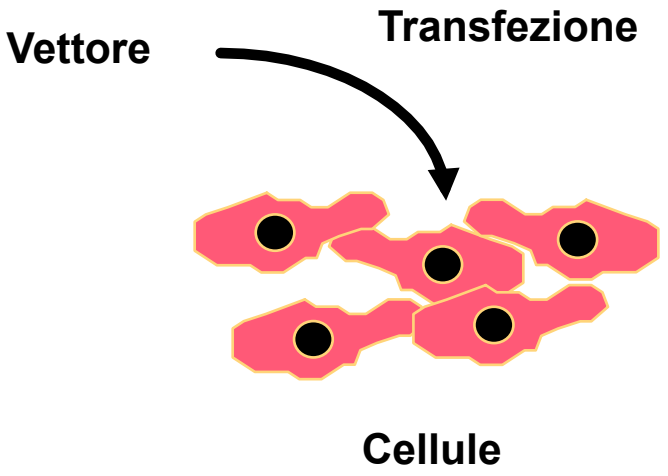


Conservazione amminoacido

									p.Tyr49His	
Homo sapiens	T	G	Q	A	P	G	Y	S	Y	T
Pab troglodytes	T	G	Q	A	P	G	Y	S	Y	T
Mus musculus	T	G	Q	A	P	G	F	S	Y	T
S. cerevisiae	S	G	Q	V	K	G	Y	S	Y	T
Magnaporthe grisea	T	G	S	V	D	G	Y	A	Y	T
Arabidosis thaliana	S	G	T	T	P	G	Y	S	Y	S
P. falciparum	S	G	D	S	D	_	F	P	Y	S

SAGGIO FUNZIONALE - MODELLO ANEMIA DI FANCONI (FA)





Cellule wt + vettore vuoto

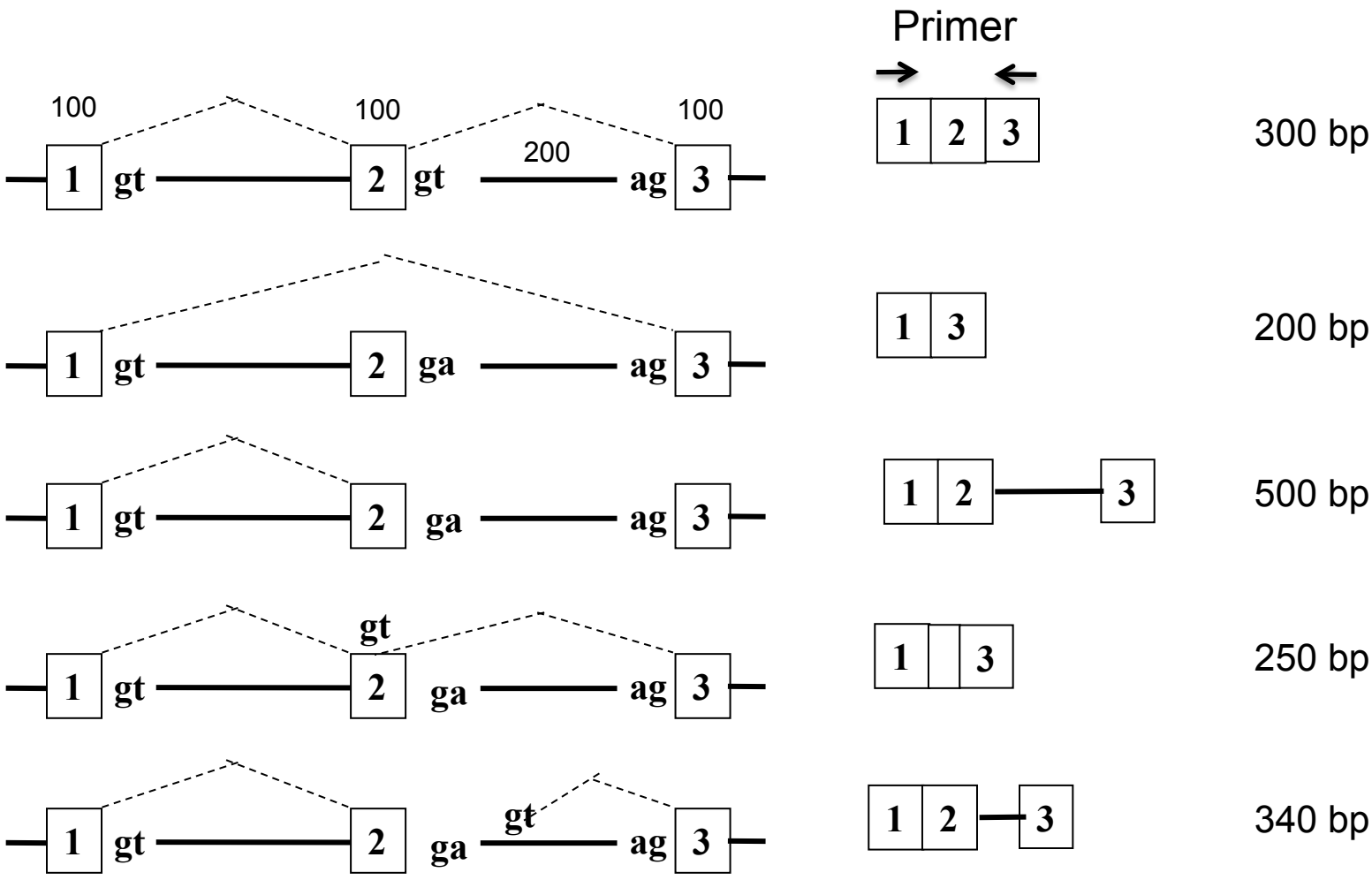
Cellule FA + vettore vuoto

Cellule FA + cDNA wt del gene

Cellule FA + cDNA mutato

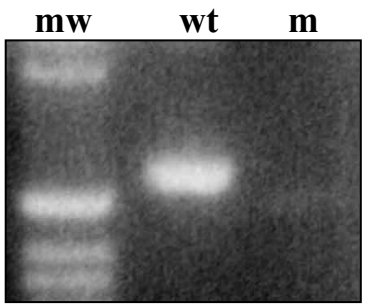
Cellule FA + cDNA mutato

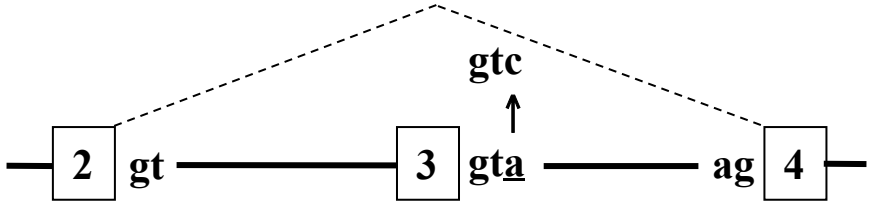
RT-PCR to identify effects of splicing mutations



Effetti mutazioni di splicing: esempi di RT-PCR su cDNA del gene FANCA

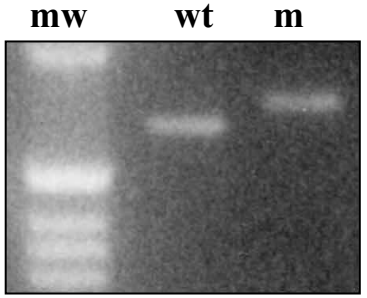
Mutazione in omozigosi





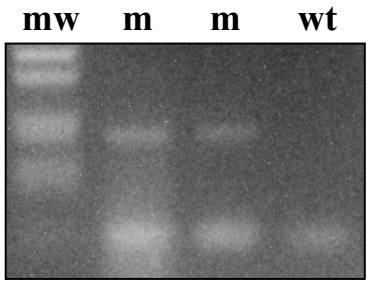
Exon skipping: banda mutata (m) 487 bp
banda wt 581 bp

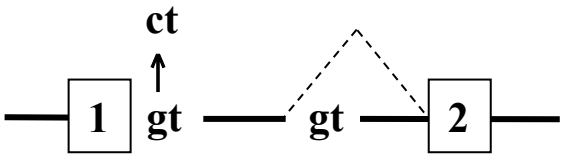
Mutazione in omozigosi



Ritenzione introne: banda mutata (m) 733 bp
banda wt 733 bp

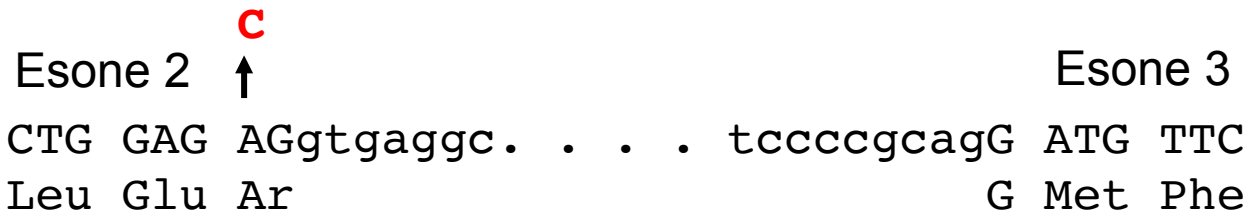
Mutazione in eterozigosi



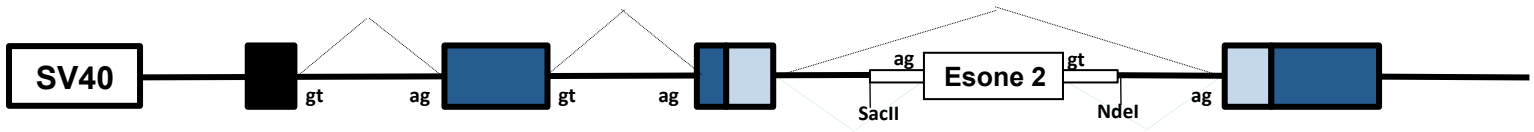


Riconoscimento sito criptico di splicing:
banda mutata (m) 199 bp, banda wt 70 bp

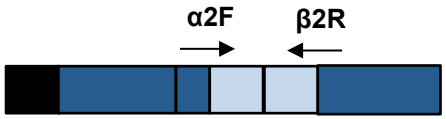
Sistema del minigene per valutare gli effetti delle mutazioni sullo splicing



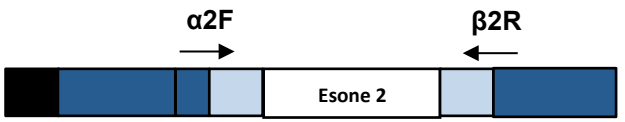
Clonaggio esone 2 nel sistema minigene



Prodotto di splicing vettore controllo senza esone 2



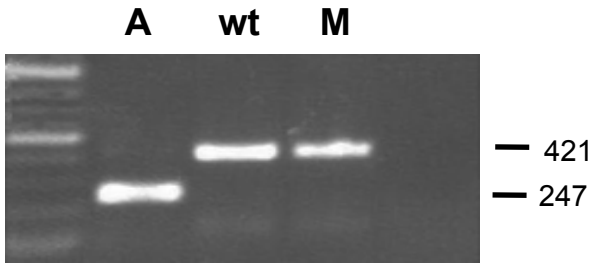
247 bp



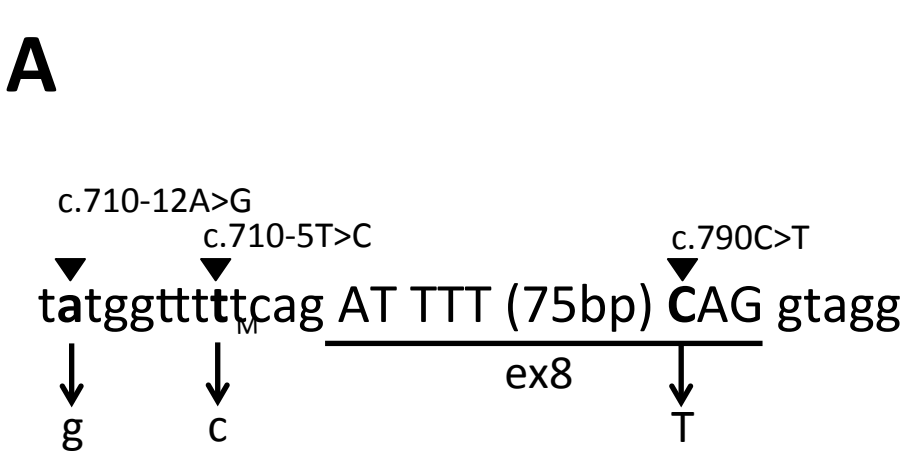
421 bp

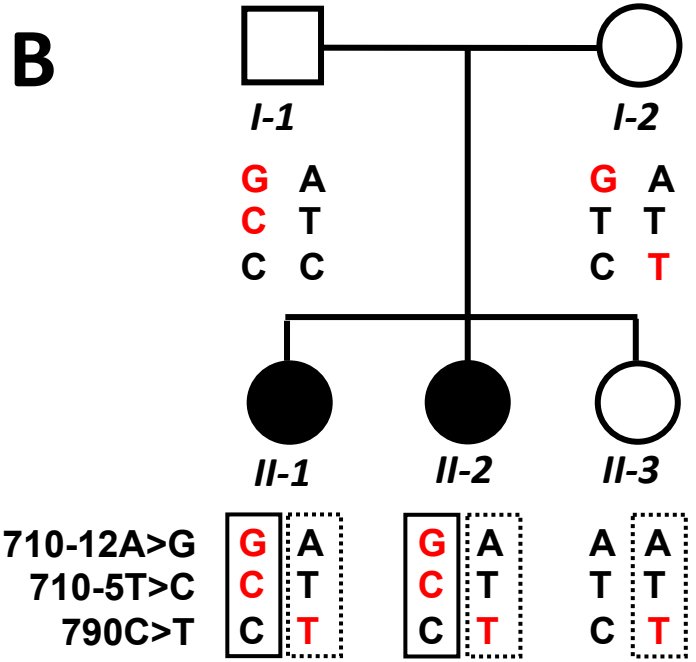
Prodotto di splicing vettore con esone 2 wt

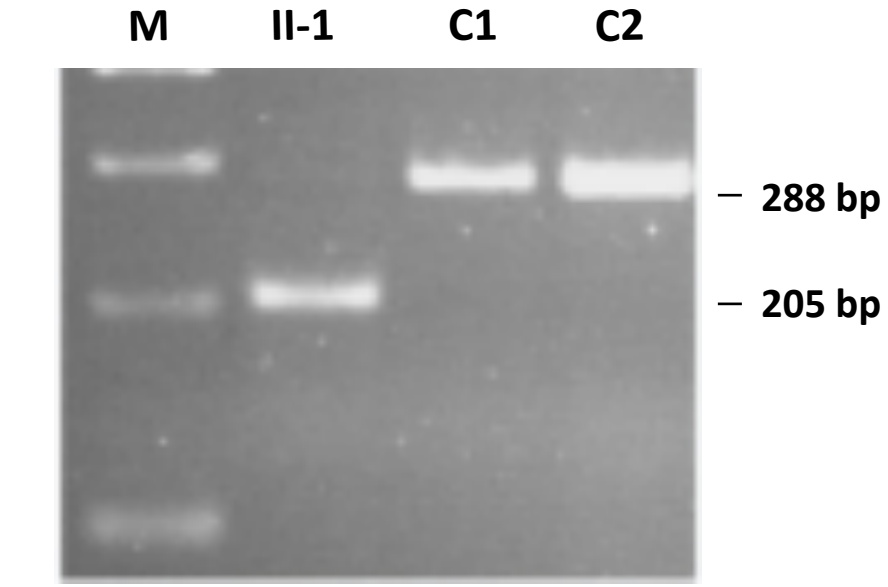
RT-PCR

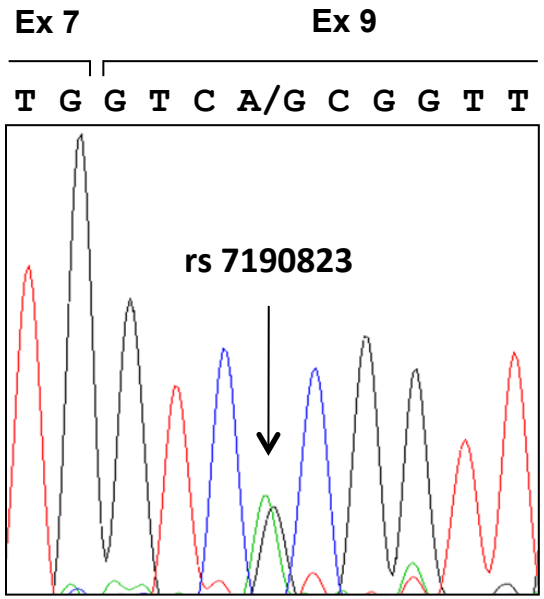


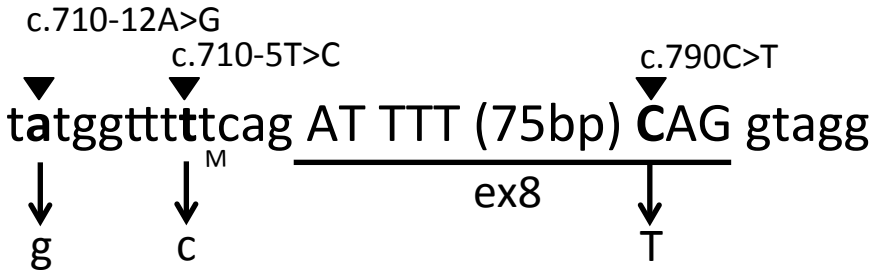
Prodotto mutato (M) uguale al prodotto wt: nessun effetto sullo splicing da parte della mutazione A>C

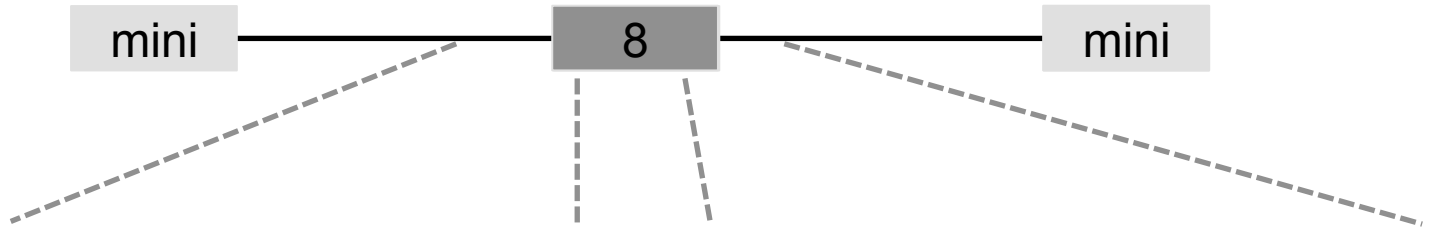

A
c.710-12A>G
c.710-5T>C
c.790C>T
tatggttttcag AT TTT (75bp) CAG gtagg
M
ex8
g
c
T


B
I-1
I-2
G A
C T
C C
G A
T T
C T
II-1
II-2
II-3
710-12A>G
710-5T>C
790C>T
G A
C T
C T
G A
C T
C T
A A
T T
C T


C
M
II-1
C1
C2
288 bp
205 bp


D
Ex 7
Ex 9
T G G T C A/G C G G T T
rs 7190823



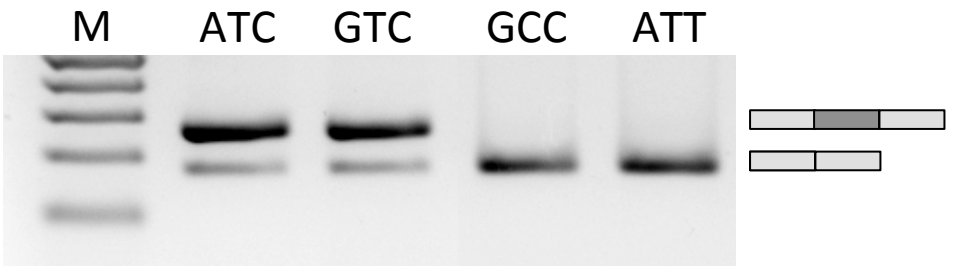


ATC ggtttatggttttcagATTTTGTTCA.....AAAAATGCCGCAGgtaggaggaaaagtcaa

GTC ggtttgtggttttcagATTTTGTTCA.....AAAAATGCCGCAGgtaggaggaaaagtcaa

GCC ggtttgtggtttctcagATTTTGTTCA.....AAAAATGCCGCAGgtaggaggaaaagtcaa

ATT ggtttatggttttcagATTTTGTTCA.....AAAAATGCCGTAGgtaggaggaaaagtcaa



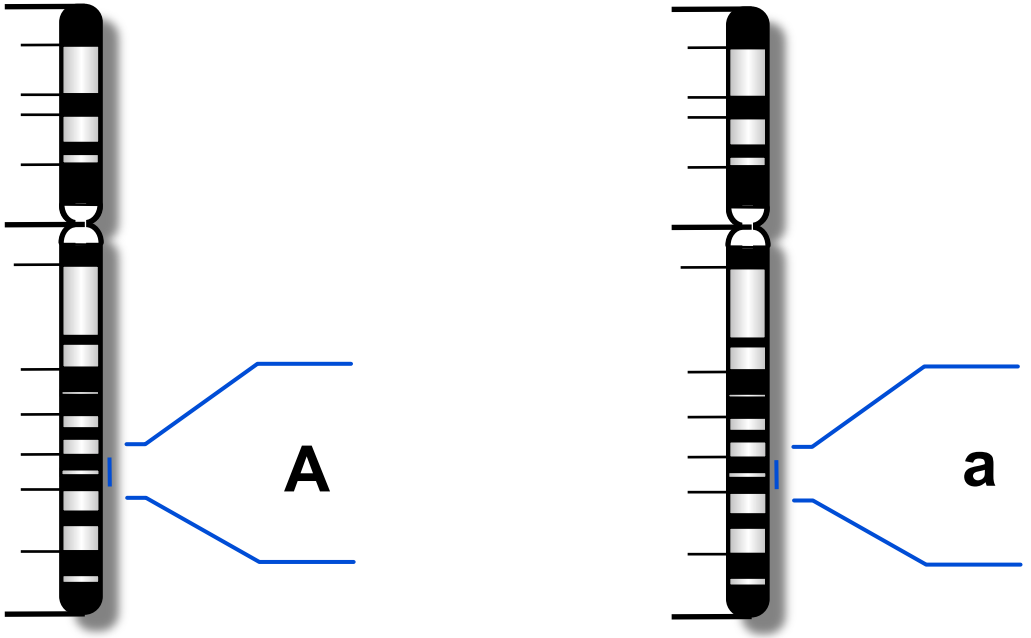
Non solo mutazioni …

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

… ma anche polimorfismi …

- Allele più raro con frequenza ≥ 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

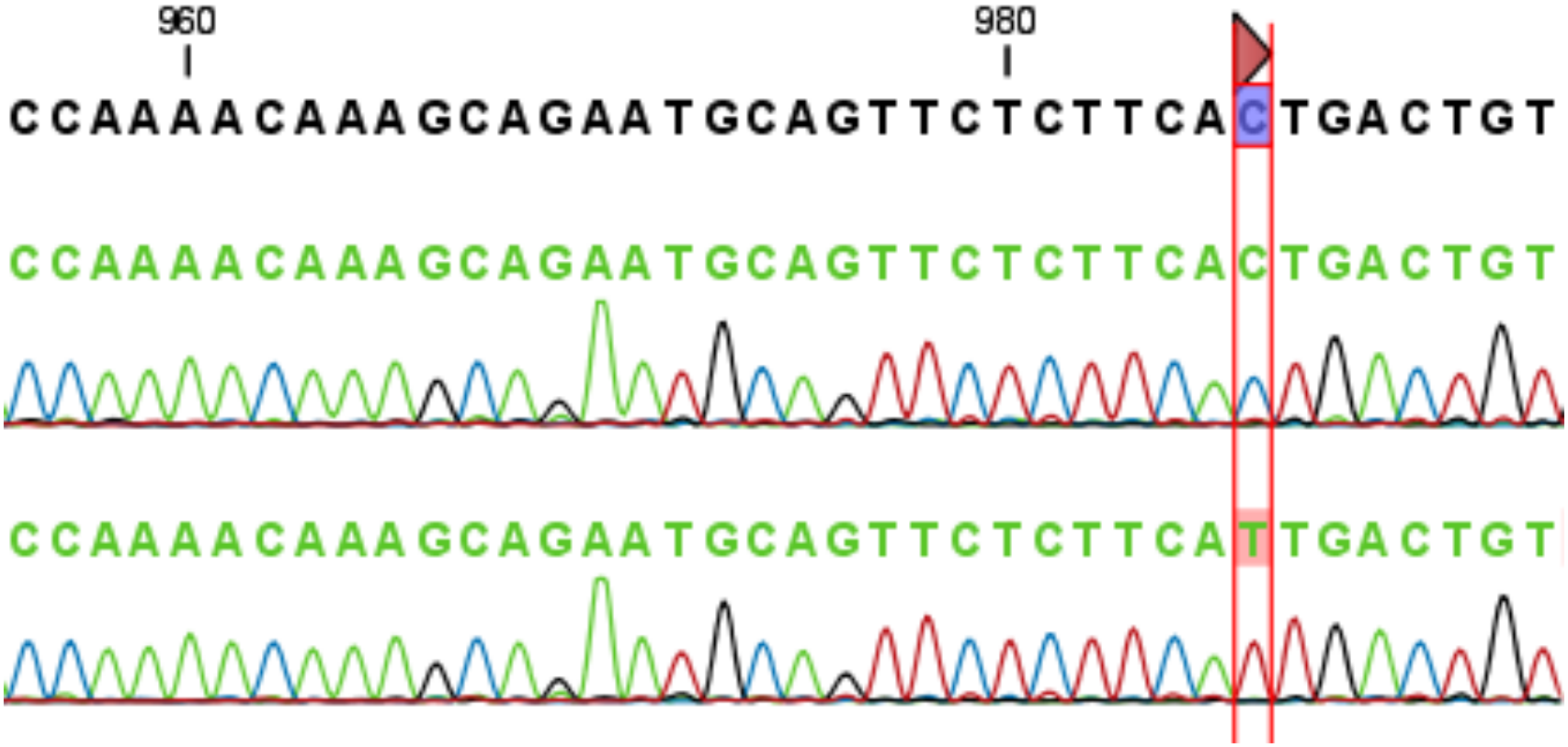
F(A1) = p = 7000/10000 = 0,70
F(A2) = q = 2000/10000 = 0,20
F(A3) = r = 1000/10000 = 0,10

p + q + r = 1

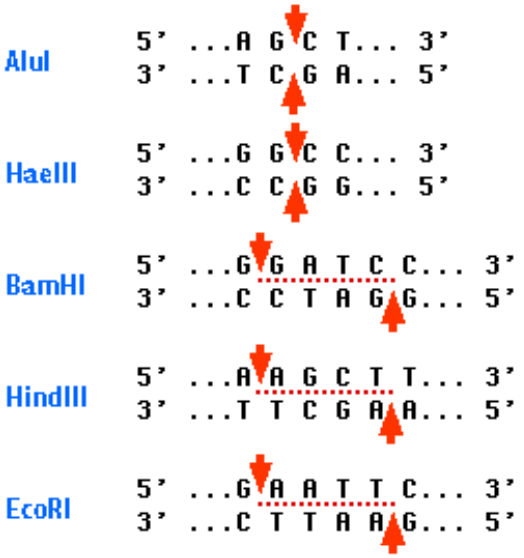
Diversi tipi di polimorfismi

- Polimorfismi del DNA
  - SNP (single nucleotide polymorphism): sostituzioni singoli nucleotidi;
  - RFLP (restriction full lenghth 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


960
980
CCAAAACAAAGCAGAATGCAGTTCTCTTCACTGACTGT
CCAAAACAAAGCAGAATGCAGTTCTCTTCACTGACTGT
CCAAAACAAAGCAGAATGCAGTTCTCTTCATTGACTGT

SNP/RFLP



AluI and HaeIII produce blunt ends

BamHI HindIII and EcoRI produce "sticky" ends

BamHI

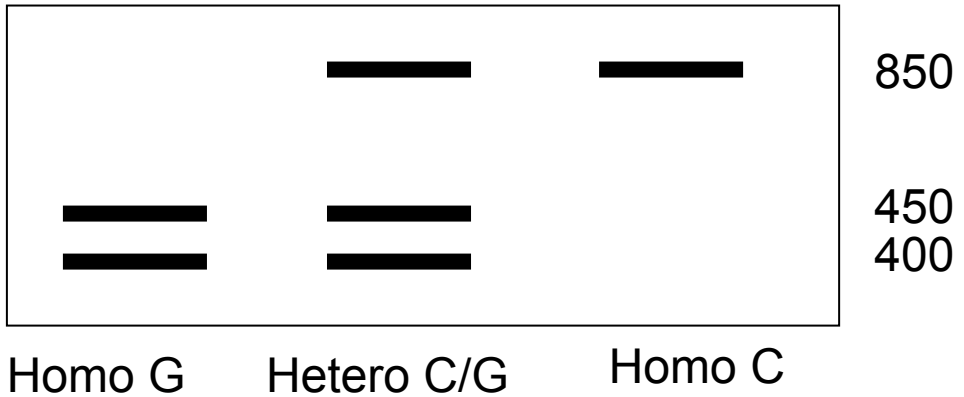
ACTGGGTACGGATCCATTTCA

450 bp 400bp

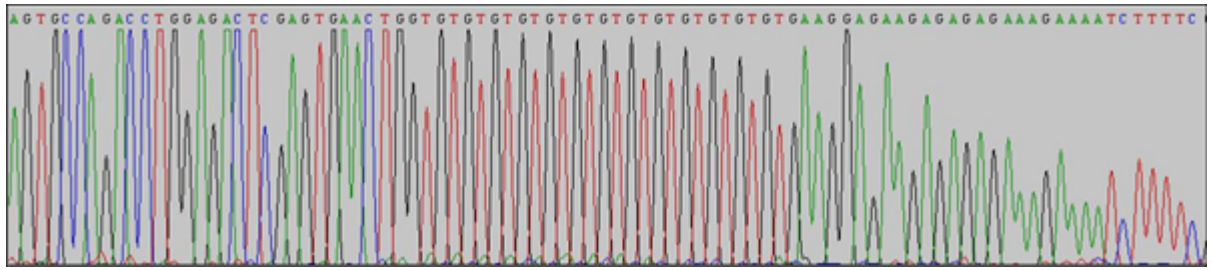
ACTGGGTACGCATCCATTTCA

850 bp

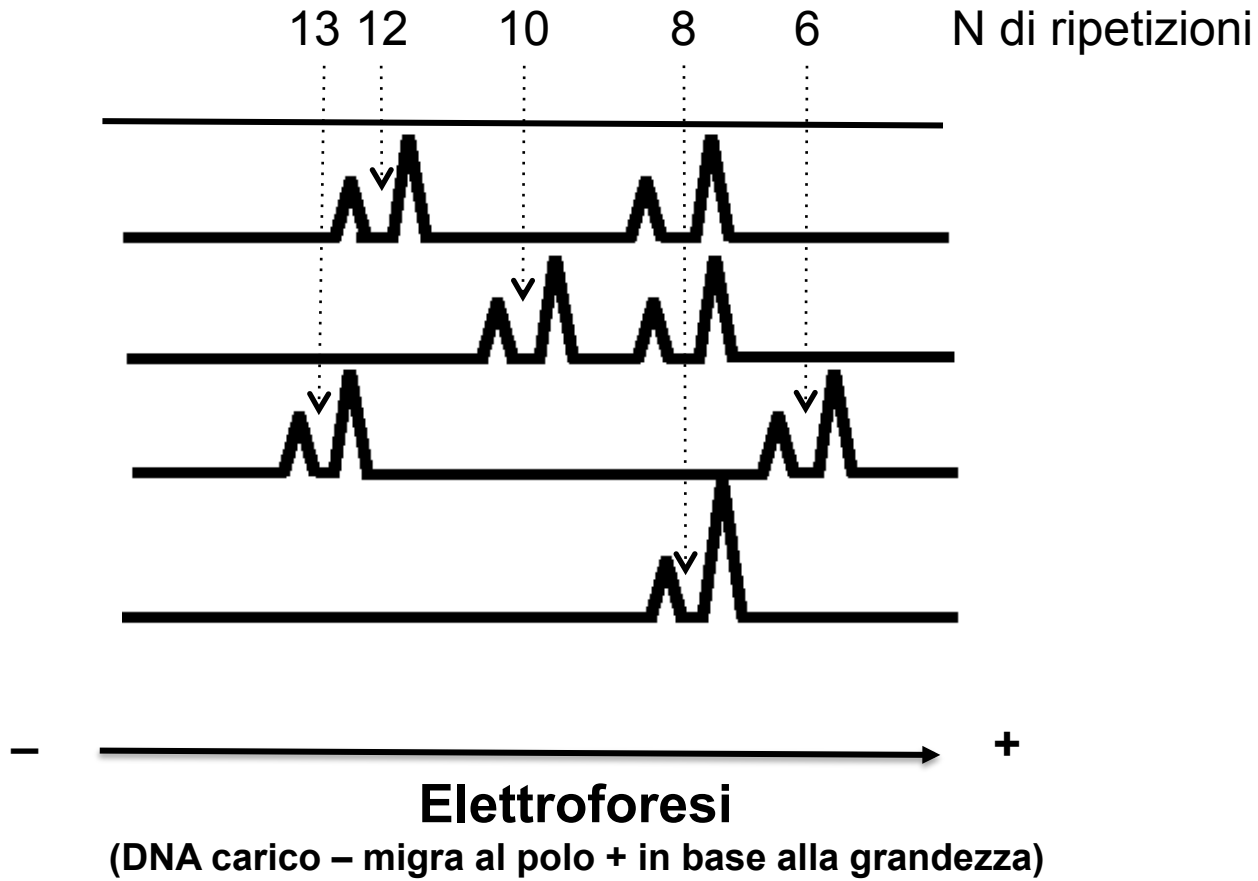
Corsa elettroforetica



Microsatelliti (ripetizioni di dinucleotidi)



(GT)n

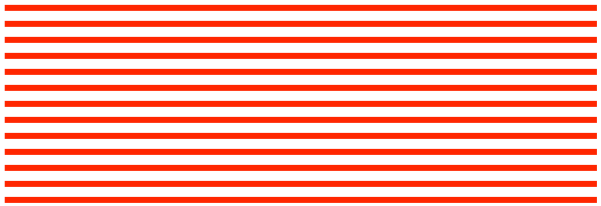


Genotipizzazione di loci polimorfici

Un locus

PCR

Amplificazione di DNA
(sintesi in vitro di molecole del frammento prescelto)



Analisi di Sequenza (altre metodiche) → SNP

Digestione con Enzimi di restrizione E elettroforesi → RFLP

Elettroforesi capillare → Microsatelliti

Più loci

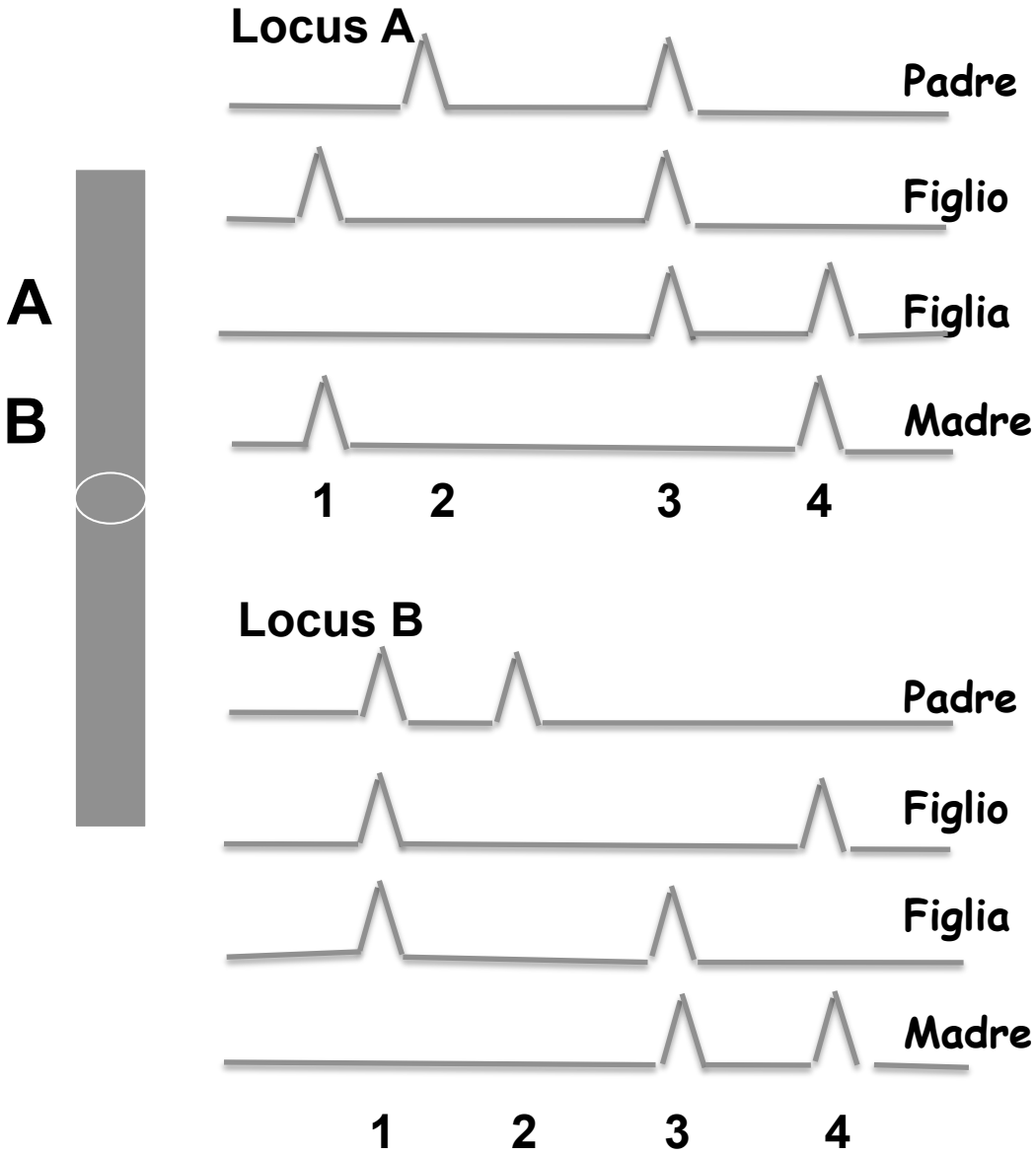
1) SNP array

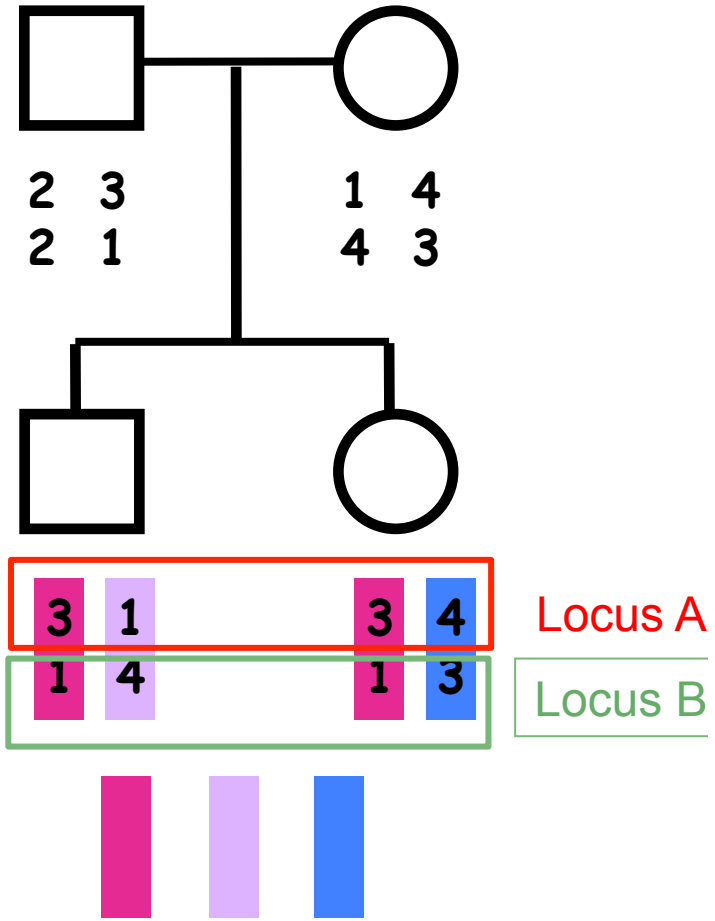
2) NGS
- Genoma
- Esoma

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 segueneti)

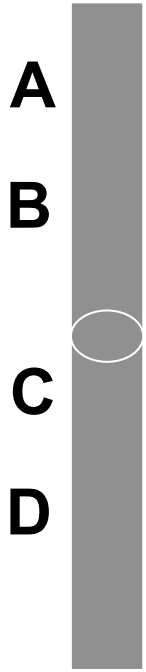
Analisi di microsatelliti e costruzione dell'aplotipo

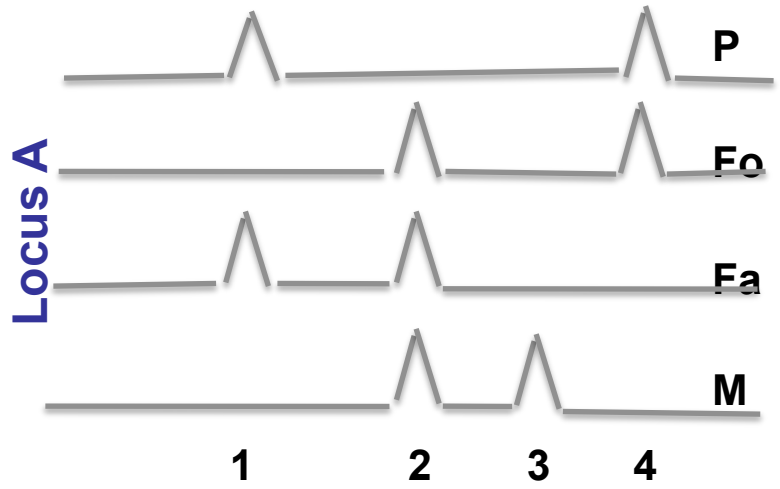


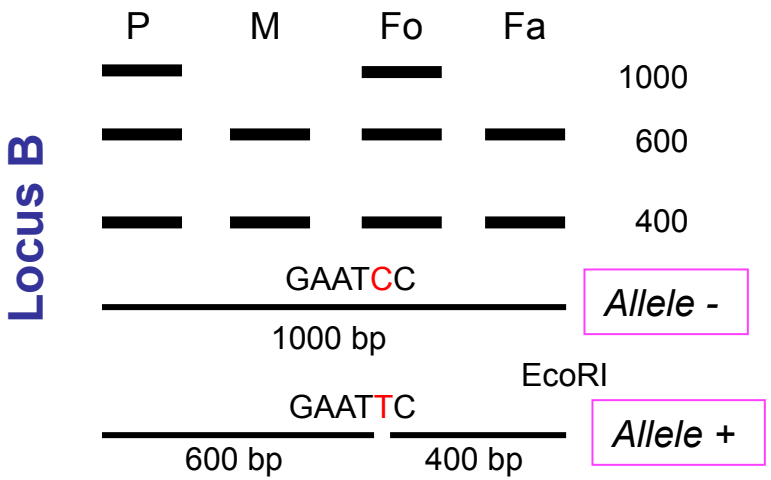


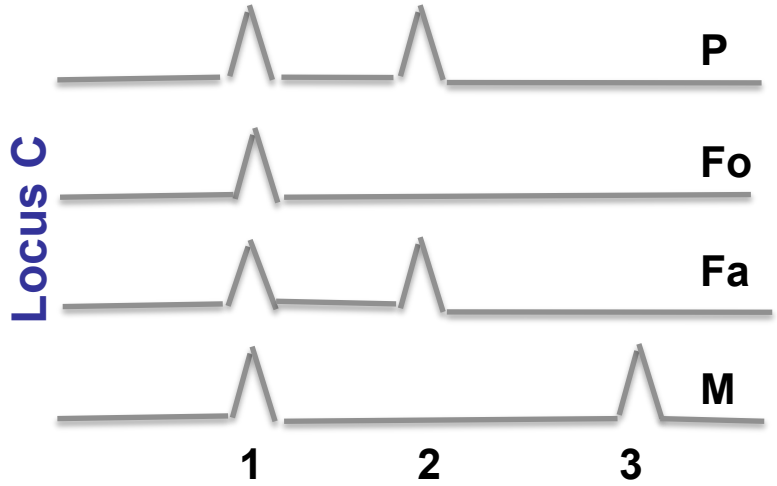
Aplotipo:
assetto allelico su cromosomi

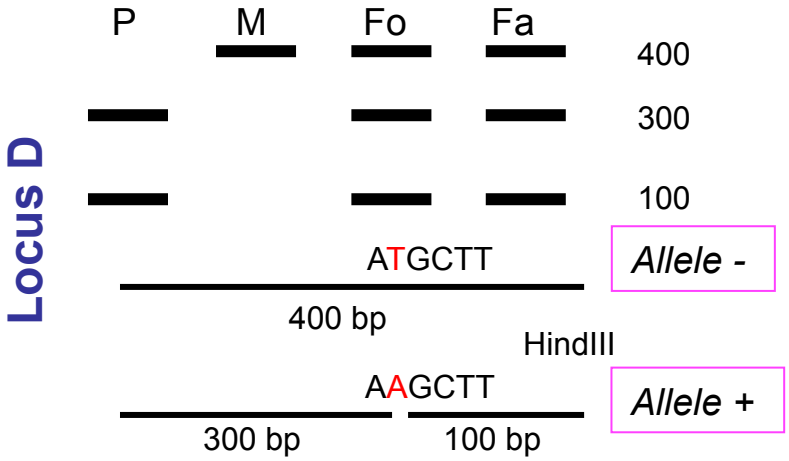
Dati i seguenti 4 loci, costruire l'aplotipo
(vedi soluzione diapo successiva)



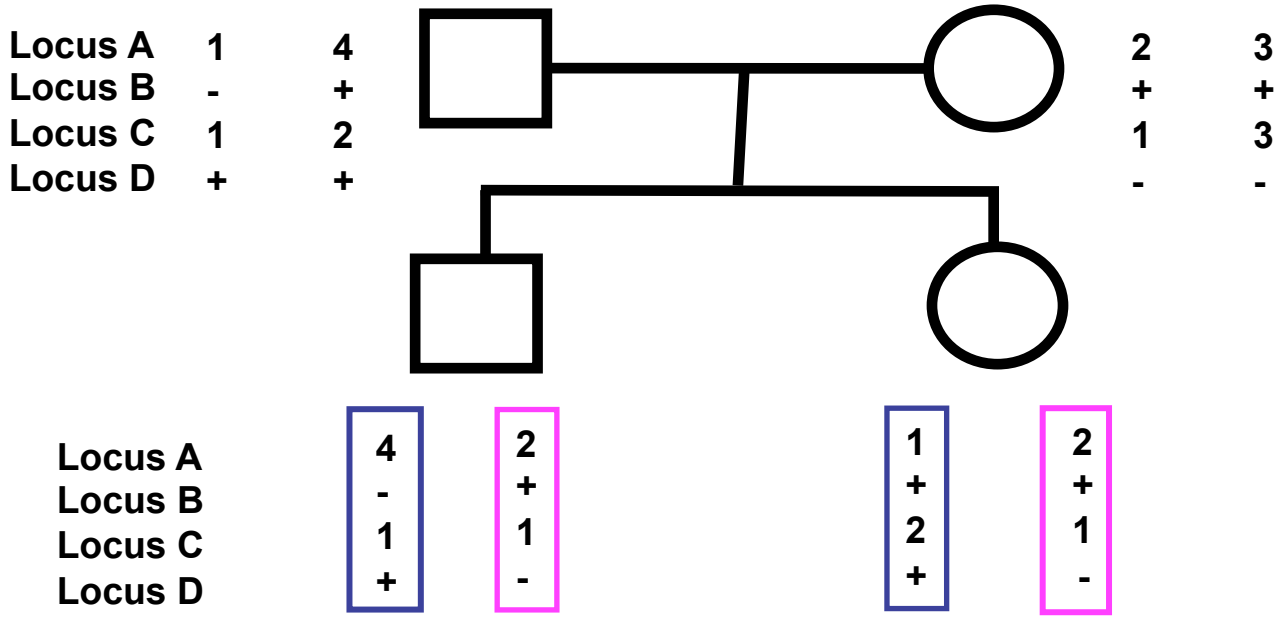








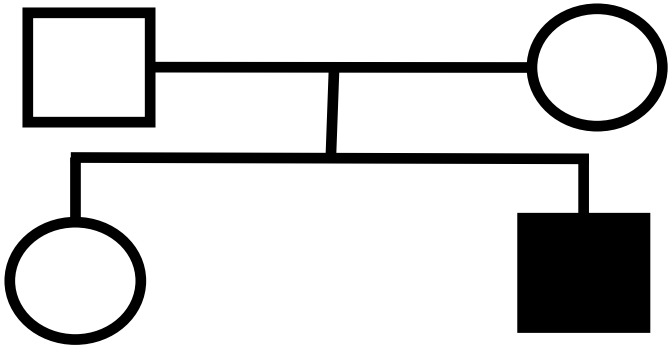
Locus	P	M	Fo	Fa
A	1 4	2 3	2 4	1 2
B	+ -	+ +	+ -	+ +
C	1 2	1 3	1 1	1 2
D	+ +	- -	+ -	+ -



Utilizzo loci polimorfici: (1) 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



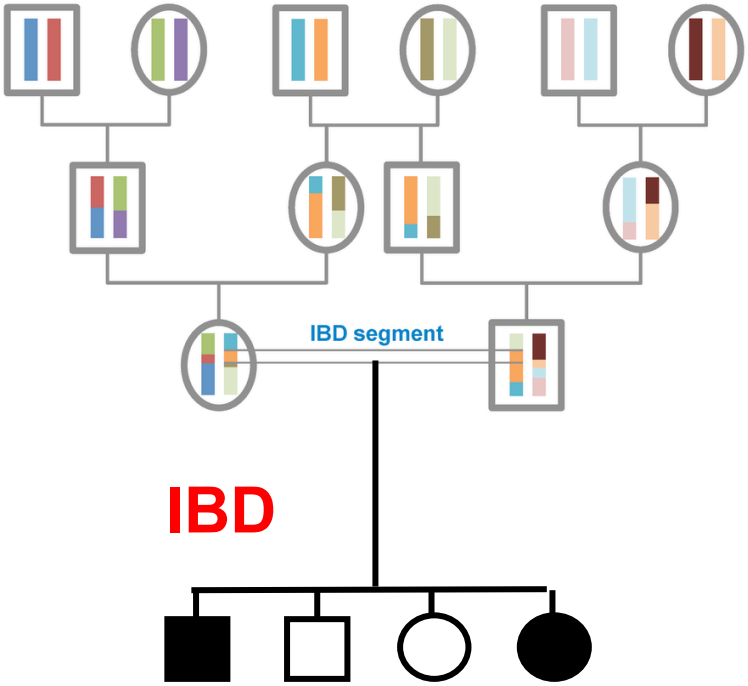
Sospetto
Sindrome di Williams (7q)

Marker A	**135**	**131**	**133**	**131**
Marker B	**150**	**154**	**160**	**-**
Marker C	**250**	**252**	**244**	**-**
Marker D	**179**	**185**	**183**	**-**
Marker E	**234**	**224**	**228**	**224**
Marker F	**120**	**120**	**127**	**120**
Marker G	**157**	**167**	**169**	**167**

Delezione Emizigosità ai loci B, C e D

Utilizzo loci polimorfici: (2) Determinazione dello stato di omozigosità:

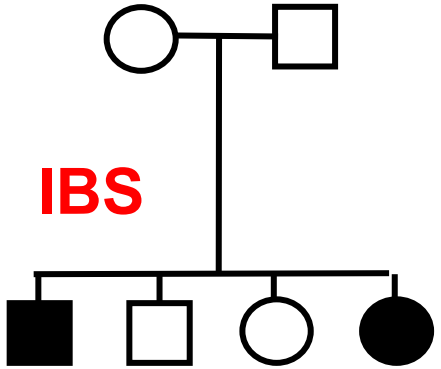
a) alleli identici per discendenza (IBD: *identical by descendant*; autozigosità)
b) Alleli identici per stato (IBS: *identical by state*)



IBD

1 1	1 2	3 2	1 1
1 1	1 2	3 2	1 1
a_1 a_1	a_1 A	A A	a_1 a_1
1 1	1 2	3 2	1 1
1 1	1 2	3 2	1 1

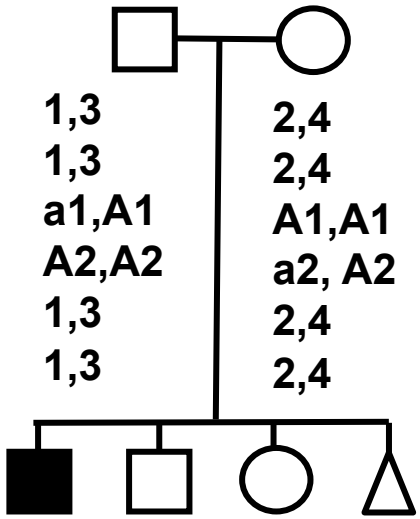
Locus 1
Locus 2
CFTR
Locus 3
Locus 4



IBS

1 2	1 3	3 4	1 2
1 2	1 3	3 4	1 2
a_1 a_1	a_1 A	A A	a_1 a_1
1 2	1 3	3 4	1 2
1 2	1 3	3 4	1 2

Utilizzo loci polimorfici: (3) Diagnosi indiretta



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

Gravidanza a rischio:
Feto sano o affetto?

1) Analisi mutazione nota a_1

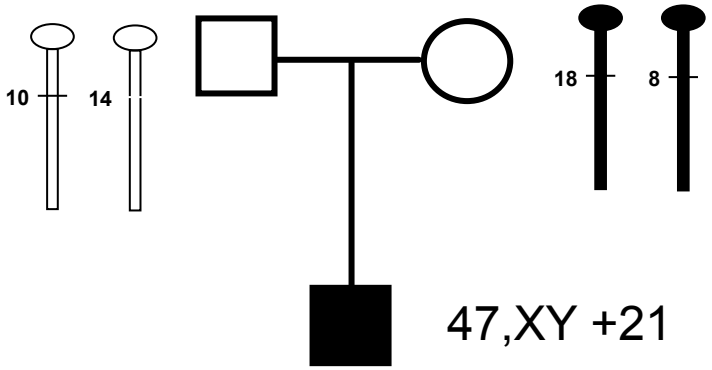
   a) Se presente > genotipo Aa_1 oppure a_1a_2
   b) Se assente > genotipo Aa_2 oppure AA

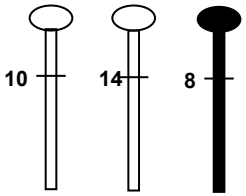
2) Se presente, procedere con l'analisi dellaplotipo paterno

   a) 4242 > affetto a_1a_2
   b) 2424 > portatore Aa_2

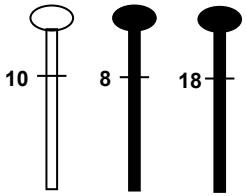
Utilizzo loci polimorfici: (4) Come determinare l'origine parentale della non-disgiunzione (ND) in I o II divisione meiotica (sindrome di Down)

Esempio con un marcatore (microsatellite)

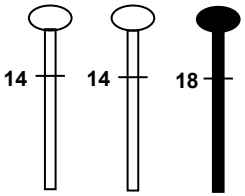




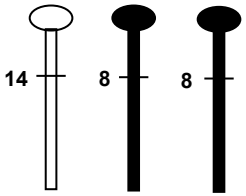
ND I
paterna



ND I
materna



ND II
paterna



ND II
materna

Utilizzo loci polimorfici: (5) 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