



# Introduzione alla Genomica

Prof. Myriam Alcalay

*Genomica Funzionale, Dpt.o Oncologia Sperimentale, Istituto Europeo di Oncologia  
Patologia Generale, Dpt.o Medicina, Chirurgia e Odontoiatria, Università degli Studi di Milano*

# Schema della lezione

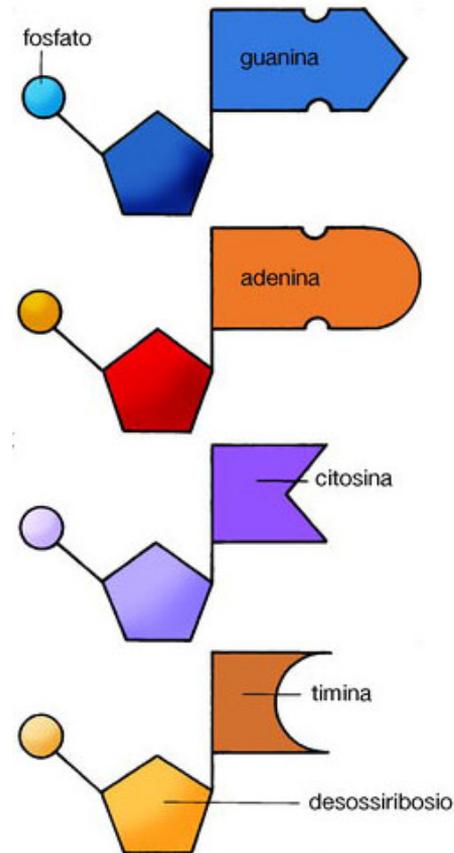
- ✚ Struttura del genoma
- ✚ Variabilità genetica
- ✚ Tecniche di analisi
- ✚ Applicazioni

# Cos'è il genoma?

- ✚ Il genoma è l'intero patrimonio genetico di un organismo vivente. Si può paragonare ad un'enorme enciclopedia in cui sono contenute le istruzioni che regolano lo sviluppo e il funzionamento dell'organismo
- ✚ Il genoma è "scritto" in un composto chimico chiamato DNA (Deoxyribo-Nucleic Acid = acido desossiribonucleico)
- ✚ *Il genoma è identico in tutte le cellule di un individuo, ad eccezione delle cellule germinali (ouovo/spermatozoo)*

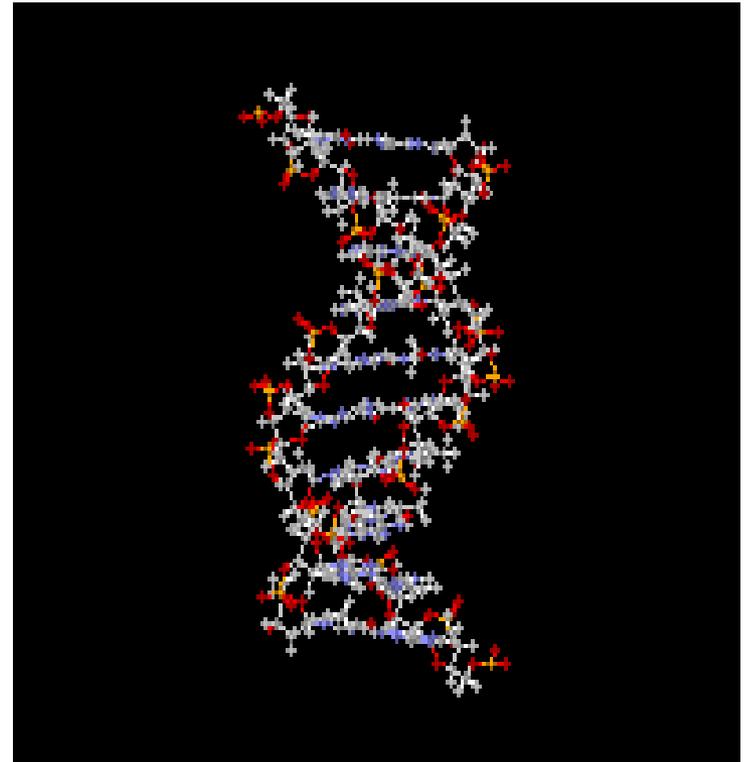
# Struttura del DNA

- ✚ I nucleotidi sono le unità che compongono una molecola di DNA, e hanno 3 componenti:
  - Uno zucchero (desossiribosio)
  - Un gruppo fosfato
  - Una base azotata (G, A, C, T)



# Struttura del DNA

- ✚ I nucleotidi si uniscono fra loro a formare lunghi filamenti di DNA.
- ✚ Il DNA non è presente come singolo filamento, ma come una coppia di filamenti saldamente associati tra loro, che s'intrecciano per formare una doppia elica.

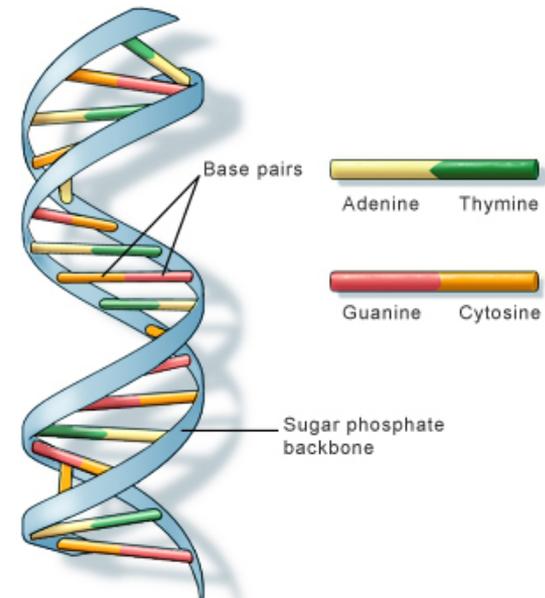


# Struttura del DNA

- La doppia elica del DNA è stabilizzata da legami che si instaurano tra le basi azotate presenti sui due filamenti.
- Ogni tipo di base forma un legame con una base posta sul filamento opposto (*appaiamento complementare*).

G → C

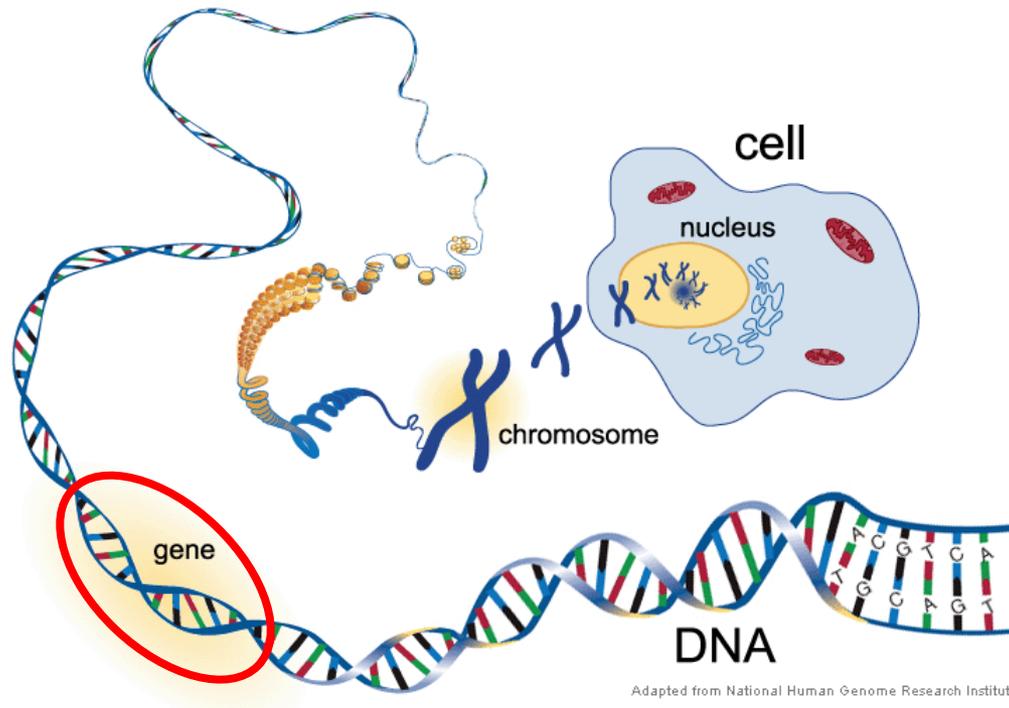
A → T



U.S. National Library of Medicine

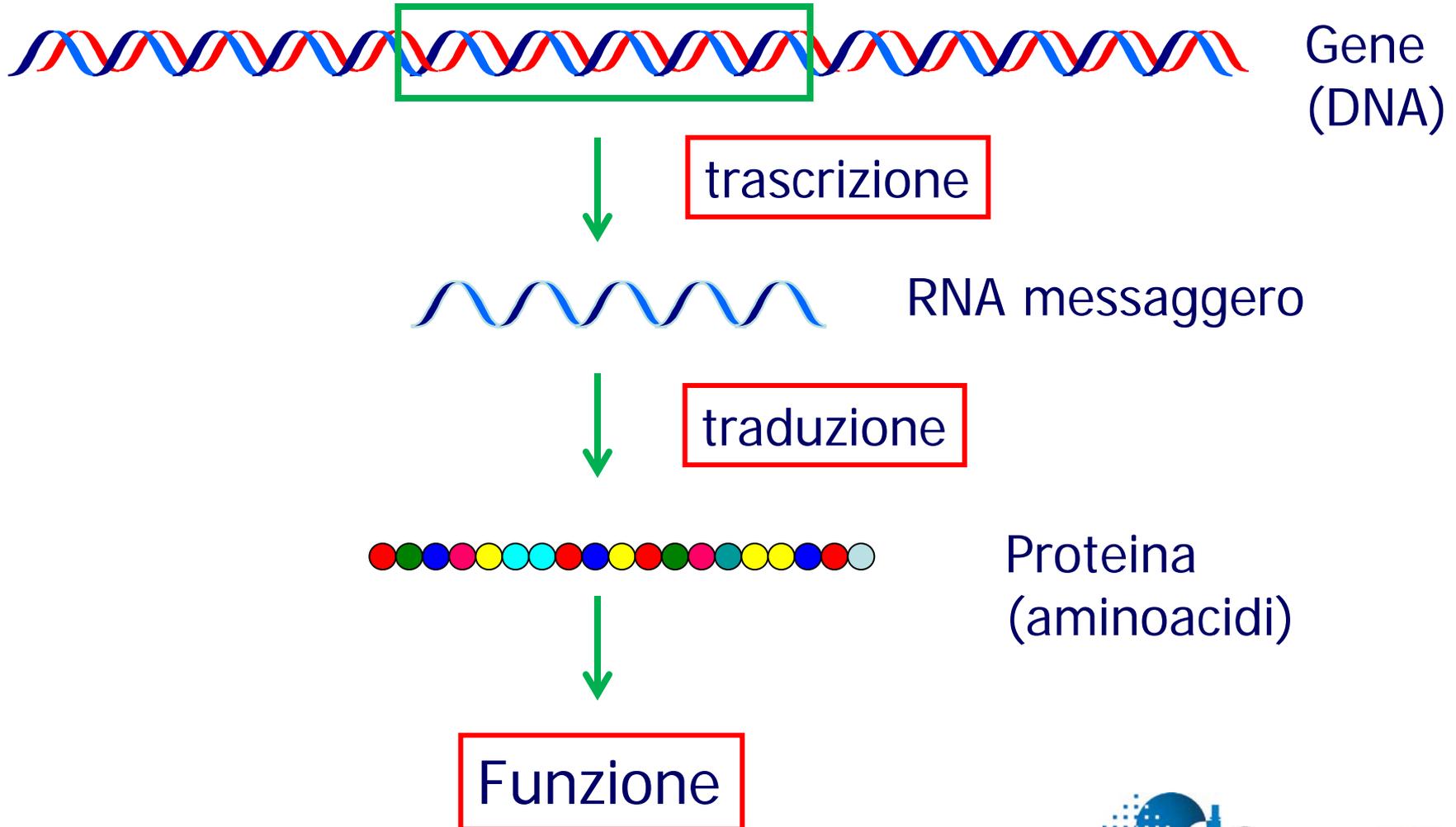
# Il gene

- Un gene è una regione del DNA che codifica per un prodotto funzionale.



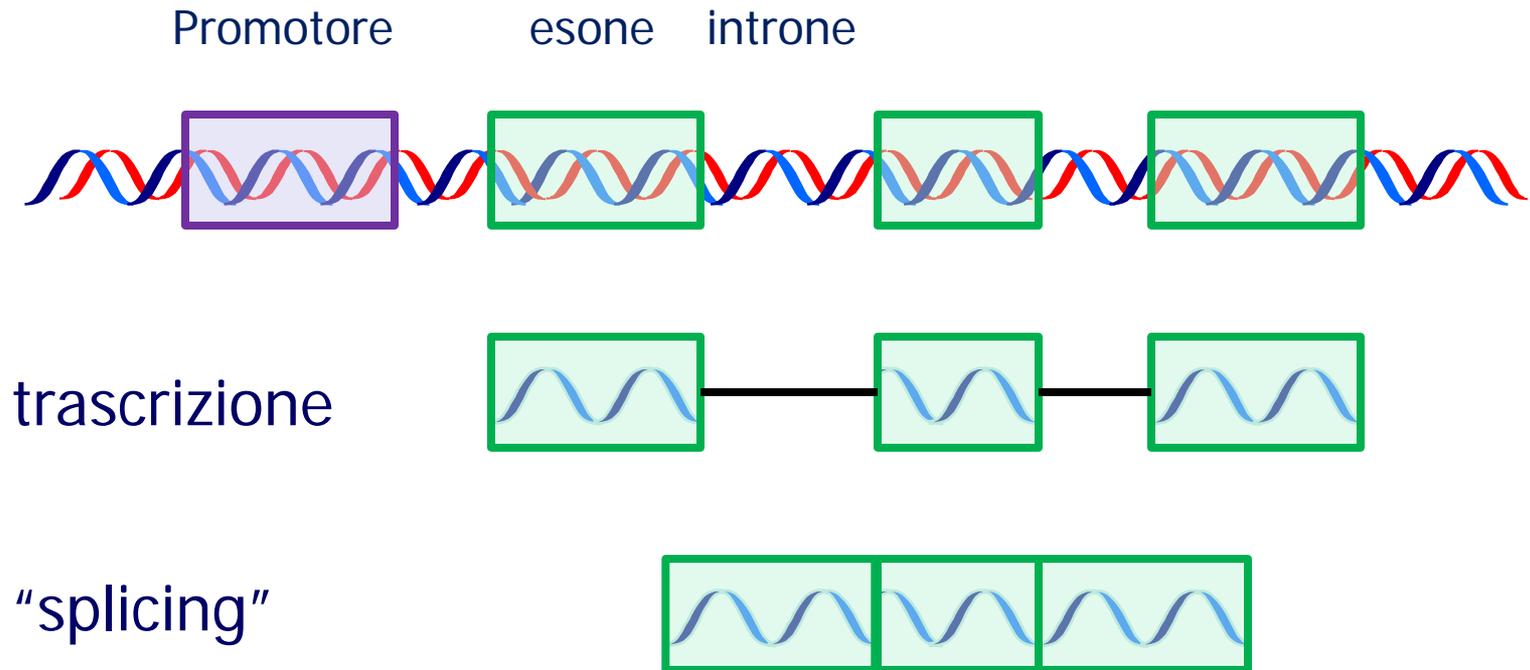
2-3% del genoma!

# Flusso dell'informazione genetica

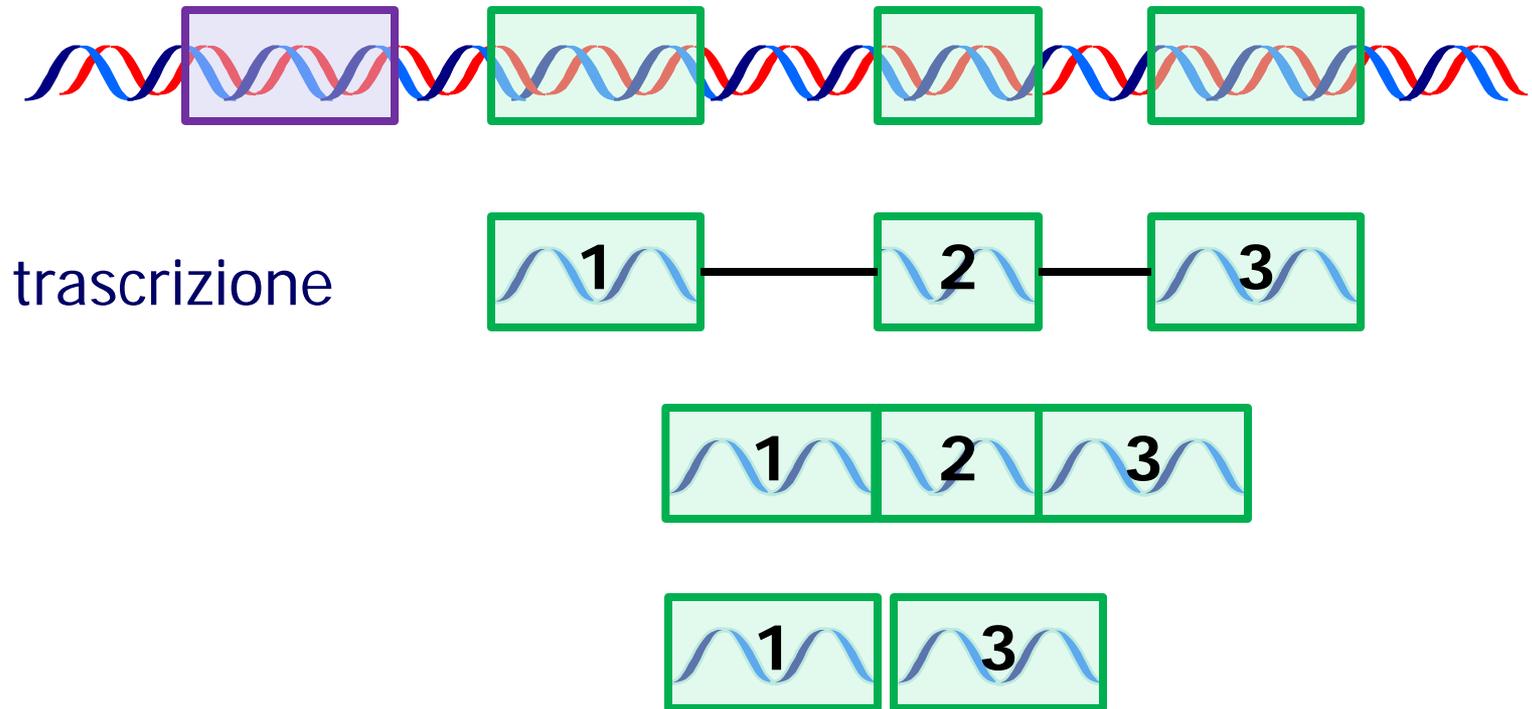


# Il gene - struttura

- I geni sono composti da sequenze codificanti e da regioni non codificanti:



# Splicing alternativo



1 gene → più di una proteina

# Splicing alternativo

ilxhbuvmiojcu fvu dentista cndicecnchebcwcb mangiaresjvbibvtroppendcaramelle vfevfammmlenai vvdenti

ilxhbuvmiojcu fvu dentista cndicecnchebcwcb mangiaresjvbibvtroppendcaramelle vfevfammalenai vvdenti

Il mio dentista dice che mangiare troppe caramelle fa male ai denti

Il mio dentista dice che mangiare troppe caramelle fa male ai denti

Il mio dentista dice che mangiare caramelle fa male ai denti

Il mio dentista dice che mangiare fa male ai denti

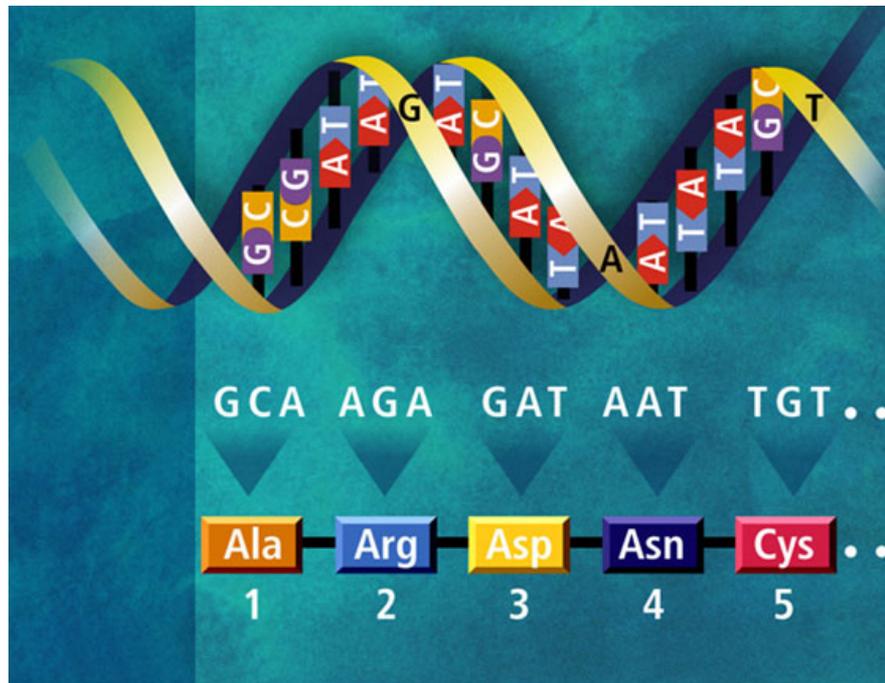
Il mio dentista dice che mangiare fa male

Il mio dentista mangia troppe caramelle

Il dentista fa male ai denti

# Il Codice Genetico

- Il codice genetico (= lo schema attraverso cui la cellula *traduce* una sequenza nucleotidica in una sequenza di amminoacidi) viene letto a triplette.



# Il Codice Genetico

- ✚ Ci sono 4 nucleotidi, quindi 64 diverse triplette ( $4^3$ ).
- ✚ Ci sono solo 20 aminoacidi  
 → il codice genetico è *degenerato* (triplette diverse codificano per lo stesso aminoacido)

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

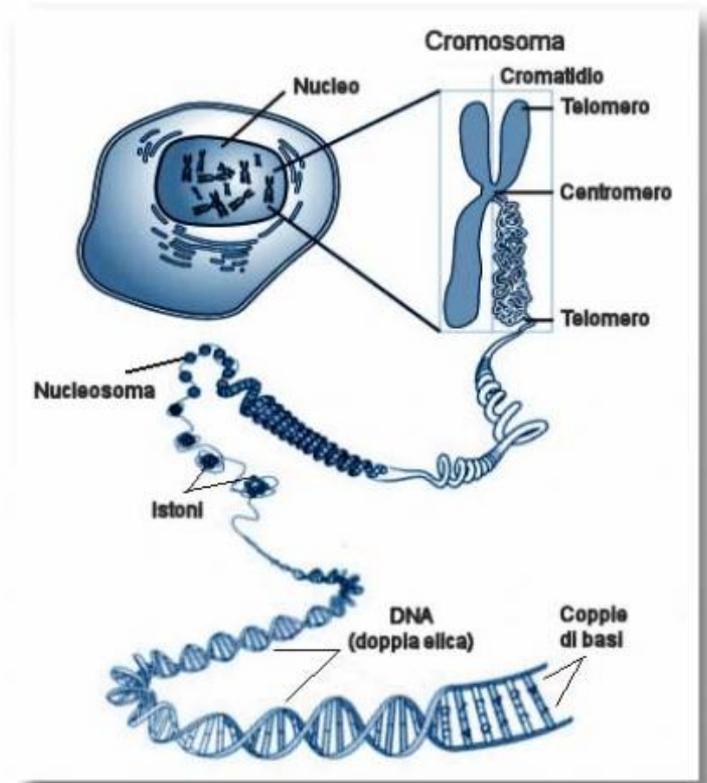
# Organizzazione del Genoma

- ✚ Il genoma umano è composto da circa 3 miliardi di nucleotidi, che disposti uno in fila all'altro formerebbero un filamento di 2 metri.
- ✚ Come fanno 2 metri di materiale genetico a entrare nel nucleo di una cellula?

→ cromosomi

# I Cromosomi

- Il DNA si avvolge intorno a proteine chiamate "istoni", che si comprimono a formare delle strutture chiamate cromosomi.
- Questi vengono aperti soltanto per leggere l'informazione genetica.



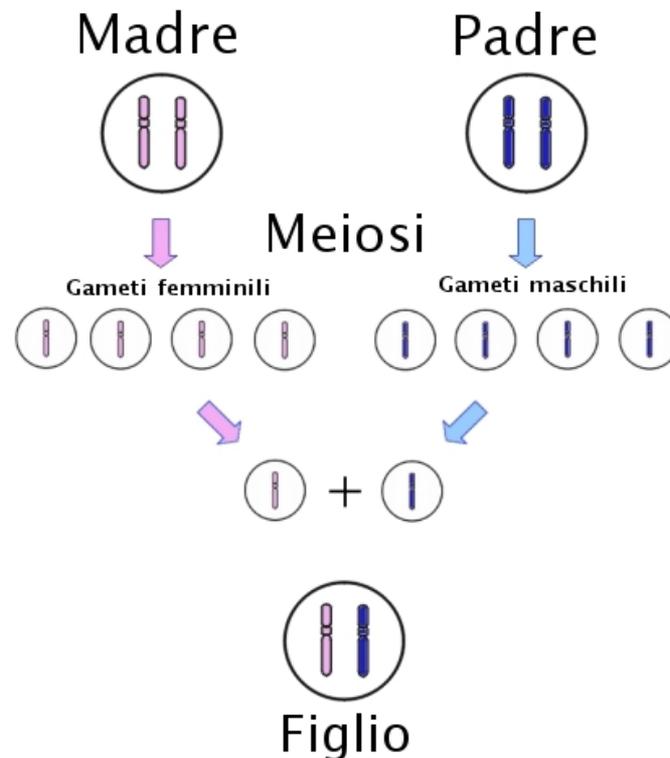
# I Cromosomi

- ✚ Ogni specie ha un numero diverso di cromosomi.
- ✚ Gli umani hanno 23 paia di cromosomi (=46 totali)

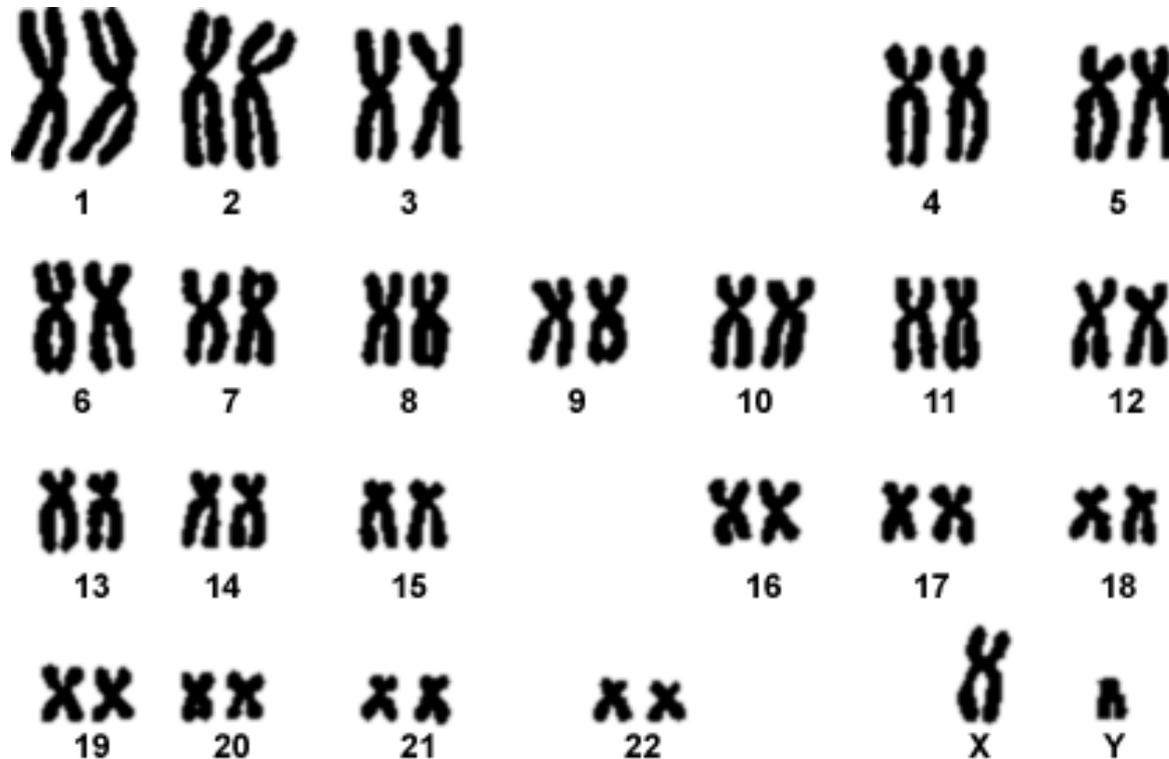
Organismo	Numero totale di cromosomi
Ameba	12
Ananas	50
Cane	78
Carpa	22
Cavallo	64
Cotone	52
Elefante	56
Fagiolo	22
Formica	2
Gamberetto	254
Gatto	38
Moscerino	8
Mucca	60
Pollo	78
Scimpanzé	48
Topo	40

# I Cromosomi

- Ciascun cromosoma è presente in duplice copia  
= genoma diploide

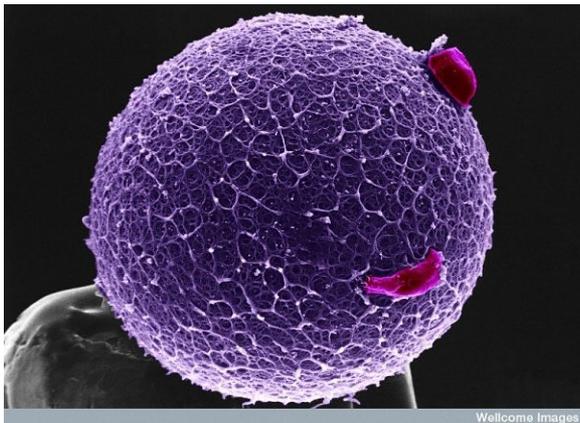
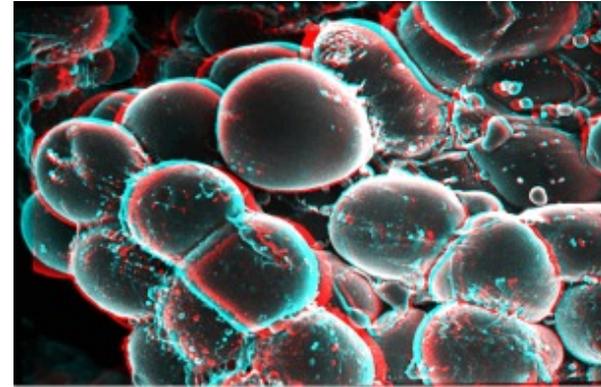
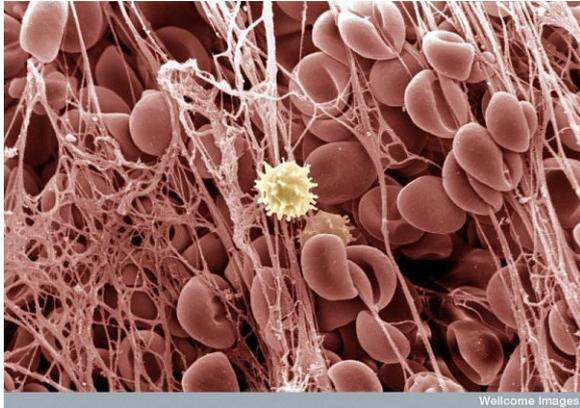


# I Cromosomi



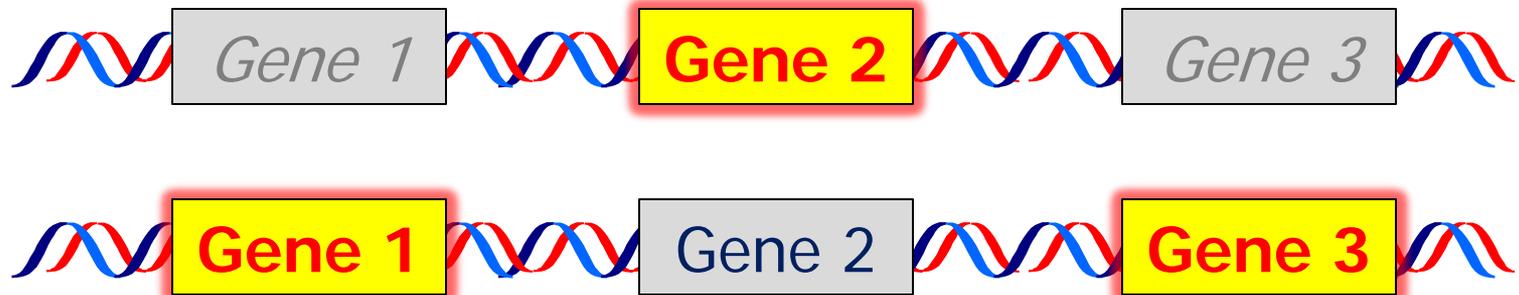
L'informazione genetica è uguale in tutte le cellule

→ le cellule sono molto diverse come struttura e come funzione



# Il gene - attività

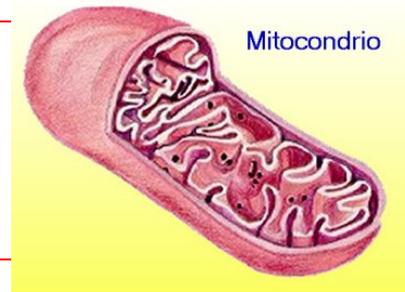
- ✚ Il genoma è uguale in tutte le cellule.
- ✚ Non tutti i geni sono attivi allo stesso tempo



# Riassumendo

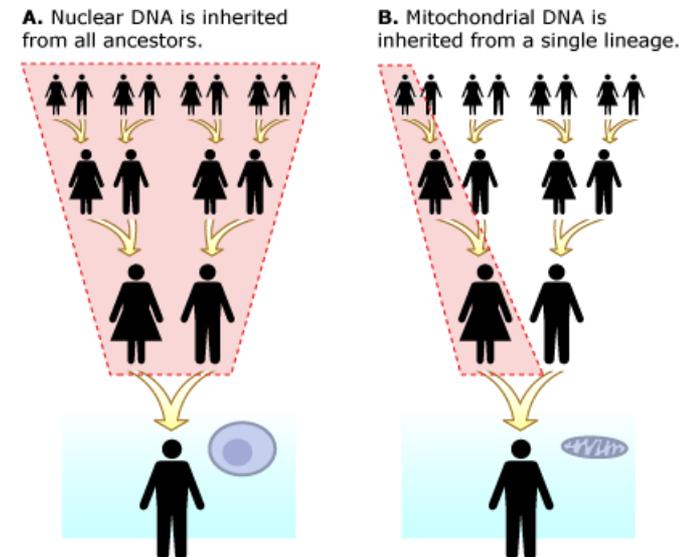
- ✚ Tutte le funzioni del nostro organismo sono svolte da proteine, codificate in specifiche regioni del DNA denominate “geni”.
- ✚ La struttura delle proteine è determinata dalla sequenza dei nucleotidi che compongono il gene.
- ✚ Per ogni proteina del nostro organismo, c'è una duplice copia del gene codificante.
- ✚ Le funzioni di una cellula sono determinate dai geni che sono attivamente trascritti.

# DNA mitocondriale

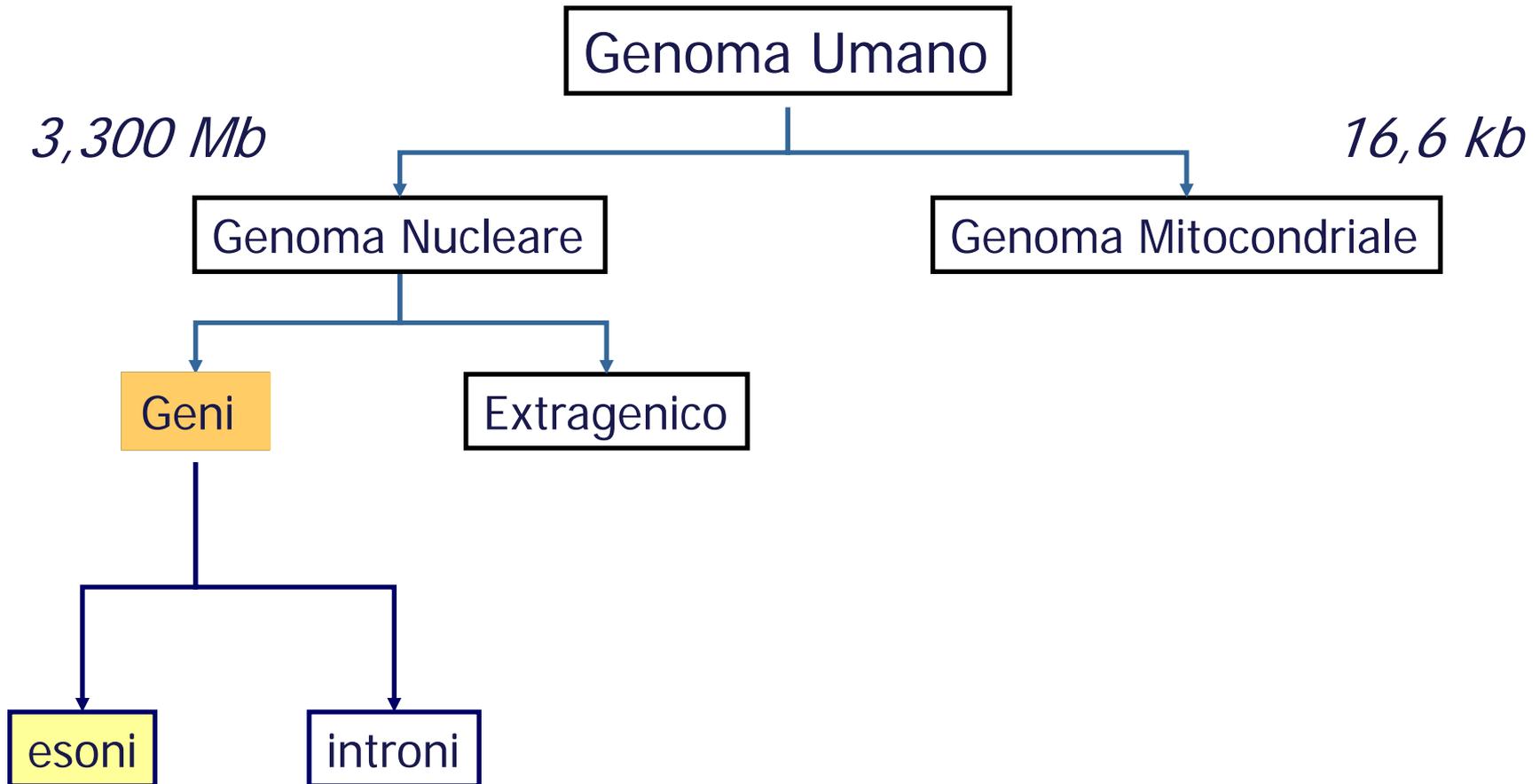


- ✚ Non tutto il DNA sta nel nucleo delle cellule. Una piccolissima porzione (17,000 nucleotidi codificanti per 37 geni) è contenuta in un organello citoplasmatico deputato alla produzione di energia, il mitocondrio.

- ✚ Il DNA mitocondriale è:
  - Di derivazione solo materna
  - Presente in molteplici copie in ogni cellula
  - Più piccolo → meno suscettibile di degradazione



# Genoma Umano



# Variabilità genetica

- Il gene che codifica per una proteina può esistere in diverse versioni:

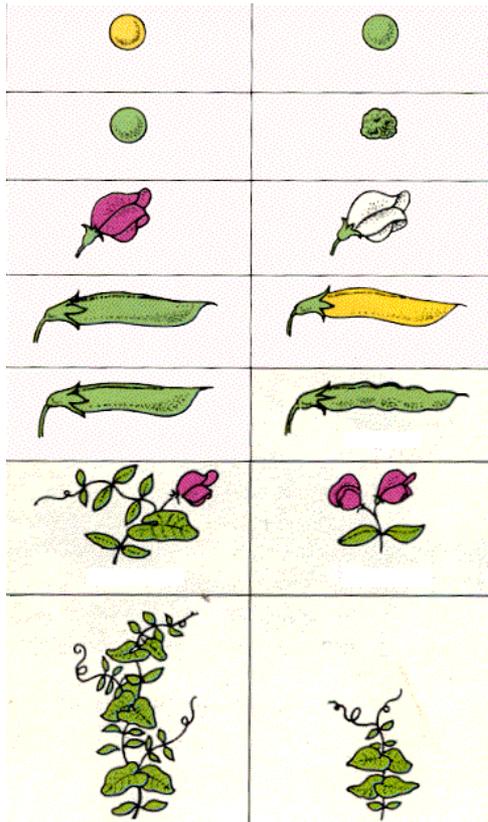


=



# Ripensiamo a Mendel...

Mendel studiò diversi caratteri



Colore del seme

Forma del seme

Colore del fiore

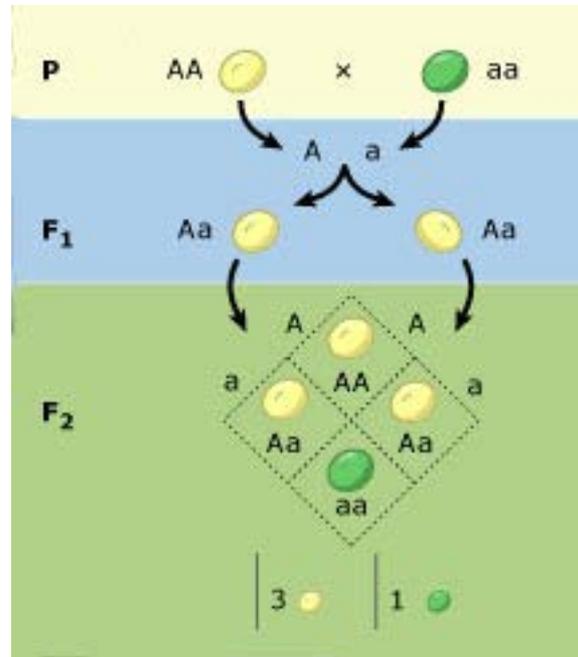
Colore del baccello

Forma del baccello

Posizione del fiore

Lunghezza del fusto

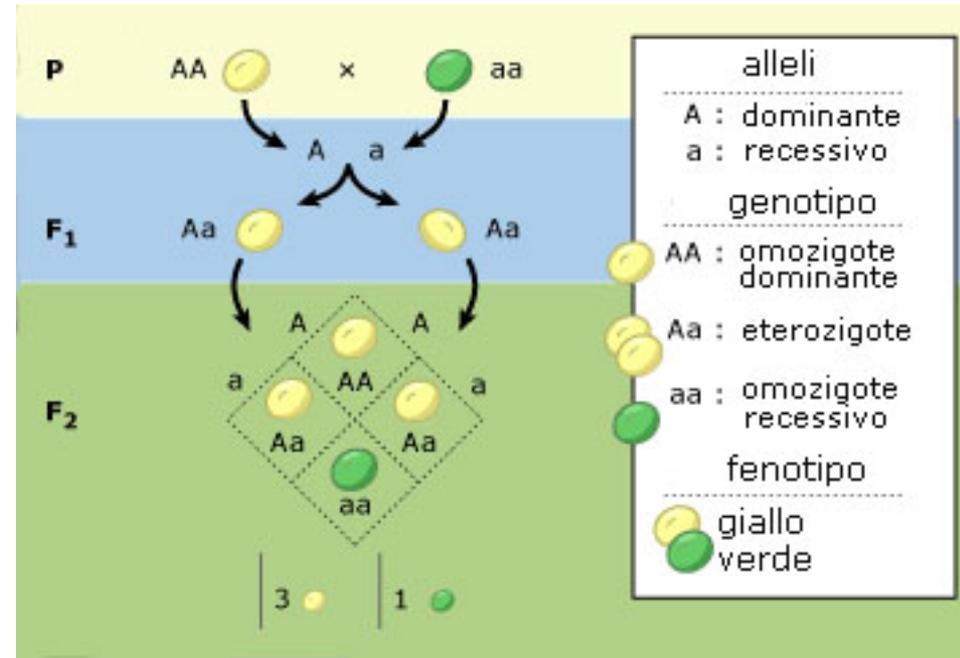
# Caratteri dominanti/recessivi



Giallo è DOMINANTE  
Verde è RECESSIVO

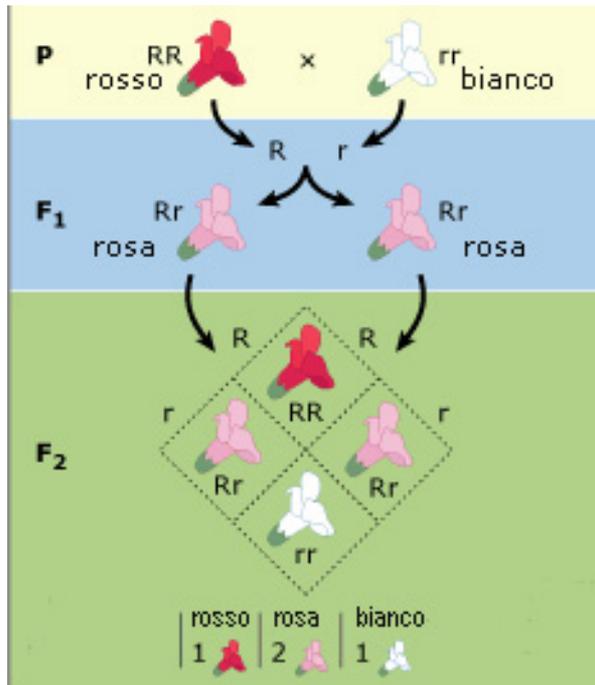
# Genotipo/Fenotipo

- ✚ **Allele:** Una delle varie forme di un gene.
- ✚ **Genotipo:** I due alleli presenti in un individuo. Definiti omozigote o eterozigote a seconda che i due alleli siano identici o diversi.
- ✚ **Fenotipo:** carattere visibile risultante dal genotipo



# Non è così semplice....

- ✚ Dominanza incompleta:  
il fenotipo dell'eterozigote è intermedio tra quelli dei due omozigoti



- ✚ Co-dominanza:  
l'eterozigote manifesta il fenotipo di entrambe le situazioni omozigoti

gruppi sanguigni nell'uomo	A	B	O
A	AA (AA)	AB (AB)	AO (A)
B	AB (AB)	BB (B)	BO (B)
O	AO (A)	BO (B)	OO (O)

# Caratteri monogenici

- ✚ il fenotipo dipende dall'azione di un solo gene, che è necessario e sufficiente all'espressione del carattere



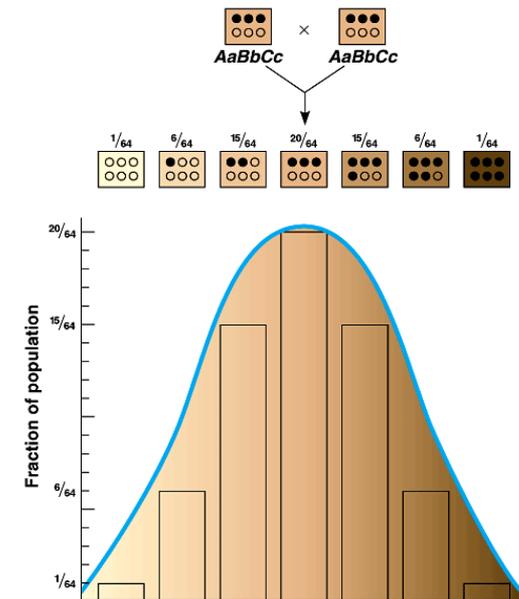
Lobo attaccato/staccato



attaccatura a "V"

# Caratteri poligenici

il carattere è determinato da più geni, ciascuno dei quali contribuisce in una certa misura all'espressione della stessa caratteristica fenotipica  
→ Es. colore della pelle (genotipo e fototipo)



# Caratteri multifattoriali

- ✚ Risultanti dall'interazione fra le condizioni ambientali e un numero di geni  $>1$  partecipanti alla determinazione del fenotipo.  
Es. → colore della pelle (fenotipo)



# Variabilità genetica



*"Gli uomini sono vicini per natura, ma in pratica diventano molto distanti."*



*Confucio,  
Filosofo cinese  
551 BC – 479 BC*

# Variabilità genetica



Ogni individuo è al 99.9% geneticamente identico ad ogni altro. Lo 0.1% di differenze è dovuto a **variazioni genetiche**.

Le variazioni genetiche sono responsabili della varianza nella specie e possono avere un impatto sulla risposta individuale a patologie, insulti ambientali e farmaci.

Questo rende le variazioni genetiche estremamente importanti nella ricerca biomedica e nello sviluppo farmacologico.

# Variazioni Genetiche Umane

- ✦ Variazioni quantitative (inserzioni o delezioni di uno o più nucleotidi)
  - Tandem Repeat Polymorphisms
  - Insertion/Deletion Polymorphisms (INDEL)
  - Copy Number Variations (CNV)
- ✦ Variazioni qualitative: sostituzione di un singolo nucleotide con un altro
  - Single Nucleotide Polymorphisms (SNP)

# Tandem Repeat Polymorphisms

Circa il 40% del genoma è rappresentato da sequenze ripetute.

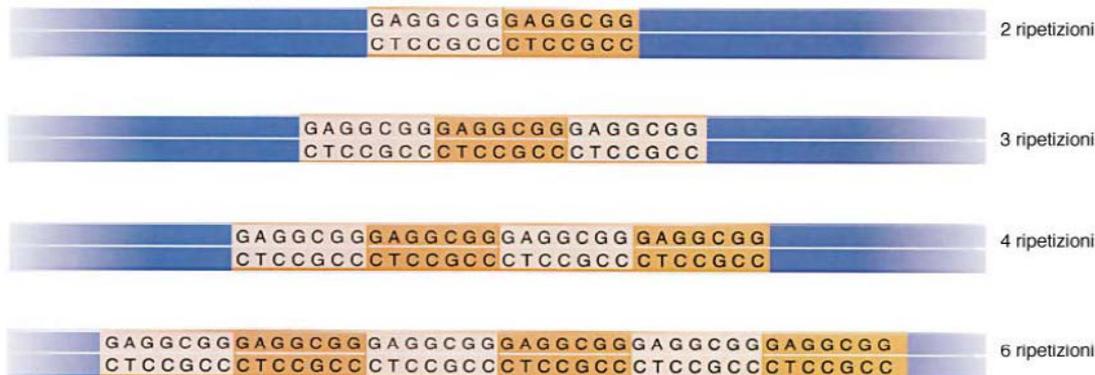
2 tipi di sequenze ripetute presentano alta variabilità:

Minisatelliti – Variable Number of Tandem Repeats (VNTR)

✚ sequenze di 8 - 80 nucleotidi ripetute in tandem numerose volte (centinaia o anche migliaia di volte) e in più regioni cromosomiche.

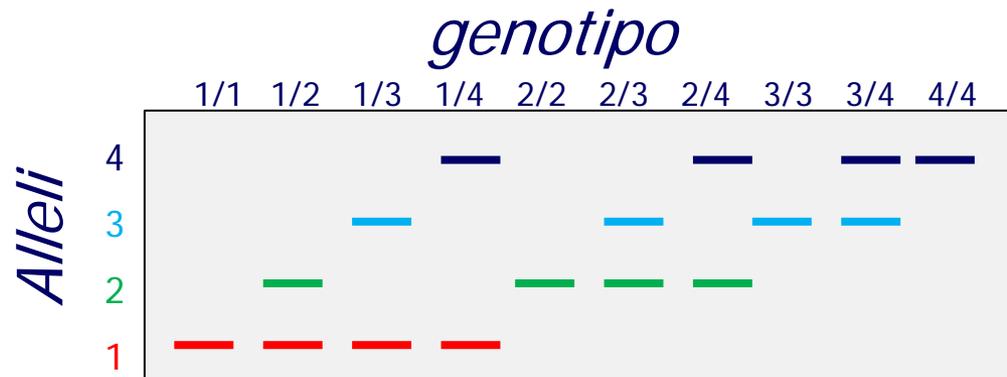
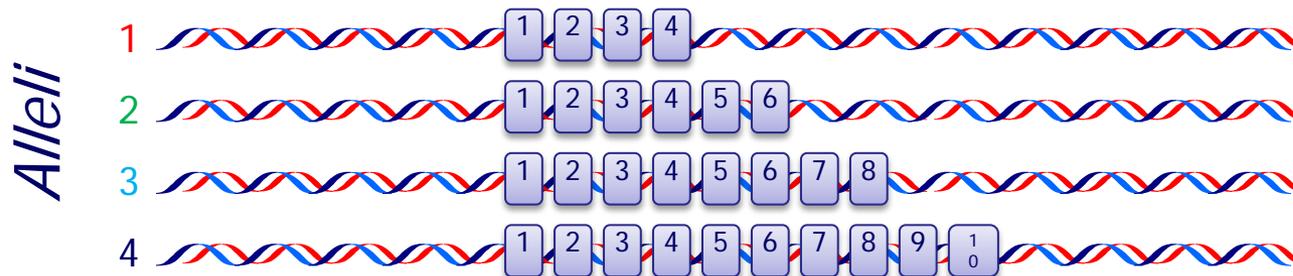
Microsatelliti – Short Tandem Repeat (STR)

✚ l'unità ripetuta in tandem è più breve: 2-7 nucleotidi



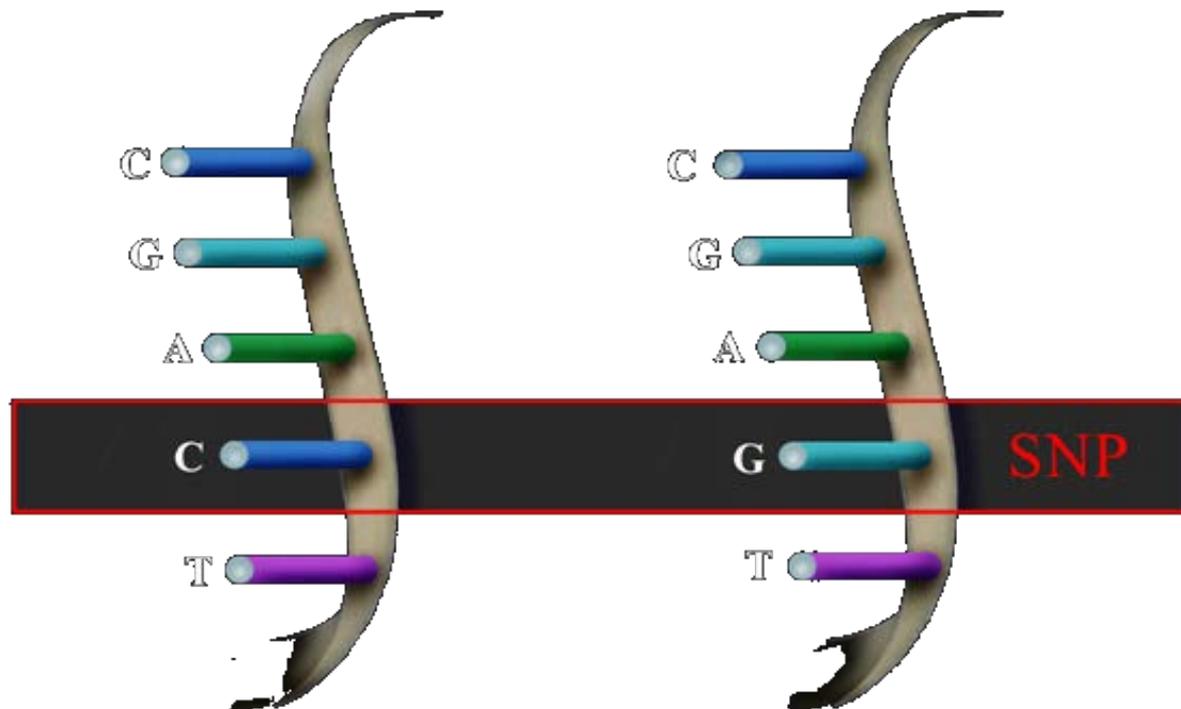
# Tandem Repeat Polymorphisms

- Ci sono decine di migliaia di VNTR e STR nel genoma, e per ciascuno esistono in molteplici varianti alleliche → ogni individuo ha una costituzione unica.



# Single Nucleotide Polymorphisms

SNP = variazione di sequenza con cambiamento di un singolo nucleotide: es - A**A**GGCTAA → AT**T**GGCTAA.



# Classificazione degli SNPs

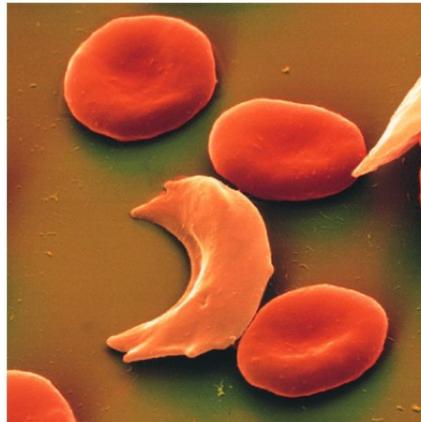
- ✚ Gli SNPs possono essere localizzati in qualsiasi punto del genoma. In base a dove si trovano in rapporto alla struttura dei geni: intronici, esonici, promotoriali, ecc.
- ✚ Gli SNPs che cadono in regioni codificanti possono essere ulteriormente suddivisi in:
  - ✚ **Sinonimi**: quando la sostituzione nucleotidica **non** dà origine ad un cambiamento dell'aminoacido codificato
  - ✚ **Non-sinonimi**: quando la sostituzione nucleotidica dà origine ad un cambiamento dell'aminoacido codificato

# SNPs e malattie

- ✚ Talvolta la variazione di un nucleotide porta alla formazione di un prodotto proteico anomalo o non funzionale → associa ad una specifica patologia.

Sickle cell anemia:

					<u>GLU</u>				
HBB	...	CAC	CTG	ACT	CCT	GTG	GAG	AAG	TCT...
HBS	...	CAC	CTG	ACT	CCT	GAG	GAG	AAG	TCT...
						<u>VAL</u>			

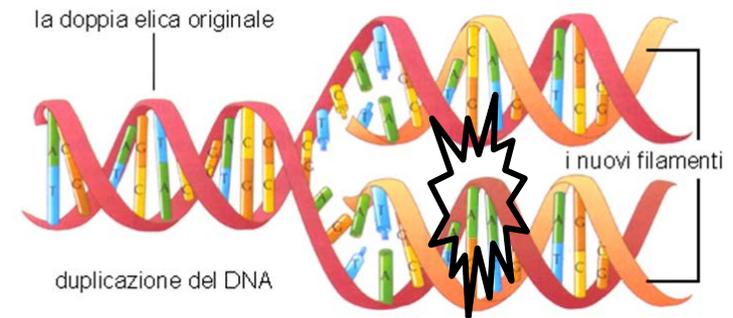
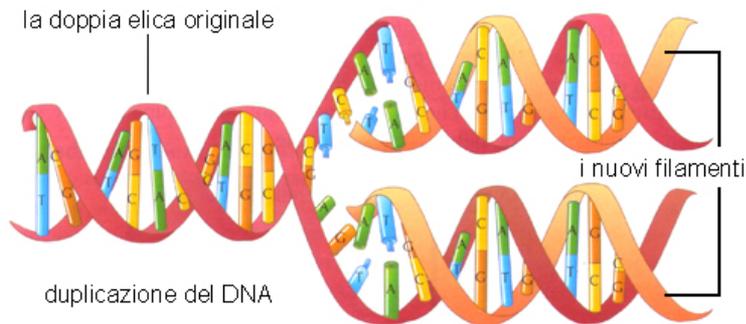


# SNPs o Mutazioni ?

- ✚ Entrambi i termini indicano una variazione di un singolo nucleotide. La differenza è definita dalla **frequenza**:
  - Una variazione che avviene in una popolazione con una frequenza  $>1\%$  si chiama **SNP**.
  - Una variazione che avviene in una popolazione con una frequenza  $<1\%$  viene definita **mutazione**.
- ✚ quando la variazione è collegata a malattia, si chiama mutazione, altrimenti è SNP.

# Meccanismo molecolare

- ✚ Gli SNP (mutazioni) avvengono per l'inserzione di nucleotidi "sbagliati" durante la duplicazione del DNA o la riparazione di danni al DNA

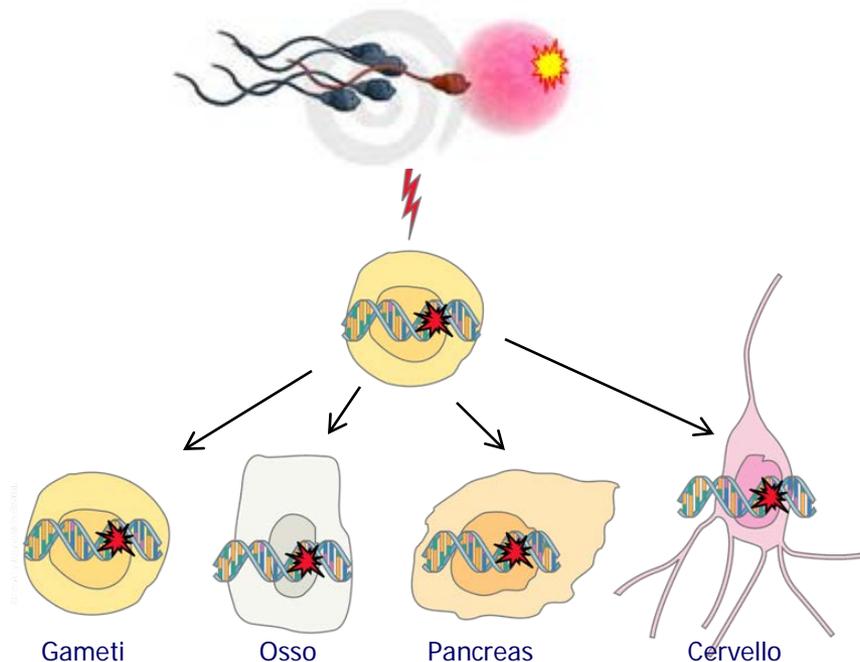


# Conseguenze

- ✦ La maggior parte delle variazioni nucleotidiche avviene in regioni non codificanti del genoma → nessuna proteina viene alterata.
- ✦ Le variazioni nelle regioni codificanti possono generare proteine diverse, ma nella maggior parte dei casi non arrecano danni all'individuo (variabilità individuale)
- ✦ Quando invece portano a mutazioni con effetti dannosi sulla funzione della proteina codificata, le conseguenze dipendono dal tipo cellulare colpito:
  - Cellule germinali (o sessuali): uovo, spermatozoo
  - Cellule somatiche: tutte le altre

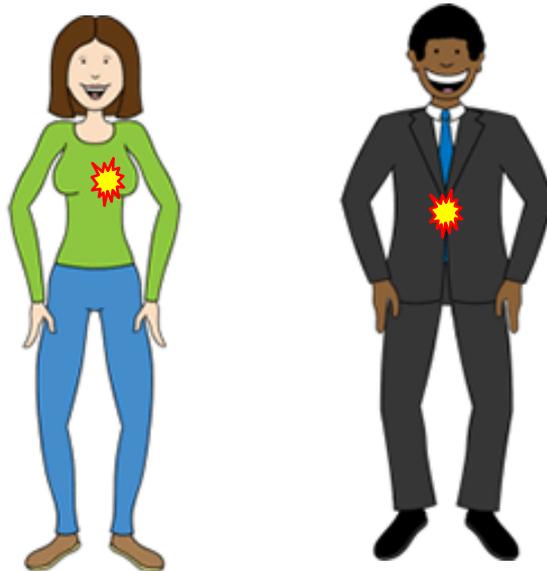
# Mutazioni germinali

- ✚ Se la mutazione avviene nelle cellule germinali (o nelle primissime fasi dello sviluppo embrionale):
  - presente in tutte le cellule dell'individuo
  - trasmissibile alla progenie (ereditabile)



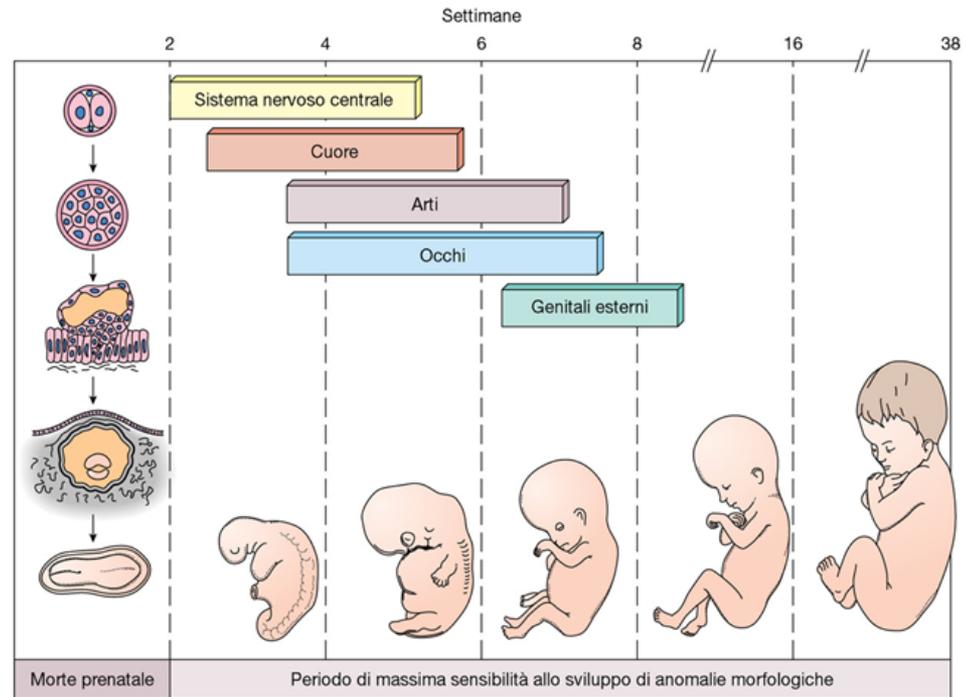
# Mutazioni somatiche

- ✚ Se la mutazione avviene nelle cellule somatiche:
  - presente soltanto nelle cellule che derivano da quella nella quale e' avvenuta la mutazione
  - NON trasmissibile alla progenie



# Ereditario $\neq$ Congenito

- ✚ Mutazioni genetiche possono avvenire nell'organismo in via di sviluppo, ed essere quindi presenti alla nascita.
- ✚ Tanto più precocemente avvengono, maggiore sarà il numero di cellule/organi colpiti
- ✚ Se la mutazione non colpisce le cellule germinali, non sarà trasmissibile.



# Tecniche di analisi

Ci sono moltissime metodiche per analizzare le varianti genetiche, e si possono eseguire in maniera focalizzata (i.e. analizzare soltanto porzioni del genoma) oppure globale (i.e. analizzare genomi interi).

Varianti quantitative (VNTR, STR):

+ RFLP

+ PCR

Varianti qualitative (SNP):

+ sequenziamento del DNA

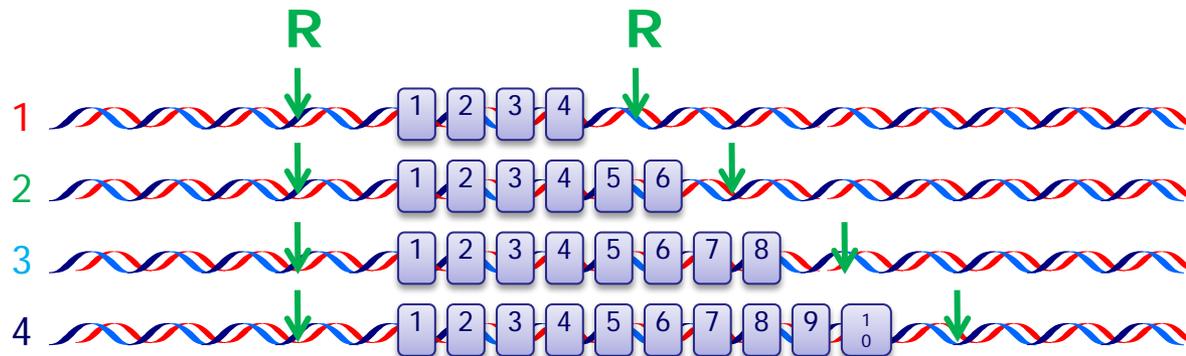


# Varianti quantitative - RFLP

= Restriction Eragment Length Polymorphism

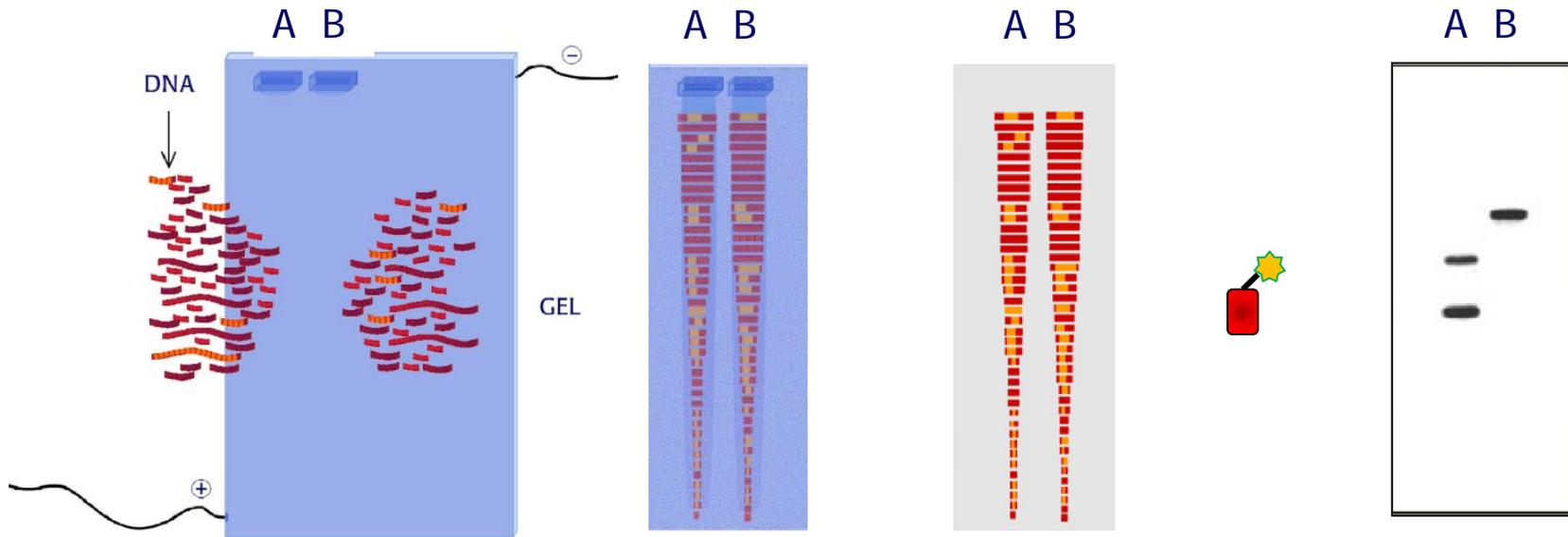
Ideata da Sir Alec Jeffreys nel 1984.

Sfrutta la capacità di alcuni enzimi (*enzimi di restrizione*) di riconoscere specifiche sequenze di nucleotidi, e di tagliare i filamenti di DNA in corrispondenza dei siti di riconoscimento



Il trattamento con enzimi di restrizione che hanno siti di riconoscimento limitrofi a VNTR genera frammenti di lunghezze diverse

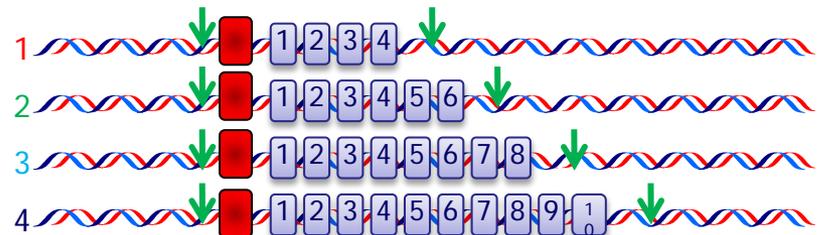
# Varianti quantitative - RFLP



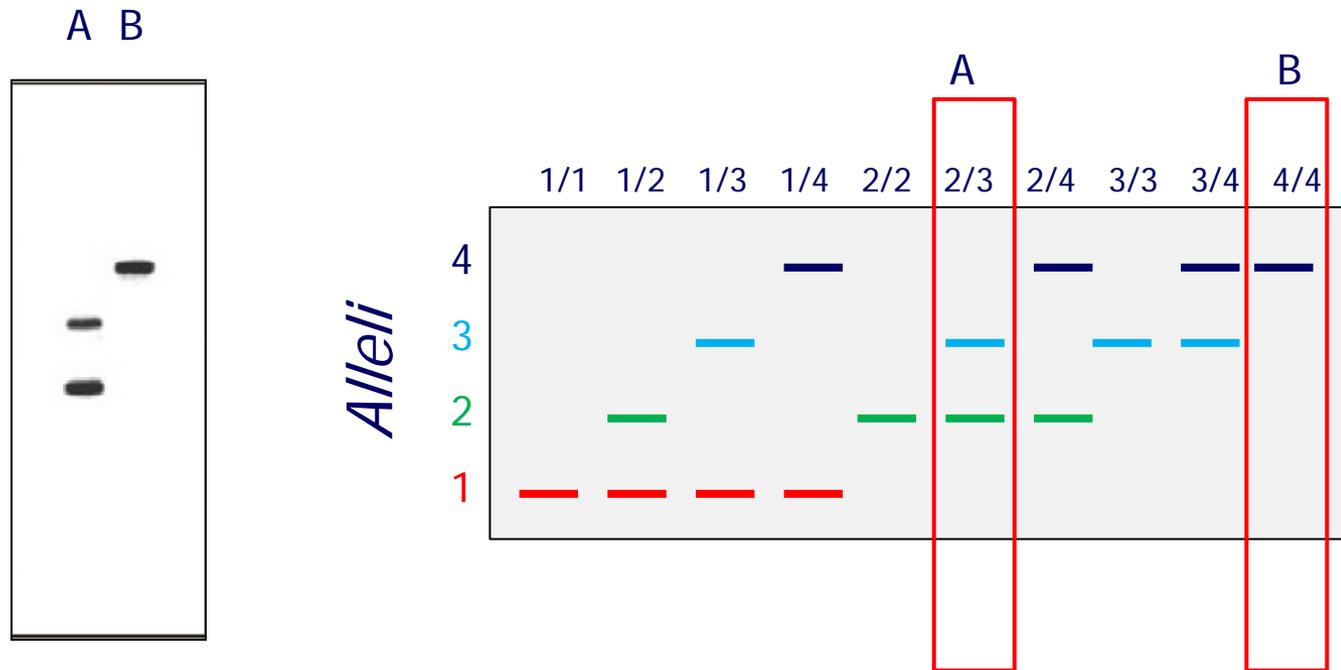
I frammenti vengono separati in base alla dimensione all'interno di un supporto semi-solido (elettroforesi su gel di agarosio)

Vengono trasferiti su un supporto poroso

I frammenti d'interesse vengono identificati mediante "ibridazione" (la proprietà di appaiarsi tipica delle basi azotate degli acidi nucleici)



# Varianti quantitative - RFLP

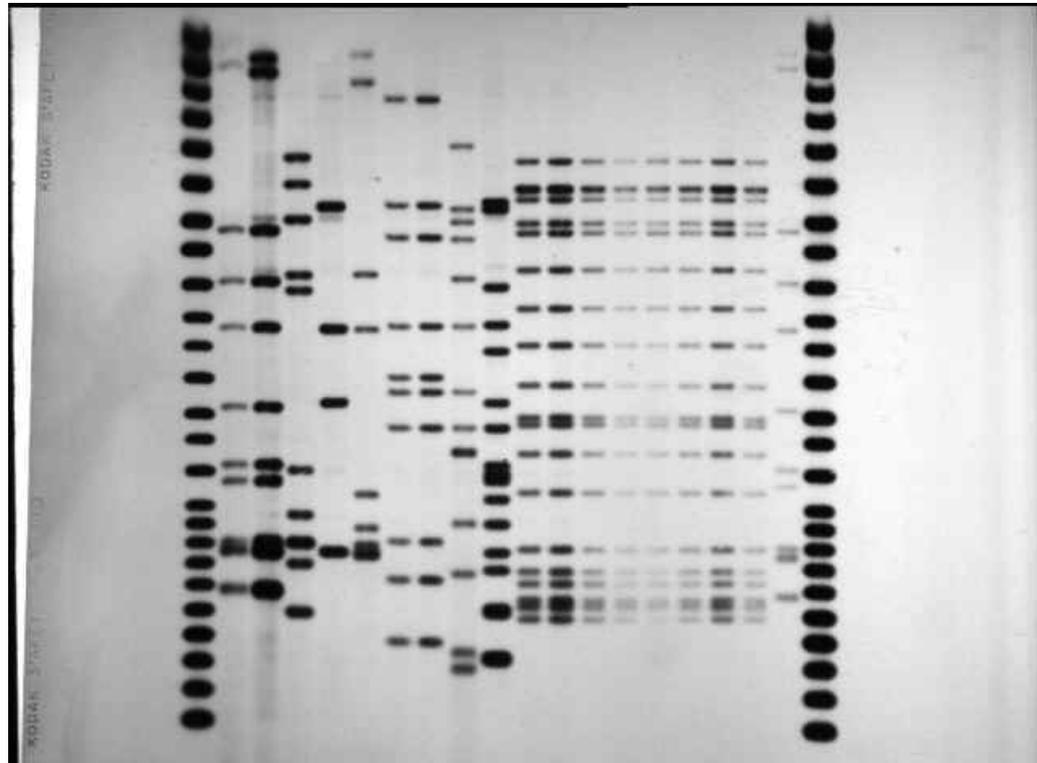
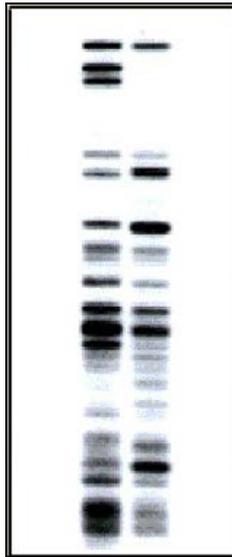


Individuo A = genotipo 2/3

Individuo B = genotipo 4/4

# Varianti quantitative - RFLP

✚ Guardando più VNTR per volta:

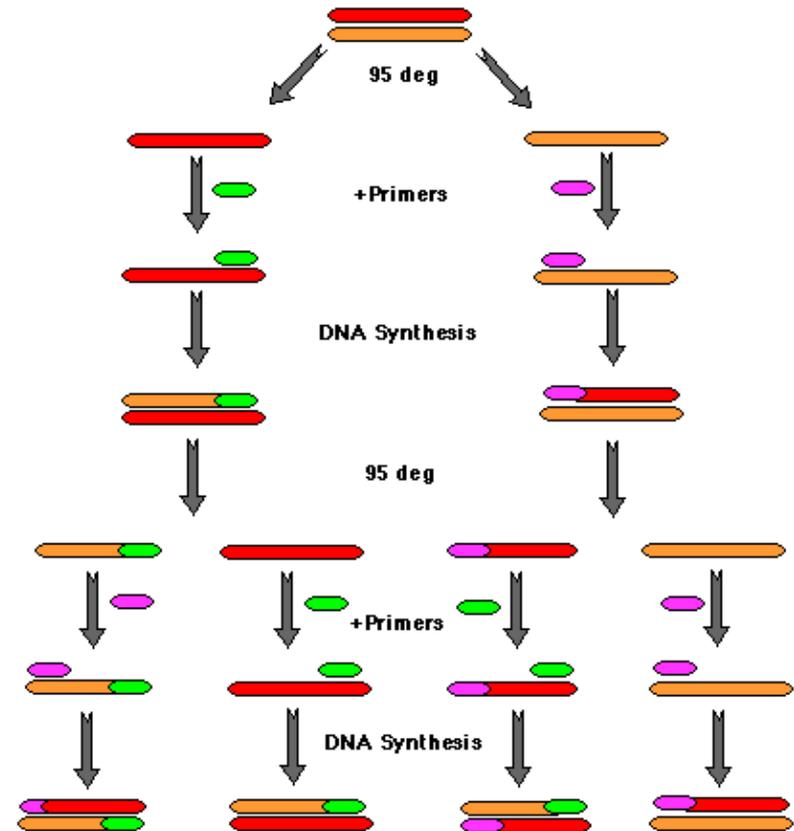


# Varianti quantitative - PCR

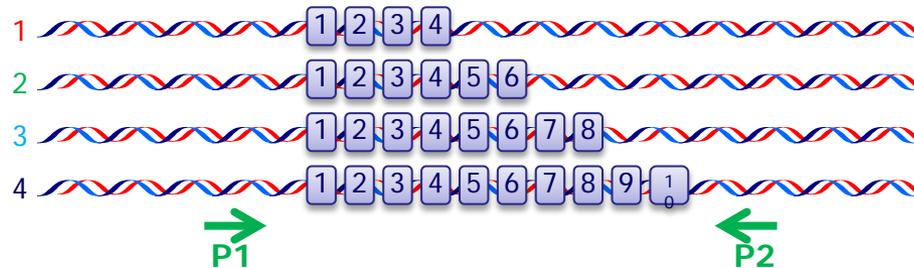
= Polymerase Chain  
Reaction

Inventata nel 1983 da Kary Mullis  
→ 1993, Premio Nobel per la  
Chimica.

Sfrutta la DNA polimerasi, enzima  
deputato alla duplicazione del  
DNA, per amplificare specifiche  
regioni del genoma.



# Varianti quantitative - PCR



Primer = breve sequenza di nucleotidi, sintetizzata in laboratorio, complementare (G→A, T→C) all'estremità della regione d'interesse.

Si amplifica il DNA utilizzando primers "colorati" ai lati del STR  
→ Si amplificano frammenti la cui lunghezza dipende dall'allele.

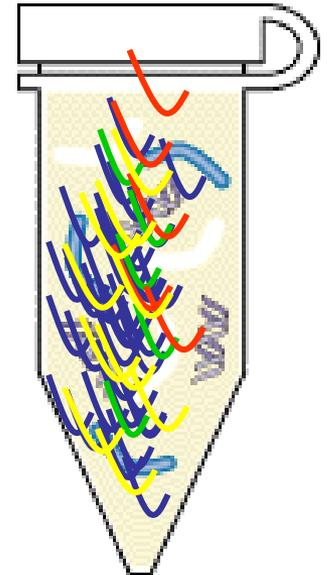
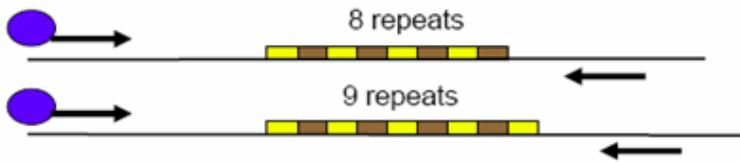
# Varianti quantitative - PCR

Si possono analizzare più regioni contemporaneamente ("multiplex PCR")  
Per potere distinguere i prodotti di amplificazione dei singoli STR, i primers per l'amplificazione sono legati a diverse molecole fluorescenti (generalmente da 3 a 5 colori).

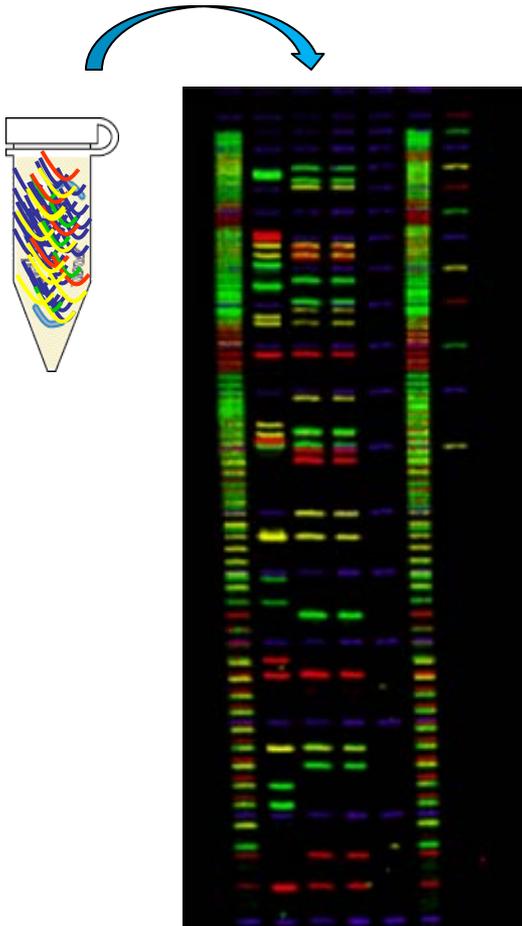
Regione 1



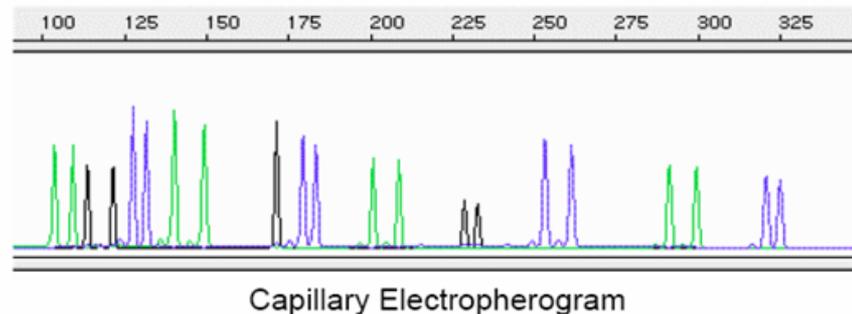
Regione 2



# Varianti quantitative - PCR



- ✚ I frammenti vengono separati per elettroforesi in base alla grandezza.
- ✚ Primers che legano una stessa molecola fluorescente amplificano regioni di diversa dimensione e quindi non confondibili.
- ✚ Il risultato viene trasformato in un grafico (elettroferogramma)



# RFLP versus PCR

RFLP:

Svantaggi: richiede quantità discrete di DNA (50-500ng)  
(→ non sempre utilizzabile a partire da tracce biologiche), tempi abbastanza lunghi.

PCR:

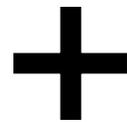
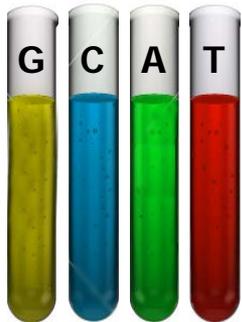
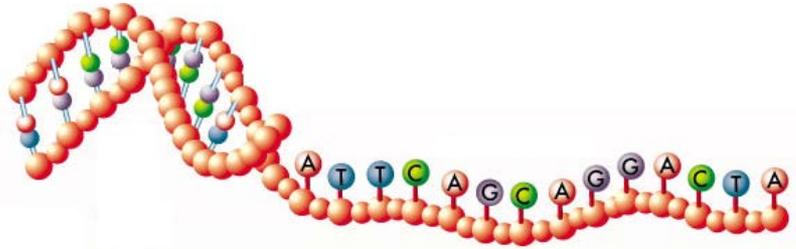
Vantaggi: più rapido, più sensibile, più efficace nei casi di campioni biologici misti

# Sequenziamento del DNA

Il metodo ancora in uso fu messo a punto da Frederic Sanger nel 1977 (vinse il Premio Nobel per la Chimica nel 1980, dopo che lo aveva vinto nel 1958 per la caratterizzazione dell'insulina).

Sfrutta l'attività DNA polimerasi, sostituendo ai normali desossinucleotidi, analoghi modificati.

# Sequenziamento del DNA

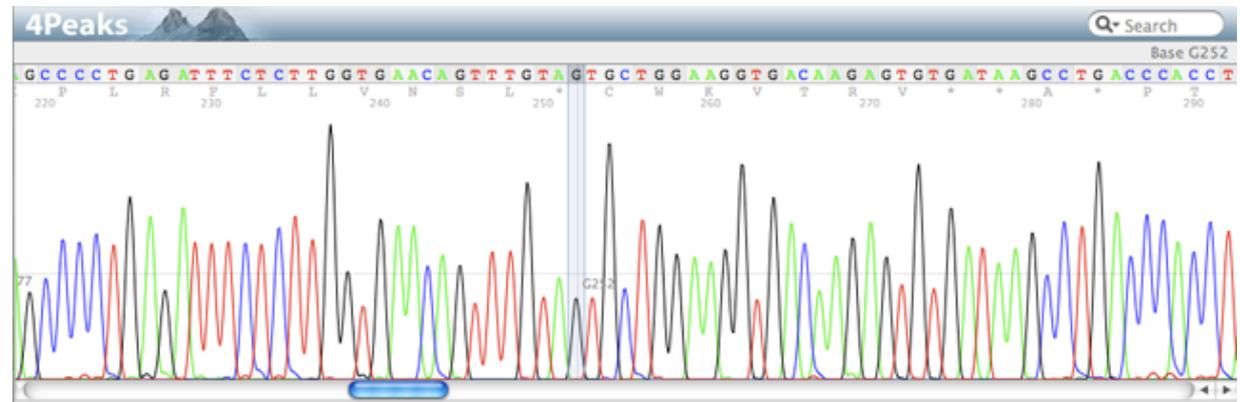
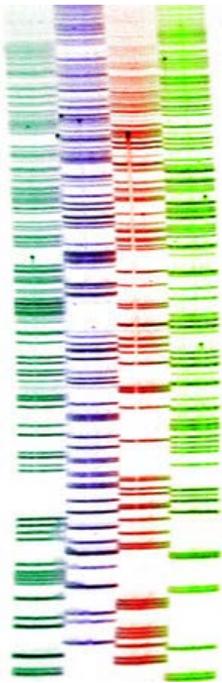


DNA polimerasi

Gel:

● G	GCGAATGCGTCCACACGCTACAGGT <b>G</b>
● T	GCGAATGCGTCCACACGCTACAGGT
● G	GCGAATGCGTCCACACGCTACAG <b>G</b>
● G	GCGAATGCGTCCACACGCTACAG
● A	GCGAATGCGTCCACACGCTAC <b>A</b>
● C	GCGAATGCGTCCACACGCTAC
● A	GCGAATGCGTCCACACGCT <b>A</b>
● T	GCGAATGCGTCCACACGCT
● C	GCGAATGCGTCCACAC <b>G</b>
● G	GCGAATGCGTCCACACG
● C	GCGAATGCGTCCACAC
● A	GCGAATGCGTCCAC <b>A</b>
● A	GCGAATGCGTCCAC <b>A</b>
● C	GCGAATGCGTCCAC
● A	GCGAATGCGTCC <b>A</b>
● C	GCGAATGCGTCC
● C	GCGAATGCGT <b>C</b>
● T	GCGAATGCGT
● G	GCGAATG <b>G</b>
● C	GCGAAT <b>C</b>
● G	GCGAAT <b>G</b>
● T	GCGAAT

# Sequenziamento del DNA



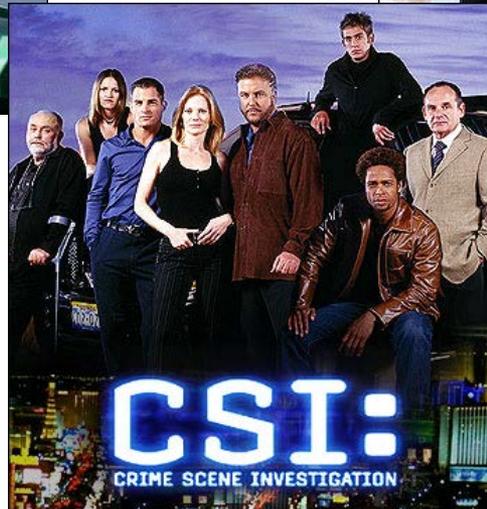
I frammenti di DNA contenenti nucleotidi marcati vengono separati per elettroforesi e convertiti in un elettroferogramma rappresentativo della sequenza nucleotidica

# Applicazioni

L'analisi delle varianti genetiche ha moltissime implicazioni, e viene utilizzata prevalentemente in quattro ambiti:

- ✚ **Analisi forensi**
- ✚ Studi antropologici
- ✚ Correlazione genotipo-malattia
- ✚ Farmacogenomica

# Analisi Forense



→ identificazione

# Fasi nel Processamento di un campione di DNA

Campione prelevato dalla  
scena del crimine

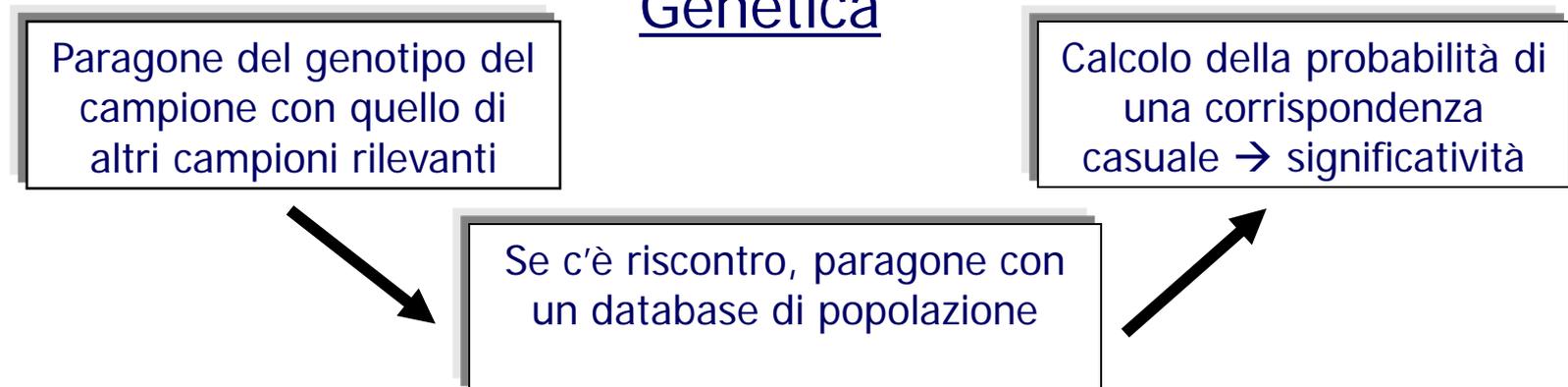
## Biologia



## Tecnologia



## Genetica

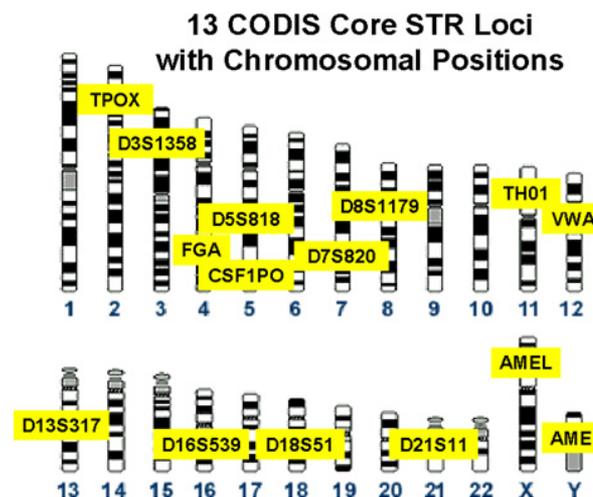


# Banche Dati

- ✚ In alcuni paesi, si stanno generando banche dati contenenti i risultati dei genotipi raccolti in corso di indagini.

**CODIS** = **C**ombined **D**NA **I**ndex **S**ystem del FBI

- ✚ Utilizzato in tutti i casi irrisolti, per identificare eventuali criminali recidivi.
- ✚ Lanciato a ottobre 1998
- ✚ Copre i 50 stati
- ✚ Richiede 13 marcatori STR

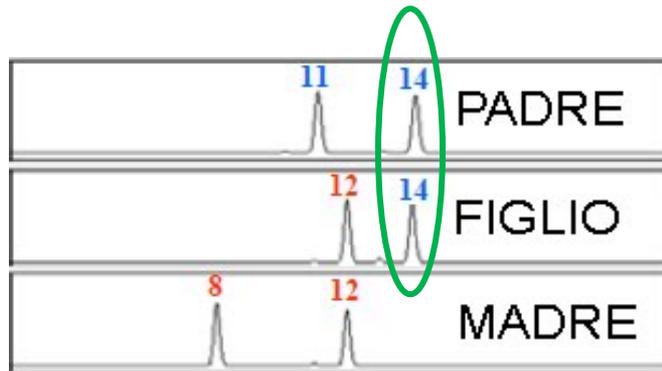


# Analisi di paternità

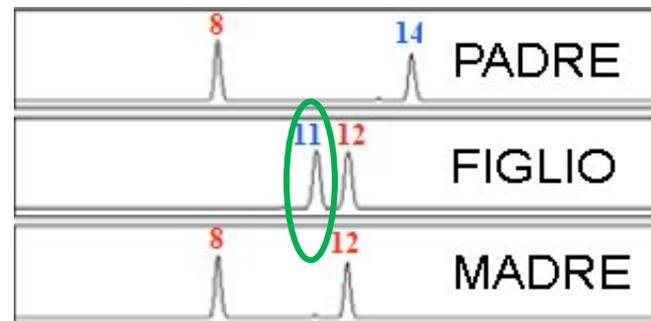


- 1) Valutazione dell'insieme delle caratteristiche che costituiscono il profilo genetico del figlio e della madre.
- 2) Tutte le caratteristiche genetiche del figlio che non sono presenti nella madre devono essere state obbligatoriamente ereditate dal padre biologico.

## Attribuzione



## Esclusione



# Applicazioni

L'analisi delle varianti genetiche ha moltissime implicazioni, e viene utilizzata prevalentemente in quattro ambiti:

✚ Analisi forensi

✚ Studi antropologici

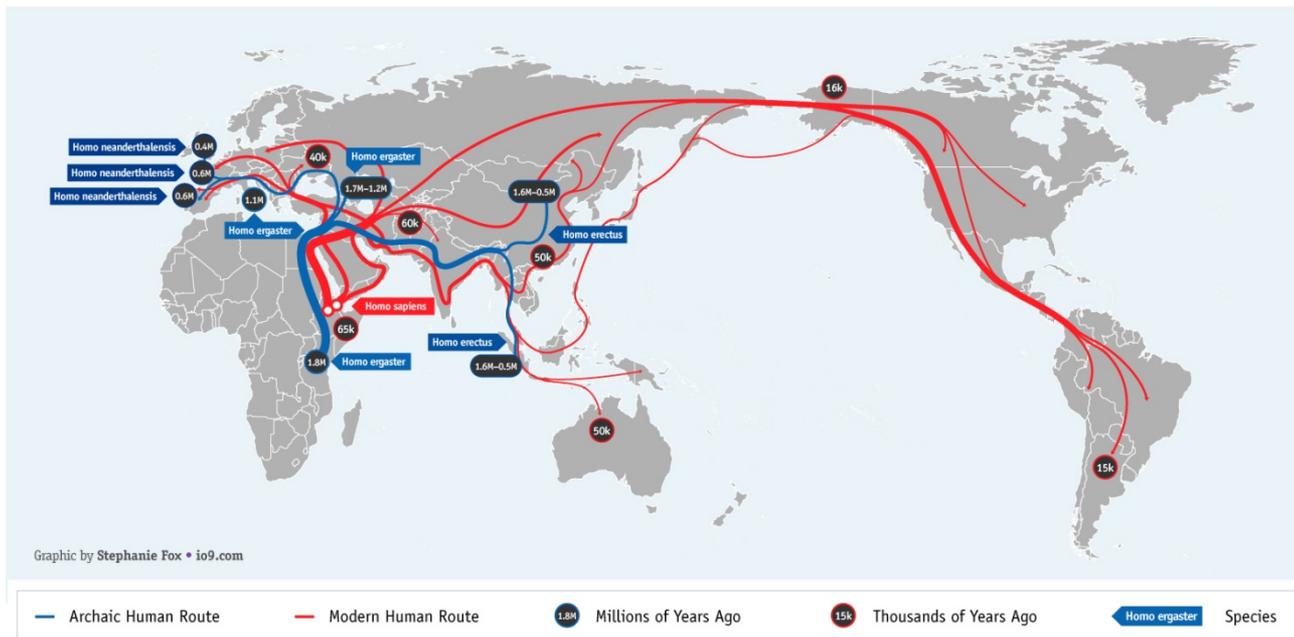
✚ Correlazione genotipo-malattia

✚ Farmacogenomica

# Antropologia

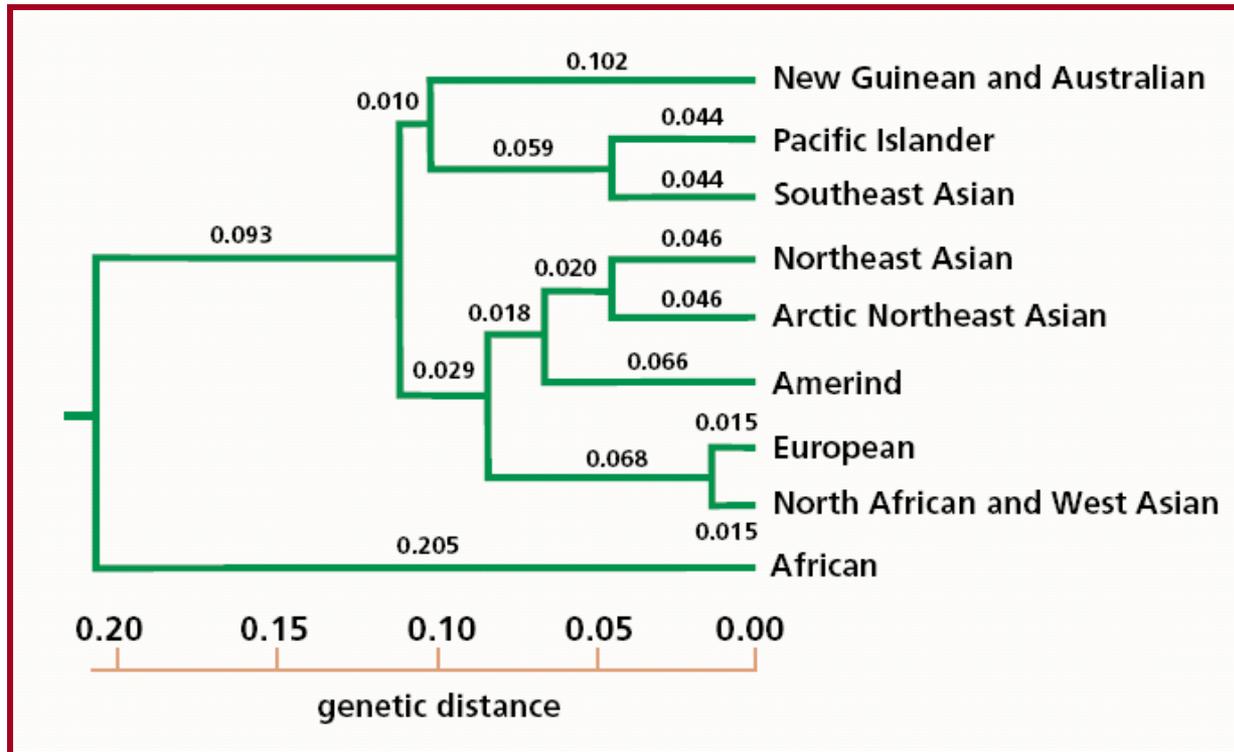
## 🚧 Evoluzione/migrazione dell'uomo

### *The Human Diaspora*



# Popolazione mondiale

- ✚ Polimorfismi di 120 geni
- ✚ 1,915 popolazioni



# Evoluzione delle specie

- ✚ L'analisi di specifici polimorfismi durante l'evoluzione può dare informazioni molecolari riguardanti caratteri complessi
- ✚ Esempio: il linguaggio è un carattere specifico degli umani:

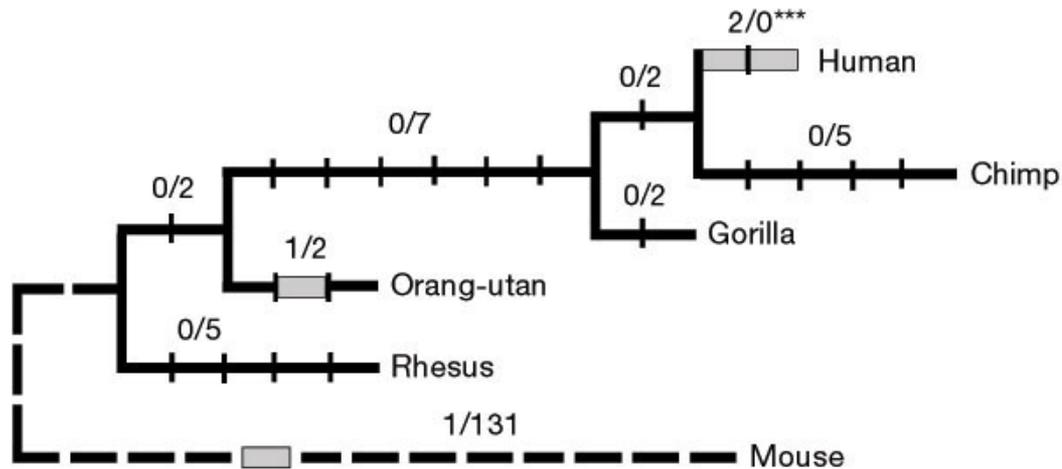
Una mutazione, anche in eterozigosi, del gene FOXP2 si associa con difficoltà marcate nel linguaggio.

→ analisi di FOXP2 durante l'evoluzione.

# Evoluzione delle specie

Sequenza nucleotidica di FOXP2 in 6 specie: Umano, Scimpanzè, Orango, Macaco e topo.

✚ solo 2 variazioni nucleotidiche specifiche dell'uomo.  
Mutazioni in uno di questi due nucleotidi → difficoltà di linguaggio

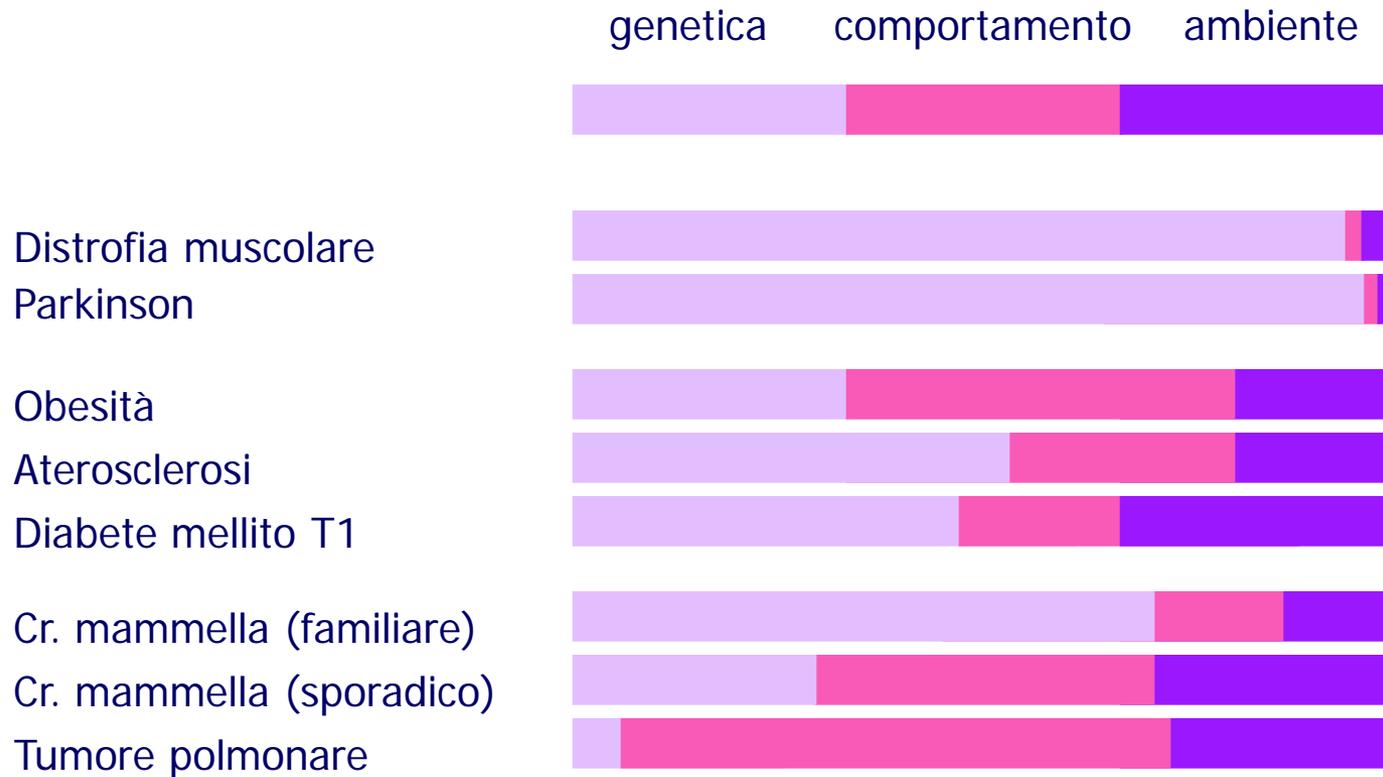


# Applicazioni

L'analisi delle varianti genetiche ha moltissime implicazioni, e viene utilizzata prevalentemente in quattro ambiti:

- ✚ Analisi forensi
- ✚ Studi antropologici
- ✚ **Correlazione genotipo-malattia**
- ✚ Farmacogenomica

# Correlazione genotipo-malattia



Questa regola si applica anche a comportamenti sociali:



# La componente genetica

**Table 1 • Relative frequency of types of mutations underlying disease phenotypes\***

Change	Number	% of total
deletion	6,085	21.8
insertion/duplication	1,911	6.8
complex rearrangement	512	1.8
repeat variations	38	0.1
missense/nonsense	16,441	58.9
splicing	2,727	9.8
regulatory	213	0.8
total	27,027	100.0

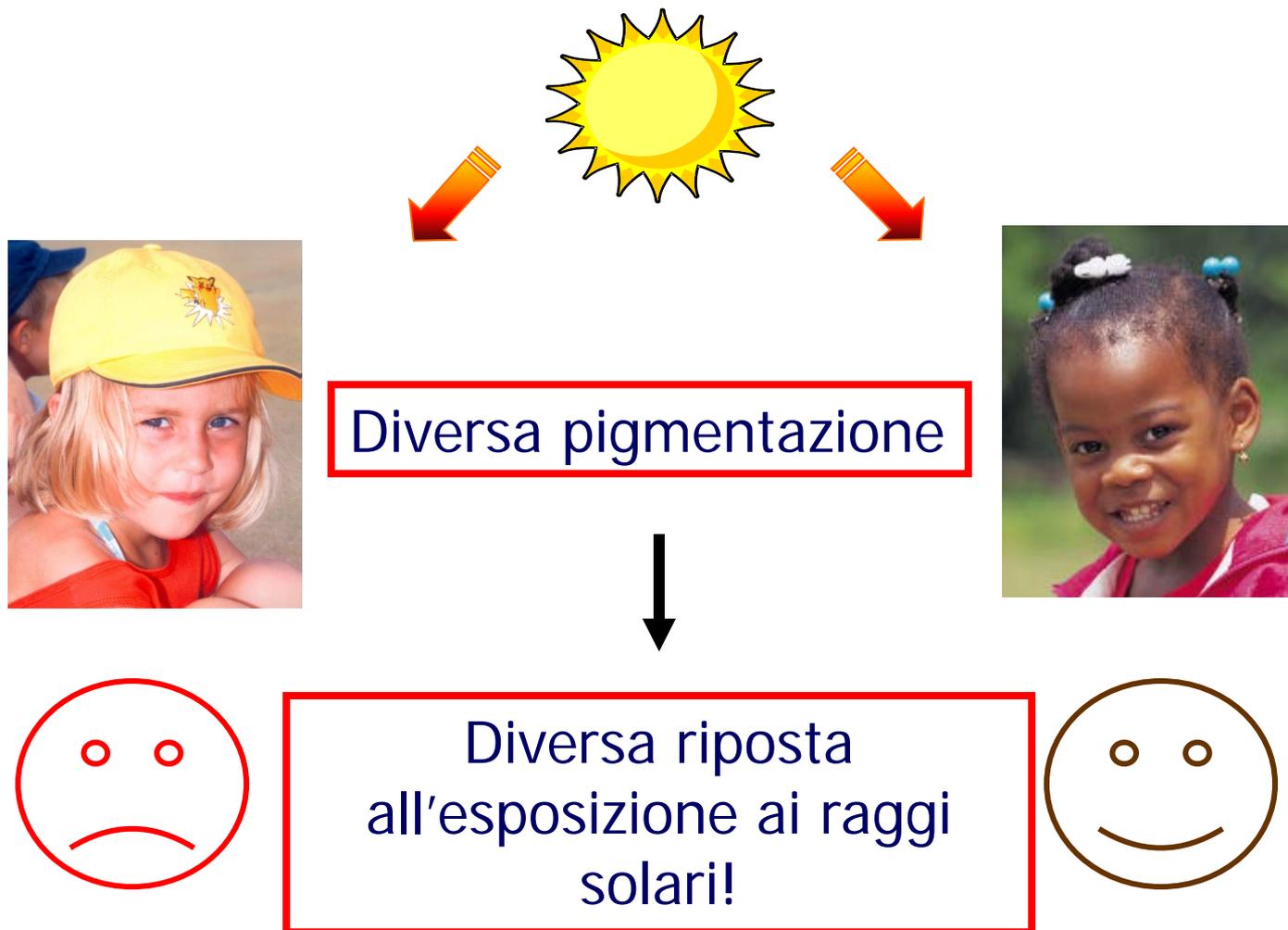
\*Data are from the Human Gene Mutation Database (June 2002).

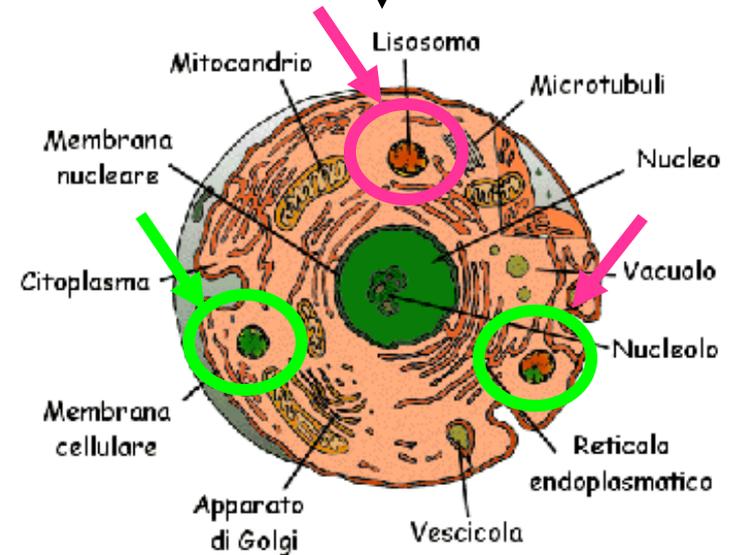
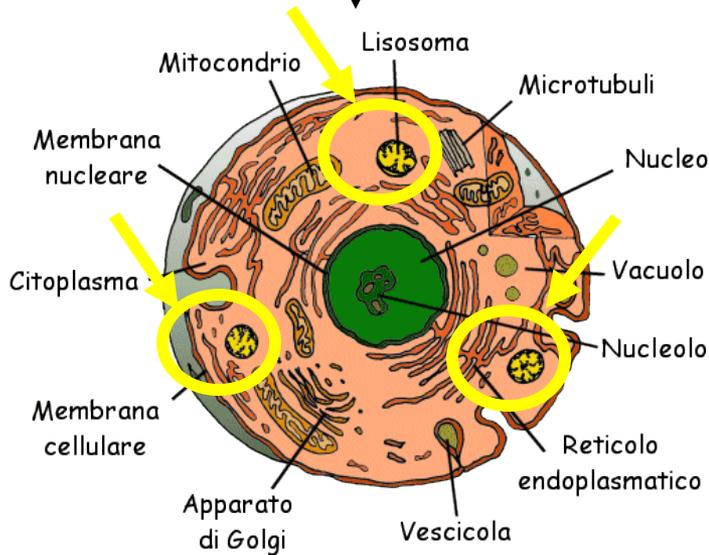
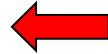
Prevalentemente variazioni di singoli nucleotidi

# Predisposizione genetica

- ✚ Le malattie multifattoriali sono le più frequenti, e nella maggior parte dei casi i geni coinvolti sono tutt'ora ignoti.
- ✚ Conoscere se si è portatori di varianti genetiche che predispongono a specifiche malattie sarebbe un potente strumento per mettere in atto protocolli di prevenzione mirata.

# Predisposizione genetica





Il tipo di risposta ad agenti nocivi è scritta nel DNA  
(quindi è prevedibile)

# Diabete Mellito Tipo 2

Il Diabete Mellito è un gruppo di malattie caratterizzate da un patologico aumento della glicemia e dalla presenza di glucosio nelle urine (glicosuria).

L'elemento caratteristico del DM-T2 è un blocco dell'utilizzo del glucosio dovuto a resistenza all'azione dell'insulina.

Forti fattori predisponenti sono: obesità, ipertensione, ipertrigliceridemia, la presenza di parenti di primo grado affetti dalla malattia.



# Genetica del DM-T2

## Studi sui gemelli

- 90% di concordanza nei monozigoti
- 15% di concordanza nei dizigoti

## Studi sulle famiglie

- 25% dei diabetici ha un familiare diabetico
- Sviluppano il diabete tipo 2:
  - + 15% dei figli di un genitore diabetico
  - + 50% dei figli di 2 genitori diabetici
  - + 15% dei fratelli di 1 diabetico

# DM-T2

Il T2DM è stato definito “l’incubo del genetista”:

- Forte componente genetica, ma fattori genetici ignoti.
- Forte correlazione con l’ambiente, in particolare con l’obesità.

- ✚ Nel 2007 sono stati pubblicati 6 lavori che descrivono l’identificazione di SNP associate a DM-T2 su un totale di circa 55,000 individui di cui 18,000 affetti da T2DM.
- ✚ Gli autori si sono poi scambiati i dati, e sono emerse 11 varianti geniche associate.

# Varianti associate a DM-T2

Table 1 | Details of 11 type 2 diabetes gene regions

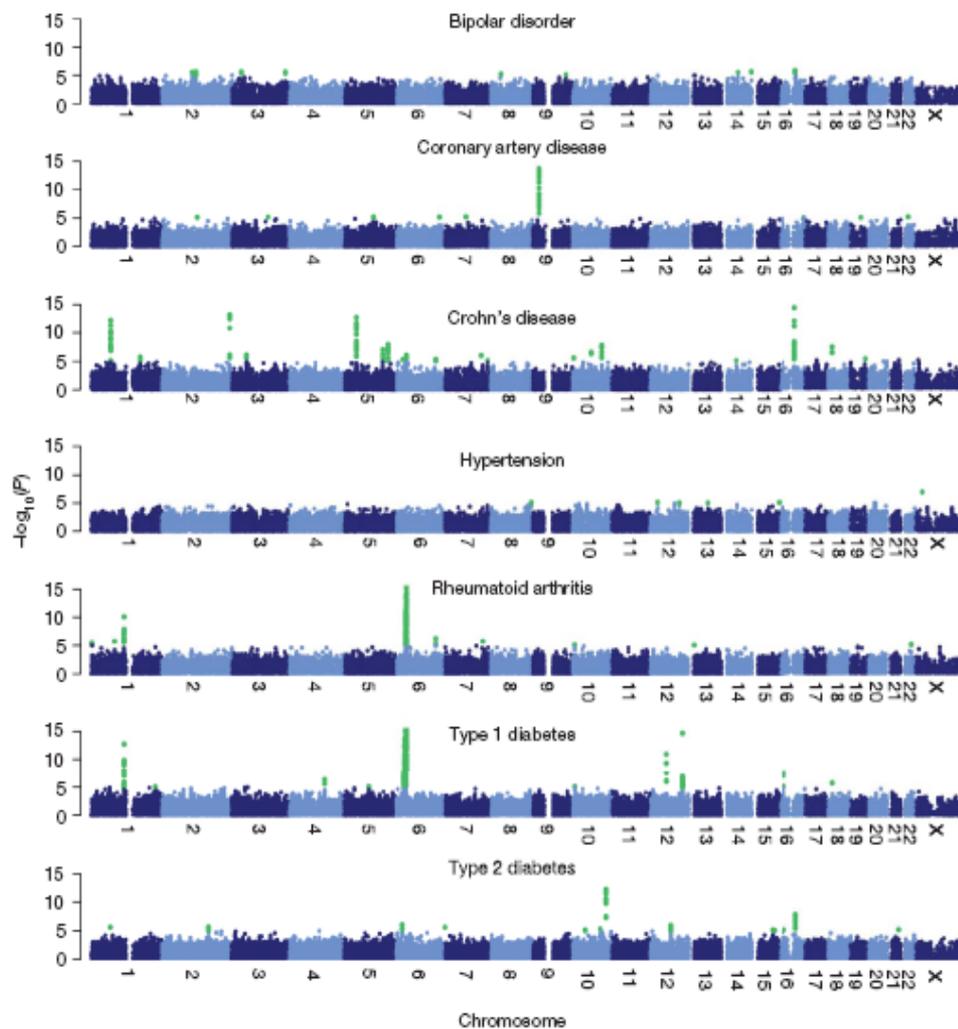
Example variant	Closest gene	Mode of identification	Previous evidence	Current evidence (p value)*	Additional evidence from human physiology	Odds ratio (per allele)*	RAF (UK)	N <sup>‡</sup>	Rank in UK GWAS <sup>§</sup>
rs1801282 (P12A)	PPARG	Candidate	Monogenic + drug target	2×10 <sup>-6</sup>	Nothing consistent	1.14 (1.08–1.20)	0.87	>20,000	786
rs5215 (E23K)	KCNJ11	Candidate	Monogenic + drug target	5×10 <sup>-11</sup>	Alters insulin secretion in general population	1.14 (1.10–1.19)	0.35	15,600	799
rs7901695	TCF7L2	Region-wide	None	1×10 <sup>-44</sup>	Alters insulin secretion in general population	1.37 (1.31–1.43)	0.31	2,760	2
rs4430796	TCF2	Candidate	Monogenic	8×10 <sup>-10</sup>	Nothing consistent	1.10 (1.07–1.14)	0.47	>20,000	N/C
rs10010131	WFS1	Candidate	Monogenic	1×10 <sup>-7</sup>	Nothing consistent	1.11 (1.08–1.16)	0.60	>20,000	26017
rs1111875	HHEX-IDE	Genome-wide	Some, e.g. HHEX KO mouse has disrupted pancreatic development	7×10 <sup>-17</sup>	Early studies indicate altered insulin secretion in general population	1.15 (1.10–1.19)	0.65	12,800	32
rs13266634	SLC30A8	Genome-wide	None	1×10 <sup>-19</sup>	Early studies indicate altered insulin secretion in general population	1.15 (1.12–1.19)	0.69	14,400	N/C
rs10946398	CDKAL1	Genome-wide	None	2×10 <sup>-18</sup>	Early studies indicate altered insulin secretion in general population	1.14 (1.11–1.17)	0.32	16,200	65
rs10811661	CDKN2A-2B	Genome-wide	Some—CDKN2A KO mouse has reduced islet proliferation	8×10 <sup>-15</sup>	Nothing consistent	1.20 (1.14–1.25)	0.83	12,400	272
rs4402960	IGF2BP2	Genome-wide	Some—binds insulin-like growth factor mRNA	9×10 <sup>-16</sup>	Nothing consistent	1.14 (1.11–1.18)	0.32	16,200	1,026
rs8050136	FTO	Genome-wide	None	1×10 <sup>-12</sup>	Alters BMI in general population	1.17 (1.12–1.22)	0.40	10,400	12

Di particolare interesse è la correlazione tra DM-T2 e FTO

# FTO, obesità e DM-T2

- ✚ Varianti genetiche di FTO si associano ad un aumento dell'indice di massa corporea (IMC), considerato come indice di obesità.
- ✚ Anche l'aumento del IMC è considerato un tratto **multifattoriale**
- ✚ Studio di 38,759 individui, molti in comune con gli studi sul DM-T2
- ✚ L'associazione tra DM-T2 e varianti di FTO era in realtà dovuta all'associazione tra DM-T2 e obesità.

# Utilizzo di SNP analisi per lo studio di patologie multifattoriali



Disordine Bipolare

Coronaropatia

M. Di Crohn

Ipertensione

Artrite Reumatoide

T1DM

T2DM

 → l'analisi dei dati di associazione tra varianti genetiche e patologie multifattoriali è complessa, e deve tenere conto dei vari componenti di un fenotipo (es. FTO si associa a obesità, che è a sua volta frequentemente correlata con DM-T2).

# Applicazioni

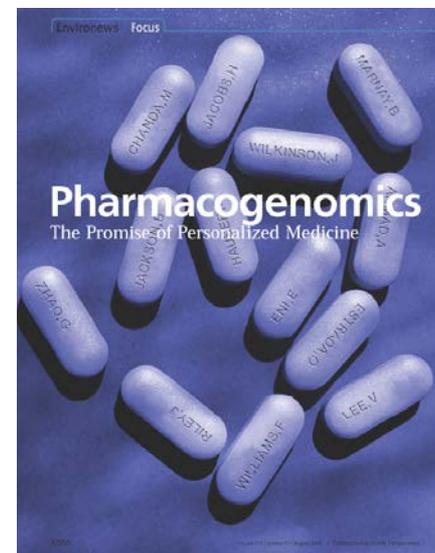
L'analisi delle varianti genetiche ha moltissime implicazioni, e viene utilizzata prevalentemente in quattro ambiti:

- ✚ Analisi forensi
- ✚ Studi antropologici
- ✚ Correlazione genotipo-malattia
- ✚ **Farmacogenomica**

# Farmacogenomica

“**La farmacogenomica** è una scienza che studia le variazioni genetiche di quei geni che determinano la risposta ai farmaci ed esplora la possibilità di utilizzare tali variazioni per prevedere se un paziente trattato con un farmaco avrà una risposta soddisfacente, scarsa o nulla.”

Obiettivo: medicina personalizzata



# Cos'è la Medicina Personalizzata?

La medicina personalizzata è quella pratica che metterà i clinici in condizioni di prescrivere:

- ✚ Il farmaco *giusto*
- ✚ Alla dose *giusta*
- ✚ Per la malattia *giusta*
- ✚ Al paziente *giusto*

E di saperlo PRIMA che il paziente prenda il farmaco

# Reazioni avverse ai Farmaci (ADR)

✚ Evento estremamente frequente



Table 3.—ADR Incidence According to ADR Severity\*

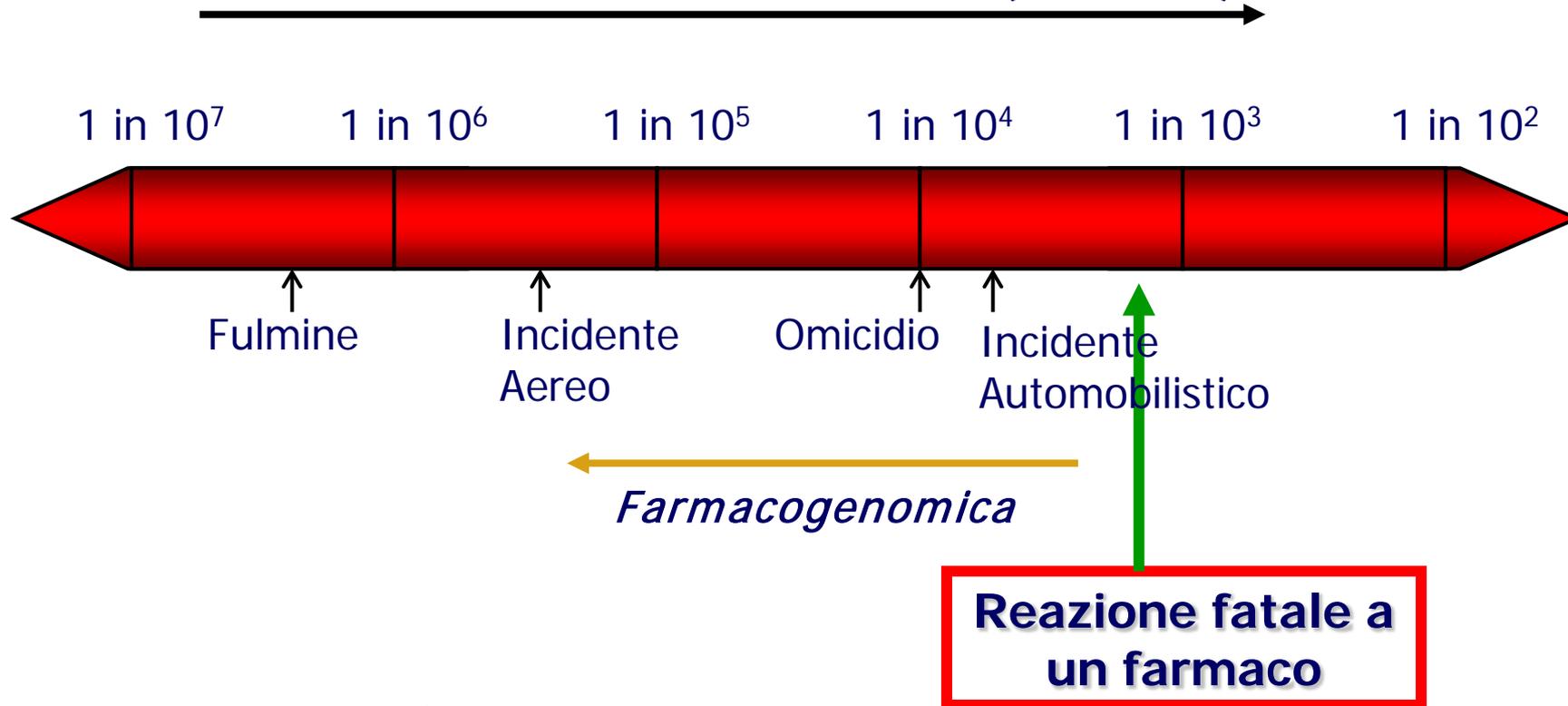
ADR Group	No. of Studies	Total Patients Studied	Incidence of ADRs, %	95% CI
<b>ADRs in Patients While in the Hospital (ADRIn)</b>				
All severities	18	34 463	10.9	7.9-13.9
Serious	12	22 502	2.1	1.9-2.3
Fatal	10	28 872	0.19	0.13-0.26
<b>Patients Admitted to the Hospital Due to an ADR (ADRAd)</b>				
Serious†	21	28 017	4.7	3.1-6.2
Fatal	6	17 753	0.13	0.04-0.21
<b>Overall ADR Incidence (ADRIn + ADRAd)‡</b>				
All severities	39	62 480	15.1	12.0-18.1
Serious	33	50 519	6.7	5.2-8.2
Fatal	16	46 625	0.32	0.23-0.41

\*ADR indicates adverse drug reaction; ADRIn, an ADR occurring in patients while in the hospital; CI, confidence interval; and ADRAd, an ADR causing admission to the hospital.

†By definition, all ADRAd are serious, hence there is no “All Severities” category for ADRAd.

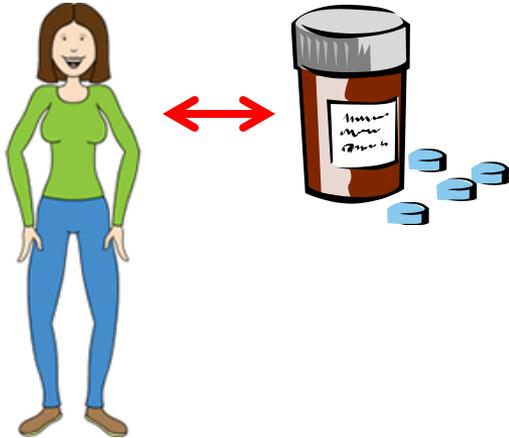
‡Overall incidence is adjusted to avoid double counting (see “Methods” section).

## Aumento Rischio Mortalità (annuale)



Da: Consumer Reports, 9/99

# Metabolismo dei Farmaci



Farmaco inattivo (profarmaco)

- Attivato da enzimi
- Trasportato in vari tessuti
- Interagisce con molecole
- +/- inattivato
- Eliminato

Farmaco attivo

- Trasportato in vari tessuti
- Interagisce con molecole bersaglio
- Inattivato
- Eliminato

# Metabolismo dei Farmaci

**Farmacocinetica** = **ciò che l'organismo fa al farmaco.**

- ✚ Cinetica del farmaco nel sangue e nei tessuti
- ✚ Enzimi che metabolizzano il farmaco
- ✚ Molecole che trasportano il farmaco

**Farmacodinamica** = **ciò che il farmaco fa all'organismo.**

- ✚ Include sia effetti benefici che tossici (molecole bersaglio del farmaco: volute e non)

Variazioni genetiche possono influenzare le risposte ai farmaci

# Varianti genetiche e risposta ai farmaci

- ✚ Nel maggio 1977, cinque membri del laboratorio di Robert Smith del St Mary's Hospital Medical School a Londra hanno assunto la *debrisochina* nel contesto di un progetto di ricerca sull'ipertensione.

Smith si è sentito male, ed è collassato per una crisi ipotensiva.



collaboratori



Capo

# Varianti genetiche e risposta ai farmaci

- ✚ Normalmente, la debrisoquina viene rapidamente trasformata in metaboliti inattivi che sono eliminati attraverso le urine.
- ✚ L'urina del Prof. Smith era priva di metaboliti della debrisoquina.
- ✚ Lo studio è stato esteso ad altri membri della famiglia di Smith, alcuni dei quali hanno presentato analoghe reazioni avverse al farmaco.

Questo lavoro ha portato all'identificazione del gene che è responsabile del metabolismo della debrisoquina, CYP2D6, e alla caratterizzazione dei polimorfismi che eliminano l'attività della CYP2D6.

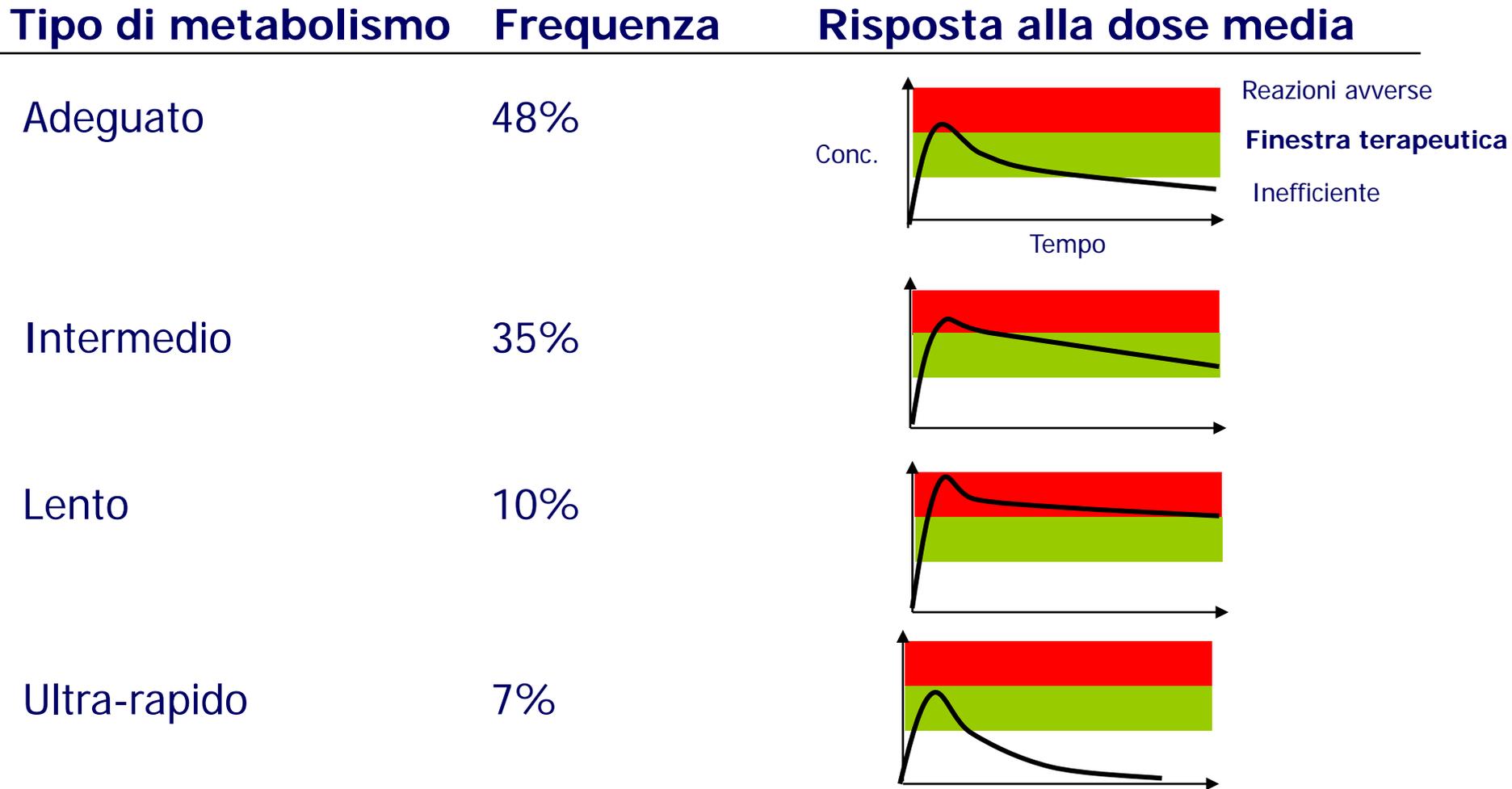
# CYP2D6

- ✚ Enzima coinvolto nel metabolismo di circa il 25% dei farmaci
- ✚ Sono state finora identificate >70 varianti del gene CYP2D6, incluse mutazioni, delezioni e duplicazioni.

# Genotipo → Fenotipo

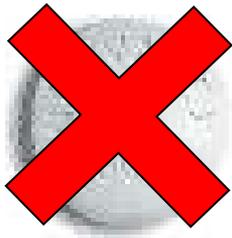
- ✦ Metabolizzatori adeguati (EM) – Possiedono almeno un allele funzionante
- ✦ Metabolizzatori Intermedi (IM) – Possiedono un allele con attività ridotta e un allele nullo
- ✦ Metabolizzatori Lenti (PM) – possiedono due alleli nulli (assenza di attività enzimatica)
- ✦ Metabolizzatori Ultrarapidi (UM) – possiedono multipli alleli funzionali (3-13) e hanno un eccesso di attività enzimatica

# Concentrazione di farmaco e genotipo



# Tamoxifen

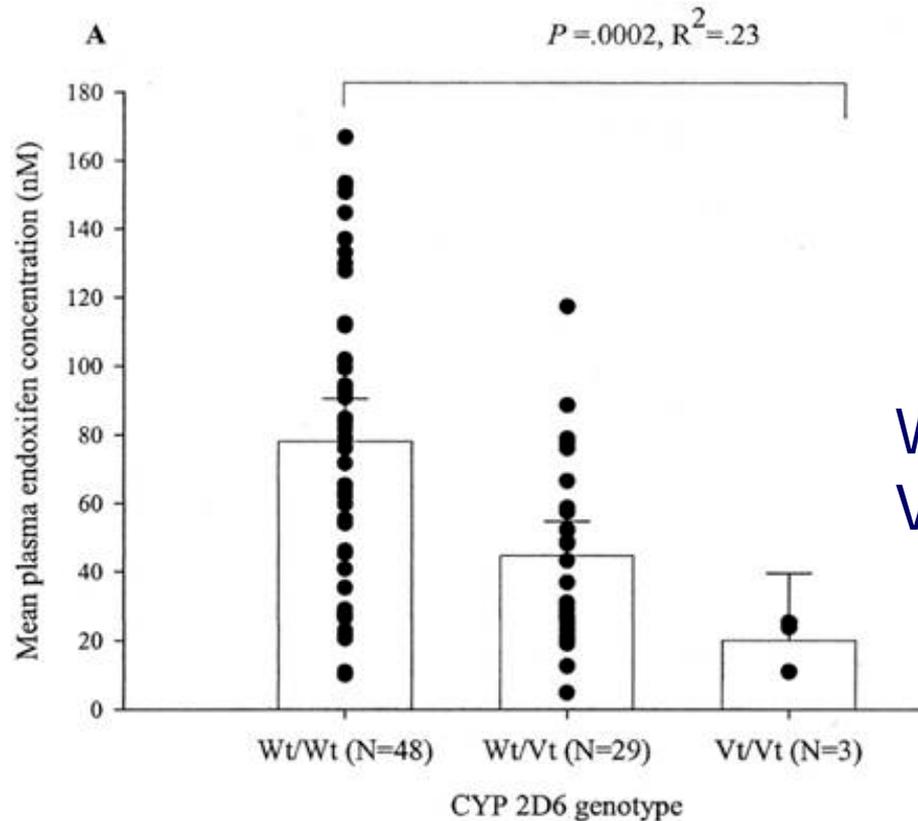
- ✚ Farmaco utilizzato in donne già operate per tumore al seno per prevenire la ripresa della malattia (recidiva).



CYP2D6



# Dosaggio di farmaco attivo



Wt = CYP2D6 nativo  
Vt = CYP2D6 variante

# Tamoxifen

- ✚ Questi studi sono stati presentati e discussi presso la **Food and Drug Administration**, con conseguente raccomandazione unanime di cambiare l'etichetta del tamoxifen. Questo cambiamento dovrebbe includere:
  - Informazione sull'aumentato rischio di interazioni con altri farmaci metabolizzati da CYP2D6
  - Menzione della possibilità di effettuare test di genotipo CYP2D6 prima d'intraprendere terapia con tamoxifen.

# Varanti genetiche associate a risposta ai farmaci

Gene	Name	Allele*	Associated phenotype
<b>Drug target/pathway protein</b>			
ACE	Angiotensin-I converting enzyme	Ins/del	Del/Del associated with decreased proteinuria in response to ACE inhibitor treatment for renal disease
ADRB1	β-1 adrenergic receptor	Arg389Gly	Arg/Arg homozygotes have greater response to β-adrenergic receptor antagonists
ADRB2	β-2 adrenergic receptor	Arg16Gly	Gly allele associated with decreased response to albuterol and salbutamol
AGT	Angiotensinogen	Met236Thr	Thr allele associated with greater reduction of blood pressure and greater decrease in left ventricular mass with anti-hypertensive treatment
AGTR1	Angiotensin-II receptor type 1	A1166C	C allele associated with greater response to angiotensin-II receptor antagonists
ALOX5	Arachidonate 5-lipoxygenase	Promoter VNTR (wt = 5; non-wt = 3,4 or 6)	Non-wt homozygotes have decreased response to 5-lipoxygenase inhibitor and leukotriene receptor antagonist
BDKRB2	Bradykinin receptor B2	C-58T	T allele associated with ACE inhibitor-related cough
CETP	Cholesteryl ester transfer protein	TaqIB polymorphism B1/B2	B1/B1 associated with increased response to pravastatin and atorvastatin
DRD2	Dopamine D2 receptor gene	3' UTR Taq1A A1/A2	A1 associated with greater positive response to the antipsychotics haloperidol and remonapride
DRD3	Dopamine D3 receptor gene	Ser9Gly	Gly associated with response to clozapine; Ser/Ser with non-response
DRD4	Dopamine D4 receptor gene	Exon 3 VNTR (twofold to sevenfold repeat)	7 allele significantly less frequent than 4 allele in patients responding to typical neuroleptics (compared with patients responding to clozapine); 4/4 homozygotes have a higher rate of good response to neuroleptics.
GNB3	Guanine nucleotide-binding protein (G protein), β-polypeptide 3	Exon 10 C825T	CC associated with response to antidepressants
GRN2B	Glutamate receptor, ionotropic, N-methyl D-aspartate 2B	C2864T	CC associated with higher mean clozapine dosage for schizophrenia
HTR2A	5-hydroxytryptamine (serotonin) receptor 2A	T102C	CC or C associated with susceptibility to tardive dyskinesia in response to antipsychotic drugs
HTR2A	5-hydroxytryptamine (serotonin) receptor 2A	His452Tyr	Tyr associated with reduced response to the antipsychotic clozapine
LPC	Hepatic lipase	C-614T	CC genotype associated with increased response to statins
MTHFR	5,10-methylene-tetrahydrofolate reductase	C677T	TT patients have increased toxicity for methotrexate therapy
SLC6A3	Dopamine transporter	3' VNTR 10-repeat allele	10/10 genotype associated with poor response to methylphenidate for treatment of attention deficit/hyperactivity disorder
SLC6A4	Serotonin transporter (5-HTT)	Promoter Ins/del (long/short)	long/long better response to fluoxetine or paroxetine than short/short
TPH1	Tryptophan hydroxylase1	A218C	The A218C polymorphism has been associated with response to anti-depressants
TYMS	Thymidylate synthase	TSER*2/*3	*3 reduced response to 5-fluorouracil treatment compared with *2; *3/*3 requires a higher dose
<b>Drug transporter</b>			
ABCB1	MDR1, P-glycoprotein 1	C3435T	TT patients are less likely to have drug-resistance epilepsy than CC and have increased immune recovery after the initiation of antiretroviral treatment

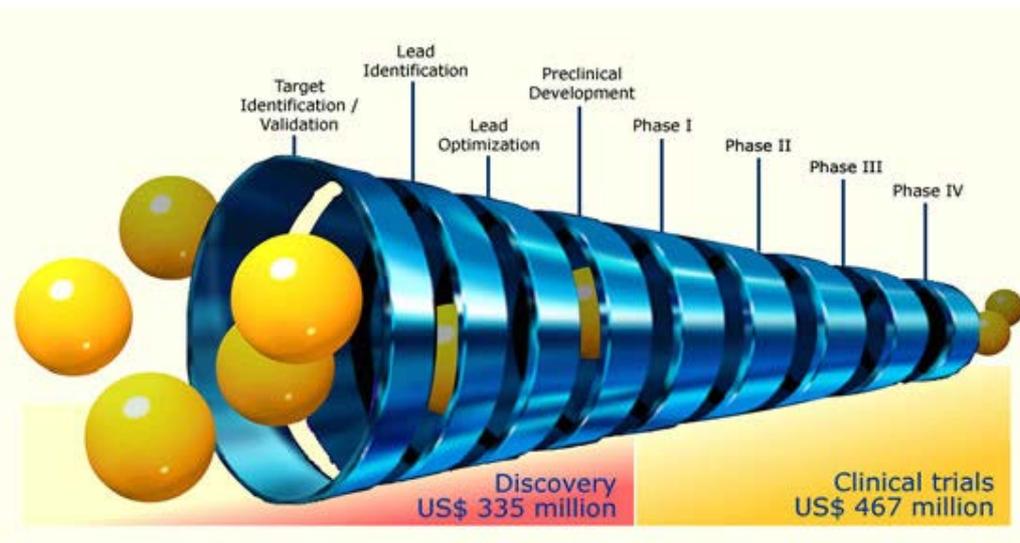
Gene	Name	Allele*	Associated phenotype
<b>Metabolism</b>			
BCHE	Butyrylcholinesterase	Several mutations including Asp70Gly (dibucaine or atypical variant) and Ala539Thr (K allele)	Variants associated with adverse effects in response to succinylcholine
COMT	Catechol O-methyltransferase	Val108/135Met	Met associated with higher daily neuroleptic dosage and poor response
CYP2C19	Cytochrome p450 2C19	*2,*3	Extensive metabolizers have decreased response to omeprazole for <i>Helicobacter pylori</i> infection
CYP2C9	Cytochrome p450 2C9	*2,*3	Non-wt alleles associated with reduced warfarin daily dose requirement
CYP2D6	Cytochrome p450 2D6	Many inactive alleles, such as *3, *4 and *5	Non-wt alleles associated with susceptibility to tardive dyskinesia in response to antipsychotics
DPYD	Dihydropyrimidine dehydrogenase (DPD)	DPYD*2A (VS14+1G>A)	*2A associated with severe toxicity and fatal outcomes for 5-fluorouracil treatment
GSTM1	Glutathione S-transferase M1	GSTM1-null	Null carriers have increased survival time and a progression-free interval following paclitaxel and cisplatin treatment for ovarian cancer; decreased risk of relapse for cytotoxic therapy for leukaemia
GSTM3	Glutathione S-transferase M3 (brain)	GSTM3*A/GSTM3*B	GSTM1*0/GSTM3*A haplotype less likely to show a beneficial response to D-penicillamine in rheumatoid arthritis; *3A has increased risk of cisplatin ototoxicity
GSTP1	Glutathione S-transferase-α	Ile105Val	Val associated with increased survival for 5-fluorouracil and oxaliplatin therapy for colorectal cancer, and following therapy for multiple myeloma
GSTT1	Glutathione S-transferase-γ1	GSTT1-null (homozygote frequent; deletion)	GSTT1 associated with susceptibility to tacrine hepatotoxicity and troglitazone hepatotoxicity in combination with GSTM1-null allele
NAT2	N-acetyltransferase 2	Slow acetylator alleles include NAT2*5B, NAT2*6A, NAT2*7A or B, NAT2*14A or B	Slow-acetylator status of NAT2 increased risk of antituberculous drug-induced hepatotoxicity
TPMT	Thiopurine methyltransferase	TPMT*2, TPMT*3A, TPMT*3C	Homozygotes for non-wt alleles at high risk of severe haematopoietic toxicity after thiopurine treatment; heterozygotes intermediate risk of dose-limiting toxicity
UGT1A1	UDP-glucuronosyltransferase 1A1	UGT1A1*28	*28 associated with increased chance of developing diarrhoea and leucopenia during irinotecan therapy
<b>Other</b>			
ADD1	Adducin 1 (α)	Gly460Ttp	Ttp associated with increased response to diuretics in hypertensives (hydrochlorothiazide)
APOE	Apolipoprotein E	E4	E4 allele associated with lesser response to statins
FCGR3A	FcγRIIIa receptor protein	Phe158Val	Val/Val patients greater response to rituximab for non-Hodgkin lymphoma and higher complete remission rate with immunosuppressive medication for idiopathic thrombocytopenic purpura
HLA-B	Major histocompatibility complex class I-B	HLA-B*5701	Associated with hypersensitivity to abacavir
IL10	Interleukin 10	A-1082G	GG has better response to prednisone in leukaemia patients and sustained response in antiviral therapy for chronic hepatitis C infection (as part of haplotype (108 base pairs)-(-2575T)-(-2763C)-(-1082A)-(-819T)-(-582A))
TNF	Tumour-necrosis factor (TNF superfamily, member 2)	G-308A	A associated with good response to immunosuppressive therapy in patients with aplastic anaemia and carbamazepine hypersensitivity
XPC1	DNA-repair protein XPC1	Arg399Gln	Gln associated with resistance to 5-fluorouracil and oxaliplatin chemotherapy, Gln/Gln individuals less likely to develop therapy related acute myeloblastic leukaemia

# Perchè utilizzare test genetici nella pratica farmacologica?



# Inoltre:

- ✚ Il costo di sviluppo di un nuovo farmaco è di ~ \$ 800,000,000, per la maggior parte dovuti alla necessità di effettuare trials clinici estesi → i farmaci sono cari.



- ✚ Malgrado ciò, negli USA nell'ultimo trentennio 28 farmaci sono stati ritirati dal commercio a causa di ADR.

# Le promesse della Farmacogenomica

- ✚ Aumentare l'efficacia della terapia
- ✚ Maggiore sicurezza nella terapia
- ✚ Ridurre i costi di produzione dei farmaci mediante trials selettivi su pazienti scelti
- ✚ Evitare di ritirare dal commercio farmaci utili per molti, ma tossici o fatali per una minoranza di pazienti.

La medicina del futuro sarà una **medicina “personalizzata”**, dove, con l’aumentare delle conoscenze, e l’aiuto delle nuove tecnologie, si potrà:

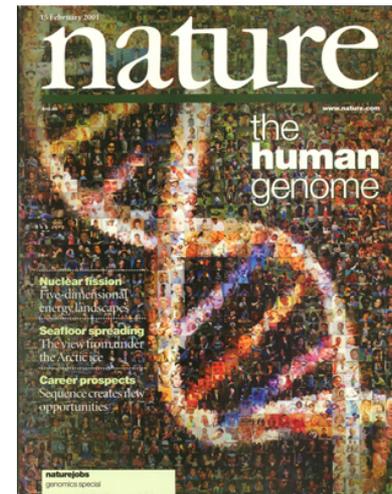
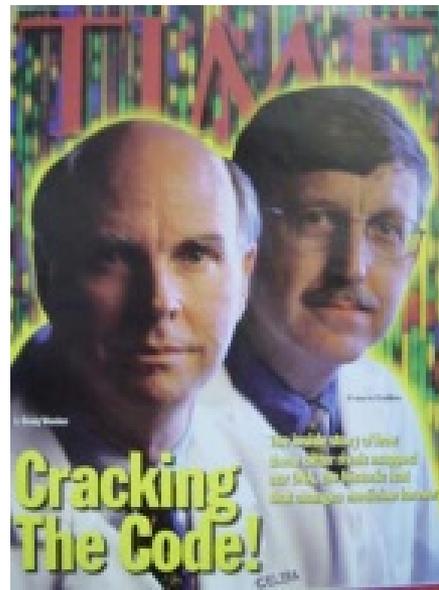
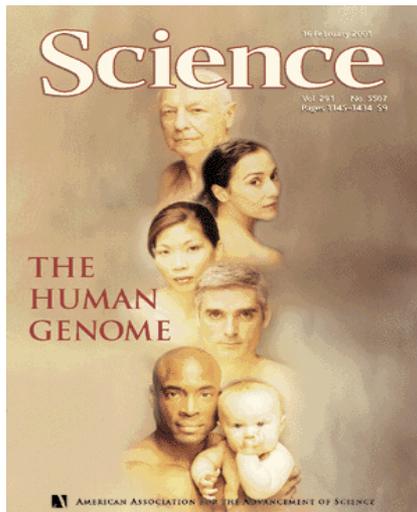
- ✚ Determinare la predisposizione di un individuo ad una specifica malattia.

- ✚ Selezionare la terapia medica più efficace, e il dosaggio adeguato

- ✚ Prevedere reazioni avverse ai farmaci.

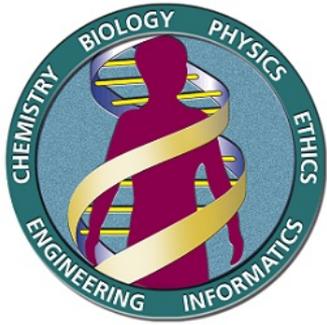
# Il futuro

## 2001: Sequenza del Genoma Umano



↓  
Craig Venter  
Celera

↓  
Francis Collins  
HGP-NIH



# Progetto Genoma Umano

- ✚ 1986: si annuncia l'iniziativa di sequenziare il genoma umano
- ✚ 1990: ha inizio un progetto con proiezione 15 anni
- ✚ 1999: primo miliardo di nucleotidi sequenziati
- ✚ 2000: completata la prima "bozza del genoma umano"
- ✚ 2003: Dichiarata completa la sequenza del genoma umano

Tempo: 13 anni

Costo: ~ \$ 1 miliardo

# Sequenziamento di genomi umani

Dopo il 2000, evolve drammaticamente la tecnologia di sequenziamento del DNA:

- ✦ 2004-7: 3 anni e \$ 70 milioni per sequenziare il genoma di Craig Venter
- ✦ 2007-8: 2 mesi e \$ 1 milione per sequenziare il genoma di James Watson
- ✦ 2008: 2 mesi e \$ 250,000 per sequenziare altri 4 genomi umani completi...

# Il genoma fai-da-te

- ✚ Come conseguenza dell'evoluzione tecnologica, sono nate una serie di servizi commerciali che offrono, a costi accessibili, analisi genomiche (23andMe, Navigenics, deCODE, GATC)
- ✚ Il cliente invia via posta un tampone di saliva, e riceve via web il suo genotipo.

# Il genoma "fa da te": 23andMe.com

https://www.23andme.com/ourservice/

23andMe genetics just got personal. [sign in](#) | [help](#)  
[have a claim code?](#)

[our service](#) [genetics 101](#) [for the experts](#) [store](#) [about us](#)

discover your genome at 23andMe [order now](#)

[ replay animation ]

1866: Gregor Mendel discovers the laws of inheritance.

200,000 years ago: *Homo sapiens* walks the Earth. 2003: The Human Genome Project maps a single person's genome.

2007: 23andMe introduces the first Personal Genome Service.

Unlock the secrets of your own DNA. Today.

175,000 years ago: The mother of all present-day humans is born in Africa.

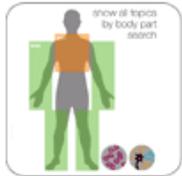
1953: Watson and Crick uncover the double-helix structure of DNA.

Welcome to 23andMe, a web-based service that helps you read and understand your DNA. After providing a saliva sample using an at-home kit, you can use our interactive tools to shed new light on your distant ancestors, your close family and most of all, yourself.

news

Welcome to the Future, *Portfolio*, November 2007

## Gene Journal



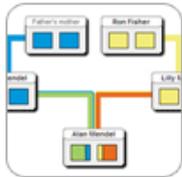
What do your genes say about you?

## Ancestry



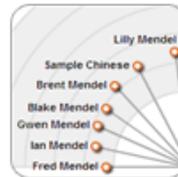
Who were your ancient ancestors?

## Family Inheritance



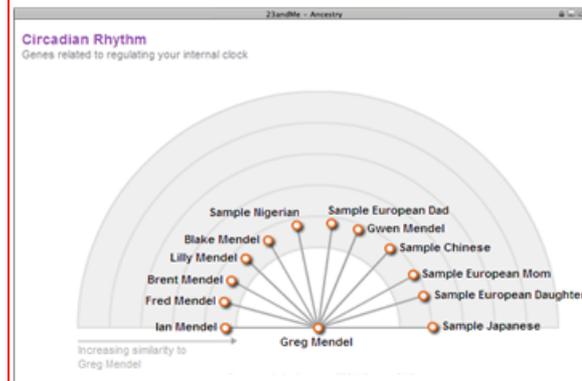
Do you have your mother's sense of taste?

## Genome Labs



Would you like to search your genome?

Compare your genetic blueprint to your friends and family.



## Genetic Comparisons: One-To-Many

23andMe's Genetic Comparisons tool allows you to see where you fall genetically with respect to your friends and family. You can compare yourself to others across all your genetic data, or look only at groups of genes affecting certain traits. You can line yourself up to just one other person or all of your participating friends and family at once.

# Social Issues

---

Genetic information is  
personal  
powerful  
potentially predictive  
pedigree-sensitive  
permanent  
prejudicial



Y-GA 99-1395

From US Department of Energy Human Genome Program, <http://www.ornl.gov/hgmis>