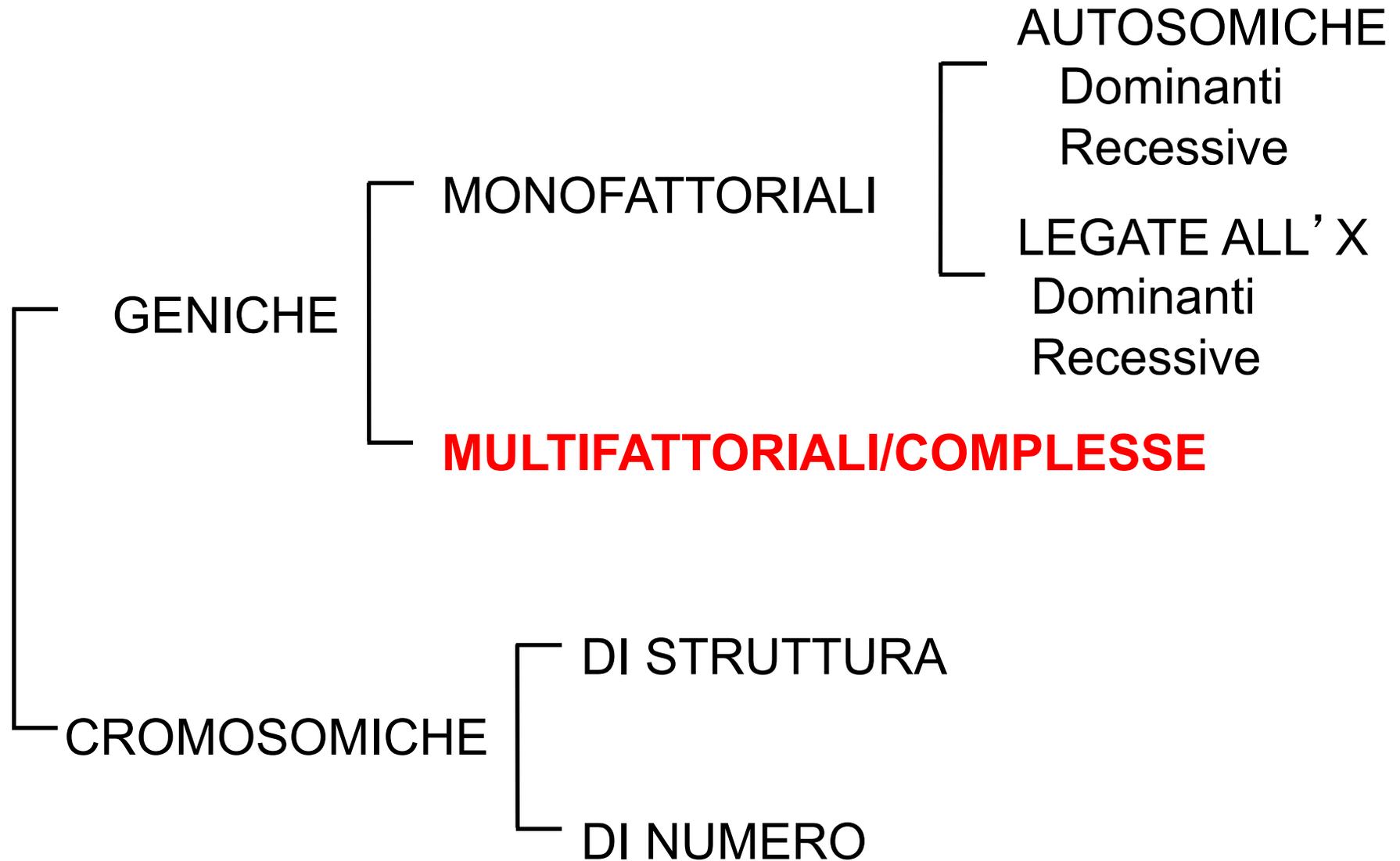
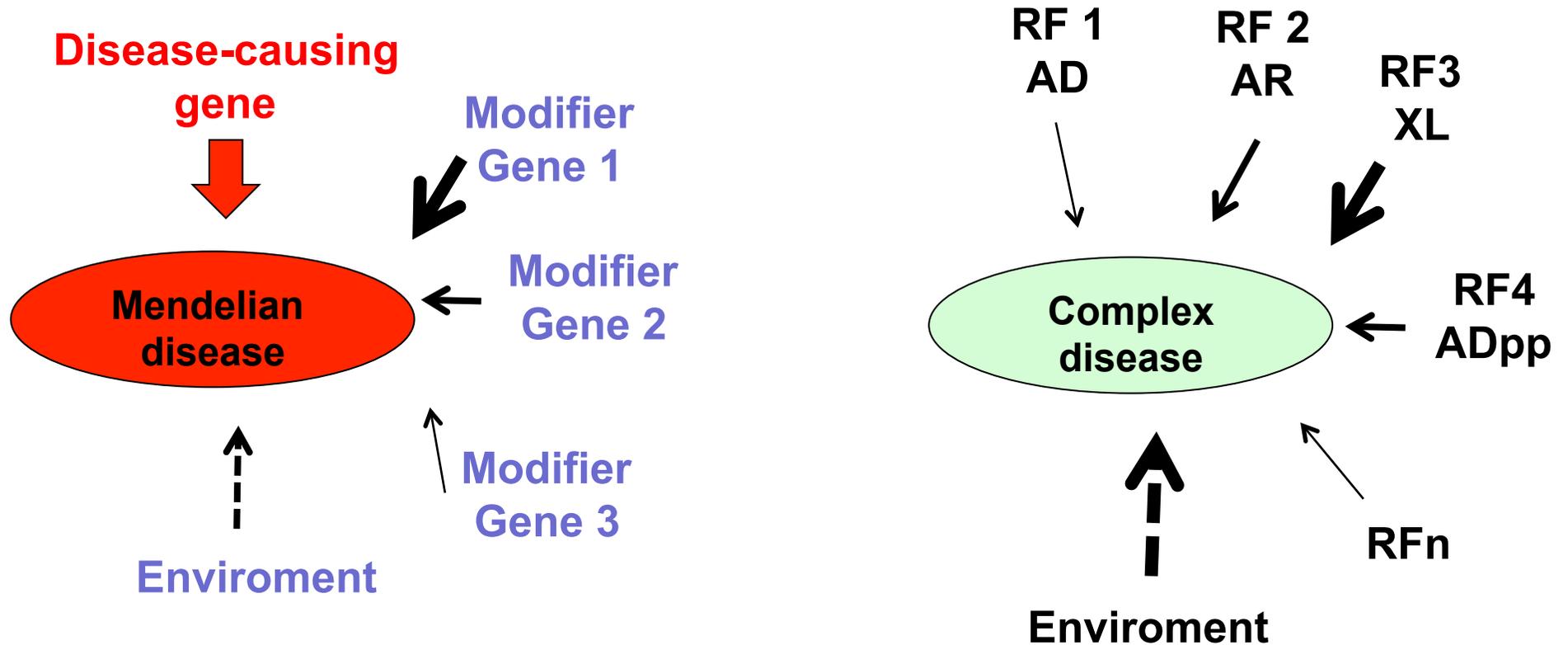


MALATTIE GENETICHE



Mendelian vs complex diseases

Penetrance



Modulation of expressivity

Do not confuse with genetic heterogeneity

Carattere poligenico: fenotipo determinato dall'azione combinata di numerosi loci, ciascuno con un piccolo effetto. **Ereditarietà poligenica**

Carattere multifattoriale: fenotipo determinato dalla combinazione di fattori ambientali e genetici. **Carattere complesso**

Malattie mendeliane



Malattie complesse

Ereditarietà

Trasmissione mendeliana

Autosomica recessiva

Autosomica dominante

X-linked

Familiarità

Malattia è più comune tra i parenti di un individuo affetto
che nella popolazione in generale

Trasmissione non-mendeliana

Rare

- fibrosi cistica
- sordità
- distrofia muscolare di Duchenne

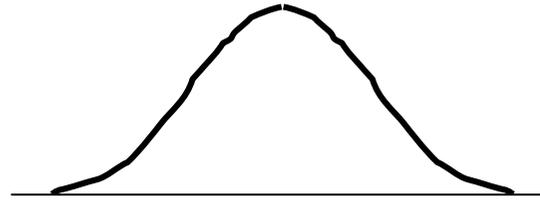
Comuni

- ipertensione
- asma
- diabete mellito
- malattia trombo-embolica
- infarto miocardio
- depressione
- cefalea
- osteoporosi
- patologia tiroidea

Caratteri

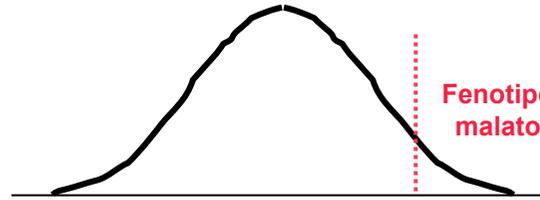
Ereditarietà

Quantitativi/continui



Esempio altezza

*A “soglia”
(discontinui)*

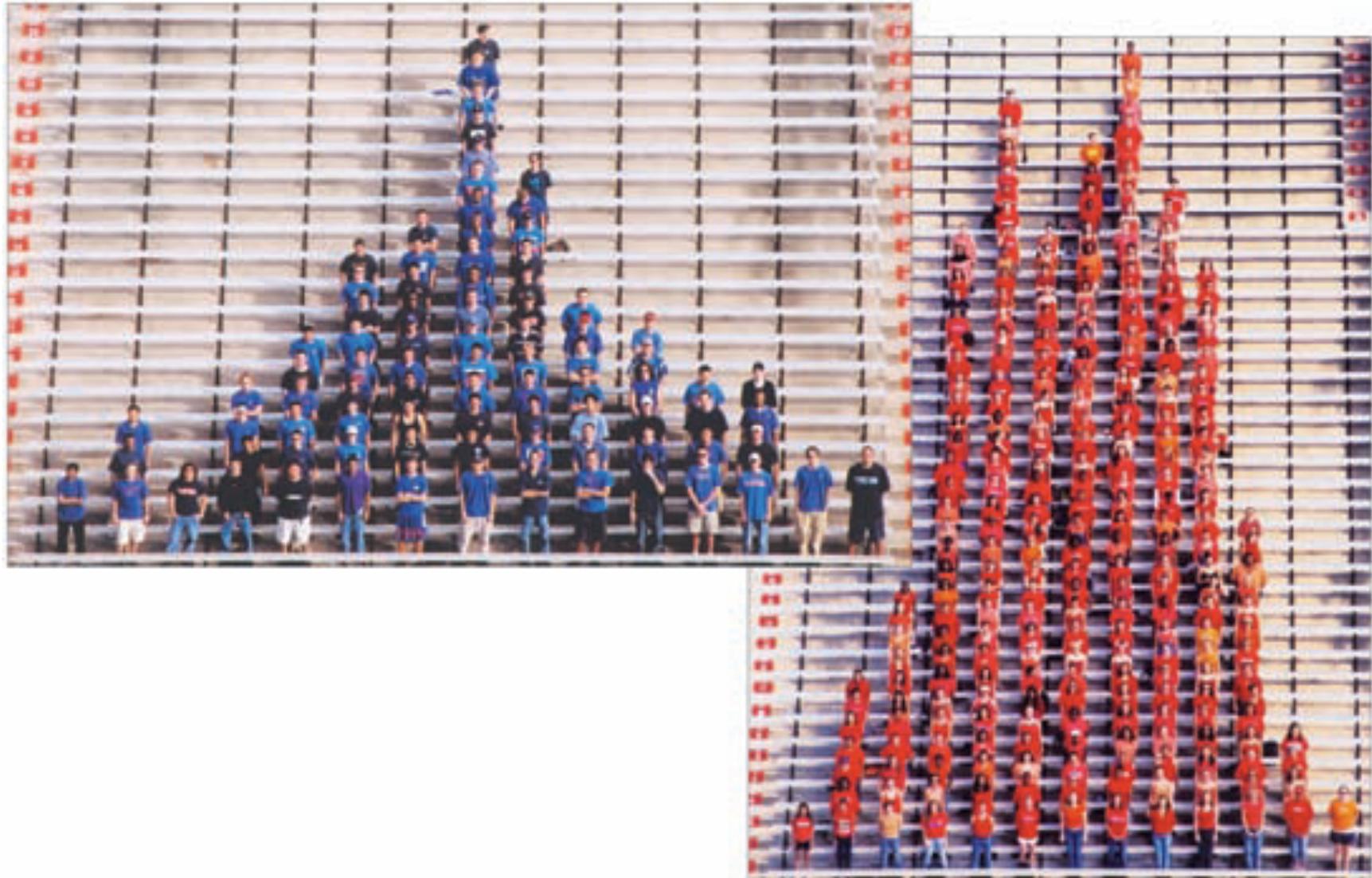


Esempio predisposizione

Teoria poligenica

Penetranza ed espressività variano con il numero di loci mutati e in relazione all'effetto delle diverse mutazioni, alle modalità di ereditarietà (recessiva, dominante, codominante) e alla mescolanza di componenti poligeniche e non.

Variazione continua: distribuzione delle altezze



Teoria poligenica: Caratteri quantitativi

QTL (quantitative trait loci)

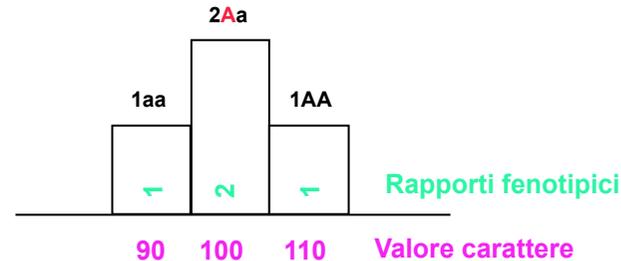
Effetto additivo degli alleli

Ipotetico carattere con valore medio = 100

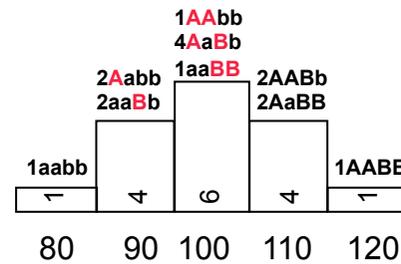
Effetto = 10
Per ogni allele dominante

Locus biallelico
 $p = q = 0,5$

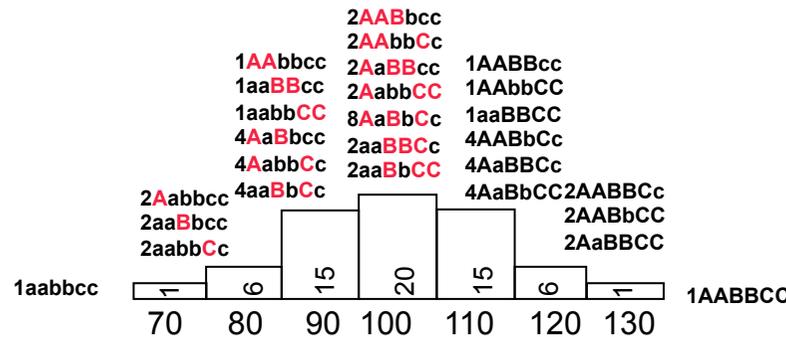
Effetto co-dominante alleli



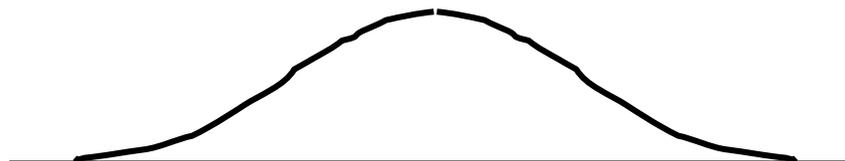
Unico locus



Due loci



Tre loci



N loci: poligeni

Solo cambiando un parametro del box rosso ...

Ipotetico carattere
con valore medio = 100

Effetto = 10
Per ogni allele dominante

Locus biallelico
p = 0,8
q = 0,2

Effetto co-dominante
alleli



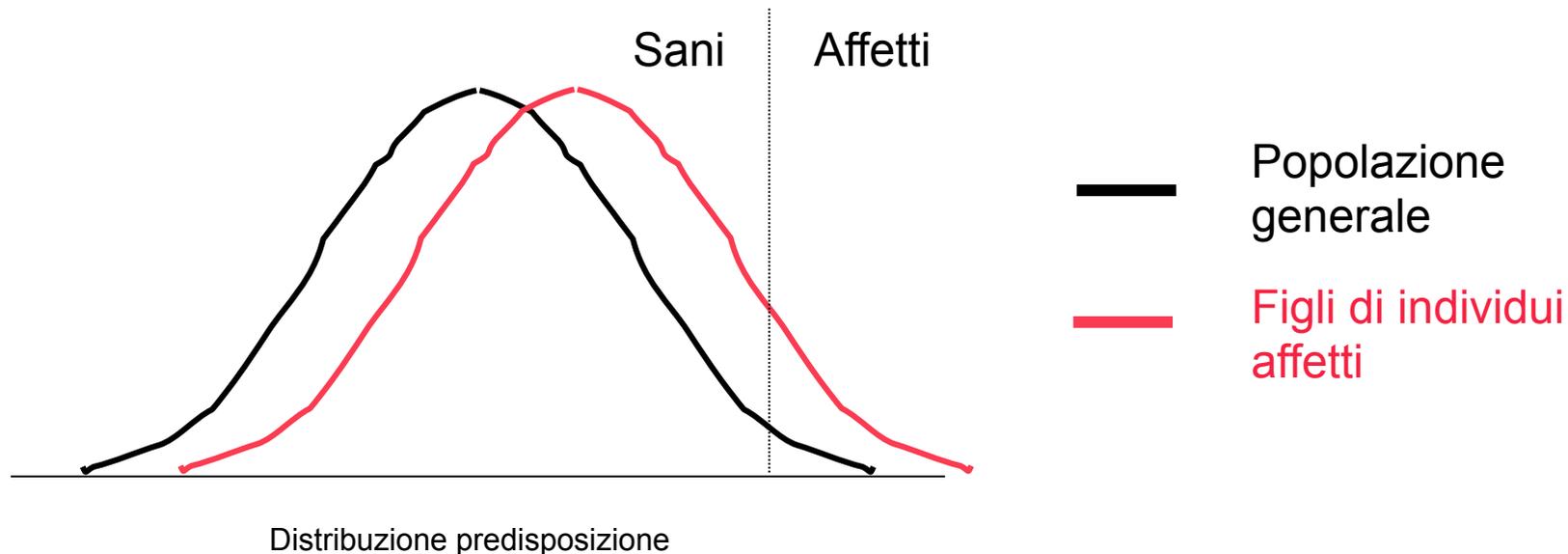
... per ogni locus i parametri del box rosso cambiano

Numero di genotipi e fenotipi attesi
al variare del numero delle coppie alleliche
che segregano indipendentemente e che hanno effetti additivi e uguali

Numero loci (2 alleli)	Alleli segreganti	Numero di genotipi	Numero fenotipi
1	2	$3^1 = 3$	3
2	4	$3^2 = 9$	5
3	6	$3^3 = 27$	7
4	8	$3^4 = 81$	9
n	2n	3^n	2n+1

Teoria poligenica: caratteri discontinui

- Tutti gli individui hanno un qualche grado di predisposizione a manifestare un carattere
- I fattori di suscettibilità (genetici e ambientali) sono molteplici
- Il numero dei fattori è distribuito secondo la curva gaussiana
- La comparsa del fenotipo dipende da una soglia di suscettibilità
- I consanguinei avranno un valore medio di fattori di predisposizione maggiore



Caratteri discontinui: rischio di ricorrenza λ_R
 $\lambda_R = 1$ (rischio uguale a quello della popolazione generale)

Rischio di ricorrenza si basa su osservazioni empiriche

- Numero dei malati in una famiglia (maggiore è il numero di individui malati nella famiglia maggiore è λ_R)
- Gravità della malattia nel consanguineo affetto (individui colpiti in modo grave hanno in media un numero maggiore di fattori di predisposizione)
- Incidenza della malattia nella popolazione (tanto maggiore è la malattia nella popolazione tanto minore è il peso dei fattori genetici)
- Incidenza diversa nei sessi

Esempio: stenosi del piloro

	Figli	Figlie	Fratelli	Sorelle
Maschio affetto	6,5	2,5	2,2	2,0
Femmina affetta	23	11,5	11	9

Studi per definire la componente genetica (in blu le criticità degli studi)

- **Ricorrenza nelle famiglie:** rischio di un parente di sviluppare la malattia rispetto alla popolazione generale (condivisione dell'ambiente potrebbe essere un'alternativa alla causa genetica)
- **Studi su gemelli "identici" monozigoti (MZ)** per valutare la concordanza (% di coppie di gemelli che condividono un carattere). Se il carattere è controllato geneticamente, è maggiore nei MZ rispetto a quella dei gemelli dizigotici (DZ) (condivisione ambiente, genere diverso tra i DZ)
- **Studi su gemelli MZ affidati in adozione:** condividono il genoma ma non l'ambiente (la separazione però non è totale)
- **Studi su figli adottati:** valutare se la malattia che colpisce i figli adottati è presente o meno nelle famiglie adottive o naturali; viceversa valutare la malattia in figli dati in adozione quando la malattia è presente nei loro genitori.

Esempi di concordanza

% di coppie di gemelli che condividono un carattere

Malattia	Concordanza (%)	
	MZ	DZ
Epilessia	70	6
Sclerosi multipla	18	2
Diabete tipo 1	40	5
Schizofrenia	53	15
Artrite reumatoide	12	3,5
Psoriasi	72	15
Palatotoschisi	30	5

Più i valori si differenziano
Maggiore è l'influenza genetica

Fattori di rischio genetici: caratteristiche

- Varianti comuni nella popolazione (es. SNP)
- Azione non “neutra” sulla funzione di un prodotto genetico (effetti/variazioni lievi)
- Esempi
 - SNP nella regione codificante di un gene che influisce sulla funzione della proteina (es. riduzione dell’attività del 5%)
 - SNP in regioni regolatrici che interferiscono con i normali livelli di espressione della proteina (ridotti/aumentati di una piccola percentuale)

Studio di associazione

Studio casi-controlli. Consiste nel confronto della frequenza di un determinato allele ad un locus in due gruppi: presenza (casi) o assenza (controlli) del carattere. Se un allele mostra una frequenza significativamente maggiore nei “casi” rispetto ai “controlli” si dice che esso è associato al fenotipo in esame.

L'allele A1 al locus A è associato alla malattia D se gli individui affetti da D hanno A1 ad una frequenza maggiore di quella attesa sulla base della frequenza di D e di A1 nella popolazione.

Esempio: allele HLA-DR4 ha una frequenza del 36% nella popolazione (UK) generale ma del 78% in individui con l'artrite reumatoide.

Studio casi-controlli

Vantaggi: facilità di reclutamento della casistica.
Approccio utilizzato per valutare il coinvolgimento di un gene candidato nella predisposizione di un carattere.

Aspetti limitanti:

- i) campione “casi” omogeneo
- ii) numerosità del campione (tanto più frequente è un polimorfismo tanto minore sarà l'effetto),
- iii) selezione adeguata del campione “controlli”: **sexso, età, origine geografica per limitare eventuali differenze delle frequenze alleliche tra le popolazioni.**

Studi di associazione caso-controllo

Calcolo dell'**Odds Ratio** (OR, rapporto di probabilità)
per valutare la forza di associazione

Allele "x"	Pazienti (N=100)	Controlli (N=120)
Presente	90 (a)	40 (b)
Assente	10 (c)	80 (d)

$$\frac{\frac{a}{a+c}}{\frac{c}{a+c}} \quad \text{Probabilità di avere l'allele "x" rispetto a non averlo nei pazienti}$$

$$\frac{\frac{b}{b+d}}{\frac{d}{b+d}} \quad \text{Probabilità di avere l'allele "x" rispetto a non averlo nei controlli}$$

$$\frac{a d}{b c} \quad \text{Utilizzo } \chi^2 \text{ per la significatività}$$

Studio di associazione

Studio di associazione con controlli interni: transmission disequilibrium test (TDT) prevede l'analisi del probando e di entrambi i genitori. Può essere utilizzato solo quando i genitori sono disponibili

Vantaggi: si evita il problema della selezione del campione controllo

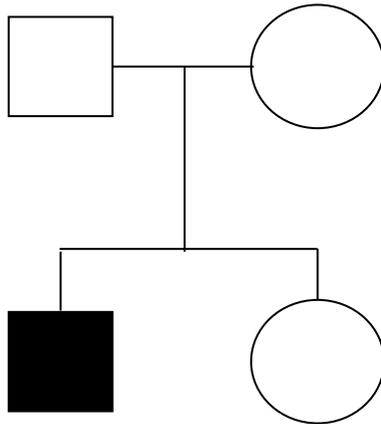
Aspetti limitanti: genitori non sempre disponibile. Malattie sono spesso ad esordio tardivo

TDT (transmission disequilibrium test)

Uso di CONTROLLI INTERNI, gli stessi usati per ricavare i dati

La trasmissione di un allele si discosta significativamente dal valore atteso del 50%

A1A2



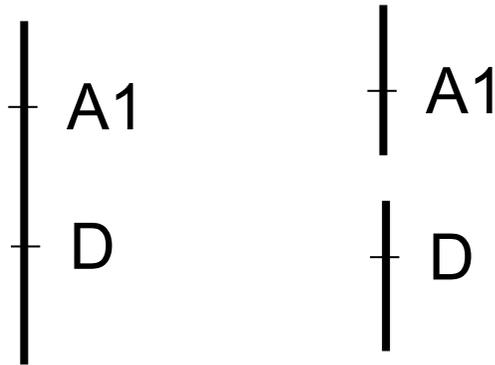
- Determinare genotipo dei probandi e dei genitori
- Selezionare genitori eterozigoti (possono essere o non essere ammalati)
- “a” è il numero di volte che un genitore eterozigote trasmette l’ allele A1 ai figli ammalati, e “b” è il numero di volte che trasmette l’ altro allele. Se c’ è una differenza significativa tra “a” e “b”, c’ è associazione.

$$(a - b)^2 / (a + b)$$

Distribuzione χ^2 con 1 grado di libertà

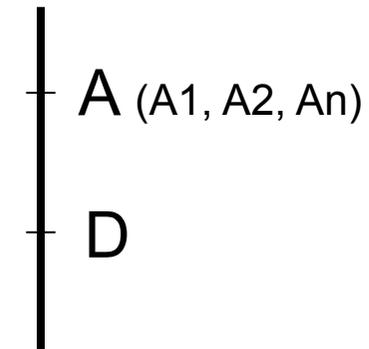
Linkage \neq Associazione

Associazione



Relazione tra alleli
(Osservazione statistica)

Linkage



Relazione tra loci (genetica)

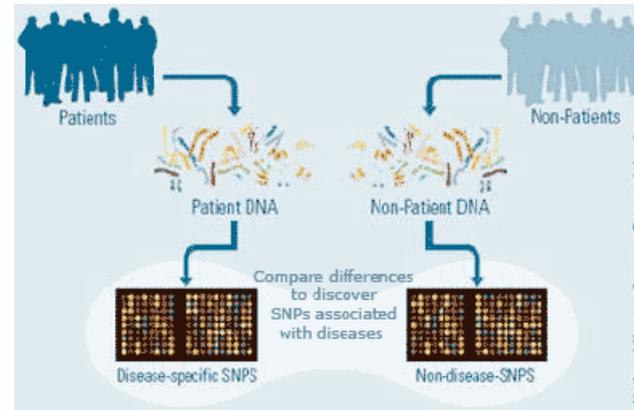
- distribuzione degli alleli A in un gruppo di pazienti è uguale a quella della popolazione generale (almeno che non ci sia un effetto fondatore);
- all' interno di una famiglia affetta che segrega con il locus A i malati condividono lo stesso allele;
- linkage crea associazione all' interno delle singole famiglie ma non tra individui non imparentati.

Studio di associazione

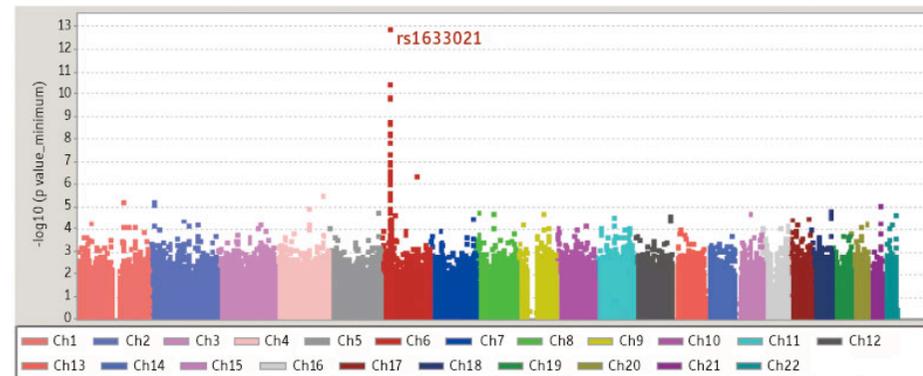
- Gene candidato: analisi di SNP ad esso associati
- Genome wide association study: analisi SNP su tutto il genoma

Genome Wide Association Study (GWAS)

**Piattaforme da
500.000/
1.000.000 SNPs**



**Rappresentazione
“Manhattan”**



- Identificazione di varianti con una frequenza maggiore (statisticamente significativa) nel gruppo di pazienti rispetto ai controlli
- Rispetto agli studi del gene candidato in cui si analizzano alcuni SNP nel gene stesso, nel GWAS si analizzano SNPs distribuiti su tutto il genoma.

Perché si osserva associazione?

- **Alleli direttamente o indirettamente implicati nella malattia**

1) Causa diretta: presenza dell' allele A aumenta la predisposizione. Effetto patogenetico dell' allele.

2) Linkage disequilibrium: alleli sono associati l' uno all' altro ad una frequenza significativamente più elevata rispetto a quella attesa calcolata sulla base delle frequenze alleliche e l' associazione random. Allele è in stretta vicinanza alla variante patogenetica.

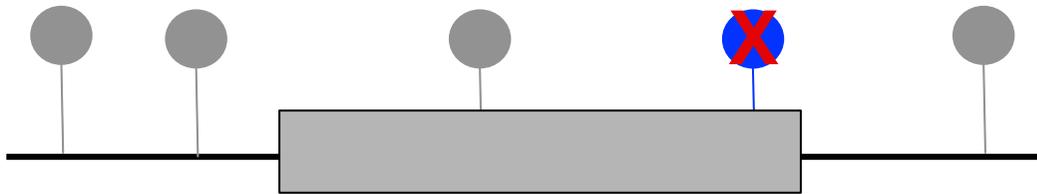


SNP analizzati

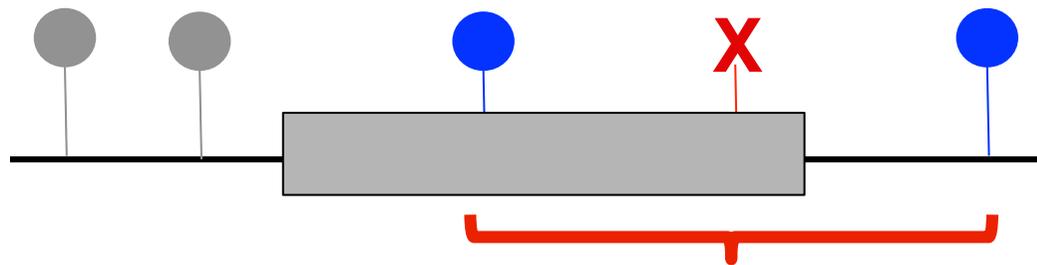


SNP in associazione

X Fattore di rischio



Causa diretta



Causa indiretta

Linkage disequilibrium

Linkage disequilibrium (LD)

Teoricamente, la distribuzione delle combinazioni alleliche nella popolazione dovrebbe seguire le leggi della segregazione causale e della probabilità.

Dati 2 loci, A e B, ciascuno con due alleli

le cui frequenze sono:

$$F(A) = p(A)$$

$$F(a) = q(a)$$

$$F(B) = p(B)$$

$$F(b) = q(b)$$

le frequenze attese delle combinazioni (aplotipi) dipendono dalle frequenze alleliche, come indicato in tabella.

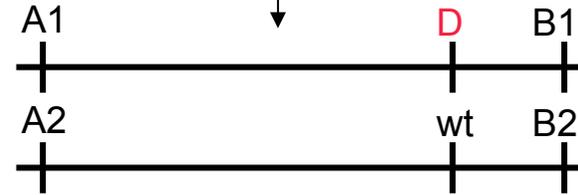
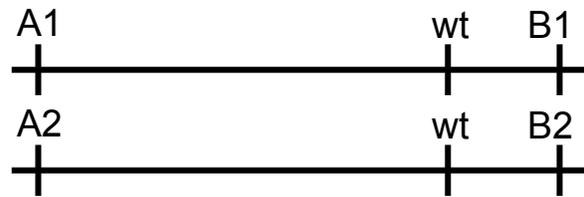
	p(A)	q(a)
p(B)	AB $p(A)p(B)$	aB $q(a)p(B)$
q(b)	Ab $p(A)q(b)$	ab $q(a)q(b)$

Linkage disequilibrium: deviazione delle frequenze aplotiche osservate rispetto a quelle attese. Ciò accade quando i loci sono strettamente concatenati. Per effetto del crossing-over, prima o poi l'equilibrio dovrebbe essere raggiunto con una velocità che dipende dalla distanza tra i loci.

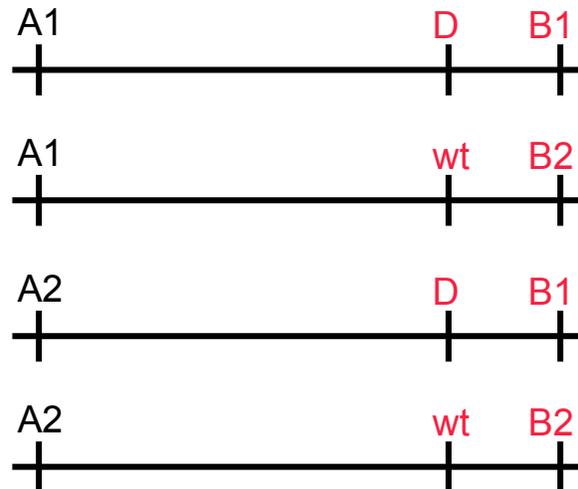
LD = Frequenza aplotipo osservata - Frequenza aplotipo attesa

$$LD = F(AB) - [p(A)p(B)]$$

$$LD = 0, \text{ frequenze alleliche in equilibrio}$$



Crossing over



Crossing over
più frequente
tra locus A e D



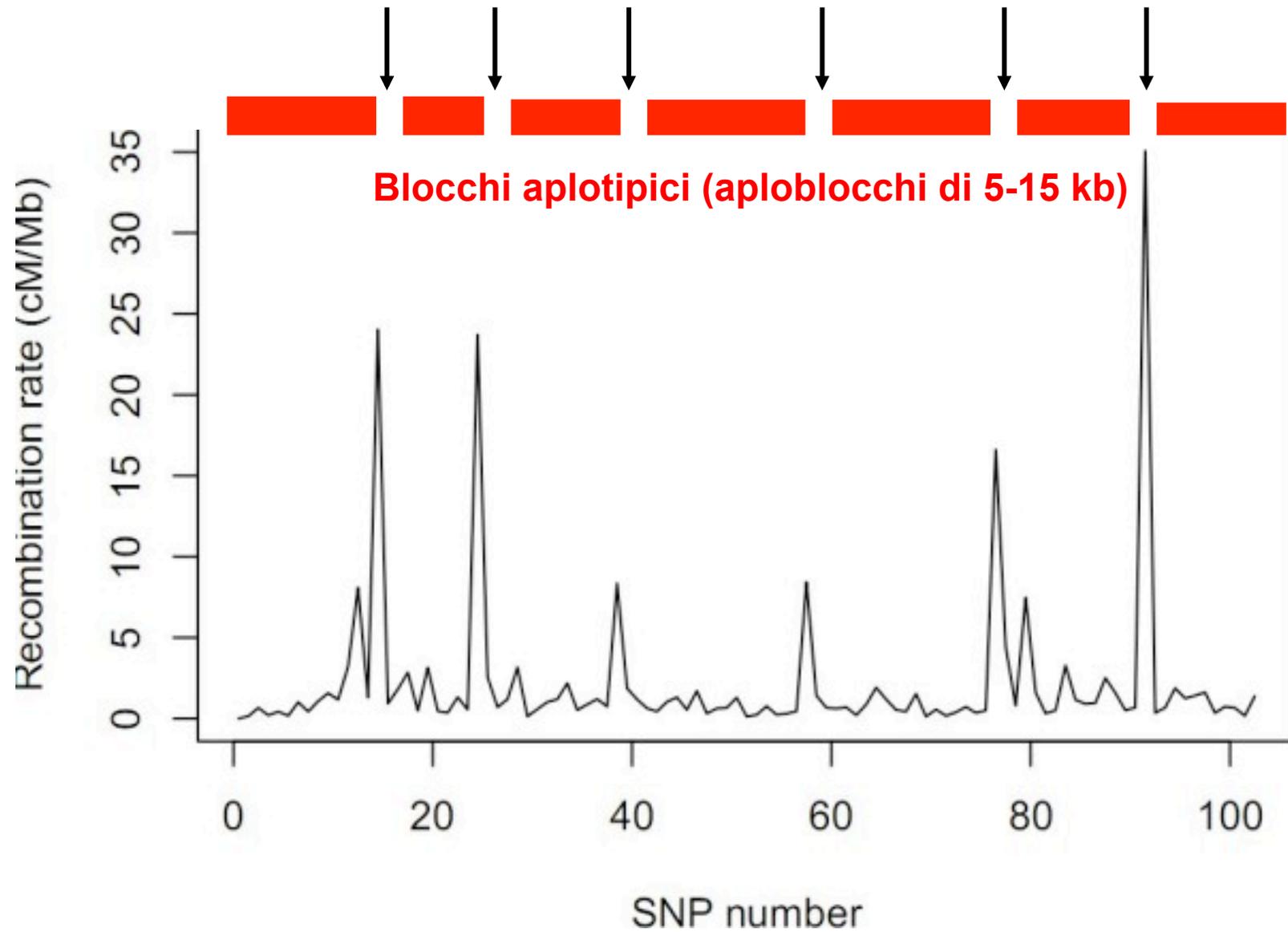
Equilibrium
tra gli alleli
al locus A e
al locus D

Crossing over
meno frequente
tra locus D e B

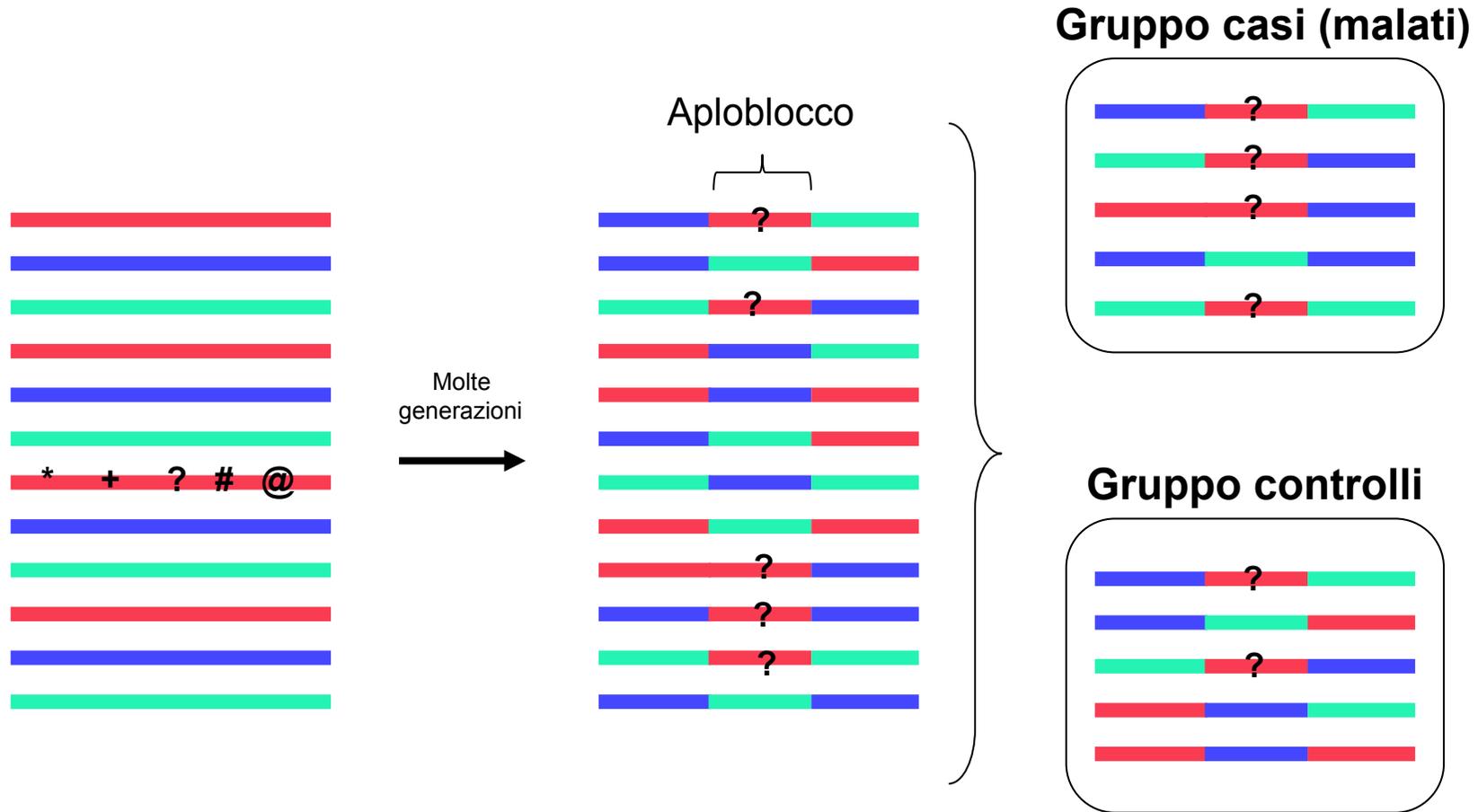


Linkage disequilibrium
(non equilibrio causato dal linkage)
Allele B1 tenderà
ad essere trasmesso
insieme all'allele D

Hotspot di ricombinazione

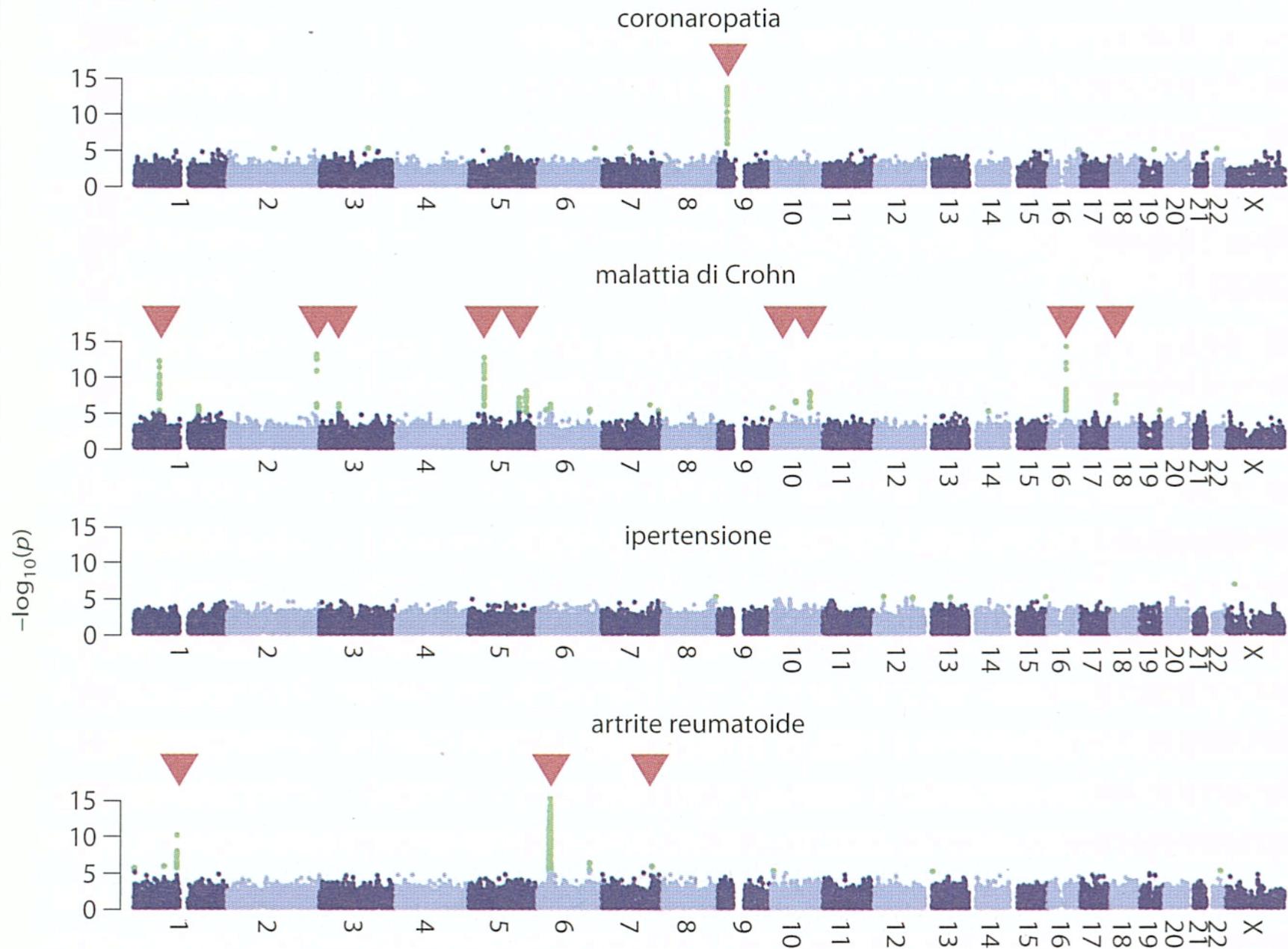


Associazione aplotipica



- Alleli ai diversi loci dell'aploblocco presentano una frequenza più elevata nel gruppo "casi" rispetto a quello "controlli";
- Se nessun allele ai diversi loci è causa diretta, la variata causativa (fattore di rischio) va ricercata nella regione rappresentata dall'aploblocco

Esempi di studi di associazione



Considerazioni:

- Più di 4.000 SNP sono associati a diverse malattie complesse (<http://www.genome.gov/GWASStudies>)
- Si conoscono le varianti patogenetiche solo in rari casi (es uno SNP del gene di IRF6 che predispone alla schisi del labbro)
- Molto difficile identificare le varianti causative nelle malattie complesse (bassa penetranza e concomitanza di numerosi altri fattori genetici) contrariamente a quanto si verifica nelle malattie mendeliane che sono determinate da varianti (mutazioni) a completa (o quasi completa) penetranza (altri fattori genetici contribuiscono all'espressività)