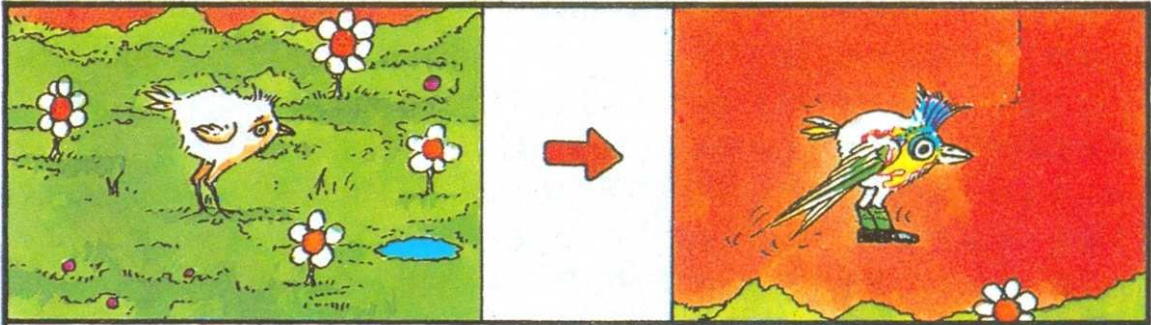
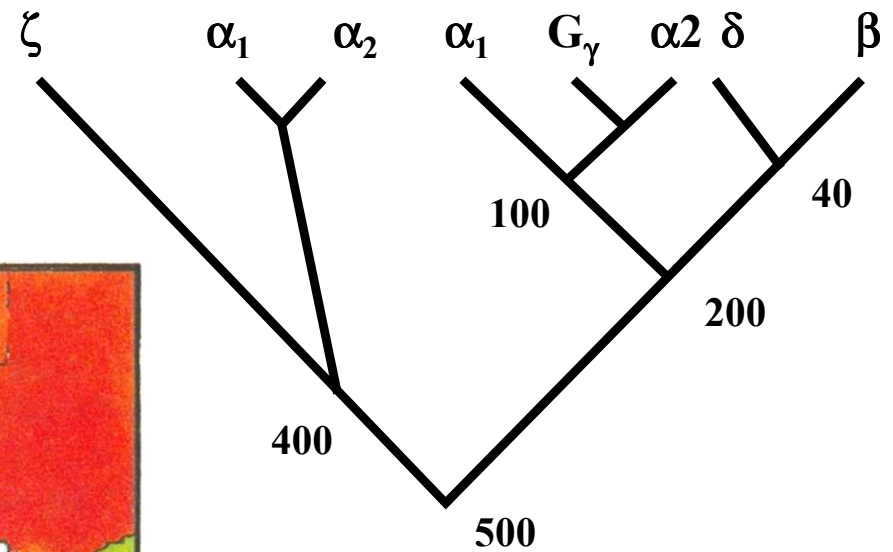
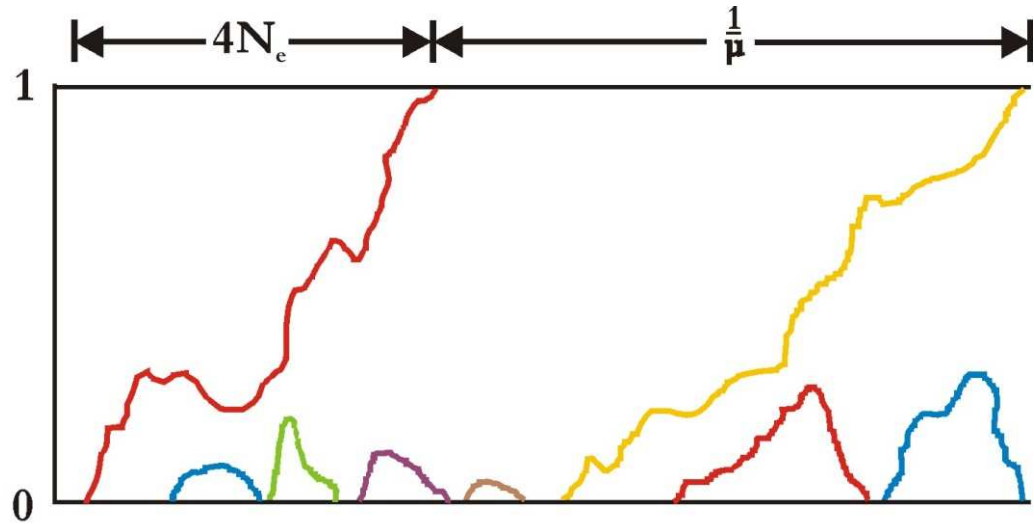


# MOLEKULÁRNÍ EVOLUCE



# NEUTRÁLNÍ TEORIE MOLEKULÁRNÍ EVOLUCE

## Genetická zátěž a selekční náklady

jestliže ne všichni jedinci v populaci mají optimální fitness,  $w_{av} < w_{max}$

**genetická zátěž populace**  $L$ :  $L = 1 - w_{av}$

jestliže  $w_{av} = w_{max}$ ,  $L = 0$

měří, do jaké míry je průměrný jedinec v populaci méně zdatný než nejlepší genotyp

vyjadřuje pravděpodobnost, že průměrný jedinec zemře před svou reprodukcí

jestliže se jedinec během svého života nerozmnoží, označujeme to jako jeho **genetickou smrt**

vznik nevýhodných  
alel

Genetická zátěž může mít několik forem, např.  
**mutační zátěž**  
**substituční zátěž**  
**segregační zátěž**

zvýhodnění heterozygotů  
→ náklady na homozygoty

# Substituční zátěž

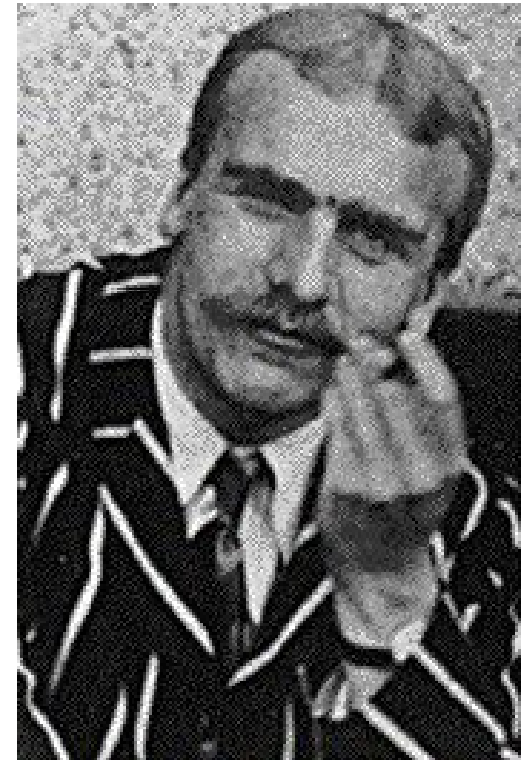
Haldane (1957):

prospěšná mutace → fixace výhodné alely a nahrazení alely nevýhodné

dokud původní alela existuje v populaci, průměrná fitness nižší než fitness optimální

$$L = \frac{w_{\max} - \bar{w}}{w_{\max}}$$

jestliže  $w_{\max} = 1$ ,  $L = 1 - \bar{w}$



J.B.S. Haldane

## Selekční náklady:

Nahrazení jedné alely v populaci za druhou si můžeme představit jako „selektivní smrt“ původní alely.

Čím je intenzita selekce vyšší, tím větší množství původní (nevýhodnější) alely je v každé generaci z populace vyřazeno („zemře“).

Pokud by selekce byla příliš silná, mohla by způsobit extinkci celé populace.

Nová, výhodnější alela  $A$ : frekvence  $p$ , fitness  $w = 1$   
původní, nevýhodnější alela  $A'$ : frekvence  $1-p = q$ ,  $w = 1-s$   
→ poměr nepřeživších/přeživších v každé generaci =  $sq/(1 - sq)$

⇒ potomstvo navíc

např. jestliže poměr 0,1/0,9 každý přeživší → 1 1/9 potomstva,  
ale jestliže poměr 0,999/0,001 → 1000 potomstva navíc

celkové selekční náklady = suma přes všechny lokusy:

$$C = \sum \frac{sq}{1 - sq}$$

**Horní limit selekčních nákladů:**

Haldane: substituce 1 genu/300 generací

# Neutrální teorie molekulární evoluce

Moderní syntéza: debata selekce vs. drift

začátek 60. let 20. stol. → sekvence aminokyselin v proteinech

1966: **Lewontin & Hubby** - *D. pseudoobscura*; **Harris** - člověk → rozsáhlý polymorfismus



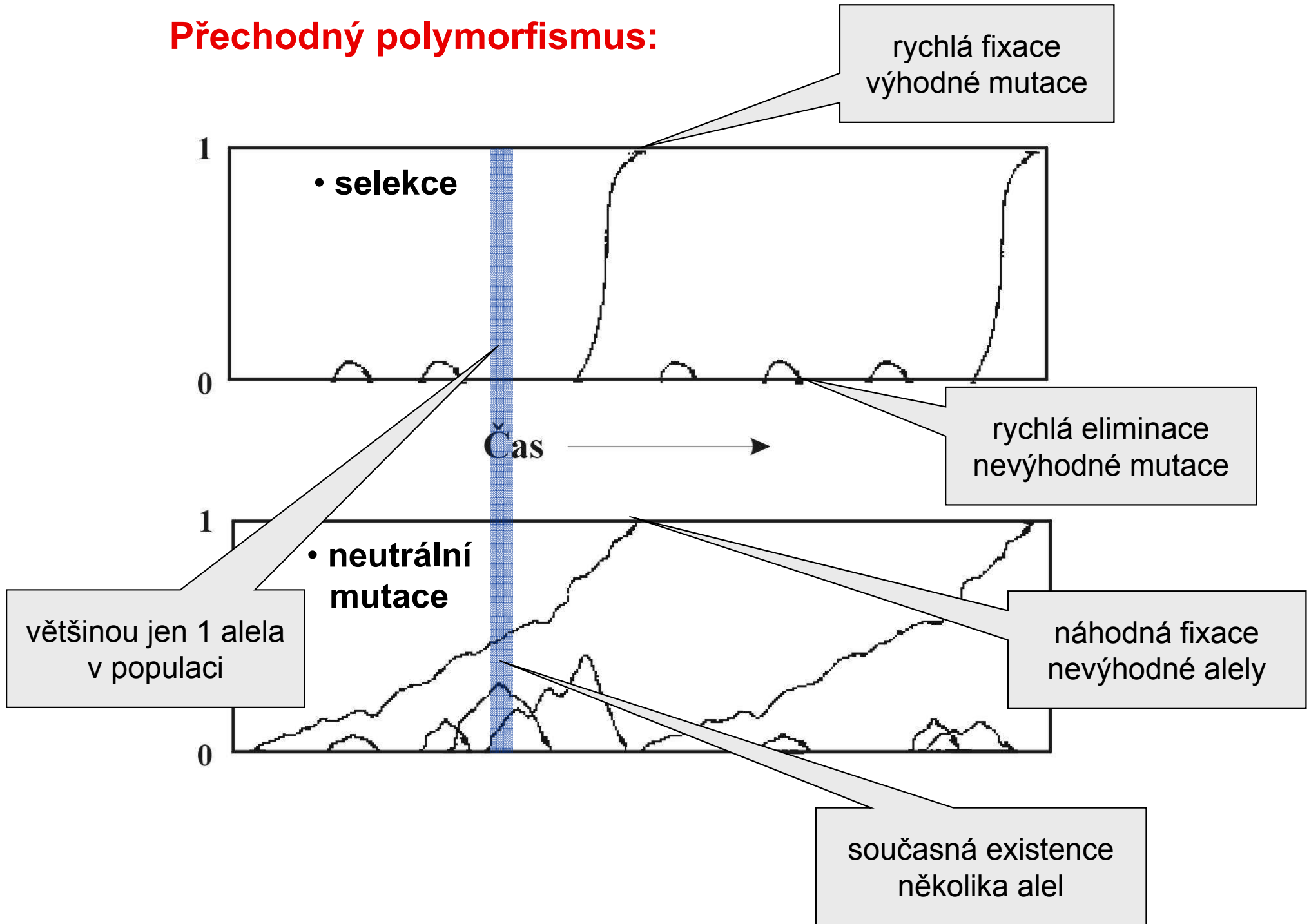
Rychlost molekulární evoluce příliš vysoká (vysoké selekční náklady)

Rozsah genetické proměnlivosti v populacích příliš vysoký (vysoká substituční zátěž ⇒ polymorfismus neudržován selekcí – **přechodný polymorfismus**)

Konstantnost molekulární evoluce

Vyšší evoluční rychlost u funkčně méně důležitých částí molekuly

# Přechodný polymorfismus:



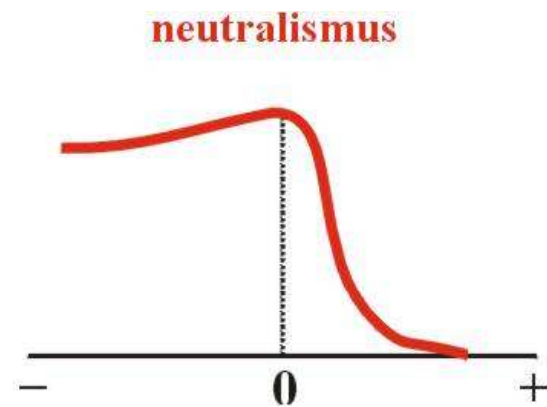
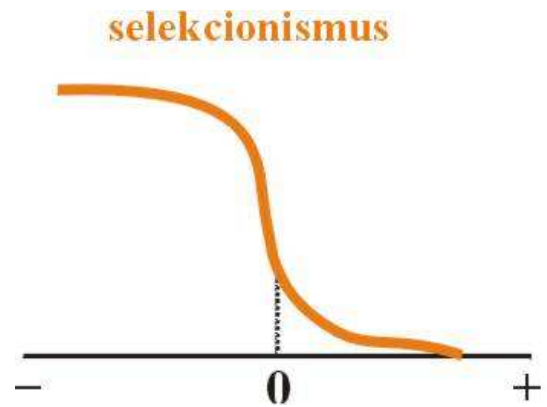
Motoo Kimura (1968)  
J.L. King & T.H. Jukes (1969)



M. Kimura

**neutrální teorie:**

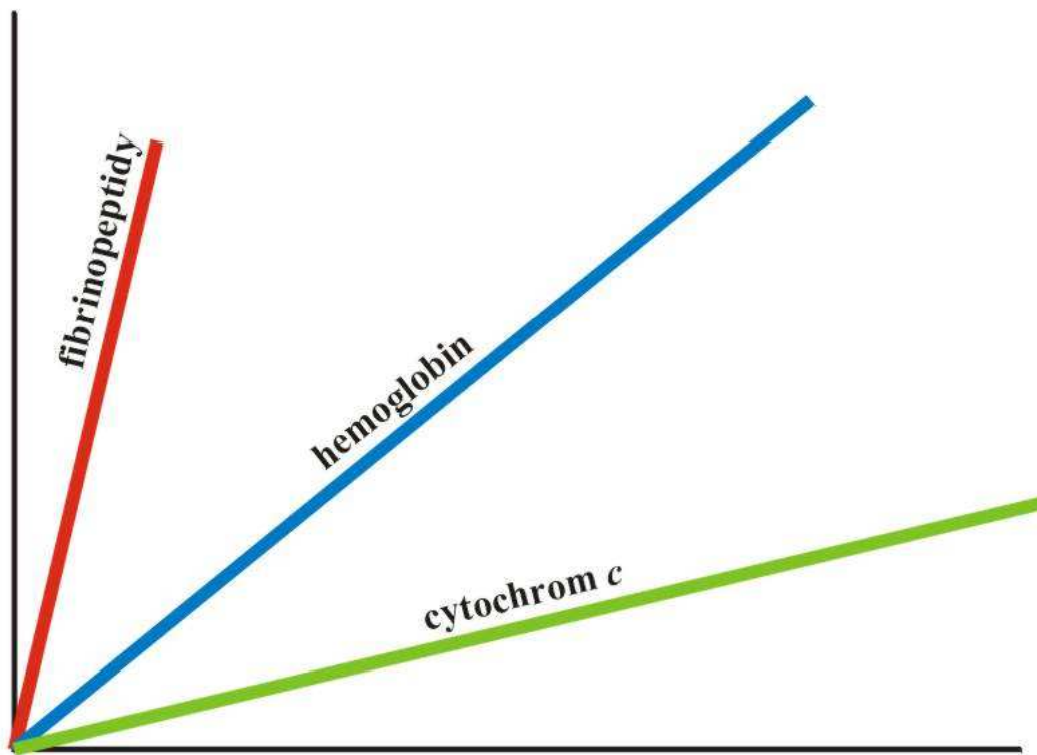
1. většina mutací, které se projeví v evoluci, je neutrální ( $\Rightarrow$  drift)



## neutrální teorie:

### 2. rozdílná evoluční rychlost u různě důležitých proteinů

Počet substitucí aminokyselin na 100 molekul



Doba divergence (miliony let)

fibrinopeptidy	8,3
pankreatická ribonukleáza	2,1
lyzozym	2,0
alfa-globin	1,2
inzulin	0,44
cytochrom c	0,3
histon H4	0,01

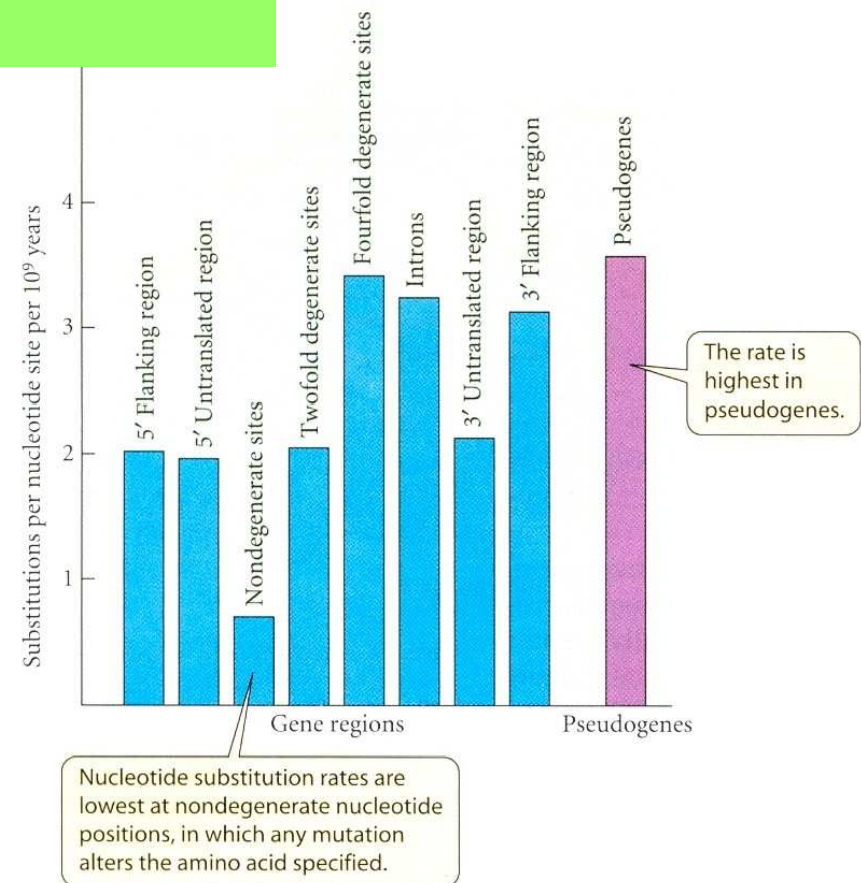
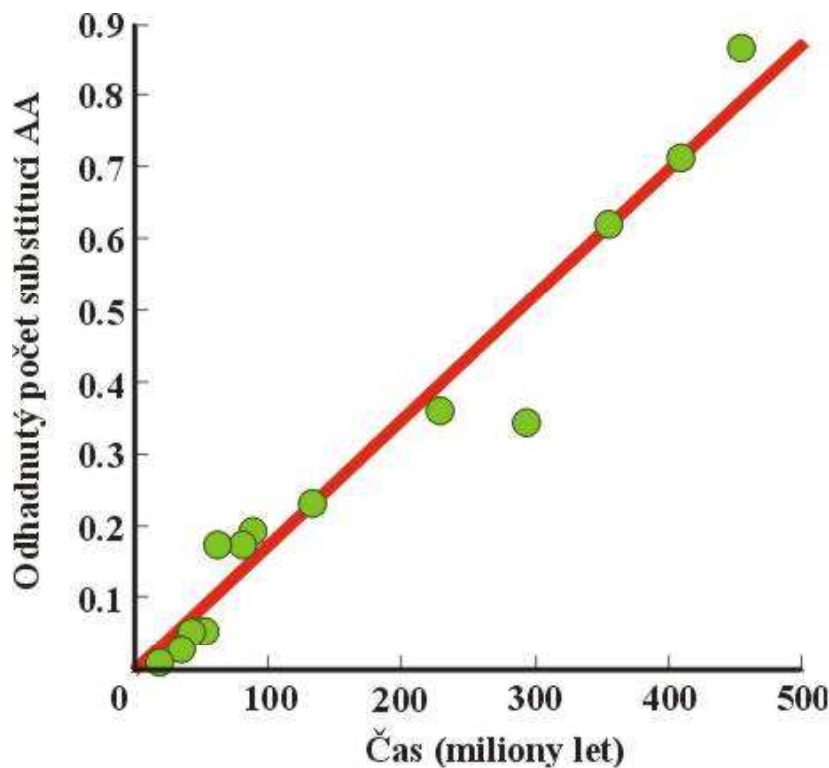


## neutrální teorie:

3. rozdílná evoluční rychlost na různých částech proteinu (vazebná místa × strukturní oblasti)

4. rozdílná evoluční rychlost na jednotlivých místech kodonu

5. rychlost evoluce daného proteinu u různých druhů přibližně konstantní



převážně se netýká morfologických, fyziologických a behaviorálních znaků

nemůže vysvětlit vznik adaptací

mnoho škodlivých mutací, ty však rychle eliminovány selekcí

selekce působí i na molekulární úrovni, avšak většina mutací má velmi malý účinek na fitness  $\Rightarrow$  důležitá role driftu

**Haldaneův odhad selekčních nákladů nadhodnocený:**

selekce většinou *měkká*

*frekvenčně závislá* selekce místo superdominance

selekce *nepůsobí* na jednotlivé lokusy *odděleně*

# Teoretické principy neutrální teorie:

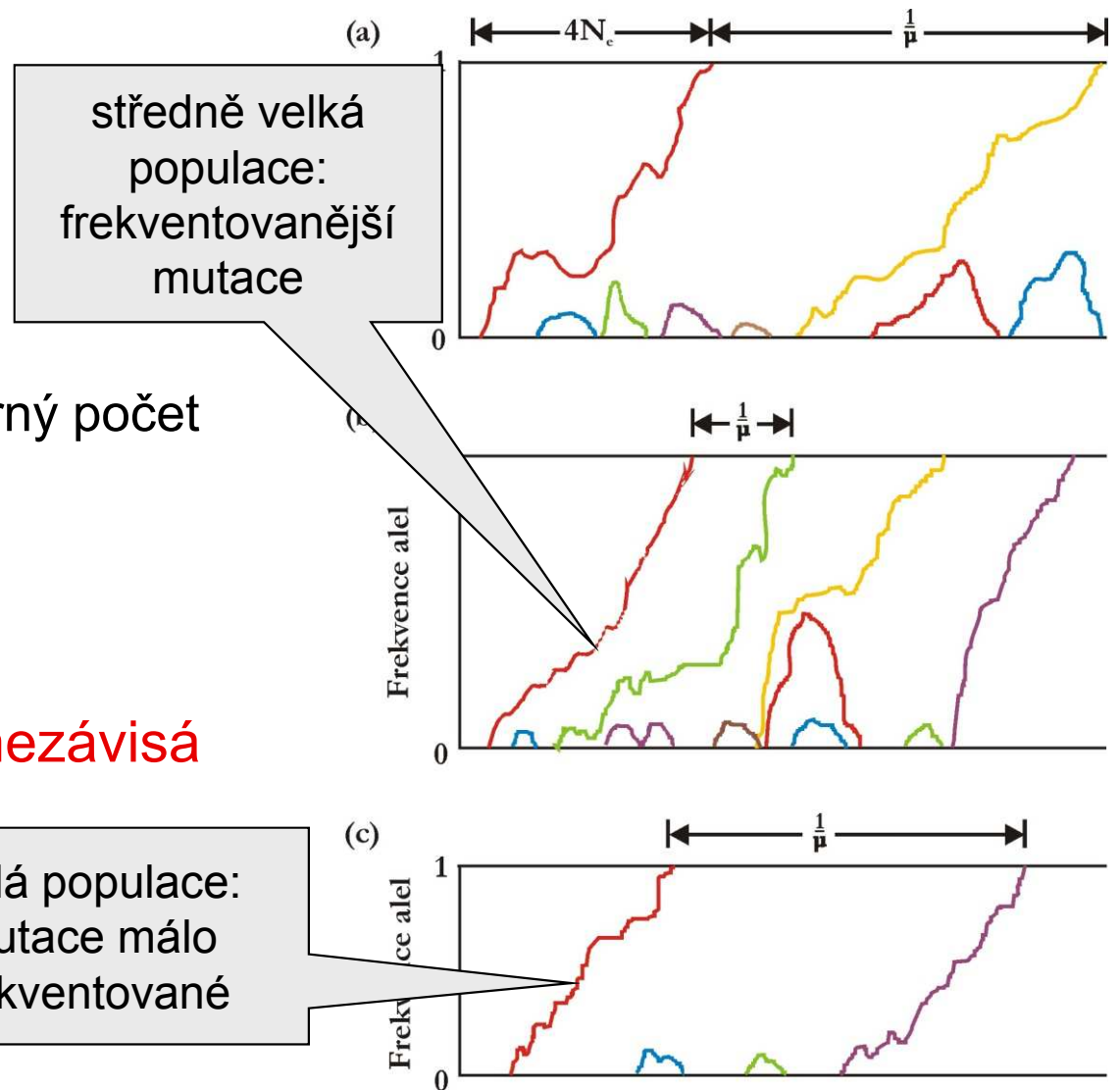
1. Pravděpodobnost fixace nové mutace =  $1/(2N_e)$

2. Frekvence nahrazení jedné alely za jinou v populaci:

pravděpodobnost fixace  $\times$  průměrný počet neutrálních mutací:

$$1/(2N_e) \times 2N_e\mu = \mu$$

$\Rightarrow$  rychlost neutrální evoluce nezávisá na  $N_e$ , ale jen na  $\mu$ !



## Teoretické principy neutrální teorie:

3. Průměrná doba mezi následujícími neutrálními mutacemi =  $1/\mu$

4. Doba fixace =  $4N_e$  generací

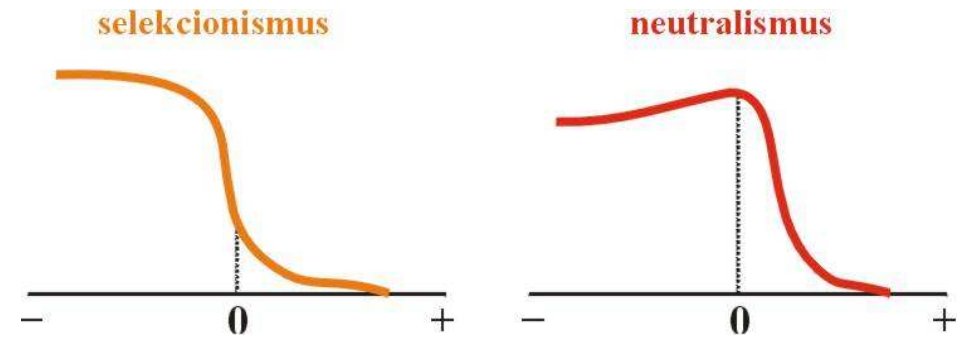
5. Průměrná rovnovážná heterozygotnost:  $\frac{\theta}{\theta + 1}$ , kde  $\theta = 4N_e\mu$

větší populace  $\Rightarrow$  vyšší heterozygotnost

neustálý vznik nových mutací  $\Rightarrow$  zvýšení proměnlivosti  $\times$  její eroze driftem  
 $\Rightarrow$  neustálé nahrazování jedné alely za druhou

$\rightarrow$  rovnováha mutace a driftu  $\Rightarrow$  polymorfismus  
(na rozdíl od rovnováhy mutace a selekce je **přechodný**)

# Frekvence neutrálních mutací:

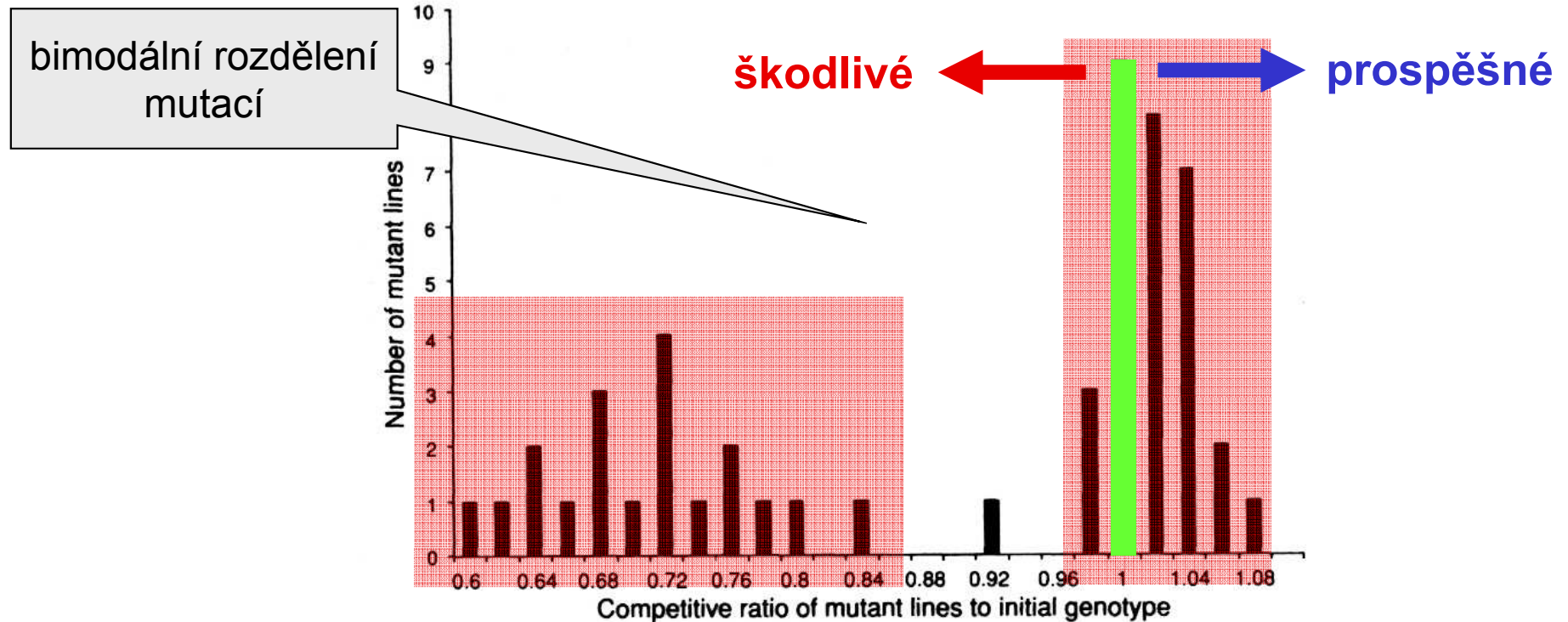


Zeyl & DeVisser (2001):

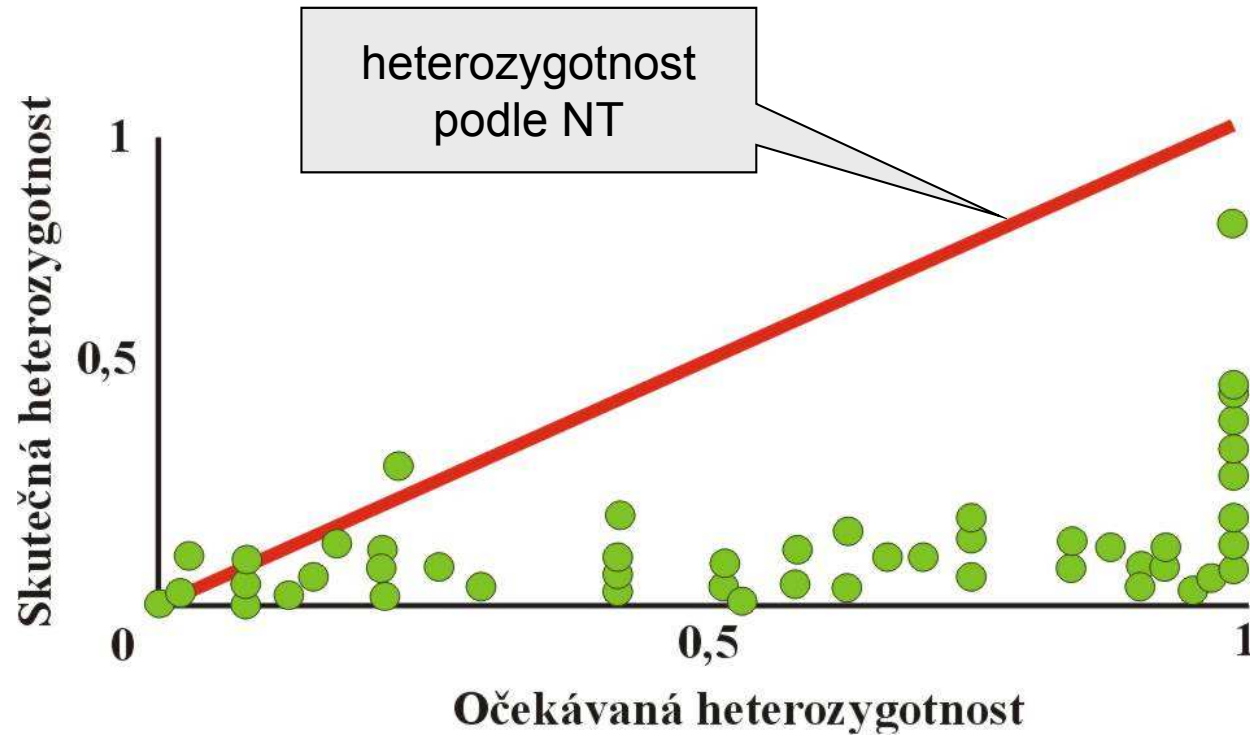
kvasinka *Saccharomyces cerevisiae*

50 replikací populace

experiment nezachycuje extrémně škodlivé mutace

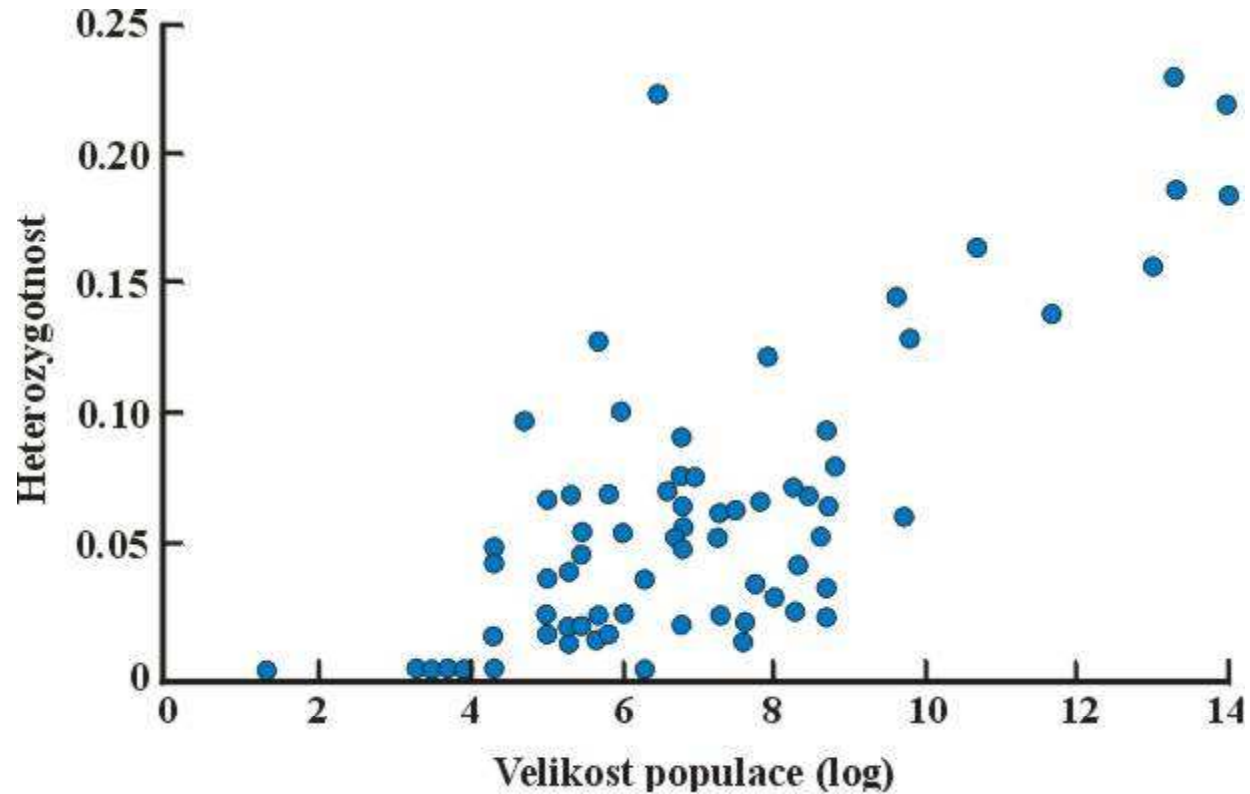


## Test neutrální teorie: rozsah heterozygotnosti



Skutečná **heterozygotnost nižší**, než předpokládá NT

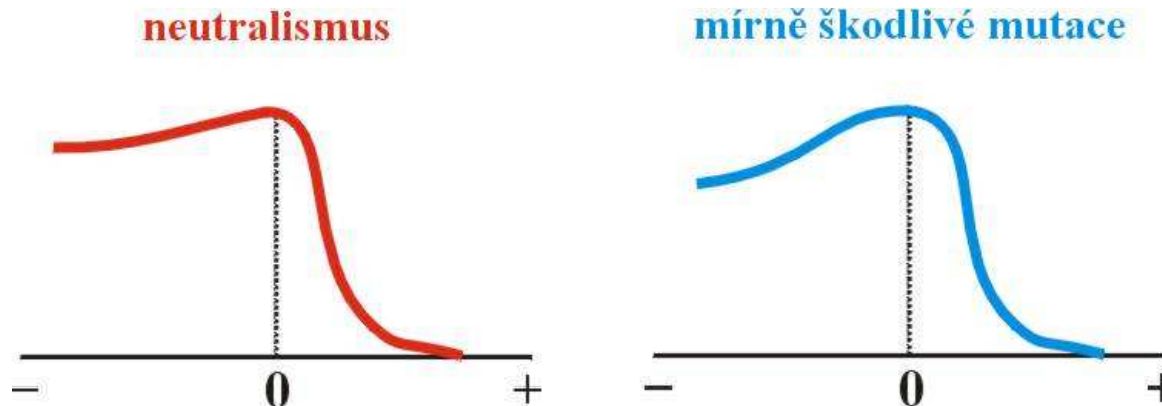
## Test neutrální teorie: rozsah heterozygotnosti



Vzhledem k obrovskému rozsahu populačních velikostí,  
rozsah heterozygotností příliš malý

Odchytky měření rozsahu heterozygotnosti od predikcí se snažila vysvětlit Tomoko Ohtová:

mírně škodlivé mutace (slightly deleterious mutations, SDM): v malých populacích se chovají jako neutrální (= efektivně neutrální alely)



Pravděpodobnost fixace neutrální, výhodné a škodlivé mutace:

$$P = \frac{1 - e^{-4N_e s q}}{1 - e^{-4N_e s}}$$

$q$  = frekvence alely,  $s$  = selekční koeficient

pravděpodobnost fixace neutrální mutace =  $q$   
nebo  $1/(2N)$  v době vzniku

pravděpodobnost fixace výhodné mutace ve velké populaci  $\approx 2s$



Př.: Jaká je pravděpodobnost fixace mutace v populaci o  $N_e = 1000$ ?

neutrální mutace ( $s = 0$ ):

$$P = 0,05\%$$

výhodná mutace ( $s = 0,01$ ):

$$P = 20\%$$

výhodná mutace ( $s = 0,001$ ):

$$P = 2\%$$

škodlivá mutace ( $s = -0,001$ )

$$P = 0,004\%$$

čím víc  $s \rightarrow 0$ ,  
tím vyšší  
„neutralita“

Z toho plyne, že

všechny výhodné mutace nemusí být v populaci zafixovány a i škodlivé mutace mohou být s malou pravděpodobností zafixovány

Př.: Jaká je pravděpodobnost fixace mutace v populaci o  $N_e = 10\ 000$ ?

neutrální mutace ( $s = 0$ ):

$$P = 0,005\%$$

výhodná mutace ( $s = 0,01$ ):

$$P = 20\%$$

výhodná mutace ( $s = 0,001$ ):

$$P = 2\%$$

škodlivá mutace ( $s = -0,001$ )

$$P = 2 \cdot 10^{-17}\%$$

ve velké populaci je  $P$   
výhodné alely stejná jako v  
malé, ale pro škodlivou  
alelu  $P \rightarrow 0$

Z uvedeného plyne:

ve velkých populacích hraje mnohem větší roli selekce

existuje nepřímá úměra mezi škodlivostí mutace a velikostí populace: čím se škodlivost alely blíží nule, tím větší může být populace, ve které se může fixovat (drift převýší negativní selekci) a naopak, čím je selekce proti škodlivé mutaci silnější, tím menší musí být populace, aby drift hrál určující roli

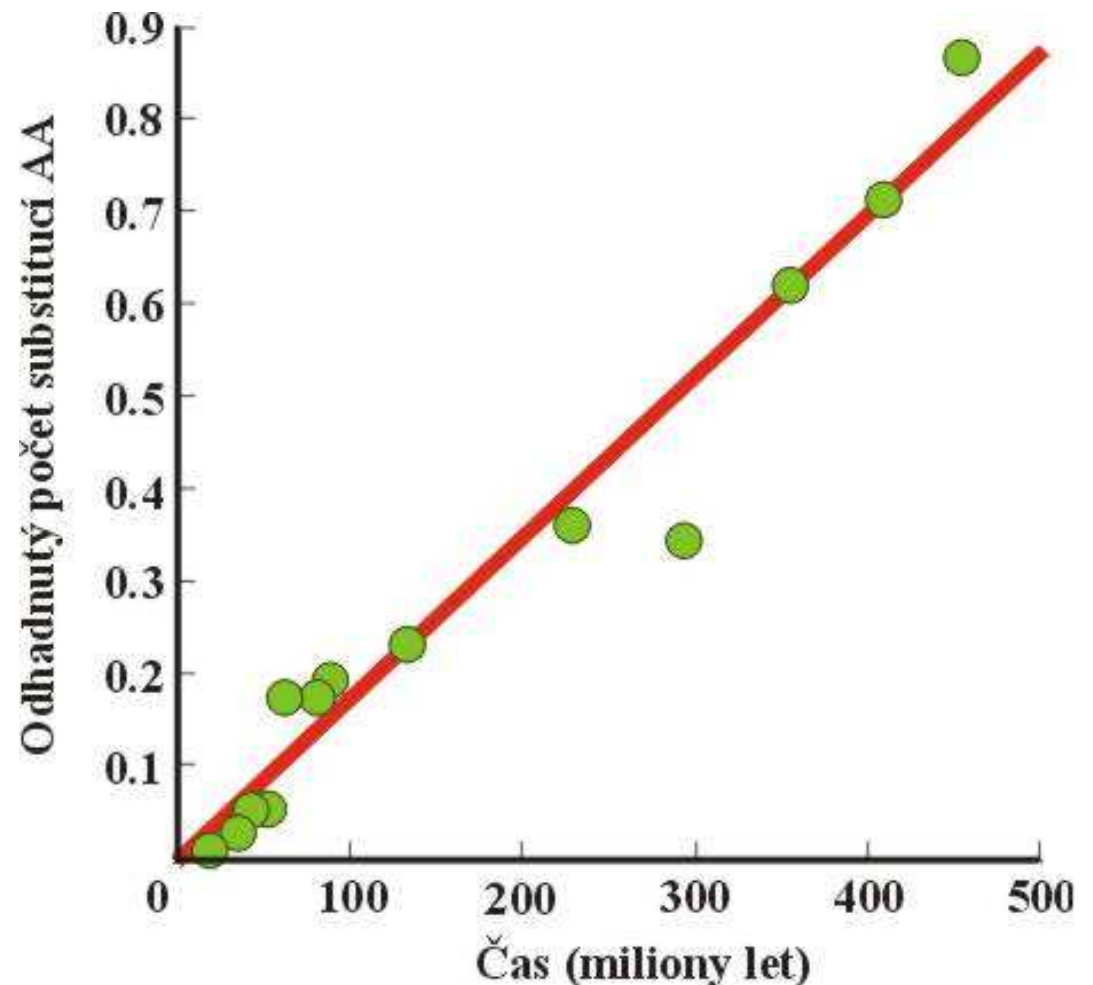
v malých populacích se mírně škodlivé mutace chovají jako  
**efektivně neutrální**

# Molekulární hodiny

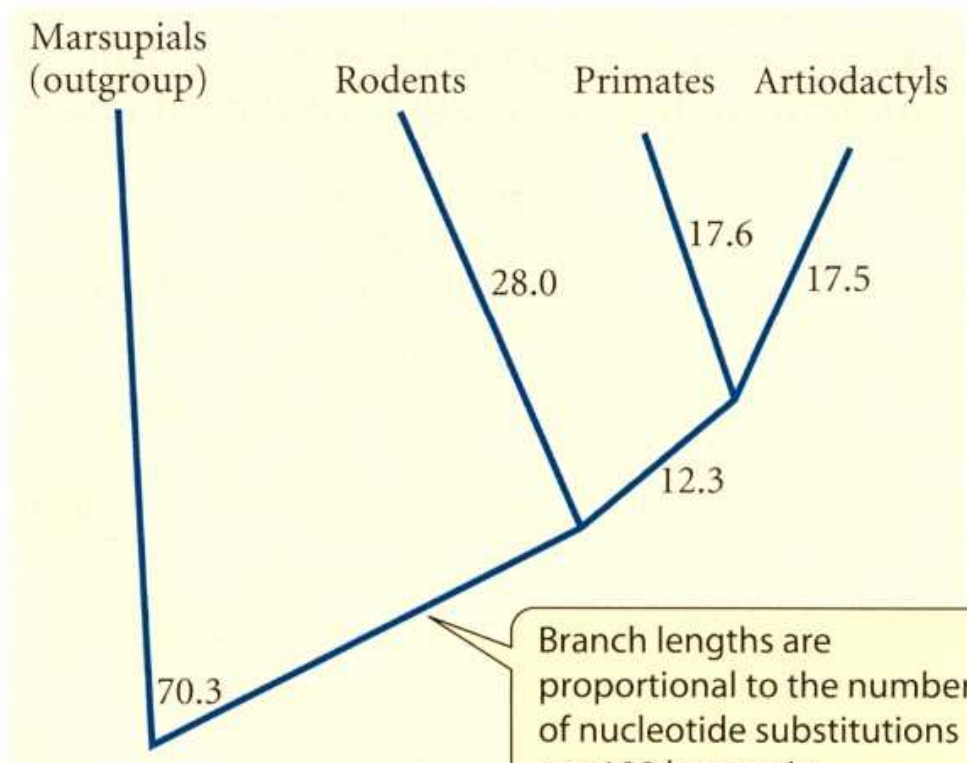
Zuckermandl & Pauling (1962-65)

rychlost substitucí AA nebo nukleotidů je konstantní

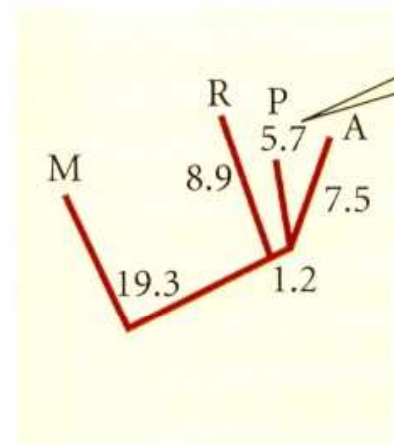
efekt generační doby:  
závislost na absolutním  
nebo generačním čase?



# Molekulární hodiny



Branch lengths are proportional to the number of nucleotide substitutions per 100 base pairs.



Nonsynonymous substitution has been slowest in the primate lineage.

# EVOLUCE GENOMU

Velikost genomu a cytoplazmatický poměr (C-value):  
C-value = množství DNA v haploidním genomu (pg, bp)

## Prokaryota:

$6 \times 10^5 - 10^7$  bp (20×)

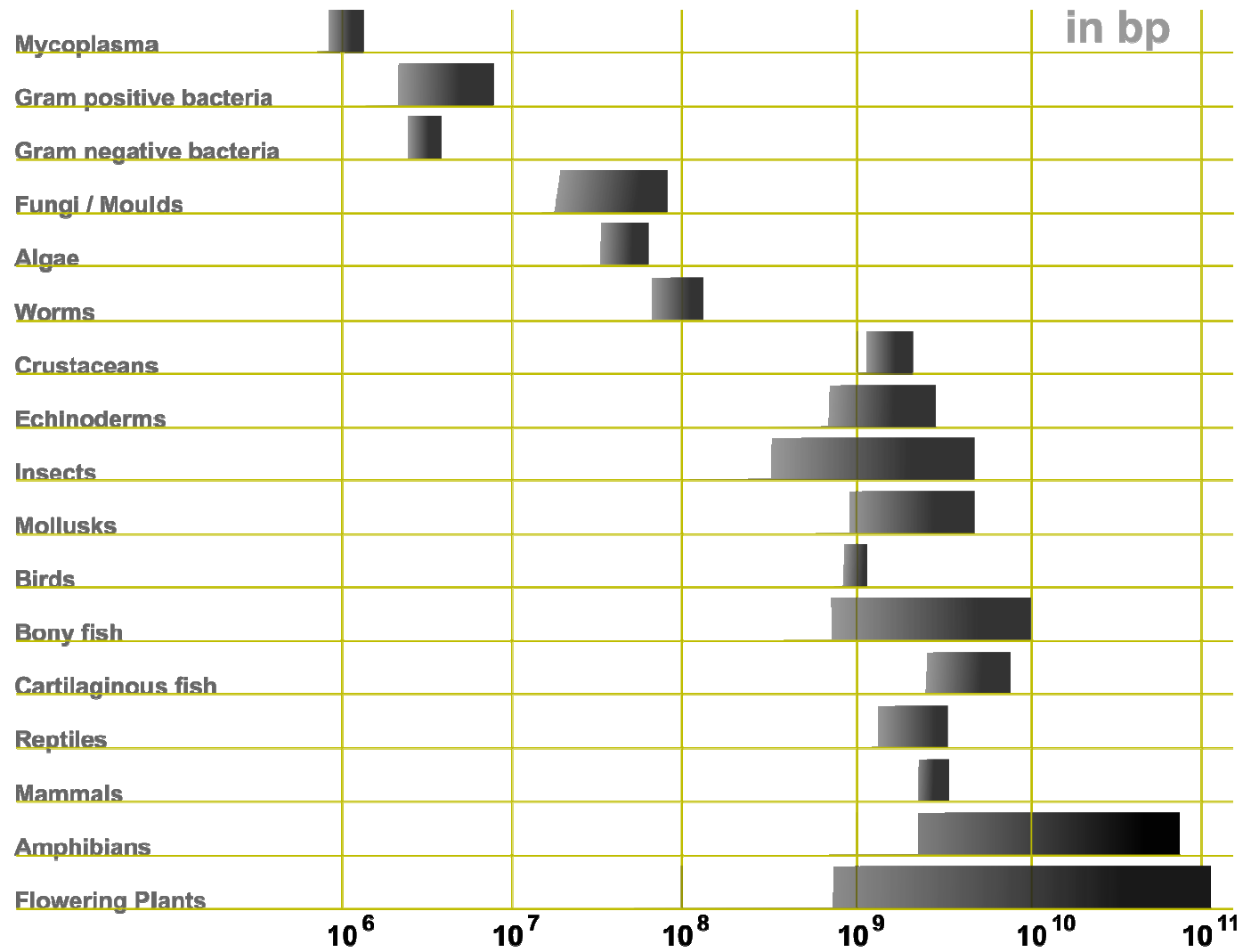
nejmenší: *Mycoplasma* (celkem ca. 400 genů)

největší: někt. G+ bakterie, sinice

## Eukaryota:

$8,8 \times 10^6 - 6,9 \times 10^{11}$  bp (80 000×)

# EVOLUCE GENOMU



# EVOLUCE GENOMU

Velikost genomu a cytoplazmatický poměr (C-value):

C-value = množství DNA v haploidním genomu (pg, bp)

žádný vztah ke složitosti organismu nebo počtu genů

velké rozdíly i u příbuzných organismů:

*Paramecium caudatum* (8 600 000 kb) × *P. aurelia* (190 000 kb)

člověk: ca.  $6 \times 10^9$  bp (~ 6,5 pg DNA)

× *Amoeba proteus*:  $2,9 \times 10^{11}$  bp

*Polychaos dubium* (*Amoeba dubia*):  $6,7 \times 10^{11}$  bp

⇒ C-value paradox (C-value enigma)

## Kolik genů obsahuje genom člověka?

před 2001 (hrubá verze sekvence lidského genomu) odhady od 50 000 po > 140 000 (max. 212 278) genů

Int. Human Genome Sequencing Consortium 2001: 30 000 – 40 000 protein kódujících genů

IHGSC 2004: 20 000 – 25 000 protein kódujících genů

Ensembl – květen 2012: 21 065 kódujících genů

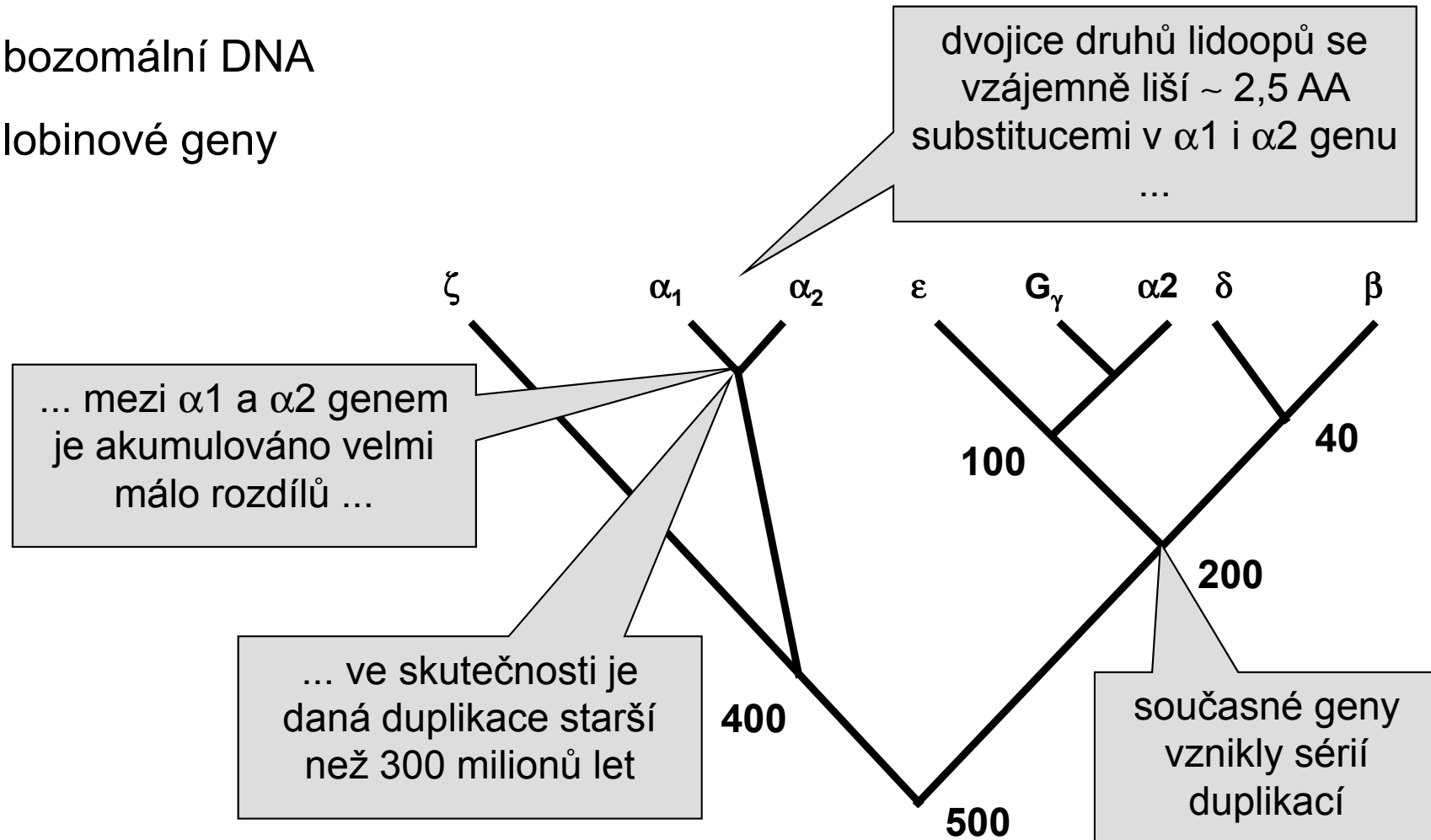
Ensembl – leden 2013: 20 848 genů

Ensembl – únor 2014: **20 805 genů**



# Spojená evoluce a molekulární tah

ribozomální DNA  
globinové geny

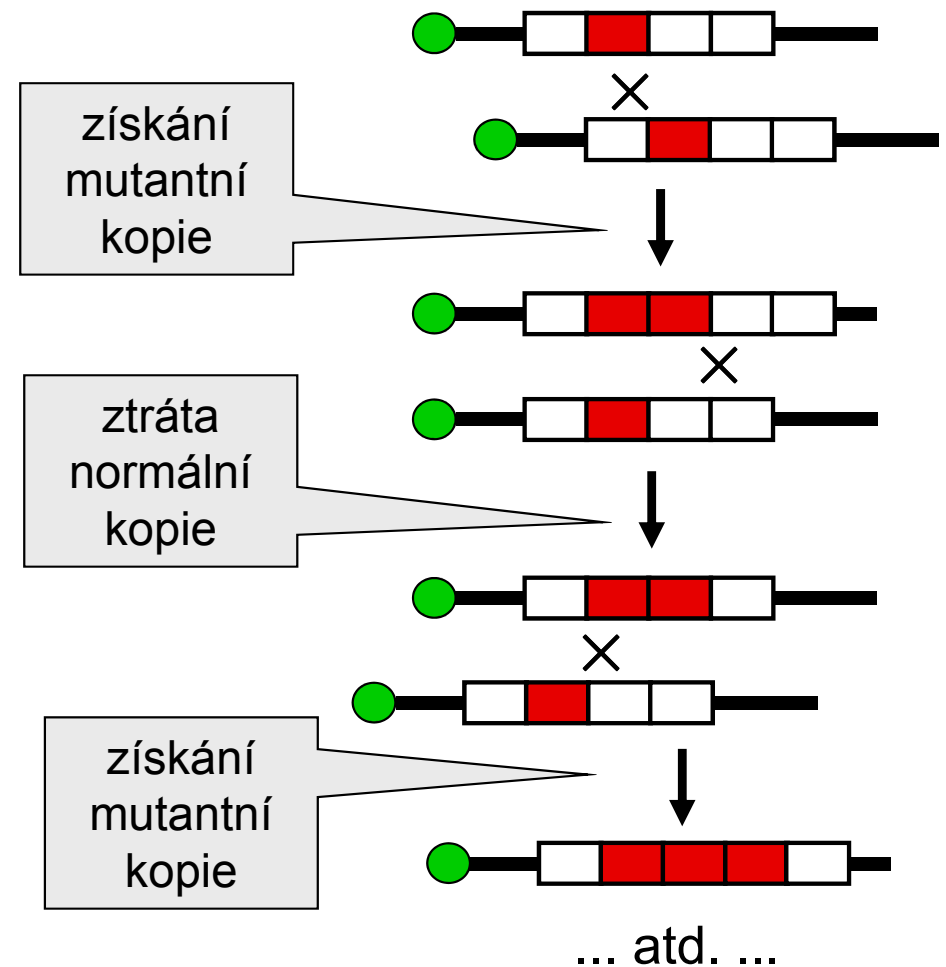
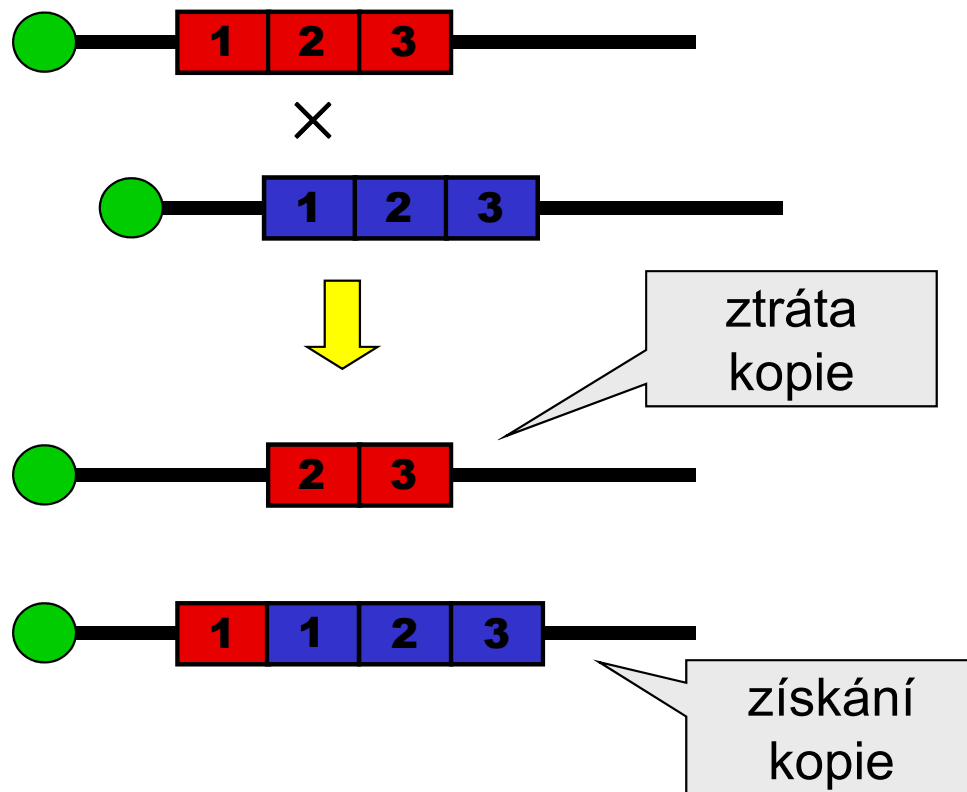


⇒ molekulární hodiny v tomto případě neplatí, geny se nevyvíjí nezávisle – evoluce je **spojená**

Gabriel Dover (1982): **Molekulární tah (molecular drive)**  
mechanismus odlišný od selekce a driftu

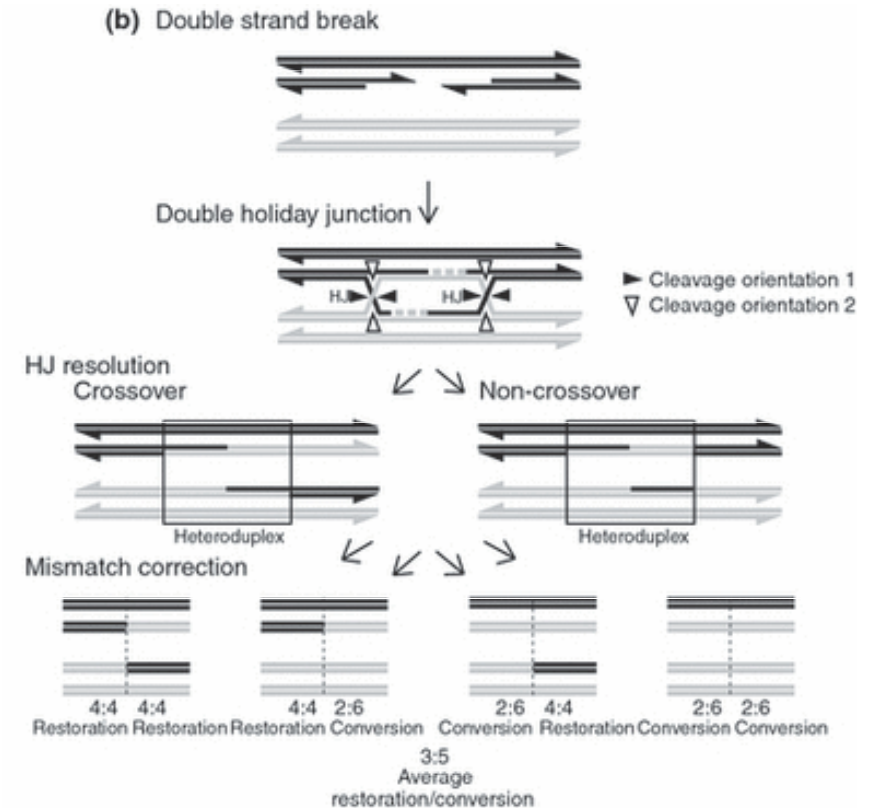
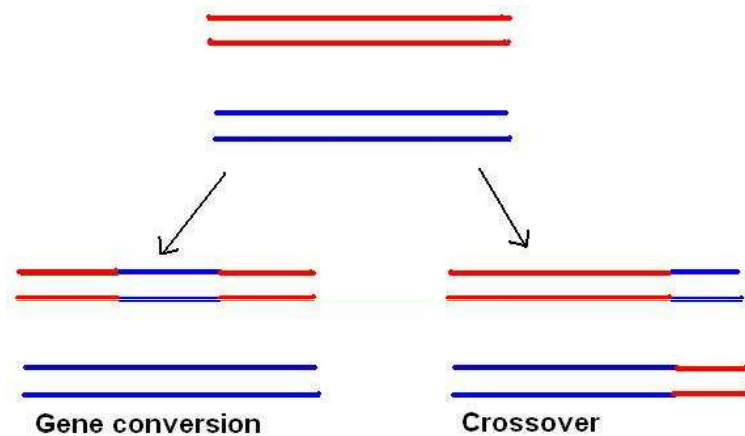
## Mechanismy spojené evoluce:

### 1. nestejný crossing-over



## 2. sklouznutí nukleotidového řetězce (slippage)

## 3. genová konverze



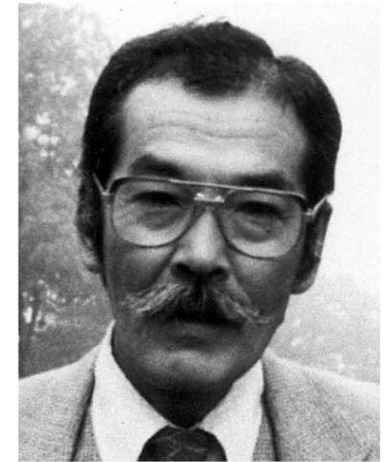
Závěr:

důsledkem nestejněměrného crossing-overu a sklouznutí řetězce je  
změna počtu kopií

důsledkem nestejněměrného c-o a genové konverze je  
homogenizace sekvencí

# Repetitivní DNA:

1. Vysoce repetitivní = satelitní
2. Středně repetitivní = minisatelity, mikrosatelity
3. Transpozabilní elementy, retroelementy (SINE, LINE)



*Susumu Ohno*

Proč existuje repetitivní DNA?

nějaká funkce

Doolittle a Sapienza, Orgel a Crick (1980): repetitivní DNA je „sobecká“  
Susumu Ohno (1972): pojem „**junk DNA**“

význam pojmu „junk“ = „harampádí“ (např. garáž plná h.), ne „odpad“ („garbage“) ... François Jacob (1977): evoluční dráteničina (fušeřina) = „evolutionary tinkering“

⇒ v budoucnu může nabýt funkci