

# Dinamica evolutiva delle popolazioni

Camilla Corfini  
Relatore Prof. G. Marchesini

18 febbraio 2010

# Indice

<b>1</b>	<b>Introduzione</b>	<b>2</b>
<b>2</b>	<b>Legge di Hardy-Weinberg</b>	<b>5</b>
2.1	Conservazione delle frequenze geniche . . . . .	5
2.2	Significato della Legge di Hardy-Weinberg . . . . .	11
2.3	Alcune applicazioni . . . . .	11
2.3.1	L'alcaptonuria e il problema della dominanza . . . . .	11
2.3.2	I geni legati all'X . . . . .	12
<b>3</b>	<b>I processi evolutivi</b>	<b>14</b>
3.1	Fitness e selezione . . . . .	14
3.1.1	La Pressione Selettiva . . . . .	15
3.1.2	La selezione in alcuni casi particolari . . . . .	20
3.1.3	Analisi della legge di F-H-W nel caso di selezione debole . . . . .	32
3.2	La Mutazione . . . . .	33
3.3	Equilibrio tra Selezione e Mutazione . . . . .	35
<b>A</b>	<b>Alcune nozioni di Genetica</b>	<b>38</b>
A.1	La nascita della Genetica . . . . .	38
A.2	I cromosomi . . . . .	38
A.3	Il DNA . . . . .	39
A.4	I geni . . . . .	40
	<b>Bibliografia</b>	<b>43</b>

# Capitolo 1

## Introduzione

La biologia moderna si fonda essenzialmente sul concetto di evoluzione che con Darwin (1809-1882) è stato oggetto di una trattazione coerente e approfondita, anche se l'idea di evoluzione era stata considerata molti anni prima. Già nel sesto secolo a.C. la scuola del greco Anassimandro aveva introdotto una teoria evolucionistica.

L'evento che, più di ogni altro, contribuì alla nascita e allo sviluppo delle idee evolucionistiche in Darwin, fu il viaggio sul brigantino Beagle. Nell'ottobre del 1838 Darwin ebbe modo di leggere il *Saggio sul principio della popolazione*, pubblicato nel 1798 dall'economista Thomas Robert Malthus (1766-1834), in cui sostiene che in assenza di ostacoli al suo sviluppo, una popolazione cresce secondo una progressione geometrica mentre le risorse disponibili sono suscettibili ad un andamento aritmetico. Egli sostiene quindi che è sbagliato cercare di migliorare le condizioni di vita dei meno abbienti, bisogna piuttosto imporre il controllo sulle nascite. Pur applicando il principio di Malthus dei fattori limitanti all'evoluzione, Darwin non ne accetterà mai l'applicazione nel contesto sociale.

Darwin adottò il concetto, sviluppato da Herbert Spencer (1820-1903), di *sopravvivenza del più adatto*, secondo cui gli individui di una specie che avessero ereditato variazioni positive avrebbero avuto maggiore probabilità di lasciare discendenti, di conseguenza la loro frequenza sarebbe aumentata.

Le teorie fondamentali sviluppate da Darwin sono: *L'evoluzione come fatto oggettivo*, ossia il reale susseguirsi delle specie nel tempo; egli sostiene che la vita ha avuto un'unica origine, differenziandosi successivamente ed ininterrottamente originando nuove specie, quindi, *per discendenza comune*. Egli spiega l'immensa varietà delle specie che si sono succedute e di quelle attualmente viventi con la *moltiplicazione delle specie*; in base a questa teoria le specie si suddividono in specie figlie nate da variazioni genetiche avvenute in popolazioni isolate geograficamente. Per Darwin il *motore dell'evoluzione*

è la *selezione naturale*, essa agisce sulle variazioni che si verificano negli individui, con un processo lento e continuo, senza brusche variazioni improvvise. L'evoluzione è quindi un processo *graduale*.

Oltre a queste teorie, Darwin sviluppò alcune teorie corollarie alle precedenti: selezione sessuale, pangenesi, effetto dell'uso e non uso, divergenza dei caratteri, ecc. . .

Mentre la biologia evoluzionistica sconvolgeva il pensiero del tempo, sarebbe stato possibile dare risposta a molti interrogativi sull'ereditarietà se gli esperimenti effettuati dal monaco ceco Gregor Mendel (1822-1884) fossero venuti alla luce. Egli concentrò i suoi studi sulla trasmissione ereditaria di singoli caratteri applicando ai suoi esperimenti il massimo rigore matematico. Ibridando fra loro delle piante di pisello, e concentrandosi su singoli caratteri Mendel arrivò a formulare le sue tre leggi fondamentali:

1. *Legge della dominanza* (o legge della omogeneità di fenotipo);
2. *Legge della segregazione* (o legge della disgiunzione);
3. *Legge dell'assortimento indipendente* (o legge di indipendenza dei caratteri).

Nel 1865 egli pubblicò i risultati del suo lavoro che passano del tutto inosservati nel mondo scientifico. Soltanto nel 1900 la scoperta del processo di meiosi e di altre componenti biologiche, come i cromosomi, ha reso possibile la verifica pratica delle leggi enunciate da Mendel e ha portato a una rivalutazione dei suoi scritti e delle sue teorie.

Tali leggi infatti sembrarono inizialmente una confutazione del darwinismo in quanto facevano riferimento alle mutazioni come ad un fenomeno drastico, sufficiente ad originare ogni volta una nuova specie, opponendosi quindi al concetto di mutazione legato ad una variazione continua sviluppato da Darwin.

La pubblicazione dei risultati di Gregor Mendel avvenne successivamente al 1859, anno in cui Charles Darwin pubblicò il suo "Sull'origine delle specie". Darwin non poté perciò avvalersi dei risultati di Mendel.

Fu proprio su tale argomento che agli inizi del Novecento si acuì la polemica fra genetisti da una parte e naturalisti dall'altra. Fra il 1937 e il 1950 vi fu una riunificazione dei campi di indagine, in passato assurdamente separati. Il primo passo fu compiuto dai genetisti Fisher e Haldane i quali si accorsero che le mutazioni sostanziali (mendeliane), rappresentano l'eccezione in natura e che si erano sottovalutate notevolmente la frequenza e l'utilità delle piccole mutazioni.

Oggi sappiamo che le piccole mutazioni avvengono frequentemente ed è su queste che opera la selezione naturale. Le basi dell'ereditarietà avrebbero potuto essere intuite da studi ben precedenti alle teorie di Mendel, egli ebbe

l'intuizione di scegliere caratteri qualitativi, mentre studi precedenti avevano preso in oggetto l'ereditarietà dei tratti quantitativi che tendono ad essere molto più complessi da spiegare.

Maupertuis, geografo e fisico, famoso per il *principio di minima azione* (1744-1747), già nel 1700 studiò l'incidenza del raro tratto qualitativo della polidattilia<sup>1</sup> in una famiglia, osservata in varie generazioni. Studiando l'ereditarietà di questa caratteristica egli intuì che il tratto genetico era causato da certi fattori ereditati indistintamente sia dalla madre che dal padre e conservati nelle generazioni successive. Il tratto è determinato da *due* fattori, ereditati singolarmente da entrambi i genitori; ogni tratto ha eguale probabilità di essere passato alla progenie e il fatto che sia espresso uno piuttosto che l'altro è dato dalla *dominanza* (o *recessività*) dello stesso; vale a dire che c'è un carattere che domina sull'altro e che si manifesta.

Questo lavoro presenta alcuni modelli di Genetica Evoluzionistica delle Popolazioni, nell'intento di studiare la dinamica di alcuni modelli di popolazioni.

Nella parte iniziale si presentano i risultati ottenuti dagli studi del matematico britannico Godfrey Harold Hardy e dal medico tedesco Weinberg riguardo alla conservazione delle frequenze geniche.

Tali risultati porteranno allo studio della dinamica di una popolazione che raggiunge l'equilibrio sotto ipotesi molto restrittive e difficilmente riscontrabili in una popolazione reale.

Il risultato ottenuto nella legge di Hardy e Weinberg corrisponde ad un'assenza di cambiamento, di evoluzione, che nella realtà non è quasi mai possibile.

Nella seconda parte del lavoro sulla Genetica Delle Popolazioni, che considera le basi genetiche per l'evoluzione in una popolazione, tenendo in considerazione gli effetti della riproduzione sessuale e della dominanza genetica, si procederà con lo studio di una popolazione senza alcune delle ipotesi imposte dalla legge di Hardy e Weinberg e si arriveranno a valutare le cause che portano una popolazione ad evolversi; si parlerà in particolare di *selezione* e *mutazione*.

---

<sup>1</sup>La polidattilia (dal greco antico *polus* "molto" + *daktulos* "dito"), anche conosciuta come iperdattilia, è un'anomalia congenita, nella quale le mani o i piedi dei soggetti affetti presentano un eccesso di dita rispetto alla normale conformazione.

## Capitolo 2

# Legge di Hardy-Weinberg

### 2.1 Conservazione delle frequenze geniche

Hardy e Weinberg applicarono le leggi di Mendel alle popolazioni che si accoppiano in modo casuale. In questo lavoro, per semplicità, viene assunto per ipotesi che le frequenze degli alleli siano le stesse per i due sessi. Supponiamo che ciò che succede in un certo *locus* sia indipendente da quanto succeda in ogni altro e ci focalizzeremo sui cambiamenti del singolo locus. Supponiamo che ci siano solo due alleli A e B che possano occupare tale locus. Si avranno quindi tre *genotipi*: gli omozigoti AA o BB o l'eterozigote AB; per i due alleli A e B di un gene autosomico la **legge di Hardy-Weinberg** può essere enunciata in due parti:

1. Nelle cosiddette condizioni di Hardy-Weinberg (vedi sotto), le frequenze genotipiche di una popolazione diventano, per *una sola* generazione, le seguenti:

$$\begin{aligned}\text{freq}(\mathbf{AA}) &= \mathbf{p^2} \\ \text{freq}(\mathbf{AB}) &= \mathbf{2pq} \\ \text{freq}(\mathbf{BB}) &= \mathbf{q^2}\end{aligned}$$

dove  $p$  e  $q$  sono le frequenze alleliche:  $p = \text{freq}(A)$  e  $q = \text{freq}(B)$ , ovviamente si ha  $p + q = 1$ . Queste equazioni indicano che è possibile prevedere le frequenze genotipiche sulla base delle frequenze alleliche.

2. Fintanto che valgono le condizioni di Hardy-Weinberg, le frequenze genotipiche non cambiano; vale a dire che, generazione dopo generazione, le frequenze dei tre genotipi rimangono costanti sui valori  $p^2$ ,  $2pq$  e  $q^2$ . Tale costanza nel tempo viene definita uno **stato di equilibrio**, a significare che non si verifica alcun cambiamento netto nelle frequenze. In assenza di mutazioni genetiche, di cui si parlerà in seguito, i fattori che determinano i tratti sopravvivono *inalterati* di generazione in

generazione; inoltre si fa l'ipotesi semplificativa di considerare le generazioni discrete e non sovrapposte nel tempo, così il tempo sarà una variabile discreta e dovremmo arrivare ad equazioni alle differenze.

Le condizioni sotto cui la legge di Hardy-Weinberg è applicabile sono:

- (a) che gli accoppiamenti siano pressoché casuali, ciò significa che il numero atteso di incontri tra il genotipo 1 e il genotipo 2 deve essere proporzionale al prodotto delle loro frequenze;
- (b) che il rapporto numerico tra i sessi sia indipendente dal genotipo;
- (c) che le mutazioni relative agli alleli A e B siano in quantità trascurabile;
- (d) che le migrazioni all'interno ed all'esterno della popolazione siano trascurabili;
- (e) che la fertilità sia indipendente dal genotipo, ciò significa che non ci sia differenza nel numero atteso di gameti prodotti da individui geneticamente diversi né nel numero atteso di zigoti prodotti da incroci geneticamente differenti;
- (f) che la probabilità di sopravvivenza fino all'età riproduttiva sia indipendente dal genotipo.

Non esiste una popolazione reale che risponda esattamente ai requisiti del modello matematico, ma esistono in molte popolazioni svariati geni i quali sottostanno a tali condizioni in modo sufficiente da permettere che la legge di Hardy-Weinberg possa diventare valida ed utile.

Siano  $x$ ,  $y$  e  $z$  le frequenze dei genotipi AA, AB e BB:  $x = freq(AA)$ ,  $y = freq(AB)$  e  $z = freq(BB)$ ; si ha che

$$p = x + \frac{1}{2}y \quad \text{e} \quad q = z + \frac{1}{2}y . \quad (2.1)$$

Assumendo che gli incontri tra i gameti siano casuali possiamo riassumere le probabilità dell'unione tra essi con il “*quadrato di Punnett*”, dal nome dell'omonimo biologo:

<b>Quadrato di Punnett</b>			
		Freq. gameti femminili	
		<i>A</i>	<i>B</i>
Freq. gameti maschili		<i>p</i>	<i>q</i>
<i>A</i>	<i>p</i>	<i>p</i> <sup>2</sup>	<i>pq</i>
<i>B</i>	<i>q</i>	<i>pq</i>	<i>q</i> <sup>2</sup>

Significa che l'unione di un gamete femminile A con un gamete maschile B ha frequenza  $pq$  e viceversa, quindi la frequenza di un eterozigote AB della generazione successiva sarà  $2pq$ ; mentre la frequenza degli omozigoti sarà  $p^2$ , per AA e  $q^2$ , per BB. Segue che:

$$p_{n+1} = p_n^2 + \frac{1}{2}2p_nq_n = p_n(p_n + q_n) = p_n ,$$

$$q_{n+1} = q_n^2 + \frac{1}{2}2p_nq_n = q_n(p_n + q_n) = q_n .$$

dove  $n$  si riferisce alla "generazione  $n$ -esima" e quindi con  $n + 1$  si intende "la generazione successiva". Si è dimostrato che le frequenze  $p_n$  e  $q_n$  sono costanti indipendenti da  $n$ , quindi chiamiamole semplicemente  $p$  e  $q$ ; si avrà che per  $n \geq 1$ , ma non necessariamente per  $n = 0$ ,  $x_n = p^2$ ,  $y_n = 2pq$  e  $z_n = q^2$ .

### Prova

Nelle condizioni di H-W si parte da un assunto che spesso corrisponde alla realtà, cioè l'**accoppiamento casuale**, ovvero l'accoppiamento che avviene indipendentemente dal genotipo o fenotipo di un individuo. In particolare questo assunto è valido per caratteri umani quali il gruppo sanguigno. Ciò comporta che ciascun individuo ha eguale probabilità di accoppiarsi con ciascun individuo dell'altro sesso e che le frequenze di tale accoppiamento dipendono *solamente* da quelle dei due genotipi in questione. Consideriamo una popolazione parentale in cui

$$\begin{aligned} \text{frequenza del genotipo } AA &= x \\ \text{frequenza del genotipo } AB &= y \\ \text{frequenza del genotipo } BB &= z \end{aligned}$$

Accoppiamenti	Freq.di acc.	Prob.relativa		
		AA	AB	BB
$AA \times AA$	$x^2$	1	0	0
$AB \times AB$	$y^2$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$
$BB \times BB$	$z^2$	0	0	1
$AA \times AB$	$2xy$	$\frac{1}{2}$	$\frac{1}{2}$	0
$BB \times AB$	$2zy$	0	$\frac{1}{2}$	$\frac{1}{2}$
$AA \times BB$	$2xz$	0	1	0

Le frequenze dei genotipi fra tutti i discendenti sono:

$$\begin{aligned}
 freq(AA) &= (1)x^2 + \left(\frac{1}{4}\right)y^2 + \left(\frac{1}{2}\right)2xy = x^2 + xy + \left(\frac{1}{4}\right)y^2 = \left(x + \frac{1}{2}y\right)^2 = p^2 ; \\
 freq(AB) &= \left(\frac{1}{2}\right)y^2 + \left(\frac{1}{2}\right)2xy + \left(\frac{1}{2}\right)2zy + (1)2xz = 2\left(xz + \frac{1}{2}xy + \frac{1}{2}zy + \frac{1}{4}y^2\right) = \\
 &= 2\left(x + \frac{1}{2}y\right)\left(z + \frac{1}{2}y\right) = 2pq ; \\
 freq(BB) &= \left(\frac{1}{4}\right)y^2 + (1)z^2 + \left(\frac{1}{2}\right)2zy = z^2 + zy + \left(\frac{1}{4}\right)y^2 = \left(z + \frac{1}{2}y\right)^2 = q^2 .
 \end{aligned}$$

Osserviamo ad esempio che la frequenza di BB fra i discendenti, ottenuta considerando tutti gli accoppiamenti che avrebbero prodotto tale progenie, è:  $freq(BB) = \left(z + \frac{1}{2}y\right)^2$  e osserviamo che  $z + \frac{1}{2}y$  è la frequenza dell'allele  $B$  nei genitori, cioè  $q$ ; abbiamo perciò mostrato che in una popolazione soggetta alle condizioni di Hardy-Weinberg la frequenza di  $BB$  diventa  $q^2$  in una sola generazione. Allo stesso modo, come si è visto sopra, si arriva a concludere che la frequenza di  $AA$  diventa  $p^2$  in una sola generazione, e la frequenza di  $AB$  diventa  $2pq$ .

La legge di Hardy-Weinberg afferma che la distribuzione dei genotipi  $p^2 : 2pq : q^2$  è stabile nel tempo e ciò può essere provato ripetendo il procedimento fatto sopra, o più semplicemente riconoscendo che la distribuzione delle frequenze che è stata ricavata è del tutto generale e permette di conoscere la distribuzione  $p^2 : 2pq : q^2$  nella generazione della progenie prescindendo dalle frequenze genotipiche della generazione parentale. Ciò avviene grazie all'assunzione di avere frequenze alleliche uguali nei due sessi.

Dalla figura (2.1), osservando l'andamento delle curve relative alle frequenze genotipiche, possiamo notare che, per frequenze  $p$  basse (ossia nel caso di un *allele raro*), la curva  $2pq$  cresce in misura maggiore della curva  $p^2$ . In altre parole un *allele raro* si manifesta nettamente di più negli eterozigoti. Per capire cosa implica questo fatto prendiamo come esempio il caso dell'*allele recessivo e raro* della fibrosi cistica.

Il fatto che esso sia recessivo implica che la malattia si manifesti solo nel caso di genotipi omozigoti, mentre nel caso degli eterozigoti si tratta solo di individui portatori di tale malattia.

Poichè la fibrosi cistica è considerata una malattia letale<sup>1</sup> si ha che gli omozigoti per tale allele scompariranno a causa di essa. Nonostante questo fatto, dato che si tratta di un allele raro, la predominanza di esso negli eterozigoti implica che gli effetti della selezione naturale su tale allele siano drasticamente ridotti.

---

<sup>1</sup>In questi casi si parla di *allele letale* nella forma omozigotica.

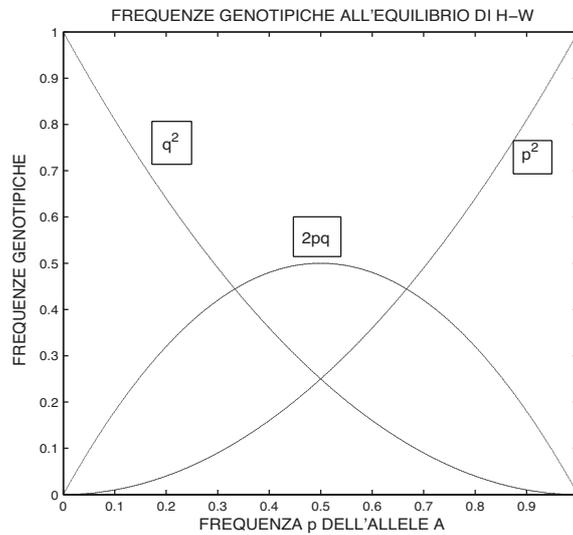


Figura 2.1: Frequenze genotipiche all'equilibrio di Hardy-Weinberg

In conclusione l'equazione di Hardy-Weinberg si applica ai loci singolarmente: una popolazione può essere in equilibrio per un locus, ma non per altri. Inoltre in una popolazione in cui avviene accoppiamento casuale si produce una distribuzione di equilibrio dei genotipi dopo soltanto una generazione, per cui la variabilità viene mantenuta (Equilibrio di H-W), rimane costante nel tempo, generazione dopo generazione. La conseguenza è che indipendentemente dalla distribuzione dei genotipi nella generazione parentale, in assenza di altre perturbazioni, nella generazione successiva la distribuzione dei genotipi risultante da accoppiamenti casuali è completamente specificata dalla frequenza allelica; in particolare per un locus biallelico è completamente specificata dalla frequenza  $p$ . Se una popolazione è in *equilibrio* si ha che le *frequenze genotipiche* dipendono esclusivamente dalle *frequenze alleliche* della generazione precedente, le quali non cambiano attraverso le generazioni.

### Raggiungimento dell'equilibrio per due o più loci

Il caso considerato di due soli alleli è un caso particolare; per geni che presentano **più di due alleli per locus**, dobbiamo adottare metodi differenti a seconda che ci interessino: le frequenze genotipiche determinate da un solo allele, i valori di equilibrio per i genotipi di due alleli, o infine i valori di

equilibrio per tre o più alleli. L'unica differenza rispetto al caso di due soli alleli per locus sarà il fatto che la somma delle singole frequenze alleliche dipenderà dalle frequenze dei genotipi; di conseguenza non sarà più 1 e quindi nemmeno il suo quadrato (somma delle frequenze genotipiche), infatti a causa della presenza degli altri alleli, la frequenza totale di questi genotipi sarà diversa da 1. In ogni caso, siccome ogni *gamete aploide* contiene *un solo allele per ogni locus*, le possibili combinazioni nello zigote di tali gameti aploidi dipenderanno soltanto dalle frequenze degli alleli (come nel caso particolare di due soli alleli per locus) e quindi anche in questi tre casi elencati, si stabilisce l'equilibrio in una sola generazione di accoppiamento casuale.

In generale è corretto parlare di *raggiungimento dell'equilibrio in una sola generazione* fintanto che consideriamo ogni locus separatamente, senza curarsi di ciò che accade negli altri; se invece consideriamo contemporaneamente i prodotti di due loci, ad esempio *Aa* e *Bb*, il numero dei possibili genotipi che si avranno aumenta a  $3^2$  (*AABB*, *AABb*, *AaBB*, *AaBb*, *AAbb*, *Aabb*, *aaBB*, *aaBb*, *aabb*). Avremo 4 frequenze corrispondenti ai 4 alleli, chiamiamole ad esempio *p q r s* e il rapporto di equilibrio dei loro genotipi è dato da  $(pr + ps + qr + qs)^2$  che dipende dalle frequenze *pr*, *ps*, *qr* e *qs* dei genotipi *AA*, *Ab*, *aB* e *ab* rispettivamente.

Quindi il raggiungimento dell'equilibrio riguarda essenzialmente il tempo impiegato dalle frequenze gametiche a raggiungere questi valori.

Per comprendere il significato di queste conclusioni prendiamo ad esempio come popolazione iniziale con i genotipi *AaBb* × *AaBb*, cioè di solo eterozigoti, in cui le *frequenze di tutti gli alleli* siano *uguali* (ad esempio  $p = q = r = s = 0.5$ ) si ha che tutti i 4 tipi di gameti (*AB*, *Ab*, *aB* e *ab*) vengono prodotti con le frequenze d'equilibrio e l'equilibrio dei genotipi si raggiunge in una generazione; ciò che è stato detto sopra accade *solo nel caso di frequenze alleliche iniziali uguali*. Consideriamo un caso limite che mostra come mai il raggiungimento dell'equilibrio per i gameti e i genotipi sarà ritardato di un certo periodo: prendiamo una popolazione iniziale avente i genotipi *AABB* e *aabb*, tale popolazione produrrà soltanto i tipi di gameti *AB* e *ab* e l'equilibrio per *tutti* i genotipi non può essere raggiunto nella generazione successiva dal momento che mancano numerosi genotipi (*aaBB*, *AAbb* ...).

## 2.2 Significato della Legge di Hardy-Weinberg

La condizione di equilibrio prevista dalla legge di Hardy-Weinberg equivale ad un'affermazione di assenza di ogni cambiamento, di ogni evoluzione.

Il fatto che gli alleli si mantengano su frequenze diverse da zero significa anche che la variabilità genetica all'interno di una popolazione non si modifica nelle generazioni successive. Questo è proprio il concetto mendeliano di unità ereditarie discrete. Le teorie moderne sull'evoluzione affondano le proprie radici proprio sulla genetica delle popolazioni: "un individuo non subisce evoluzione, una popolazione sì", e ciò accade a causa proprio di quelle variabili che si sono trascurate a fronte delle condizioni di Hardy-Weinberg: la *non* casualità dell'accoppiamento, le mutazioni, le migrazioni e specialmente la *selezione naturale* (intesa da Darwin come il risultato della maggiore efficienza riproduttiva degli individui meglio adatti). Tali meccanismi modificano le frequenze alleliche e di conseguenza i cambiamenti del fenotipo medio di una popolazione; inoltre l'accumularsi di tali modificazioni nel tempo porta ai profili evolutivi studiati e documentati dai paleontologi.

Quando i meccanismi dell'evoluzione sono trascurabili, la frequenza di un allele rimane costante, sia esso dominante o recessivo; infatti la frequenza di un allele all'equilibrio dipende principalmente dall'influenza che esso esercita sulla capacità riproduttiva degli individui.

## 2.3 Alcune applicazioni

Esaminiamo alcuni esempi numerici reali in cui la legge di Hardy-Weinberg possa fornire risultati interessanti.

### 2.3.1 L'alcaptonuria e il problema della dominanza

Fra 250 000 bambini scelti a caso forse soltanto uno sarà affetto da alcaptonuria avendo genotipo  $k/k$ . La diagnosi viene effettuata osservando semplicemente l'imbrunimento dell'urina, ma non esiste un test che ci consenta di distinguere tra i rimanenti 249 999 bambini i portatori sani  $K/k$  dagli omozigoti normali  $K/K$

Genotipo	Numero	Frequenza
$K/-$	249 999	$0,999\ 996 \approx 1$
$k/k$	1	$0,000\ 004 = 4 \times 10^{-6}$
totale	250 000	1

Se cerchiamo di calcolare la frequenza dell'allele responsabile dell'alcaptonuria  $q = freq(k)$ , sorge un problema: noi non sappiamo quale sia il numero degli eterozigoti, in quanto la malattia è determinata da un allele recessivo e perciò non possiamo usare l'equazione (2.1). Si può procedere

alternativamente facendo qualche assunzione sulla popolazione: siccome la malattia è molto rara e relativamente benigna è ragionevole supporre che la distribuzione dei genotipi si avvicini alla formula di Hardy-Weinberg; ci aspettiamo quindi di trovare

$$q^2 = \text{freq}(k/k) = 4 \times 10^{-6} \text{ quindi } q = \sqrt{4 \times 10^{-6}} = 2 \times 10^{-3} = 0,002$$

Da qui possiamo stimare la frazione degli eterozigoti presenti nella popolazione

$$\text{freq}(K/k) = 2pq = 2(1 - 0,002)0,002 \approx 0,004$$

In pratica stiamo eguagliando una frequenza *osservata* ( $4 \times 10^{-6}$ ) ad una frequenza *attesa* ( $q$ ) sulla base di ipotesi fatte riguardo alla popolazione presa in esame, le quali se non fossero valide ci porterebbero ad un errore di procedimento nel calcolare tale frequenza. Infatti la stima è assolutamente grossolana ma in un caso del genere non c'è altro modo di procedere e quindi accetteremo il valore ricavato con la legge di H-W e lo considereremo come il migliore valore di stima delle frequenze alleliche.

### 2.3.2 I geni legati all'X

Per i caratteri legati al sesso, il numero dei genotipi possibili aumenta, dal momento che vi è una differenza nel numero dei cromosomi sessuali tra il sesso omogametico (femminile) e il sesso eterogametico (maschile). I geni legati al cromosoma X saranno presenti in doppia copia nelle femmine ma soltanto in singola copia nei maschi. Per tale motivo le frequenze di equilibrio sotto le condizioni di Hardy-Weinberg devono essere stabilite separatamente per i due sessi:

	<b>Genotipo</b>	<b>Freq. di eq. di H-W</b>
femmine	$H/H$	$p^2$
	$H/h$	$2pq$
	$h/h$	$q^2$
maschi	$H/(Y)$	$p$
	$h/(Y)$	$q$

In questo caso le lettere  $H$  e  $h$  vengono usate per indicare i due alleli di un qualsiasi gene legato ad X (per esempio quello dell'emofilia);  $p$  e  $q$  rappresentano rispettivamente le frequenze di  $H$  e di  $h$  in ciascun sesso. Fra le femmine le frequenze determinate secondo la legge di H-W sono le stesse che ci dovremmo aspettare per un gene autosomico. Invece nei maschi è *uno solo* l'allele che ne determina il fenotipo legato all'X,  $H$  oppure  $h$ . Perciò le frequenze dei genotipi o dei fenotipi nei maschi coincidono con le frequenze alleliche.

Come si desume dalla tabella riportata sopra, se si verifica un accoppiamento casuale fra questi genotipi, risulteranno delle frequenze genotipiche stabili sia per i maschi che per le femmine.

Questi valori di equilibrio sono basati sul fatto che le frequenze geniche per  $H$  e  $h$ , cioè  $p$  e  $q$ , sono identiche in entrambi i sessi. Considerando ad esempio che 1 maschio su  $10^4$  è affetto da emofilia classica ed ha quindi un genotipo  $h/(Y)$ , si ha che  $q = freq(h) = 1/10^4 = 10^{-4}$ ; rispetto a questi dati la formula di H-W fa capire perchè il carattere sia estremamente raro, pur se non assente, nelle femmine. Infatti una femmina malata deve avere genotipo  $h/h$ , e la frequenza attesa è  $q^2 = (10^{-4})^2 = 10^{-8}$  vale a dire un caso su 100 milioni. Se la popolazione totale degli Stati Uniti (oltre 200 milioni di individui) fosse un'enorme popolazione soggetta alla legge di H-W, ci dovremmo aspettare di trovare soltanto alcune femmine emofiliache. In effetti, si sono riportati solo pochi casi di femmine omozigoti malate i cui padri erano a loro volta affetti e le madri portatrici.

## Capitolo 3

# I processi evolutivi

### 3.1 Fitness e selezione

L'idea di evoluzione precedette il lavoro di Charles Darwin di molti anni. Agli inizi del 1800, per esempio, il francese Jean-Baptiste de Lamarck sostenne l'evoluzione e propose, in modo abbastanza logico per il suo tempo, che essa fosse dovuta all'eredità dei cosiddetti caratteri acquisiti. Prove convincenti a sostegno di questa teoria non furono fornite né da Lamarck né da nessun altro. Darwin invece raccolse prove molto importanti circa il fatto che l'evoluzione avviene per mezzo della selezione naturale. Questo meccanismo evolutivo, insieme alle leggi dell'eredità pubblicate nel 1865 da Mendel ma sconosciute a Darwin e generalmente sottovalutate fino ai primi del '900, può render conto dei cambiamenti adattativi, successivi e casuali, caratteristici dell'evoluzione delle popolazioni nello spazio e nel tempo.

Esistono altri meccanismi evolutivi -mutazione, migrazione e deriva- ma questi fattori da soli non sono capaci di guidare la selezione. Haldane, Fisher, Wright ed altri si occuparono di studiare la loro interazione, sia fra loro che in relazione alla selezione. In questo lavoro mostreremo una parte dei risultati che ottennero, in particolare vedremo come la mutazione e la selezione agiscano nel meccanismo evolutivo.

Importante è il concetto darwiniano di **fitness** letteralmente traducibile ne "l'essere adatto", inteso in senso genetico come il numero medio di progenie prodotta da un individuo o da una classe di individui. Quando un genotipo può produrre più discendenti di un altro nello stesso ambiente si dice che la sua *adattabilità* è superiore. In questo senso, possiamo anche dire che certi genotipi hanno *valori adattativi* o *valori selettivi* diversi, altri termini per indicare il successo riproduttivo. Se indichiamo il valore adattativo  $W$  del genotipo, come la porzione di discendenza che sopravvive alla maturità relativamente a quella di altri genotipi,  $W$  può essere tale da cadere in un intervallo tra 1,00 per i genotipi più "adatti", e 0 per i letali. La forza che agisce su ciascun genotipo riducendone il valore adattativo, viene definita

coefficiente di selezione  $s$ . Quindi le relazioni tra  $W$  ed  $s$  sono semplicemente  $W = 1 - s$  o  $s = 1 - W$ .

Ciò che caratterizza la fitness sono le differenze in termini di *fertilità* e *capacità di sopravvivenza*. Maggiore è la fitness di individui con un particolare fenotipo, più i loro geni tendono ad essere rappresentati nelle generazioni future. I valori di fitness sono dipendenti dalle condizioni ambientali prevalenti e, se queste cambiano, la fitness degli individui può cambiare. La *fitness assoluta* di un genotipo si definisce in termini del suo successo riproduttivo; per generazioni distinte e non sovrapposte è il contributo in termini di copie di ogni gene che un individuo di quel genotipo si aspetta di apportare al **pool genico** (il complesso dei geni di una popolazione) della generazione successiva. La *fitness relativa* di un genotipo è il rapporto tra la sua fitness assoluta e quella di un genotipo di riferimento.

### 3.1.1 La Pressione Selettiva

Le fasi di un ciclo di vita (partendo dal tempo  $t$  e giungendo a  $t + 1$ ) si succedono nel seguente modo: *fase zigotica-fase riproduttiva-fase gametica-fase zigotica*; tra la fase zigotica e quella riproduttiva e tra la fase gametica e la nuova fase zigotica si può considerare un diverso tasso di sopravvivenza dei geni, mentre tra la fase gametica e quella zigotica si osservano differenti coefficienti di fertilità.

Al contrario di quanto si assume nelle legge di H-W consideriamo che ci sia un qualche vantaggio o svantaggio selettivo per l'allele  $A$ , cioè assumiamo che sia diversa la probabilità di sopravvivenza del genotipo  $AA$  dalla fase zigotica a quella riproduttiva. Manteniamo valida l'assunzione di riproduzione casuale tra i genotipi, ciò ci permetterà di usare il quadrato di Punnett e ottenere le equazioni unicamente in termini di frequenze alleliche. Siano  $p_n$  e  $q_n$  le frequenze alleliche alla fine della fase gametica della generazione  $n$ . All'inizio della fase zigotica siano date le frequenze genotipiche

$$x_n = p_n^2, y_n = 2p_nq_n \text{ e } z_n = q_n^2$$

e siano date le probabilità di sopravvivenza dalla fase zigotica a quella riproduttiva nel seguente rapporto  $w_x : w_y : w_z$  rispettivamente di  $AA$ ,  $AB$  e  $BB$ . In questa situazione si ha che i valori relativi di selezione e fitness si equivalgono. Alla fase riproduttiva il rapporto tra i genotipi  $AA$ ,  $AB$  e  $BB$  si sarà modificato in

$$w_x p_n^2 : w_y 2p_nq_n : w_z q_n^2$$

si avrà quindi che le frequenze alleliche sono ora nel rapporto

$$w_x p_n^2 + w_y p_nq_n : w_y p_nq_n + w_z q_n^2 .$$

Poichè non si ha diversa fertilità né diversa sopravvivenza dei gameti, gli alleli sono nello stesso rapporto anche alla fine della fase gametica della generazione  $n + 1$ . In frequenza si ha:

$$p_{n+1} = f(p_n) = \frac{(w_x p_n + w_y q_n) p_n}{w_x p_n^2 + w_y 2p_n q_n + w_z q_n^2} \quad (3.1)$$

$$q_{n+1} = h(q_n) = \frac{(w_y p_n + w_z q_n) q_n}{w_x p_n^2 + w_y 2p_n q_n + w_z q_n^2} \quad (3.2)$$

questo è il sistema non lineare che governa il processo di *selezione naturale*; il numeratore della (3.1) rappresenta la probabilità di sopravvivenza alla fase riproduttiva dell'allele  $A$  alla generazione  $n + 1$  e il denominatore la somma delle probabilità di sopravvivenza di tutti i genotipi alla generazione  $n + 1$  (assunte uguali in tutte le generazioni).

Facciamo qualche conto per riscrivere la (3.1):

moltiplichiamo per  $(p_n + q_n) = 1$  il numeratore e otteniamo

$$p_{n+1} = f(p_n) = p_n \frac{(w_x p_n + w_y q_n)(p_n + q_n)}{w_x p_n^2 + w_y 2p_n q_n + w_z q_n^2} = \frac{p_n (w_x p_n^2 + w_y p_n q_n + w_x p_n q_n + w_y q_n^2)}{w_x p_n^2 + w_y 2p_n q_n + w_z q_n^2}$$

ora aggiungendo e togliendo le quantità  $w_y p_n q_n$  e  $w_z q_n^2$  al numeratore, otteniamo

$$p_{n+1} = \frac{p_n (w_x p_n^2 + w_y 2p_n q_n + w_z q_n^2 + w_x p_n q_n - w_y p_n q_n + w_y q_n^2 - w_z q_n^2)}{w_x p_n^2 + w_y 2p_n q_n + w_z q_n^2}$$

con analoghi calcoli sulla (3.2) giungiamo infine al sistema di equazioni:

$$\begin{cases} p_{n+1} = p_n + g_1(p_n) = p_n + p_n q_n \frac{(w_x - w_y) p_n + (w_y - w_z) q_n}{w_x p_n^2 + w_y 2p_n q_n + w_z q_n^2} \\ q_{n+1} = q_n + g_2(q_n) = q_n + p_n q_n \frac{(w_y - w_x) p_n + (w_z - w_y) q_n}{w_x p_n^2 + w_y 2p_n q_n + w_z q_n^2} \end{cases} \quad (3.3)$$

Osserviamo che se fossimo nel caso di probabilità di sopravvivenza uguali per i tre genotipi,  $w_x = w_y = w_z = w$ , verificherebbero che la popolazione raggiungerebbe l'equilibrio dopo una sola generazione, com'era stato dedotto nella legge di Hardy-Weinberg. In simboli avremmo

$$p_{n+1} = p_n \frac{w(p_n + q_n)}{w(p_n^2 + 2p_n q_n + q_n^2)} = p_n \frac{w(p_n + q_n)}{w(p_n + q_n)^2} = p_n .$$

La prima equazione del sistema (3.3) è l'**equazione di Fisher-Haldane-Wright** che dà una completa descrizione dei cambiamenti che avvengono nelle frequenze geniche come risultato del processo di selezione.

Studiamo la stabilità dei punti di equilibrio del sistema di equazioni (3.1) e (3.2). Si osservano innanzitutto che i due stati di equilibrio corrispondenti a  $(p, q) = (0, 1)$  e  $(p, q) = (1, 0)$  corrispondono rispettivamente, il primo ad

avere l'allele  $A$  fissato, il secondo ad avere fissato nella popolazione l'allele  $B$ .

Consideriamo il sistema di equazioni:

$$\begin{cases} p_{n+1} = f(p_n) = \frac{w_x p_n^2 + w_y p_n q_n}{w_x p_n^2 + w_y 2p_n q_n + w_z q_n^2} = f_1(p, q) \\ q_{n+1} = h(q_n) = \frac{w_y p_n q_n + w_z q_n^2}{w_x p_n^2 + w_y 2p_n q_n + w_z q_n^2} = f_2(pq) \end{cases}$$

e calcoliamo la matrice Jacobiana :

$$J = \begin{bmatrix} \frac{\partial f_1}{\partial p} & \frac{\partial f_1}{\partial q} \\ \frac{\partial f_2}{\partial p} & \frac{\partial f_2}{\partial q} \end{bmatrix}$$

$$\begin{aligned} \frac{\partial f_1}{\partial p} &= \frac{(2pw_x + w_yq)(w_x p^2 + w_y 2pq + w_z q^2) - (w_x p^2)(2w_x p + 2w_y q)}{(w_x p^2 + w_y 2pq + w_z q^2)^2} \Big|_{(0,1)} = \frac{w_y}{w_z} \\ \frac{\partial f_1}{\partial q} &= \frac{w_y p(w_x p^2 + w_y 2pq + w_z q^2) - (w_x p^2 + w_y pq)(2w_y p + 2w_z q)}{(w_x p^2 + w_y 2pq + w_z q^2)^2} \Big|_{(0,1)} = 0 \\ \frac{\partial f_2}{\partial p} &= \frac{w_y q(w_x p^2 + w_y 2pq + w_z q^2) - (w_y pq + w_z q^2)(2w_x p + 2w_y q)}{(w_x p^2 + w_y 2pq + w_z q^2)^2} \Big|_{(0,1)} = -\frac{w_y}{w_z} \\ \frac{\partial f_2}{\partial q} &= \frac{(w_y p + 2w_z q)(w_x p^2 + w_y 2pq + w_z q^2) - (w_y pq + w_z q^2)(2w_y p + 2w_z q)}{(w_x p^2 + w_y 2pq + w_z q^2)^2} \Big|_{(0,1)} = 0 \end{aligned}$$

si ha che la matrice Jacobiana valutata per  $p = 0$  e  $q = 1$  è:

$$J_{(0,1)} = \begin{bmatrix} \frac{w_y}{w_z} & 0 \\ -\frac{w_y}{w_z} & 0 \end{bmatrix}_{(p=0, q=1)}$$

gli autovalori sono  $\lambda_1 = \frac{w_y}{w_z}$  e  $\lambda_2 = 0$  e l'equilibrio  $(0,1)$  è asintoticamente stabile se  $\left| \frac{w_y}{w_z} \right| < 1$ .

Allo stesso modo calcoliamo la matrice Jacobiana valutata per  $p = 1$  e  $q = 0$ :

$$\frac{\partial f_1}{\partial p} \Big|_{(1,0)} = 0, \quad \frac{\partial f_1}{\partial q} \Big|_{(1,0)} = -\frac{w_y}{w_x}, \quad \frac{\partial f_2}{\partial p} \Big|_{(1,0)} = 0, \quad \frac{\partial f_2}{\partial q} \Big|_{(1,0)} = \frac{w_y}{w_x};$$

quindi si ha che:

$$J_{(1,0)} = \begin{bmatrix} 0 & -\frac{w_y}{w_x} \\ 0 & \frac{w_y}{w_x} \end{bmatrix}_{(p=1, q=0)}$$

gli autovalori sono  $\lambda_3 = \frac{w_y}{w_x}$  e  $\lambda_4 = 0$  e l'equilibrio  $(1,0)$  è asintoticamente

stabile se  $\left| \frac{w_y}{w_x} \right| < 1$ .

Questo ci porta a concludere che, perchè gli equilibri siano stabili, gli eterozigoti devono avere *minore* probabilità di sopravvivenza rispetto a quella degli omozigoti.

L'analisi dell'equazione di Fisher-Haldane-Wright può essere resa più semplice se consideriamo il rapporto  $u = \frac{p}{q}$  o  $v = \frac{q}{p}$ , ottenendo

$$u' = \frac{p'}{q'}, v' = \frac{q'}{p'} \text{ ricordando che } p' = \frac{w_p p}{\bar{w}} \text{ e } q' = \frac{w_q q}{\bar{w}}$$

si ha che:

$$\delta u = u' - u = \frac{w_p p}{w_q q} - \frac{p}{q} = \left( \frac{w_p}{w_q} - 1 \right) \frac{p}{q}$$

quindi

$$\delta u = \frac{w_p - w_q}{w_q} u = \frac{(w_y - w_z) + (w_x - w_y)u}{w_z + w_y u} u . \quad (3.4)$$

Allo stesso modo otteniamo

$$\delta v = v' - v = \frac{w_q q}{w_p p} - \frac{q}{p} = \left( \frac{w_q}{w_p} - 1 \right) \frac{q}{p}$$

e quindi

$$\delta v = \frac{w_q - w_p}{w_p} v = \frac{(w_y - w_x) + (w_z - w_y)v}{w_x + w_y v} v . \quad (3.5)$$

Con le espressioni (3.4) e (3.5) si può arrivare agli stessi risultati ricavati in precedenza.

Definiamo ora la *fitness media*  $w_p$  di A prendendo la media pesata del contributo di tutti gli omozigoti e metà del contributo degli eterozigoti portatori dell'allele A:

$$w_p = \frac{w_x p^2 + w_y p q}{p^2 + p q} = w_x p + w_y q \quad (3.6)$$

la *fitness media*  $w_q$  di B:

$$w_q = \frac{w_y p q + w_z q^2}{p q + q^2} = w_y p + w_z q \quad (3.7)$$

la *fitness media totale*  $\bar{w}$  come:

$$\bar{w} = \frac{w_x p^2 + w_y 2 p q + w_z q^2}{p^2 + 2 p q + q^2} = p w_p + q w_q \quad (3.8)$$

L'equazione (3.1), togliendo i pedici "n" a  $p$  e  $q$  e usando  $p'$  per intendere la generazione  $n + 1$ , diventa

$$p' = \frac{w_p p}{\bar{w}} \quad (3.9)$$

e la (3.3) diventa

$$\delta p = p' - p = \frac{\alpha_p p}{\bar{w}} = \frac{(w_p - \bar{w})p}{\bar{w}} \quad (3.10)$$

la quantità  $\alpha_p = w_p - \bar{w}$  si chiama *fitness d'eccesso medio* dell'allele A. Questa equazione può essere interpretata dicendo che la quantità relativa  $\frac{\delta p}{p}$  con la quale un genotipo A cresce in una generazione è dato dalla fitness d'eccesso nella popolazione relativamente alla fitness della stessa popolazione:

$$\delta p = p' - p = pq \frac{w_p - w_q}{\bar{w}} = pq \frac{(w_x - w_y)p + (w_y - w_z)q}{\bar{w}} \quad (3.11)$$

Uno stato di equilibrio intermedio  $p^* = 1 - q^*$  esiste se A e B sono ugualmente adatti, cioè se  $w_{p^*} = w_{q^*}$ , che equivale a dire

$$(w_x - w_y)p^* + (w_y - w_z)q^* = 0 ,$$

il che è biologicamente realistico unicamente se si intende che  $(w_y - w_x)$  e  $(w_y - w_z)$  abbiano lo stesso segno, cioè che sia

$$\frac{p^*}{q^*} = \frac{w_y - w_z}{w_y - w_x} > 0$$

ossia

$$\frac{p^*}{q^*} = \frac{1 - q^*}{q^*} = \frac{1}{q^*} - 1 > 0 \quad \text{cioè} \quad q^* < 1$$

conseguenza di ciò è il fatto che sia  $p^* > 0$ ; in altre parole, la condizione appena descritta sta a significare che gli eterozigoti debbano essere o più o meno adatti rispetto ad entrambi gli omozigoti.

### 3.1.2 La selezione in alcuni casi particolari

Nella maggior parte degli organismi superiori, sia animali che vegetali, la selezione si verifica soprattutto allo stato diploide, cioè dopo la formazione dello zigote. Nei diploidi vi sono tre possibili genotipi ( $AA$ ,  $Aa$ ,  $aa$ ), così che l'efficacia della selezione dipende anche dalle relazioni di dominanza.

Se esiste dominanza completa e se la selezione è soltanto contro il recessivo  $aa$ , in generale l'eterozigote si dice "protetto" dalla selezione, così che i geni recessivi e dannosi o addirittura letali possono persistere in tale condizione per generazioni. A titolo di esempio consideriamo in questo paragrafo alcuni esempi che ci chiariscano le idee su come agisce la selezione dipendentemente dai diversi rapporti di dominanza tra i due alleli.

Parleremo di selezione "*contro*" o "*verso*" un tipo di allele, intendendo il fatto che essa risulti essere a svantaggio o a vantaggio della presenza di questi alleli nel pool genico della popolazione considerata.

La selezione *contro* un allele *dominante* equivale ad una selezione a favore di un allele *recessivo*.

Faremo distinzione sui tipi di alleli: "*vantaggiosi*", "*dannosi*" e "*letali*".

Essi si riferiscono: i primi ad alleli che favoriscono l'adattamento dell'individuo all'ambiente, i secondi al contrario lo sfavoriscono e gli ultimi portano alla scomparsa dell'individuo.

Negli ultimi due esempi tratteremo di casi particolari in cui la selezione avviene tra alleli che, o non sono né completamente dominanti né completamente recessivi (caso di semidominanza), o tali che si abbia un vantaggio adattativo degli eterozigoti superiore a quello di entrambi gli omozigoti (caso di sovradominanza).

#### 1. Selezione contro allele recessivo $a$ dannoso, non letale nell'omozigote

Quando gli omozigoti recessivi non sono letali, cioè quando per il coefficiente di selezione vale  $s < 1$ , la modificazione della frequenza genica in una generazione si può derivare nel modo seguente:

**Selezione contro allele  $a$  dannoso, non letale nell'omozigote**

	<b>AA</b>	<b>Aa</b>	<b>aa</b>	<b>totale</b>	<b>freq(a)</b>
freq. iniziale	$p^2$	$2pq$	$q^2$	1	$q$
fitness	1	1	$1 - s$		
freq. dopo selez.	$p^2$	$2pq$	$q^2(1 - s)$	$p^2 + 2pq + q^2 - sq^2$	
freq. relativa	$\frac{p^2}{1 - sq^2}$	$\frac{2pq}{1 - sq^2}$	$\frac{q^2(1 - s)}{1 - sq^2}$		$\frac{q(1 - sq)}{1 - sq^2}$

$$\Delta q = f_{fin} - f_{in} = \frac{q(1 - sq)}{1 - sq^2} - q = \frac{q - sq^2 - q + sq^3}{1 - sq^2} = \frac{-sq^2(1 - q)}{1 - sq^2}$$

La situazione che ha determinato la selezione può cambiare, tanto che l'allele dominante possa essere a sua volta sottoposto a selezione e venga così favorito l'allele recessivo. Quando ciò si verifica la selezione sarà ovviamente più efficace, dal momento che il gene dominante è presente in due genotipi su tre; se addirittura un allele dominante diventa letale, la sua frequenza si riduce a zero in una sola generazione.

**2. Selezione contro allele dominante  $A$  dannoso, non letale nell'omozigote**

Nel caso in cui ci sia selezione contro un allele dominante con frequenza genica  $p$  si ha:

**Selezione contro allele  $A$  dannoso dominante**

	<b>AA</b>	<b>Aa</b>	<b>aa</b>	<b>totale</b>	<b>freq(A)</b>
freq. iniziale	$p^2$	$2pq$	$q^2$	1	$p$
fitness	$1 - s$	$1 - s$	1		
freq. dopo selez.	$p^2(1 - s)$	$2pq(1 - s)$	$q^2$	$1 - sp(2 - p)$	
freq. relativa	$\frac{p^2(1 - s)}{1 - sp(2 - p)}$	$\frac{2pq(1 - s)}{1 - sp(2 - p)}$	$\frac{q^2}{1 - sp(2 - p)}$		$\frac{p - sp}{1 - sp(2 - p)}$

$$\begin{aligned}\Delta q = f_{fin} - f_{in} &= \frac{p - sp}{1 - sp(2 - p)} - p = \frac{p - sp - p + sp^2(2 - p)}{1 - sp(2 - p)} = \\ &= \frac{-sp(1 - 2p + p^2)}{1 - sp(2 - p)} = \frac{-sp(1 - p)^2}{1 - sp(2 - p)}\end{aligned}$$

Quando manca la dominanza (lo vedremo più avanti), gli eterozigoti mostreranno l'effetto dei geni sfavoriti dalla selezione; dal momento che essi portano solo uno di tali geni, tuttavia, il loro coefficiente di selezione sarà esattamente metà di quello degli omozigoti che portano due degli stessi geni svantaggiosi.

### 3. La selezione verso un allele dominante e *vantaggioso*

Consideriamo l'allele A **dominante** e vantaggioso, assumiamo che  $\mathbf{w}_z = \mathbf{1}$ , abbiamo che  $\mathbf{w}_x = \mathbf{w}_y = \mathbf{1} + \mathbf{s}$  con  $s$  *coefficiente di selezione* e misuriamo la forza della selezione a favore di A.

Sostituendo nel sistema (3.3), si ha

$$p_{n+1} = p_n + p_n q_n \frac{(1 + s - 1)q_n}{(1 + s)p_n^2 + (1 + s)2p_n q_n + q_n^2}$$

e

$$q_{n+1} = q_n + p_n q_n \frac{(1 - 1 - s)q_n}{(1 + s)p_n^2 + (1 + s)2p_n q_n + q_n^2}$$

il sistema di equazioni (3.3) di F-H-W diventa quindi

$$\begin{cases} p_{n+1} = p_n + sp_n q_n \frac{q_n}{1 + s(p_n^2 + 2p_n q_n)} = \tilde{f}_1(p, q) \\ q_{n+1} = q_n - sp_n q_n \frac{q_n}{1 + s(p_n^2 + 2p_n q_n)} = \tilde{f}_2(p, q) \end{cases} \quad (3.12)$$

Si hanno ancora gli stati di equilibrio  $(p, q) = (0, 1)$  e  $(p, q) = (1, 0)$ ;

$$\begin{aligned}\frac{\partial \tilde{f}_1}{\partial p} &= 1 + \frac{sq_n^2(1 + s(p_n^2 + 2p_n q_n)) - sp_n q_n^2(2sp_n + 2sq_n)}{(1 + s(p_n^2 + 2p_n q_n))^2} \\ \frac{\partial \tilde{f}_1}{\partial q} &= \frac{2sp_n q_n(1 + s(p_n^2 + 2p_n q_n)) - sp_n q_n^2(2sp_n)}{(1 + s(p_n^2 + 2p_n q_n))^2} \\ \frac{\partial \tilde{f}_2}{\partial p} &= \frac{-sq_n^2(1 + s(p_n^2 + 2p_n q_n)) + sp_n q_n^2(2sp_n + 2sq_n)}{(1 + s(p_n^2 + 2p_n q_n))^2} \\ \frac{\partial \tilde{f}_2}{\partial q} &= 1 + \frac{-2sp_n q_n(1 + s(p_n^2 + 2p_n q_n)) + sp_n q_n^2(2sp_n)}{(1 + s(p_n^2 + 2p_n q_n))^2}\end{aligned}$$

il sistema (3.12) ha matrici Jacobiane valutate in  $(0,1)$  e  $(1,0)$

$$J_{(0,1)} = \begin{bmatrix} 1+s & 0 \\ -s & 1 \end{bmatrix}_{(p=0,q=1)}$$

e

$$J_{(1,0)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}_{(p=1,q=0)}$$

in entrambi i casi abbiamo un autovalore di modulo unitario: nel primo caso abbiamo la presenza di un altro autovalore di modulo maggiore di 1 (  $|1+s| > 1$  poichè  $s > 0$  ) e quindi possiamo concludere che l'equilibrio  $(p, q) = (0, 1)$  è instabile per il sistema non lineare; mentre per quanto riguarda il secondo caso, avendo entrambi gli autovalori di modulo unitario, non possiamo concludere nulla sulla stabilità del sistema non lineare mediante linearizzazione.

Possiamo però osservare che  $(p_n)$  è una successione monotona crescente limitata inferiormente da 0 e superiormente da 1, tale che  $p = 0$  è equilibrio instabile e  $p = 1$  stabile.

Infatti l'andamento nell'intorno di  $p = 0$  è esponenziale e mano a mano che ci si avvicina a  $p = 1$  diventa algebrico. Si ha in particolare che per  $p$  prossimo a 0,  $q$  è vicino a 1 e viceversa quindi l'equazione per  $p \approx 0$  è

$$p_{n+1} = (1+s)p_n$$

cioè  $p_n \approx p_0(1+s)^n$ . Per  $p = 1$  usiamo il rapporto  $u = \frac{p}{q}$  per rendere le cose più semplici; l'equazione diventa

$$\delta u = u' - u = \frac{su}{1 + (1+s)u}.$$

Al primo ordine per  $q$  piccolo, cioè per  $u$  grande, abbiamo che

$$u' - u = \frac{s}{\frac{1}{u} + (1+s)} \frac{u}{u} \approx \frac{s}{1+s}$$

si ha quindi l'equazione approssimata

$$u_{n+1} = u_n + \frac{s}{1+s}$$

cosicché per  $q_0$  piccolo, si ha  $u_0 \approx \frac{1}{q_0}$  e le equazioni approssimate diventano

$$u_n = \frac{1}{q_n} = \frac{(1+s)u_0 + ns}{1+s}$$

quindi essendo

$$q_n = \frac{1+s}{(1+s)\frac{1}{q_0} + ns} = \frac{(1+s)q_0}{1+s+nsq_0},$$

si ha che  $q_n \rightarrow 0$ , per  $n \rightarrow \infty$ .

#### 4. La selezione verso un allele recessivo e *vantaggioso*

Ora invece supponiamo che A sia **recessivo** e vantaggioso, come prima assumiamo che  $\mathbf{w}_z = \mathbf{1}$  e  $\mathbf{w}_x = \mathbf{1} + \mathbf{s}$  ma ora si ha  $\mathbf{w}_y = \mathbf{1}$ ; sostituendo nel sistema (3.3) abbiamo:

$$p_{n+1} = p_n + p_n q_n \frac{(1+s-1)p_n + (1-1)q_n}{(1+s)p_n^2 + 2p_n q_n + q_n^2}$$

e

$$q_{n+1} = q_n - p_n q_n \frac{(1+s-1)p_n + (1-1)q_n}{(1+s)p_n^2 + 2p_n q_n + q_n^2}$$

il sistema di equazioni (3.3) di F-H-W diventa quindi

$$\begin{cases} p_{n+1} = p_n + sp_n q_n \frac{p_n}{1+sp_n^2} = \tilde{h}_1(p, q) \\ q_{n+1} = q_n - sp_n q_n \frac{p_n}{1+sp_n^2} = \tilde{h}_2(p, q) . \end{cases} \quad (3.13)$$

Abbiamo stati di equilibrio in  $(p, q) = (0, 1)$  e  $(p, q) = (1, 0)$ ;

$$\begin{aligned} \frac{\partial \tilde{h}_1}{\partial p} &= 1 + \frac{2sp_n q_n (1+sp_n^2) - sp_n^2 q_n (2sp_n)}{(1+sp_n^2)^2} \\ \frac{\partial \tilde{h}_1}{\partial q} &= \frac{sp_n^2 (1+sp_n^2)}{(1+sp_n^2)^2} \\ \frac{\partial \tilde{h}_2}{\partial p} &= \frac{-2sp_n q_n (1+sp_n^2) + sp_n^2 q_n (2sp_n)}{(1+sp_n^2)^2} \\ \frac{\partial \tilde{h}_2}{\partial q} &= 1 + \frac{-sp_n^2 (1+sp_n^2)}{(1+sp_n^2)^2} \end{aligned}$$

e il sistema (3.13) ha matrici Jacobiane valutate in (0,1) e (1,0)

$$J_{(0,1)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}_{(p=0,q=1)}$$

e

$$J_{(1,0)} = \begin{bmatrix} 1 & \frac{s}{1+s} \\ 0 & \frac{1}{1+s} \end{bmatrix}_{(p=1,q=0)}$$

per il primo equilibrio, avendo entrambi gli autovalori di modulo unitario, non possiamo concludere nulla sulla stabilità del sistema non lineare mediante linearizzazione; per il secondo equilibrio abbiamo la presenza di un autovalore di modulo minore di 1 (  $\left| \frac{1}{1+s} \right| < 1$  ) e uno di modulo unitario, per questa ragione nuovamente non possiamo concludere nulla sulla stabilità del sistema.

Si ha ancora che  $(p_n)$  è una successione monotona crescente limitata inferiormente da 0 e superiormente da 1, tale che  $p = 0$  è equilibrio instabile e  $p = 1$  stabile. A differenza del caso precedente, a partire da  $p = 0$  si ha un andamento algebrico, mentre si arriva a  $p = 1$  con andamento esponenziale. Per un dato valore di  $s$ , si ha selezione debole per una frequenza bassa di A al contrario di quanto accadeva nel caso di A dominante; ciò è dovuto al fatto che gli eterozigoti, che hanno molte copie dell'allele A se è raro, non traggono il beneficio del vantaggio selettivo di A nel caso in cui esso sia recessivo.

## 5. La selezione verso un allele recessivo e *letale*

Per un allele recessivo e letale i valori di fitness e le frequenze dei genotipi sono:

	<b>AA</b>	<b>Aa</b>	<b>aa</b>	<b>totale</b>
popolaz. iniziale	$p^2$	$2pq$	$q^2$	1
fitness	1	1	0	
popolaz. che si riproduce	$p^2$	$2pq$	0	$p^2 + 2pq = 1 - q^2$

per fitness di  $aa$  uguale a zero si intende specificare che il genotipo  $aa$  non si riproduce. Interessante è osservare come a differenza del caso di un allele dominante e letale, gli alleli recessivi  $a$ , potenzialmente letali, presenti negli eterozigoti, risultano essere temporaneamente protetti dalla selezione.

Assumiamo che la popolazione sia soggetta a tutte le condizioni di H-W a meno della selezione; supponiamo inoltre che gli individui  $aa$  muoiano prima di riprodursi. Avremo che alla prima generazione, detta  $freq(a) = q$  la frequenza iniziale dell'allele  $a$ :

$$freq(a) = q' = \frac{0 + \frac{1}{2}(2pq)}{1 - q^2} = \frac{pq}{(1 - q)(1 + q)} = \frac{q}{1 + q}$$

che è la frequenza assoluta dell'allele  $a$ ; se  $freq(a)$  iniziale è bassa, il cambiamento da  $q$  a  $\frac{q}{1+q}$  sarà conseguentemente piccolo e più raro diventerà il recessivo letale, meno efficace sarà la selezione nel ridurre ulteriormente la sua frequenza. Ciò dipende essenzialmente dal fatto che, man mano che un allele diventa più raro, il numero degli eterozigoti non si riduce così velocemente come quello degli omozigoti. Si dice che il numero degli eterozigoti fenotipicamente normali *mascherano* il letale recessivo nei confronti della selezione molto di più di quello degli omozigoti che *espongono* il letale alla selezione.

## 6. La selezione nel caso di *semidominanza*

Esiste anche il caso in cui un allele  $A$  non sia né totalmente dominante né totalmente recessivo, avviene in generale nei casi di alleli che portano l'informazione di tratti quantitativi (come l'altezza) piuttosto che qualitativi, ed è chiamato **caso additivo** o **semi-dominante**. In questo caso gli eterozigoti avrebbero un vantaggio selettivo intermedio rispetto a quello degli omozigoti; in particolare se  $A$  fosse vantaggioso varrebbe  $w_x > w_y > w_z$ , viceversa se  $A$  fosse svantaggioso varrebbe  $w_x < w_y < w_z$ . Se né  $A$  né  $B$  sono dominanti, ma entrambi semi-dominanti, essi portano lo stesso contributo al vantaggio selettivo degli eterozigoti ed è perciò che si parla di *assenza di effetto di dominanza* riguardo al caso additivo.

In questo caso possiamo scrivere il sistema (3.3) considerando  $w_z = 1$ ,  $w_x = 1 + 2s$  e  $w_y = 1 + s$ , nel modo seguente:

$$\begin{cases} p_{n+1} = p_n + \frac{sp_nq_n}{1+2sp_n} = \tilde{f}_1(p, q) \\ p_{n+1} = p_n + \frac{sp_nq_n}{1+2sp_n} = \tilde{f}_2(p, q) \end{cases} \quad (3.14)$$

Calcoliamo le matrici Jacobiane valutate nei due stati di equilibrio  $(p, q) = (0, 1)$  e  $(p, q) = (1, 0)$ :

$$\begin{aligned}\frac{\partial \tilde{f}_1}{\partial p} &= 1 + \frac{q_n s(1 + 2sp_n) - p_n q_n s(2s)}{(1 + 2sp_n)^2} \\ \frac{\partial \tilde{f}_1}{\partial q} &= \frac{p_n s(1 + 2sp_n)}{(1 + 2sp_n)^2} \\ \frac{\partial \tilde{f}_2}{\partial p} &= \frac{-q_n s(1 + 2sp_n) + p_n q_n s(2s)}{(1 + 2sp_n)^2} \\ \frac{\partial \tilde{f}_2}{\partial q} &= 1 + \frac{-p_n s(1 + 2sp_n)}{(1 + 2sp_n)^2}\end{aligned}$$

il sistema (3.14) ha le seguenti matrici Jacobiane valutate in  $(0,1)$  e  $(1,0)$ :

$$J_{(0,1)} = \begin{bmatrix} 1+s & 0 \\ -s & 1 \end{bmatrix}_{(p=0, q=1)}, \quad J_{(1,0)} = \begin{bmatrix} 1 & \frac{s}{1+2s} \\ 0 & \frac{1+s}{1+2s} \end{bmatrix}_{(p=1, q=0)}$$

in entrambi i casi abbiamo un autovalore di modulo unitario: nel primo caso abbiamo la presenza di un altro autovalore di modulo maggiore di 1 (  $|1+s| > 1$  poichè  $s > 0$  ) e quindi possiamo concludere che l'equilibrio  $(p, q) = (0, 1)$  è instabile per il sistema non lineare; mentre per quanto riguarda il secondo caso, avendo entrambi gli autovalori di modulo unitario, non possiamo concludere nulla sulla stabilità del sistema non lineare mediante linearizzazione.

Analizzando l'andamento della successione  $(p_n)$  si possono fare discorsi analoghi a quelli fatti nei casi **3** e **4**.

In particolare analizzando tale successione si osserva che si ha un andamento esponenziale sia a partire da  $p = 0$  sia al tendere di  $p$  a 1.

7. La selezione verso entrambi gli omozigoti (Il vantaggio degli eterozigoti)

In quasi tutti gli esempi di selezione che abbiamo considerato finora si va sempre in una direzione, cioè verso l'eliminazione dell'allele dannoso e la stabilizzazione o fissazione dell'allele favorito. In questi casi fin tanto che il coefficiente di selezione rimane costante, un equilibrio tra alleli favoriti e alleli sfavoriti è impossibile senza il verificarsi di una *mutazione* (di cui parleremo più avanti).

C'è un caso particolare in cui è possibile lo stabilirsi di un equilibrio per cui gli alleli possono rimanere *indefinitamente* all'interno della popolazione, a patto che i coefficienti di selezione rimangano costanti. Tale condizione è nota con il nome di *sovradominanza* e si verifica quando l'eterozigote ha una fitness riproduttiva superiore a quella di entrambi gli omozigoti.

In generale se l'eterozigote  $Aa$  ha un valore adattativo 1, mentre la fitness dell'omozigote  $AA$  è ridotta dal coefficiente di selezione  $s$  e quella dell'omozigote  $aa$  dal coefficiente di selezione  $t$ , calcoliamo la modificazione della frequenza ( $\Delta q$ ) di  $a$  in una singola generazione:

	<b>AA</b>	<b>Aa</b>	<b>aa</b>	<b>totale</b>	<b>freq(a)</b>
freq. iniziale	$p^2$	$2pq$	$q^2$	1	q
fitness	$1 - s$	1	$1 - t$		
freq. dopo selez.	$p^2(1 - s)$	$2pq$	$q^2(1 - t)$	$1 - p^2s - q^2t$	$\frac{q(1-qt)}{1-p^2s-q^2t}$

$$\begin{aligned}
 \Delta q &= f_{fin} - f_{in} = \frac{q(1-qt)}{1-p^2s-q^2t} - q = \frac{q(1-qt) - q(1-p^2s-q^2t)}{1-p^2s-q^2t} = \\
 &= \frac{q - q^2t - q - qp^2s - q^3t}{1-p^2s-q^2t} = \frac{qp^2s + q^2t(1-p) - q^2t}{1-p^2s-q^2t} = \\
 &= \frac{qp^2s + q^2t - pq^2t - q^2t}{1-p^2s-q^2t} = \frac{qp^2s - pq^2t}{1-p^2s-q^2t} = \frac{pq(ps - qt)}{1-p^2s-q^2t} .
 \end{aligned}$$

Vi è raggiungimento dell'equilibrio nel momento in cui sarà  $\Delta q = 0$  e non vi sarà un'ulteriore modificazione delle frequenze geniche. Vi sono tre condizioni in cui ciò accade:

- (a) le prime due si verificano quando o  $p$  o  $q$  sono uguali a zero;
  - tuttavia in queste condizioni entrambi i geni non saranno contemporaneamente presenti nella popolazione e quindi non vi sarà in ogni caso equilibrio;
- (b) si può però verificare che  $ps = qt$ ; in tali condizioni deriviamo le seguenti relazioni:
  - aggiungendo  $qs$  ad entrambi i termini, si ottiene:

$$\begin{aligned} ps + qs &= qt + qs \\ s(p + q) &= q(s + t) \\ q &= \frac{s}{s + t} \end{aligned}$$

- invece, aggiungendo  $pt$  ad entrambi i termini, si ottiene:

$$\begin{aligned} ps + pt &= qt + pt \\ p(s + t) &= t(p + q) \\ p &= \frac{t}{s + t} \end{aligned}$$

Si comprende che se  $s$  e  $t$  sono valori costanti, sia  $p$  che  $q$  raggiungeranno un equilibrio stabile; nel senso che se ad esempio  $q$  si spostasse dai valori di equilibrio, la pressione di selezione lo forzerà a tornare verso l'equilibrio. se  $\Delta q > 0$  significa che la frequenza genica  $q$  aumenta, in caso contrario, se  $\Delta q < 0$ ,  $q$  diminuisce. Quindi il segno positivo o negativo di  $\Delta q$  dipende dal fatto che  $q$  sia al di sopra o al di sotto del suo valore di equilibrio.

Entrambi gli alleli rimangono quindi nella popolazione con frequenze  $\hat{p}$  e  $\hat{q}$ , fin tanto che questi coefficienti di selezione conferiscono agli eterozigoti una fitness superiore a quella di entrambi gli omozigoti.

Il mantenimento di questi genotipi costituisce un esempio di “*polimorfismo bilanciato*”, termine introdotto da Ford per descrivere la conservazione della *variabilità genetica* (polimorfismo) attraverso la selezione.

Per stabilità si intende appunto la tendenza della selezione a riportare le frequenze alleliche ai valori di equilibrio, qualora queste se ne allontanino.

Il *vantaggio dell'eterozigosi* è uno dei modi per mezzo dei quali gli alleli di un gene si bilanciano fra loro, stabilizzandosi su frequenze intermedie in seguito a forze selettive opposte.

Non tutti i polimorfismi sono tuttavia bilanciati in modo stabile o in equilibrio permanente. Essi si considerano instabili nel momento in cui un qualche disturbo nelle frequenze causi la fissazione o la perdita di uno dei geni.

Un esempio di equilibrio instabile è possibile quando ad esempio la *selezione* agisce *contro gli eterozigoti*, nel caso semplice di un locus con due alleli.

Se entrambi gli omozigoti hanno uguale fitness e l'eterozigote ha una fitness inferiore, si avrà condizione di equilibrio soltanto nel caso in cui la frequenza di entrambi gli alleli sia  $p = q = \frac{1}{2}$ . In queste condizioni si ha che gli alleli sono perfettamente bilanciati, poichè la stessa porzione di ciascuno dei due viene sottratta in quanto presente nell'eterozigote: cioè  $0.25AA$ ,  $0.50Aa$ ,  $0.25aa$ .

Ogni leggero scostamento da questo valore farà sì che *l'allele meno frequente* avrà proporzionalmente più geni nell'eterozigote rispetto all'*allele più frequente*; esso comincerà perciò a perdere proporzionalmente più geni, mentre l'altro allele andrà verso la fissazione.

### Applicazione (Anemia falciforme in Africa)

Il modello sopra discusso è ben illustrato dal polimorfismo dell'anemia falciforme<sup>1</sup>. In molte regioni dell'Africa, ove è diffusa la malaria, gli individui *sani* omozigoti per l'allele normale (non del tipo appunto "falciforme"),  $Hb^A/Hb^A$ , vengono infettati più facilmente dal parassita rispetto agli individui eterozigoti, *portatori* dell'anemia falciforme,  $Hb^A/Hb^S$ .

Ciò accade perchè sembra che il parassita della malaria<sup>2</sup>, che trascorre parte del suo ciclo vitale all'interno dei globuli rossi, prosperi meno quando queste cellule contengono l'emoglobina tipica dell'anemia falciforme.

Gli individui *affetti* da anemia falciforme,  $Hb^S/Hb^S$ , pur essendo resistenti alla malaria, spesso muoiono per complicazioni prima che possano riprodursi, per cui la loro resistenza ha conseguenze minime.

---

<sup>1</sup>L'anemia *drepanocitica* è una malattia genetica del sangue in cui i globuli rossi circolanti, in condizioni di bassa tensione di ossigeno, assumono una forma irregolarmente cilindrica spesso ricurva che assomiglia ad una mezzaluna o una falce, per questo motivo l'anemia drepanocitica è anche detta anemia falciforme.

<sup>2</sup>Il *Plasmodium falciparum*, agente eziologico della malaria, che ha un ciclo di vita molto lungo e complesso, non riesce a riprodursi negli eritrociti dei soggetti portatori del gene mutato (sia omozigoti che eterozigoti). Questo succede in quanto gli eritrociti contenenti l'emoglobina mutata E6V hanno una emivita più breve degli eritrociti normali.

Sebbene sia difficile stimare con precisione dai dati disponibili sulle popolazioni quali siano gli svantaggi selettivi dei due omozigoti, succede che circa il 15% degli individui  $Hb^A/Hb^A$  ed il 90% degli  $Hb^S/Hb^S$ , in confronto agli eterozigoti, muore entro il periodo compreso tra la nascita e l'età riproduttiva.

Dati questi valori avremo:

	Genotipo		
	$Hb^A/Hb^A$	$Hb^A/Hb^S$	$Hb^S/Hb^S$
svantaggio selettivo	$s = 0.15$	0	$t = 0.90$
fitness	0.85	1	0.10

Calcoliamo le frequenze di equilibrio:

$$\hat{p} = freq_{eq}(Hb^A) = \frac{t}{s+t} = \frac{0.90}{(0.15+0.90)} = 0.86$$

$$\hat{q} = freq_{eq}(Hb^S) = \frac{s}{s+t} = \frac{0.15}{(0.15+0.90)} = 0.14$$

Si tratta di frequenze alleliche caratteristiche di  $Hb^A$  e  $Hb^S$  in alcune popolazioni africane a contatto con il parassita della malaria.

Haldane, sulla base delle correlazioni geografiche tra la malaria ed un altro difetto dell'emoglobina (la talassemia), ipotizzò per primo che esistesse un'interazione tra la malaria infettiva e quella genetica. Esistono molte prove che confermano il fatto che l'eterozigote (portatore sano) di anemia falciforme in ambienti malarici ha una fitness più alta degli omozigoti sani.

### 3.1.3 Analisi della legge di F-H-W nel caso di selezione debole

L'equazione di FHW è un'equazione alle differenze non lineare a partire dalla quale è possibile ottenere risultati qualitativi e soluzioni analitiche *nell'intorno degli stati di equilibrio*, ma da cui in generale non è possibile ricavare una soluzione per via analitica. Possiamo però partire dall'assunto di essere nel caso di **selezione debole**, ciò risulterà abbastanza veritiero in quanto ciò che ci dicono gli studi paleontologici è che grandi cambiamenti in termini evolutivi avvengono nella misura di migliaia di generazioni, per cui  $s$  sarebbe circa 0.001, il che ci suggerisce di essere in condizioni di buona approssimazione.

Siano  $w_x = w_y = w_z = 1 + O(s)$  con  $O(s)$  termini con andamento lineare in  $s$ , intesa come misura della forza della selezione. Facciamo riferimento all'equazione scritta nella forma

$$\delta p = p' - p = \frac{\alpha_p p}{\bar{w}} = \frac{(w_p - \bar{w})p}{\bar{w}}$$

abbiamo  $w_p = 1 + O(s)$ ,  $\bar{w} = 1 + O(s)$  e  $w_p - \bar{w} = O(s)$ , se  $s$  è piccolo possiamo trascurare i termini  $O(s)$  al denominatore ed approssimare la funzione alle differenze con un'equazione differenziale:

$$\dot{p} = \alpha_p p = (w_p - \bar{w})p$$

esplicitando i termini  $O(s)$  si hanno  $w_x = 1 + hs$ ,  $w_y = 1 + ks$  e  $w_z = 1$  l'equazione diventa un'equazione a variabili separabili

$$\dot{p} = spq((h - k)p + kq)$$

possiamo ricavare il tempo per ogni cambiamento nelle frequenze alleliche da  $p_0$  a  $p_1$  con la seguente integrazione

$$t = \frac{1}{s} \int_{p_0}^{p_1} \frac{dp}{p(1-p)((h-k)p + k(1-p))}$$

Se fossimo nel caso additivo con A semi-dominante e vantaggioso, prendendo  $k = 1$  e  $h = 2$  avremmo la curva logistica

$$\frac{dp}{dt} = sp(1-p) .$$

Fisher ne studiò una versione comprendente un termine dipendente dalla variabile spazio  $x$ :

$$\frac{\partial p}{\partial t} = sp(1-p) + D \frac{\partial^2 p}{\partial x^2} \quad (3.15)$$

dove il termine dipendente dalla variabile spaziale  $x$  rappresenta il moto casuale degli alleli nello spazio inteso come la distribuzione dei geni nella popolazione.

## 3.2 La Mutazione

Nella maggior parte dei casi non è possibile determinare la causa delle mutazioni che avvengono in natura poiché la maggior parte di esse sono rappresentate da variazioni di alleli normali in alleli *recessivi* e *deleterii*; deleteri nel senso che alterano la funzione già adattata all'ambiente. Essi sono assolutamente fondamentali nei processi evolutivi in quanto introducono diversità genetica sulla quale la selezione possa agire. In effetti un gene mutato ha una piccola probabilità di sopravvivenza, inoltre le probabilità che un carattere originato per mutazione venga perduto, aumenta di generazione in generazione, a partire dal momento in cui è avvenuta la mutazione. Fisher dimostrò che dopo  $n$  generazioni la probabilità che un gene possa sopravvivere è approssimativamente  $2/n$ , supposto che  $n$  sia molto grande.

Il fatto che molti nuovi caratteri persistano e anzi aumenti la loro frequenza si spiega con il fatto che una delle cause che portano alla modificazione delle frequenze geniche è proprio la frequenza di mutazione stessa.

Prendiamo ad esempio la mutazione di un gene  $A$  in  $a$  e ipotizziamo che non esista la trasformazione inversa, se siamo in una popolazione di dimensioni costanti e abbiamo un tasso di mutazione persistente, dopo un lungo periodo di tempo  $a$  può arrivare addirittura a sostituirsi ad  $A$ .

Sia  $p_0 = \text{freq}(A)$  la frequenza iniziale di  $A$  e  $u$  tasso di mutazione di  $A$  in  $a$ , dopo la prima generazione si avrà:

$$\begin{aligned}\text{freq}(a) &= u = p_0 \\ \text{freq}(A) &= p_0 - up_0 = p_0(1 - u)\end{aligned}$$

nella generazione successiva compaiono nuovi geni  $a$  con frequenza

$$\begin{aligned}\text{freq}(a) &= u \times p_0(1 - u) = p_0(u - u^2) \\ \text{freq}(A) &= p_0(1 - u) - p_0(u - u^2) = p_0 + 2p_0u + p_0u^2 = p_0(1 - u)^2\end{aligned}$$

alla generazione  $n$ -esima

$$\text{freq}(A) = p_0(1 - u)^n$$

quindi se anche il tasso di mutazione  $u$  di  $A$  in  $a$  fosse piccolo, dopo un tempo abbastanza grande il fattore  $(1 - u)^n$  tenderà a zero, vale a dire che  $A$  tenderà a scomparire.

Non sempre il tasso di mutazione esiste solo in una direzione, possiamo quindi considerare che l'allele  $a$  possa retromutare in  $A$  con la frequenza  $v$ ; siano ora  $p_0 = \text{freq}(A)$  e  $q_0 = \text{freq}(a)$  le frequenze iniziali, si ha che una singola generazione di mutazione produrrà:

$$\begin{aligned}freq(A) &= p_0 + vq_0 \\freq(a) &= q_0 + up_0\end{aligned}$$

Osserviamo che  $a$  ha guadagnato la frazione  $up_0$ , ma ha perso la frazione  $vq_0$ , in altre parole la frequenza di  $a$  subisce una variazione di  $\Delta q = up_0 - vq_0$ ; se abbiamo che  $p$  è relativamente grande e di conseguenza  $q$  è relativamente piccolo,  $\Delta q$  è elevato e  $q$  comincia ad aumentare rapidamente e quindi a far diminuire  $p$  facendo diminuire anche  $\Delta q$ ; c'è un punto in cui  $p$  e  $q$  sono bilanciati nelle loro frequenze di mutazione così che  $\Delta q = 0$  ed è detto **punto di equilibrio di mutazione**:

$$\begin{aligned}\Delta q = 0 &\iff u\hat{p} - v\hat{q} = 0 \\ \Rightarrow u\hat{p} &= v\hat{q} \\ u(1 - \hat{q}) &= v\hat{q} \\ u &= \hat{q}(u + v)\end{aligned}$$

si ottengono in tal modo

$$\hat{q} = \frac{u}{u + v} \quad \text{e con lo stesso procedimento} \quad \hat{p} = \frac{v}{u + v}$$

che ci permettono di notare che quando i tassi di mutazione sono molto simili, cioè  $u = v$ , le frequenze geniche all'equilibrio  $\hat{p}$  e  $\hat{q}$  saranno uguali.

Assumendo di essere in una popolazione verificante le condizioni di Hardy-Weinberg, e ricordando che  $u$  è la probabilità che  $A$  muti in  $a$  e  $v$  la probabilità di "retromutazione", abbiamo che della frazione  $p$  dei gameti  $A$  che ci si aspetta solo  $(1 - u)p$  sono in effetti  $A$ , e della frazione  $q$  dei gameti  $a$  che ci si aspetta solo  $vq$  sono in effetti  $a$ , con  $u$  e  $v$  costanti; quindi dopo una generazione di accoppiamenti casuali si ha,

$$p_{n+1} = (1 - u)p_n + vq_n$$

dove  $(1 - u)p_n$  rappresenta la frazione dei gameti  $A$  che *non* sono mutati, e  $vq_n$  sono i gameti  $a$  retromutati in  $A$ ; ricordando che  $q_n = 1 - p_n$  e  $v = \hat{p}(u + v)$ , possiamo riscriverla come

$$p_{n+1} - \hat{p} = (1 - u - v)(p_n - \hat{p})$$

che ha soluzione

$$p_n = \hat{p} + (p_0 - \hat{p})(1 - u - v)^n .$$

Lo stato di equilibrio  $\hat{p} = \frac{v}{u+v}$  è stabile per valori di  $u$  e  $v$  dell'ordine di  $10^{-6}$  e quindi si può mostrare che il tasso con cui viene raggiunta la frequenza di equilibrio della mutazione è molto lento:

$$(1 - u - v)^n = \frac{p_n - \hat{p}}{p_0 - \hat{p}}$$

passando ai logaritmi si ha:

$$n = \log_{(1-u-v)} \frac{p_n - \hat{p}}{p_0 - \hat{p}} = \frac{\ln\left(\frac{p_n - \hat{p}}{p_0 - \hat{p}}\right)}{\ln(1 - u - v)}$$

dove  $n$  è il numero di generazioni richieste per raggiungere una frequenza  $p_n$  partendo da una frequenza  $p_0$ .

### 3.3 Equilibrio tra Selezione e Mutazione

Finora abbiamo considerato gli effetti derivanti da selezione e mutazione che si hanno sulle modificazioni delle frequenze geniche separatamente, mentre nella realtà i valori delle frequenze geniche sono influenzati da entrambi questi fattori simultaneamente. Ad esempio si consideri che se anche un gene recessivo fosse sfavorito nella condizione omozigote, esso potrebbe tuttavia persistere in una popolazione grazie alla sua frequenza di mutazione; ciò che avviene in questi casi è che viene raggiunto un certo punto di equilibrio nel quale il numero di geni che vengono tolti con la perdita degli omozigoti vengono sostituiti dallo stesso numero di geni introdotti negli eterozigoti in seguito ad un processo di mutazione.

Se ora includiamo la **selezione** (intesa come differenza nei tassi di sopravvivenza dei genotipi) nella nostra trattazione riguardante il fenomeno della mutazione, arriveremo a determinare le frequenze di equilibrio tra *mutazione e selezione*.

Ci basta considerare il rapporto delle frequenze geniche all'inizio della fase riproduttiva come  $w_x p^2 + w_y pq : w_y pq + w_z q^2$  o equivalentemente come  $w_p p : w_q q$  invece che semplicemente  $p : q$  e tenere conto dell'effetto della mutazione mediante i coefficienti  $(1 - u)$  e  $v$ :

l'equazione (2.1) diventa

$$p' = (1 - u) \frac{w_p p}{\bar{w}} + v \frac{w_q q}{\bar{w}}$$

quindi

$$\begin{aligned} \delta p &= p' - p = (1 - u) \frac{w_p p}{\bar{w}} + v \frac{w_q q}{\bar{w}} - p = \\ &= \frac{w_p p}{\bar{w}} - u \frac{w_p p}{\bar{w}} + v \frac{w_q q}{\bar{w}} - \frac{\bar{w} p}{\bar{w}} = \\ &= \frac{\alpha_p p}{\bar{w}} - u \frac{w_p p}{\bar{w}} + v \frac{w_q q}{\bar{w}} \end{aligned}$$

Quindi si può dire che quando siamo nel caso di selezione debole, quindi avendo  $w_i = 1 + O(s)$  per ogni genotipo  $i$  e per  $s$  piccola, nell'ipotesi che anche i coefficienti di mutazione  $u$  e  $v$  siano molto piccoli possiamo riscrivere l'espressione mettendo in evidenza piccole variazioni:

$$\delta p = \delta p_{selez} + \delta p_{mutaz}$$

dove i contributi di selezione e mutazione appaiono separati.

Si può dire che uno stato di equilibrio cresce nel momento in cui gli effetti di selezione e mutazione si eliminano a vicenda, vediamo come: consideriamo il caso di selezione contro un allele recessivo e svantaggioso  $a$  con frequenza  $q$ , abbiamo visto che la modificazione della sua frequenza genica ( $\Delta q$ ) per una generazione corrisponde alla perdita di

$$\frac{sq^2(1-q)}{1-sq^2}.$$

Se  $s$  è piccolo si può approssimare a 1 il denominatore, quindi si avrà  $\Delta q = sq^2(1-q)$ ; si ha che la frequenza dei geni  $a$  neomutati sarà  $u(1-q)$ , dove  $u$  è il tasso di mutazione di  $A \rightarrow a$  e  $1-q$  la frequenza di  $A$ . Quindi la perdita dei geni  $a$  data dalla selezione è esattamente bilanciata dal guadagno dei geni  $a$  dato dalla mutazione, ossia:

$$\begin{aligned} sq^2(1-q) &= u(1-q) \\ sq^2 &= u \\ q &= \sqrt{\frac{u}{s}} \end{aligned}$$

Risulta evidente che quando il tasso di mutazione aumenta, la frequenza genica aumenta anch'essa, mentre quando il tasso di selezione aumenta la frequenza genica diminuisce.

Per un carattere recessivo letale ( $s = 1$ ) si avrà che

$$q = \sqrt{u} \quad \text{cioè} \quad q^2 = u$$

in pratica la frequenza degli omozigoti letali all'equilibrio è circa uguale alla frequenza dei nuovi geni introdotti con la mutazione.

Nel caso di selezione contro un allele dominante e svantaggioso con frequenza  $p$ , si aveva che (approssimando a 1 il denominatore)  $p$  si riduce di un fattore di

$$sp(1-p)^2 ,$$

quindi all'equilibrio si avrà:

$$\begin{aligned} sp(1-p)^2 &= v(1-p) \\ p(1-p) &= \frac{v}{s} \end{aligned}$$

dal momento che  $q$  (o  $1-p$ ) è la frequenza dell'allele recessivo e il tasso di mutazione  $v$  di  $a \rightarrow A$  produce nuovi alleli dominanti con frequenza  $v(1-p)$ . Di solito ci si aspetta che i valori di  $p$  siano piccoli quando il gene dominante viene sottoposto a selezione contraria, quindi  $1-p$  può essere approssimato con 1 e quindi il valore di equilibrio è

$$p = \frac{v}{s} .$$

Se siamo nel caso di dominante letale ( $s = 1$ ) la frequenza dell'allele dominante nella popolazione è uguale al suo tasso di mutazione. Come prima si ha che se il tasso di selezione aumenta, la frequenza genica diminuisce, mentre quando il tasso di mutazione aumenta, la frequenza genica aumenta proporzionalmente ad essa.

# Appendice A

## Alcune nozioni di Genetica

### A.1 La nascita della Genetica

La genetica è nata nel 1900 con la riscoperta degli esperimenti di Gregor Mendel da parte di alcuni ibridatori di specie vegetali. Sebbene fosse stato pubblicato nel 1866, il fondamentale lavoro di questo monaco ceco era stato completamente ignorato per più di quarant'anni.

Mendel aveva incrociato alcune linee pure di pisello che presentavano una serie di caratteri opposti (seme liscio o rugoso, pianta alta o nana, fiore bianco o rosa ecc.) e in base ai risultati ottenuti in questi esperimenti era arrivato a formulare le regole in base alle quali questi caratteri vengono trasmessi alla generazione successiva. In particolare osservò che i fattori responsabili dei caratteri sono sempre presenti in coppia, ma solo uno dei due membri della coppia viene trasmesso alla generazione successiva; inoltre concluse che essi vengono ereditati come unità separate, ognuno indipendentemente dagli altri. I fattori ereditari discreti ipotizzati da Mendel vennero in seguito chiamati geni.

### A.2 I cromosomi

Tutti gli organismi viventi sono formati da cellule; si calcola che in ognuno di noi ne esistano circa 100 mila miliardi. In ogni cellula è presente un nucleo, all'interno del quale sono racchiusi i cromosomi. Ogni cromosoma è formato da un lunghissimo filamento di DNA.

I cromosomi sono strutture visibili al microscopio come piccoli bastoncini formati da filamenti di DNA avvolto. Normalmente ciascuna cellula umana possiede 46 cromosomi, di cui 23 sono ereditati dalla madre e 23 dal padre. Il nostro corredo cromosomico è composto come segue:

1. Due cromosomi sessuali: il cromosoma X e il cromosoma Y. Le femmine possiedono due copie del cromosoma X (XX) mentre i maschi possie-

dono un cromosoma X e un cromosoma Y (XY), quindi il cromosoma Y è per forza di cose di origine paterna.

2. Ventidue copie di cromosomi non sessuali, detti *autosomi*, uguali a due a due. Questo fa sì che ciascun gene dell'organismo sia presente in due copie, una di origine materna e l'altra di origine paterna. Ognuna delle due copie è detta allele.

Non sempre gli alleli sono uguali fra loro, anzi molto spesso presentano delle differenze. Gli alleli sono quindi versioni diverse dello stesso gene. Fanno eccezione i geni contenuti nel cromosoma X e Y, che sono presenti in una sola copia nel maschio, in due copie nella femmina.

### A.3 Il DNA

Il DNA, o acido deossiribonucleico, è la sostanza che costituisce il patrimonio genetico ed è contenuto in tutte le cellule dello stesso individuo. È formato da due filamenti disposti a spirale a formare una doppia elica. Ciascun filamento è formato dalla successione di 4 diversi elementi chiamati basi nucleotidiche: Adenina, Guanina, Citosina, Timina (abbreviate A, G, C, T) chiamate anche basi azotate. Queste 4 basi sono le lettere che formano l'alfabeto della vita in quanto il nostro patrimonio genetico è formato dalla successione delle lettere A, G, C, T. Nel genoma umano queste 4 lettere si succedono per ben 3,2 miliardi di volte. Le 4 basi azotate sono localizzate su filamenti opposti e si accoppiano sempre in maniera fissa, cioè la guanina sempre con la citosina e la timina sempre con l'adenina. Ogni coppia viene definita un paio di basi di DNA. Le basi azotate, insieme ad altre sostanze chimiche, costituiscono i nucleotidi (unità degli acidi nucleici, costituita da uno zucchero, una base azotata e un gruppo fosfato). Il DNA funziona come il codice binario di un computer il quale, basandosi sulle combinazioni dei numeri 1 e 0, consente di compiere innumerevoli operazioni. Allo stesso modo, i possibili abbinamenti delle 4 basi azotate del DNA sono in grado di codificare tutte le diverse proteine. Il DNA ha la capacità di fungere da stampo per replicare se stesso: ognuno dei filamenti della doppia elica infatti è complementare all'altro, come una foto e il proprio negativo. Quando una cellula si duplica, i due filamenti della doppia elica si separano e ognuno fa da stampo per la sintesi di un filamento complementare. Alla fine del processo, avremo così due molecole di DNA, identiche fra loro.

## A.4 I geni

I geni sono l'unità più piccola di materiale ereditario e contengono l'informazione necessaria per determinare un singolo carattere di un individuo, come ad esempio il colore degli occhi, la struttura delle ossa, il gruppo sanguigno ecc... Ogni gene è un segmento di DNA (generalmente di oltre 1000 paia di basi) e contiene le istruzioni per fabbricare una determinata proteina. È dunque chiaro che all'interno di ciascun cromosoma si trovano migliaia di geni. I geni sono le unità funzionali del patrimonio genetico: se paragoniamo il genoma ad un'enciclopedia, ogni gene rappresenta una singola frase. Su ogni cromosoma sono localizzati molti geni, la posizione dei quali su ciascun cromosoma è detta *locus*. Il Locus indica la posizione in cui risiede un gene su un cromosoma. Il braccio corto del cromosoma è definito "p" (petit) e quello lungo "q" (queue). Ogni braccio è poi suddiviso in regioni che sono indicate come p1, p2, p3 oppure q1, q2, q3 ecc... a seconda che ci si riferisca al braccio corto o a quello lungo, con progressione numerica a partire dal centromero (la regione centrale del cromosoma) verso il telomero (l'estremità del cromosoma). Ogni regione è divisa in bande (p11, p12 ecc...), sottobande (p11.1, p12.1 ecc...), sotto-sotto bande (p11.11, p11.22 ecc...) e così via. Il processo per il quale una cellula madre si divide in due cellule figlie identiche si chiama *mitosi* e comporta la duplicazione di ciascun cromosoma e la separazione delle due copie nelle due cellule figlie. Questo processo assicura che ciascuna delle due cellule figlie abbia lo stesso numero di cromosomi e, dunque, di geni della cellula madre. Gli organismi superiori si riproducono invece per via sessuata e derivano dall'unione di due gameti, che vengono prodotti da una variante della divisione mitotica: la *meiosi*. Questa differisce dalla mitosi per il fatto che a ciascuna delle nuove cellule viene trasmesso solo un singolo cromosoma di ciascuna coppia. In questo modo ciascun gamete viene a contenere metà dei cromosomi presenti nelle cellule somatiche. Quando avviene la fecondazione, i due gameti si uniscono a formare lo zigote e viene così restaurato il numero intero di cromosomi tipico della specie, proveniente per metà da un genitore e per metà dall'altro.

### La trasmissione dei geni

L'unione dei gameti riunisce due corredi genici. Per ciascun locus genico, responsabile di un particolare carattere (ad esempio, il colore degli occhi), nel nucleo di tutte le cellule somatiche di ogni individuo sono presenti due copie, una di origine materna e una paterna (eccezioni a questa regola sono rappresentate dai cromosomi sessuali e dai caratteri legati al sesso). Ciascuna copia si trova nella stessa posizione, su ciascuno dei due cromosomi dello stesso tipo che generalmente stanno anche appaiati.

## Omozigosi ed Eterozigosi

Quando le due copie del gene sono identiche, si dice che l'individuo è omozigote per quel particolare gene. Molto spesso, tuttavia, queste due copie sono diverse: ad esempio, quando un genitore ha gli occhi azzurri e l'altro ha gli occhi marroni, nella discendenza possono essere presenti due copie diverse del gene, una per ciascun colore. Se questo è il caso, si dice che l'individuo è eterozigote per quel gene. Le diverse copie di uno stesso gene sono dette alleli e occupano lo stesso locus sui cromosomi omologhi. Per ciascun gene può esistere un gran numero di alleli.

## Alleli dominanti e recessivi

Nonostante nelle cellule somatiche per ogni gene siano sempre presenti due alleli, generalmente solo uno dei due è manifesto e maschera l'altro; il primo viene detto dominante e il secondo recessivo. Come dimostrato da Mendel, il carattere recessivo ricompare, nelle generazioni successive, negli individui omozigoti per l'allele considerato. Per convenzione gli alleli sono indicati da una singola lettera, maiuscola per indicare l'allele dominante (ad esempio A) e minuscola (ad esempio a) per l'allele recessivo. Gli eterozigoti (Aa) e gli omozigoti (AA) per un determinato gene mostrano l'effetto dell'allele dominante, mentre gli omozigoti (aa) mostrano l'effetto dell'allele recessivo. Da quanto descritto, è quindi molto importante distinguere tra l'apparenza esterna di un organismo e i geni e gli alleli che esso trasporta. L'insieme dei caratteri espressi costituisce il fenotipo dell'organismo, mentre il corredo genetico (comprendente anche gli alleli recessivi mascherati dai dominanti) rappresenta il genotipo. Ciascun figlio di due genitori eterozigoti (Aa) ha una probabilità su quattro di essere omozigote (AA), una su due di essere eterozigote (Aa) e una su quattro di essere omozigote (aa). Si noti che dire che ciascun figlio ha un quarto di probabilità di essere omozigote recessivo non equivale a dire che un quarto dei figli sarà omozigote recessivo dal momento che le nascite sono, in senso probabilistico, eventi indipendenti.

## Ereditarietà dei caratteri quantitativi ed ereditarietà poligenica

Il *genotipo* rappresenta la costituzione genetica dell'individuo e viene rivelato dall'analisi di tipo molecolare, mentre il *fenotipo* l'insieme delle caratteristiche morfologiche e fisiologiche di un individuo. Nella stragrande maggioranza dei casi i caratteri fenotipici non sono controllati da un singolo gene: solitamente un gene può influire su più di un carattere e un carattere può dipendere dall'azione di più geni. Ad esempio, per produrre il pigmento porpora nel fiore del pisello odoroso occorre l'azione di almeno due geni dominanti. Caratteri come ad esempio il peso, l'altezza, il grado di pigmentazione, che nei diversi individui presentano una gamma di variazioni quantitative continua e molto estesa, in genere dipendono da un gran numero di geni, i singoli effetti

dei quali sembrano sommarsi gli uni agli altri. L'altezza di una pianta, ad esempio, potrebbe essere determinata da una serie di quattro geni: A, B, C e D. Si supponga che una pianta abbia un'altezza media di 25 cm quando il suo genotipo è aabbccdd e che ogni sostituzione con una coppia di alleli dominanti aumenti l'altezza media di circa 10 cm; in quel caso una pianta che è AABBccdd sarà alta 45 cm e una che è AABBCCDD sarà alta 65 cm. In realtà, i risultati sono raramente così regolari: geni differenti possono dare contributi differenti al valore totale e alcuni geni possono interagire in modo che il contributo di uno dipenda dalla presenza di un altro. L'eredità di caratteri quantitativi dipendenti da più geni viene detta eredità poligenica. Inoltre, quando il fenotipo è determinato, oltre che dal materiale genetico, anche da una componente ambientale, l'eredità viene detta multifattoriale.

# Bibliografia

- [1] N.F. Britton. *Essential Mathematical Biology*. Springer Ed. SUMS.
- [2] Monroe W. Strickberger, *Trattato di Genetica*. Ed. Piccin.
- [3] Arthur P. Mange e Elaine Johansen Mange, *Genetica e l'Uomo*. Ed. Zanichelli.