

Genetica della Conservazione

Lezione 2

Lo studio della diversità genetica

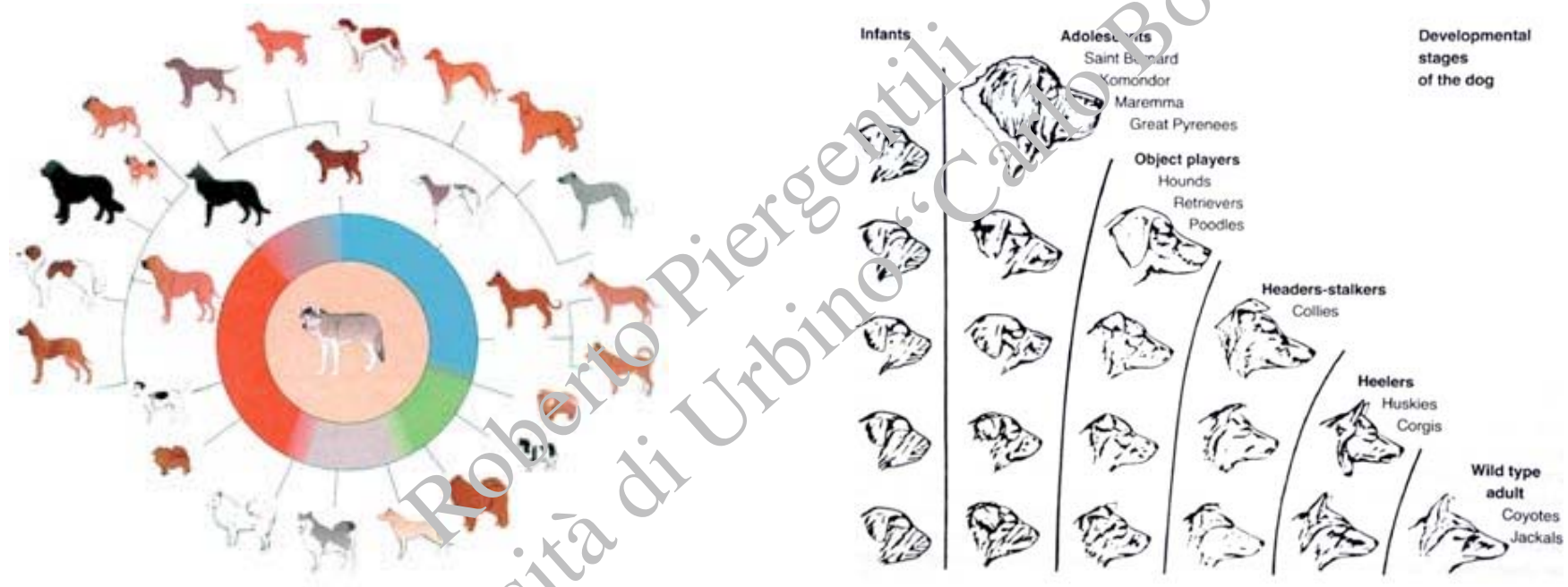
Roberto Piergentili
Università di Urbino "Carlo Bo"

La diversità genetica

E' la **varietà di alleli e genotipi** presenti in un gruppo in esame (popolazione, specie, gruppo di specie).

Viene descritta mediante l'analisi di polimorfismi, eterozigosità media e diversità allelica.

La diversità genetica



Un caso di elevata diversità genetica è quello delle razze canine, tutte derivanti dal lupo e con esso interfeconde.

Definizione: polimorfismo

Un **locus** è detto **polimorfico** se, in una specie, sono presenti **due o più alleli**. Di solito l'allele più frequente in questi casi ha una frequenza inferiore al 99% (diminuita al 95% per minimizzare i problemi con dimensioni diverse dei campioni analizzati). Se è presente **un solo allele**, il locus è **monomorfo** la diversità genetica per quel locus è zero.

La **proporzione di loci polimorfici (P)** è il rapporto tra numero di loci polimorfici e numero totale di loci campionati. Esempio: se su 10 loci campionati, 3 sono polimorfici, si ha $P=3/10=0,3$.

Si usa per calcolare l'**eterozigosità media**.

Definizione: eterozigosità media (H)

È la somma delle proporzioni degli eterozigoti per tutti i loci fratto il numero totale di loci campionati.

Esempio: le proporzioni di eterozigoti per 10 loci in una popolazione (calcolati come detto prima) è: 0,2-0,4-0,1-0-0-0-0-0-0-0 (cioè $P=0,3$). Quindi si ha:

$$H=(0,2+0,4+0,1+0+0+0+0+0+0+0)/10=0,07$$

He: eterozigosità media attesa meno sensibile alla dimensione del campione

Ho: eterozigosità media osservata dipende dal campione

In popolazioni con accoppiamento casuale $H_e=H_o$

Definizione: diversità allelica (A)

È il numero medio di alleli per locus.

Esempio: in 10 loci analizzati sono stati trovati i seguenti numeri per gli alleli riscontrati nel campione: 2-3-2-1-1-1-1-1-1-1 (cioè sempre $P=0,3$).

$$A=(2+3+2+1+1+1+1+1+1+1)/10=1,4$$

Definizione: popolazione mendeliana

Una **popolazione** si dice **di tipo mendeliano** se è costituita da un gruppo di individui interfertili che condividono una serie di geni la cui trasmissione segue le leggi di Mendel. I geni condivisi dagli individui di una popolazione mendeliana costituiscono il **pool genico**.

L'equilibrio di Hardy-Weinberg

Le frequenze alleliche, e di conseguenza quelle genotipiche, non variano nel tempo. Si instaura così un equilibrio.

La legge può essere applicata quando la popolazione soddisfa **contemporaneamente** le seguenti cinque condizioni:

deve essere **infinitamente grande**;

gli **accoppiamenti** al suo interno devono essere **casuali** (*panmissia*);

ogni allele deve avere la **stessa *fitness***, ovvero la stessa capacità di essere trasferito alle generazioni successive;

le velocità di **acquisizione e perdita di nuovi alleli** devono essere uguali;

non deve esserci **migrazione**, né in uscita né in entrata, con altre popolazioni.

Calcolo delle frequenze alleliche

Nel caso sia possibile riconoscere in qualunque modo gli individui eterozigoti (codominanza, dominanza incompleta, ecc.), e se si considerano solo due alleli ad un *locus*, se per esempio la nostra popolazione, su 100 individui, ne avesse $X/100$ di tipo AA , $Y/100$ di tipo Aa e $Z/100$ di tipo aa , allora la **frequenza p** dell'allele A sarà $(2X+Y)/200$ e quella **q** di a sarà $(2Z+Y)/200$, o anche **$q=1-p$** *. Si usano $2X$ e $2Z$ perché ciascun omozigote, essendo diploide, porta due copie di ciascun allele, mentre gli eterozigoti ne portano uno solo. Inoltre, per lo stesso motivo, 100 individui porteranno in tutto 200 alleli, due per ciascun individuo.

* $q=1-p$ perché $p+q=1$, cioè la somma dei due alleli dà la totalità di alleli segreganti in quel *locus*, in quella popolazione.

L'esempio dei gruppi sanguigni M/N

Nell'uomo i gruppi sanguigni M/N sono **codominanti** ed esistono solo due alleli al *locus* L. Sarà allora:

gruppo sanguigno	genotipo	N° di individui	N° di alleli L ^M	N° di alleli L ^N
M	L ^M L ^M	49	98	---
MN	L ^M L ^N	42	42	42
N	L ^N L ^N	9	---	18
totali		100	140	60

$$p = 98+42/200 = 0,7 \quad q = 1-p = 0,3$$

Le frequenze genotipiche

Possono venire calcolate in base alle frequenze alleliche.

Sia p la frequenza dell'allele A e q la frequenza dell'allele a

In totale si avrà:

individui $AA = p^2$

individui $Aa = 2pq$

individui $aa = q^2$

ovvero $(p+q)^2!$

Eggs:

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

Sperm:

Dalle frequenze genotipiche alle frequenze alleliche,

Riconsiderando l'esempio precedente, si avrà:

uova

$$L^M(0.7) = p \quad L^N(0.3) = q$$

spermi

L^M (0.7) = p	$L^M L^M$ (0.49) = p^2	$L^M L^N$ (0.21) = pq
L^N (0.3) = q	$L^M L^N$ (0.21) = pq	$L^N L^N$ (0.09) = q^2

generazione
successiva

$$L^M L^M = 0.49 \quad L^M L^N = 0.42 \quad L^N L^N = 0.09$$

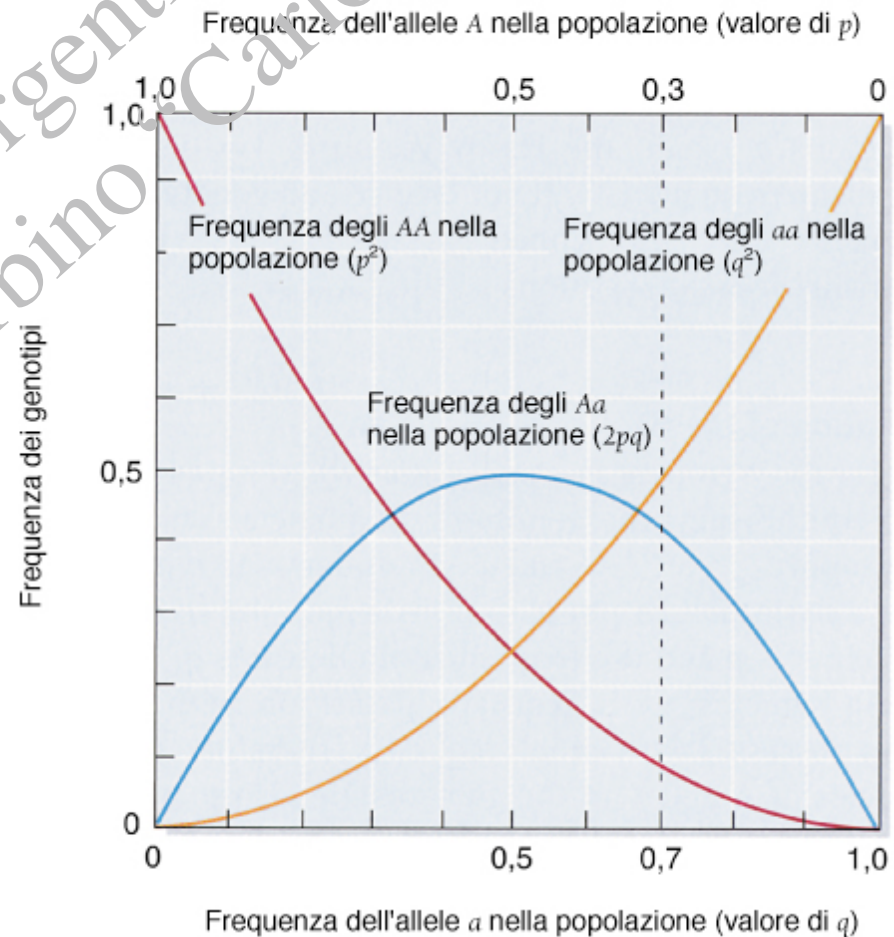
Relazione alleli / genotipi

Se si considerano **DUE** alleli:

1. quando questi tendono ad avere la stessa frequenza, il genotipo più frequente è l'eterozigote;
2. la frequenza degli eterozigoti non può MAI essere $>50\%$!

Figura 22.3

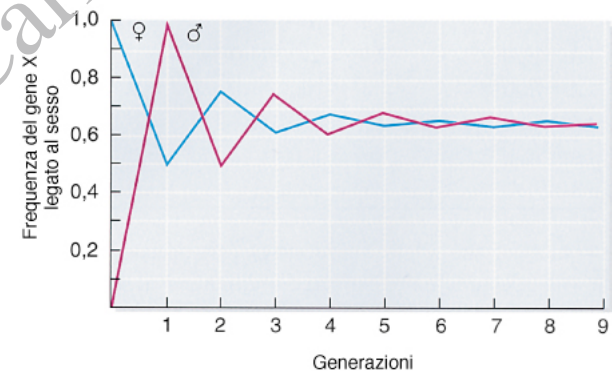
Relazione tra le frequenze dei genotipi AA , Aa ed aa e le frequenze degli alleli A ed a (in termini dei rispettivi valori p [ascissa superiore] e q [ascissa inferiore]) in una popolazione che rispetti le assunzioni della legge di Hardy-Weinberg. Ogni singola popolazione è definita da una linea verticale, come quella indicata in figura caratterizzata da $p = 0,3$ e $q = 0,7$.



I geni legati al sesso

La formula precedente $(p+q)^2$ vale sia per la popolazione intera, sia per ciascuna sottopopolazione considerata in un dato momento in base ad un dato criterio: ad esempio varrà sia se consideriamo solo i maschi sia che consideriamo solo le femmine, e le frequenze alleliche e genotipiche per geni autosomici saranno identiche nei 2 sessi. Tuttavia i maschi di mammifero (e di molte specie animali) sono **emizigoti**, quindi **non possono essere eterozigoti**. Questo significa che **nei maschi le frequenze alleliche sono uguali alle frequenze genotipiche** perché ogni volta che un allele c'è, si manifesta anche se recessivo. Per questo motivo i caratteri recessivi legati all'X sono più frequenti tra i maschi che non tra le femmine.

Figura 22.4
Rappresentazione dell'avvicinamento graduale all'equilibrio di un gene associato al cromosoma X con una frequenza iniziale di 1 nelle femmine e di 0 nei maschi.



Quando la popolazione esce dall'equilibrio

La legge di Hardy-Weinberg dice che, all'equilibrio, **le frequenze alleliche e genotipiche non cambiano nel tempo**. Ma se per un qualunque motivo tale equilibrio venisse alterato, non appena la popolazione torna alle condizioni che soddisfano la legge, essa si posizionerà, **nel giro di una sola generazione, in un nuovo equilibrio** (altrettanto costante) i cui valori dipenderanno dagli eventi accaduti durante le cause dello squilibrio. Quando le condizioni di equilibrio non sono rispettate, le frequenze alleliche cambiano e si va verso l'**evoluzione** (cioè il *cambiamento*) della popolazione.

Diversità attesa nelle popolazioni

La diversità genetica per un singolo locus è caratterizzata da:

- Eterozigosità attesa (H_e)
- Eterozigosità osservata (H_o)
- Diversità allelica (A)

Se gli alleli sono due, allora $H_e=2pq$ (**eterozigosità genica**).

Se gli alleli sono più di due, $H_e=1-(\text{somma dei quadrati delle frequenze alleliche})$ secondo la formula:

$$H_e = 1 - \sum_{i=1}^{\text{\#alleli}} p_i^2$$

restano nel calcolo solo gli eterozigoti!

Un esempio pratico

Dati 3 alleli ad un locus ($A1-A2-A3$), con frequenze rispettivamente di 0,364-0,352-0,284, calcolare l'eterozigosità attesa secondo l'equilibrio di Hardy-Weinberg.

Sappiamo che per 3 alleli si ha: $p^2+q^2+r^2+2pq+2qr+2pr=1$

$$H_e=2pq+2qr+2pr=1-(p^2+q^2+r^2)$$



$$H_e=1-(0,364^2+0,352^2+0,284^2)=0,663$$

Come confronto, si verifica facilmente che $H_o=0,659$, perciò

$H_e=H_o$ la popolazione è ad accoppiamento casuale.

Calcolo delle frequenze partendo da un allele recessivo

Poiché gli individui con un carattere recessivo sono genotipicamente a/a , e poiché la loro frequenza è q^2 , calcolare i valori delle frequenze alleliche e genotipiche risulta relativamente semplice. Tuttavia facendo ciò si dà per scontato che (1) gli accoppiamenti siano casuali, (2) sia assente la selezione, (3) sia assente la migrazione, ecc., quindi **questo sistema non dovrebbe mai essere usato per i loci con i genotipi in qualunque modo (per esempio, molecolare) distinguibili.**

I caratteri quantitativi

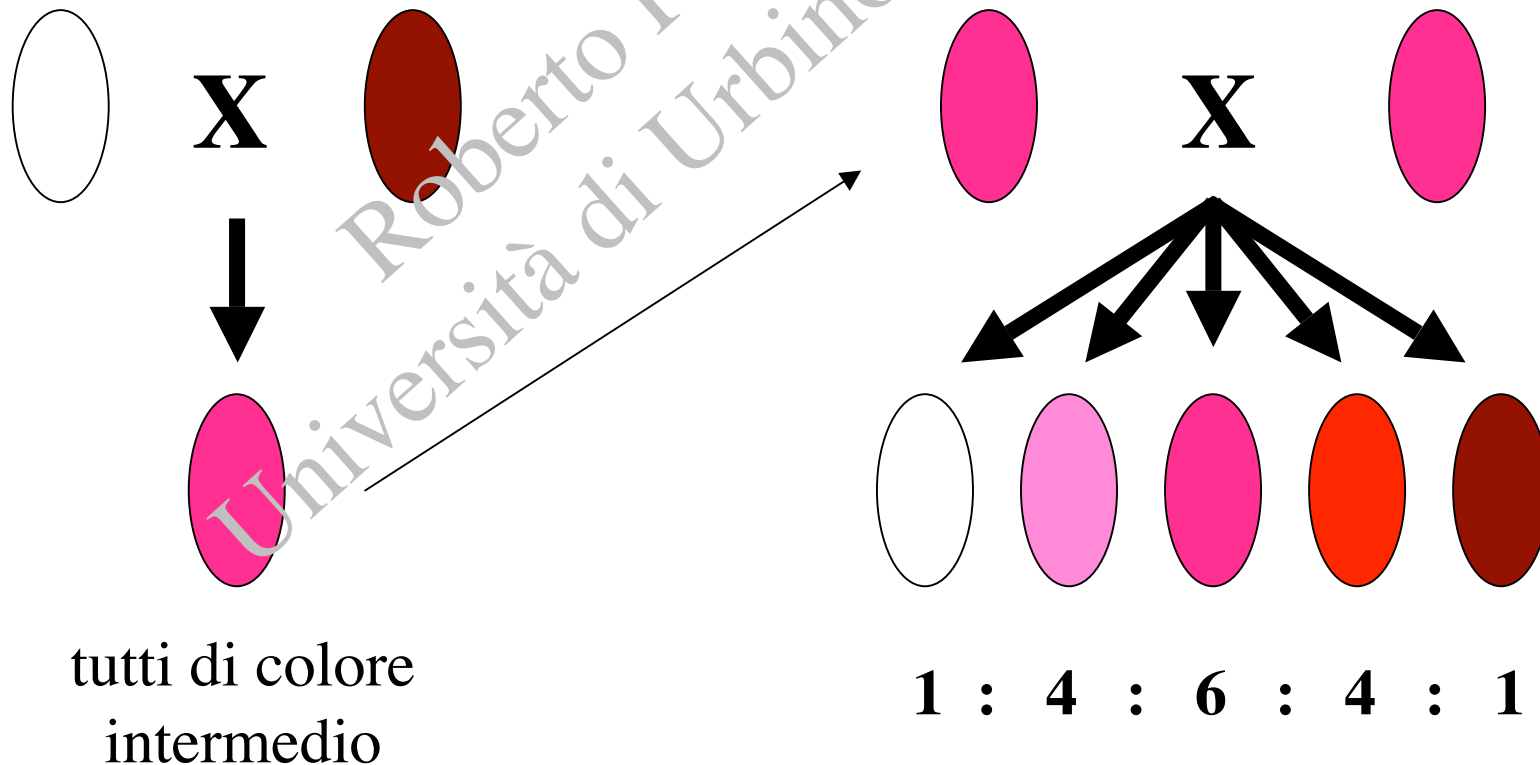
I caratteri quantitativi, o *complessi*, sono caratteri per cui la variazione fenotipica è distribuita in maniera **continua** nelle popolazioni naturali. Sono caratterizzati da una distribuzione statistica molto vicina alla distribuzione *normale* (gaussiana). I caratteri quantitativi includono caratteri morfologici (peso, altezza), fisiologici (pressione sanguigna, forza muscolare), comportamentali (aggressività, intelligenza), ma anche molecolari (livello di espressione genica, livelli intraematici di macromolecole quali il colesterolo, quantità di melanina).

Variabilità quantitativa

- I caratteri quantitativi sono determinati da **molti loci** (QTL, Quantitative Trait Loci), sono influenzati dal background genetico (altri loci, sesso, ecc.) e dalla sensibilità all'ambiente.
- Ogni singolo QTL normalmente influisce poco sul fenotipo finale (eredità per mescolamento, pre-mendeliana). In generale, **non c'è una relazione evidente tra genotipo e fenotipo.**
- La variabilità quantitativa è spesso alla base dell'evoluzione delle popolazioni naturali e della selezione artificiale fatta dall'uomo (vedi razze canine).

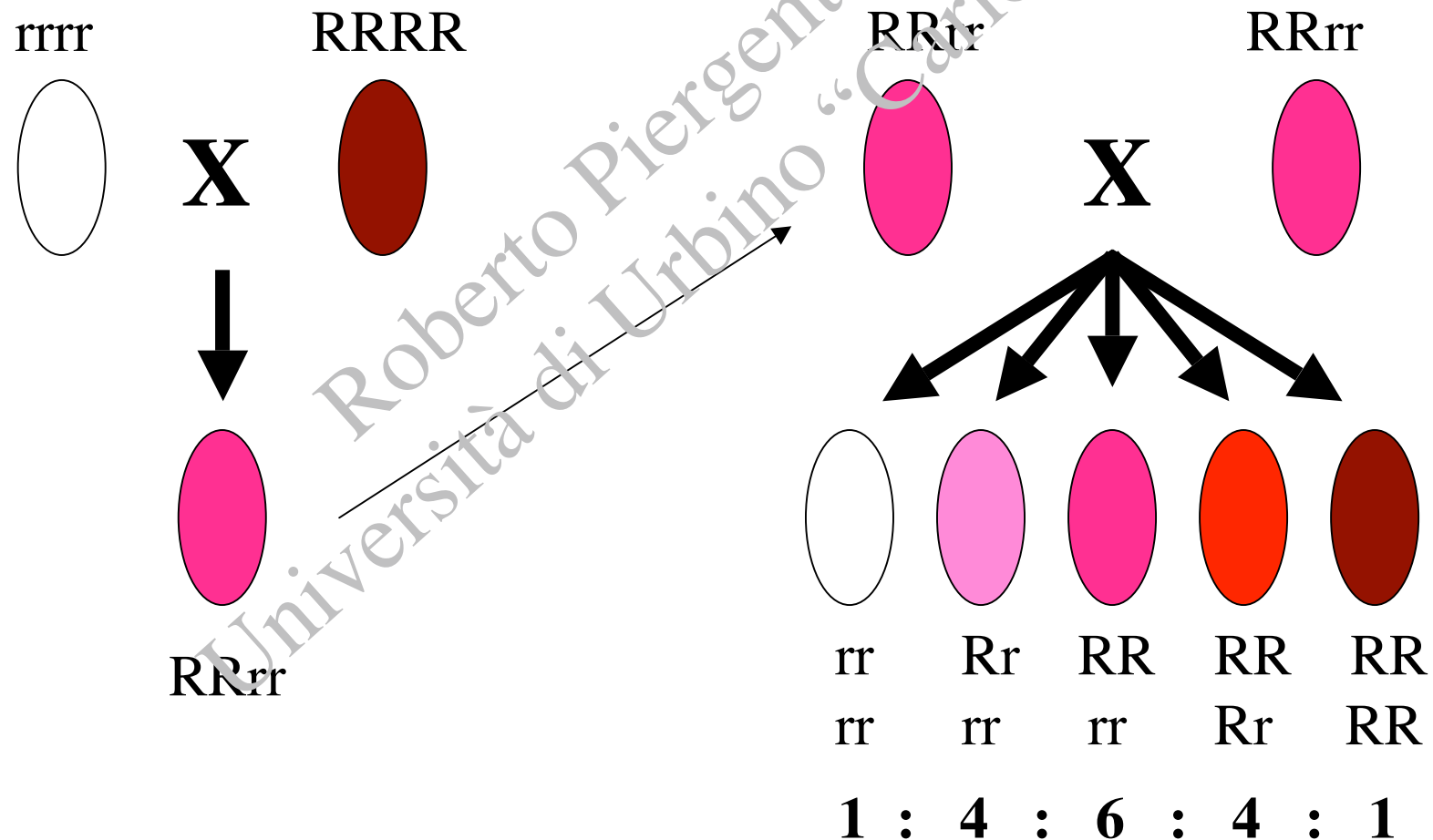
Un esempio pratico

La cariosside di grano può essere bianca, rossa intensa, oppure di sfumature intermedie. Nel 1908 Nilsson-Ehle ottenne i seguenti risultati:



Spiegazione: duplicazione genica

L'allele R contribuisce per il colore in maniera quantitativa (somma), l'allele r no. Per cui si ha:



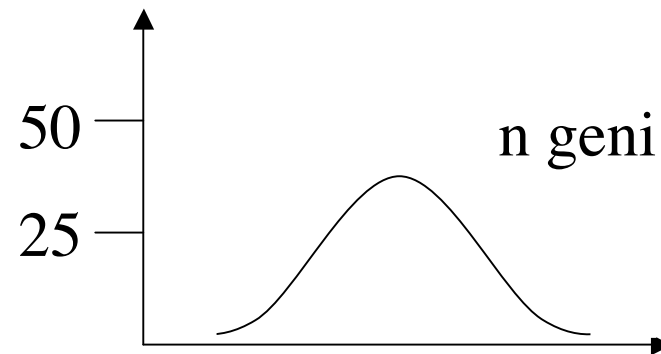
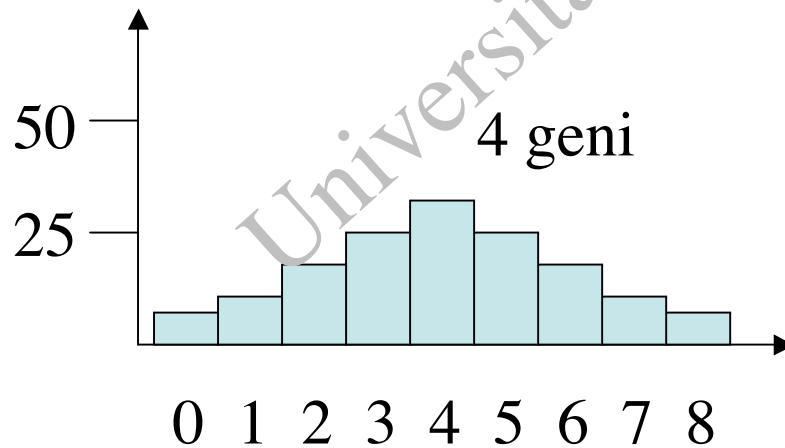
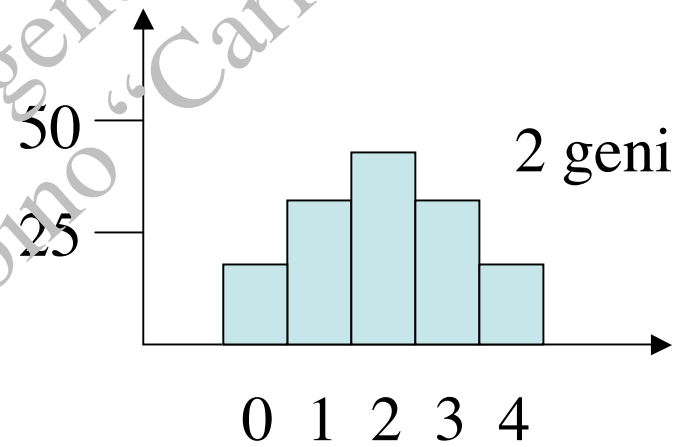
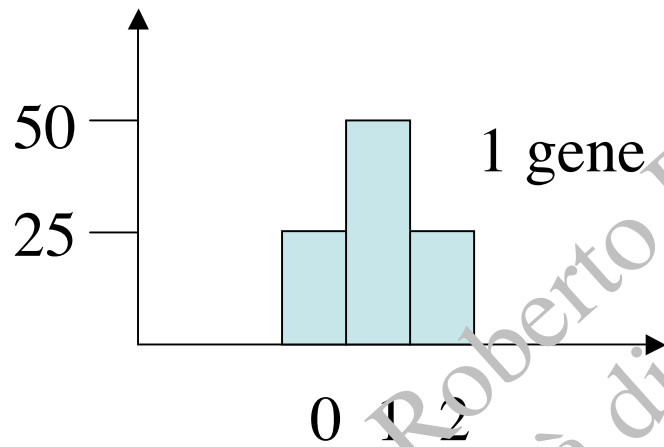
Infatti...

gameti	r1r2	R1r2	r1R2	R1R2
r1r2	r1r1 r2r2	R1r1 r2r2	r1r1 R2r2	R1r1 R2r2
R1r2	R1r1 r2r2	R1R1 r2r2	R1r1 R2r2	R1R1 R2r2
r1R2	r1r1 R2r2	R1r1 R2r2	r1r1 R2R2	R1r1 R2R2
R1R2	R1r1 R2r2	R1R1 R2r2	R1r1 R2R2	R1R1 R2R2

Generalizzando...

Asse X: numero di geni che incrementano il carattere;

Asse Y: frequenze del fenotipo, in percentuale



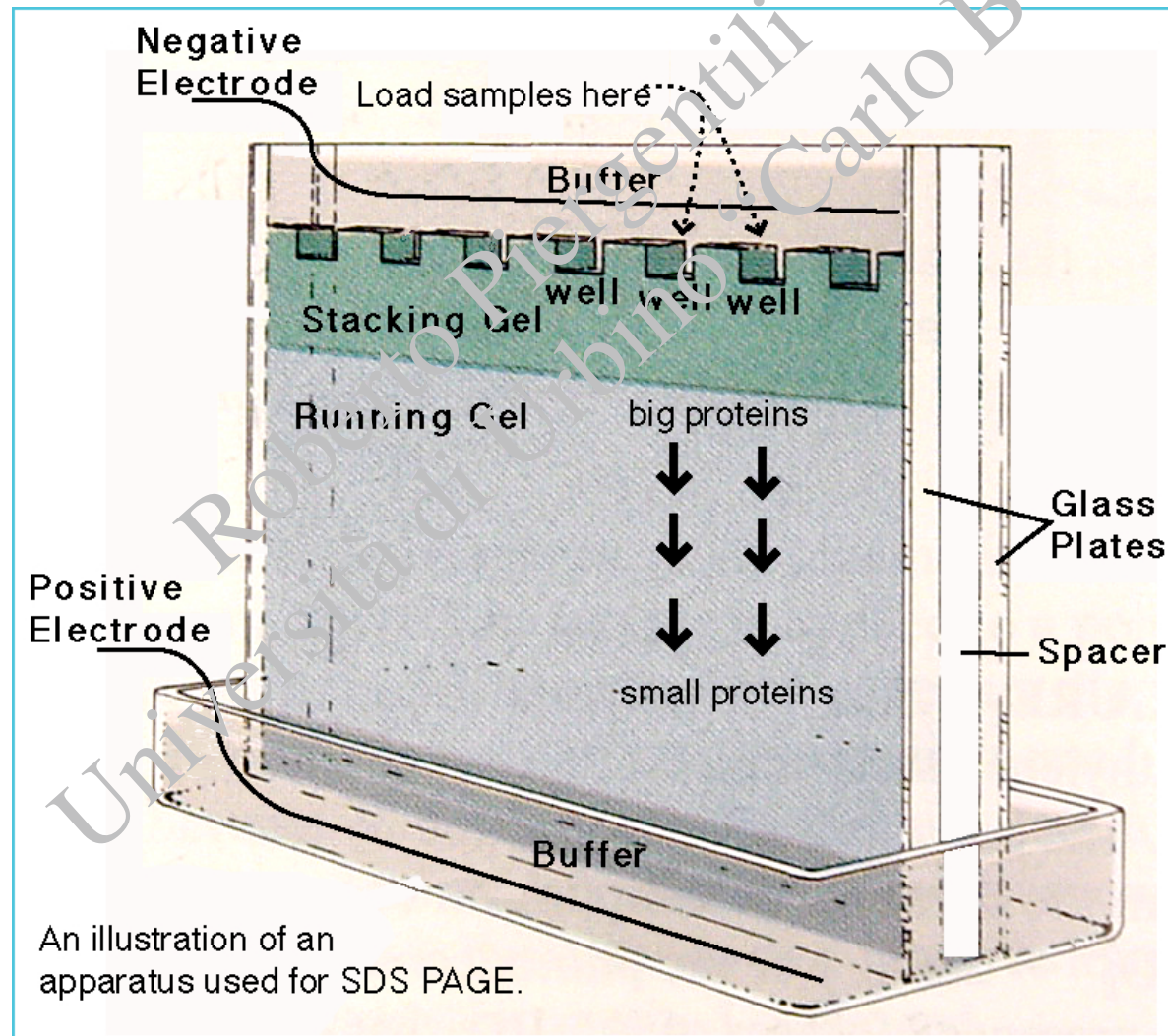
Gli alleli deleteri

La diversità genetica dovuta ad alleli deleteri è critica nella Genetica della Conservazione in quanto essi riducono la vitalità e la fitness riproduttiva se si trovano in omozigosi in seguito a inincrocio.

Gli alleli deleteri vengono continuamente creati *ex novo* per mutazione, e altrettanto continuamente persi per selezione naturale (il bilancio nelle popolazioni di tipo mendeliano resta zero).

Il numero di alleli deleteri rari fissati nella popolazione (**carico mutazionale**) per le popolazioni grandi che praticano esoincrocio è di norma al di sotto dell'1%.

Misurazione della diversità genetica a livello proteico



Le popolazioni
di grandi
dimensioni
mostrano grande
diversità
allozimica

Per esempio nell'uomo, su 104
loci analizzati, il 32% è
polimorfico, con una
eterozigosità media (H) del 6%.

TAXON	H (%)
vertebrati	
totale	6,4
mammiferi	5,4
uccelli	5,4
rettili	9,0
anfibi	9,4
pesci	5,4
invertebrati	
totale	11,3
insetti	12,2
crostacei	6,3
molluschi	12,1
piante	
gimnosperme	11,3
monocotiledoni	14,4
dicotiledoni	9,6

Diversità genetica nel genoma

Lo studio del gene *Adh* (alcol deidrogenasi) in *D. melanogaster* ha mostrato che su 11 campioni del gene (lungo 2379 nt) 43 siti erano polimorfici. Tuttavia **solo uno** di questi dava un cambiamento aminoacidico (alozima); gli altri mappavano negli introni del gene.

Regola generale

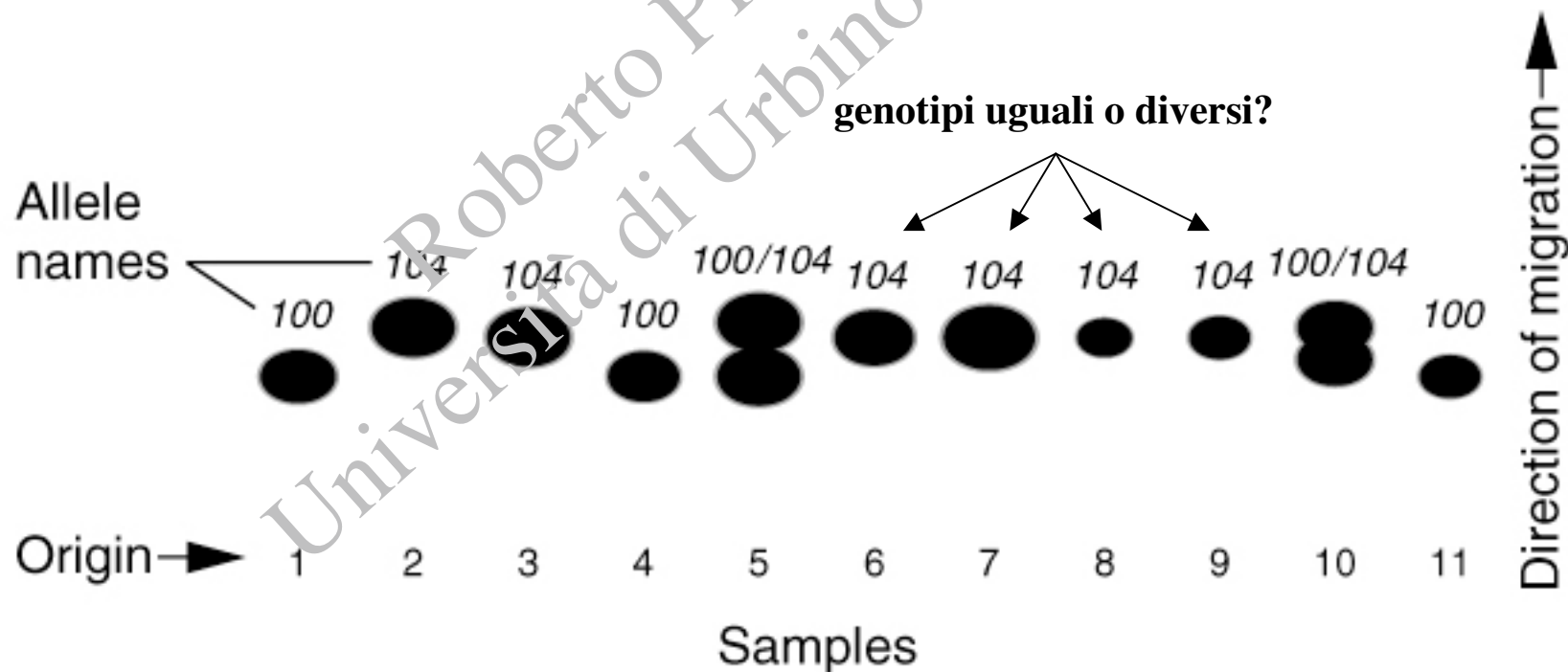
Le variazioni nucleotidiche si accumulano in sequenze con scarso significato funzionale, ma sono rare in regioni funzionalmente importanti delle molecole (per esempio, i siti attivi degli enzimi) perché eliminate per selezione naturale.

Eccezioni

- Il sistema maggiore di istocompatibilità (MHC) negli animali
- I loci di auto-incompatibilità nelle piante

Limiti dell'elettroforesi proteica

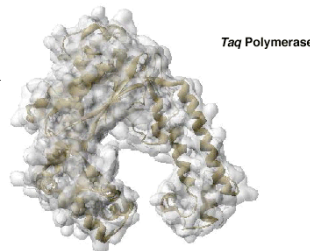
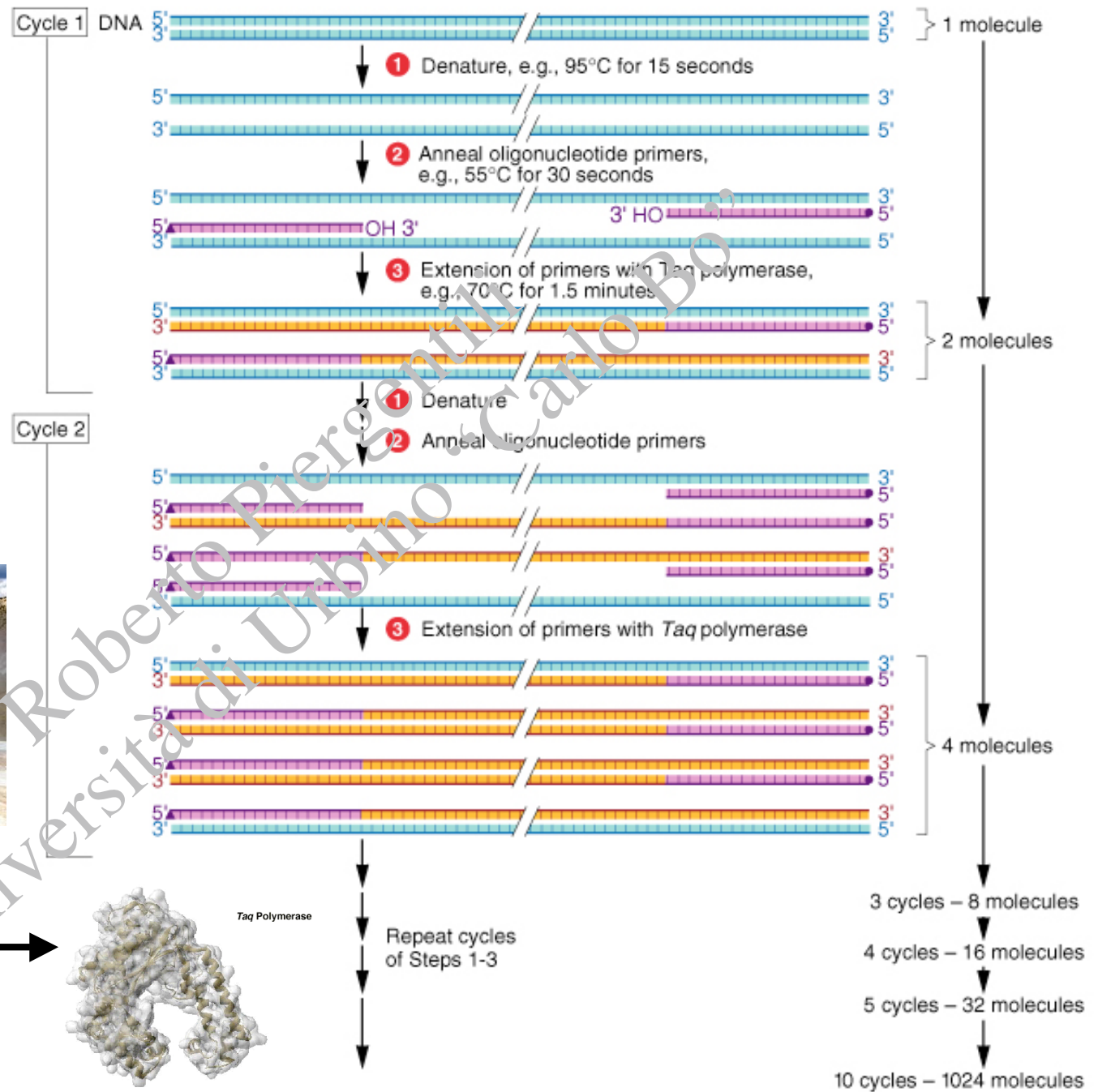
Omozigoti ed eterozigoti possono essere identificati a livello molecolare tramite elettroforesi, ma solo circa il 30% dei cambiamenti nel DNA determina un cambiamento nelle proteine **sottostima** significativa del **vero** livello di diversità genetica.



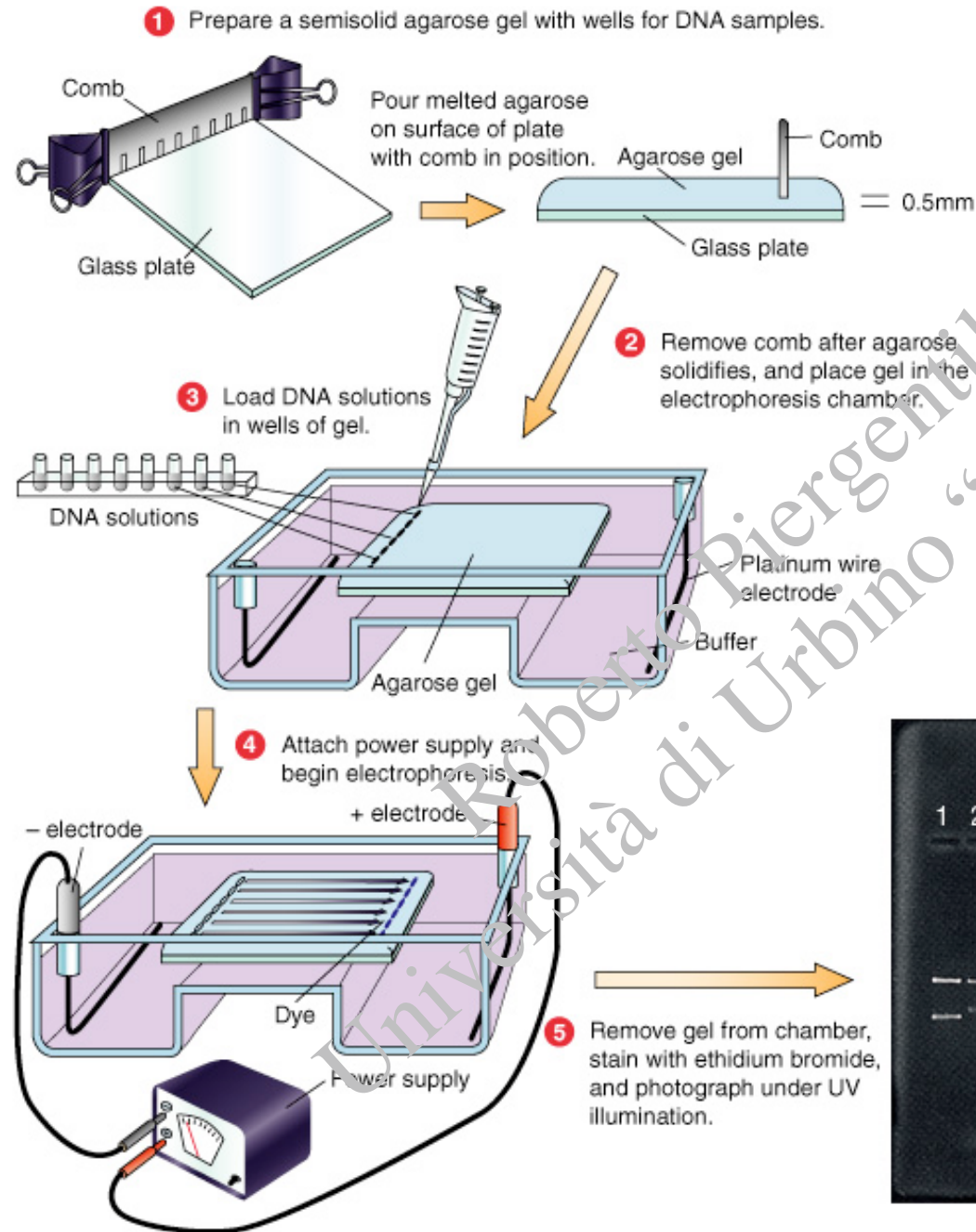
La Polymerase Chain Reaction



La TAQ polimerasi è estratta dal *Termophilus aquaticus*.



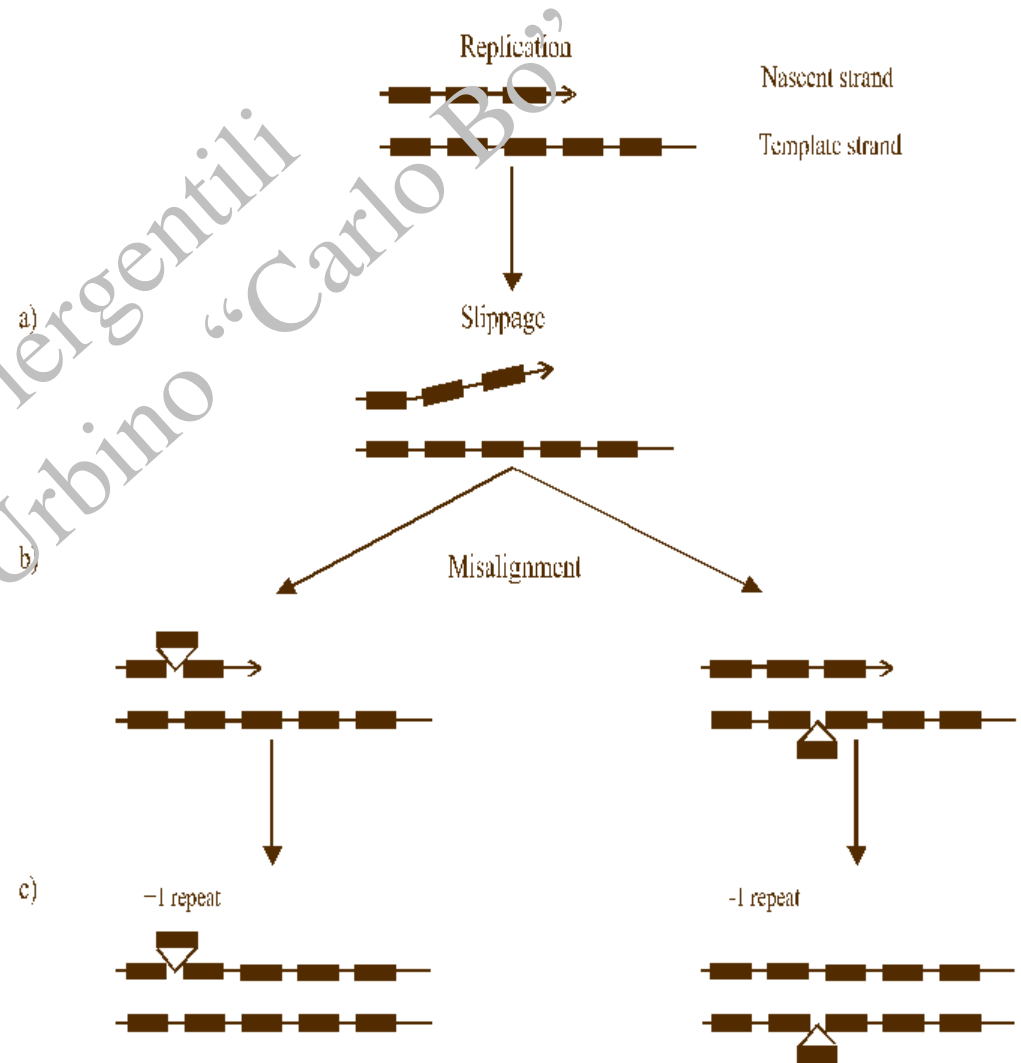
L'elettroforesi del DNA su gel di agarosio



Esiste una relazione **logaritmica** tra altezza della banda e sua lunghezza (ovvero peso molecolare).

I microsatelliti

I microsatelliti sono sequenze genetiche corte (1-5 nt) ripetute *in tandem* più volte. Il numero delle ripetizioni è altamente variabile a causa dello *slittamento della polimerasi* durante la replicazione del DNA, quindi si prestano bene per la misura della diversità genetica.



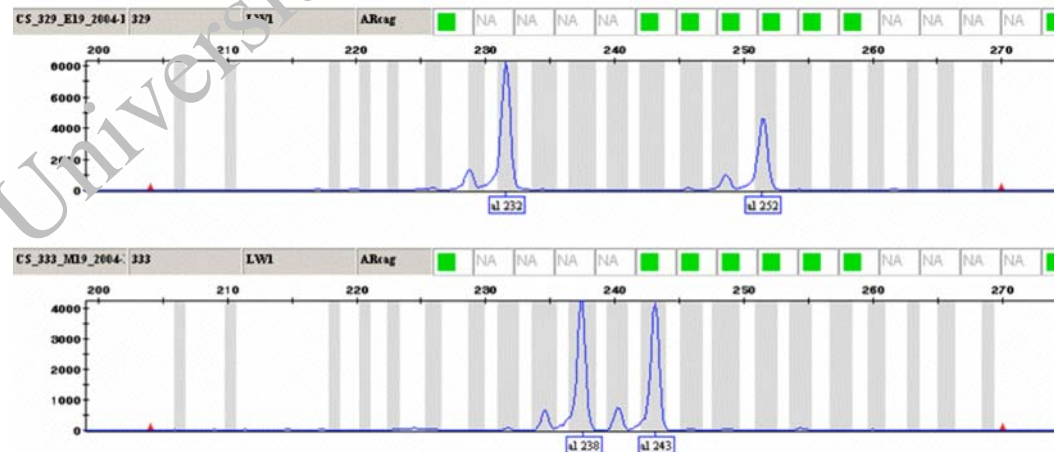
I microsatelliti

Vantaggi:

- misurano la variazione del DNA in tempi brevi;
- i genotipi individuali sono facilmente deducibili;
- la tipizzazione degli individui avviene attraverso campionamenti non invasivi.

Svantaggi:

- i primer per la PCR sono specie-specifici, quindi variano da caso a caso.



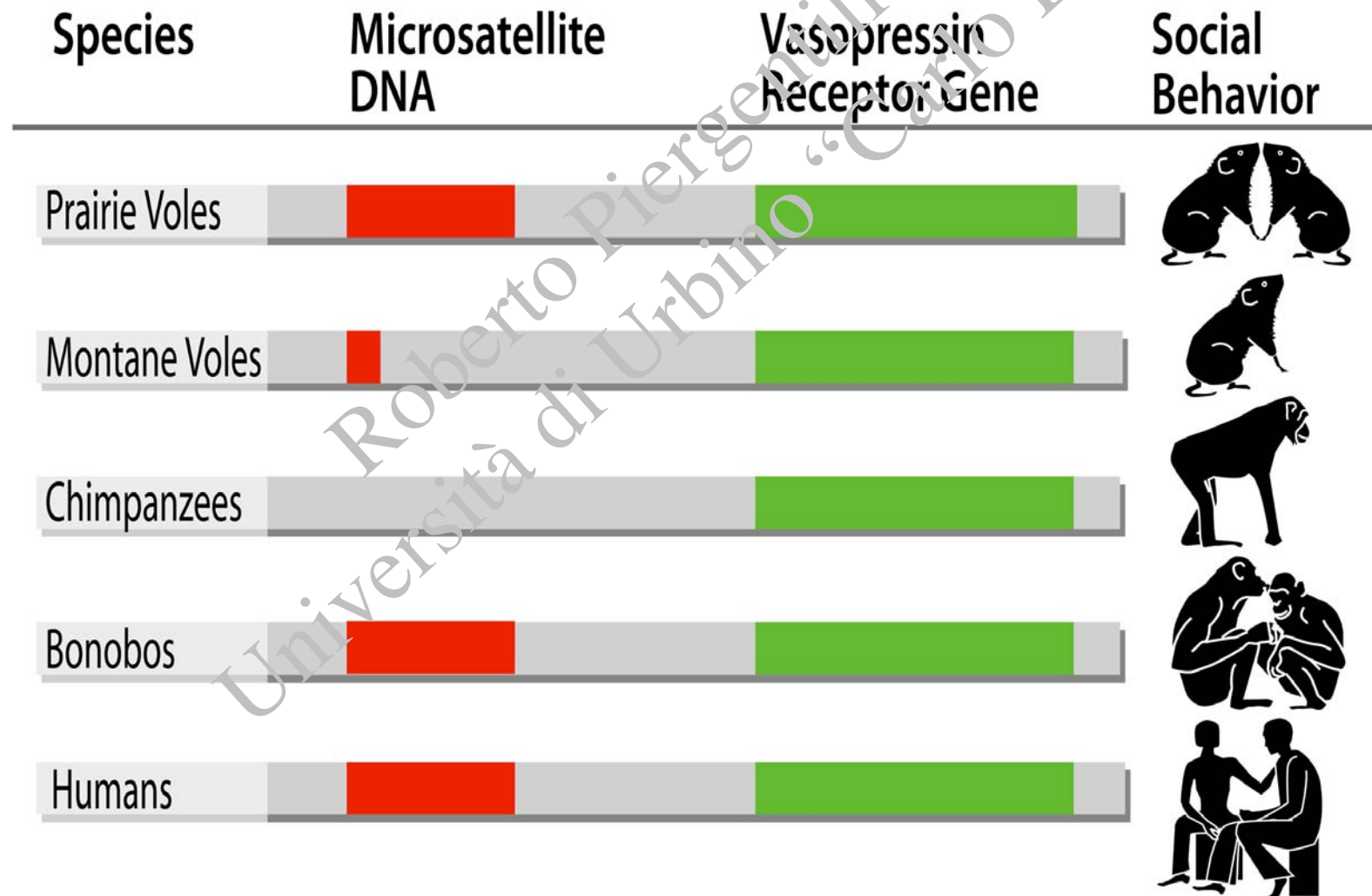
L'esempio dell'arvicola

Il comportamento sociale (gene quantitativo) dell'arvicola è determinato dalla lunghezza di un microsatellite a monte di un gene.



Random mutations in the length of the microsatellite DNA regions modify vole social behavior

L'influenza del microsatellite è evolutivamente conservata



Il DNA mitocondriale

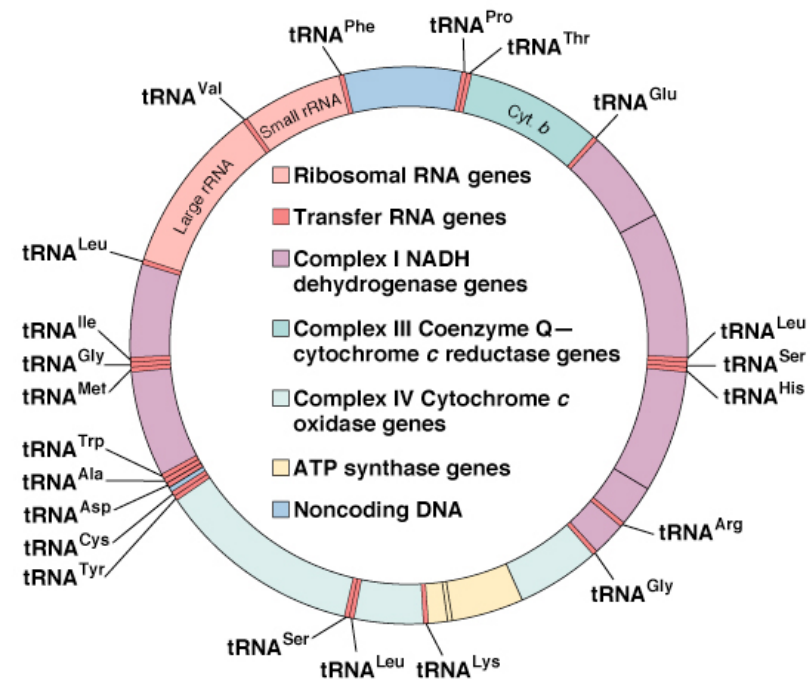
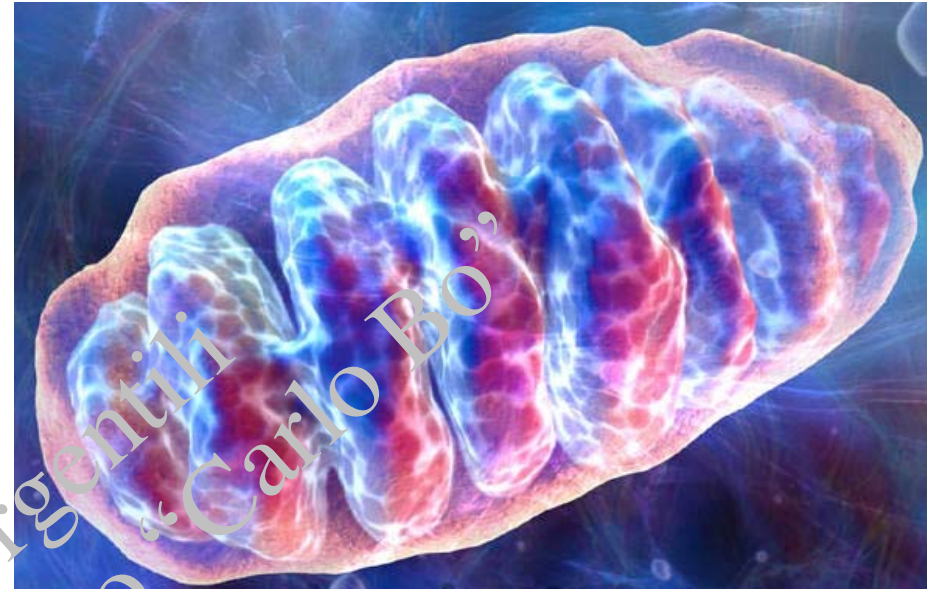
Anche il mtDNA si presta bene per l'analisi della diversità genetica.

Vantaggi:

- non ricombina, perché è ereditato esclusivamente per via materna;
- ha un elevato tasso di mutazione.

Svantaggi:

- dà informazioni solo sull'eredità matroclina.



Le popolazioni piccole hanno ridotta diversità genetica

Le popolazioni passate attraverso un *collo di bottiglia* presentano sia una ridotta variazione allozimica che per i microsatelliti, se confrontate con specie che non hanno subito riduzioni numeriche. Come si vede in tabella, le specie minacciate hanno perso il 40% circa della diversità genetica rispetto a quelle non a rischio.

Specie in pericolo	A	H%	Specie non a rischio	A	H%
Rinoceronte nero	4,2	69	Bufalo cafro	8,6	73
Lupo del Messico	2,7	42	Lupo grigio	4,5	62
Lupo etiopico	2,4	21	Coyote	5,9	68
Licaone	3,5	56	Cane domestico	6,4	73
Ghepardo	3,4	39	Leone africano	4,3	66
Cornacchia delle Marianne	1,8	16	Cornacchia americana	6,0	68
Gheppio delle Mauritius	1,4	10	Gheppio	5,5	68
Gheppio delle Seychelles	1,3	12	Gheppio africano m.	4,5	59
Falco pellegrino	4,1	48	Falco grillaio	5,4	70
Vombato <i>L. krefftii</i>	2,1	32	Vombato <i>L. latifrons</i>	5,9	71
<i>Potorus longipes</i>	3,7	56	Koala	8,0	81
Wallaby dalle briglie	11,6	83	Wallaby <i>P. assimilis</i>	12	86
Varano di Komodo	4,0	31	Alligatore americano	8,3	67
Mogano (pianta)	9,7	55	<i>P. arboreum</i>	9,3	67

Definizione: la fitness riproduttiva

Rappresenta il numero di **progenie fertile** con cui ciascun individuo di una popolazione contribuisce alla generazione successiva, e che a sua volta raggiunge la maturità sessuale, ha la capacità di accoppiarsi e di allevare la prole.

La **fitness media** è il rappporto tra il numero di individui nella generazione filiale e quello nella parentale.

Il **potenziale evolutivo** è (prevalentemente) misurato dalla variazione genetica quantitativa per la fitness riproduttiva.

