



Struttura genetica delle popolazioni

“Vi sono pochi dubbi che il sistema genetico in grado di promuovere una maggiore flessibilità evolutiva, basato sulla fecondazione incrociata, rappresenti quello ancestrale e che i tipi caratterizzati da autofecondazione siano invece quelli sviluppatisi ripetutamente a vari livelli nel regno vegetale, verosimilmente come risposta alla pressione selettiva in favore di un adattamento immediato.”

George Ledyard Stebbins (1950)

Le caratteristiche del materiale ereditario possono essere studiate a diversi livelli: singolo individuo, gruppo di individui di una progenie o di una popolazione. Le proprietà del gene nell'individuo sono indicate dal carattere che il gene stesso controlla, come ad esempio la sintesi di un particolare enzima, la presenza o l'assenza di un pigmento o la forma di un organo. A livello di progenie, le proprietà di un gene possono riguardare le relazioni di dominanza o recessività, i modelli e le frequenze di segregazione, lo stato autosomico o legato al sesso, la posizione cromosomica rispetto ad altri geni. Nelle popolazioni è invece possibile studiare la frequenza caratteristica di un particolare gene: tale frequenza è strettamente dipendente dalla presenza del gene in molti o in pochi degli individui che compongono la popolazione. La conoscenza delle frequenze di singoli geni in più popolazioni di ambienti simili o differenti consente di confrontare le frequenze che caratterizzano le varie popolazioni e di associare la variazione riscontrata nelle frequenze geniche con determinate caratteristiche degli individui che compongono la popolazione o che vivono in condizioni ambientali peculiari.

La genetica di popolazione si occupa dell'eredità dei caratteri controllati da uno o più geni in gruppi di individui, basandosi sull'applicazione dei principi mendeliani a livello di popolazione. In particolare, studia la frequenza dei geni e dei genotipi a livello di popolazioni, e le variazioni che tali frequenze possono subire nel corso di generazioni successive. La genetica di popolazioni fornisce, pertanto, i modelli di distribuzione della variabilità all'interno di gruppi di individui, cioè la struttura genetica delle popolazioni, e di ripartizione della variabilità tra popolazioni, ma anche come questi modelli cambiano cioè si evolvono nel tempo e nello spazio.

Le acquisizioni teoriche conseguenti alla formulazione della legge dell'equilibrio da parte del matematico inglese Godfrey H. Hardy e del medico tedesco Wilhelm Weinberg risultarono determinanti per la nascita della genetica di popolazioni. Nel 1908, Hardy e Weinberg dimostrarono indipendentemente che in una popolazione le frequenze geniche ad un particolare locus si mantengono costanti nel corso delle generazioni e che le frequenze genotipiche possono essere previste a partire da quelle geniche qualora vengano però rispettate un insieme di condizioni. Ronald A. Fisher, J.B.S. Haldane, Sewall Wright studiarono invece l'evoluzione dal punto di vista genetico: nel decennio 1920–30, Fisher, Haldane e Wright elaborarono relazioni matematiche in grado di esprimere la teoria della selezione naturale, individuando varie modalità alla base dei processi selettivi.

11.1 Struttura genetica delle popolazioni naturali

Una popolazione naturale rappresenta una entità continua nello spazio e nel tempo. Una popolazione, nell'ambito di una specie, è infatti costituita da un insieme di individui che occupano un certo spazio e la cui successione nel tempo è assicurata dal susseguirsi delle generazioni. Tale continuità dipende dalla riproduzione, cioè dalla capacità di un individuo di originare nuovi individui.

Nel passaggio da una generazione all'altra si osservano due fenomeni caratteristici che si possono indicare come **eredità** e **variabilità**. La trasmissione dei caratteri genetici assicura che gli individui della discendenza siano in una certa misura somiglianti ai genitori e tra di loro. Tuttavia, un individuo può mostrare, rispetto ai genitori ed agli altri individui della discendenza, alcune differenze più o meno marcate. L'entità della variabilità che scaturisce dalla riproduzione e quindi la struttura genetica delle popolazioni è fortemente influenzata dal sistema riproduttivo degli individui che la compongono e dal tipo di unioni che si verificano al loro interno durante la riproduzione. Sotto questo punto di vista, le popolazioni vegetali di specie prevalentemente autogame sono fondamentalmente differenti da quelle di specie prevalentemente allogame, oltre che dalle popolazioni di specie che si moltiplicano per apomissia o che si propagano per parti vegetative (**Fig. 11.11**).

Il sistema riproduttivo (→ Cap. 10) influenza la struttura genetica delle popolazioni naturali determinando il modo e il grado con cui il potenziale di ricombinazione viene accumulato, mantenuto e liberato sotto l'influenza dell'ambiente; condiziona il raggiungimento di un compromesso tra capacità adattativa immediata di una specie nel suo ambiente e flessibilità verso future modificazioni dell'ambiente stesso. Dalla struttura genetica delle popolazioni dipendono, quindi, le strategie di miglioramento genetico che possono essere applicate, i tipi di varietà che possono essere costituite e gli schemi di selezione conservatrice e produzione commerciale del seme che debbono essere attuati.

La genetica di popolazione è la disciplina che studia la frequenza dei geni e dei genotipi in gruppi di individui, e la variazione di tali frequenze in generazioni successive, basandosi sull'applicazione dei principi mendeliani a livello di popolazione. Un gruppo di individui in cui gli alleli a singoli loci seguono le leggi dell'eredità mendeliana

costituisce una **popolazione mendeliana**. Le caratteristiche di dominanza e recessività di un gene, così come il suo modello di segregazione o la sua posizione sul cromosoma rispetto ad altri geni, nonché tutte le altre caratteristiche che vengono rilevate a livello di un singolo individuo e di una singola discendenza, si mantengono anche a livello della popolazione. Considerando, ad esempio, un locus con due possibili alleli, le relazioni tra questi possono essere di dominanza e recessività o di additività (assenza di dominanza), così come questi possono trovarsi allo stato omozigote oppure eterozigote, tanto nel singolo individuo che nel gruppo di individui. Il modello più semplice di popolazione mendeliana prende in considerazione un unico locus con due alleli, che possono combinarsi per formare tre genotipi, due omozigoti per alleli diversi e uno eterozigote. In presenza di dominanza gli

alleli possono essere indicati con una lettera maiuscola e minuscola dell'alfabeto (ad esempio, A e a), mentre in assenza di dominanza con una lettera maiuscola o minuscola e un simbolo o un indice numerico (ad esempio, A e A' o a e a' , oppure A_1 e A_2 o a_1 e a_2). Una parte consistente dei modelli adottati nella genetica delle popolazioni si riferisce proprio ad un unico locus con due alleli, ammettendo implicitamente che quanto è valido per un unico locus sia valido per la maggioranza dei loci.

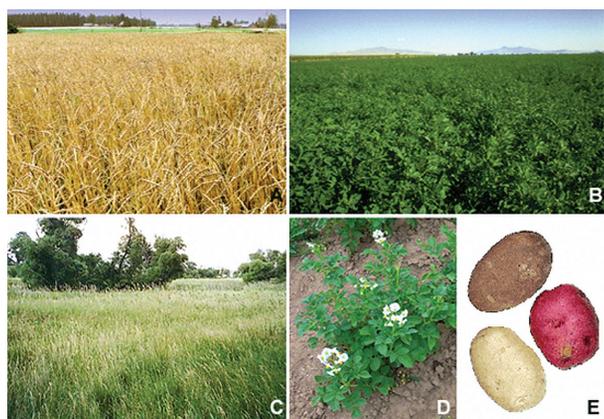


Fig. 11.1 – Esempi di popolazioni coltivate di specie autogame (A, farro), allogame (B, erba medica), apomittiche (C, poa pratense) moltiplicate per seme e di specie (D, patata) propagata per parti vegetative (E, tuberi).

Secondo la definizione di T. Dobzhansky, una popolazione si identifica, più precisamente, con un gruppo di individui sessuati ed interfertili che condividono un **pool genico**, cioè un insieme di alleli comune. Per studiare la composizione di una tale popolazione dal punto di vista genetico è quindi necessario descrivere quantitativamente il pool genico calcolando le frequenze genotipiche e le frequenze alleliche che la caratterizzano. Poiché una popolazione può essere considerata come un complesso di genotipi che sono derivati da una collezione di gameti maschili e femminili, per un dato locus è possibile calcolare sia il numero e la frequenza relativa dei genotipi che il numero e la frequenza relativa degli alleli.

Le popolazioni naturali di specie propagate vegetativamente presentano una variabilità piuttosto contenuta poiché sono generalmente costituite da singoli cloni o da una mescolanza di cloni. In queste popolazioni nuova variabilità può comunque insorgere continuamente attraverso il seme, a meno che la riproduzione sessuale non risulti completamente assente, come nel caso delle specie apomittiche obbligate.

Le popolazioni naturali di specie anfimittiche sono invece caratterizzate da una grande variabilità. In tali specie, la struttura genetica delle popolazioni naturali è comunque fortemente influenzata dal tipo di unioni che si verificano al loro interno. Ciò significa che le popolazioni naturali di specie autogame, cioè che si riproducono prevalentemente mediante autofecondazione (come, ad esempio, frumento, tabacco e pisello) sono profondamente diverse da quelle di specie allogame, che invece si riproducono prevalentemente mediante fecondazione incrociata (ad esempio, mais, erba medica, patata e radicchio). Tali popolazioni si differenziano, infatti, sia in termini di costituzione genotipica dei singoli individui che di ripartizione della variabilità genetica tra individui entro singole popolazioni e tra popolazioni distinte. Dal punto di vista fenotipico, la variabilità esistente entro le popolazioni di specie allogame è generalmente più grande di quella riscontrabile entro popolazioni di specie autogame. D'altronde, nelle autogame l'autofecondazione rappresenta un potente ed efficace fattore di isolamento tra individui, mentre nelle allogame gli individui possono incrociarsi liberamente.

11.2 Dinamica delle popolazioni

Oltre al tipo di riproduzione, un altro elemento che ha notevoli ripercussioni sulla struttura genetica delle popolazioni concerne la variazione che il numero dei componenti delle popolazioni subisce in generazioni successive per effetto della lunghezza del ciclo vitale caratteristico della specie. Questo aspetto di dinamica delle popolazioni è molto importante per lo studio della struttura genetica delle popolazioni in quanto le variazioni a carico delle frequenze geniche e genotipiche sono spiegabili soprattutto in funzione di variazioni nella frequenza degli individui portatori di certi geni o riconducibili a determinati genotipi. Il numero di individui che compongono una popolazione può variare nel tempo e nello spazio, cioè in generazioni diverse e in ambienti diversi. I processi che provocano tali cambiamenti sono gli elementi della dinamica di popolazioni e l'aspetto più importante è rappresentato dalla lunghezza del ciclo vitale: in questo senso le specie annuali si distinguono nettamente da quelle poliennali. Gli individui di specie annuali nascono, si sviluppano e dopo la riproduzione muoiono senza sovrapposizione del ciclo vitale dei parentali con quello della propria discendenza (popolazioni a generazioni separate). Nel caso delle specie poliennali si ha invece sovrapposizione del ciclo vitale tra parentali e individui della discendenza anche nello stesso ambiente e per periodi di tempo che possono essere anche piuttosto lunghi (popolazioni a generazioni sovrapposte). Assumendo un tasso di riproduzione (w) costante è possibile definire dei modelli di dinamica delle popolazioni a generazioni

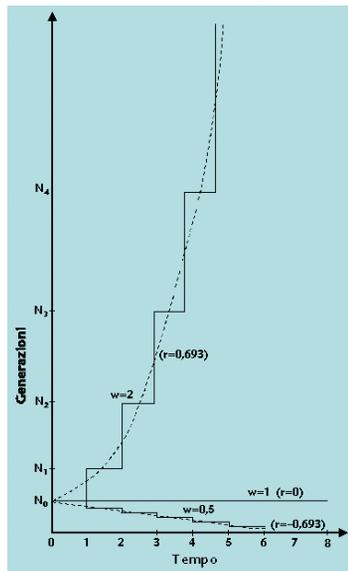


Fig. 11.2 – Dinamica di popolazioni a generazioni discontinue (—) e continue (---). w è il tasso di riproduzione (nelle tre popolazioni considerate pari a 2, 1 e 0,5), mentre r è il parametro malthusiano uguale al logaritmo naturale di w .

separate e sovrapposte (**Fig. 11.2**). Il tasso di riproduzione, che equivale al valore adattativo e alla capacità riproduttiva di un individuo, si indica col termine *fitness* (w) ed è misurato come contributo di individui dato alla discendenza. La conoscenza della fitness media (\bar{w}) di una popolazione formata da un gruppo numeroso di individui consente di formulare ipotesi circa la sua dinamica nel tempo: la popolazione si estingue quando $\bar{w} < 1$, aumenta in maniera esponenziale se $\bar{w} > 1$ e rimane di numerosità costante quando il tasso di riproduzione $\bar{w} = 1$, nel caso di generazioni separate (o discontinue), oppure quando $\bar{w} = 0$, nel caso di generazioni sovrapposte (continue). Il modello di popolazioni a generazioni separate è valido per molti organismi, quali ad esempio le piante annuali, tuttavia non è il modello più generale in quanto in numerose specie vegetali individui parentali e discendenze di diverse generazioni possono coesistere, per un periodo di tempo anche lungo, nello stesso ambiente. Ambedue i modelli richiedono il verificarsi in natura di condizioni poco realistiche poiché assumono che il tasso di riproduzione sia uguale per tutti gli individui e resti costante per tutte le generazioni. Nelle popolazioni naturali di specie vegetali l'estinzione è un fenomeno che si verifica piuttosto raramente e quando ciò accade è spesso riconducibile a cambiamenti ambientali consistenti che pregiudicano la capacità di adattamento degli individui. Anche l'esplosione, cioè l'aumento considerevole e illimitato degli individui di una popolazione, è un fenomeno poco frequente come conseguenza della competizione che si instaura tra individui di uno stesso ambiente nel corso delle generazioni. La tendenza delle popolazioni naturali è quella di diminuire la velocità di crescita con l'aumentare della numerosità della popolazione stessa.

L'obiettivo della genetica di popolazioni è lo studio della variazione delle frequenze geniche e genotipiche. La capacità riproduttiva degli individui ha influenze notevoli sulla dinamica e sulla struttura genetica delle popolazioni in quanto la frequenza di singoli geni in una popolazione è determinata dalla fitness dei genotipi che li portano. Altro aspetto importante in grado di condizionare la struttura genetica delle popolazioni è connesso al tipo di riproduzione e di generazioni che caratterizzano una specie: nel caso di specie autogame, la differenza tra generazioni discontinue e continue non è molto influente poiché gli individui di una popolazione possono solo autofecondarsi, mentre nel caso di specie allogame la differenza è consistente dato che gli individui possono incrociarsi liberamente. Nelle popolazioni a generazioni continue può verificarsi anche l'unione tra individui parentali e individui appartenenti a una o più discendenze diverse, mentre questo non è possibile nelle popolazioni a generazioni discontinue dove le unioni interessano unicamente gli individui di una stessa discendenza.

11.3 Struttura genetica delle popolazioni di specie apomittiche e a propagazione vegetativa

Le popolazioni di specie apomittiche e di specie propagate vegetativamente sono composte da uno o più gruppi di piante geneticamente identiche tra loro, derivate da un unico capostipite per mezzo di semi apomittici o di parti vegetative (ad esempio, rizomi, bulbi, tuberi e stoloni). Le popolazioni naturali di queste specie sono generalmente costituite da una mescolanza di **cloni** (popolazioni multiclionali o policlionali) e il numero complessivo di genotipi clonali può essere anche molto alto (**Fig. 11.3**). Sia la riproduzione asessuata che la propagazione vegetativa permettono la clonazione e la fissazione dei genotipi materni e rappresentano quindi un sistema genetico che consente la rapida selezione degli individui con elevata adattabilità all'ambiente, come del resto consente anche l'autofecondazione. Queste popolazioni dal punto di vista temporale presentano variabilità ambientale sia tra che entro cloni e variabilità

genetica unicamente tra cloni, salvo l'insorgenza di mutazioni spontanee. Infatti, sia con la riproduzione apomittica che con la propagazione vegetativa non viene rilasciata variabilità genetica benché tale variabilità possa in realtà essere presente nella popolazione allo stato potenziale, cioè in condizione eterozigote. Le popolazioni di specie apomittiche e di specie propagate vegetativamente sono composte di genotipi altamente eterozigoti e segregano ampiamente: il rilascio di variabilità presente nel clone avviene, nel caso di specie a propagazione vegetativa, tutte le volte che si ricorre al seme (sia in seguito ad autofecondazione che incrocio) mentre, nel caso di specie apomittiche, in seguito ad eventi sessuali occasionali (fecondazione incrociata, partenogenesi aploide, ecc.). In questi casi, comunque, il processo riproduttivo comporta anche la formazione di una quota considerevole di individui con bassa vitalità o adattabilità all'ambiente.

11.4 Struttura genetica delle popolazioni di specie prevalentemente autogame

Le popolazioni naturali di specie autogame sono normalmente costituite da una mescolanza di linee omozigoti (linee pure) strettamente imparentate che, sebbene viventi l'una accanto all'altra, rimangono più o meno indipendenti nella riproduzione. In queste popolazioni tutti gli individui sono il risultato di infinite generazioni di autofecondazione e si possono pertanto supporre omozigoti a tutti i loci, anche se per alleli diversi. Ciascun individuo possiede la capacità di autofecondarsi ed è quindi in grado di produrre una discendenza uniforme, cioè una **linea pura**. Il processo di autofecondazione continuata assicura l'omozigosi e fornisce discendenze per lo più omogenee: la variabilità genetica è quindi concentrata tra le linee ed è pertanto dovuta a differenze tra i genotipi omozigoti. Le piante tollerano l'unione tra individui imparentati (*inbreeding*) e presentano un ottimo adattamento all'ambiente in cui si trovano, ma sono poco flessibili nei confronti di ulteriori cambiamenti. L'incrocio occasionale e la mutazione spontanea, seguiti dalla segregazione e dalla ricombinazione, sono i due principali fattori che si oppongono al raggiungimento della situazione limite di omozigosi a tutti i loci. Tali fattori, creando nuova variabilità, danno infatti alla popolazione la possibilità di evolversi e di adattarsi a condizioni ambientali variabili.

Studi sulle mutazioni hanno comunque portato a concludere che la frequenza di gameti mutati ad un singolo locus è molto bassa e che la maggior parte dei tipi mutati viene eliminata dalla pressione selettiva naturale quando si manifestano. Poiché è improbabile che entrambi gli alleli ad un locus subiscano lo stesso cambiamento, risulta che la mutazione insorge allo stato eterozigote ed essendo di norma recessiva – l'allele selvatico (+) si comporta in genere da dominante – si esprimerà solo quando raggiunge la condizione omozigote. In **Tab. 11.1** è schematizzata la riduzione dell'eterozigosi ad

Generazione	Individui omozigoti (%)	Individui eterozigoti (%)
F ₁	0,00	100,00
F ₂	50,00	50,00
F ₃	75,00	25,00
F ₄	87,50	12,50
F ₅	93,75	6,25

		1600 +m		
	400 ++	800 +m	400 mm	
	400 ++	200 ++	400 +m	200 mm
	600 ++	100 ++	200 +m	100 mm
	700 ++	50 ++	100 +m	50 mm

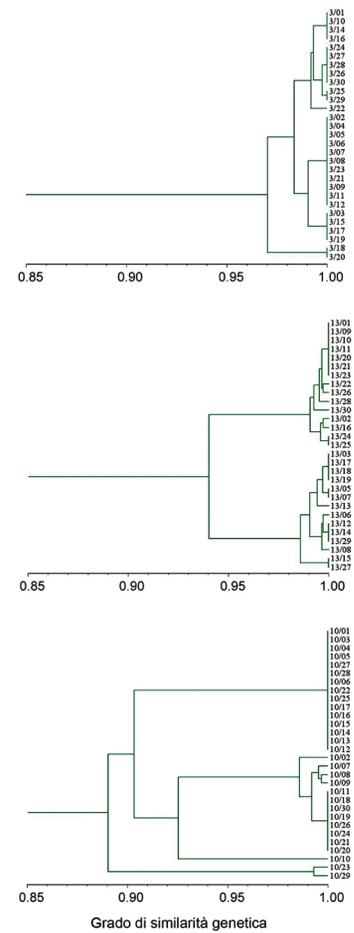


Fig. 11.3 – Struttura genetica di ecotipi di una specie apomittica facoltativa (iperico). Le tre popolazioni risultano multiclonali, non dominate da un singolo genotipo, e sono caratterizzate da diverso grado di similarità genetica media (97%, 94% e 89%).

Tab. 11.1 – Riduzione dell'eterozigosi, determinata ad un dato locus da un evento di mutazione, nel corso di successive generazioni di autofecondazione.

Tab. 11.2 – Frequenza di alcune mutazioni spontanee a carico di caratteri dell'endosperma di mais.

Geni	No. di gameti osservati	No. di mutazioni osservate	No. di mutazioni per milione di gameti
R (fattore del colore)	554.786	273	492,0
I (inibitore del colore)	265.391	28	106,0
P (violaceo)	647.102	7	11,0
Su (zuccherino)	1.678.736	4	2,4
Y (giallo)	1.745.280	4	2,2
Sh (grinzoso)	2.469.285	3	1,2
Wx (ceroso)	1.503.744	0	0,0

un singolo locus in successive generazioni di autofecondazione. Ammettendo che la mutazione non abbia alcuna influenza sulla capacità riproduttiva degli individui, ad ogni generazione la quota di eterozigoti si dimezza e quella degli omozigoti aumenta invece progressivamente, così che ad esempio dopo cinque generazioni di autofecondazione gli individui eterozigoti rappresentano il 6,25% e quelli omozigoti per la mutazione e per l'allele selvatico costituiscono nel loro complesso il 93,75% della popolazione. La quota di eterozigoti dopo n generazioni di autofecondazione è quindi pari a: $x = (1/2)^n$. La mutazione può essere considerata un processo ricorrente che, in ogni specie coltivata, ha portato alla formazione di serie alleliche multiple. Nel corso del processo evolutivo le mutazioni sono sottoposte all'azione della selezione naturale.

Uno dei primi ricercatori a fornire informazioni su questo argomento è stato L.J. Stadler che nel 1942 studiò la frequenza di alcune mutazioni spontanee a carico di geni che controllano specifici caratteri dell'endosperma di mais, un materiale particolarmente favorevole perché l'esame delle cariossidi di una sola spiga è sufficiente a definire parecchie centinaia di genotipi (Tab. 11.2). I dati suggeriscono che la frequenza delle mutazioni spontanee dipende dal carattere considerato e comunque per un singolo locus è generalmente molto bassa, con alcune eccezioni. Attualmente, prendendo in considerazione unicamente i geni per i quali sono disponibili molti dati, il tasso mutazionale per gli eucarioti risulta variabile tra 10^{-4} e 10^{-7} per generazione. Sebbene la loro insorgenza sia rara e la loro frequenza risulti bassa, nel lungo periodo l'accumulo delle mutazioni è in grado di spiegare la variabilità che si riscontra tra linee pure nelle popolazioni naturali di specie autogame.

L'effetto dell'incrocio rimane invece a lungo nella popolazione, anche in considerazione del fatto che gli eterozigoti presentano sempre un certo vantaggio selettivo sugli omozigoti in termini di *fitness*, cioè di capacità riproduttiva. In una popolazione dove una particolare coppia allelica è presente con frequenze pari a $p(A)$ e $q(a)$, se la quota di autofecondazione è s e quella di incrocio è $c = (1-s)$, la proporzione degli

Tab. 11.3 – Frazione degli eterozigoti presenti in condizioni di equilibrio in popolazioni autogame con diverse frequenze geniche (p) e quote diverse di incrocio casuale (c) (da: R. Allard, 1968).

$c \backslash p$	0,05	0,10	0,20	0,30	0,40	0,50
0,1	0,0173	0,0327	0,0582	0,0764	0,0783	0,0909
0,2	0,0317	0,0600	0,1067	0,1400	0,1600	0,1667
0,3	0,0439	0,0831	0,1477	0,1939	0,2216	0,2308
0,4	0,0543	0,1029	0,1829	0,2610	0,2743	0,2857
0,5	0,0633	0,1200	0,2133	0,2800	0,3200	0,3333
0,6	0,0713	0,1350	0,2400	0,3150	0,3600	0,3750
0,7	0,0783	0,1482	0,2636	0,3459	0,3953	0,4118
0,8	0,0844	0,1600	0,2844	0,3733	0,4266	0,4444
0,9	0,0900	0,1705	0,3022	0,3969	0,4548	0,4737
1,0	0,0950	0,1800	0,3200	0,4200	0,4800	0,5000

eterozigoti che viene teoricamente mantenuta nella popolazione dopo moltissime generazioni ($n = \infty$), in condizioni di equilibrio è la seguente:

$$x = 2pq \cdot \frac{2c}{(2-s)}$$

Ad esempio, con frequenze geniche $p = q = 0,5$ e una quota di incrocio $c = 0,1$ (10%), la proporzione di eterozigoti che viene mantenuta nella popolazione in condizioni di equilibrio è pari a 0,09 (9%). Le frazioni teoriche degli eterozigoti relative a diverse quote di incrocio e frequenze geniche sono riportate in **Tab. 11.3**. In realtà, la situazione in natura può discostarsi da quella descritta a causa della constatazione che un aumento della frequenza degli eterozigoti nella popolazione spesso determina una loro minore capacità riproduttiva: tale capacità è molto alta quando gli eterozigoti costituiscono il 2–4% della popolazione, mentre sopra il 10% non differisce dalla capacità riproduttiva degli omozigoti.

Studiando la manifestazione di caratteri morfologici evidenti, nel 1970 Jain e Allard hanno dimostrato che in una popolazione di orzo, avente una quota di incrocio del 2%, la fitness degli omozigoti a singoli loci è inferiore rispetto a quella degli eterozigoti corrispondenti (**Tab. 11.4**). Infatti, ponendo pari ad 1,00 la fitness media degli eterozigoti A_1A_2 agli otto loci considerati, la fitness media degli omozigoti risulta pari a 0,81 per A_1A_1 e 0,82 per A_2A_2 . Il vantaggio medio degli eterozigoti rispetto agli omozigoti è dato dal rapporto $0,18/0,82=0,22$, cioè pari al 22%. Tuttavia, nel 1966 Harding e collaboratori hanno dimostrato che la fitness degli eterozigoti diminuisce con l'aumentare della loro frequenza nella popolazione (**Fig. 11.4**). I dati acquisiti in fagiolo dimostrano chiaramente che quando la frequenza degli eterozigoti si aggira attorno al 2–3% la loro fitness è addirittura tripla rispetto a quella degli omozigoti, ma che quando la frequenza degli eterozigoti raggiunge il 10–15% la loro fitness è simile a quella degli omozigoti.

In generale, nelle popolazioni naturali di specie autogame gli eterozigoti hanno una frequenza molto contenuta e presentano sempre un certo vantaggio selettivo rispetto agli omozigoti.

11.4.1 Autofecondazione e omozigosi

Nelle specie strettamente autogame, la mutazione spontanea e l'incrocio occasionale sono i principali fattori che provocano l'insorgenza della condizione ibrida, rispettivamente, a singoli loci o a un numero consistente di loci. Tali fattori, creando eterozigosi, fanno sì che la popolazione vada incontro a processi di segregazione e ricombinazione durante le generazioni riproduttive.

Nelle popolazioni ibride di piante autogame l'eterozigosi diminuisce nel corso di successive generazioni di autofecondazione, mentre la frazione di omozigoti ai loci inizialmente ibridi segue una precisa espressione matematica. Considerando un genotipo ibrido con n coppie alleliche in condizione eterozigote, la frazione di piante omozigoti a tutti i loci dopo m generazioni segreganti è pari a:

$$x = \left[\frac{(2^m - 1)}{2^m} \right]^n$$

Tab. 11.4 – Fitness degli omozigoti e degli eterozigoti a singoli loci che controllano caratteri morfologici evidenti in una popolazione di orzo.

Locus	Fitness dei genotipi		
	A_1A_1	A_1A_2	A_2A_2
B/b	1,06	1,00	1,31
S/s	0,81	1,00	0,96
G/g	1,04	1,00	0,82
E/e	0,47	1,00	0,59
B1/b1	0,61	1,00	0,54
R/r	0,82	1,00	0,68
Bt/bt	0,96	1,00	1,06
Sh/sh	0,71	1,00	0,63

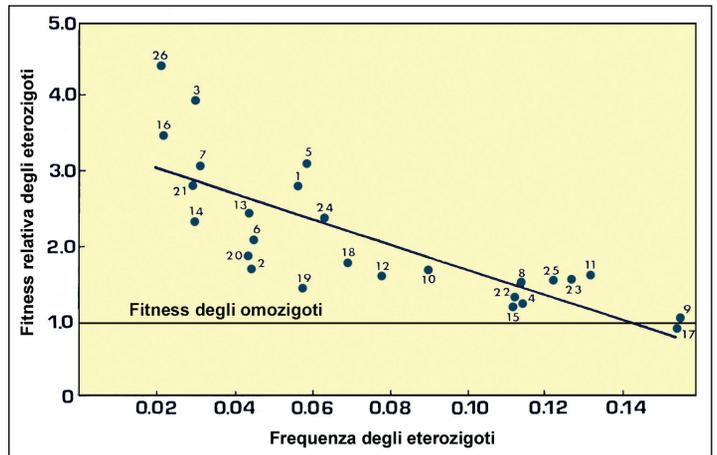


Fig. 11.4 – Relazione tra frequenza nella popolazione e fitness relativa degli eterozigoti al locus *S/s* in *Phaeolus limensis*.

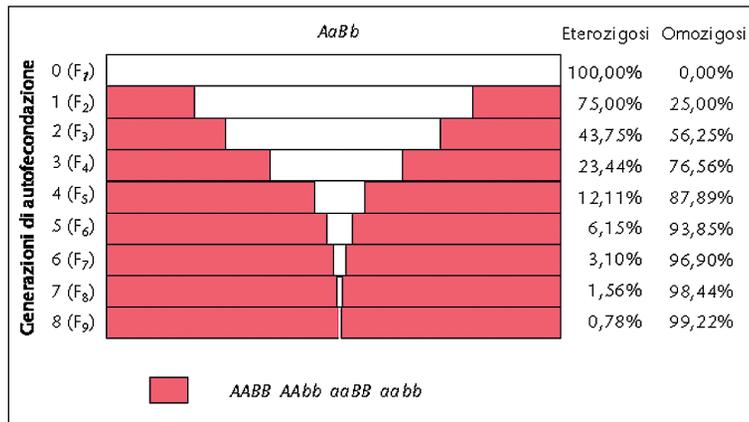
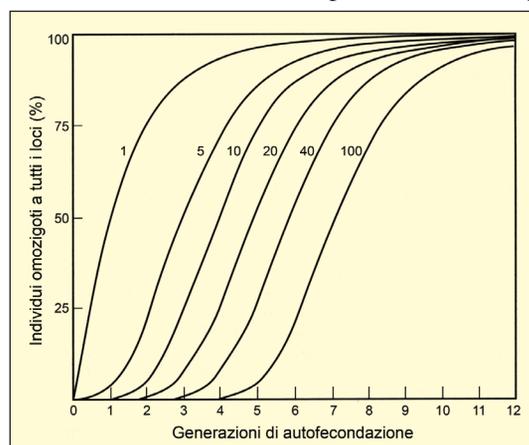


Fig. 11.5 – Aumento della frequenza degli omozigoti attraverso ripetute autofecondazioni a partire da un diibrido *AaBb*.

l'autofecondazione continuata fa ottenere, dopo un numero molto alto di generazioni, una popolazione costituita quasi esclusivamente da omozigoti. La percentuale di individui omozigoti a tutti i loci dopo varie generazioni di autofecondazione in relazione alla complessità dell'ibrido iniziale può essere rappresentata graficamente ipotizzando un diverso numero di coppie alleliche inizialmente in condizione eterozigote ed ereditate in modo indipendente (**Fig. 11.6**). Dopo infinite generazioni di autofecondazione la popolazione risulterà costituita da un numero di linee pure omozigoti per alleli diversi, cioè di famiglie omozigoti per caratteri diversi, pari a $x = 2^n$ dove n indica il numero di loci inizialmente in condizione ibrida. Tale valore corrisponde ai tipi di gameti che può formare l'ibrido F_1 . Così per un solo locus ($n=1$), corrispondente al monoibrido, le linee pure possibili sono due, per il tetraibrido ($n=4$) sono 16 e così via. Per un poliibrido con $n=100$ le possibili linee pure risultano alcuni miliardi! In **Tab. 11.5** sono riassunte le percentuali di individui omozigoti a tutti i loci nel corso di successive generazioni di autofecondazione di individui con numero progressivamente crescente di loci in condizione eterozigote. In conseguenza di ripetute autofecondazioni, dopo una decina di generazioni segreganti, anche gli ibridi molto complessi risultano costituiti prevalentemente da genotipi omozigoti. Le popolazioni naturali di specie autogame, che si riproducono spontaneamente e da tempo illimitato per autofecondazione, risultano pertanto costituite esclusivamente da una mescolanza di linee omozigoti per alleli diversi. Tuttavia, le linee pure presenti nella popolazione sottoposta ad infinite generazioni di autofecondazione non saranno tutte ugualmente rappresentate, ma prevarranno le combinazioni parentali su quelle ricombinanti. Tra le discendenze prodotte per autofecondazione a partire da un incrocio iniziale tra due linee pure saranno infatti prevalenti gli individui caratterizzati da combinazioni di caratteri parentali in conseguenza dell'associazione dei geni corrispondenti nei singoli cromosomi dei parentali, secondo le proporzioni volute dalla segregazione indipendente. Gli individui aventi nuove combinazioni di caratteri per effetto del crossing-over e quindi in conseguenza della ricombinazione dei geni entro i singoli cromosomi saranno invece più rari. In ogni caso, la proporzione di ciascuna linea pura entro una popolazione derivante da infinite generazioni di autofecondazione dipende dalla sua fitness relativa.

Fig. 11.6 – Proporzione di individui omozigoti a tutti i loci dopo varie generazioni di autofecondazione con un numero crescente di coppie alleliche inizialmente ibride ereditate in maniera indipendente.



rispondenti nei singoli cromosomi dei parentali, secondo le proporzioni volute dalla segregazione indipendente. Gli individui aventi nuove combinazioni di caratteri per effetto del crossing-over e quindi in conseguenza della ricombinazione dei geni entro i singoli cromosomi saranno invece più rari. In ogni caso, la proporzione di ciascuna linea pura entro una popolazione derivante da infinite generazioni di autofecondazione dipende dalla sua fitness relativa.

No. di loci eterozigoti	No. di generazioni segreganti								
	1 (F ₂)	2 (F ₃)	3 (F ₄)	4 (F ₅)	5 (F ₆)	6 (F ₇)	7 (F ₈)	8 (F ₉)	9 (F ₁₀)
1	50,00	75,00	87,50	93,75	96,88	98,44	99,22	99,61	99,80
2	25,00	56,25	76,56	87,89	93,85	96,90	98,44	99,22	99,61
5	3,13	23,73	51,29	72,42	85,32	92,43	96,15	98,06	99,03
10	0,10	5,63	26,31	55,45	72,80	85,43	92,46	96,16	98,06
20	0,00	0,32	6,92	27,51	52,99	72,98	85,48	92,47	96,17
40	0,00	0,001	0,48	7,57	28,08	53,26	73,07	85,51	92,48
100	0,00	0,000	0,00	0,16	4,18	20,70	45,64	67,61	82,24

Tab. 11.5 – Percentuale di individui omozigoti a tutti i loci dopo varie generazioni di autofecondazione a partire da individui inizialmente ibridi ad un numero crescente di loci.

In generale, nelle popolazioni naturali di specie autogame è presente soprattutto una variabilità dovuta a differenze tra genotipi omozigoti e l'associazione tende a conservare le combinazioni geniche esistenti, limitando così la segregazione dei caratteri. Nelle specie che hanno una antica tradizione colturale, dove l'attività di selezione ha portato alla costituzione di varietà aventi geni utili a molti dei loci ubicati in uno stesso cromosoma, questa situazione deve essere considerata positivamente poiché consente di mantenere nel tempo le combinazioni geniche favorevoli ed è quindi un fattore di stabilità genetica. In effetti, le linee pure una volta fissate tendono a rimanere immutabili nel tempo. L'esperienza ha tuttavia dimostrato che l'incrocio occasionale e la mutazione spontanea sono due fattori che si oppongono al raggiungimento della situazione limite di omozigosi a tutti i loci. Tali fattori sono in grado di creare variabilità genetica all'interno delle linee pure, innescando nelle varietà coltivate fenomeni di degenerazione e offrendo alle popolazioni naturali la possibilità di evolversi continuamente attraverso i processi di segregazione e ricombinazione.

11.5 Struttura genetica delle popolazioni di specie prevalentemente allogame

Quando la popolazione è formata da individui a sessi separati (specie dioiche), oppure da individui ermafroditi con barriere riproduttive che impediscono l'autofecondazione, l'incrocio rappresenta la unica forma possibile di riproduzione o comunque quella prevalente. Qualora non esista unione preferenziale tra i gameti, tali popolazioni sono dette panmistiche poiché nell'ambito di una popolazione ognuno dei gameti maschili ha la stessa probabilità di accoppiarsi con uno dei possibili gameti femminili. In condizioni di **panmissia** le unioni sessuali tra piante sono totalmente casuali (*random mating*) e ciascuna cellula uovo di un individuo può essere potenzialmente fecondata da uno qualsiasi dei nuclei spermatici di un altro individuo della popolazione.

Nelle popolazioni naturali di specie allogame tutti gli individui sono eterozigoti ad un gran numero di loci e ciascun genotipo è differente in una certa misura dagli altri. La variabilità genetica è ampia ed è distribuita fra tutti gli individui. Per i caratteri quantitativi, la maggior parte della variabilità entro popolazioni è imputabile ad azioni geniche di tipo additivo, mentre più modesti sono gli effetti di dominanza ed epistasia.

La fecondazione incrociata determina la produzione di progenie geneticamente differenziate che solitamente includono anche una frazione con scarsa vitalità, non limitando tuttavia la flessibilità evolutiva e la capacità adattativa della specie a condizioni ambientali variabili. Tali popolazioni condividono un pool di geni comune, geni che si combinano in varia maniera nel corso delle generazioni successive. La frequenza dei singoli geni nel pool comune è determinata dalla capacità riproduttiva dei genotipi: quando i geni contribuiscono a produrre genotipi dotati di notevole valore

adattivo la loro frequenza tenderà a crescere, viceversa se hanno effetti negativi sul genotipo la loro frequenza tenderà a diminuire.

Le popolazioni naturali di tali specie sono caratterizzate, di norma, da unioni sessuali casuali che consentono di mantenere inalterate nel tempo le frequenze geniche e le frequenze genotipiche. Tali popolazioni vengono comunemente dette in **equilibrio Hardy-Weinberg** nel senso che la composizione genetica di una popolazione sufficientemente numerosa rimane costante, relativamente a ciascun gene, nel corso delle generazioni in presenza di unioni casuali ed in assenza di fattori di disturbo (migrazione, mutazione differenziale e selezione).

11.5.1 Frequenze geniche e genotipiche

Una popolazione mendeliana per un dato locus può evidenziare individui appartenenti ad uno dei tre possibili genotipi AA , Aa e aa che derivano dalle combinazioni a coppie degli alleli A e a portati singolarmente dai gameti. Il pool genico di una popolazione può essere caratterizzato ricorrendo al calcolo delle frequenze geniche e genotipiche. In questo modo è possibile evidenziare la relazione che esiste tra le frequenze degli alleli e dei genotipi nella popolazione in esame.

Il termine **frequenza genotipica** si usa per indicare la proporzione (o percentuale) di un certo genotipo in una data popolazione. Supponendo di avere una popolazione di N individui che abbia ad un dato locus la costituzione genotipica AA , Aa e aa oppure $A'A'$, $A'A$ e AA , le frequenze genotipiche D , H e R riportate in **Tab. 11.6** rappresentano le frequenze assolute, cioè il numero di individui che presentano un dato genotipo, così che la loro somma equivale alla dimensione della popolazione ($D+H+R=N$). In presenza di relazione di dominanza e recessività per il carattere considerato, i genotipi omozigoti dominanti si possono distinguere da quelli eterozigoti soltanto ricorrendo a particolari analisi genetiche o molecolari. Le frequenze genotipiche relative possono essere calcolate come rapporto tra la frequenza assoluta e il numero totale di individui della popolazione nel modo seguente:

$$\begin{aligned} f(AA) &= D/N = d && \rightarrow && \text{frequenza dei genotipi omozigoti } AA \\ f(Aa) &= H/N = h && \rightarrow && \text{frequenza dei genotipi eterozigoti } Aa \\ f(aa) &= R/N = r && \rightarrow && \text{frequenza dei genotipi omozigoti } aa \end{aligned}$$

In assenza di dominanza per il carattere considerato, le frequenze genotipiche relative sono calcolate in modo analogo, ma l'analisi della popolazione è più semplice poiché i genotipi omozigoti (AA e $A'A'$) sono entrambi distinguibili da quelli eterozigoti ($A'A$) anche a livello fenotipico. Naturalmente, in ogni caso, si ha la relazione: $d+h+r=1$ (Tab. 11.6). Tali frequenze si possono esprimere anche come percentuali. Quando si conoscono le frequenze genotipiche ed il modo con cui il carattere viene ereditato (presenza o assenza di dominanza) è possibile determinare il tipo e il numero relativo degli alleli che hanno dato origine alla popolazione.

La **frequenza genica** (sinonimo di frequenza allelica) definisce la frequenza di un gene ad un determinato locus ed equivale alla proporzione (o percentuale) di un certo allele sul totale degli alleli possibili al locus considerato nella popolazione. Le frequenze alleliche possono essere pertanto ricavate dal rapporto tra il numero di copie di un dato allele nella popolazione e il numero totale di alleli esistenti nella popolazione al locus considerato:

$$\begin{aligned} f(A) &= \sum A / \sum (A+a) = p && \rightarrow && \text{frequenza dell'allele } A \\ f(a) &= \sum a / \sum (A+a) = q && \rightarrow && \text{frequenza dell'allele } a \end{aligned}$$

Tab. 11.6 – Possibili genotipi ad un locus con terminologia corrispondente alla presenza o all'assenza di dominanza, e frequenze genotipiche assolute e relative.

Genotipo		Frequenza	
presenza di dominanza	assenza di dominanza	assoluta	relativa
AA	$A'A'$	D	d
Aa	$A'A$	H	h
aa	AA	R	r
Totale		N	1

È evidente che come proprietà fondamentale delle frequenze alleliche si ha sempre che $p+q=1$. Poiché la somma delle frequenze alleliche ad un locus è uguale all'unità, ne deriva che $p=1-q$ e analogamente $q=1-p$.

Mentre le frequenze genotipiche ad un singolo locus sono utili per analizzare gli effetti di specifici processi evolutivi in una popolazione, i genetisti preferiscono impiegare le frequenze alleliche per descrivere il pool genico di una popolazione. In effetti, le frequenze alleliche presentano alcuni importanti vantaggi rispetto a quelle genotipiche: i) negli organismi sessuati, i genotipi si riducono agli alleli al momento della formazione dei gameti (meiosi); ii) conseguentemente, sono gli alleli, e non i genotipi, a venire trasmessi da una generazione all'altra attraverso il processo della fecondazione; iii) inoltre, in una popolazione esistono sempre meno alleli che non genotipi e ciò permette di caratterizzare il pool genico utilizzando meno parametri. Si consideri, ad esempio, un locus con tre alleli (a_1 , a_2 e a_3), le frequenze alleliche possibili sono tre [$f(a_1)$, $f(a_2)$ e $f(a_3)$], mentre quelle genotipiche sono sei [$f(a_1a_1)$, $f(a_1a_2)$, $f(a_1a_3)$, $f(a_2a_2)$, $f(a_2a_3)$ e $f(a_3a_3)$].

Per il calcolo delle frequenze alleliche p e q si possono seguire due diverse vie, una basata sul numero dei genotipi e l'altra sulla frequenza dei genotipi nella popolazione. Indipendentemente dal parametro impiegato, si parla comunque di metodo generale per il calcolo delle frequenze geniche (**Quadro 11.1**).

Quando il calcolo si basa sul numero dei genotipi, è necessario considerare che al locus esaminato in una popolazione di N individui il totale degli alleli è uguale a $2N$ poiché ognuno di questi individui deriva dall'unione di due gameti. Ciò significa che ciascuno degli individui con genotipo AA porta due alleli dello stesso tipo e che i relativi D individui della popolazione derivano pertanto da $2D$ gameti tutti con l'allele A . Analogamente, ciascuno degli individui con genotipo aa porta due alleli dello stesso tipo e quindi i relativi R individui della popolazione derivano da $2R$ gameti tutti con l'allele a . I genotipi eterozigoti (Aa), invece, includendo alleli di tipo diverso derivano da H gameti con l'allele A e H gameti con l'allele a . Il numero totale di alleli A coinvolti nella formazione degli N individui è pari a $2D+H$ mentre quello degli alleli a è pari a $2R+H$. Le frequenze degli alleli al locus considerato sono pertanto:

$$f(A)=p=(2D+H)/2N=(D+1/2H)/N$$

$$f(a)=q=(2R+H)/2N=(R+1/2H)/N$$

In altri termini, p è uguale al rapporto tra la somma del totale degli omozigoti AA e della metà degli eterozigoti Aa , e il numero complessivo di individui della popolazione, mentre q è dato dal rapporto tra la somma del totale degli omozigoti aa e della metà degli eterozigoti Aa , e il numero complessivo di individui della popolazione.

Quando il calcolo si basa sulle frequenze dei tre possibili genotipi al locus considerato, cioè $f(AA)=D/N=d$, $f(Aa)=H/N=h$ e $f(aa)=R/N=r$, le frequenze geniche sono le seguenti:

$$f(A)=p=d+1/2h$$

$$f(a)=q=r+1/2h$$

In altri termini, p è uguale alla somma della frequenza degli omozigoti AA e della metà della frequenza degli eterozigoti Aa , mentre q è data dalla somma della frequenza degli omozigoti aa e della metà della frequenza degli eterozigoti Aa .

Quadro 11.1 – Metodo generale per il calcolo delle frequenze geniche: esempio numerico

L'esempio di calcolo delle frequenze geniche prende in considerazione un caso classico di carattere per il quale è stata riscontrata assenza di dominanza: il colore del fiore in *Mirabilis jalapa* (bella di notte). In questo caso, gli eterozigoti sono riconoscibili rispetto ad entrambi gli omozigoti: le piante a fiore rosso hanno genotipo $R'R'$, quelle a fiore bianco hanno genotipo RR , mentre gli eterozigoti hanno fiore rosa e genotipo $R'R$. Ciò significa che l'incrocio tra piante omozigoti a fiore rosso e a fiore bianco produce una F_1 uniforme composta di sole piante a fiori rosa, mentre l'incrocio tra due di queste piante ibride produce una discendenza F_2 di piante con fiori nel rapporto 1 rosso : 2 rosa : 1 bianco.

Assumendo di avere una popolazione di 100 piante con frequenze assolute pari a 48 (piante a fiori rossi, genotipi $R'R'$), 44 (piante a fiori rosa, genotipi $R'R$) e 8 (piante a fiori bianchi, genotipi RR), le frequenze geniche risultano pari a $p(R')=0,7$ e $q(R)=0,3$, indipendentemente dal procedimento che viene usato per il loro calcolo. Si noti che una volta determinata la frequenza di un allele, la frequenza dell'altro allele si può ottenere per differenza all'unità (**Fig. 11.7**).

Fig. 11.7 – Metodo generale per il calcolo delle frequenze geniche: esempio numerico.

Carattere: colore dei fiori in *Mirabilis jalapa*

Popolazione = 100 piante = N

$R'R'$ → rossi 48 piante = D

$R'R$ → rosa 44 piante = H

RR → bianchi 8 piante = R



1) Numero dei genotipi:

$$p = \frac{D + H}{N} = \frac{48 + 44}{100} = 0,92$$

$$q = \frac{R + H}{N} = \frac{8 + 44}{100} = 0,52$$

2) Frequenze dei genotipi:

$$f(R'R') = \frac{D}{N} = d = 48/100 = 0,48$$

$$f(R'R) = \frac{H}{N} = h = 44/100 = 0,44$$

$$f(RR) = \frac{R}{N} = r = 8/100 = 0,08$$

$$p = d + h = 0,48 + 0,44 = 0,92$$

$$q = r + h = 0,08 + 0,44 = 0,52$$

$$p + q = 1 \rightarrow q = 1 - p = 1 - 0,92 = 0,08$$

11.6 Legge dell'equilibrio genetico di Hardy-Weinberg

Una volta determinate le frequenze geniche e le frequenze genotipiche di una particolare generazione di una ipotetica popolazione, è possibile prevedere quali saranno le frequenze geniche e le frequenze genotipiche nelle generazioni successive. Le popolazioni di specie allogame sono di norma caratterizzate da unioni sessuali casuali che consentono in particolari condizioni di mantenere inalterate nel tempo le frequenze geniche e le frequenze genotipiche. Tali popolazioni vengono comunemente considerate in equilibrio.

Nel 1908 Hardy e Weinberg osservarono in maniera del tutto indipendente che, in una popolazione ed in un particolare locus, le frequenze geniche si mantengono costanti attraverso le generazioni qualora vengano rispettate un insieme di condizioni: i) la popolazione sia molto grande, cioè composta da un numero elevato di individui; ii) non agisca la selezione naturale, cioè che nessuno dei tre possibili genotipi AA , Aa e aa sia avvantaggiato rispetto agli altri; iii) non vi sia mutazione differenziale di un allele in un altro, nel senso che il tasso di mutazione $A \rightarrow a$ sia uguale a quello $a \rightarrow A$; iv) non si verifichi migrazione di A o a da o verso popolazioni contigue; v) le unioni fra gli individui della popolazione siano del tutto casuali. La legge che unisce matematicamente le frequenze geniche con le frequenze genotipiche in una popolazione caratterizzata da incroci casuali porta il nome di entrambi: legge dell'equilibrio di Hardy-Weinberg. Nel 1903 il primo ad evidenziare la relazione tra frequenze geniche e genotipiche fu W.E. Castle che tuttavia non riuscì a formulare tale relazione in termini matematici (**Quadro 11.2**).

La legge di Hardy-Weinberg è una espressione matematica dell'equilibrio genetico delle popolazioni panmittiche: descrive cosa avviene alle frequenze geniche e genotipiche di una popolazione mentre i geni vengono trasmessi nelle generazioni successive in assenza di forze evolutive. Qualora il pool genico non subisca né perdite

A Popolazione di genotipi $Aa \rightarrow p(A)=q(a)=0,5$			B Popolazione di genotipi AA, Aa e $aa \rightarrow p(A)=0,7$ $q(a)=0,3$		
	A (0,5)	a (0,5)		A (0,7)	a (0,3)
A (0,5)	AA (0,25)	Aa (0,25)	A (0,7)	AA (0,49)	Aa (0,21)
a (0,5)	Aa (0,25)	aa (0,25)	a (0,3)	Aa (0,21)	aa (0,09)
	$f(AA) = 0,5 \times 0,5 = 0,25 \rightarrow p^2$			$f(AA) = 0,7 \times 0,7 = 0,49 \rightarrow p^2$	
	$f(Aa) = 2(0,5 \times 0,5) = 0,5 \rightarrow 2pq$			$f(Aa) = 2(0,7 \times 0,3) = 0,42 \rightarrow 2pq$	
	$f(aa) = 0,5 \times 0,5 = 0,25 \rightarrow q^2$			$f(aa) = 0,3 \times 0,3 = 0,09 \rightarrow q^2$	

Tab. 11.7 – Quadrato di Punnett sulle predizioni della legge di Hardy-Weinberg nel caso di popolazione composta di soli eterozigoti Aa (A) e di popolazione con frequenze geniche $p(A)=0,7$ e $q(a)=0,3$ (B).

né aggiunte di alleli e le unioni tra individui siano casuali, le popolazioni di specie allogame sono dette in equilibrio poiché la composizione genetica rimane costante nel corso delle generazioni. Nelle popolazioni che non sono in equilibrio questo si raggiunge in una sola generazione di unioni casuali, indipendentemente dalle frequenze geniche iniziali. In una popolazione ipotetica con frequenze geniche pari a p per A e q per a , attraverso la meiosi ogni individuo trasmette ai gameti gli alleli che possiede con uguale frequenza. Pertanto anche nei gameti le frequenze degli alleli A e a sono p e q . Poiché i gameti sono i componenti del pool genico, la formazione casuale di uno zigote con un determinato genotipo deriva dalla combinazione di due gameti e la sua frequenza è pari al prodotto delle frequenze relative dei gameti o analogamente delle frequenze alleliche. Ammettendo che i gameti maschili e femminili prodotti dalla popolazione siano nelle proporzioni di p per A e q per a , in presenza di unioni perfettamente casuali, le combinazioni genotipiche e le relative frequenze possono essere derivate applicando il quadrato di Punnett. Qualunque siano le frequenze geniche iniziali, dopo una generazione di unioni casuali, le frequenze genotipiche diventano p^2 per AA , $2pq$ per Aa e q^2 per aa . La popolazione è detta in equilibrio poiché tali frequenze genotipiche rimangono costanti finché p e q rimangono invariati.

Assumendo costanti le frequenze geniche e supponendo casuali le unioni è quindi possibile prevedere il destino dei geni e dei genotipi nel corso delle generazioni. Nella **Tab. 11.7** è presa in considerazione una situazione limite corrispondente a una popolazione composta di soli genotipi eterozigoti (Aa) tale che $p(A)=q(a)=0,5$. Dato che alle unioni casuali tra genotipi corrispondono unioni casuali tra gameti, l'equilibrio si raggiunge ai valori $p^2(AA)=0,25$, $2pq(Aa)=0,50$ e $q^2(aa)=0,25$ che rispettano le proporzioni attese nella discendenza F_2 di un monobrido. Nelle allogame, la fecondazione incrociata casuale fa sì che queste proporzioni vengano mantenute, mentre nelle autogame, l'autofecondazione determinerà la fissazione degli omozigoti e la scomparsa degli eterozigoti. Allo stesso modo, considerando un'altra situazione, ad esempio una popolazione con frequenze geniche $p(A)=0,7$ e $q(a)=0,3$, è possibile prevedere che le frequenze genotipiche raggiungono l'equilibrio ai valori $p^2(AA)=0,49$, $2pq(Aa)=0,42$ e $q^2(aa)=0,09$ ed ipotizzare che in presenza di unioni casuali e in assenza di fattori di disturbo queste proporzioni rimangono costanti ad ogni generazione. La **Fig. 11.8** mette a confronto l'equazione di Hardy-Weinberg con il quadrato di Punnett. L'equazione dell'equilibrio deriva dal modo con cui gli alleli portati dai gameti si combinano casualmente l'uno con l'altro per produrre gli zigoti della discendenza. Si noti che la somma delle frequenze geniche è pari ad uno ($p+q=1$) così come la somma delle frequenze genotipiche ($p^2+2pq+q^2=1$).

La legge di Hardy-Weinberg può essere enunciata nel modo seguente: data la costanza delle frequenze geniche, in una popolazione numerosa che abbia frequenze iniziali per la coppia allelica al locus considerato pari $p(A)$ e $q(a)$, le frequenze genotipiche raggiungono, in una sola generazione di unioni casuali, un equilibrio ai valori seguenti: p^2 per AA , $2pq$ per Aa e q^2 per aa . Quando i loci considerati sono più di uno, l'equilibrio viene raggiunto in alcune generazioni.

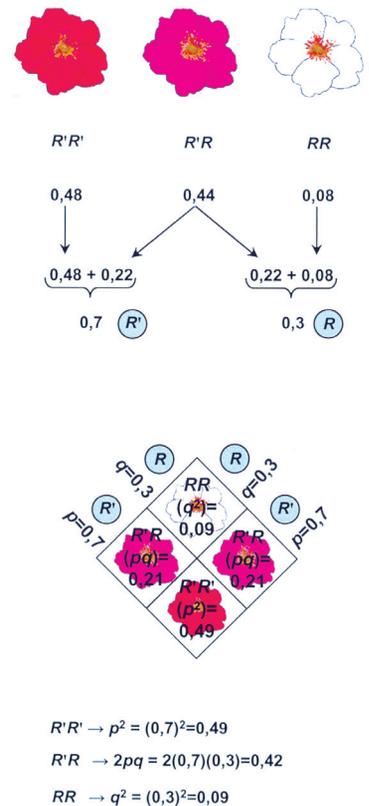


Fig. 11.8 – Confronto tra l'equazione di Hardy-Weinberg e il quadrato di Punnett.

Legge di Hardy-Weinberg

Un insieme di assunzioni:

- (1) Costanza delle frequenze geniche (popolazione numerosa, assenza di mutazione differenziale, migrazione e selezione) e presenza di unioni casuali.

Due predizioni principali:

- (2) Le frequenze alleliche p e q per gli alleli A e a non variano con il tempo;
- (3) Le frequenze genotipiche si stabiliscono sulle proporzioni p^2 per i genotipi AA , $2pq$ per i genotipi Aa e q^2 per i genotipi aa .

Quadrato di Punnett

	$A (p)$	$a (q)$
$A (p)$	$AA (p^2)$	$Aa (pq)$
$a (q)$	$Aa (pq)$	$aa (q^2)$

La somma delle frequenze genotipiche $p^2+2pq+q^2=1$

Fig. 11.9 – Legge dell'equilibrio di Hardy-Weinberg.

Tale legge può essere suddivisa in tre parti principali, nel senso che prevede un insieme di assunzioni, cioè (1) la popolazione deve essere sufficientemente numerosa e non presentare mutazione differenziale, migrazione e selezione naturale, e permette due predizioni principali: (2) le frequenze alleliche p e q per A e a della popolazione non variano con il tempo, rimanendo costanti nel corso delle generazioni; (3) le frequenze genotipiche della popolazione si stabiliscono sulle proporzioni p^2 per i genotipi AA , $2pq$ per i genotipi Aa e q^2 per i genotipi aa (**Fig. 11.9**).

In sostanza, la struttura genetica di una popolazione di specie allogama può venire rappresentata negli stadi zigotici come frequenze genotipiche [$p^2(AA)+2pq(Aa)+q^2(aa)$] e in quelli gametici come frequenze alleliche [$p(a)$ e $q(a)$]. Ciascuna generazione di individui di origine zigotica produce per meiosi gameti con alleli A e a nelle proporzioni p e q , i gameti maschili e femminili si uniscono con la fecondazione per formare zigoti con genotipo AA , Aa e aa nelle proporzioni p^2 , $2pq$ e q^2 . Tale ciclo si ripete per un numero indefinito di generazioni fino a quando permangono le assunzioni della legge di Hardy-Weinberg e ciò equivale a dire che la composizione genetica della popolazione rimane costante di generazione in generazione rispetto al locus considerato (**Fig. 11.10**).

È possibile dimostrare algebricamente la validità della legge di Hardy-Weinberg. In una popolazione ipotetica con frequenze geniche pari a p per A e q per a , e con genotipi AA , Aa e aa nelle frequenze, rispettivamente, di p^2 , $2pq$ e q^2 , i possibili tipi di unioni tra individui sono quelli elencati in **Tab. 11.8**. L'unione casuale tra individui implica che la frequenza di una particolare unione tra due genotipi sia uguale al prodotto delle rispettive frequenze genotipiche. Ad esempio, la frequenza di un incrocio

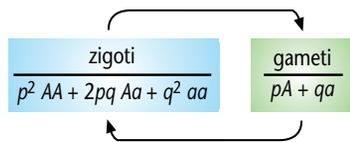


Fig. 11.10 – Rappresentazione schematica della legge dell'equilibrio di Hardy-Weinberg applicata alla successione di stadi zigotici e gametici in una popolazione mendeliana.

Tab. 11.8 – Dimostrazione algebrica della legge di Hardy-Weinberg.

Tipi di unioni	Frequenza delle unioni	Frequenze genotipiche nelle discendenze		
		AA	Aa	aa
$p^2 AA \times p^2 AA$	p^4	p^4	—	—
$p^2 AA \times 2pq Aa$ $2pq Aa \times p^2 AA$	$4p^3q$	$2p^3q$	$2p^3q$	—
$p^2 AA \times q^2 aa$ $q^2 aa \times p^2 AA$	$2p^2q^2$	—	$2p^2q^2$	—
$2pq Aa \times 2pq Aa$	$4p^2q^2$	p^2q^2	$2p^2q^2$	p^2q^2
$2pq Aa \times q^2 aa$ $q^2 aa \times 2pq Aa$	$4pq^3$	—	$2pq^3$	$2pq^3$
$q^2 aa \times q^2 aa$	q^4	—	—	q^4
Totali	$(p^2+2pq+q^2)^2=1$	$p^2(p^2+2pq+q^2)=p^2$	$2pq(p^2+2pq+q^2)=2pq$	$q^2(p^2+2pq+q^2)=q^2$

$AA \times AA$ sarà pari a p^4 , così come quella di un incrocio $aa \times aa$ sarà pari a q^4 , mentre l'incrocio tra due eterozigoti sarà $4p^2q^2$, e così via. In questo modo possono calcolarsi le frequenze di tutte le unioni possibili (Tab. 11.8). Stabilite le unioni possibili e determinate le relative frequenze, in base ai rapporti di segregazione mendeliana attesi per tutte le combinazioni si può calcolare la frequenza di ogni possibile genotipo nella discendenza. Ad esempio, l'incrocio $AA \times Aa$ e l'incrocio reciproco $Aa \times AA$, con frequenza complessiva $4p^3q$, forniscono discendenze con $1/2$ di genotipi AA e $1/2$ di genotipi Aa , e pertanto la probabilità di ottenere ciascuno di questi genotipi è pari a $2p^3q$. Analogamente, sapendo che un incrocio tra eterozigoti $Aa \times Aa$ fornisce una discendenza con genotipi AA , Aa e aa nel rapporto di 1:2:1, la probabilità di ottenere questi genotipi è di p^2q^2 per ciascuno degli omozigoti e $2p^2q^2$ per l'eterozigote. Le frequenze genotipiche nelle discendenze di tutte le possibili unioni sono indicate nella Tab. 11.8. Sommando e risolvendo la frequenza totale di ogni genotipo, si ha che dopo una generazione di unioni casuali le frequenze genotipiche sono ancora p^2 per AA , $2pq$ per Aa e q^2 per aa . In una situazione di questo tipo è evidente che anche le frequenze alleliche p di A e q di a rimangono costanti.

In conclusione, la legge di Hardy-Weinberg stabilisce che se la popolazione rimane numerosa, ad unioni casuali e non subisce l'effetto delle forze evolutive (mutazione, migrazione differenziale e selezione), le frequenze alleliche e le frequenze genotipiche rimangono costanti generazione dopo generazione al locus considerato. Quando i loci considerati sono due o più, l'equilibrio viene raggiunto in alcune generazioni.

Quadro 11.2 – Hardy, Weinberg e Castle

Qualche anno dopo la riscoperta delle leggi di Mendel, un matematico inglese e un medico tedesco osservarono in maniera del tutto indipendente che in una popolazione le frequenze geniche ad un particolare locus si mantengono costanti nel corso delle generazioni e che le frequenze genotipiche possono essere previste a partire da quelle geniche qualora vengano però rispettate un insieme di condizioni. Nel 1908 G.H. Hardy pubblicò un breve lavoro, *Mendelian proportions in a mixed population*, che descriveva la relazione tra i genotipi ed i fenotipi di una popolazione e poche

settimane più tardi W. Weinberg stabilì in modo analogo la stessa relazione (*Über den nachweis der vererbung beim menschen*, è il titolo del suo lavoro ben più consistente). La legge che unisce matematicamente le frequenze geniche con le frequenze genotipiche in una popolazione caratterizzata da incroci casuali porta il nome di entrambi: legge dell'equilibrio di Hardy-Weinberg. Tuttavia, qualche anno prima, nel 1903, il primo ad evidenziare la relazione tra frequenze geniche e genotipiche fu lo statunitense William E. Castle (1867-1962), dell'Università di Harvard, che tuttavia non riuscì a formulare tale relazione in termini matematici.

11.6.1 Accertamento dell'equilibrio Hardy-Weinberg

Al fine di verificare l'equilibrio Hardy-Weinberg in una popolazione di piante è necessario per prima cosa calcolare le frequenze geniche. Dopo di che si prosegue calcolando le frequenze genotipiche attese e confrontando queste con le frequenze genotipiche osservate. Infine, può essere richiesto l'uso di un semplice test statistico per stabilire se le deviazioni tra frequenze genotipiche attese ed osservate possono essere attribuite al caso oppure sono dovute all'assenza di equilibrio nella popolazione.

A titolo di esempio, può essere preso in considerazione il locus che controlla il colore del fiore in *Mirabilis jalapa*, carattere per il quale si ha assenza di dominanza: le piante $R'R'$ sono a fiore rosso, quelle RR sono a fiore bianco e gli eterozigoti $R'R$ sono a fiore rosa. In una popolazione composta di 100 individui (N) si ipotizzi che siano state contate 70 piante a fiore rosso, 20 piante a fiore rosa e 10 piante a fiore bianco. Le frequenze genotipiche assolute sono pertanto pari a: $D=70$, $H=20$ e $R=10$. Le frequenze alleliche possono essere calcolate a partire dalle frequenze genotipiche assolute nel modo seguente: $p(R')=(D+1/2H)/N=(70+10)/100=0,80$ e $q(R)=(R+1/2H)/N=(10+10)/100=0,20$. Le frequenze alleliche possono anche essere calcolate a partire dalle frequenze genotipiche relative. Nella popolazione considerata tali frequenze

Popolazione:		
RR' (rossi)	→ 70 piante (D)	
$R'R$ (rosa)	→ 20 piante (H)	
RR (bianchi)	→ 10 piante (R)	
100 piante (N)		
Frequenze alleliche:		
$p(R) = \frac{D + 1/2H}{N} = \frac{70 + 10}{100} = 0,8$		
$q(R) = 1 - p(R) = 1 - 0,8 = 0,2$		
Frequenze genotipiche:		
osservate	attese	
$f(R'R) = D/N = 70/100 = 0,7$	$p^2(R'R) = (0,8)^2 = 0,64$	
$f(R'R) = H/N = 20/100 = 0,2$	$2pq(R'R) = 2(0,8 \times 0,2) = 0,32$	
$f(RR) = R/N = 10/100 = 0,1$	$q^2(RR) = (0,2)^2 = 0,04$	
Numero di genotipi:		
osservati	attesi	Chi-quadrato:
RR' 70	$p^2 \times N = 64$	
$R'R$ 20	$2pq \times N = 32$	
RR 10	$q^2 \times N = 4$	
		$\chi^2 = \sum_1^n \frac{(X_{oss} - X_{att})^2}{X_{att}}$
		$\chi^2_{calc} \gg \chi^2_{tab} (1 \text{ g.l., } P=1\%)$

Fig. 11.11 – Esempio di calcolo per verificare l'equilibrio Hardy-Weinberg.

sono pari a: $f(R'R')=d=D/N=70/100=0,70=70\%$, $f(R'R)=H/N=h=20/100=0,20=20\%$ e $f(RR)=R/N=r=10/100=0,10=10\%$. Secondo questa procedura, $p(R)=(d+1/2h)=0,70+0,10=0,8$ e $q(R)=(r+1/2h)=0,10+0,10=0,20$. Poiché $p+q=1$, ne consegue che una volta noto $p(R)$, $q(R)$ può essere ottenuto anche per differenza: $1-0,8=0,2$. A questo punto possono essere derivate le frequenze genotipiche relative attese in caso di equilibrio Hardy-Weinberg: $f(R'R')=p^2=(0,8)^2=0,64$, $f(R'R)=2pq=2(0,8 \times 0,2)=0,32$ e $f(RR)=q^2=(0,2)^2=0,04$. Tuttavia, il confronto tra dati osservati e attesi non può essere effettuato usando le frequenze relative, ma servono le frequenze genotipiche assolute. Le frequenze genotipiche osservate sono ovviamente note, $D=70$, $H=20$ e $R=10$, mentre quelle attese vanno calcolate moltiplicando le frequenze relative per il numero totale di individui della popolazione: $p^2 \times N=0,64 \times 100=64$, $2pq \times N=0,32 \times 100=32$ e $q^2 \times N=0,04 \times 100=4$. Ciò significa che in una popolazione in equilibrio 64 piante dovrebbero essere a fiore rosso, 32 piante a fiore rosa e 4 piante a fiore bianco.

Per controllare se le frequenze differiscono in modo significativo da quelle previste dall'equilibrio Hardy-Weinberg, si effettua il calcolo del chi-quadrato, che è il saggio statistico usato per stabilire se le

deviazioni delle frequenze genotipiche osservate da quelle attese sono imputabili al caso. In questo esempio, il valore stimato di χ^2 è superiore rispetto a quello teorico e quindi si deve supporre che la popolazione non sia in equilibrio Hardy-Weinberg (**Fig. 11.11**). Qualora le unioni tra le piante avvengano a caso e non intervengano fattori di disturbo delle frequenze geniche, la popolazione raggiungerà l'equilibrio al locus considerato in una sola generazione e le frequenze genotipiche diventeranno circa: 64% ($R'R'$), 32% ($R'R$) e 4% (RR).

L'accertamento dell'equilibrio Hardy-Weinberg di una qualsiasi popolazione di piante rispetto al locus considerato è un aspetto molto importante poiché consente, in presenza di equilibrio, di presumere che la sua composizione genetica rimarrà costante nel corso delle generazioni e, in assenza di equilibrio, di prevedere quale sarà la sua composizione genetica dopo una sola generazione con unioni a caso, indipendentemente dalle frequenze alleliche iniziali.

11.6.2 Relazione tra frequenze geniche e genotipiche, e rappresentazione dell'equilibrio genetico

La relazione di base tra frequenze geniche e frequenze genotipiche nelle popolazioni all'equilibrio è data dalla seguente espressione:

$$(p+q)^2=p^2+2pq+q^2$$

Tale espressione dimostra che le frequenze genotipiche all'equilibrio derivano dall'espansione del binomio delle frequenze geniche.

Le frequenze genotipiche possono quindi essere previste a partire da quelle geniche, dimostrando così l'esistenza di una relazione univoca tra frequenze geniche e genotipiche. Poiché la somma delle frequenze geniche $p+q=1$, si ha che $p=1-q$ e $q=1-p$. Note le frequenze geniche, le frequenze genotipiche possono essere ricavate come p^2 e q^2 per i due omozigoti e $2pq$ per l'eterozigote. Analogamente, la somma delle frequenze genotipiche $p^2+2pq+q^2=1$. In definitiva, in una popolazione in equilibrio ad una determinata frequenza allelica corrisponde una sola frequenza di ciascuno dei genotipi possibili ad un dato locus. In **Tab. 11.9** sono riportati alcuni esempi numerici della relazione esistente tra frequenze geniche e genotipiche.

Tab. 11.9 – Relazione univoca tra frequenze alleliche e genotipiche.

Frequenze alleliche		Frequenze genotipiche		
$p(A)$	$q(a)=1-p(A)$	$f(AA)=p^2$	$f(Aa)=2pq$	$f(aa)=q^2$
0,05	0,95	0,0025	0,095	0,9025
0,1	0,9	0,01	0,18	0,81
0,2	0,8	0,04	0,32	0,64
0,3	0,7	0,09	0,42	0,49
0,5	0,5	0,25	0,50	0,25

Tale relazione può anche essere rappresentata ricorrendo ad un grafico cartesiano che preveda le frequenze dei genotipi $p^2(AA)$, $2pq(Aa)$ e $q^2(aa)$ nell'asse delle ordinate e le frequenze degli alleli $p(A)$ e $q(a)$, rispettivamente, nell'asse superiore e inferiore delle ascisse con orientamento opposto (Fig. 11.12). Le caratteristiche più importanti che possono essere messe in evidenza con questo grafico riferito ad un singolo locus di popolazioni in equilibrio Hardy–Weinberg sono le seguenti: i) la frequenza massima dell'eterozigote è 0,5 e si riscontra quando $p(A)=q(a)=0,5$; ii) se le frequenze alleliche sono comprese tra 0,33 e 0,66 l'eterozigote rappresenta il genotipo più numeroso; iii) quando la frequenza di un allele è bassa, il più raro dei genotipi è l'omozigote per quell'allele e l'allele è presente nella popolazione prevalentemente in condizione eterozigote. Ad esempio, se $p(A)=0,8$ e $q(a)=0,2$, la frequenza del genotipo omozigote per A è del 64% (p^2) e quella del genotipo omozigote per a è del 4% (q^2), mentre l'eterozigote Aa ha un frequenza pari al 32% ($2pq$); iv) ogni popolazione in equilibrio può essere definita da una linea verticale (Fig. 11.12).

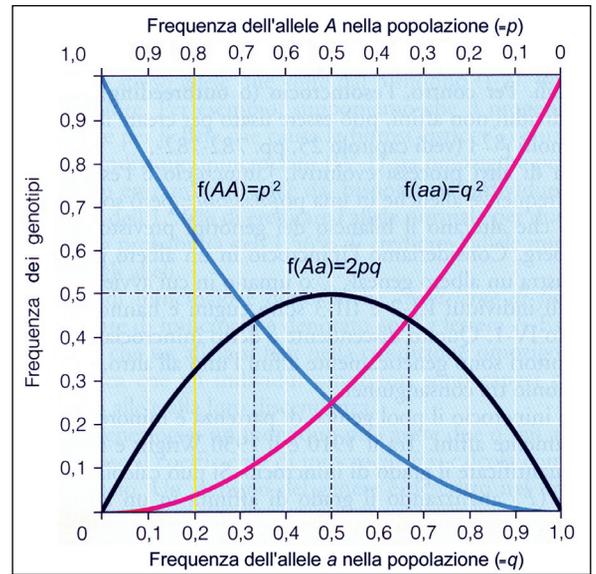


Fig. 11.12 – Rappresentazione grafica della relazione esistente tra frequenze geniche e genotipiche.

Una importante applicazione della legge di Hardy-Weinberg è rappresentata dal calcolo delle frequenze alleliche nel caso di caratteri che presentano dominanza completa. In queste situazioni, i fenotipi dominanti aventi una situazione di eterozigosi al locus considerato non possono essere distinti da quelli con genotipo omozigote per l'allele dominante. Tuttavia, assumendo che la popolazione in esame sia in equilibrio, la frequenza genotipica degli individui omozigoti recessivi può essere utilizzata per il calcolo delle frequenze alleliche. Tale frequenza consente infatti di ricavare agevolmente la frequenza dell'allele recessivo estraendo la radice quadrata della frequenza relativa dei fenotipi recessivi. Poiché $f(aa)=q^2$, si ha che $\sqrt{f(aa)}=q(a)$ e quindi $p(A)$ è uguale a $1-q(a)$. Una volta che sono state derivate le frequenze alleliche, la frequenza genotipica degli omozigoti dominanti $f(AA)$ non può che essere pari a p^2 e quella degli eterozigoti $f(Aa)$ pari a $2pq$ (Fig. 11.13). È necessario sottolineare che questo metodo per il calcolo delle frequenze alleliche e per la derivazione delle frequenze genotipiche si basa sulle assunzioni della legge di Hardy-Weinberg. Qualora le condizioni alla base dell'equilibrio genetico di una popolazione non vengano soddisfatte, le frequenze alleliche stimate secondo questo metodo non saranno accurate, così come le frequenze genotipiche non potranno essere considerate affidabili. Nel caso di caratteri che non presentano dominanza, i fenotipi possibili ad un determinato locus risultano invece tutti evidenti e quindi le frequenze genotipiche coincideranno con quelle fenotipiche: $f(A'A')=p^2$, $f(A'A)=2pq$ e $f(AA)=q^2$, mentre le frequenze alleliche potranno essere calcolate direttamente come $\sqrt{f(AA)}=p$ e $\sqrt{f(aa)}=q$.

Assumendo che una data popolazione sia in equilibrio Hardy-Weinberg, la frequenza genotipica degli omozigoti recessivi (o minus) può essere utilizzata per il calcolo delle frequenze alleliche:

$$f(aa) = q^2 \rightarrow \sqrt{f(aa)} = q$$

$$p + q = 1 \rightarrow p = 1 - q$$

$f(AA) = p^2$ frequenza genotipica degli omozigoti dominanti (o plus)

$f(Aa) = 2pq$ frequenza genotipica degli eterozigoti

Fig. 11.13 – Uso della legge dell'equilibrio di Hardy-Weinberg.

La struttura genetica di una popolazione ed, in particolare, la sua composizione genotipica rispetto all'equilibrio previsto dalla legge di Hardy-Weinberg può essere

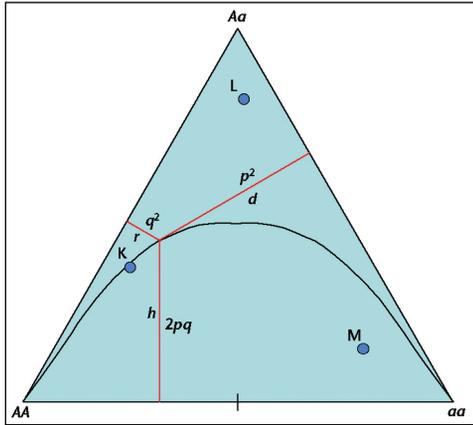


Fig. 11.14 – Rappresentazione grafica di popolazioni mendeliane dialleliche in coordinate triangolari secondo il diagramma di De Finetti.

descritta utilizzando un grafico basato sulle coordinate triangolari. Tali coordinate sono costituite da un sistema formato da un triangolo equilatero avente l'altezza uguale ad uno. Dato che la somma delle frequenze relative dei tre genotipi possibili ad un locus con due alleli è uguale ad uno, si può individuare un punto all'interno del triangolo in modo che le sue distanze dai lati siano uguali alle frequenze relative di ciascun genotipo (**Fig. 11.14**). La somma di queste tre distanze è uguale all'altezza del triangolo che è posta uguale ad uno. Il punto in cui convergono i tre assi perpendicolari ai lati del triangolo equilatero, che tra loro formano angoli di 120° , rappresenta la posizione della popolazione. Il metodo delle coordinate triangolari, applicato per la prima volta in genetica di popolazioni dall'italiano Bruno De Finetti, consente di rappresentare qualsiasi popolazione in equilibrio Hardy-Weinberg. Qualora una popolazione risulti costituita da un solo genotipo, il punto con cui si identifica la popolazione stessa sarà posizionato in uno dei vertici del triangolo. Ad esempio, una popolazione costituita unicamente da eterozigoti Aa sarà rappresentata dal punto che occupa il vertice del triangolo, mentre una popolazione formata da due soli genotipi sarà rappresentata da un punto che giace su uno dei lati. Il punto rappresentativo della popolazione si trova all'interno del triangolo quando tutti e tre i genotipi sono presenti. Le frequenze geniche si possono ricavare dalla proiezione del punto che rappresenta la popolazione sulla base del triangolo. Tale proiezione divide infatti il lato delimitato dai vertici AA e aa in due parti proporzionali alle frequenze geniche $p(A)$ e $q(a)$.

Nella Fig. 11.14 è riportata una popolazione in equilibrio con frequenze genotipiche pari a $p^2(A)=0,5$, $2pq(Aa)=0,4$ e $q^2(aa)=0,1$. Il punto rappresentativo di ogni popolazione in equilibrio Hardy-Weinberg si trova sull'arco di parabola che, passando per i vertici di base AA e aa , interseca l'altezza nel punto medio in corrispondenza del quale $2pq(Aa)=0,5$.

Tale rappresentazione è molto utile per visualizzare la posizione relativa di qualsiasi popolazione rispetto all'equilibrio, rendendo possibile anche il confronto tra diverse popolazioni. Nella stessa figura sono riportate tre diverse situazioni per altrettante popolazioni: la popolazione K si trova vicino alla parabola e può essere pertanto considerata in equilibrio, mentre le popolazioni L e M si trovano lontane dalla parabola, in prossimità dei vertici del triangolo. Nel primo caso, ad esempio, c'è un eccesso di eterozigoti poiché il punto rappresentativo si trova vicino al vertice Aa .

11.6.3 Estensioni della legge di Hardy-Weinberg: loci con allelismo multiplo e loci di cromosomi sessuali

La legge di Hardy-Weinberg è valida anche per loci con più di due alleli e può essere estesa a qualsiasi numero di alleli a condizione però che questi vengano campionati a due a due così come accade negli individui diploidi. Un esempio classico è rappresentato dalle serie alleliche ($S_1, S_2, S_3, \dots, S_n$) che causano l'incompatibilità nelle piante (trifoglio, radichio, ecc.). L'incompatibilità è una barriera riproduttiva connessa al fatto che il polline altrimenti normale è incapace di funzionare in maniera appropriata su certi pistilli a causa di ostacoli fisiologici e morfologici che si oppongono alla fecondazione impedendo la formazione del seme anche in presenza di ovuli funzionali (\rightarrow Cap. 10). Un altro esempio di allelismo multiplo è quello riguardante i gruppi sanguigni nell'uomo (**Quadro 11.3**).

Nel caso di loci con allelismo multiplo, la frequenza genica di ogni allele può essere calcolata come segue:

$$f(A_i) = \frac{\sum A_i}{\sum (A_i + A_2 + A_3 + A_n)}$$

Le frequenze genotipiche sono date dall'espressione multinomiale:

$$(p+q+r+\dots n)^2$$

dove ogni lettera rappresenta la frequenza di un dato allele. Così, ipotizzando un locus con tre alleli, ad esempio A_1 , A_2 e A_3 , le frequenze dei sei possibili genotipi all'equilibrio saranno date dal quadrato del trinomio costituito dalla somma delle frequenze alleliche:

$$(p+q+r)^2 = p^2(A_1A_1) + 2pq(A_1A_2) + q^2(A_2A_2) + 2pr(A_1A_3) + 2qr(A_2A_3) + r^2(A_3A_3)$$

Qualora le frequenze geniche siano, ad esempio, $p(A_1)=0,5$, $q(A_2)=0,3$ e $r(A_3)=0,2$, assumendo una condizione di equilibrio per la popolazione in esame, le frequenze genotipiche attese sono pari a: $p^2(A_1A_1)=0,25$, $2pq(A_1A_2)=0,30$, $q^2(A_2A_2)=0,09$; $2pr(A_1A_3)=0,20$, $2qr(A_2A_3)=0,12$ e $r^2(A_3A_3)=0,04$.

Qualora non siano, invece, note le frequenze geniche, queste possono essere calcolate dalle frequenze genotipiche assumendo sempre una condizione di equilibrio. Ad esempio, tenendo presente che la frequenza combinata dei genotipi A_2A_2 , A_2A_3 e A_3A_3 è uguale a $q^2+2qr+r^2=(q+r)^2$, la radice quadrata di questo ultimo binomio fornisce una stima di $(q+r)$, per cui la frequenza genica di A_1 è pari a $p=1-\sqrt{(q+r)^2}$ dato che $p(A_1)+q(A_2)+r(A_3)=1$. In modo analogo, la frequenza combinata dei genotipi A_1A_1 , A_1A_3 e A_3A_3 è uguale a $p^2+2pr+r^2=(p+r)^2$ e quindi frequenza genica di A_2 è pari a $q=1-\sqrt{(p+r)^2}$. La frequenza genica di A_3 può essere stimata come $r=\sqrt{r^2}$ oppure usando la relazione $1-p-q=r$.

La legge di Hardy-Weinberg è valida anche per loci di cromosomi sessuali e può essere quindi estesa ai caratteri controllati da alleli ubicati su questi cromosomi. Nella maggior parte delle specie vegetali uno stesso individuo produce sia gameti maschili che femminili, mentre la separazione dei sessi (dioicismo) è un fenomeno piuttosto raro. Nelle piante la presenza di cromosomi eteromorfi è stata accertata nei generi *Rumex* e *Viscum*: in tutti i casi i cromosomi sessuali eteromorfi sono comunque quelli maschili e il sesso femminile prevede dunque cromosomi omomorfi, analogamente a quanto accade per la quasi totalità dei mammiferi (sistema XY-XX).

Nel caso di geni ubicati sul cromosoma X, gli individui del sesso eterogametico possiedono per tali geni un solo locus e di conseguenza le frequenze genotipiche (e fenotipiche) della popolazione coincidono con le frequenze geniche: $p(X^AY)$ e $q(X^aY)$. Nel sesso omogametico, le frequenze alleliche di loci ubicati sul cromosoma sessuale sono le stesse che per i loci di cromosomi autosomici: $p^2(X^AX^A)$, $2pq(X^AX^a)$ e $q^2(X^aX^a)$. Per questo motivo, i caratteri recessivi associati al cromosoma X sono più frequenti tra i maschi che non tra le femmine. In generale, per i caratteri legati al sesso, la frequenza delle femmine omozigoti recessive è pari al quadrato della frequenza fenotipica dei maschi che manifestano il carattere:

$$f(X^aX^a) = f(X^aY)^2 = \sqrt{q^2(X^aX^a)} = q(X^aY)$$

Quando gli alleli sono ubicati sul cromosoma X e i sessi hanno frequenze alleliche diverse, l'equilibrio è raggiunto solo dopo alcune generazioni in quanto gli individui maschili ricevono il cromosoma X da un solo parentale, mentre quelli femminili ricevono un cromosoma X da entrambi i parentali. Quindi, negli individui maschili della discendenza la frequenza di un allele ad un locus del cromosoma X avrà la stessa frequenza di quell'allele nei parentali femminili, mentre la frequenza negli individui femminili sarà intermedia tra quelle dei parentali maschile e femminile. In una popolazione caratterizzata da unioni casuali e in assenza di fattori di disturbo, le

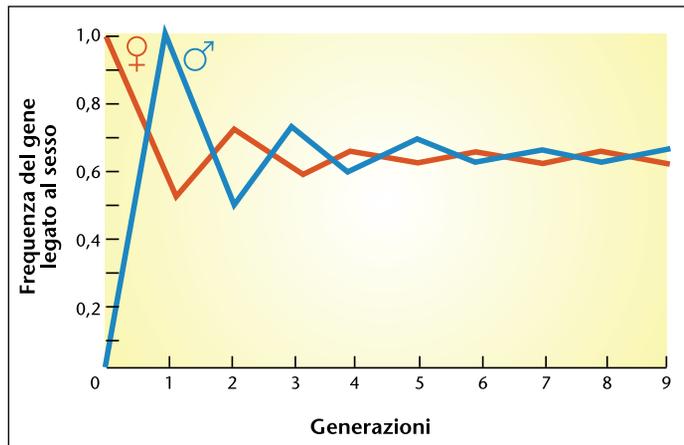


Fig. 11.15 – Rappresentazione dell'avvicinamento graduale all'equilibrio di un gene associato al cromosoma X con una frequenza iniziale di 1 nelle femmine e 0 nei maschi.

luzione. La mutazione ha frequenze talmente basse che non è ritenuta determinante nello spostare le frequenze geniche, come pure la migrazione da o verso popolazioni contigue è un fenomeno generalmente piuttosto limitato. Il vantaggio di un genotipo rispetto agli altri così come la dimensione della popolazione e il tipo di unioni che si verificano al suo interno sono invece molto importanti.

La legge dell'equilibrio Hardy-Weinberg è stata formulata sulla base dei principi della genetica mendeliana e di numerose assunzioni teoriche. Pertanto è lecito chiedersi se le condizioni teoriche di equilibrio possano essere riscontrate nelle popolazioni reali. Tale legge è infatti soggetta a numerose restrizioni. Innanzitutto, la popolazione deve essere infinitamente grande, uniformemente distribuita e non deve essere interessata da fenomeni di mutazione e migrazione differenziale. Inoltre, non deve esservi selezione naturale, nel senso che i vari genotipi devono contribuire alla generazione successiva originando lo stesso numero di discendenti. Infine, non sono ammesse restrizioni alla perfetta casualità delle unioni, cioè non devono esserci unioni preferenziali tra genotipi e nemmeno unioni tra individui imparentati. Si tratta di un lungo elenco di assunzioni che nella realtà possono non verificarsi, dato che le popolazioni di individui sono finite, i genotipi in un determinato ambiente hanno vitalità, capacità adattativa e fertilità differenziati, le mutazioni sono rare ma possibili e continue, così come è possibile anche la migrazione tra popolazioni, che nelle piante può realizzarsi mediante polline o seme. L'equilibrio di Hardy-Weinberg è pertanto difficilmente riscontrabile in natura poiché corrisponde ad una situazione non soggetta ad alcuna forza evolutiva capace di agire sulle frequenze geniche e genotipiche. L'equilibrio di Hardy-Weinberg rappresenta comunque uno strumento molto potente ed efficace per studiare le deviazioni rispetto alla situazione attesa secondo il modello, che può quindi essere usato come "ipotesi zero" contro la quale confrontare la struttura genetica di qualsiasi popolazione e per mezzo della quale individuare l'assunzione che è stata violata.

Per i caratteri qualitativi, controllati da singoli geni, la situazione di equilibrio è comunque sperimentalmente verificabile. Qualora una popolazione risulti in equilibrio ad un determinato locus, si può supporre che le assunzioni della legge siano soddisfatte e che la composizione genetica possa mantenersi costante durante le generazioni. In caso contrario, qualora nella popolazione siano intervenuti fattori di disturbo che hanno portato ad una rottura dell'equilibrio, si può ipotizzare che la popolazione raggiunga un nuovo equilibrio in una sola generazione di unioni casuali.

Di conseguenza, la legge di Hardy-Weinberg consente, da un lato, di accertare se una popolazione rimarrà stabile nelle generazioni successive, per quanto riguarda la sua composizione genetica e, dall'altro, permette di calcolare in modo molto rapido le frequenze geniche nelle popolazioni che hanno già raggiunto l'equilibrio.

frequenze alleliche nei due sessi oscillano, aumentando e diminuendo ad ogni generazione, e la differenza tra le frequenze alleliche nei due sessi subisce un dimezzamento ad ogni generazione (Fig. 11.15). Quando le frequenze alleliche negli individui maschili e femminili si equivalgono, le frequenze genotipiche raggiungeranno le proporzioni volute dall'equilibrio Hardy-Weinberg dopo un'ulteriore generazione di unioni casuali.

11.6.4 Fattori che disturbano l'equilibrio

Fisher, Haldane e Wright si resero conto che la violazione di una o più delle assunzioni dell'equilibrio di Hardy-Weinberg può costituire la causa di variazione delle frequenze geniche, cioè del primo gradino dell'evol-

Il fatto che l'equilibrio possa essere raggiunto solo dopo alcune generazioni quando i loci considerati sono più di uno ha conseguenze importanti per i caratteri quantitativi controllati da una pluralità di geni. In questi casi il raggiungimento dell'equilibrio genetico può richiedere anche un numero molto elevato di generazioni. Ad esempio, considerando i caratteri adattativi di un ecotipo, la rottura dell'equilibrio determinata da flusso genico implica molte generazioni di unioni casuali affinché venga ripristinato un nuovo equilibrio e cioè un nuovo stato di coadattamento.

La presenza di fattori di disturbo dell'equilibrio quali mutazione differenziale, migrazione e selezione determina variazioni delle frequenze geniche e quindi rottura dell'equilibrio Hardy-Weinberg. Tali fattori, come del resto le unioni non casuali e le variazioni casuali delle frequenze geniche (deriva), alterano la composizione genetica delle popolazioni e possono pertanto essere considerati come le cause principali dell'evoluzione.

I processi evolutivi possono essere considerati di due tipi: neutrali e adattativi. I **processi evolutivi neutrali** modificano le frequenze geniche in modo del tutto casuale, cioè indipendentemente dalle capacità di sopravvivenza dell'individuo. I principali processi neutrali sono la migrazione e la deriva genetica. I **processi evolutivi adattativi** modificano, invece, le frequenze geniche in conseguenza del fatto che certi alleli sono in grado di conferire una maggiore sopravvivenza e un maggiore successo riproduttivo. La selezione naturale rappresenta il principale processo adattativo in quanto favorisce gli individui che hanno una più alta probabilità di sopravvivere e riprodursi contribuendo così al pool genico della generazione successiva. Le mutazioni possono essere sia neutrali che adattative (vantaggiose oppure deleterie).

Quadro 11.3 – Esempio di allelismo multiplo nell'uomo: gruppi sanguigni ABO

Nelle analisi genetiche vengono solitamente presi in considerazione coppie di alleli responsabili di caratteri antagonisti. L'allele più frequente nelle popolazioni naturali rappresenta l'allele selvatico mentre quello alternativo costituisce l'allele mutante. Tuttavia, in una popolazione di individui possono esistere non soltanto due alleli ma molti alleli di un dato gene: uno selvatico e più varianti. In queste situazioni si parla di **allelismo multiplo**. Benché di un gene in una data popolazione possano esistere serie alleliche numerose, è evidente che un singolo individuo diploide ad un dato locus può possedere al massimo due alleli, uno su ciascuno dei due cromosomi omologhi.

Un esempio di allelismo multiplo è costituito dal sistema di gruppo sanguigno ABO dell'uomo, scoperto da Karl Landsteiner all'inizio del 1900. La serie di alleli di questo sistema è responsabile di una importante caratteristica dei globuli rossi. Nell'effettuare trasfusioni di sangue o trapianti di organo bisogna combinare attentamente i gruppi sanguigni dei donatori e dei riceventi, dato che gli alleli che determinano il gruppo specificano antigeni cellulari che si trovano attaccati alla superficie esterna dei globuli rossi. Infatti, un determinato individuo possiede a livello di cellule e tessuti un gran numero di antigeni, molti dei quali sono estranei per un altro individuo. Quando gli antigeni sono riconosciuti come molecole estranee stimolano la produzione di specifiche proteine (anticorpi): l'anticorpo reagisce con l'antigene responsabile della sua formazione provocandone l'agglutinazione. Dato che esiste incompatibilità tra alcuni gruppi sanguigni, si rende necessaria un'attenta valutazione del gruppo sanguigno nelle trasfusioni di sangue e nei trapianti di organo affinché siano evitate le reazioni antigene-anticorpo.

Secondo il sistema ABO è possibile classificare tutte le persone in base agli antigeni presenti nei loro globuli rossi in quattro gruppi sanguigni: A, B, AB e O. I quattro fenotipi derivano dalle diverse combinazioni dei tre possibili alleli (I^A , I^B e i) del sistema ABO, per un totale di sei genotipi (**Tab. 11.10**). Le persone omozigoti per l'allele recessivo i sono di gruppo O. Sia l'allele I^A che l'allele I^B sono dominanti su i . Gli individui $I^A I^A$ e quelli $I^A i$ sono perciò di gruppo A, mentre gli individui $I^B I^B$ e quelli $I^B i$ sono di gruppo B. Gli individui eterozigoti $I^A I^B$ sono, invece, di gruppo AB, poiché manifestano entrambi i gruppi sanguigni A e B contemporaneamente.

La tipizzazione sanguigna di un individuo e l'analisi dell'ereditarietà dei gruppi vengono a volte usate nei casi di controversa paternità o maternità. In tali casi, i dati genetici non possono provare l'identità del genitore ma possono comunque essere usati per dimostrare che un individuo non è il genitore di un dato bambino: ad esempio, un bambino di fenotipo AB (genotipo $I^A I^B$) non potrebbe essere il figlio di un genitore di gruppo O (genotipo ii). Attualmente possono essere usati dei test più moderni, compresa la tipizzazione del DNA basata sulla rilevazione di marcatori molecolari (→ Cap. 17).

Tab. 11.10 – Possibili gruppi sanguigni dell'uomo e relativi genotipi.

Gruppo sanguigno (Fenotipo)	Antigene nei globuli rossi	Anticorpi del siero	Genotipo
O	–	α e β	ii
A	A	β	$I^A I^A$ oppure $I^A i$
B	B	α	$I^B I^B$ oppure $I^B i$
AB	A e B	–	$I^A I^B$

L'allele I^A determina l'antigene A sui globuli rossi e il siero del sangue delle persone di gruppo A ($I^A I^A$ o $I^A i$) contiene anticorpi che si formano spontaneamente contro l'antigene B (chiamati anticorpi β o anti-B), ma nessun anticorpo contro l'antigene A. Tali anticorpi faranno aggregare tutti i globuli rossi che possiedano l'antigene B. Al contrario, le persone di gruppo B ($I^B I^B$ o $I^B i$) hanno sui globuli rossi l'antigene B e nel siero del sangue anticorpi contro l'antigene A (chiamati anticorpi α o anti-A), ma nessun anticorpo contro l'antigene B. Sui globuli rossi delle persone di gruppo AB ($I^A I^B$) si trovano sia l'antigene A che l'antigene B, per cui nel siero delle persone di questo gruppo non vi sono né gli anticorpi anti-A né quelli anti-B. Nelle persone di gruppo O (*ii*), i globuli rossi non hanno né l'antigene A né l'antigene B e perciò il siero contiene sia gli anticorpi anti-A che quelli anti-B. Queste relazioni antigene-anticorpo sono illustrate nella **Fig. 11.16**. L'aggregazione (o agglutinazione) dei globuli rossi si osserva in tutti i casi in cui un anticorpo interagisce con l'antigene specifico. Dato che le cellule agglutinate non possono muoversi attraverso i capillari sottili, l'agglutinazione può portare ad insufficienza nella funzione di organi, anche molto grave da determinare la morte.

Le trasfusioni sicure tra individui con diverso gruppo sanguigno sono le seguenti: i) individui di gruppo sanguigno A producono l'antigene A, per cui il loro sangue può essere trasfuso solo a riceventi che non abbiano l'anticorpo anti-A, vale a dire a individui di gruppo A e AB; ii) individui di gruppo B producono l'antigene B, per cui il loro sangue può essere trasfuso solo a riceventi che non abbiano l'anticorpo anti-B, vale a dire a persone di gruppo B a AB; iii) individui AB producono entrambi gli antigeni A e B, per cui il loro sangue può essere trasfuso solo a riceventi che non abbiano né l'anticorpo anti-A né l'anticorpo anti-B, vale a dire a persone di gruppo AB; iv) individui O non producono né l'antigene A né quello B, per cui il loro sangue può essere trasfuso in qualsiasi ricevente, vale a dire a persone del gruppo A, B, AB e O.

Siero del gruppo sanguigno	Cellule del gruppo sanguigno			
	O	A	B	AB
O				
A				
B				
AB				

Fig. 11.15 – Reazioni antigeniche caratteristiche dei gruppi sanguigni dell'uomo: il siero del sangue di ciascuno dei quattro possibili gruppi sanguigni è stato mescolato con le cellule del sangue dei quattro tipi. In alcune delle possibili combinazioni, l'anticorpo del siero reagisce con l'antigene presente nei globuli rossi che ha determinato la sua formazione provocandone l'agglutinazione.

Si noti che gli individui del gruppo AB possono ricevere trasfusioni di sangue da individui di qualsiasi gruppo e per questo motivo sono definiti "accettori universali". Analogamente, il sangue di individui di gruppo O può essere trasfuso a qualsiasi ricevente e sono pertanto definiti "donatori universali".

11.6.4.1 Mutazione

La mutazione è una modificazione ereditaria del materiale genetico e rappresenta una sorgente di variabilità (**Fig. 11.17**). Il genetista sovietico S. Tshetverikov, nel 1926, fu il primo ad ipotizzare che la variabilità genetica conseguente alle mutazioni spontanee costituissero un elemento essenziale per i processi adattativi alla base dell'evoluzione. La mutazione intesa come alterazione a carico delle sequenze geniche può fornire nuovi alleli e comportare cambiamenti delle frequenze di quelli già presenti in una popolazione. Gli effetti delle mutazioni sulla struttura genetica delle popolazioni possono essere diversi a seconda che la mutazione di un allele in un altro ad un determinato locus sia un evento raro oppure ricorrente. La mutazione rara di solito non comporta sostanziali cambiamenti delle frequenze geniche, a meno che questa non risulti particolarmente vantaggiosa dal punto di vista della sopravvivenza, dell'adattamento e della riproduzione degli individui che la portano. Le mutazioni ricorrenti esercitano invece un'influenza più consistente sulla struttura genetica di una popolazione. Tuttavia, anche in questo caso i cambiamenti delle frequenze geniche dipendono dall'effetto che la mutazione esercita sulla sopravvivenza e sul potenziale riproduttivo dell'individuo che la eredita. Una nuova mutazione può infatti essere vantaggiosa, neutrale o deleteria: le mutazioni neutrali e deleterie tendono a insorgere secondo una frequenza molto superiore a quella delle mutazioni vantaggiose. Nel caso dei geni strutturali che codificano proteine, è stato accertato che le delezioni, le inserzioni e le sostituzioni di basi che provocano cambiamenti del codice a triplette o slittamenti del modulo di lettura portano molto spesso a proteine prive di funzionalità o comunque meno efficienti rispetto alla

proteina di tipo selvatico. Tali mutazioni sono pertanto deleterie, mentre le mutazioni puntiformi a carico delle sequenze non codificanti (introni) oppure di siti specifici della sequenza codificante che non provocano cambiamenti di senso, cioè non influenzano la sequenza amminoacidica della proteina, e quindi la sua funzionalità sono considerate neutrali. Le mutazioni vantaggiose costituiscono eventi piuttosto infrequenti poiché le alterazioni a carico della sequenza nucleotidica devono riflettersi in cambiamenti della sequenza amminoacidica di una proteina capaci di migliorarne l'efficienza e la funzionalità. Il tasso di mutazione, cioè la probabilità che un gene venga modificato in un altro gene da un evento mutazionale, si ritiene compreso soprattutto tra 1:100.000 e 1:1.000.000 per ogni generazione. Dato che queste stime sono tipicamente basate sulla valutazione di conversione di un gene selvatico (funzionale) in uno deleterio (non funzionale) è verosimile che il tasso delle mutazioni che producono alleli vantaggiosi sia notevolmente inferiore.

Per conoscere come e quanto il tasso di mutazione influenzi le frequenze alleliche di una popolazione nel corso del tempo, si può considerare un caso semplice di un allele A_1 che muta nell'allele A_2 con una frequenza pari a u ad ogni generazione. Se in una generazione la frequenza dell'allele originario è $f(A_1)=p_0$, la sua frequenza nella generazione successiva sarà $f(A_1)=p_0-up_0$. Se la mutazione avviene anche nella direzione opposta ad ogni generazione, assumendo che l'allele A_2 muti nell'allele A_1 con una frequenza pari a v , si ha che in una generazione la frequenza dell'allele mutato è $f(A_2)=q_0$, la frequenza dell'allele originario nella generazione successiva sarà $f(A_1)=p_0+vq_0$. Qualora la mutazione di A_1 , con frequenza iniziale p_0 , in A_2 avvenga con tasso pari a u e la retromutazione di A_2 , con frequenza iniziale q_0 , in A_1 avvenga con tasso pari a v , dopo una generazione si avrà un variazione di frequenza genica pari a:

$$\Delta q = up_0 - vq_0$$

dove up_0 rappresenta l'aumento della frequenza dell'allele A_2 dovuto alla mutazione e vq_0 rappresenta la diminuzione della frequenza dello stesso allele in seguito alla retromutazione nell'allele A_1 . Il cambiamento delle frequenze alleliche proseguirà fino a quando non viene raggiunto l'equilibrio tra il numero di alleli che mutano in A_2 e il numero di quelli che mutano in A_1 : il punto di equilibrio si ottiene infatti con $\Delta q=0$ e cioè quando $up_0=vq_0$, equazione che può anche essere scritta come $p_0/q_0=v/u$. Sostituendo in questa equazione a p_0 l'equivalente $1-q_0$ si ha che $q_0=u/(u+v)$. Analogamente, sostituendo a q_0 l'equivalente $1-p_0$, si ha che $p_0=v/(u+v)$.

Le frequenze dei due alleli in condizioni di equilibrio sono pertanto date dal rapporto tra singolo tasso di mutazione e tasso di mutazione complessivo:

$$p_i = \frac{v}{(u+v)} \quad q_i = \frac{u}{(u+v)}$$

Qualora la retromutazione verso l'allele selvatico A_1 ($A_1 \leftarrow A_2$) sia trascurabile e quindi la mutazione avvenga solo in una direzione ($A_1 \rightarrow A_2$), si ha che l'incremento della frequenza dell'allele mutato A_2 dopo una generazione è pari a: $\Delta q=up_0$ e pertanto $q_{n+1}=q_0+\Delta q$ e $p_{n+1}=p_0-\Delta q$. In realtà, le mutazioni non sono in grado di modificare in modo consistente le frequenze alleliche in una singola generazione. È possibile calcolare la variazione delle frequenze alleliche di mutazioni neutrali dopo un numero arbitrario di generazioni usando la seguente espressione: $(1-u)^t = p_t/p_0$, dove u è il tasso di mutazione dell'allele A_1 in A_2 , t è il numero di generazioni, p_0 è la frequenza dell'allele A_1 nella generazione iniziale e p_t è la frequenza di questo stesso allele dopo t generazioni.

Il sequenziamento del DNA genomico ha rivelato che la variazione genetica presente nelle popolazioni di specie vegetali è dovuta in parte a mutazioni neutrali

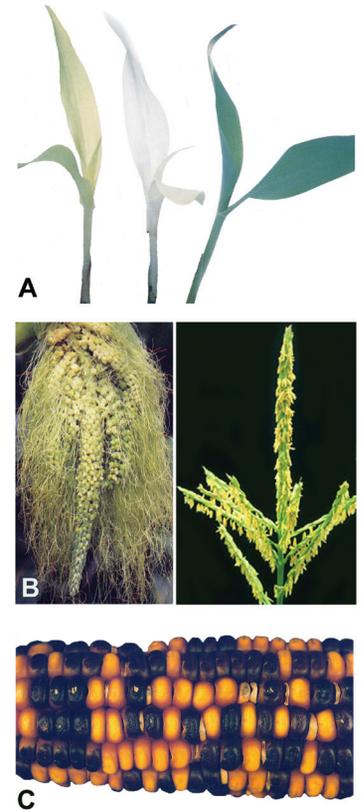


Fig. 11.17 – Mutazioni geniche evidenti a livello fenotipico: (A) mutante clorofilliano *l1* (*luteus*), (B) mutante del pennacchio *ts1* (*tassel seed*), (C) mutante per cariossidi difettive *dek2* (*defective kernels*) (da: *Mutants of maize*).

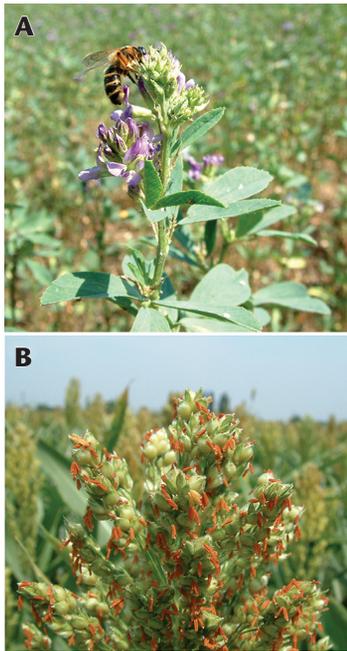


Fig. 11.18 – Flusso genico dovuto alla dispersione del polline: fiori di specie prevalentemente allogama, e autogama, ad impollinazione entomofila (A, erba medica) ed anemofila (B, sorgo).

che sono state fissate per effetto della deriva genetica casuale e in parte a mutazioni vantaggiose per la sopravvivenza che sono state quindi incrementate per effetto della selezione naturale. In generale, le mutazioni neutrali fanno variare molto lentamente le frequenze alleliche di una popolazione anche dopo un numero elevato di generazioni. Altri processi, invece, come il flusso genico, la deriva genetica e la selezione naturale hanno effetti molto più consistenti sulle frequenze alleliche e sulla struttura genetica delle popolazioni naturali.

11.6.4.2 Migrazione e flusso genico

La migrazione consiste in uno scambio di geni (**flusso genico**) che può avvenire tra due popolazioni e che può provocare cambiamenti della loro struttura genetica quando le popolazioni interessate sono caratterizzate da differenti frequenze alleliche e quando il numero di individui coinvolti è sufficiente. Nelle specie vegetali la migrazione si realizza attraverso il polline (introduzione di singoli alleli) oppure mediante i semi (introduzione di combinazioni alleliche) ed in entrambi i casi può comportare variazioni delle frequenze geniche e quindi cambiamenti del pool genico della popolazione ricevente (**Fig. 11.18**). Supponendo che una popolazione sia costituita da una proporzione di m individui introdotti da quella donatrice e di $1-m$ individui di quella originaria, e che la frequenza di un certo allele sia p_d nella popolazione donatrice e p_o in quella originaria, la frequenza dell'allele considerato nella popolazione mista sarà:

$$p_m = mp_d + (1-m)p_o = m(p_d - p_o) + p_o$$

Disponendo di queste informazioni si può calcolare la variazione della frequenza genica tra la popolazione mista e quella originaria come $\Delta p = p_m - p_o$, che equivale alla espressione seguente:

$$\Delta p = m(p_d - p_o)$$

L'entità del cambiamento delle frequenze geniche al locus in questione dipende pertanto dalla consistenza del flusso genico e dalla differenza tra la frequenza genica della popolazione originaria e di quella donatrice.

Dal momento che implica scambio di geni o movimento di genotipi tra popolazioni, la migrazione tende a limitare la divergenza genetica di queste. Per capire l'effetto omogeneizzante della migrazione vengono usualmente applicati modelli semplici di struttura di popolazioni. Uno di questi è il **modello dell'isola**, secondo il quale una grande popolazione può essere suddivisa in molte sottopopolazioni disperse geograficamente. Considerando due alleli A e a con frequenze medie pari, rispettivamente, a p e q , bisogna innanzitutto supporre che la migrazione avvenga in maniera tale che gli individui coinvolti dalla migrazione stessa siano rappresentativi della sottopopolazione per quanto riguarda le frequenze alleliche.

La quantità di migrazione è misurata dal valore di m che corrisponde alla probabilità che un gene qualsiasi in ogni sottopopolazione provenga da un individuo introdotto. Prendendo in esame un gene scelto a caso in ciascuna sottopopolazione alla generazione t , questo allele alla generazione $t-1$ poteva appartenere ad un individuo della stessa sottopopolazione (con probabilità $1-m$) e, quindi, essere un allele A con frequenza p_{t-1} equivalente alla frequenza di A nella sottopopolazione in questione alla generazione $t-1$, oppure appartenere ad un individuo introdotto della generazione $t-1$ (con probabilità m) e quindi essere un allele A con frequenza \bar{p} pari a quella media delle sottopopolazioni assunta invariabile nel corso delle generazioni. Nel complesso, la frequenza allelica alla generazione t equivale a $p_t = p_{t-1}(1-m) + \bar{p}m$ la cui soluzione

in termini di p_0 è la seguente:

$$p_t = \bar{p} + (p_0 - \bar{p})(1 - m)^t$$

dove p_0 è la frequenza iniziale di A della sottopopolazione presa in esame. Come esempio di applicazione di tale equazione, si suppongano solamente due popolazioni con frequenze iniziali di A uguali rispettivamente a 0,2 e 0,8, con $m = 0,10$: ciò significa che il 10% degli individui, contraddistinti da una frequenza allelica media di A pari a $\bar{p} = (0,2+0,8)/2 = 0,5$, migrano da ciascuna popolazione ad ogni generazione. Al fine di calcolare la frequenza allelica di A nelle due popolazioni dopo dieci generazioni, per la popolazione con frequenza allelica iniziale uguale a 0,2, bisogna sostituire $p_0 = 0,2$, $\bar{p} = 0,5$ e $m = 0,10$ nella equazione corrispondente ottenendo così $p_{10} = 0,5 + (0,2 - 0,5)(1 - 0,10)^{10} = 0,395$, mentre per l'altra popolazione bisogna sostituire $p_0 = 0,8$, $\bar{p} = 0,5$ e $m = 0,10$ perciò $p_{10} = 0,5 + (0,8 - 0,5)(1 - 0,10)^{10} = 0,605$.

Un altro esempio sull'uso della equazione idonea a calcolare il cambiamento nel tempo delle frequenze geniche in seguito ad eventi di migrazione è mostrato in **Fig. 11.19**, con riferimento a cinque sottopopolazioni con frequenze iniziali uguali a 1, 0,75, 0,50, 0,25 e 0, ancora con $m = 0,10$. È interessante notare con quanta rapidità le frequenze alleliche convergono verso lo stesso valore medio di frequenza genica, in questo caso uguale a 0,5. Sebbene l'equazione riguardante la migrazione sia matematicamente simile a quella riguardante la mutazione, le implicazioni biologiche dei due fenomeni a livello di popolazione sono completamente differenti. Poiché i tassi di migrazione sono tipicamente molto più elevati dei tassi di mutazione, i cambiamenti di frequenze alleliche sono generalmente più veloci con la migrazione.

11.6.4.3 Unioni casuali in piccole popolazioni: deriva genetica

L'introduzione in un nuovo ambiente di un gruppo relativamente poco numeroso di individui, così come la frammentazione di una popolazione in sottogruppi di dimensioni ridotte può dare origine a nuove popolazioni aventi una composizione genetica modificata rispetto a quella originaria a causa della deriva genetica casuale. Affinché possa essere raggiunto e mantenuto un equilibrio genetico è infatti necessario che la popolazione sia molto numerosa, cioè composta da un numero infinitamente grande di individui. Nessuna popolazione soddisfa però tale caratteristica, sebbene quelle molto grandi si comportino in modo simile a quelle teoricamente infinite. In una popolazione di dimensioni ridotte, invece, le fluttuazioni casuali delle frequenze alleliche al momento del campionamento dei gameti possono causare, nel corso delle generazioni, deviazioni dei rapporti genotipici rispetto a quelli attesi. In altri termini, ad ogni generazione le frequenze genotipiche diventano indipendenti dalle frequenze geniche della generazione precedente. Qualora le unioni siano casuali all'interno di popolazioni che rimangono piccole per un periodo di tempo prolungato, il pool genico della generazione parentale che contribuisce alla generazione filiale può non essere sufficientemente ampio e le frequenze geniche possono così variare sensibilmente di generazione in generazione per il semplice effetto del campionamento casuale dei gameti che combinandosi portano alla formazione degli individui della generazione successiva. Le popolazioni di specie vegetali distribuite in una vasta area geografica che risultano suddivise in piccoli gruppi di piante isolati tra loro possono andare incontro a questo fenomeno, chiamato deriva genetica. In queste condizioni, le frequenze geniche sono soggette perciò a cambiamenti continui e casuali con il succedersi delle generazioni. La deriva genetica ha un effetto tanto più marcato quanto più piccola è la dimensione della popolazione e quanto più bassa è la frequenza di un allele nella

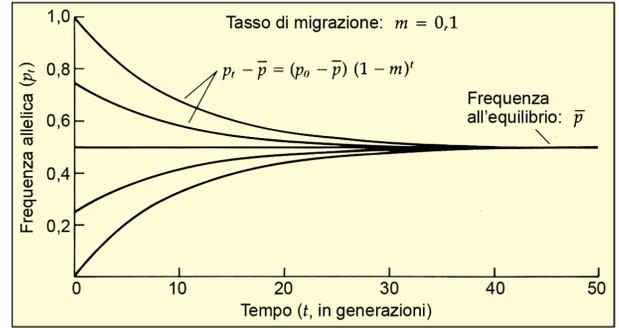


Fig. 11.19 – Cambiamento nel tempo delle frequenze geniche in cinque sottopopolazioni interessate da migrazione di individui con un tasso $m = 0,1$ per generazione.

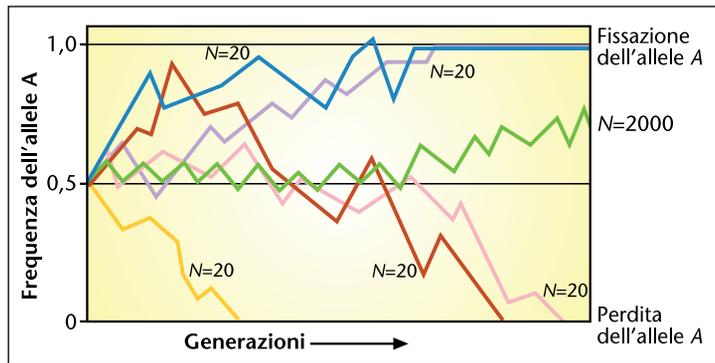


Fig. 11.20 – Variazione ipotetica della frequenza allelica per effetto della deriva genetica.

zione in generazione in conseguenza del campionamento casuale, fino a portare alla eliminazione di un allele e alla fissazione dell'altro allele. Quando l'allele è stato fissato significa che tutti gli individui sono in condizione omozigote al locus relativo e che le frequenze alleliche non possono più modificarsi. Le frequenze alleliche nella popolazione più numerosa subiscono, invece, piccole oscillazioni e la deriva genetica potrà condurre anche in questo caso all'omozigosi, ma solo dopo moltissime generazioni. Supponendo che le popolazioni considerate rimangano numericamente costanti nel corso delle generazioni, i gameti complessivamente coinvolti nella fase riproduttiva dovranno necessariamente essere 4.000, con riferimento alla popolazione più grande, e 40 nelle varie popolazioni più piccole. Tali gameti dovrebbero essere prodotti, rispettivamente, nel rapporto di 2.000 A : 2.000 a e 20 A : 20 a. Per il solo effetto del campionamento dei gameti si possono però avere degli scostamenti in termini di contributi alla generazione successiva. Tuttavia, l'entità degli scostamenti rispetto ai rapporti attesi è inversamente proporzionale alla dimensione della popolazione: ciò significa che è tanto maggiore quanto minore è il numero totale degli alleli. Ad esempio, un rapporto 23:17 nelle popolazioni più piccole è sicuramente più frequente di un rapporto 2300:1700, come può essere verificato con il semplice calcolo del χ^2 , ed ha conseguenze più consistenti. Il valori calcolati di χ^2 (0,9 contro 90) suggeriscono, infatti, che nel primo caso lo scostamento è probabile e può essere considerato casuale, mentre nel secondo caso lo scostamento rispetto al rapporto atteso è del tutto improbabile. Risulta quindi evidente che la deriva genetica ha un effetto tanto più marcato quanto più piccola è la dimensione della popolazione. La differenza tra le frequenze geniche (δq) di generazioni successive causata dalla dimensione ridotta della popolazione può essere misurata attraverso la varianza riferita a tutti i potenziali valori che q può assumere nel numero di potenziali campioni di gameti prodotti. La quota di variabilità tra popolazioni di N individui che risulta dalla deriva genetica è pertanto misurata dalla varianza della frequenza allelica:

$$\sigma^2 = \frac{q(1-q)}{2N}$$

Poiché la varianza è una funzione di q e di N , è evidente che la fluttuazione è tanto maggiore quanto più piccolo è il numero degli individui, ed inoltre, assumendo una numerosità costante della popolazione, è tanto maggiore quanto più q è prossimo a 0,5 fino ad essere massima quando le frequenze di A e a sono uguali.

La popolazione di gameti prodotti da N individui ha una dimensione molto grande, indipendentemente dal numero di parentali considerati. Di questa popolazione solo un campione di $2N$ gameti (N maschili e N femminili) corrisponde a quello che dà origine alla generazione filiale. Assumendo che la dimensione della popolazione resti pressoché costante nel corso delle generazioni, la frequenza genica nel campione può variare casualmente da 0 a 1 con probabilità definite dalla distribuzione binomiale i cui

popolazione. Un ruolo di primo piano nello sviluppo e nella formulazione del concetto di deriva genetica casuale fu svolto da Sewall Wright intorno al 1930.

La deriva genetica favorisce la perdita oppure la fissazione di un allele in una popolazione. Nella **Fig. 11.20** sono illustrate le potenziali variazioni delle frequenze alleliche per effetto della deriva genetica in una popolazione grande ($N=2.000$) e alcune popolazioni piccole ($N=20$), assumendo che le frequenze alleliche iniziali siano uguali: $p(A)=q(a)=0,5$. Nelle popolazioni di numerosità molto ridotta le frequenze alleliche possono fluttuare ampiamente di genera-

termini sono dati dallo sviluppo del binomio $(p + q)^{2N}$. In **Fig 11.21** è riportata la distribuzione normale, che nel caso di N sufficientemente elevato, interpola con buona approssimazione la distribuzione binomiale. L'ampiezza di questa distribuzione, definita dalla deviazione standard, dipende dalla dimensione del campione dei gameti ($2N$). In altri termini, l'entità dei cambiamenti delle frequenze geniche causate dalla dimensione ridotta della popolazione può anche essere misurata usando la deviazione standard della distribuzione binomiale generata da $(p+q)^{2N}$ che risulta essere:

$$\sigma = \sqrt{\frac{pq}{2N}}$$

dove p e q sono le frequenze alleliche e $2N$ il numero totale degli alleli nel gruppo di individui considerato. Ad esempio, in una popolazione di 2.000 individui (N) con frequenze $p(A)=0,8$ e $q(a)=0,2$ si ha che in oltre il 95% dei casi ($\mu \pm 2\sigma$) la frequenza genica di A fluttuerà tra 0,8013 e 0,7987 e quella di a oscillerà tra 0,2013 e 0,1987. Nel caso, invece, di una popolazione di appena 20 individui caratterizzata dalle stesse frequenze geniche, la frequenza di A potrà variare tra 0,9265 e 0,6735 e quella di a tra 0,3265 e 0,0740, per il solo effetto del caso. Nelle popolazioni poco numerose che vanno incontro a deriva genetica, le frequenze geniche sono pertanto soggette a cambiamenti casuali di generazione in generazione e il gruppo di individui che compone una popolazione risulta caratterizzato da frequenze geniche che ad ogni generazione variano in modo continuo e secondo una direzione imprevedibile. La conseguenza più estrema della deriva genetica, cioè la scomparsa di un allele e quindi la fissazione dell'altro allele al locus relativo, si ottiene tanto più velocemente quanto maggiore è la differenza tra le frequenze relative nella popolazione. Ad esempio, nel caso di alleli rari con $q(A) \leq 0,01$, in una popolazione le frequenze genotipiche all'equilibrio saranno le seguenti: $p^2(A'A')=98,01\%$, $2pq(A'A)=1,98\%$ e $q^2(AA)=0,01\%$. Qualora contribuiscano soltanto pochi individui alla generazione successiva, per effetto del caso sarebbe molto probabile che nessuno di quelli omozigoti per l'allele A risulti compreso nel gruppo parentale, determinando così la scomparsa di questo dal pool genico e la fissazione dell'allele A' che nella popolazione assumerebbe una frequenza pari a $p(A')=1,0$.

La probabilità di fissazione di un allele dipende dalla sua frequenza nella popolazione. Nel caso di comparsa di una mutazione neutra nella popolazione, la probabilità che questo nuovo allele venga fissato o eliminato per effetto della deriva genetica casuale può essere calcolata: la probabilità di fissazione è data da $1/(2N)$, mentre quella di eliminazione è $1 - [1/(2N)]$. Quando il gruppo di individui è poco numeroso, cioè N è piccolo, la possibilità che insorgano nuove mutazioni è molto bassa, ma qualora ciò avvenisse la probabilità che queste vengano fissate è relativamente alta. Anche il numero di generazioni richieste affinché un nuovo allele possa essere fissato dipende dal numero di individui che compongono la popolazione. Ipotizzando che gli individui maschili e femminili contribuiscano in uguale misura a ciascuna generazione, il numero medio di generazioni è uguale a: $t=4N$.

Le dimensioni delle popolazioni influenzano quindi l'entità della deriva genetica casuale: l'effetto sarà minimo nelle popolazioni grandi, mentre può costituire la componente evolutiva principale nelle piccole popolazioni. Tale effetto della dimensione di una popolazione si valuta misurando il grado di eterozigosi nel corso delle generazioni. Il modello di riferimento definito dall'equilibrio Hardy-Weinberg presuppone che la popolazione sia costituita da una comunità molto grande ($N \rightarrow \infty$) ed isolata di individui tra i quali le unioni sessuali siano totalmente casuali. In queste condizioni le frequenze alleliche ad un locus rimangono costanti e la frequenza

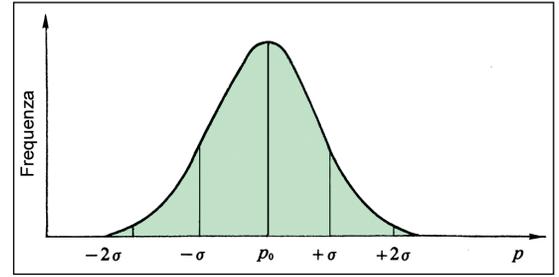


Fig. 11.21 – Distribuzione binomiale di p .

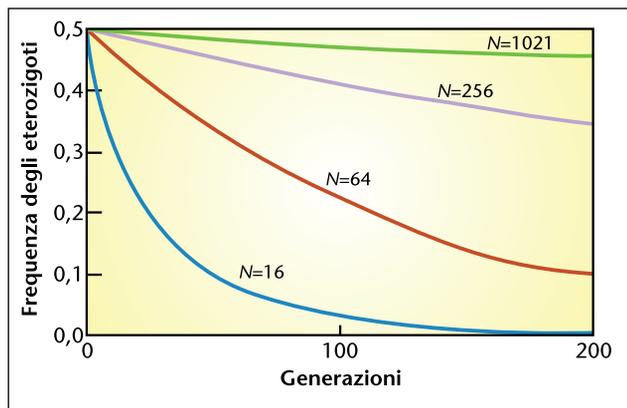


Fig. 11.22 – Andamento della diminuzione della frequenza degli eterozigoti dovuta a deriva genetica casuale in popolazioni di dimensioni N diverse assumendo frequenze geniche iniziali $p=q=0,5$.

Il fattore $1/(2N)$ rappresenta la diminuzione di eterozigosità che si verifica in una generazione a causa della deriva genetica. L'equazione che esprime l'effetto cumulativo della deriva genetica sulla diminuzione di eterozigosità nel corso di t generazioni è la seguente:

$$H' = \left(1 - \frac{1}{2N}\right)H$$

$$H'_t = \left(1 - \frac{1}{2N}\right)^t H$$

Dopo molte generazioni il grado di eterozigosi sarà nullo, la popolazione avrà perso tutta la variabilità genetica in quanto possederà solo un allele e quindi le frequenze geniche saranno $p=1$ e $q=0$ oppure $p=0$ e $q=1$. Gli effetti teorici della deriva genetica sulle frequenze di alleli neutrali attese in un campione di popolazioni secondo il modello di A. Fisher e S. Wright sono riportati in **Fig. 11.23**. Ipotizzando che ciascuna popolazione sia composta da 16 individui e che abbia frequenze alleliche iniziali pari a 0,5, con il procedere delle generazioni le frequenze divergono fino a risultare dopo 19 generazioni pari a 0 o 1 nella maggior parte delle popolazioni. La validità di questo modello teorico venne dimostrata nel 1956 con un esperimento classico condotto da P. Buri: per 19 generazioni successive analizzò le frequenze di due alleli (bw e $+$) al locus che determina il colore dell'occhio in *Drosophila melanogaster* impiegando 107 popolazioni aventi una frequenza iniziale di bw pari a 0,5. La dimensione effettiva di ogni popolazione era sempre di 16 individui: in ogni generazione venivano infatti scelti casualmente 8 maschi e 8 femmine che potevano incrociarsi liberamente per fornire la generazione successiva. Le frequenze alleliche nel corso delle prime generazioni risultarono intorno a 0,5 in tutte le popolazioni, mentre con il procedere delle generazioni le frequenze alleliche nelle popolazioni divergevano, tanto che alla generazione 19 la frequenza osservata di bw risultava 0 o 1 nella maggior parte delle popolazioni esaminate. I risultati osservati in questo esperimento, relativamente alle oscillazioni delle frequenze alleliche, concordarono pienamente con quelli attesi secondo il modello teorico, in termini di $\mu \pm \sigma$.

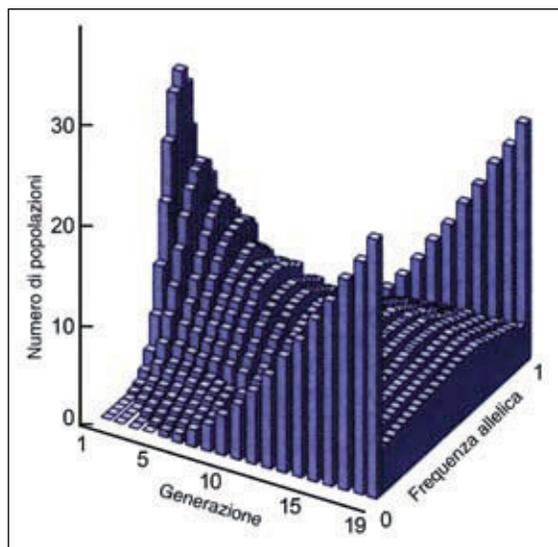


Fig. 11.23 – Risultati teorici attesi secondo il modello di Fisher-Wright relativamente al destino di alleli neutrali in un campione di popolazioni.

specie quando alcuni gruppi poco numerosi di individui originano popolazioni isolate dal punto di vista riproduttivo. Le frequenze alleliche di queste popolazioni sono suscettibili di deriva genetica e poiché si tratta di un processo casuale generalmente

accade che la composizione genetica delle popolazioni isolate risulti molto diversificata rispetto a quella delle altre popolazioni della stessa specie.

Un caso particolare di deriva genetica, noto come **effetto collo di bottiglia**, si verifica quando in seguito ad eventi naturali particolarmente avversi, come inondazioni o siccità, le dimensioni di una popolazione di una vasta area geografica possano ridursi in modo consistente ed improvviso. In queste situazioni, la perdita di buona parte degli individui della popolazione originaria fa sì che la popolazione residua venga dominata dalla deriva genetica poiché, qualora persista la condizione di bassa numerosità, potranno verificarsi pochissimi eventi riproduttivi. Le conseguenze dipendono dalla composizione genotipica degli individui sopravvissuti e dal tipo di fluttuazioni subite dalle frequenze geniche. Di fatto la popolazione che è stata interessata dall'effetto collo di bottiglia può tornare alle dimensioni originali, ma la sua variabilità genetica potrà risultare anche inferiore rispetto a quella originaria e certi alleli potranno perfino non essere più presenti nel suo pool genico (**Fig. 11.24**).

Un altro caso di deriva genetica è quello conosciuto come **effetto del fondatore**. Tale fenomeno si verifica quando una popolazione viene inizialmente stabilita da un piccolo gruppo di individui che si separano da una popolazione più grande e che si riproducono tra di loro in una nuova località. Sebbene la popolazione dopo un certo numero di generazioni possa risultare costituita da molti individui, il suo pool genico sarà sempre quello derivato dall'insieme di alleli portati dagli individui fondatori. Ad esempio, alcuni semi provenienti da una grande popolazione possono essere introdotti in un nuovo ambiente ed originare così una popolazione altrettanto numerosa: in una situazione di questo tipo la popolazione fondatrice in genere evidenzia una variabilità genetica inferiore a quella della popolazione originaria ed inoltre le frequenze alleliche possono risultare anche molto diverse tra le due popolazioni. In alcuni casi l'effetto del fondatore è stato ritenuto responsabile dell'evoluzione di nuove specie.

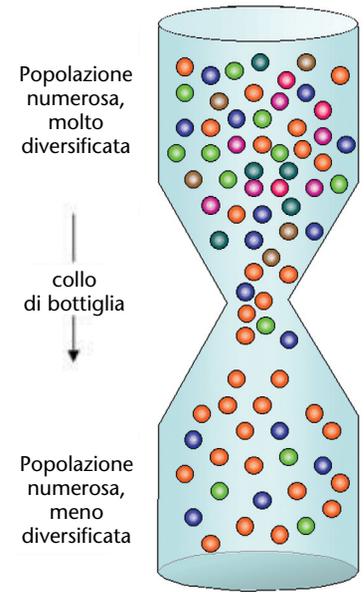


Fig. 11.24 – Rappresentazione grafica dell'effetto collo di bottiglia.

11.6.4.4 Selezione naturale

Nella metà del XIX secolo C. Darwin e A.R. Wallace proposero indipendentemente la teoria della selezione naturale secondo la quale le condizioni esistenti in natura determinano la sopravvivenza e la riproduzione degli individui più adattati all'ambiente. Gli individui con le caratteristiche che li rendono più adattati all'ambiente sarebbero in sostanza quelli aventi la maggiore probabilità di sopravvivere e con la maggiore fecondità e capacità di riprodursi, potendo così contribuire in misura maggiore con i loro discendenti alla generazione successiva.

Col termine di selezione naturale si indica pertanto l'insieme dei fattori che tendono a favorire o sfavorire un dato genotipo e quindi ad aumentare o diminuire a livello di popolazione le frequenze degli alleli che lo compongono. La capacità di sopravvivenza e di riproduzione di un individuo, o *fitness*, misurata come contributo alla generazione successiva in termini di individui, è indicata con la lettera w . La legge di Hardy-Weinberg assume assenza di selezione naturale, vale a dire che affinché questa legge sia valida è necessario ipotizzare che tutti gli individui abbiano lo stesso valore adattativo e riproduttivo, contribuendo nella stessa misura alla generazione successiva. Tuttavia, gli individui che costituiscono una popolazione possono in realtà diversificarsi per vitalità e fertilità: per il genotipo con la capacità riproduttiva più alta si assume $w=1$. Qualora le differenze in termini di capacità adattativa e riproduttiva risultino associate con la presenza o l'assenza di determinati geni, l'effetto della selezione naturale si traduce in un cambiamento delle frequenze alleliche e delle frequenze genotipiche nel corso delle generazioni.

La selezione naturale agisce in funzione del fenotipo individuale che a sua volta è determinato dal genotipo e dalla sua interazione con l'ambiente. Relativamente ai

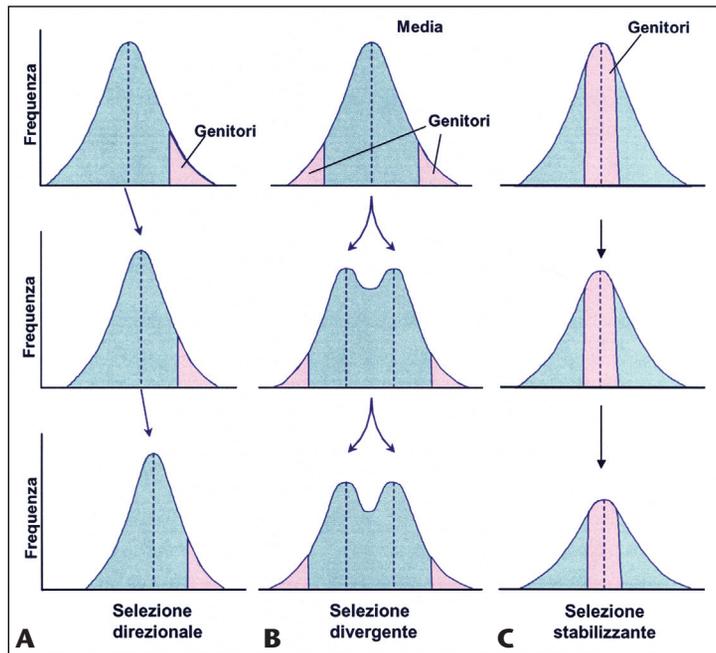


Fig. 11.25 – Tipologie di selezione con effetti sulla media e sulla variazione quantitativa: (A) selezione direzionale; (B) selezione divergente; (C) selezione stabilizzante.

condizioni ambientali contrastanti. Ad esempio, la dimensione dell'apparato radicale può diversificarsi tra gruppi di individui della stessa popolazione che si trovano a vivere in ambienti siccitosi e umidi. La **selezione stabilizzante** favorisce invece la sopravvivenza e la riproduzione degli individui con valore fenotipico intermedio. Per esempio, quando le piante con il colore del fiore più appariscente sono preferite dai pronubi responsabili dell'impollinazione (**Fig. 11.26**).

L'azione quantitativa della selezione naturale dipende pertanto dalla fitness di un individuo, equivalente ad una sorta di idoneità biologica del suo fenotipo ad uno specifico ambiente: da questo punto di vista la **fitness darwiniana** non è altro che la probabilità relativa che un determinato fenotipo ha di sopravvivere, riprodursi e contribuire al pool genico della generazione successiva rispetto agli altri fenotipi della stessa popolazione. Ciò significa che in una popolazione mendeliana, i due geni possibili ad un dato locus sottoforma di alleli *A* e *a* possono determinare un valore di fitness distinto per ognuna delle tre classi genotipiche. Indicando con w_{AA} , w_{Aa} e w_{aa} la fitness dei genotipi *AA*, *Aa* e *aa*, rispettivamente, la fitness media della popolazione sarà la seguente:

$$\bar{w} = p^2(AA)w_{AA} + 2pq(Aa)w_{Aa} + q^2(aa)w_{aa}$$

Tuttavia, in una popolazione soggetta all'azione della selezione naturale tale somma non è uguale ad uno, come invece accadrebbe in caso di equilibrio Hardy-Weinberg. Dividendo i singoli membri per la fitness media della popolazione si ha che:

$$p^2(AA)w_{AA}/\bar{w} + 2pq(Aa)w_{Aa}/\bar{w} + q^2(aa)w_{aa}/\bar{w} = 1$$

Le frequenze genotipiche possono essere derivate a partire da questa equazione nel modo seguente:

$$\begin{aligned} f(AA) &= p^2(AA)w_{AA}/\bar{w} \\ f(Aa) &= 2pq(Aa)w_{Aa}/\bar{w} \\ f(aa) &= q^2(aa)w_{aa}/\bar{w} \end{aligned}$$

caratteri quantitativi, la selezione naturale può agire secondo diverse modalità a seconda delle conseguenze che determina a carico delle frequenze fenotipiche della popolazione. La selezione può essere direzionale, divergente e stabilizzante: ognuno di questi tipi di selezione ha conseguenze diverse sul valore fenotipico medio di un carattere quantitativo e sulla sua variazione a livello di popolazioni (**Fig. 11.25**). La **selezione direzionale** favorisce la sopravvivenza e la riproduzione degli individui di una frazione estrema della distribuzione normale determinando lo spostamento del valore fenotipico medio della popolazione in tale direzione. Per esempio, in caso di competizione per la luce sarebbero avvantaggiate le piante più alte. La **selezione divergente** favorisce gli individui di entrambe le classi fenotipiche estreme della curva di distribuzione che tende pertanto ad assumere una forma bimodale: secondo questo modello nell'ambito della specie possono diversificarsi popolazioni con caratteristiche fenotipiche divergenti che rendono massimo l'adattamento in

Pertanto, le frequenze alleliche in seguito alla selezione naturale sono le seguenti:

$$f(A) = p' = f(AA) + \frac{1}{2}f(Aa) = p^2(AA)w_{AA} / \bar{w} + pq(Aa)w_{Aa} / \bar{w}$$

$$f(a) = q' = f(aa) + \frac{1}{2}f(Aa) = q^2(aa)w_{aa} / \bar{w} + pq(Aa)w_{Aa} / \bar{w}$$

Le variazioni delle frequenze alleliche dovute alla selezione naturale possono essere calcolate come $\Delta p = p' - p$ per l'allele A e $\Delta q = q' - q$ per l'allele a.

Il metodo generale per la determinazione della variazione delle frequenze genotipiche e alleliche in seguito a selezione naturale è riassunto in **Tab. 11.11**. La formula generale della selezione per il calcolo della variazione della frequenza genica è la seguente:

$$\Delta = \frac{pq[p(w_{Aa} - w_{AA}) + q(w_{aa} - w_{Aa})]}{\bar{w}}$$

A parità di condizioni di fitness dei genotipi, l'entità del cambiamento della frequenza dipende dal prodotto delle frequenze geniche ed è quindi massimo quando $p=q=0,5$. L'effetto della selezione è invece nullo quando $q=0$ e di conseguenza quando $p=1$ e ciò equivale a dire che la selezione è inefficace in assenza di variabilità genetica. La quantità $p(w_{Aa} - w_{AA}) + q(w_{aa} - w_{Aa})$ rappresenta l'effetto additivo delle sostituzioni alleliche sulla fitness: $w_{Aa} - w_{AA}$ misura l'effetto della sostituzione $Aa \rightarrow AA$, mentre $w_{aa} - w_{Aa}$ quello della sostituzione $aa \rightarrow Aa$. Ognuna di queste sostituzioni è ponderata in accordo con le frequenze alleliche p e q . Inoltre, la formula generale suggerisce che il cambiamento della frequenza genica è inversamente proporzionale alla fitness media della popolazione: maggiore è il valore adattativo medio della popolazione e minori sono i cambiamenti delle frequenze geniche.

Gli effetti della selezione naturale possono essere diversi a seconda delle relazioni tra le fitness dei possibili genotipi. Quando il valore adattativo dei tre genotipi è lo stesso ($w_{AA} = w_{Aa} = w_{aa}$) significa che non c'è selezione naturale, mentre quando è diverso significa che agisce la selezione naturale. La selezione può essere di più tipi:

- i) selezione contro l'allele recessivo, quando il valore adattativo del genotipo eterozigote equivale a quello del genotipo omozigote dominante ed è maggiore di quello del genotipo omozigote recessivo ($w_{AA} = w_{Aa} = 1$ e $w_{aa} < 1$);
- ii) selezione contro l'allele dominante, quando il genotipo eterozigote possiede un valore adattativo uguale all'omozigote dominante, ma inferiore a quello dell'omozigote recessivo ($w_{AA} = w_{Aa} < 1$ e $w_{aa} = 1$);
- iii) selezione in assenza di dominanza, quando l'eterozigote ha un valore adattativo intermedio rispetto ai genotipi omozigoti ($w_{AA} > w_{Aa} > w_{aa}$ con $w_{AA} = 1$ oppure $w_{AA} < w_{Aa} < w_{aa}$ con $w_{aa} = 1$);

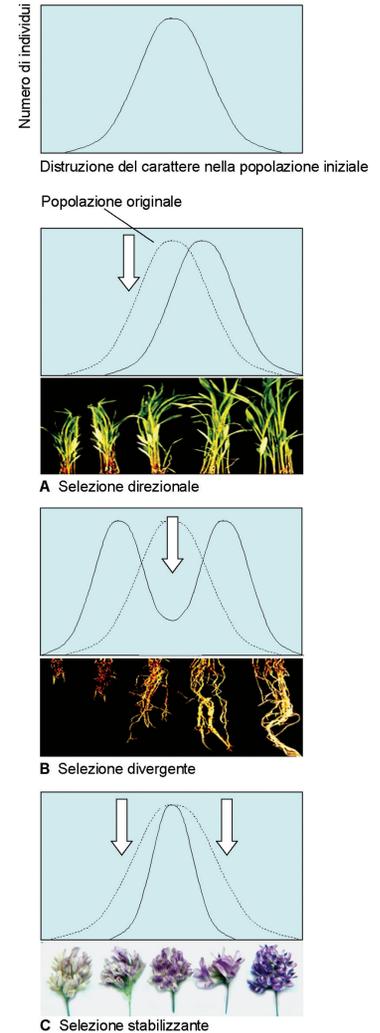


Fig. 11.26 – Tipologie di selezione in relazione alla fitness degli individui, esempi connessi a lunghezza degli steli (A), ampiezza delle radici (B) e colore dei fiori (C).

	Genotipi		
	A ¹ A ¹	A ¹ A ²	A ² A ²
Frequenze genotipiche iniziali	p^2	$2pq$	q^2
Fitness	W_{11}	W_{12}	W_{22}
Frequenze genotipiche dopo selezione	$p^2 W_{11}$	$2pq W_{12}$	$q^2 W_{22}$
Frequenze genotipiche relative dopo selezione	$p' = \frac{p^2 W_{11}}{\bar{W}}$	$H' = \frac{2pq W_{12}}{\bar{W}}$	$Q' = \frac{q^2 W_{22}}{\bar{W}}$

Frequenza allelica in seguito a selezione: $p' = p + 1/2(H)$ $q' = 1 - p'$
 Variazione della frequenza allelica dovuta alla selezione: $\Delta p = p' - p$
 Fitness media: $\bar{W} = p^2 W_{11} + 2pq W_{12} + q^2 W_{22}$

Tab. 11.11 – Metodo generale per la determinazione della variazione delle frequenze genotipiche e alleliche in seguito a selezione naturale.

- iv) selezione a favore dell'eterozigote, quando l'eterozigote ha un valore adattativo superiore a quello di entrambi i genotipi omozigoti ($w_{Aa}=1$ e $w_{AA} \neq w_{aa} < 1$);
- v) selezione contro l'eterozigote, quando l'eterozigote ha un valore adattativo inferiore a quello di entrambi i genotipi omozigoti ($w_{Aa} < 1$ e w_{AA} e/o $w_{aa}=1$).

In presenza di dominanza, sia la selezione contro l'allele recessivo che quella contro l'allele dominante rappresentano esempi classici di selezione direzionale poiché portano alla eliminazione o comunque ad una consistente riduzione di uno degli

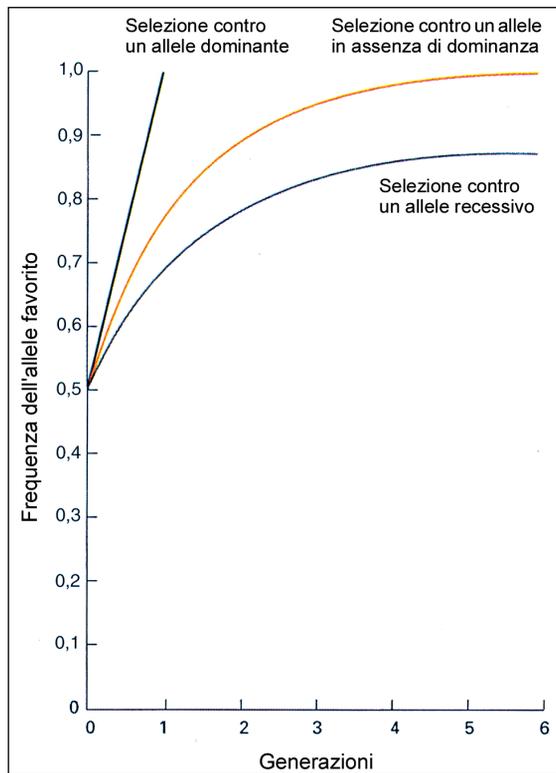


Fig. 11.27 – Variazioni delle frequenze geniche in caso di selezione contro un allele dominante, contro un allele recessivo e contro un allele in assenza di dominanza.

alleli, così come avviene con la selezione in assenza di dominanza. Il confronto tra le variazioni delle frequenze alleliche in presenza e in assenza di dominanza è riportato in **Fig. 11.27**. La selezione a favore dell'eterozigote è invece un esempio di generazione stabilizzante poiché ha come risultato il raggiungimento di un equilibrio stabile che una volta raggiunto è in grado di bloccare ogni ulteriore processo evolutivo. Infine, la selezione contro l'eterozigote può rappresentare un esempio di selezione disruptiva capace di determinare una variazione divergente delle frequenze alleliche.

Le pressioni selettive naturali possono agire contro gli individui meno adatti ed aventi un determinato genotipo con diversa intensità relativa, usualmente misurata dal **coefficiente di selezione** (s) che equivale al complemento ad uno della *fitness*: $s=1-w$. Il coefficiente di selezione del genotipo più favorito si assume per convenzione pari a $s=0$, mentre i genotipi sfavoriti dalla selezione hanno valori di $s>0$ e comunque $s \leq 1$. Un caso estremo si ha quando un allele recessivo è letale o causa sterilità: il genotipo omozigote per questo allele avrà un valore di $s=1$ e quindi $w=0$. Qualora, invece, la selezione agisca contro i recessivi diminuendone la fitness senza però comprometterne la sopravvivenza o impedirne la riproduzione, il valore di w risulterà inferiore rispetto a quella degli altri due possibili genotipi, ma maggiore di zero. Considerando una sola coppia allelica e una popolazione costituita da genotipi AA , Aa e aa aventi frequenze pari rispettivamente a p^2 , $2pq$ e q^2 , la fitness dei genotipi AA e Aa può essere assunta pari a $w=1$ e quella dei genotipi aa uguale a $w=1-s$

(**Tab. 11.12**). Le frequenze genotipiche dopo l'azione della selezione contro un carattere recessivo si ottengono moltiplicando le frequenze genotipiche iniziali per i rispettivi valori di fitness: $p^2 \times 1 + 2pq \times 1 + q^2 \times (1-s)$. Dato che $p^2 + 2pq + q^2 = 1$, la fitness media della popolazione in seguito alla selezione è uguale a:

$$\bar{w} = 1 - sq^2$$

Le frequenze genotipiche relative dopo l'azione della selezione si ottengono dividendo le frequenze iniziali per la fitness media:

$$\begin{aligned} f(AA) &= p^2 / (1 - sq^2) \\ f(Aa) &= 2pq / (1 - sq^2) \\ f(aa) &= q^2(1-s) / (1 - sq^2) \end{aligned}$$

Tab. 11.12 – Selezione contro l'omozigote recessivo.

Genotipo	Frequenza	Fitness
AA	p^2	1
Aa	$2pq$	1
aa	q^2	$1-s$

La frequenza di a nella generazione successiva a quella di azione della selezione (q') si ottiene quindi sommando la frequenza degli omozigoti aa a metà della frequenza degli eterozigoti. Sviluppando si ha che:

$$q' = (q - sq^2) / (1 - sq^2)$$

Pertanto, il cambiamento della frequenza allelica di a dovuta alla selezione è la seguente:

$$\Delta q = q' - q = [-sq^2(1-q)] / (1-sq^2) = -spq^2 / (1-sq^2)$$

Dato che il coefficiente di selezione così come le frequenze alleliche sono sempre positivi, o al limite nulli, Δq è sempre negativo oppure uguale a zero. Quando $\Delta q = 0$ non avviene più alcuna variazione delle frequenze alleliche. Comunque, l'efficacia della selezione contro un carattere recessivo dipende non soltanto dalla fitness, ma anche dalle frequenze dell'allele recessivo e quindi dalle proporzioni dei genotipi eterozigoti e omozigoti recessivi nella popolazione. Infatti, quando la frequenza di un gene recessivo è relativamente alta, nella popolazione saranno presenti molti omozigoti recessivi che causano cambiamenti consistenti della frequenza dell'allele recessivo. Quando invece la frequenza di un gene recessivo è bassa, nella popolazione gli omozigoti recessivi saranno rari e i cambiamenti della frequenza dell'allele recessivo saranno lievi.

Una formula che consente di misurare il progressivo declino della frequenza di un allele recessivo letale è quella di Li:

$$q_{n+1} = \frac{q_n}{(1 + q_n)}$$

dove q_n rappresenta la frequenza genica alla generazione n e q_{n+1} è la frequenza alla generazione successiva. Considerando un allele recessivo raro, come ad esempio quello che determina l'albinismo nel mais e che rappresenta un carattere letale allo stadio di plantula, la formula di Li consente di calcolare quante generazioni sono richieste per dimezzare la frequenza di questi fenotipi in una popolazione naturale. Assumendo per il genotipo omozigote recessivo una frequenza q^2 pari a $1/10.000$, la frequenza genica

sarà: $q = \sqrt{1/10.000} = 0,01$. In questo caso sono necessarie 100 generazioni, cioè 100 anni per portare la frequenza genica da tale valore a 0,005. Ciò dimostra quanto sia limitata l'efficacia della selezione nel caso di alleli rari.

La selezione naturale non ha sempre un effetto direzionale a carico delle frequenze alleliche associato ad una diminuzione della variabilità genetica. In alcuni casi, infatti, la selezione naturale agisce favorendo il mantenimento della variabilità genetica: la forma più conosciuta di selezione naturale con effetto stabilizzante sulla popolazione è quella dovuta al vantaggio dell'eterozigote. Quando l'eterozigote è favorito rispetto a tutti e due gli omozigoti, entrambi gli alleli vengono mantenuti nella popolazione. Il modello relativo prevede che la fitness dei genotipi Aa sia pari a $w=1$, quella dei genotipi AA uguale a $w=1-s_1$ e quella dei genotipi aa uguale a $w=1-s_2$ (**Tab. 11.13**). La selezione contro il genotipo AA è pertanto diversa da quella contro aa e la fitness dei due genotipi omozigoti è inferiore a quella degli eterozigoti Aa . In una popolazione costituita da genotipi AA , Aa e aa aventi frequenze pari rispettivamente a p^2 , $2pq$ e q^2 , le frequenze genotipiche dopo l'azione della selezione sono pertanto date dal prodotto delle frequenze genotipiche iniziali per i rispettivi valori di fitness: $p^2 \times (1-s_1) + 2pq \times 1 + q^2 \times (1-s_2)$. La fitness media della popolazione in seguito alla selezione è data pertanto da:

$$\bar{w} = 1 - s_1 p^2 - s_2 q^2$$

mentre la frequenza di a nella generazione successiva è:

$$q' = (q - s_2 q^2) / (1 - s_1 p^2 - s_2 q^2)$$

Tab. 11.13 – Selezione a favore dell'eterozigote.

Genotipo	Frequenza	Fitness
AA	p^2	$1 - s_1$
Aa	$2pq$	1
aa	q^2	$1 - s_2$

Tab. 11.14 – Formule per il calcolo della variazione della frequenza allelica dopo una generazione di selezione.

Tipo di selezione	Fitness dei genotipi A^1A^1 A^1A^2 A^2A^2	Fitness media della popolazione	Variazione della frequenza genica
Selezione contro l'omozigote recessivo	1 1 1-s	$\bar{W}=p^2W_{11}+2pqW_{12}+q^2(1-s)W_{22}$	$= \frac{-spq^2}{1-sq^2}$
Selezione contro un allele dominante	1-s 1-s 1	$\bar{W}=p^2(1-s)W_{11}+2pq(1-s)W_{12}+q^2W_{22}$	$= \frac{-spq^2}{1-s+sq^2}$
Selezione in assenza di dominanza	1 (1-s/2) 1-s	$\bar{W}=p^2W_{11}+2pq(1-s/2)W_{12}+q^2(1-s)W_{22}$	$= \frac{-spq^2}{1-sq}$
Selezione a favore dell'eterozigote	1-s ₁ 1 1-s ₂	$\bar{W}=p^2(1-s_1)W_{11}+2pqW_{12}+q^2(1-s_2)W_{22}$	$= \frac{pq(s_1p-s_2q)}{1-s_1p^2-s_2q^2}$
Selezione contro l'eterozigote	1 1-s 1	$\bar{W}=p^2W_{11}+2pq(1-s)W_{12}+q^2W_{22}$	$= \frac{spq(q-p)}{1-2spq}$
Formula generale	W_{11} W_{12} W_{22}	$\bar{W}=p^2W_{11}+2pqW_{12}+q^2W_{22}$	$= \frac{pq[p(W_{12}-W_{11})+q(W_{22}-W_{12})]}{\bar{W}}$

Il cambiamento della frequenza allelica di a dovuto alla selezione è quindi il seguente:

$$\Delta q = q' - q = pq(s_1p - s_2q) / (1 - s_1p^2 - s_2q^2)$$

Le frequenze alleliche cambieranno come risultato della selezione fino a che non verrà raggiunto nuovamente l'equilibrio. Il valore di Δq si annulla quando $s_1p = s_2q$. Si può dimostrare per via algebrica che questo avviene quando: $f(A) = p = s_2 / (s_1 + s_2)$ e $f(a) = q = s_1 / (s_1 + s_2)$. Se la selezione contro i due omozigoti è ugualmente efficace, cioè $s_1 = s_2$, allora le frequenze alleliche all'equilibrio sono 0,5 e la popolazione conterrà il 50% di eterozigoti. Se la selezione contro i due omozigoti è invece asimmetrica, le frequenze alleliche all'equilibrio si sposteranno verso quella dell'allele più adattato. In ogni caso, la selezione in favore dell'eterozigote non consente la fissazione del genotipo più adattato.

I modelli e le formule per il calcolo della fitness media della popolazione e della variazione della frequenza allelica dopo una generazione di selezione e secondo diversi tipi di selezione sono riassunte nella **Tab. 11.14**.

Un caso particolare di selezione è quello contro un carattere dominante letale. Qualora la dominanza sia completa e l'allele dominante sia letale per l'individuo o renda l'individuo sterile ($s=1$), la selezione naturale produce il suo effetto massimo in una sola generazione determinando la fissazione dei genotipi omozigoti recessivi e portando alla scomparsa dell'allele dominante dalla popolazione. Il cambiamento della frequenza genica è $\Delta q = 1 - q = p$ cioè q diventa uguale a 1.

Dal punto di vista molecolare, gli effetti della selezione naturale sono connessi all'esistenza nelle popolazioni di variabilità genetica riconducibile a polimorfismi a carico delle sequenze nucleotidiche degli alleli. Alleli distinti possono codificare per proteine aventi diverse sequenze amminoacidiche capaci di condizionare in positivo o in negativo la sopravvivenza e la riproduzione di un individuo. Gli individui della popolazione con alleli vantaggiosi che codificano per proteine più efficienti hanno una maggiore probabilità di sopravvivere e una maggiore capacità di riprodursi, contribuendo così in modo più consistente al pool genico della generazione successiva. Nel corso di molte generazioni le frequenze alleliche e le caratteristiche fenotipiche di una popolazione possono pertanto modificarsi per il solo effetto della selezione naturale. In questo modo la selezione naturale rende massimo l'adattamento e il successo riproduttivo di una popolazione in un determinato ambiente.

11.6.4.5 Unioni non casuali tra individui geneticamente simili: inbreeding

La legge di Hardy-Weinberg assume che le unioni tra genotipi avvengano con la stessa proporzione dei genotipi nella popolazione: la probabilità che due genotipi si incrocino tra loro deve essere quindi uguale al prodotto delle loro frequenze genotipiche.

Qualora individui geneticamente simili o dissimili abbiano la tendenza ad unirsi tra loro più spesso di quanto voluto dal caso l'equilibrio viene disturbato. In particolare, le unioni tra individui geneticamente dissimili (*outbreeding*) portano ad un mantenimento o addirittura ad un aumento della variabilità genetica della popolazione. Nei casi di esoincrocio, infatti, aumenta la probabilità che le progenie ereditino geni differenti dai due individui parentali determinando così lo stato eterozigote in un certo numero di loci. Le unioni tra individui geneticamente simili (*inbreeding*) hanno, invece, conseguenze opposte poiché determinano una diminuzione della variabilità genetica della popolazione. In questi casi l'inincrocio aumenta, infatti, la probabilità che le progenie ereditino gli stessi geni da entrambi gli individui parentali portando così ad una loro fissazione allo stato omozigote in un certo numero di loci. Nelle loro applicazioni più estreme, l'esoincrocio e l'inincrocio controllati rappresentano strategie molto utili per sviluppare linee pure omozigoti a tutti i loci o per creare ibridi eterozigoti a molti loci. Nelle popolazioni naturali, invece, tali modalità di unione tra piante portano ad una alterazione della composizione genotipica e ad un cambiamento delle proporzioni tra genotipi rispetto a quanto previsto dall'equilibrio di Hardy-Weinberg.

Particolarmente importante per le conseguenze genetiche che determina e le ripercussioni che ha sui metodi di miglioramento genetico e di costituzione varietale è l'*inbreeding*. L'*inbreeding* comprende le situazioni relative alle popolazioni di specie allogame per le quali le unioni tra individui geneticamente affini e strettamente imparentati avvengono con una frequenza maggiore di quanto voluto dal caso. La forma più estrema di *inbreeding* è ovviamente l'autofecondazione che è possibile unicamente nelle piante ermafrodite e che costituisce la forma di riproduzione naturale delle specie prevalentemente autogame. Le popolazioni di specie prevalentemente allogame, invece, non tollerano o tollerano male l'*inbreeding*. In linea generale, la fecondazione incrociata tra individui imparentati non provoca un cambiamento delle frequenze geniche,

Tab. 11.15 – Composizione genotipica e distribuzione genica attesa in progenie di mais in caso di inincrocio e esoincrocio.

Combinazioni genotipiche	Tipo di incrocio	Progenie			
<i>Inbreeding</i>					
I × I	Cc × Cc	1CC	2Cc	1cc	
I × I	Cc × Cc	1CC	2Cc	1cc	
II × II	CC × CC	4CC			
II × II	CC × CC	4CC			
III × II	cc × cc				4cc
III × III	cc × cc				4cc
Composizione dei genotipi		10CC	4Cc	10cc	
Proporzione dei geni				24C	24c
<i>Outbreeding</i>					
I × II	Cc × CC	2CC	2Cc		
II × I	CC × Cc	2CC	2Cc		
II × III	CC × cc				4Cc
III × II	cc × CC				4Cc
I × III	Cc × cc	2Cc	2cc		
III × I	cc × Cc	2Cc	2cc		
Composizione dei genotipi		4CC	16Cc	4cc	
Proporzione dei geni				24C	24c

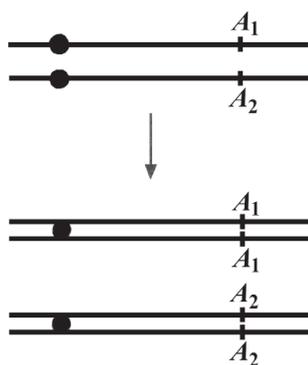


Fig. 11.28 – Origine degli alleli identici.

qualora tali unioni non siano accompagnate da selezione, ma determina un aumento delle frequenze dei genotipi omozigoti. Uno degli effetti più vistosi dell'*inbreeding* è proprio connesso alla comparsa di fenotipi recessivi che risultano così esposti all'azione della selezione. A titolo di esempio, nella **Tab. 11.15** è riportata la distribuzione attesa di geni in progenie di mais in caso di inincrocio e esoincrocio quando la proporzione iniziale dei geni è la stessa nei parentali e le progenie hanno la stessa numerosità. Considerando tutte le possibili combinazioni di incrocio e i corrispondenti reciproci, in entrambi i sistemi le proporzioni dei geni al locus considerato rimangono invariate, mentre le combinazioni dei genotipi subiscono sostanziali cambiamenti a favore di quelli omozigoti (in caso di *inbreeding*) oppure di quelli eterozigoti (in caso di *outbreeding*). In termini generali, in assenza di processi evolutivi, il sistema di unione tra individui non ha ripercussioni sulle frequenze geniche quanto invece sulle frequenze genotipiche, condizionando di conseguenza i rapporti fenotipici della popolazione. Qualora i fenotipi non abbiano la stessa fitness allora l'ambiente è in grado di esercitare una azione selettiva determinando così variazioni delle frequenze geniche nel corso di generazioni successive.

Tra il 1910 e il 1930, Wright e Fisher elaborarono metodi per quantificare il grado di *inbreeding* che, come effetto a livello di un singolo individuo, si traduce in un aumento della probabilità che l'individuo sia omozigote per **alleli identici**, cioè per alleli aventi la stessa origine e risultanti quindi dalla replicazione di uno stesso allele inizialmente presente in un individuo (**Fig. 11.28**). La separazione degli alleli identici si verifica durante la meiosi con la segregazione dei cromosomi omologhi e la conseguente formazione dei gameti. Una volta che tali gameti originano per fecondazione nuovi zigoti, gli alleli identici andranno a finire in individui diversi della stessa discendenza e attraverso un sistema di unioni casuali sarà molto improbabile che possano ritrovarsi insieme nello stesso individuo. Qualora invece le unioni non siano casuali, ma coinvolgano individui imparentati è possibile che alleli identici portati da gameti diversi vadano a finire insieme nello stesso zigote e tale evento sarà tanto più probabile quanto più alto è il grado di affinità tra gli individui. Il **coefficiente di *inbreeding* (F)** è stato assunto come misura quantitativa del grado di affinità genealogica ed equivale alla probabilità che un individuo possieda alleli identici. Tale coefficiente può anche essere chiamato coefficiente di fissazione poiché corrisponde alla probabilità che un allele venga fissato in condizione omozigote dato che gli individui con tale costituzione genotipica sono capaci di trasmettere soltanto un tipo di allele alla loro discendenza.

Si possono considerare alcuni casi a titolo di esempio. Prendendo in esame lo schema in **Fig. 11.29A** relativo ad un individuo III-1 derivante dall'unione tra individui (II-1 e II-2) di una stessa discendenza, per calcolare il suo coefficiente di *inbreeding* è innanzitutto necessario identificare gli ascendenti comuni (individui I-1 e I-2). La probabilità che gli individui II-1 e II-2 ereditino alleli identici è pari ad $1/2$; ipotizzando infatti che gli individui I-1 e I-2 siano eterozigoti A_1A_2 e A_3A_4 , gli individui II-1 e II-2 potranno ricevere o l'uno o l'altro allele al locus in esame, ma non entrambi. Quindi la probabilità che II-1 erediti l'allele A_1 quando anche II-2 eredita l'allele A_1 è pari a $1/4 = (1/2 \times 1/2)$ e che II-1 erediti l'allele A_2 quando anche II-2 eredita l'allele A_2 è pari a $1/4 = (1/2 \times 1/2)$. Analogamente, la probabilità che II-1 erediti l'allele A_3 quando anche II-2 eredita l'allele A_3 è pari a $1/4 = (1/2 \times 1/2)$ e che II-1 erediti l'allele A_4 quando anche II-2 eredita l'allele A_4 è pari a $1/4 = (1/2 \times 1/2)$. La probabilità che II-1 e II-2 ricevano alleli identici è quindi $1/2 = (1/4 + 1/4)$, mentre la probabilità complessiva che III-1 riceva alleli identici da I-1 è pari ad $1/8$, così come quella che III-1 riceva alleli identici da I-2 è pari ad $1/8$. Cioè la probabilità che un allele particolare passi da I-1 a III-1 sia attraverso II-1 che attraverso II-2 così come la probabilità che III-1 riceva alleli identici da I-2 è uguale a $1/2 \times 1/2 \times 1/2$. Per ognuno dei due ascendenti comuni si può quindi calcolare che: i) $1/2$ è la probabilità che II-1 e II-2 ereditino alleli identici; ii) $1/2$ è la probabilità

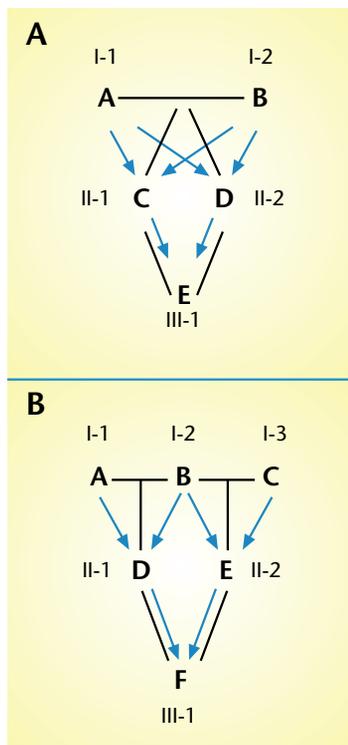


Fig. 11.29 – Genealogie di discendenze full-sib (A) e half-sib (B).

che II-1 trasmetta a III-1 l'allele in questione; iii) $1/2$ è la probabilità che II-2 trasmetta a III-1 lo stesso allele. Ammettendo che i due ascendenti comuni dell'individuo in questione non siano imparentati, il coefficiente di *inbreeding* di III-1 diventa $1/8+1/8=1/4$, cioè $F=0,25$.

In questo caso, ciascuno dei geni dell'individuo derivante dall'incrocio tra due piante della stessa progenie (incrocio tra fratello e sorella, detto *full-sib*) ha una probabilità del 25% di essere omozigote per avere ereditato lo stesso allele da uno dei due parentali iniziali. Analogamente, si può dimostrare che un individuo derivante dall'incrocio tra piante appartenenti a due progenie che hanno un parentale comune (incrocio di mezzi-fratelli con mezze-sorelle, detto *half-sib*) ha una probabilità del 12,5% di essere omozigote per avere ereditato lo stesso allele dal parentale comune iniziale (**Fig. 11.29B**).

Il coefficiente di *inbreeding* di un individuo i i cui parentali abbiano N ascendenti comuni e il cui percorso di inincrocio includa n passaggi, cioè coinvolga $n+1$ individui, può essere calcolato usando la seguente formula:

$$F_i = \sum_1^N (1/2)^{n+1} (1+F_a)$$

dove F_a è il coefficiente di inbreeding dell'ascendente comune. Un ascendente (o progenitore) comune è rappresentato da qualsiasi individuo che si incontra nella linea ascendente di entrambi i parentali (o genitori) dell'individuo in questione, mentre un percorso di un inincrocio è dato dai passaggi che collegano entrambi i parentali dell'individuo in questione all'ascendente o agli ascendenti comuni attraverso la via più breve dell'albero genealogico. Il numero di passaggi più uno coincide in realtà con il numero di individui coinvolti nel percorso di inincrocio, escluso l'individuo in esame. F varia tra 0 e 1: è nullo per gli individui delle popolazioni numerose con unioni perfettamente casuali, mentre è massimo quando le unioni sono tali che un individuo riceverà certamente due alleli identici.

Nelle popolazioni naturali di specie vegetali, così come in quelle umane ed animali, la condizione di accoppiamento casuale, necessaria affinché possa stabilirsi un equilibrio Hardy-Weinberg, viene violata frequentemente. Le conseguenze degli accoppiamenti tra individui geneticamente affini riguardo ad un carattere fenotipico possono essere analizzate ricostruendo in maniera appropriata alberi genealogici. Un **albero genealogico** consiste in una rappresentazione grafica del succedersi delle generazioni di una famiglia che eventualmente contiene le indicazioni circa la manifestazione di un carattere particolare e che serve a determinare se il gene responsabile di questo carattere è trasmesso secondo un particolare modello di eredità (**Fig. 11.30**). Negli alberi genealogici, le femmine si rappresentano con un cerchio mentre i maschi si identificano con un quadrato. Il colore bianco contraddistingue i fenotipi normali, mentre quello nero gli individui che manifestano un particolare carattere o che sono

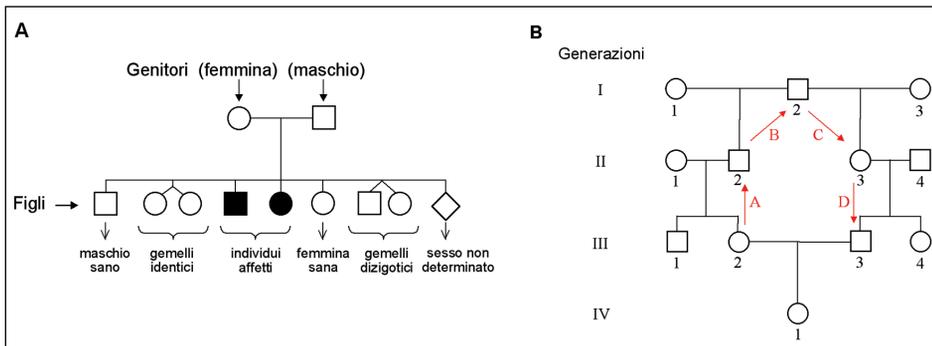


Fig. 11.30 – Significato dei simboli principali di un albero genealogico (A). Esempio di albero genealogico (*pedigree*) animale con una unione consanguinea (B).

affetti da una determinata malattia. I genitori vengono collegati tra di loro da una linea orizzontale a sua volta collegata, mediante un tratto verticale, ad una linea orizzontale sottostante alla quale sono collegati con altri tratti verticali tutti i figli derivanti dalla loro unione in ordine di nascita, da sinistra a destra. Le generazioni sono indicate con numeri romani.

Prendendo in esame lo schema di Fig. 11.30B relativo ad un individuo IV-1, per il quale nella linea ascendente compare un solo progenitore comune, rappresentato dall'individuo I-2, mentre la lunghezza del percorso di inincrocio che collega i genitori dell'individuo in questione al progenitore comune include quattro passaggi (A-B-C-D) o cinque individui (III-2, II-2, I-2, II-3, III-3). Supponendo che il coefficiente di inbreeding dell'ascendente comune sia nullo, nell'esempio considerato $F=(1/2)^5(1+0)=3,125\%$. Qualora l'esame dell'albero genealogico evidenzi due o più ascendenti comuni, il coefficiente di inbreeding dell'individuo in questione si calcola come somma di tutti gli F ottenibili considerando singolarmente i diversi ascendenti e i possibili percorsi di inincrocio che un allele può seguire per arrivare all'individuo stesso.

L'inbreeding nelle specie allogame può avere diverse conseguenze in una popolazione naturale: provoca un aumento generalizzato della frequenza degli omozigoti e quindi una diminuzione della variabilità genetica complessiva, ma anche una riduzione della capacità adattativa media di una popolazione in quanto oltre alla comparsa di omozigoti per alleli già manifesti si ha anche la comparsa di recessivi per alleli sfavorevoli o letali. Di fatto, in presenza di inbreeding le frequenze genotipiche di una popolazione cambiano rispetto a quanto voluto dall'equilibrio Hardy-Weinberg. Prendendo in esame un singolo locus con due alleli A e a , l'unione tra individui imparentati all'interno di una popolazione può determinare la formazione di genotipi omozigoti per alleli identici. Al locus considerato, la frazione di questi omozigoti nell'ambito della popolazione è proporzionale al coefficiente di *inbreeding* degli individui che compongono la popolazione stessa. Quindi le frazioni di omozigoti per alleli identici presenti nell'intera popolazione sono: pF per AA e qF per aa , mentre la frazione rimanente dei genotipi $(1-F)$ sono distribuiti tra AA , Aa e aa in accordo con la legge di Hardy-Weinberg. Le frequenze genotipiche di una popolazione soggetta ad *inbreeding* saranno pertanto:

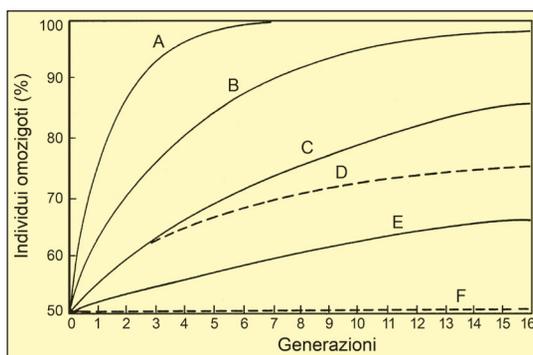


Fig. 11.31 – Variazione della percentuale di omozigoti in caso di autofecondazione e in vari sistemi di incrocio tra individui imparentati, ipotizzando una frequenza iniziale di eterozigoti del 50%: (A) autofecondazione; (B) incrocio fratello-sorella; (C) doppi primi cugini; (D) fratello-sorella con un solo genitore comune; (E) primi cugini; (F) secondi cugini.

$$\begin{aligned} f(AA) &= p^2(1-F) + pF = p^2 + pqF \\ f(Aa) &= 2pq(1-F) = 2pq - 2pqF \\ f(aa) &= q^2(1-F) + qF = q^2 + pqF \end{aligned}$$

Nella popolazione si ha pertanto un aumento della frequenza degli omozigoti misurato dal prodotto delle frequenze geniche e del coefficiente di inbreeding (pqF), mentre la riduzione della frequenza dei genotipi eterozigoti non può che essere uguale al doppio di questo prodotto ($2pqF$). L'espressione seguente:

$$(p^2 + pqF)AA + (2pq - 2pqF)Aa + (q^2 + pqF)aa = 1$$

rappresenta la **formula generale dell'equilibrio di Sewall Wright**. In presenza di unioni casuali, cioè quando $F=0$, si ritorna alla formula dell'equilibrio di Hardy-Weinberg, mentre il massimo dell'effetto *inbreeding* si ha quando $F=1$ poiché la popolazione risulta costituita unicamente da omozigoti, così come accade nelle specie autogame.

I cambiamenti nelle frequenze genotipiche delle popolazioni in funzione del grado di *inbreeding* sono riportati nella **Fig. 11.31**. Il grafico mette in evidenza l'incremento della percentuale di omozigoti in caso di autofecondazione e in vari

Tipo di unione	F
Autofecondazione	$1/2$ (0,5)
Incroci	
Genitore-progenie (<i>parent-offspring</i>)	$1/4$ (0,25)
Fratello-sorella (<i>full-sibs</i>)	$1/4$ (0,25)
Mezzi-fratelli (<i>half-sibs</i>)	$1/8$ (0,125)
Zio/a-nipote	$1/8$ (0,125)
Primi cugini	$1/16$ (0,625)
Mezzi primi cugini (terzo grado)	$1/32$ (0,03125)
Secondi cugini	$1/64$ (0,015625)

Tab. 11.16 – Valori dei coefficienti di inbreeding (F) per alcuni dei più comuni sistemi di unione.

sistemi di incrocio tra individui imparentati (fratello con sorella, zio/zia con nipote, fratello-sorella con un solo genitore comune, ecc.) ipotizzando una frequenza iniziale di eterozigoti del 50%.

È evidente che l'autofecondazione negli organismi ermafroditi, come buona parte delle piante, costituisca il sistema in grado di condurre più rapidamente all'omozigosi: $F=0,5$ significa che ad ogni locus, la metà della discendenza prodotta per autofecondazione sarà teoricamente omozigote per un gene parentale. In altri termini, la quota di eterozigoti si dimezza ad ogni ciclo di autofecondazione. Così dopo otto generazioni la quasi totalità dei geni che erano inizialmente in condizione eterozigote risulteranno allo stato omozigote. Anche gli incroci fratello-sorella determinano un consistente incremento della quota di omozigoti, benché secondo un progresso meno rapido. Con questo tipo di unioni, molto impiegato anche nel miglioramento genetico degli animali, dopo undici generazioni il 95% dei geni che erano inizialmente in condizione eterozigote raggiungono, in teoria, lo stato omozigote e solo il 5% permangono allo stato eterozigote. Quando il grado di parentela si riduce, l'incremento progressivo della percentuale di omozigosi diviene più lento ed il numero di generazioni richieste per raggiungere il limite del 100% può essere anche molto elevato. Tuttavia, ognuno dei sistemi di inbreeding considerati, fino agli incroci tra primi cugini, nel lungo periodo conduce alla piena omozigosi. Nel caso dei secondi cugini, invece, gli incroci tra questi possono portare ad un limite massimo teorico di omozigosi del 51% (Fig. 11.31).

Nella **Tab. 11.16** sono riportati i valori dei coefficienti di inbreeding (F) per alcuni dei più importanti sistemi di unione utilizzati nel miglioramento genetico delle specie di interesse agrario e zootecnico. Si può notare che nel caso degli incroci, il valore di F dipende dal grado della relazione parenterale esistente tra gli individui usati come genitori maschile e femminile. In particolare, minore è la vicinanza di tali individui rispetto ad un ascendente comune e più bassa è la probabilità che abbiano ereditato alleli identici da questo.

11.7 Dimensione effettiva della popolazione in relazione a deriva genetica e inbreeding

Il modello di riferimento definito dall'equilibrio di Hardy-Weinberg presuppone, tra gli altri, che la popolazione sia costituita da una comunità molto grande ed isolata, in cui gli individui possano riprodursi completamente a caso. L'incrocio con individui di altre popolazioni (migrazione) comporta variazioni nella costituzione genetica della

popolazione le cui intensità dipendono dal tasso di interincrocio e dalle differenze genetiche tra le popolazioni interessate.

Particolarmente interessanti per l'evoluzione, in generale, e per il miglioramento genetico, in particolare, sono gli effetti genetici determinati dalle dimensioni limitate della popolazione. In queste condizioni sono possibili variazioni delle frequenze geniche e genotipiche, indipendentemente dalla selezione e dalla mutazione. Infatti, in una popolazione composta di pochi individui solo un piccolo campione di tutti i gameti prodotti dai genitori può venire impiegato per la costituzione della generazione filiale ed è pertanto verosimile che notevoli variazioni nella composizione genetica possano verificarsi per il solo effetto del caso. Ad esempio, in una popolazione composta di soli otto individui, in cui gli alleli A e a abbiano frequenze $p = q = 0,5$, in condizione di perfetto equilibrio sono attesi due genotipi AA , quattro Aa e due aa . Questo però richiede che nella generazione dei genitori quattro gameti A e altrettanti gameti a prodotti dagli individui maschili si combinino equamente con quattro gameti A ed altrettanti a prodotti dagli individui femminili. Per il solo effetto del caso si potrebbero avere, per esempio, quattro genotipi AA , tre Aa e uno aa con conseguente spostamento delle frequenze geniche ai valori $p = 0,69$ e $q = 0,31$. Una variazione delle frequenze geniche della stessa entità è invece molto meno probabile che avvenga in una popolazione composta da centinaia di individui. Ad un dato locus le variazioni delle frequenze alleliche per effetto del campionamento possono verificarsi sia in senso positivo (aumento della frequenza genica) che in senso negativo (diminuzione della frequenza genica), in modo del tutto imprevedibile. Per questo motivo il fenomeno viene indicato col termine di deriva genetica.

Il fenomeno di deriva genetica si ripete nel corso delle generazioni e gli effetti si sommano fino alla condizione limite della completa eliminazione di un allele. Qualora, ad esempio, in una data generazione i gameti con l'allele A vengano completamente esclusi dalla riproduzione, allora è l'altro allele a allo stesso locus a venire fissato ($p = 0$ e $q = 1$). Data la natura casuale del fenomeno, gli effetti riguardanti la variazione delle frequenze geniche non sono prevedibili in termini di direzione (incremento o decremento). Viceversa è possibile determinare l'entità delle variazioni e i tempi medi necessari affinché si realizzi la fissazione.

Il parametro necessario per poter formulare queste previsioni è la dimensione effettiva (N) della popolazione, cioè il numero di individui parentali che daranno origine alla generazione filiale. I gameti complessivamente prodotti dagli N individui di una popolazione formano un campione di dimensione molto grande, indipendentemente dal numero di parentali considerati. Tuttavia, solo un campione di $2N$ gameti (N maschili e N femminili) corrisponde a quello che dà effettivamente origine alla generazione filiale. Assumendo che la dimensione della popolazione (N) resti pressoché costante nel corso delle generazioni, la frequenza genica nel campione può variare casualmente da 0 a 1 con una distribuzione di probabilità che ricalca con buona approssimazione quella normale. L'ampiezza della distribuzione della frequenza genica (p) è definita dalla deviazione standard:

$$\sigma = \sqrt{\frac{pq}{2N}}$$

Ciò significa che nel 95% dei casi il valore delle frequenze geniche del campione oscilla tra $p-2\sigma$ e $p+2\sigma$ e solo raramente (5%) supera questi limiti. L'entità di queste variazioni è pertanto definita dalla deviazione standard che a sua volta dipende in misura inversa dalla dimensione del campione ($2N$). Per esempio, considerando popolazioni di dimensioni diverse ($N = 1000-5$ individui), ma con le stesse frequenze geniche ($p = q = 0,5$), i valori di deviazione standard e gli intervalli sono quelli riportati in **Tab. 11.17**. L'intervallo entro cui sono compresi i probabili

Tab. 11.17 – Intervallo di variazione dei probabili valori campionari delle frequenze geniche con dimensioni diverse della popolazione, assumendo $p=q=0,5$.

N	σ	Intervallo	
		$p-2\sigma$	$p+2\sigma$
1000	0,0112	0,4776	0,5224
500	0,0158	0,4684	0,5316
100	0,0354	0,4292	0,5708
50	0,0500	0,4000	0,6000
10	0,1118	0,2764	0,7236
5	0,1581	0,1838	0,8162

valori campionari aumenta con la riduzione della popolazione e, conseguentemente, gli effetti della deriva genetica sono molto più marcati quando la dimensione della popolazione è limitata.

L'effetto della riduzione della dimensione della popolazione può essere studiato anche in relazione al conseguente aumento del grado di parentela tra gli individui che la compongono. Se la popolazione è costituita da un numero limitato di individui è inevitabile che dopo alcune generazioni questi risultino tra loro imparentati, con grado variabile non solo in relazione alla dimensione della popolazione, ma anche in funzione del numero di generazioni trascorse. Gli antenati di un individuo sono due nella generazione precedente alla propria, quattro in quella dei nonni e così via. Più in generale, risalendo per un numero t di generazioni, il numero degli antenati è pari a 2^t . Di fatto, risultano necessarie poche generazioni affinché il numero degli antenati sia pari alla dimensione della popolazione: in una data generazione trascorsa, tutti gli individui hanno gli stessi antenati comuni e sono pertanto imparentati. In una situazione di questo tipo è inevitabile che l'incrocio benché casuale coinvolga individui imparentati (*inbreeding*), determinando così la produzione di individui omozigoti in una quota maggiore rispetto a quella prevista dall'equilibrio. L'incremento degli omozigoti dipende dal fatto che gli individui imparentati incrociandosi tra loro hanno una certa probabilità di trasmettere alla discendenza alleli identici derivanti dalla replicazione di uno stesso gene presente nell'antenato comune. La quota di individui omozigoti per discendenza (omozigoti per alleli identici) si somma quindi a quella degli omozigoti per alleli di origine indipendente (omozigoti per alleli simili). Una misura quantitativa dell'inbreeding è data dal coefficiente di inbreeding (F), che corrisponde alla probabilità che, ad un dato locus, un individuo sia omozigote per alleli identici. Questa probabilità può essere calcolata in base alla dimensione effettiva della popolazione, con la quale risulta avere una relazione inversa. Considerando una popolazione di N genitori non imparentati, la popolazione gametica è costituita da $2N$ tipi di gameti diversi come origine ($A_1, A_2, A_3, A_4, \dots, A_{2N}$), ciascuno rappresentato da tante copie. Se l'incrocio avviene completamente a caso, un particolare gamete ha la probabilità di $1/2N$ di unirsi con uno identico. Il coefficiente di inbreeding alla prima generazione è allora uguale a $F = 1/2N$. Nella generazione successiva possono originarsi ancora $1/2N$ di zigoti identici al locus considerato, a cui bisogna aggiungere quelli della generazione precedente: i rimanenti $1-1/2N$ zigoti hanno infatti una probabilità pari a F di produrre omozigoti per alleli identici. Per cui $F_2 = 1/2N + (1-1/2N)F$. In generale alla generazione t si ha:

$$F_t = \frac{1}{2N} + \left(1 - \frac{1}{2N}\right)F_{t-1}$$

dove $1/2N$ è l'incremento di inbreeding in ogni generazione. Il valore di F ad una data generazione indica l'incremento relativo di omozigosi e, conseguentemente, il decremento percentuale di eterozigosi rispetto alla quota attesa con incrocio completamente casuale, equivalente alla condizione di panmissia. Per esempio, se $F = 0,2$, in una popolazione con frequenze geniche $p(A) = 0,6$ e $q(a) = 0,4$ per la quale la quota di eterozigoti attesa è pari a $2pq = 0,48$, l'inbreeding determina una riduzione del 20% degli eterozigoti (Aa) previsti dal momento che $2pq - 2pqF = 0,48 - 0,2(0,48) = 0,38$. Per $F = 1$, invece, la popolazione è composta unicamente da omozigoti:

$f_{AA} = p^2 + pqF = p^2 + pq = p^2 + p(1-p) = p$ così come $f_{aa} = q^2 + pqF = q^2 + pq = q^2 + (1-q)q = q$. Nella Fig. 11.32 è rappresentata la variazione della quota di eterozigoti nel corso delle generazioni in popolazioni di dimensioni (N) diverse.

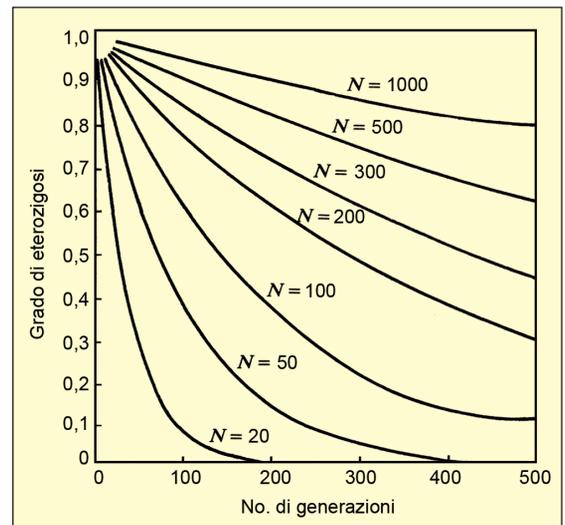


Fig. 11.32 - Variazione del grado di eterozigosi nel corso delle generazioni in popolazioni di dimensioni (N) diverse.

In conclusione, la riduzione della dimensione della popolazione porta a variazioni casuali delle frequenze geniche, con possibilità di fissazione di uno degli alleli e alla riduzione del grado di eterozigosi. L'intensità di entrambi gli effetti è legata alla dimensione effettiva della popolazione.

11.8 Principio di Hardy-Weinberg con due geni

Tab. 11.18 – Associazione casuale tra gli alleli a due loci e frequenze dei tipi gametici quando i loci sono in equilibrio.

		Alleli al locus 1	
		(p)A ₁	(q)A ₂
Alleli al locus 2	(u)B ₁	pu (A ₁ B ₁)	qu (A ₂ B ₁)
	(v)B ₂	pυ (A ₁ B ₂)	qv (A ₂ B ₂)

In condizioni di accoppiamento casuale, gli alleli a un qualsiasi locus raggiungono rapidamente l'associazione casuale nei genotipi. Considerando due loci con due alleli ciascuno, A₁ e A₂, aventi frequenze p e q, e B₁ e B₂, aventi frequenze u e v, le frequenze genotipiche al primo locus saranno p² per A₁A₁, 2pq per A₁A₂ e q² per A₂A₂, mentre le frequenze genotipiche al secondo locus saranno u² per B₁B₁, 2uv per B₁B₂ e v² per B₂B₂. Tuttavia, nei gameti i geni ad un locus possono non essere in associazione casuale con i geni all'altro locus. A un dato momento, comunque, gli alleli dei due geni raggiungeranno nei gameti l'associazione casuale, ma l'avvicinamento è graduale e la velocità di avvicinamento è tanto più bassa quanto più piccola è la frazione di ricombinazione tra i due geni. Il significato di "associazione casuale nei gameti" è rappresentato in **Tab. 11.18** dove la combinazione degli alleli ai due loci si riferisce ai gameti e non ai genotipi. Poiché p(A₁)+q(A₂)=1 così come u(B₁)+v(B₂)=1, ne consegue che p+q=u+v. Una situazione di associazione gametica casuale tra alleli di due geni diversi è detta **linkage equilibrium** (equilibrio delle fasi gametiche).

Per descrivere l'effetto che la ricombinazione svolge sull'avvicinamento al linkage equilibrium si può considerare, ad esempio, la situazione di equivalenza delle frequenze alleliche: p=q=0,5 e u=v=0,5. Se le due coppie alleliche sono indipendenti, all'equilibrio i quattro tipi di gameti si formeranno con uguale frequenza, cioè 0,25 (25%) per ciascuno dei tipi gametici A₁B₁, A₁B₂, A₂B₁ e A₂B₂, mentre le sedici possibili combinazioni genotipiche saranno presenti con frequenza pari a 0,25×0,25=0,0625 (6,25%). In sostanza, la condizione di equilibrio Hardy-Weinberg viene raggiunta quando i gameti si formano dall'assortimento casuale degli alleli e i genotipi derivano dall'unione casuale dei gameti. Qualora non siano rispettate tali situazioni e i geni non risultino in equilibrio nei gameti si parla di **linkage disequilibrium** (LD, disequilibrio di associazione). Ciò equivale alla presenza contemporanea e non casuale in una popolazione di due alleli con deviazione dall'equilibrio Hardy-Weinberg tale che la frequenza degli alleli combinati insieme sia diversa da quella derivabile semplicemente dal prodotto delle frequenze dei singoli alleli. Tale condizione può essere causata e mantenuta principalmente da fattori quali l'associazione fisica dei loci corrispondenti e la combinazione di alleli favorevoli in termini di *fitness*, ma può anche essere dovuta al mescolamento di popolazioni con frequenze geniche distinte o al campionamento gametico casuale nelle piccole popolazioni.

Il linkage disequilibrium è spesso misurato usando statistiche di associazione tra forme alleliche a coppie di loci. Considerando, ad esempio, due loci A e B ciascuno avente due alleli, A₁ e A₂, con frequenze pari a p_A e q_A, e B₁ e B₂, con frequenze pari a p_B e q_B, le frequenze aplotipiche saranno le seguenti: p_Ap_B per A₁B₁, p_Aq_B per A₁B₂, q_Ap_B per A₂B₁ e q_Aq_B per A₂B₂. Qualora la frequenza di un dato aplotipo nella popolazione sia maggiore o minore di quella attesa sulla base delle frequenze alleliche significa che due alleli tendono ad essere osservati insieme e che i due loci sono in linkage disequilibrium. Una misura comunemente usata è il parametro di linkage disequilibrium (D) definito come:

$$D = f(A_1B_1) \times f(A_2B_2) - f(A_1B_2) \times f(A_2B_1) = p_A p_B \times q_A q_B - p_A q_B \times q_A p_B$$

Genotipo	GAMETI			
	A_1B_1	A_1B_2	A_2B_1	A_2B_2
A_1B_1/A_1B_1	1	0	0	0
A_1B_1/A_1B_2	0,5	0,5	0	0
A_1B_1/A_2B_1	0,5	0	0,5	0,5
A_1B_1/A_2B_2	$(1/2)(1-r)$	$(1/2)r$	$(1/2)r$	$(1/2)(1-r)$
A_1B_2/A_1B_2	0	1	0	0
A_1B_2/A_2B_1	$(1/2)r$	$(1/2)(1-r)$	$(1/2)(1-r)$	$(1/2)r$
A_1B_2/A_2B_2	0	0,5	0	0,5
A_2B_1/A_2B_1	0	0	1	0
A_2B_1/A_2B_2	0	0	0,5	0,5
A_2B_2/A_2B_2	0	0	0	1

Tab. 11.19 – Genotipi a due loci con geni biallelici e loro frequenze gametiche (r equivale alla frazione di ricombinazione tra i loci).

La **Tab. 11.19** mostra i dieci diversi genotipi che sono possibili per due geni biallelici e per ciascun genotipo riporta le frequenze dei gameti prodotti, ipotizzando che la frequenza di ricombinazione sia uguale sia nei maschi che nelle femmine.

Il valore D per coppie di loci può essere calcolato anche ricorrendo ad altre formule: la formula di P.W. Hedrick prevede la differenza, in valore assoluto, tra la frequenza osservata dell'aplotipo A_1B_1 e il prodotto delle frequenze dei singoli alleli A_1 e B_1 (noto come coefficiente di disequilibrio standardizzato):

$$|D| = f(A_1B_1) - p_A \times p_B$$

mentre quella di B.S. Weir è un coefficiente di correlazione tra frequenze aplotipiche e geniche, uguale al rapporto tra il quadrato del coefficiente di disequilibrio standardizzato e il prodotto delle frequenze dei quattro possibili alleli ai due loci considerati:

$$r^2 = \frac{D^2}{p_A \times q_A \times p_B \times q_B}$$

Entrambi i parametri possono assumere valori compresi tra 0 (completo *linkage equilibrium*) ed 1 (completo *linkage disequilibrium*). Per il parametro D è preso in considerazione il suo valore assoluto in quanto il segno, positivo o negativo, dipende dalla scelta arbitraria della coppia di alleli ai due loci esaminati. Un'altra misura utile del disequilibrio è la seguente:

$$D' = D/D_{\max}$$

dove D_{\max} è il valore massimo che D può raggiungere con le frequenze alleliche date.

In condizione di equilibrio (*linkage equilibrium*) è atteso che le possibili associazioni alleliche (gameti A_1B_1 , A_2B_2 e gameti A_1B_2 , A_2B_1) siano ugualmente frequenti, dato che $pu \times qv = pv \times qu$, come nel caso di geni indipendenti. Nel caso risultasse che una delle due fasi sia più frequente dell'altra (*linkage disequilibrium*), è possibile che la popolazione non abbia ancora raggiunto lo stato di equilibrio o che una combinazione allelica sia più vantaggiosa delle altre. Nei casi di forte disequilibrio è invece possibile attribuire le differenze tra le classi gametiche alla associazione genetica. La **Fig 11.33** mostra la riduzione del parametro di *linkage disequilibrium* (D) nel corso delle generazioni per

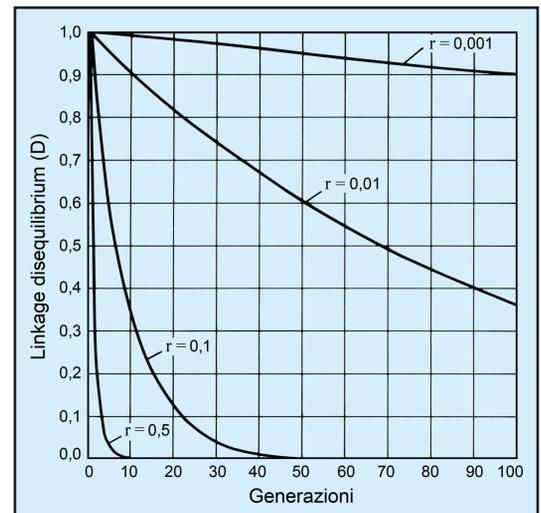


Fig. 11.33 – Andamento del *linkage disequilibrium* (D) tra i geni di una coppia di loci nel corso delle generazioni con quattro diverse frequenze di ricombinazione (r)

vari valori della frazione di ricombinazione tra i loci (r), ipotizzando presenza di unioni casuali.

I valori di D calcolati per molti geni marcatori nelle popolazioni naturali sono spesso risultati uguali a zero o molto vicini a zero, indicando quindi linkage equilibrium, con eccezioni nei casi di geni strettamente associati lungo il cromosoma. Valori altamente significativi di D sono stati comunque osservati in popolazioni naturali di piante (→ Cap. 20).

11.9 Variabilità genetica nelle popolazioni naturali

Una delle questioni più rilevanti della genetica di popolazioni è proprio quella connessa alla misura della variabilità genetica nelle popolazioni naturali e alla sua ripartizione tra popolazioni di una stessa specie. La quota di variabilità ereditabile è importante poiché condiziona i processi evolutivi e l'adattamento. Per spiegare la variabilità genetica sono stati proposti tre diversi modelli: il modello classico, il modello bilanciato e il modello neutralista.

Secondo il **modello classico**, nelle popolazioni naturali, per ogni locus esisterebbe un allele favorito dalla selezione naturale che fornisce un vantaggio adattativo. Di conseguenza, la maggior parte degli individui è omozigote per questo allele superiore o selvatico. Le mutazioni che insorgono nei vari loci portano alla formazione di nuovi alleli che sono però quasi sempre deleteri e vengono pertanto eliminati dalla selezione naturale o mantenuti nella popolazione ad una frequenza molto bassa. Qualora la mutazione risulti, invece, vantaggiosa dal punto di vista adattativo, determinando un miglioramento della capacità riproduttiva degli individui, la frequenza del nuovo allele tenderà ad aumentare fino a divenire quello prevalente, che a sua volta potrà essere indicato come selvatico. Tuttavia, tale modello non è sostenibile poiché il pool genico di una popolazione è in realtà costituito da molti possibili alleli a ciascuno dei loci.

Sulla base di questa osservazione è stato allora sviluppato il **modello bilanciato**, secondo il quale non esiste un allele superiore e più diffuso. Tale modello prevede che gli individui nella popolazione siano eterozigoti per molti loci e che l'evoluzione sia connessa alla variazione casuale della composizione allelica e ai relativi spostamenti delle frequenze alleliche. Attraverso la selezione naturale verrebbe mantenuta la variabilità genetica poiché questa agisce in modo che permanga un equilibrio tra i possibili alleli ad ogni locus, evitando che un singolo allele raggiunga frequenze elevate divenendo predominante. Questa condizione prevede pertanto la presenza in equilibrio di genotipi e fenotipi diversi nella stessa popolazione. Una forma di equilibrio bilanciato è dovuta alla sovradominanza, fenomeno per il quale l'espressione fenotipica del carattere di un eterozigote supera quella di entrambi gli omozigoti, assicurando così all'individuo eterozigote un vantaggio selettivo maggiore di ciascuno degli omozigoti in conseguenza della manifestazione di eterosi. In questa situazione, i due omozigoti, sebbene possano avere un valore selettivo molto basso, fino ad essere perfino letali, non scompaiono dalla popolazione in quanto vengono continuamente prodotti per segregazione a partire dagli individui eterozigoti. La *fitness* media più alta della popolazione si raggiunge quando la frequenza degli eterozigoti è massima e cioè quando le frequenze alleliche di A e a sono $p=q=0,5$. Ciò si verifica per effetto della selezione naturale quando gli eterozigoti manifestano eterosi e determina il mantenimento nella popolazione di omozigoti, per alleli sfavorevoli o letali, in una condizione di equilibrio. Il vantaggio dell'eterozigote e il conseguente equilibrio bilanciato ai singoli loci tra eterozigote e omozigoti hanno un ruolo determinante nel mantenimento della variabilità genetica delle popolazioni naturali, soprattutto di specie allogame. Qualora l'eterozigote non manifestasse una capacità riproduttiva

superiore, la selezione naturale condurrebbe alla fissazione dell'omozigote più adatto e conseguentemente all'esaurimento della variabilità genetica.

Esiste un terzo modello che fornisce una spiegazione alternativa all'elevata variabilità genetica osservabile nelle popolazioni naturali. Secondo il **modello neutralista** le mutazioni ricorrenti e le fluttuazioni casuali delle frequenze alleliche sono sufficienti a spiegare la variabilità genetica senza l'intervento della selezione naturale. Questo modello, sostenuto soprattutto da M. Kimura, asserisce che la maggior parte delle modificazioni a carico degli alleli sono il risultato di mutazioni selettivamente neutrali e che solo una piccolissima frazione dei cambiamenti ha significato in termini adattativi. La maggior parte delle mutazioni sarebbe infatti silente tale da lasciare il prodotto genico inalterato e quindi con effetti così piccoli sulla capacità di sopravvivenza e riproduttiva di un organismo che la selezione naturale sarebbe incapace di influenzare apprezzabilmente la loro frequenza. Le analisi molecolari hanno permesso di accertare che buona parte della variabilità nelle popolazioni naturali è in effetti dovuta ad un accumulo di mutazioni neutre che non influenzano il fenotipo dell'organismo. Di conseguenza, gli alleli neutri non vanno incontro a selezione naturale e sono mantenuti nella popolazione.

Tuttavia, non tutte le mutazioni sono neutre e quando si manifestano a livello fenotipico sono quasi sempre sfavorevoli in quanto comportano una riduzione del valore adattativo, venendo perciò eliminate dalla selezione naturale. L'insorgenza di mutazioni è indispensabile poiché creando variabilità offre alla popolazione la possibilità di evolversi e di adattarsi ai cambiamenti delle condizioni ambientali. Il mantenimento della variabilità genetica implica pertanto una diminuzione di *fitness* delle popolazioni in conseguenza della necessità di conservare geni sfavorevoli in condizione eterozigote che possono dare origine a omozigoti sfavorevoli ad ogni generazione. L'insieme dei geni recessivi deleteri o letali presenti in una popolazione viene indicato con il termine di **carico genetico**. In definitiva, la sopravvivenza della popolazione nel suo complesso è assicurata a scapito di quella dei singoli individui. Tuttavia, il carico genetico di una popolazione deve essere naturalmente contenuto entro certi limiti altrimenti la ridotta capacità di sopravvivenza e riproduttiva di un numero considerevole di individui può portare alla sua estinzione.

La variabilità fenotipica presente entro le popolazioni naturali di specie allogame è generalmente più grande di quella riscontrabile entro le popolazioni naturali di specie autogame. La relativa uniformità per molti caratteri morfologici e fisiologici di tipo poligenico consente di ipotizzare che fenotipi simili possano essere ottenuti anche con combinazioni di geni molto differenti, omozigoti o eterozigoti. Esiste tuttavia una profonda differenza tra specie autogame e allogame circa la composizione genotipica alla base della variabilità delle popolazioni naturali. Le popolazioni naturali di piante autogame sono costituite da una mescolanza di linee omozigoti strettamente imparentate che rimangono più o meno indipendenti nella riproduzione. In queste popolazioni il processo di autofecondazione continuata assicura, infatti, l'omozigosi e fornisce discendenze per lo più omogenee: la variabilità genetica è concentrata tra le linee ed è pertanto dovuta a differenze tra i genotipi omozigoti. La presenza di genotipi eterozigoti è sostanzialmente riconducibile a eventi occasionali di mutazione a singoli loci oppure a incrocio spontaneo seguito da processi di segregazione e ricombinazione ad un numero più vasto di loci. Nelle popolazioni naturali delle piante allogame tutti gli individui sono, invece, eterozigoti ad un gran numero di loci e ciascun genotipo è differente in una certa misura dagli altri. La variabilità genetica è ampia ed è distribuita fra tutti gli individui. La fecondazione incrociata determina la produzione di prole differenziata a livello genotipico. La variabilità genetica presente nelle popolazioni naturali di piante allogame può essere quindi di due tipi: libera e potenziale. La **variabilità genetica libera** è dovuta a differenze tra

Tab. 11.20 – Confronto tra i principali aspetti evolutivi delle specie vegetali prevalentemente autogame e allogame.

	Autogame	Allogame
Ciclo vitale	annuale (moltiplicazione per seme)	poliennale (propagazione vegetativa)
Destino dei mutanti recessivi	omozigosi e selezione naturale	eterozigosi e conservazione
Variabilità genetica	libera	potenziale
Flessibilità evolutiva	bassa	alta
Capacità adattativa	minima	massima

gli omozigoti, si manifesta a livello fenotipico ed è pertanto sottoposta all'azione della selezione. La **variabilità genetica potenziale** è, invece, dovuta alla presenza di genotipi eterozigoti, non si manifesta a livello fenotipico e quindi non è esposta all'azione della selezione.

La variabilità genetica dovuta alla presenza di linee omozigoti per alleli diversi, nelle autogame, e la variabilità genetica potenziale dovuta alla presenza di genotipi eterozigoti a diversi loci, nelle allogame, consentono alle popolazioni naturali di evolversi ed adattarsi (**Tab. 11.20**). A differenza delle specie autogame, il sistema riproduttivo delle specie allogame asseconda continuamente la segregazione e la ricombinazione. Inoltre, le caratteristiche del sistema riproduttivo differenziano notevolmente la flessibilità evolutiva e la capacità adattativa delle popolazioni di piante autogame e allogame. Nelle autogame, l'autofecondazione continuata assicura l'omozigosi e le piante presentano un ottimo adattamento all'ambiente in cui si trovano, ma sono poco flessibili nei confronti di ulteriori cambiamenti. Nelle allogame, la fecondazione incrociata determina la produzione di prole geneticamente differenziate che includono una frazione con scarsa vitalità, tuttavia non limita la flessibilità evolutiva e la capacità adattativa della specie a condizioni ambientali variabili. L'insorgenza di nuova variabilità attraverso le mutazioni è un fenomeno possibile nelle autogame quanto nelle allogame. Profondamente diverso è invece il destino dei mutanti recessivi nelle due situazioni: nelle popolazioni di specie autogame l'autofecondazione porta rapidamente i nuovi alleli in condizione omozigote sottoponendoli così all'azione della selezione naturale, mentre nelle popolazioni di specie allogame i nuovi alleli possono anche essere mantenuti a lungo in condizione eterozigote prima di essere sottoposti all'azione della selezione naturale.

Quadro 11.4 – Attività di miglioramento genetico e costituzione varietale in relazione al sistema riproduttivo della specie e alla struttura genetica delle popolazioni naturali

L'attività di miglioramento genetico tende sempre a ridurre la variabilità affinché le popolazioni caratterizzate da perfetta uniformità possano assicurare, nelle condizioni favorevoli create dall'agricoltore, produzioni qualitativamente omogenee e massimo rendimento. Il principale obiettivo del costitutore nel lavoro di selezione di una varietà, che è anche requisito essenziale per la sua commercializzazione, è quello di ottenere e mantenere popolazioni facilmente distinguibili, uniformi e stabili ed è evidente che tali risultati si ottengono in misura molto differente a seconda del sistema riproduttivo della specie con cui si opera.

Nelle specie che si propagano vegetativamente i problemi di costituzione varietale e selezione conservatrice sono relativamente semplici così come nelle specie apomittiche obbligate. In questi casi, infatti, la varietà si identifica con il clone cioè con un gruppo

di piante geneticamente identiche tra loro e al capostipite da cui derivano per mezzo di parti vegetative o di semi apomittici. Tali popolazioni presentano soltanto variabilità ambientale. La selezione conservatrice è comunque utile per individuare e scartare i fuori-tipo derivanti da mutazioni spontanee o eventi occasionali di incrocio.

Il lavoro di miglioramento genetico delle specie autogame inizia in genere con lo sfruttamento della variabilità presente nelle popolazioni naturali oppure della variabilità presente nelle popolazioni costituite artificialmente attraverso l'incrocio. Nel primo caso, le varietà vengono costituite attraverso selezione massale e selezione per linea pura (genealogica o individuale) mentre, nel secondo, le popolazioni segreganti vengono seguite principalmente con i metodi pedigree, per popolazione riunita, reincrocio e SSD (*single seed descent*, fondato su un seme per pianta). Tali varietà si identificano principalmente con la linea pura ad eccezione delle varietà da selezione massale che sono costituite da una pluralità di linee pure. In alcuni casi si costitui-

scono anche varietà multilinee, generalmente composte da una mescolanza di linee isogeniche che differiscono soltanto al locus o ai loci occupati dai geni per le resistenze verticali desiderate. L'eterozigosi nelle autogame è cercata o creata nella fase iniziale dei programmi di selezione.

Le varietà coltivate di specie allogame riprodotte per seme sono distinguibili in due categorie: varietà in equilibrio e varietà ibride. Le prime sono ottenute mediante selezione massale, selezione fenotipica, selezione ricorrente a partire da *landraces*, selezione basata su prove di progenie (varietà sintetiche ottenute dall'in-

terincrocio di un certo numero di cloni scelti in base all'ACG) e reincrocio e rispondono al principio dell'equilibrio di Hardy-Weinberg, mentre le seconde si basano sulla legge dell'uniformità degli ibridi F_1 di Mendel e sono ottenute con l'incrocio di linee *inbred* selezionate per l'ACS. Durante il lavoro di miglioramento genetico nelle specie prevalentemente allogame, l'eterozigosi va mantenuta o ripristinata nella sua fase finale. Ad esempio, nella costituzione degli ibridi commerciali, prodotti attraverso l'incrocio di due linee *inbred* (o di due ibridi F_1) l'incrocio è finale e consente di sfruttare il fenomeno dell'eterosi.

Sommario

Struttura genetica delle popolazioni di specie apomittiche e a propagazione vegetativa

Le popolazioni naturali di specie apomittiche e di specie propagate vegetativamente sono composte di piante geneticamente identiche tra loro, derivate da un unico capostipite per mezzo di semi apomittici o di parti vegetative. Le popolazioni di queste specie sono generalmente costituite da singoli cloni o da una mescolanza di cloni. La riproduzione asessuata permette la clonazione e la fissazione del genotipo materno e rappresenta quindi un sistema genetico che consente la rapida selezione degli individui con elevata adattabilità all'ambiente. Qualora una certa popolazione risulti policlonale (o multiclonale), come frequentemente è stato riscontrato a livello di ecotipi, la variabilità genetica è concentrata tra i cloni e la sua entità dipende dalle differenze esistenti tra i genotipi clonali. In alcuni casi è, comunque, possibile avere un genotipo clonale dominante in termini numerici che si può identificare con quello maggiormente adattato in quel determinato ambiente. In generale, le popolazioni di specie a riproduzione asessuata presentano soprattutto variabilità ambientale, salvo l'insorgenza di mutazioni spontanee o eventi occasionali di incrocio, poiché con essa non viene rilasciata variabilità genetica che può essere invece presente allo stato potenziale cioè in condizione eterozigote. Le popolazioni di specie propagate per via asessuale sono altamente eterozigoti e segregano ampiamente: il rilascio di variabilità presente nel clone avviene, nel caso di specie a propagazione vegetativa, tutte le volte che si ricorre al seme (sia in seguito ad autofecondazione che incrocio) mentre, nel caso di specie apomittiche, in seguito ad eventi sessuali occasionali (fecondazione incrociata, partenogenesi aploide, ecc.).

Struttura genetica delle specie a riproduzione sessuale

Le popolazioni naturali di specie anfimittiche sono caratterizzate da una grande variabilità. La struttura genetica delle popolazioni naturali è fortemente influenzata dal tipo di unioni che si verificano nel loro interno.

Autogame

Le popolazioni naturali di piante autogame sono costituite, di norma, da una mescolanza di linee omozigoti (linee pure) strettamente imparentate che, sebbene viventi

l'una accanto all'altra, rimangono più o meno indipendenti nella riproduzione. In queste popolazioni il processo di autofecondazione continuata assicura l'omozigosi e fornisce discendenze per lo più omogenee: la variabilità genetica è concentrata tra le linee ed è pertanto dovuta a differenze tra gli omozigoti. Tali specie tollerano bene l'*inbreeding* e presentano un ottimo adattamento all'ambiente in cui si trovano, ma sono poco flessibili nei confronti di ulteriori cambiamenti. L'incrocio occasionale e la mutazione spontanea, seguiti dalla segregazione e dalla ricombinazione, sono due fattori che si oppongono al raggiungimento della situazione limite di omozigosi a tutti i loci. Tali fattori, creando variabilità, danno alla popolazione la possibilità di evolversi e di adattarsi a condizioni ambientali variabili. Studi sulle mutazioni hanno comunque portato a concludere che la frequenza di gameti mutati ad un singolo locus è molto bassa e che la maggior parte dei tipi mutati viene eliminata dalla pressione selettiva naturale quando si manifestano. L'effetto dell'incrocio rimane invece a lungo nella popolazione, anche in considerazione del fatto che gli eterozigoti presentano sempre un certo vantaggio selettivo sugli omozigoti.

Allogame

Nelle popolazioni naturali delle piante allogame tutti gli individui sono eterozigoti ad un gran numero di loci e ciascun genotipo è differente in una certa misura dagli altri. La variabilità genetica è ampia ed è distribuita fra tutti gli individui. La fecondazione incrociata determina la produzione di progenie geneticamente differenziate che includono una frazione con scarsa vitalità, tuttavia non limita la flessibilità evolutiva e la capacità adattativa della specie a condizioni ambientali variabili.

Tali popolazioni condividono un *pool* di geni che si combinano in varia maniera nel corso delle generazioni successive. La frequenza dei singoli geni nel *pool* comune è determinata dal destino dei genotipi: se i geni contribuiscono a produrre genotipi dotati di notevole valore adattativo, o *fitness*, la loro frequenza tenderà a crescere, viceversa se hanno effetti negativi sul genotipo la loro frequenza tenderà a diminuire.

Le popolazioni di tali specie sono caratterizzate, di norma, da unioni casuali (*random mating*) che consentono di mantenere inalterate nel tempo le frequenze geniche e le frequenze genotipiche. Tali popolazioni sono dette in equilibrio Hardy-Weinberg: la composizione genetica di una popolazione numerosa rimane costante nel corso delle generazioni in presenza di unioni casuali ed in assenza di fattori di disturbo (deriva genetica, migrazione, mutazione differenziale e selezione). Nelle popolazioni che non sono in equilibrio questo si raggiunge in una sola generazione di unioni casuali, indipendentemente dalle frequenze iniziali $p(A)$ e $q(a)$ della coppia allelica considerata, ai valori seguenti: p^2 per AA , $2pq$ per Aa e q^2 per aa . Quando i loci considerati sono più di uno, l'equilibrio viene raggiunto in alcune generazioni.

Qualora individui geneticamente simili o dissimili abbiano la tendenza ad unirsi tra loro più spesso di quanto voluto dal caso l'equilibrio viene disturbato. In particolare, le unioni tra individui geneticamente simili (*inbreeding*) determinano una diminuzione della variabilità genetica della popolazione. Le unioni tra individui geneticamente dissimili (*outbreeding*) portano invece a conseguenze opposte, cioè ad un mantenimento o addirittura ad un aumento della variabilità genetica della popolazione. Le popolazioni di specie allogame non tollerano o tollerano male l'*inbreeding* e manifestano un vantaggio eterotico più marcato rispetto alle autogame e questo fatto condiziona anche i metodi di miglioramento genetico e di costituzione varietale.

Bibliografia di riferimento e approfondimento

- Barrai, I. (1977) *Genetica di popolazioni*. Piccin Editore, Padova.
- Crow, J.F., M. Kimura (1970) *An introduction to population genetics theory*. Harper & Row Publ., New York.
- Crow, J.F. (1986) *Basic concepts in population, quantitative and evolutionary genetics*. Freeman, New York.
- Dobzhansky, T. (1937) *Genetics and the origin of species*. Columbia Biological Series, Columbia University Press, New York.
- Falconer, D.S. (1983) *Introduction to quantitative genetics* (2nd ed.). Longman Group Ltd., Harlow, UK.
- Fisher, R.A. (1949) *The theory of inbreeding*. Oliver Boyd, Edinburgh.
- Fisher, R. (1958) *The genetical theory of natural selection*. Dover, New York.
- Hardy, G.H. (1908) Mendelian proportions in a mixed population. *Science*, 28: 49-50.
- Hartl, D.L., A.G. Clark (1988) *Principles of population genetics* (2nd ed.). Sinauer, Sunderland, MA.
- Hedrick, P.H. (2000) *Genetics of populations* (2nd ed.). Jones & Bartlett Publ., Boston.
- Lewontin, R.C. (1985) Population genetics. *Annual Reviews in Genetics*, 19:81-102.
- Lorenzetti, F., M. Falcinelli, F. Veronesi (1994) *Miglioramento genetico delle piante agrarie*. Edagricole, Bologna.
- Nei, M. (1987) *Molecular evolutionary genetics*. Columbia University Press, New York.
- Ottaviano, E. (1988) *Struttura genetica delle popolazioni delle piante coltivate*. In: G.T. Scarascia Mugnozza (coordinatore) *Miglioramento genetico vegetale*, pp. 103-120. Pàtron Editore, Bologna.
- Stebbins, G.L. (1950) *Variation and evolution in plants*. Columbia Biological Series, Columbia University Press, New York.
- Weinberg, T. (1908) Über den nachweis der vererbung beim menschen. *Jahresh. Ver. F. Vaterl. Naturk. In Württemberg*, 64: 369-382.
- Wright, S. (1968) *Evolution and the genetics of populations*. The University of Chicago Press, Chicago.

