

GENETICS AND MOLECULAR BIOLOGY FOR ENVIRONMENTAL ANALYSIS

MOLECULAR ECOLOGY LESSON 9: POPULATION GENETICS

**Prof. Alberto Pallavicini
pallavic@units.it**

INTRODUZIONE

Gli studi di Genetica si sono sviluppati all'inizio di questo secolo, con la riscoperta delle leggi di Mendel, e le ricerche in questa disciplina hanno avuto un grande impulso in relazione all'importanza che la comprensione dei fenomeni ereditari ha per gli studi evuzionistici.

Questi studi in un primo tempo si sono occupati delle modalità di trasmissione dei caratteri da genitori a figli, sono cioè stati condotti a **LIVELLO DELL'ORGANISMO**

Successivamente, a partire dal lavoro di Muller (1922) ci si è occupati della fonte delle variazioni del materiale genetico, cioè delle cause delle mutazioni;

http://en.wikipedia.org/wiki/Hermann_Joseph_Muller

nel 1931 Creighton e McClintock pongono le basi della citogenetica, cioè della disciplina che studia i cromosomi come sede del materiale genetico;

CYTOGENETICS

- Cytogenetics is a branch of genetics that is concerned with the study of chromosomes and cell division.
- It includes routine analysis of G-Banded chromosomes, other cytogenetic banding techniques, as well as molecular cytogenetics such as fluorescent in situ hybridization (FISH) and comparative genomic hybridization (CGH).

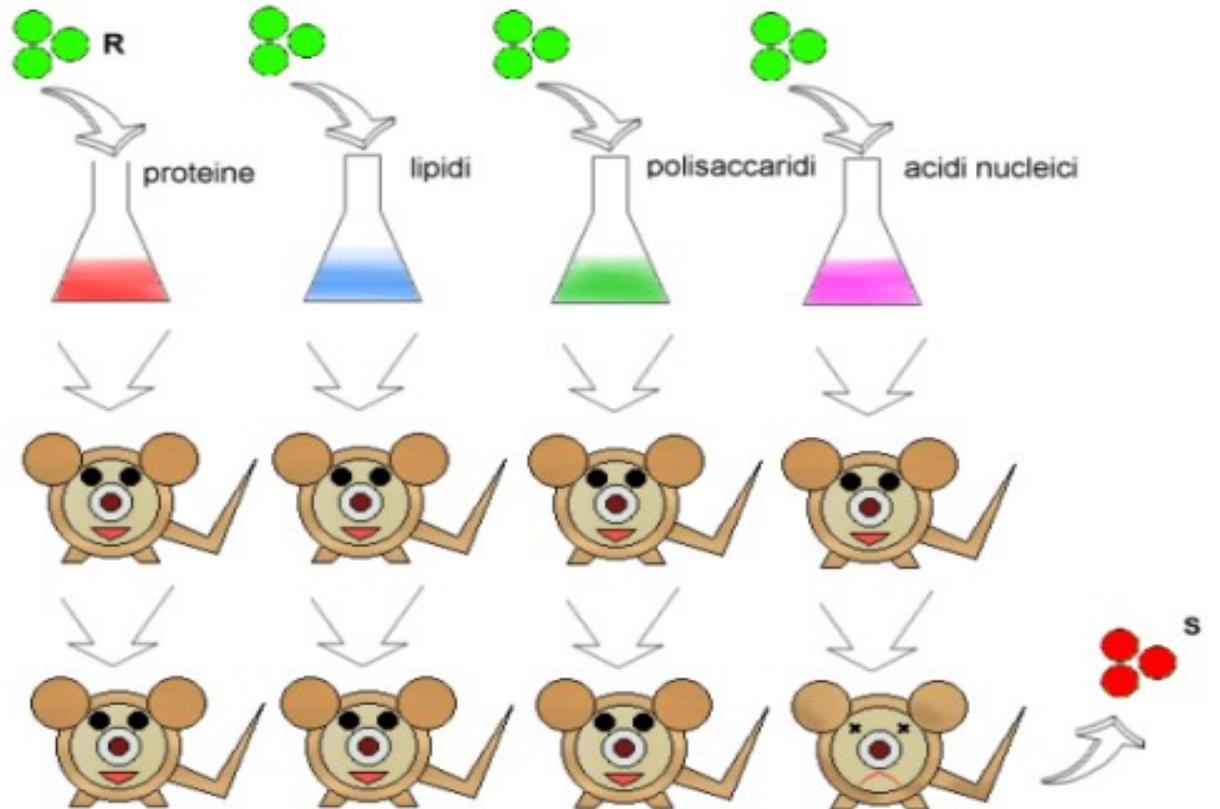


con Beadle e Tatum nel 1941 si affronta il problema delle modalità di espressione del materiale genetico (fisiogenetica).

By constructing mutant strains that required specific nutritional elements ([amino acids](#) or [vitamins](#)), they established that individual gene mutations were responsible for individual steps in the metabolism and synthesis of vital nutrients. This led, in 1941, to propose the "one gene-one enzyme hypothesis," the idea that a gene specifies a single enzyme, rather than a complex set of characteristics



con Avery e collaboratori (1944) si comincia a comprendere la natura chimica del materiale genetico.



- Gli studi genetici passano quindi da un livello organismico ad un **livello molecolare**.
- Attualmente gli studi di genetica sono massimamente orientati in questo senso, per cui si ha spesso l'impressione che la genetica sia semplicemente una delle discipline della **Biologia Molecolare** (termine coniato da Astbury nel 1945).



Popolarmente si considera la Genetica come una disciplina che studia i meccanismi con i quali determinate molecole contengono l'informazione necessaria per determinare i processi vitali ed eventualmente utile per interferire in questi processi (**ingegneria genetica**).

La Genetica ha avuto uno sviluppo non solo sul livello molecolare a partire da quello dell'organismo, ma anche uno sviluppo su un piano sopraorganismico, popolazionistico, che studia come possano modificarsi le caratteristiche genetiche in un complesso di individui, **le popolazioni.**

Possiamo considerare la Genetica delle popolazioni come l'altra linea di sviluppo della Genetica, che conserva legami con la Genetica molecolare in quanto ne utilizza le scoperte a fini di indagine e di comprensione dei processi evolutivi.

CONCETTO DI POPOLAZIONE

L'oggetto della Genetica delle popolazioni, come dell'Ecologia, non è costituito quindi da individui, ma da **insiemi di individui**.

Nel caso dell'Ecologia si tratta di individui che hanno tra loro rapporti in quanto convivono nello stesso ambiente.

Nel caso della Genetica delle popolazioni si tratta di individui che hanno tra loro rapporti di tipo genetico.



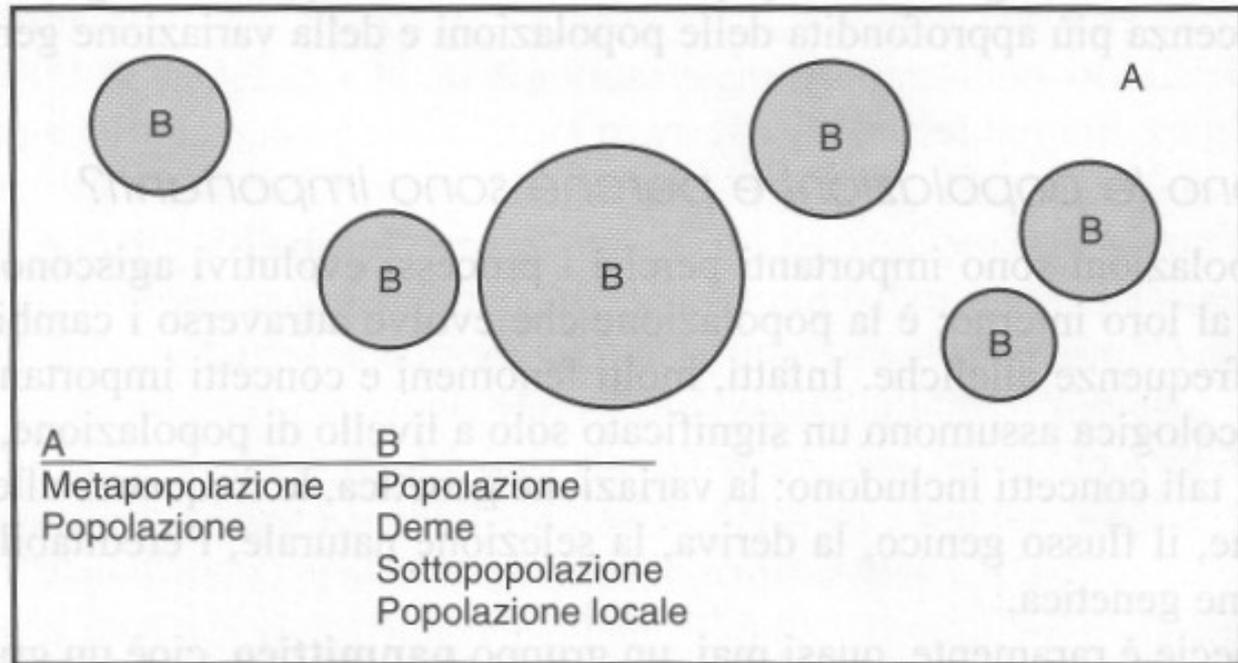
Potremmo considerare popolazione l'insieme degli individui tra cui possono avvenire scambi genetici, cioè l'insieme degli individui che si incrociano.

Tale definizione però è notevolmente ambigua.

concetto biologico della specie:

Mayr definisce le specie (1942) come gruppi di reali o potenziali popolazioni naturali che si incrociano ma sono riproduttivamente isolate da altri gruppi del genere.

Semplice rappresentazione di una popolazione strutturata nello spazio. L'area A, potenzialmente, può essere piccola come un metro cubo di terreno in cui vive un microrganismo, oppure grande come un continente per un uccello migratore.



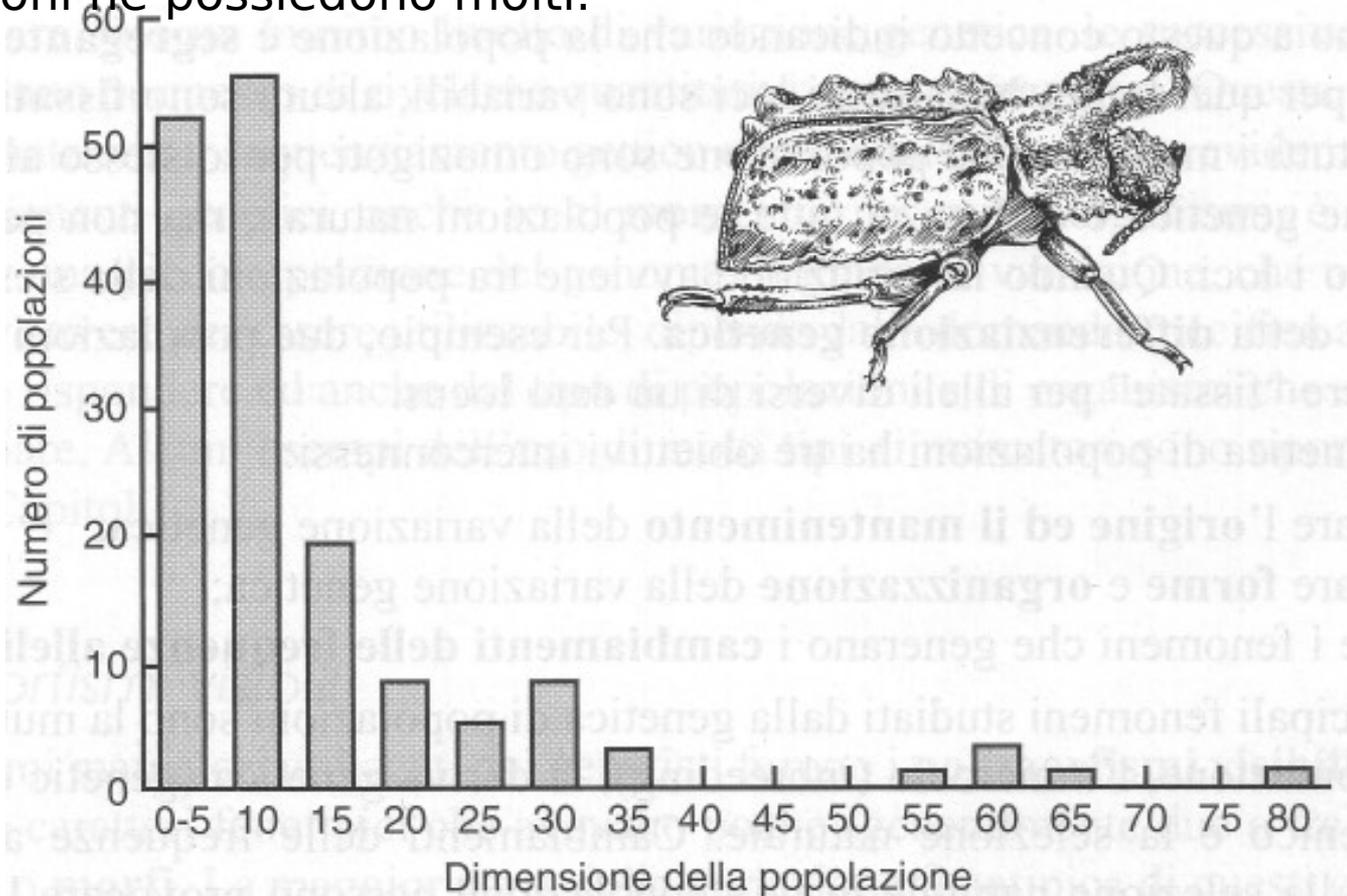
La maggior parte degli incroci avverrà all'interno delle **sottopopolazioni**, ma esisterà un certo potenziale di flusso genico anche tra sottopopolazioni di una determinata **metapopolazione**.

Verrà usata tuttavia la parola popolazione nelle discussioni che non si riferiscono a particolari strutture organizzate nello spazio.



Grafico della distribuzione della frequenza, in cui è mostrato il numero di sottopopolazioni osservate per ciascuna classe di dimensioni nella metapopolazione di coleottero dei funghi, *Bolitotherus corrmjvs*.

Molte sottopopolazioni hanno pochissimi individui mentre poche sottopopolazioni ne possiedono molti.



STRUTTURA GENETICA DELLE POPOLAZIONI

Una volta definita una popolazione dobbiamo descriverla da un punto di vista genetico.

Dobbiamo cioè descriverne la struttura genetica

composizione allelica e genotipica.

composizione allelica: identificare per i vari loci la frequenza delle loro varianti, cioè degli alleli

composizione genotipica: riconoscere la frequenza dei vari genotipi.



Nel caso di forme diploidi abbiamo individui

omozigoti ed eterozigoti, per cui se sono 2 gli alleli sono possibili 3 genotipi differenti

in generale se sono presenti **n** alleli avremo
n tipi di omozigoti e
 $n \cdot (n-1) / 2$ tipi di eterozigoti

In totale:

$$\text{n}^\circ \text{ genotipi} = n + \frac{n \cdot (n-1)}{2} = \frac{n \cdot (n+1)}{2}$$

Nel caso di due alleli, in popolazioni diploidi, il numero di genotipi possibili è 3:

AA Aa aa

Per semplicità considereremo:

un solo gene con due alleli, per forme diploidi

popolazioni a generazioni separate
(**monovoltini**)

In queste condizioni abbiamo due alleli e tre genotipi di cui considerare le frequenze:

$$p_A \quad q_a$$

$$D_{AA} \quad H_{Aa} \quad R_{aa}$$

E' evidente che trattandosi di frequenze deve essere:

$$p + q = 1$$

$$D + H + R = 1$$

Osserviamo che esiste una relazione tra i descrittori delle frequenze alleliche (p e q) e i descrittori delle frequenze genotipiche (D, H e R).

La frequenza dell'allele A sarà pari alla frequenza degli omozigoti per A (che contengono solo questo allele) più la metà della frequenza degli eterozigoti con A (che contengono sia questo che u

$$p = D + \frac{H}{2}$$

Antigeni sanguigni del sistema MN su un campione della popolazione dei groenlandesi.

due forme alleliche (M e N) codominanti;

Avremo pertanto tre fenotipi (M, N, MN), corrispondenti a tre genotipi (MM, NN, MN).

I risultati del campionamento siano stati:

	M	MN	N	totale
n°	475	89	5	569

da cui si ricava:

	DM	HMN	RN
	0.835	0.156	0.009

e pertanto

$$pM = DM + \frac{1}{2}HMN = 0.835 + 0.078 = 0.913$$

e

$$qN = RN + \frac{1}{2}HMN = 1 - pM = 0.087$$

I valori che abbiamo trovato ovviamente sono stime. Utilizzando un altro campione della popolazione groenlandese avremmo potuto trovare valori differenti.

problema statistico

si deve considerare la significatività del campione e la significatività dei valori ottenuti.

Un campionamento simile è stato condotto sulla popolazione islandese, ottenendo i seguenti risultati:

	M	MN	N	totale
n°	233	385	129	747

da cui si ricava:

	DM	HMN	RN
	0.312	0.515	0.173

e pertanto

$$pM = 0.569 \quad qN = 0.431$$

In questo caso i valori di D, H e R potevano essere direttamente calcolati dai fenotipi, in quanto non si presenta il fenomeno della **dominanza**.

Se invece si manifesta questo fenomeno abbiamo due soli fenotipi, quello dominante e quello recessivo, le cui frequenze saranno rispettivamente D e R, dove $D = D + H$.

E' evidente che in questo caso non possiamo direttamente stimare i valori di **p** e **q**.

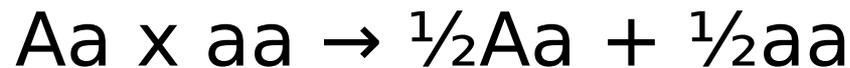
Potremmo calcolarli sperimentalmente se osservassimo un sufficiente numero di reincroci di individui a fenotipo **dominante** con individui a fenotipo **recessivo**.

Infatti in questo incrocio si possono verificare due casi:

- se l'individuo a fenotipo dominante è omozigote avremo solo discendenza a fenotipo dominante:



- se l'individuo a fenotipo dominante è eterozigote potremo avere discendenza con ambedue i fenotipi con pari probabilità:



Nel 1908 indipendentemente Hardy, inglese, e Weinberg, tedesco, giungono alle stesse conclusioni:

le frequenze dei tre genotipi possibili sono funzione diretta delle frequenze alleliche, per cui non sono necessari due parametri, ma ne è **sufficiente uno solo** per descrivere la struttura genetica di una popolazione diploide.

Supponiamo di partire da una determinata struttura genetica D_0, H_0, R_0

supponiamo che gli incroci si verifichino casualmente, indipendentemente dai genotipi dei due partner;

per semplicità supporremo anche che le frequenze genotipiche **non siano differenti** tra i due sessi (se anche lo fossero si arriverebbe comunque alle stesse conclusioni).

	D_0 AA	H_0 Aa	R_0 aa
D_0 AA	AA	$\frac{1}{2}$ AA $\frac{1}{2}$ Aa	Aa
H_0 Aa	$\frac{1}{2}$ AA $\frac{1}{2}$ Aa	$\frac{1}{4}$ AA $\frac{1}{2}$ Aa $\frac{1}{4}$ aa	$\frac{1}{2}$ Aa $\frac{1}{2}$ aa
R_0 aa	Aa	$\frac{1}{2}$ Aa $\frac{1}{2}$ aa	aa

Potremo allora calcolare i valori attesi della discendenza D_1 , H_1 , R_1 .

$$D_1 = D_0 \cdot D_0 + \frac{1}{2} \cdot D_0 \cdot H_0 + \frac{1}{2} \cdot H_0 \cdot D_0 + \frac{1}{4} \cdot H_0 \cdot H_0$$

$$H_1 = \frac{1}{2} \cdot D_0 \cdot H_0 + D_0 \cdot R_0 + \frac{1}{2} \cdot H_0 \cdot D_0 + \frac{1}{2} \cdot H_0 \cdot H_0 + \frac{1}{2} \cdot H_0 \cdot R_0 + R_0 \cdot D_0 + \frac{1}{2} \cdot R_0 \cdot H_0$$

$$R_1 = \frac{1}{4} \cdot H_0 \cdot H_0 + \frac{1}{2} \cdot H_0 \cdot R_0 + \frac{1}{2} \cdot R_0 \cdot H_0 + R_0 \cdot R_0$$

e, semplificando:

$$D_1 = D_0 \cdot D_0 + D_0 \cdot H_0 + \frac{1}{4} \cdot H_0 \cdot H_0 = \left(D_0 + \frac{1}{2} \cdot H_0\right)^2$$

$$R_1 = R_0 \cdot R_0 + R_0 \cdot H_0 + \frac{1}{4} \cdot H_0 \cdot H_0 = \left(R_0 + \frac{1}{2} \cdot H_0\right)^2$$

$$H_1 = D_0 \cdot H_0 + 2 \cdot D_0 \cdot R_0 + \frac{1}{2} H_0 \cdot H_0 + R_0 \cdot H_0 = 2 \cdot \left(D_0 + \frac{1}{2} \cdot H_0\right) \cdot \left(R_0 + \frac{1}{2} \cdot H_0\right)$$

Ma, osservando che
abbiamo:
e inoltre

$$p_0 = \mathbf{D}_0 + \frac{1}{2}\mathbf{H}_0 \quad \text{e} \quad q_0 = \mathbf{R}_0 + \frac{1}{2}\mathbf{H}_0$$

$$\mathbf{1} \quad \mathbf{D}_1 = p_0^2 \quad \mathbf{H}_1 = 2p_0q_0 \quad \mathbf{R}_1 = q_0^2$$

$$p_1 = \mathbf{D}_1 + \frac{1}{2}\mathbf{H}_1 = p_0^2 + p_0q_0 = p_0 \cdot (p_0 + q_0)$$

cioè

2

$$\boxed{p_1 = p_0}$$

per cui anche

$$q_1 = q_0$$

La **2** ci dice che le frequenze alleliche non cambiano nel tempo,

la **1** che le frequenze genotipiche sono diretta funzione delle frequenze alleliche secondo lo sviluppo binomiale:

$$(p_A + q_a)^2 = p_A p_A + 2p_A q_a + q_a q_a$$

A questo risultato si poteva arrivare anche direttamente, considerando le probabilità di incontro dei due tipi di gameti. Possiamo costruire la seguente tabella, chiamata quadrato di Punnett:

(A)

Frequenze alleliche:		Sperma	
		A p	a q
Uova	A p	AA p^2	Aa pq
	a q	Aa pq	aa q^2

Frequenze alleliche:

(B)

Frequenze alleliche:		Sperma	
		A p	a q
Uova	A p	AA p^2	Aa pq
	a q	Aa pq	aa q^2

Frequenze alleliche:

Frequenze nella progenie: $AA : P' = p^2$
 $Aa : H' = pq + pq = 2pq$
 $aa : Q' = q^2$

Figura 2.9 Quadri di moltiplicazione incrociata che mostrano graficamente come si generano, secondo i presupposti di Hardy-Weinberg, le frequenze genotipiche (aree interne) dalle frequenze alleliche dei gameti. Nel quadrato a sinistra, le frequenze alleliche sono uguali (i.e., $p = q = 0.5$) mentre, nel quadrato a destra, l'allele A è tre volte più frequente di a ($p = 0.75$; $q = 0.25$).

1- gli incroci avvengono casualmente.

Ciò significa che non vi sono accoppiamenti che avvengono preferenzialmente, cioè più frequentemente di quanto non preveda l'incontro casuale di due individui. Questo si esprime dicendo che nella popolazione si ha **panmissia**; cioè significa anche che tutti i membri della popolazione hanno la stessa probabilità di avere discendenti.

2- la frequenza di un allele nei gameti è eguale alla frequenza dell'allele nella generazione parentale;

ciò significa che non esistono fenomeni come mutazione, migrazione, selezione, che fanno cambiare di per sé il valore di p .

3- inoltre abbiamo assunto che le frequenze dei genotipi alla generazione successiva sono pari alle probabilità di incontro dei diversi gameti;
ciò significa escludere variazioni dovute puramente al caso.

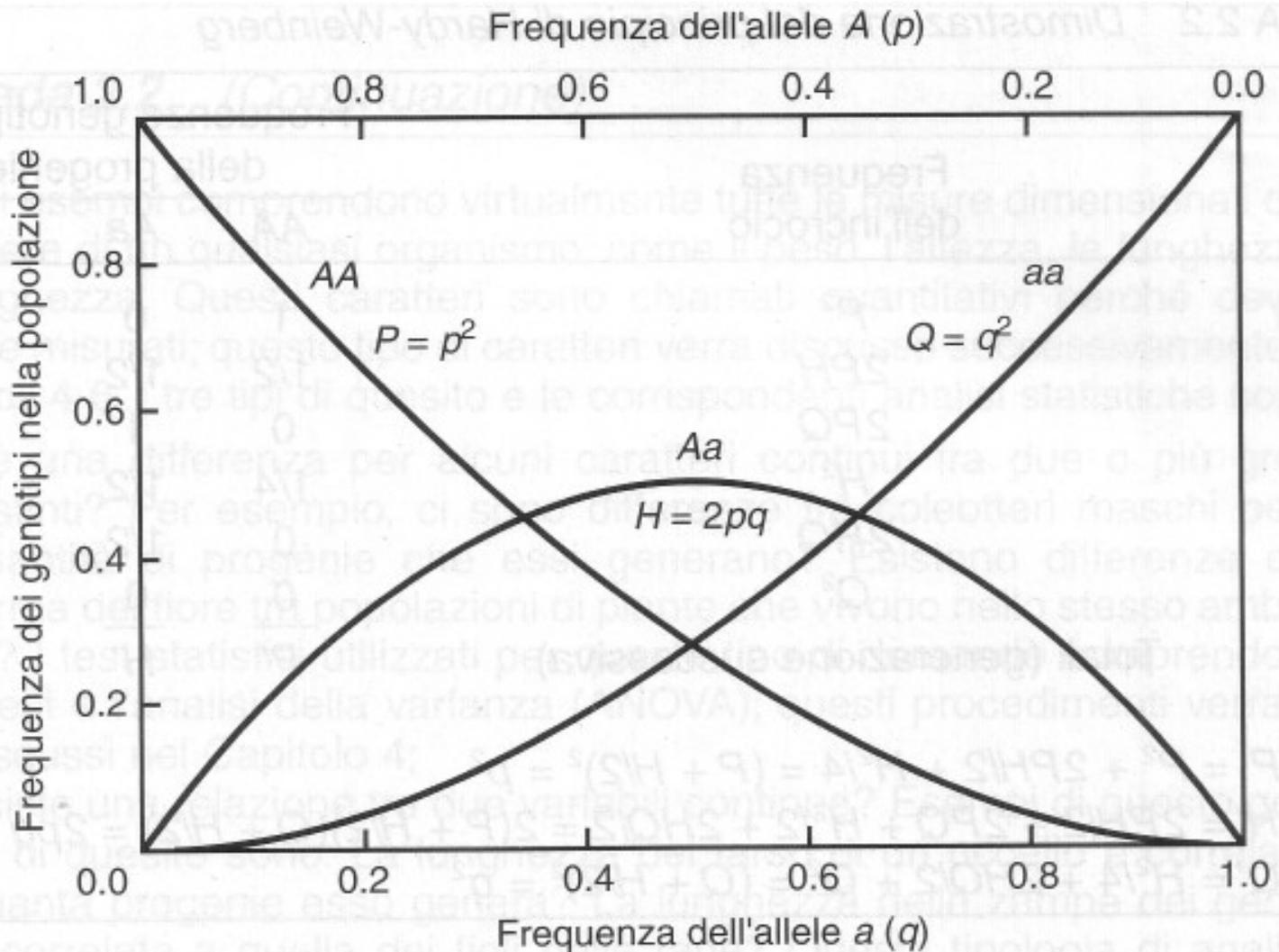


Figura 2.10 Frequenze genotipiche al variare delle frequenze alleliche per due alleli in equilibrio di Hardy-Weinberg. Nota che quando un allele è raro la frequenza degli omozigoti per quell'allele è molto più bassa di quella degli eterozigoti.

La prima condizione può non essere sempre soddisfatta.

Invece di considerare un gene consideriamo un altro elemento genetico, gli **eterocromosomi**. In forme ad eterogametia maschile sono possibili solo gli incroci $XX \times XY$, mentre non possono realizzarsi gli incroci $XX \times XX$ e $XY \times XY$.

Questo porta a far sì che, partendo da qualunque composizione di maschi e di femmine, se non si verificano altri fattori che favoriscono gli uni o le altre, si arriverà sempre ad una situazione in cui $p_X = 0.75$, $q_Y = 0.25$.

Più in generale vedremo che se sono favoriti gli accoppiamenti tra simili si arriva ad una diminuzione degli eterozigoti e ad un aumento degli omozigoti;

non è quindi rispettato il fatto che le frequenze genotipiche sono uno sviluppo binomiale delle frequenze alleliche.

Se invece sono favoriti gli incroci tra dissimili, aumentano le frequenze degli eterozigoti e si può arrivare a frequenze alleliche di valore determinato.

Mutazione e migrazione possono cambiare sistematicamente la frequenza degli alleli.

Se consideriamo solo due alleli possiamo avere la mutazione $A \rightarrow a$ e la retromutazione $a \rightarrow A$. **Aumenterà continuamente la frequenza dell'allele verso il quale si ha il tasso di mutazione maggiore.**

Se abbiamo due popolazioni e da una delle due migra costantemente un certo numero di individui verso la seconda, nel tempo la seconda raggiungerà la stessa struttura genetica della prima.

Abbiamo diversi tipi di **selezione**: contro un allele, a favore di un genotipo, contro un certo tipo di incroci.

Vedremo come i risultati possono essere vari e possono portare alla perdita di un allele oppure a situazioni in cui sono stabili solo determinati tipi di strutture genetiche.

La presenza di questi fenomeni determina quindi **cambiamenti sistematici**.

La terza condizione può essere soddisfatta solo in termini statistici.

La popolazione alla generazione successiva è costituita da un campione casuale degli alleli della generazione precedente.

Illustriamo questo concetto con un esempio banale: se ho una moneta e la lancio ho 50 % di probabilità che esca testa e 50 % che esca croce;

ma se lancio contemporaneamente 100 monete sarà eccezionale il caso che escano esattamente 50 testa e 50 croce: ci si avvicina a questo risultato solo come una media su un numero molto grande di casi;

100 eterozigoti non trasmetteranno alla discendenza esattamente 50 gameti di un tipo e 50 dell'altro, **si avrà soltanto la stessa probabilità di trasmettere l'uno o l'altro tipo di gamete.**

Ciò significa dire che possono verificarsi **cambiamenti dispersivi** della frequenza degli alleli.

L'entità di tali cambiamenti dipende dalle dimensioni di una popolazione: tanto più numerosi sono gli individui che costituiscono la popolazione, tanto più piccola sarà la probabilità che il cambiamento sia significativo.

Non si verificano cambiamenti di questo tipo solo in popolazioni di dimensioni infinite.

LEGGE DI HARDY-WEINBERG

in una popolazione, in cui non si abbiano migrazione, mutazione, selezione, che sia panmittica ed in cui non si abbiano fenomeni dispersivi, le frequenze alleliche rimangono costanti nel tempo e le frequenze genotipiche rispettano una distribuzione binomiale.



Lo studio delle differenze di struttura genetica tra popolazioni può fornire informazioni sulla loro storia ed evoluzione ma dà anche suggerimenti sui rapporti tra determinate strutture genetiche e l'ambiente.

La genetica di popolazioni non si occupa semplicemente di descrivere tale struttura, ma di vedere come possono cambiare tali strutture, in relazione agli scarti della panmissia ed ai cambiamenti di tipo sistematico e dispersivo.