

Dynamik und Schicksal von mutanten Allelen in Populationen

„Nothing in evolution makes sense,
except in the light of the population“

(Gabriel Dover, verändert nach Dobzhansky)

Thomas Hankeln, Institut für Molekulargenetik

SS 2008

JOHANNES
GUTENBERG
UNIVERSITÄT
MAINZ

Begriffe

Genort/Locus

Allel

Allelfrequenz

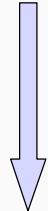
Fixierung

Polymorphismus

„effektive“ Populationsgröße

Evolution = Änderungen der Allelfrequenz mit der Zeit

Mutation



Natürliche Selektion
Zufall/ genetische Drift
Rekombination etc.
Migration

bewirken
Änderungen der
Allelfrequenz

Fixierung? Verlust?

Populationsgenetik

- > wie schnell?
- > wie wahrscheinlich?
- > welche Kräfte?

3

Nature 22. Mai 2003

brief communications

Rapid change in mouse mitochondrial DNA

Wild mice around Chicago may have switched genotype to keep pace with modern living.

We have compared the sequences of mitochondrial DNA extracted from museum skins of white-footed mice caught in the Chicago area since 1855 and from modern mice trapped alive in the same locations. We found a consistently similar directional change of mouse genotype over this period at each of five collection sites that were separated by 10–70 km. The genotype most common 100 years ago is now extremely rare, indicating that the mammalian mitochondrial genome can undergo rapid evolution.

Museums from Alaska to Switzerland were searched for specimens of the white-footed mouse, *Peromyscus leucopus*, originally collected from two counties in northeastern Illinois, USA. The white-footed mouse is the most common rodent in the deciduous forests of North America. We borrowed 61 museum specimen skins of mice collected from five locations in the Chicago area, and in 1999 we trapped 52 white-footed mice in the same locations and in three others.

Small (12 mm × 1.5 mm) strips were cut

**3 Allele
des mito-
chondrialen
CytC-Gens
untersucht
(stille Mutationen)**

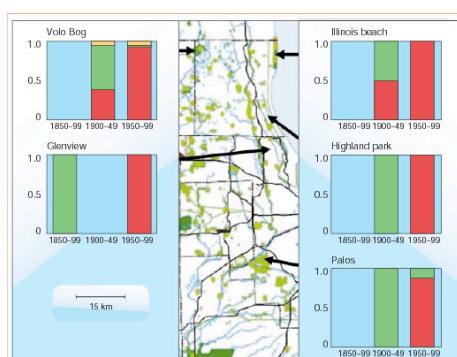


Figure 1 Modern map of the Chicago region where the change in genetic diversity of white-footed mice over 150 years was analysed. The geographic edition is shown between the five locations (arrows) from which pre-1950 museum specimens of *Peromyscus leucopus* were obtained. Major highways are shown in black; natural areas are shown in green. The frequencies (vertical scale) of three different mitochondrial haplotypes (A, green; M, red; Mix, yellow) are shown for each location for the periods 1850–99, 1900–49 and 1950–2000.

Selektion? Zufall?

4

Selektion

= differentielle Reproduktion unterschiedlicher Genotypen
(Fertilität, Viabilität)

Relative fitness ω

Häufig:	Mutation reduziert fitness	> negative/reinigende Selektion
„Manchmal“:	Mutation hat gleiche fitness	> M. ist „neutral“
Selten:	Mutation erhöht fitness	> positive/Darwin'sche Selektion

5

Selektion

Relative fitness ω zweier Genotypen

$$\omega = 1 + S \text{ (Selektionskoeffizient)}$$

- wobei $S = 0$ > neutral
 $S < 0$ > fitness-Nachteil
 $S > 0$ > fitness-Vorteil

Beispiel: bei $S = + 0,1$ hat Genotyp 10 % Fitnessvorteil

6

Selektion

Bestimmung des Fitnesswerts am Bsp. der Chondrodysplasie

- 108 Kleinwüchsige > 27 Kinder > $27/108 = 0,25$
 - 457 Kontrollpersonen > 582 Kinder > $582/457 = 1,274$
 - Chondro-Allel/ Wildtyp = $0,25 / 1,274 = 0,196$
- Relative fitness $\omega = 0,196$ ($S_c = -0,804$)

7

Selektion

Grundlage: Allelverteilung in Population nach Hardy und Weinberg (1908)

$$\begin{array}{ll} \text{Allel A1} & - \quad \text{Frequenz } p \\ \text{Allel A2} & - \quad \text{Frequenz } q \end{array} \quad \left. \right\} \quad p + q = 1$$

Genotypen	A1 A1	A1 A2	A2 A2
Frequenz*	p^2	$2pq$	q^2
Fitness	ω_{11}	ω_{12}	ω_{22}

* Gilt nur bei bestimmten Bedingungen: Panmixie, keine Migration, etc.

8

Wie wirkt sich Fitness auf Allelfrequenz in der folgenden Generation aus?

$$\text{Frequenz } G_{t+1} = \text{Frequenz } G_t \times \omega$$

d.h. die Anteile der drei Genotypen an der nächsten Generation sind:

A1 A1	$p^2\omega_{11}$
A1 A2	$2pq\omega_{12}$
A2 A2	$q^2\omega_{22}$

Für neues Allel A2 gilt:

$$q_{t+1} = \frac{pq\omega_{12} + q^2\omega_{22}}{p^2\omega_{11} + 2pq\omega_{12} + q^2\omega_{22}} \quad (1)$$

und $\Delta q = q_{t+1} - q_t = \frac{pq(p(\omega_{12} - \omega_{11}) + q(\omega_{22} - \omega_{12}))}{p^2\omega_{11} + 2pq\omega_{12} + q^2\omega_{22}}$ (2)

9

Allel-Interaktion und Selektion

Die fitness-Werte hängen von der Interaktion der Allele A1 und A2 ab!

A1 A1	A1 A2	A2 A2
1	$1 + S$	$1 + S$
1	$1 + S$	$1 + 2S$
1	$1 + S$	$1 + t$ ($S > 0, S > t$)
1	1	$1 + S$

Komplette Dominanz

Ko-Dominanz
(Het. intermediär; syn. generic selection)

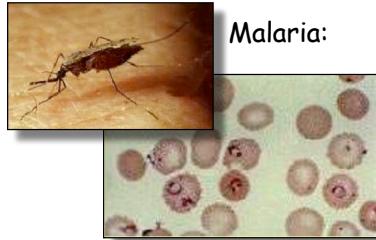
Über-Dominanz
(Het. hat höchste fitness)

Rezessivität

in Formel (2) anstelle von ω einsetzen

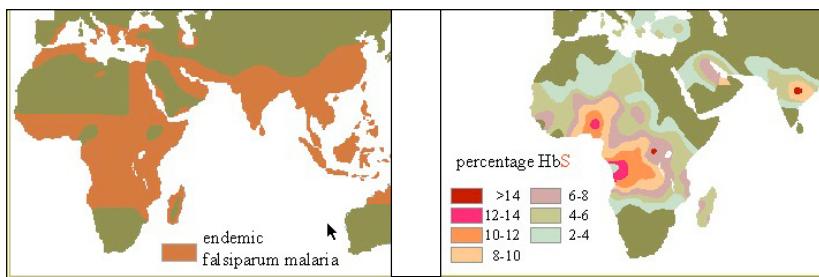
10

Beispiel: Überdominanz



1-2 Mio. Tote jährlich
übertragen durch Anopheles-Weibchen
ausgelöst durch Plasmodium falciparum
dringt in Erythrozyten ein, lebt von Hämoglobin

Sichelzell-Allel: HbS ($\beta 6$ Glu $>$ Val)
homozygot $>$ Anämie
heterozygot $>$ Vorteil in
Malariagebieten

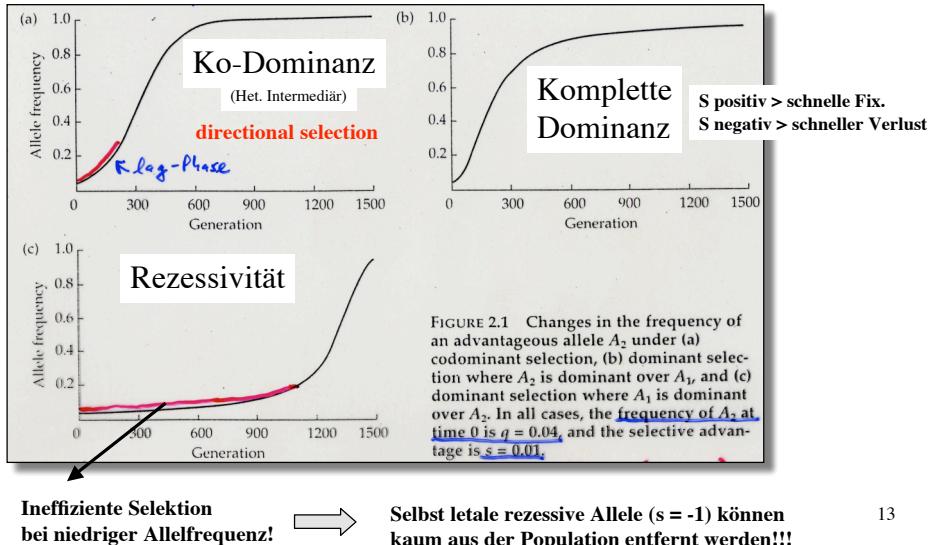


11

Wie beeinflussen die Selektion
und die verschiedenen Formen
der Allelinteraktion das Schicksal
eines neuen Allels A2 in der
Population?

12

Schicksal eines raren Allels A2 unter verschiedenen Selektionsbedingungen



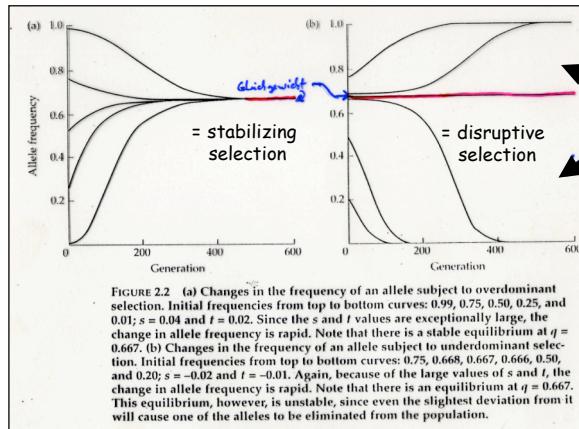
13

Schicksal eines raren Allels A2 unter verschiedenen Selektionsbedingungen

Ein rares Allel ist nur sehr selten homozygot anzutreffen!

- **Ko-Dominanz:**
 - Zunahme von A2 auf Kosten von A1 (= directional selection“)
 - Lag-Phase wegen schwächerer Selektion von A1A2-Heterozygoten
- **Dominanz:**
 - sehr schnelle directional selection
 - dominant negatives Allel würde schnell entfernt (werden daher sehr selten in Population beobachtet)
- **Rezessivität:**
 - sehr langsame directional selection; nicht sehr effizient in der Fixierung vorteilhafter bzw. in der Entfernung nachteiliger Allele!!
 - > Persistenz von z.B. Cyst. Fibrose-Allelen¹⁴

Über-Dominanz Unter-Dominanz (Heterozygoten haben größte fitness) (Heterozygoten haben geringste fitness)



Bei Verlassen des Gleichgewichts:

- ein Allel fixiert
- ein Allel verloren



Ein neues unterdominantes Allel ist rar, d.h. unterhalb des Gleichgewichts:
Würde trotz Selektionsvorteil in A₂/A₂ aus Population entfernt!

15

Die zweite Kraft: „Genetische Drift“

= Allelfrequenz-Änderungen durch den Zufall

- Auswahl von Gameten für die Befruchtung
- Auswahl von an der Reproduktion teilnehmenden Individuen

16

Genetische Drift

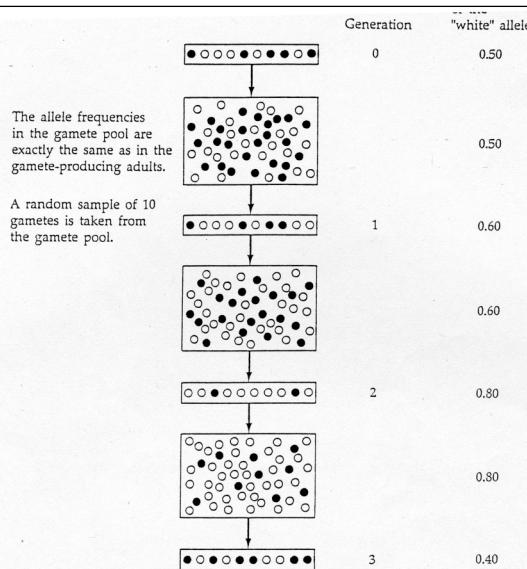


Figure 3. Random sampling of gametes. Allele frequencies in the gamete pools (large boxes) in each generation are assumed to reflect exactly the allele frequencies in the adults of the parental generation (small boxes). Since the population size is finite, allele frequencies fluctuate up and down. Modified from Bodmer and Cavalli-Sforza (1976).

17

Genetische Drift

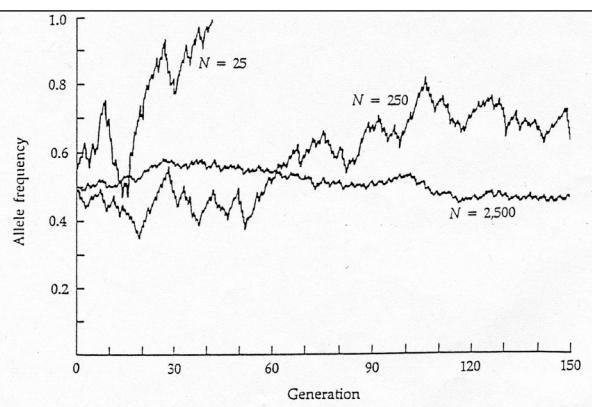


Figure 4. Changes in frequencies of alleles subject to random genetic drift in populations of different sizes. The smallest population reached fixation after 42 generations. The other two populations were still polymorphic after 150 generations but will ultimately reach fixation (allele frequency = 0% or 100%) if the experiment is continued long enough. From Bodmer and Cavalli-Sforza (1976).

18

Genetische Drift

- Die Richtung der Allelfrequenzänderungen unter Drift ist zufällig
- Die Bedeutung von Drift ist in kleinen Populationen sehr groß
- prinzipiell erreichen alle Allele unter Drift nach langer Zeit entweder Fixierung oder Verlust
(Ausnahme: balancierte Allele)

19

Wie wahrscheinlich ist die Fixierung eines mutanten Alles in der Population? > **fixation probability**

Wie lange dauert es bis zur Fixierung?
> **fixation time**

Wieviel neue Allele pro Zeiteinheit werden fixiert?
> **gene substitution rate**

20

10

Die Fixierungswahrscheinlichkeit eines Allels ist abhängig von...

- ...seiner Frequenz q
- ...seinem Selektionskoeffizienten S
- ...der effektiven Populationsgröße N_e

Kimura 1962

$$P = \frac{1 - e^{-4N_e s q}}{1 - e^{-4N_e s}} \quad (3)$$

(bei $x = \text{klein}$, gilt: $e^{-x} = 1 - x$)

21

Fixierungswahrscheinlichkeit

$$P = \frac{1 - e^{-4N_e s q}}{1 - e^{-4N_e s}} \quad (3)$$

• für S gegen 0 (Neutralität) gilt: $P \approx q$

d.h. neutrales Allel mit $q = 0,4$ hat 40%ige Fixierungswahrsch.

Für neue Mutation in diploider Population ist $q = 1 / 2N$.

d.h. für neues, neutrales Allel gilt: $P = 1 / 2N$

Je kleiner die Population, desto höher ist die Fixierungswahrscheinlichkeit eines neutralen Allels (durch Drift)

22

Fixierungswahrscheinlichkeit

Bei Einsetzen von $q = 1/2N$ in Formel (3), sowie bei $N \approx Ne$ gilt:

$$P = \frac{1 - e^{-2s}}{1 - e^{-4Ns}} \quad (4)$$

- für s (klein, positiv) und N (groß) gilt:

$$P \approx 2s$$

d.h. in großen Populationen hängt die Fix-wahrscheinlichkeit vom Selektionswert ab (1% Selektionsvorteil > 2% Fix-wahrsch.)

Bsp. $N = 1000$	neutral	$\rightarrow P = 0,05\%$
$N = Ne$	$S = +0,01$	$\rightarrow P = 2\%$
	$S = -0,001$	$\rightarrow P = 0,004\% \quad !!!$

23

Fixierungswahrscheinlichkeit

ACHTUNG!

- auch nachteilige Mutationen sind im Prinzip fixierbar!!
- vorteilhafte Mutationen müssen NICHT unbedingt fixiert werden

24

Fixierungsrate

Neutrale Mutationen ($s = 0$)

$$\begin{array}{lll} K = \text{Anzahl neuer Mut.} & X & \text{Fix. Wahrsch.} \\ K = & 2Nu & X \\ & = u \quad (\text{Mutationsrate}) & 1/2N \end{array}$$

Für neutrale Mutationen ist K proportional zur Mutationsrate und unabhängig ! von der Populationsgröße.

(in großen Populationen gibt es viele Mutationen, aber wenig Fix-Chancen;
in kleinen Populationen gibt es wenig Mutationen, aber höhere Fix-Chancen)

25

Fixierungsrate

Vorteilhafte Mutationen ($s > 0$)

$$\begin{array}{lll} K = \text{Anzahl neuer Mut.} & X & \text{Fix. Wahrsch.} \\ K = & 2Nu & X \\ K = & 4Ns u & 2s \end{array}$$

Fix-rate ist von allen 3 Parametern abhängig

26

Dynamik der Allel-Substitution

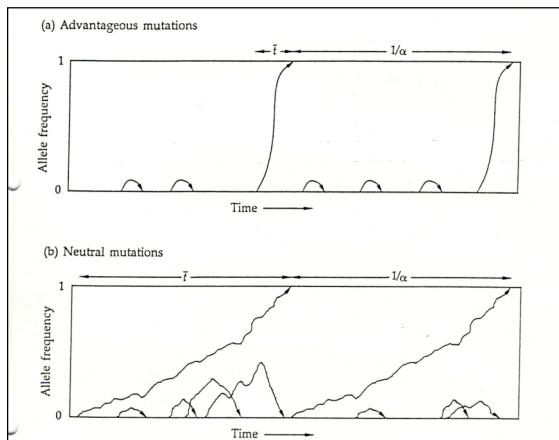


Figure 6. Dynamics of gene substitution for (a) advantageous and (b) neutral mutations. Advantageous mutations are either quickly lost from the population or quickly fixed, so that their contribution to genetic polymorphism is small. The frequency of neutral alleles, on the other hand, changes very slowly by comparison, so a large amount of transient polymorphism is generated. \bar{t} is the conditional fixation time and $1/\alpha$ is the mean time between consecutive fixation events. From Nei (1987).

Schnelle Fix. oder schneller Verlust
➤ Wenig Beitrag zum Polymorphismus innerhalb der Population

Langsame Frequenzänderungen
➤ Transienter Polymorphismus

27

Triebkräfte der Evolution

neo-darwinistische Evolutionstheorie = Selektionismus

- natürliche Selektion vorteilhafter Mutationen ist treibende Kraft jeglicher Veränderungen (positive Selektion)
- (Mutation) (Drift)
- Polymorphismen beruhen immer auf stabilisierender Selektion

Neutrale Evolutionstheorie (Kimura 1968)

- Zufallseffekte (Drift) stärker als Selektion
- Polymorphismen = Übergangszustand zwischen Fixierung oder Verlust eines Allels in der Population (Allelsubstitution ist fließender Vorgang)

beiden gemeinsam:

- die meisten Mutationen sind nachteilig und werden relativ schnell entfernt (negative Selektion)
- der Anteil „vorteilhafter Mutationen : neutralen Mutationen“ ist strittig

(experimenteller Ansatz: MacDonald-Kreitman-Test)

28

MacDonald/Kreitman-Test auf „neutrale Evolution“ eines Gens

- Frage: sind Genunterschiede zwischen zwei Arten **neutral** oder auf **positive/adaptive Selektion** zurückzuführen?

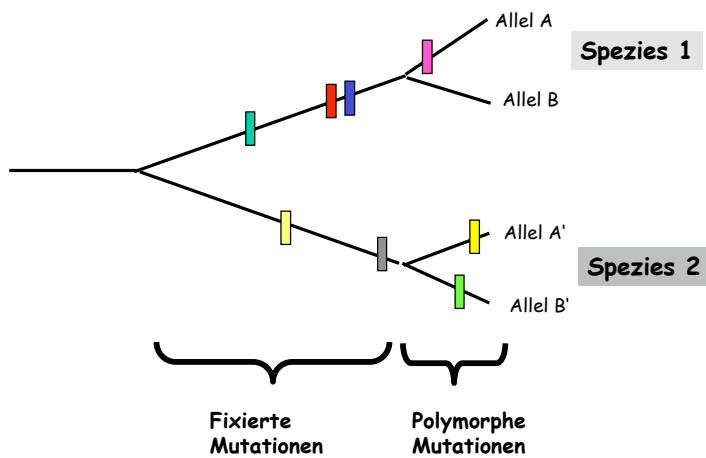
29

MacDonald/Kreitman-Test auf „neutrale Evolution“ eines Gens

- Annahmen aufgrund der „neutral theory“ als Nullhypothese:
 - > Unterschiede innerhalb der Population (= allelische Polymorphismen) sind ebenso wie zwischenartliche Unterschiede weitgehend neutral
 - > nur nicht-synonyme Nukleotidaustausche können adaptiv sein; synonyme Substitutionen sind (zumeist) neutral
 - > eine vorteilhafte Mutation wird in der Population schneller fixiert als eine neutrale, und ist daher weniger häufig noch in einem polymorphen Zustand

30

MacDonald/Kreitman-Test auf „neutrale Evolution“ eines Gens



31

MacDonald/Kreitman-Test auf „neutrale Evolution“ eines Gens

- Erwartung: bei Neutralität sollte das Mengen-Verhältnis zwischen syn/non-syn Substitutionen an innerartlich „polymorphen“ Positionen und zwischenartlich „fixierten“ Positionen etwa gleich sein

Test:	Spezies A	Spezies B
Allel 1	T G	A
Allel 2	T	A
Allel 3	T G	A
Allel 4	T	A C
	↓ ↓	↓ ↓
	fix polymorph	fix polymorph

32

**Adaptive protein evolution at the
Adh locus in *Drosophila***
John H. McDonald & Martin Kreitman
Department of Ecology and Evolutionary Biology, Princeton University,
Princeton, New Jersey 08544, USA

Nature 351 (1991) p.652ff

**MacDonald/
Kreitman-
Test**

TABLE 1 Variable nucleotides from the coding region of the *Adh* locus in *D. melanogaster*, *D. simulans* and *D. yakuba*

Con.	<i>D. melanogaster</i>										<i>D. simulans</i>						<i>D. yakuba</i>										Repl.	Fixed
	a	b	c	d	e	f	g	h	i	j	k	l	a	b	c	d	e	f	g	h	i	j	k	l				
781	G	T	T	T	T	T	T	T	T	T	T	T	-	-	-	-	-	-	C	C	C	C	C	C	Repl.	Fixed		
789	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	G	G	G	G	G	Syn.	Fixed		
808	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	C	C	C	C	C	RepL.	Fixed		
816	G	T	T	T	-	-	-	-	-	T	T	T	T	-	-	-	-	-	G	G	G	G	G	G	Syn.	Poly.		
834	T	-	-	-	-	-	-	-	-	-	-	-	C	C	-	C	-	-	G	G	G	G	G	G	Syn.	Poly.		
859	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	G	G	G	G	G	RepL.	Fixed		
867	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	G	G	G	G	G	Syn.	2 Poly.		
870	C	T	T	T	T	T	T	T	T	T	T	T	-	-	-	-	-	-	G	G	G	G	G	G	Syn.	Fixed		
950	G	-	-	-	-	-	-	-	-	-	-	-	A	-	-	-	-	-	C	C	C	C	C	C	Syn.	Poly.		
974	G	-	-	-	-	-	-	-	-	-	-	-	T	T	T	T	T	T	C	C	C	C	C	C	Syn.	Poly.		
983	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	C	C	C	C	C	Syn.	Fixed		
1019	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-	-	-	-	Syn.	Poly.		
1031	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-	-	-	-	Syn.	Poly.		
1034	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	C	C	C	C	C	Syn.	Poly.		
1043	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-	-	-	-	Syn.	Poly.		
1068	C	T	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	Syn.	Poly.		

33

MacDonald/Kreitman-Test

TABLE 9.4 Numbers of replacement and synonymous fixed differences between species and of polymorphisms within species

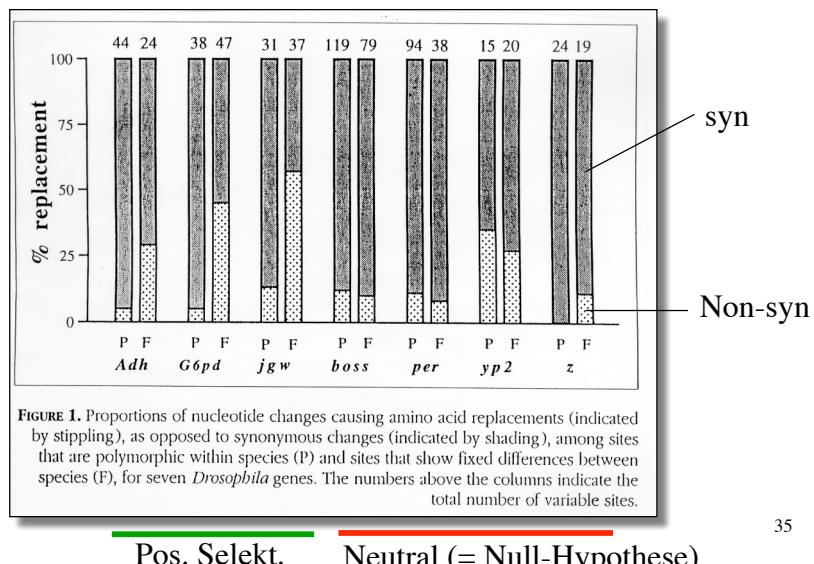
Type of change	Differences	
	Fixed	Polymorphic
<i>Adh</i> gene ^a		
Replacement	7	2
Synonymous	17	42
<i>G6pd</i> gene ^b		
Replacement	21	2
Synonymous	26	36

^aMcDonald and Kreitman (1991): 12, 6, and 24 sequences from *Drosophila melanogaster*, *D. simulans*, and *D. yakuba*, respectively.
^bEanes et al. (1993): 32 and 12 sequences from *D. melanogaster* and *D. simulans*, respectively.

34

17

MacDonald/Kreitman-Test auf „neutrale Evolution“



Pos. Selekt. Neutral (= Null-Hypothese)

35

Ist die Evolution primär durch Mutation oder Selektion angetrieben?

- Frage betrifft vor allem **nicht-syn. Nukleotidaustausche** sowie die resultierenden **Aminosäureunterschiede** zwischen Proteinen.

Akzeptiert ist, dass nicht-kodierende Genomregionen weitgehend unter Mutationseinfluss, nicht jedoch unter Selektion evolvieren.

(Ausnahme: genregulatorische Bereiche, die weitgehend noch unerkannt sind, aber möglicherweise viel ausgedehnter im Genom sind als lange angenommen; vergl. ENCODE-Projekt)

36

Ist die Evolution primär durch Mutation oder Selektion angetrieben?

Bsp. 1: die Rate der nicht-syn. Substitutionen ist in vielen Organismen positiv korreliert mit der syn. (neutralen) Subst.-Rate

TABLE 14.1 Rates of synonymous and nonsynonymous substitution in various organisms ^a			
Gene	Nonsynonymous	Synonymous	Ratio
<i>Mammals^b</i>			
Lowest (actin α)	0.01×10^{-9}	2.92×10^{-9}	0.003
Highest (interferons γ)	3.06×10^{-9}	5.50×10^{-9}	0.556
Average	0.74×10^{-9}	3.51×10^{-9}	0.210
<i>Drosophila^c</i>			
Lowest (arrestin B)	0.23×10^{-9}	11.62×10^{-9}	0.020
Highest (<i>Cp18</i> Chorion prot. 4)	5.63×10^{-9}	18.38×10^{-9}	0.306
Average	1.91×10^{-9}	15.60×10^{-9}	0.122
<i>Human immunodeficiency virus (HIV-1)^d</i>			
Lowest (<i>pol</i> gene)	1.6×10^{-3}	11.0×10^{-3}	0.145
Highest (<i>env</i> hypervariable regions)	14.0×10^{-3}	17.2×10^{-3}	0.814
Average	3.9×10^{-3}	10.3×10^{-3}	0.379

^aAll rates are in units of substitutions per site per year.

^bFrom Table 7.1, excluding histone genes because their synonymous rates are difficult to estimate.

^cFrom Table 7.6.

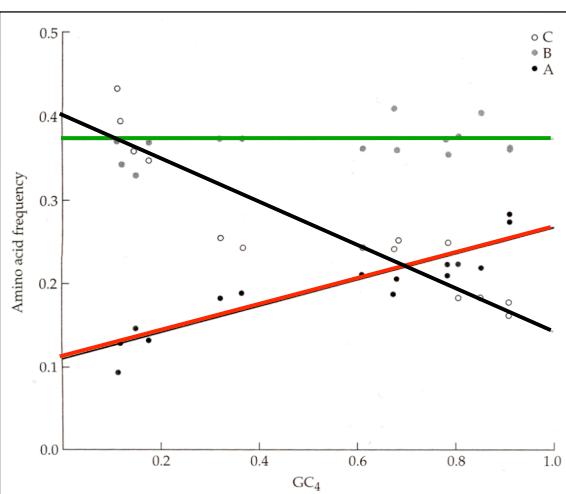
^dFrom Li et al. (1988a).

Beide Raten steigen
etwa um 10^6 von
Tieren hin zu Viren

37

Ist die Evolution primär durch Mutation oder Selektion angetrieben?

Bsp. 2: Mutationsdruck beeinflusst die Aminosäurezusammensetzung!



A = As mit GC-reichen Codons (pro, ala, gly, trp)

B = As mit Codons eines mittleren GC-Gehalts (val, thr, his, glu, asp, gln, cys, ser)

C = As mit GC-armen Codons (phe, tyr, asn, lys, ile, met)

38

Ist die Evolution primär durch Mutation oder Selektion angetrieben?

Bsp. 3: Entstehung nichtuniverseller genetischer Codes ist angetrieben durch den Mutationsdruck!

Bsp. 4: Neue Funktionen in Proteinen können zunächst mutationsbedingt entstehen

- > 2,3-DPG unterstützt O₂-Abgabe des β-Hämoglobins
- > Evolution der DPG-Bindung an β-Hb erforderte mehrere As-Veränderungen, die nur zusammen wirksam sind
- > d. h., die ersten mutativen Schritte waren noch neutral (da nicht effektiv für DPG-Bindung; erst der letzte Mutationsschritt hat DPG-Bindung ermöglicht und ist als positiv-selektioniert anzusehen)

39

Ist die Evolution primär durch Mutation oder Selektion angetrieben? ...ein Fazit

Es existieren eine Reihe guter Beispiele, dass selbst die Evolution kodierender DNA und die Evolution von Proteinen zu einem wesentlichen Teil von der **Mutation** (und nicht von der Selektion) bestimmt werden!!!

Welche Rolle spielt denn dann die Selektion?

- > **Negative (reinigende) Selektion** eliminiert unvorteilhafte Mutationen
- > **Positive (adaptive) Selektion** ist sehr wichtig für die Evolution neuartiger Eigenschaften und kann (im Gegensatz zur genetischen Drift) neue Allele rasch in der Population etablieren