

# Úvod do populační genetiky

# Populační genetik

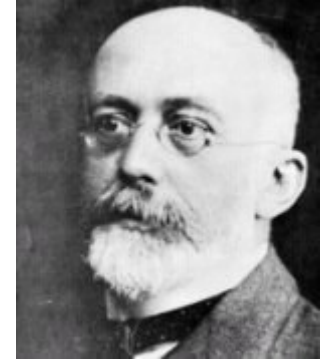
- Obor zabývající se genetickou variabilitou v populacích a procesy, které ji ovlivňují.
- Snaha vysvětlit evoluci pomocí mechanismů měnících frekvence alel v populaci.



# Hardy-Weinbergův zákon



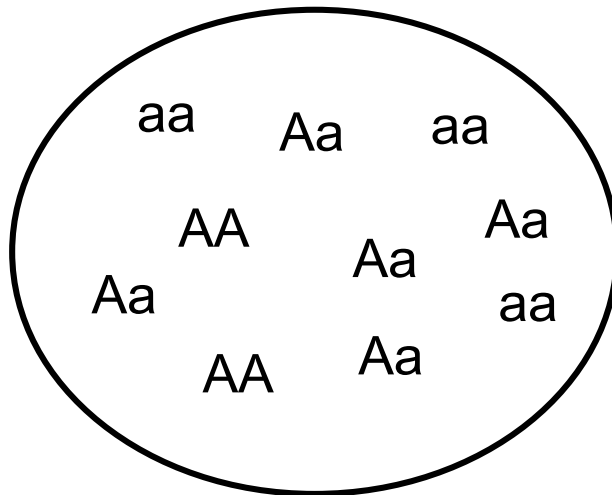
Godfrey Hardy



Wilhelm Weinberg

$$p^2 + 2pq + q^2 = 1$$

- p ..... frekvence alely A  
q ..... frekvence alely a  
p<sup>2</sup> ..... frekvence genotypu AA  
q<sup>2</sup> ..... frekvence genotypu aa  
2pq ..... frekvence genotypu Aa



$$p^2 = 2/10 = 0,2$$

$$q^2 = 3/10 = 0,3$$

$$2pq = 5/10 = 0,5$$

$$p = 9/20 = 0,45$$

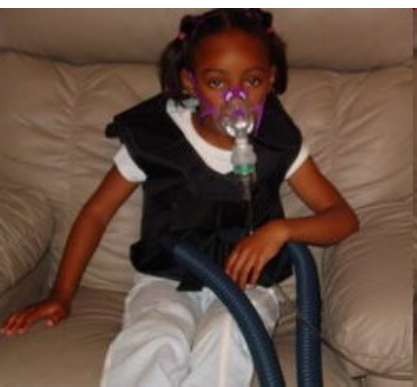
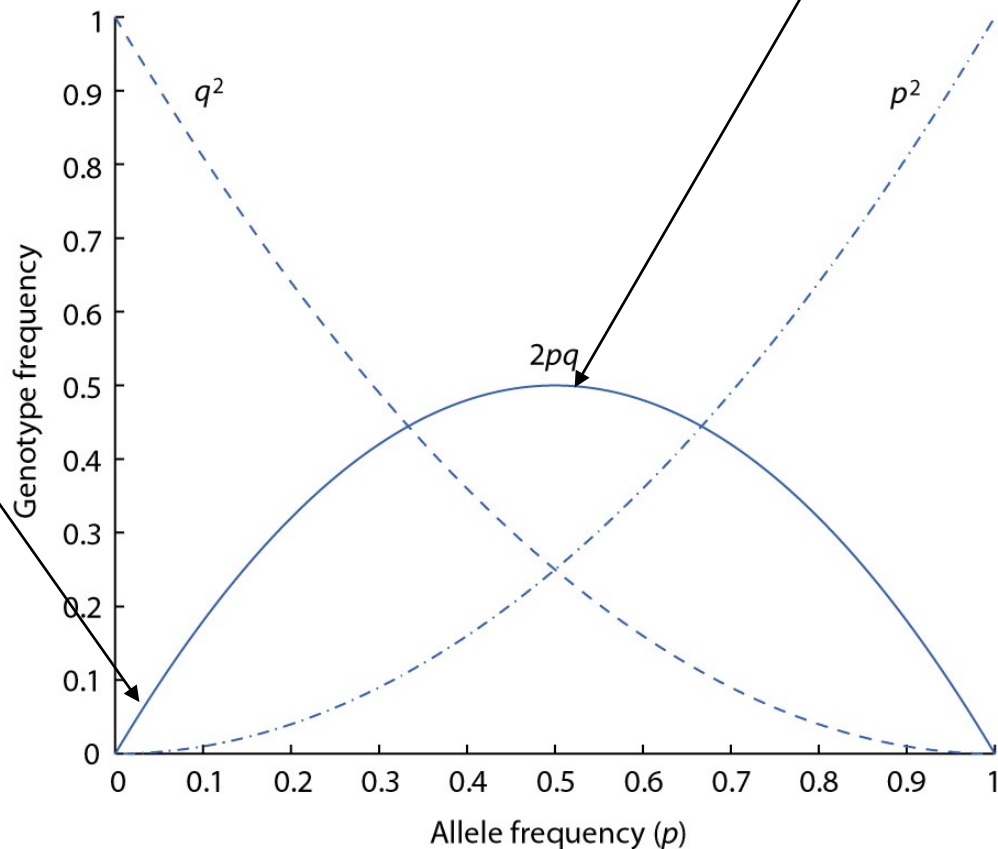
$$q = 11/20 = 0,55$$

# Hardy-Weinbergův zákon

Přetrvávání  
škodlivých  
recesivních  
mutací v  
populaci.

Vzácné alely skryty  
v heterozygotech.

Nejvíce heterozygotů v populaci  
při stejné frekvenci obou alel.



Cystická fibróza  
recesivní homozygoti 1 z 1700  
heterozygoti 1 z 24

# Hardy-Weinbergův zákon pro více alel

2 alely

$$(p + q)^2 = 1$$

$$p^2 + 2pq + q^2 = 1$$

3 alely

$$(p + q + r)^2 = 1$$

$$p^2 + q^2 + r^2 + 2pq + 2pr + 2qr = 1$$

4 alely

$$(p + q + r + s)^2 = 1$$

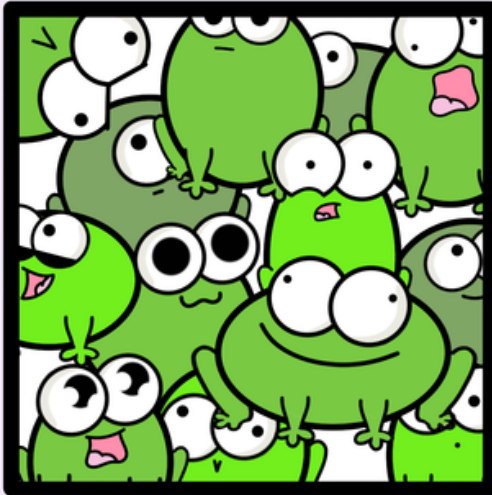
$$p^2 + q^2 + r^2 + s^2 + 2pq + 2pr + 2ps + 2qr + 2qs + 2rs = 1$$

atd.

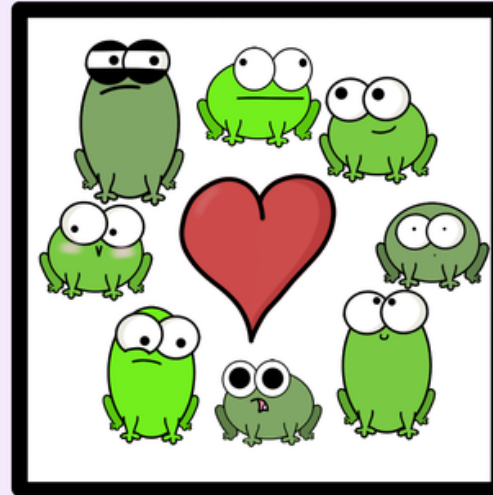
# Předpoklady Hardy-Weinbergovy rovnováhy

- Dostatečně velká populace, ve které se jedinci párují zcela náhodně bez ohledu na genotyp/fenotyp jedinců (**panmiktická populace**). Nehraje roli vzdálenost mezi jedinci. Neprobíhá selekce.
- Stačí jedna generace náhodného křížení, aby se ustavila HW rovnováha.
- To, zda je populace v HW rovnováze se může lišit pro různé geny v genomu.

4. Large Population



5. Random Mating



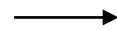
**Jak zjistíme, zda je populace v Hardy-Weinbergově rovnováze?**

<b>Genotypy</b>	<b>Pozorované počty</b>
AA	125
Aa	550
aa	310

# Testování odchylky od Hardy-Weinbergovy rovnováhy

## Odhad frekvence alel

Genotypy	Pozorované počty
AA	125
Aa	550
aa	325
<u>celkem</u>	<u>1000</u>



$$f_A = \frac{2 \times 125 + 550}{2 \times 1000} = 0,4$$
$$f_a = 1 - A = 0,6$$

## Odhad očekávaných počtů genotypů

AA	$1000 \times (0,4)^2 = 160$
Aa	$1000 \times 2(0,4)(0,6) = 480$
aa	$1000 \times (0,6)^2 = 360$



## Výpočet rozdílů pozorovaných a očekávaných počtů genotypů

AA	$125 - 160 = -35$
Aa	$550 - 480 = 70$
aa	$325 - 360 = -35$

## Chi-square statistika ( $X^2$ )

$$X^2 = \sum \frac{(\text{pozorované} - \text{očekávané})^2}{\text{očekávané}}$$

## Počet stupňů volnosti (df)

- počet porovnávaných tříd
- počet odhadovaných parametrů
- 1

## Chi-square test

$$X^2 = \frac{(-35)^2}{160} + \frac{(70)^2}{480} + \frac{(-35)^2}{360} = 21,27$$

$$df = 3 - 1 - 1 = 1$$

$$p < 0.01$$

**Populace není v HW rovnováze.**



# Mechanismy způsobující odchylku od HW rovnováhy

## Nenáhodné (asortativní) párování

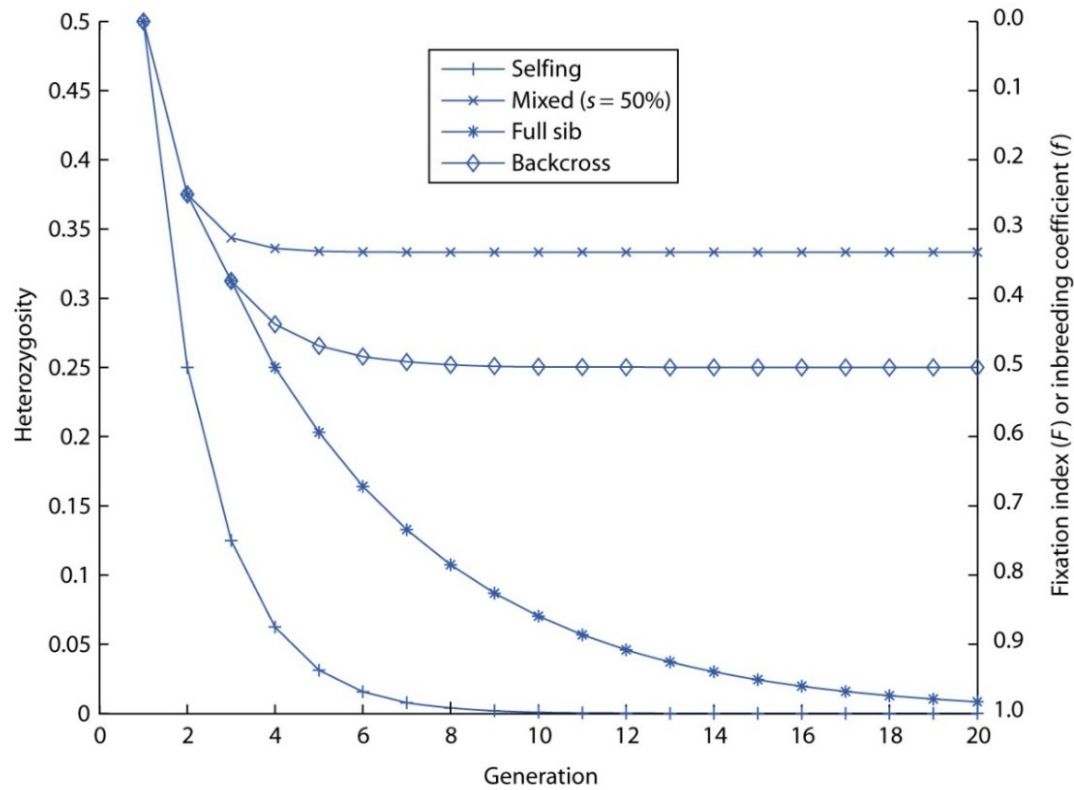
### Pozitivní asortativní párování

- Jedinci s podobnými genotypy se vzájemně párují s větší pravděpodobností. Vede k přebytku homozygotů v populaci.
- Např. příbuzenské křížení (inbreeding)
- Vede k odhalení škodlivých recesivních mutací (inbrední deprese).
- V přírodě dochází k inbreedingu také v extrémně malých populacích.



# Inbreeding

- V laboratoři se využívá při vytváření inbredních kmenů.

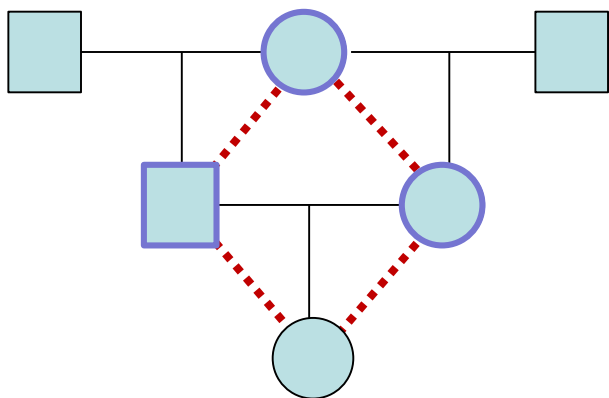


## Koeficient inbreedingu (F)

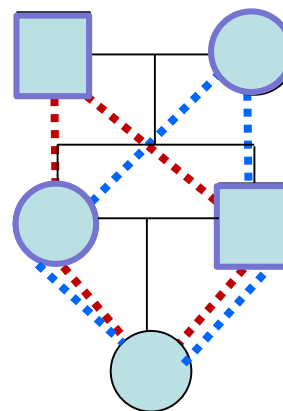
Pravděpodobnost, že jedinec zdědil obě alely téhož genu od jednoho předka.

$$F = (1/2)^n$$

n = počet jedinců v genealogii od daného jedince ke společnému předkovi



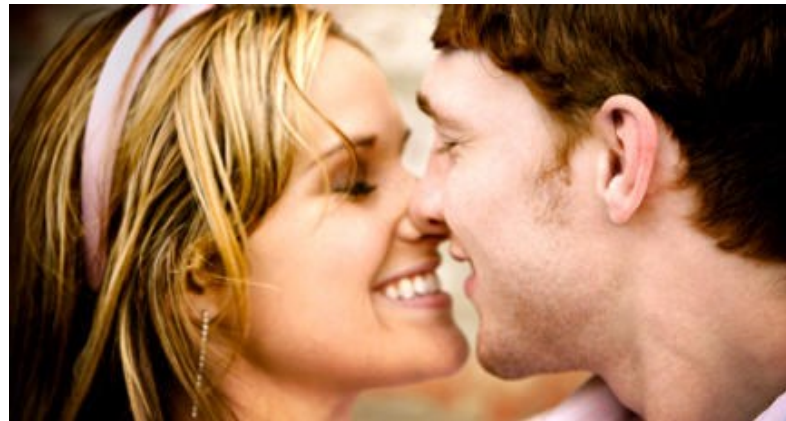
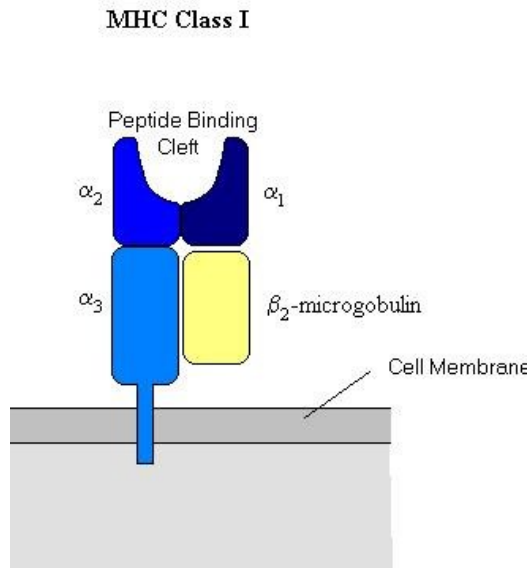
$$F = (1/2)^3 = 1/8$$



$$F = 1/8 + 1/8 = 1/4$$

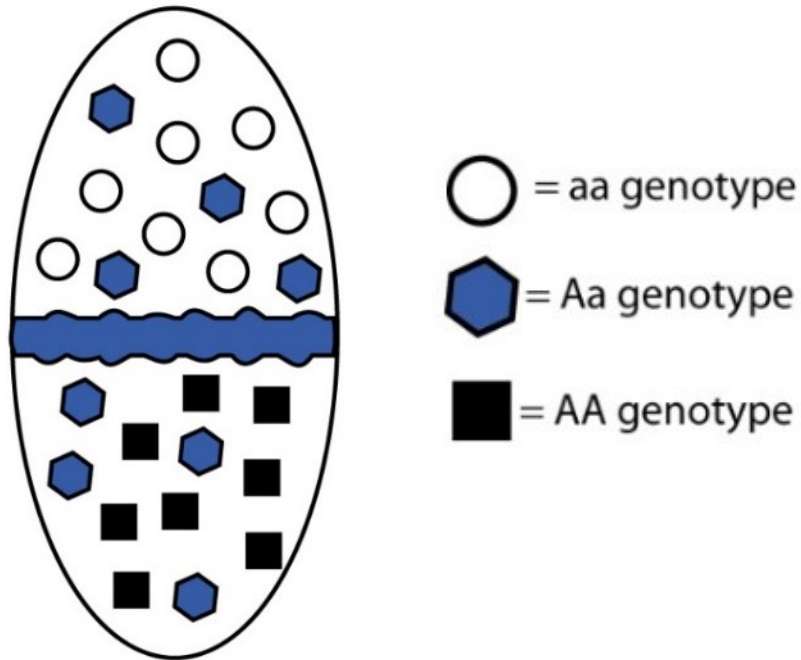
## Negativní asortativní párování

- Jedinci s odlišnými genotypy se vzájemně párují s větší pravděpodobností.
- Vede k přebytku heterozygotů.
- V přírodě nastává např. pokud jsou zvýhodněni heterozygoti.
  - MHC I (*Major Histocompatibility Complex I*) geny



## Geografická struktura populace

- Brání náhodnému párování mezi jedinci.
- Snížení celkové heterozygotnosti ve strukturované populaci = **Wahlundův princip**.



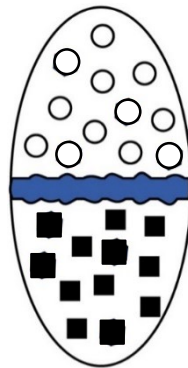
# Odhad míry geografické strukturovanosti populací

## $F_{ST}$ statistika

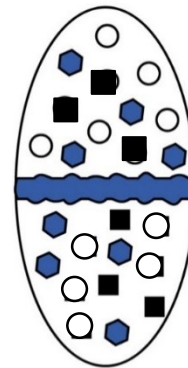
Proporce o kterou je snižená heterozygotnost subpopulací ve srovnání s očekávanou heterozygotností za předpokladu náhodného křížení mezi všemi jedinci.

$$F_{ST} = \frac{H_T - H_S}{H_T}$$

$H_T$  očekávaná heterozygotnost pro celkovou populaci  
 $H_S$  očekávaná heterozygotnost pro subpopulace



$F_{ST} = 1$



$F_{ST} = 0$

- = aa genotype
- ⬡ = Aa genotype
- = AA genotype

## Příklad

Jaké je  $F_{ST}$  mezi dvěma populacemi?

	Počty genotypů			
	AA	Aa	aa	
subpopulace 1	25	50	25	100
subpopulace 2	49	42	9	100
celková populace	74	92	34	200

Frekvence alel	
p	q
0,5	0,5
0,7	0,3
0,6	0,4

$$H_T = 2pq = 2 \cdot 0,6 \cdot 0,4 = 0,48$$

$$H_S = p_1q_1 + p_2q_2 = 0,25 + 0,21 = 0,46$$

$$F_{ST} = 0,0416$$

$$F_{ST} = \frac{H_T - H_S}{H_T}$$

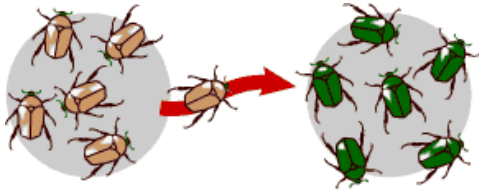
# Mechanismy měnící frekvence alel v populaci



**mutace**



**drift**



**migrace**



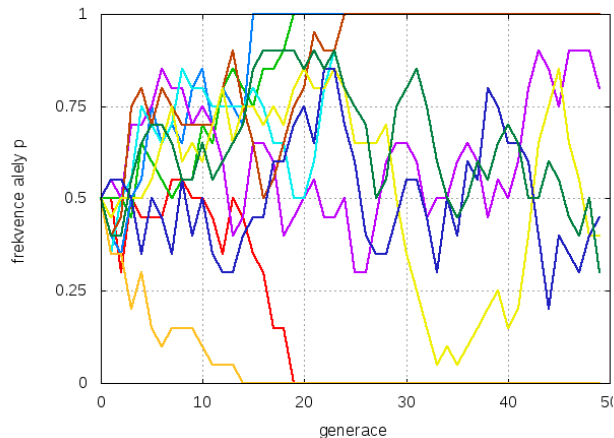
**selekce**



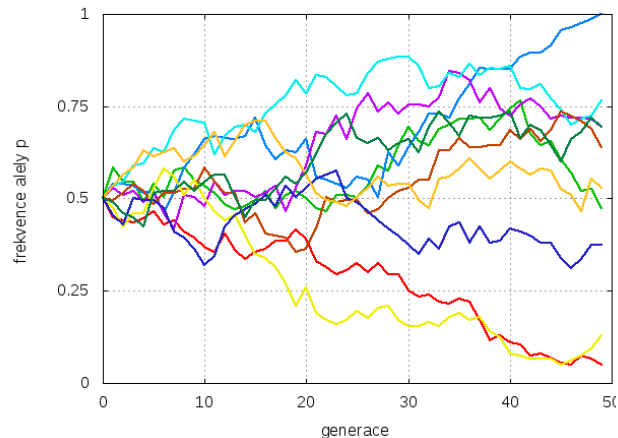
# Genetický drift

- Náhodné změny ve frekvencích alel v omezeně velkých populacích.
- Působí silněji v malých populacích.
- Ochuzuje genetickou variabilitu populace.
- Směr genetického driftu je náhodný.
- Pravděpodobnost fixace alely působením genetického driftu odpovídá její frekvenci v populaci.
- Doba fixace neutrální mutace je  $4N_e$  (generací).

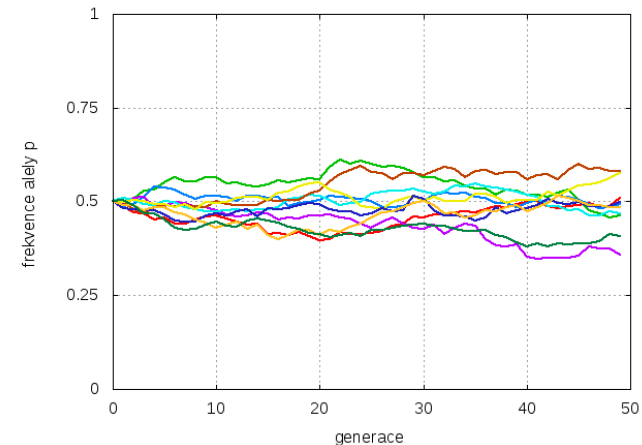
počáteční podmínky: velikost populace  $n=20$ ,  $p=0.5$



počáteční podmínky: velikost populace  $n=200$ ,  $p=0.5$

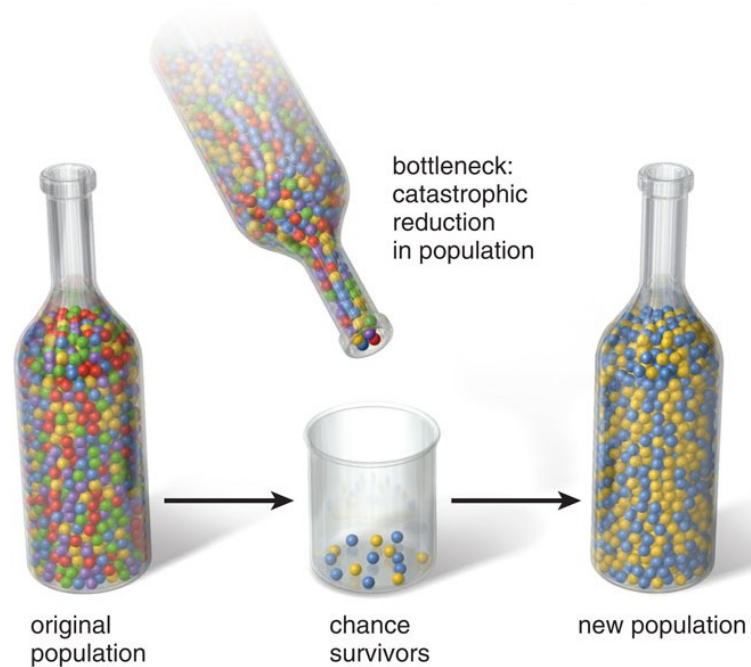
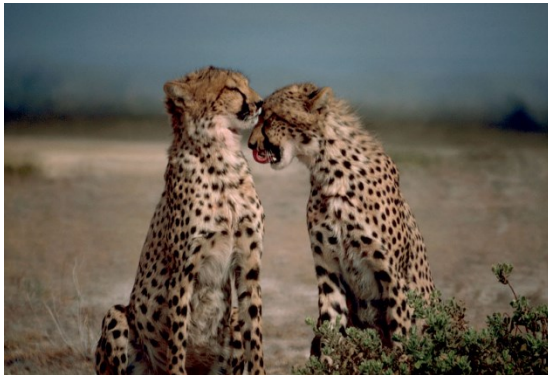


počáteční podmínky: velikost populace  $n=2000$ ,  $p=0.5$

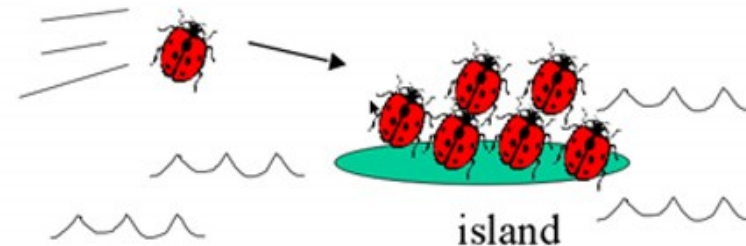


# Genetický drift působí silně i při krátkodobém zmenšení velikosti populace

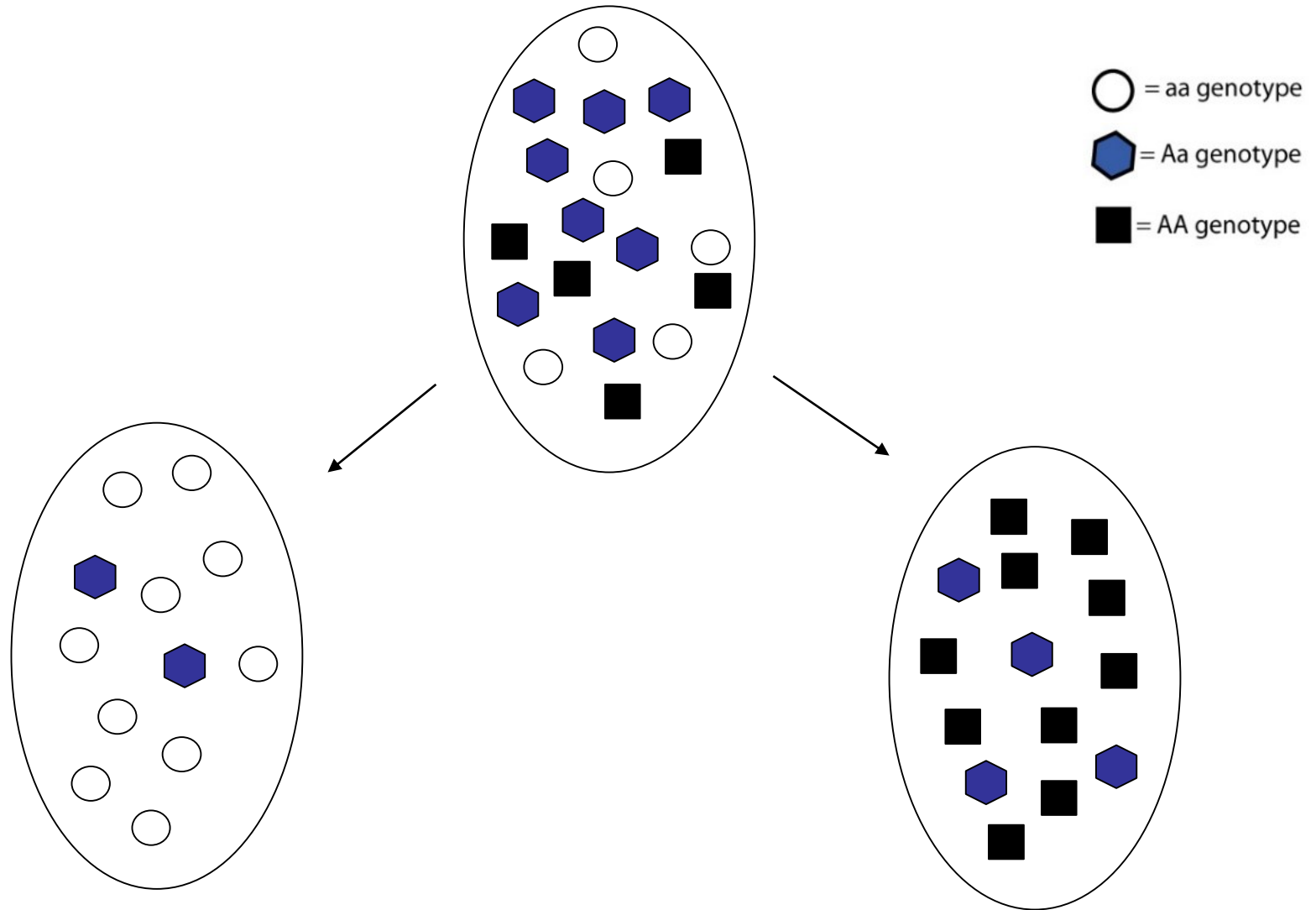
## Efekt hrdla lahve



## Efekt zakladatele



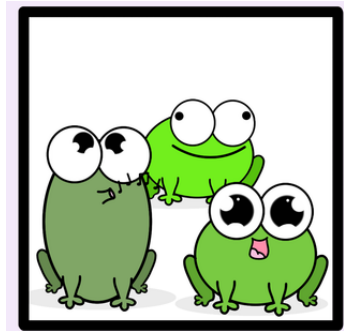
# Genetický drift způsobuje rozdílnost izolovaných populací



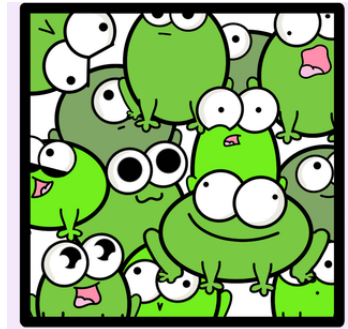
# Efektivní velikost populace ( $N_e$ )

- Vyjadřuje velikost ideální panmiktické populace, ve které genetické procesy (např. genetický drift) probíhají stejnou rychlostí jako v dané reálné populaci.

Panmiktická populace  
 $N = 100$



Reálná populace  
?

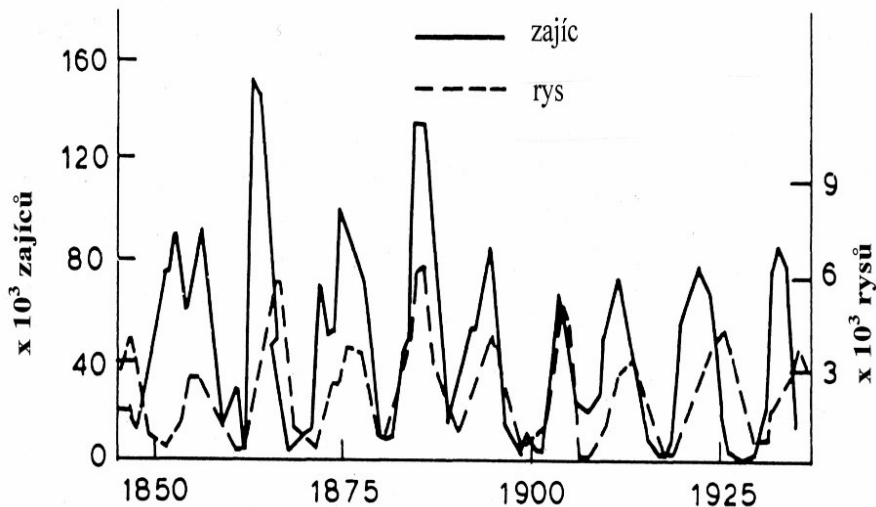


Síla genetického driftu stejná.

## Faktory ovlivňující efektivní velikost populace

### Kolísavá velikost populace, efekt hrdla lahve, efekt zakladatele

- $N_e$  odráží nejen současnou velikost populace, ale také velikost populace v minulosti.
- U populací s proměnlivou velikostí se  $N_e$  blíží spíše nižším hodnotám celkového počtu jedinců v populaci.



## Faktory ovlivňující efektivní velikost populace

### Nenáhodné křížení a rozdílný počet rozmnožujících se samic a samců

- Vede ke snížení efektivní velikosti populace.

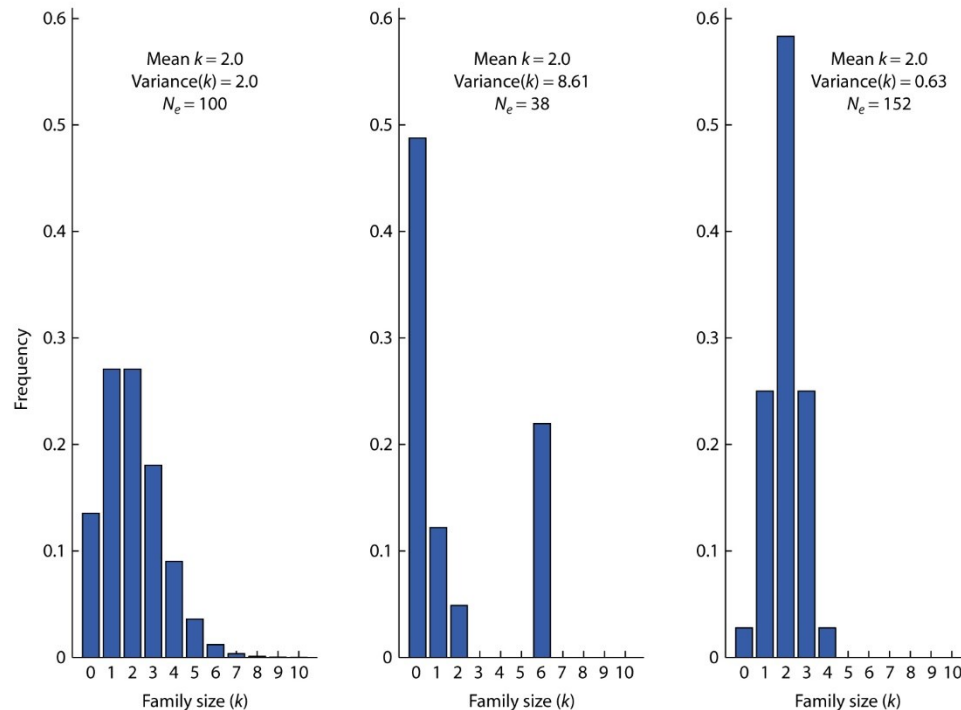
$$N_e = \frac{4N_m N_f}{N_m + N_f}$$



# Faktory ovlivňující efektivní velikost populace

## Rozptyl v počtu potomků mezi jedinci (selekce)

- $N_e$  se rovná skutečnému počtu jedinců pokud počty potomků jednotlivých jedinců mají Poisson rozložení. Tzn. rozptyl v počtu potomků je stejný jako průměrný počet potomků.
- Pokud je rozptyl větší,  $N_e$  je nižší než skutečný počet jedinců v populaci.



# Odhady efektivní velikosti populace u různých druhů

**~ 10 000**



**~ 30 000**



**~ 2 000 000**



**~ 1 000 000**

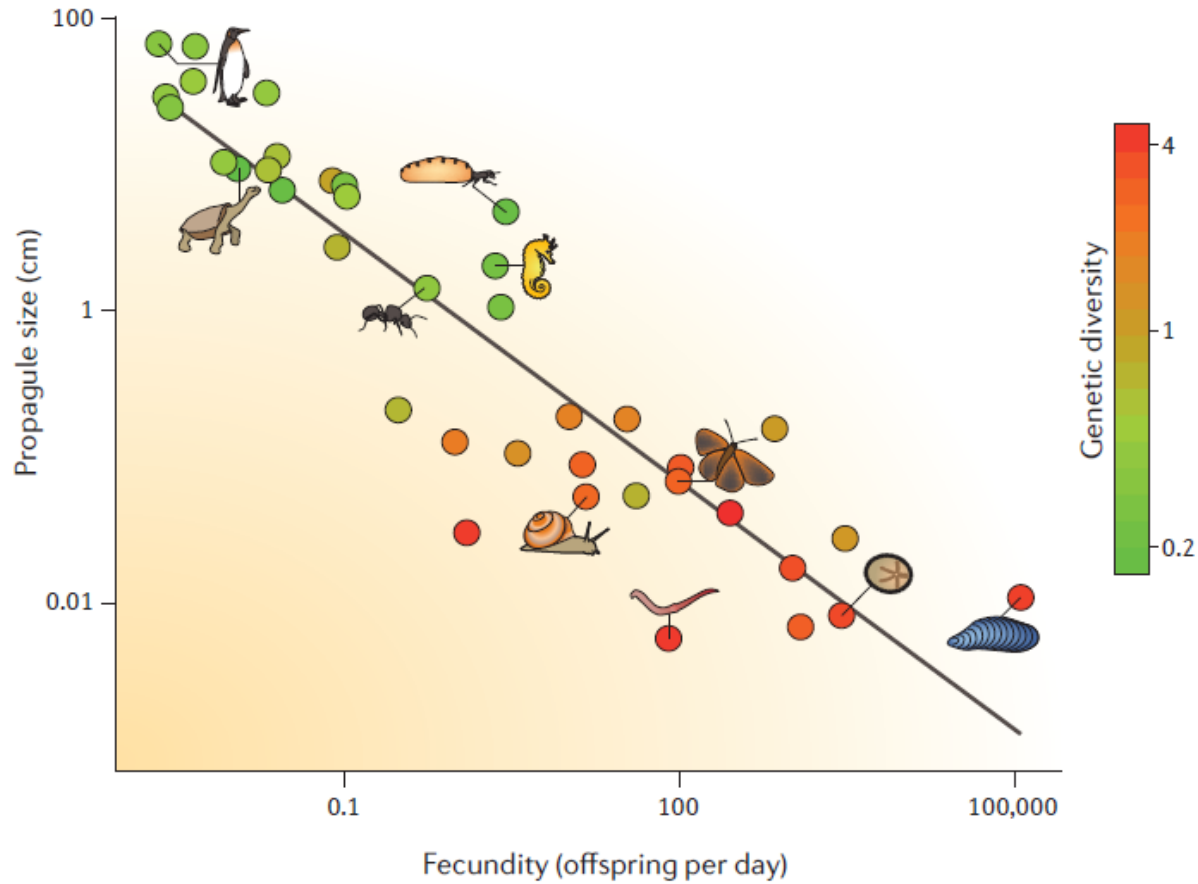


**~ 100 000**



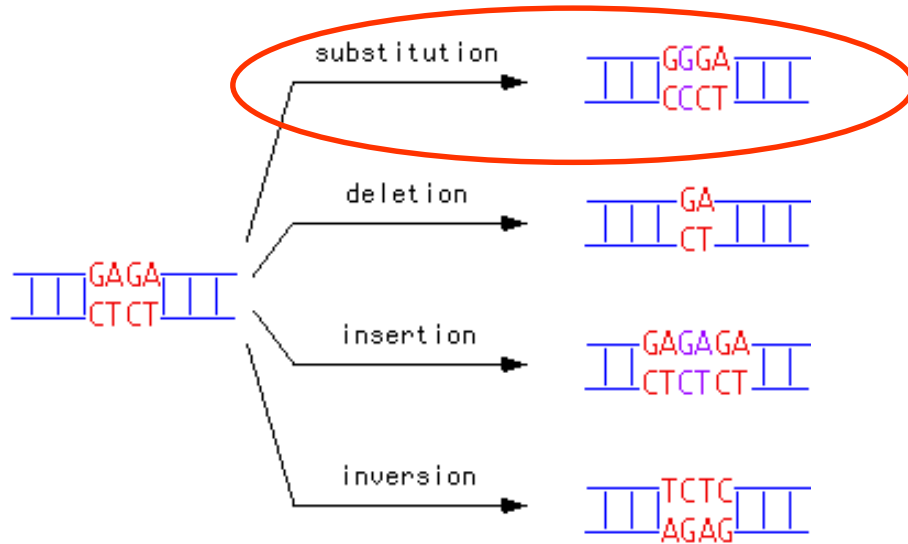


# Efektivní velikost populace závisí na životních strategiích organismů

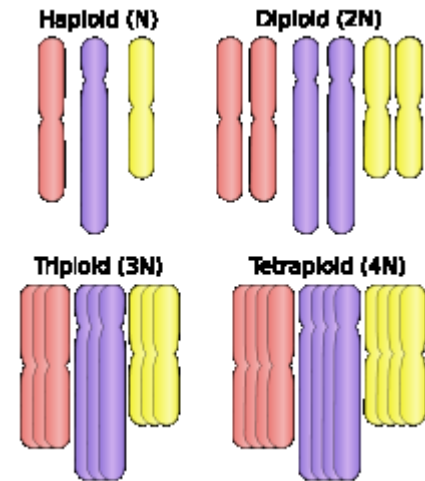


# Mutace jako zdroj genetické variability

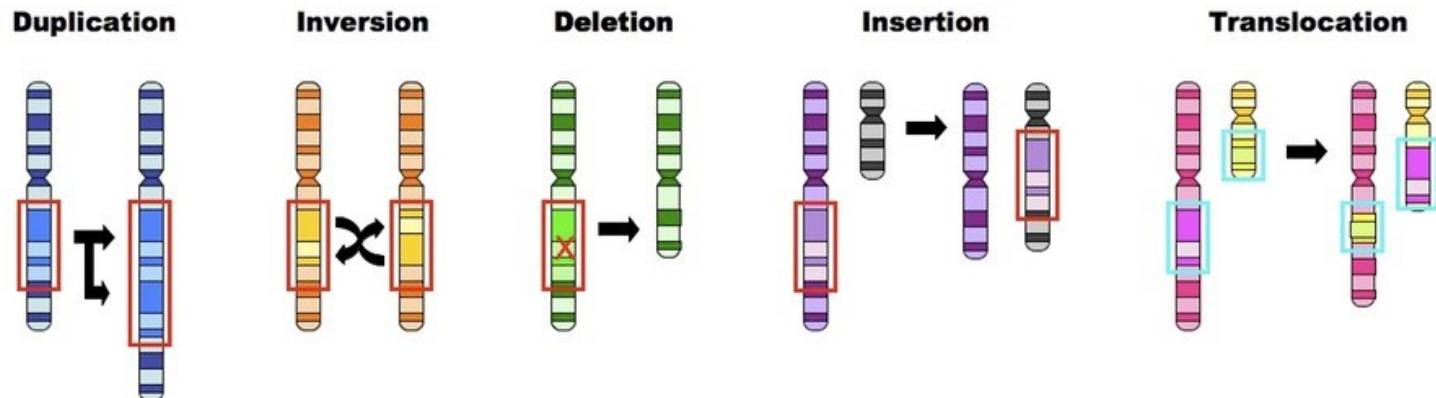
## Bodové/genové mutace



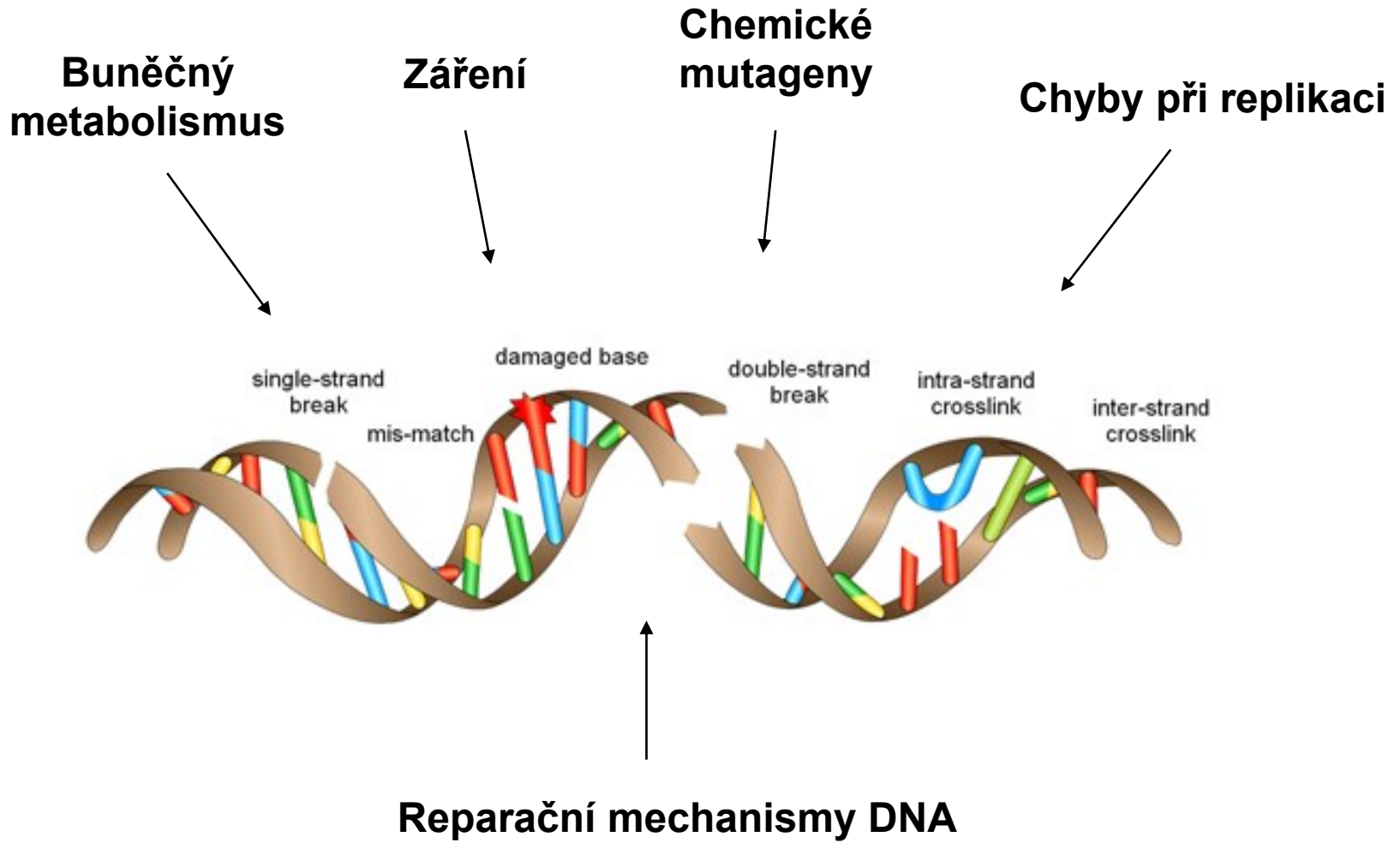
## Genomové mutace



## Chromosomové mutace

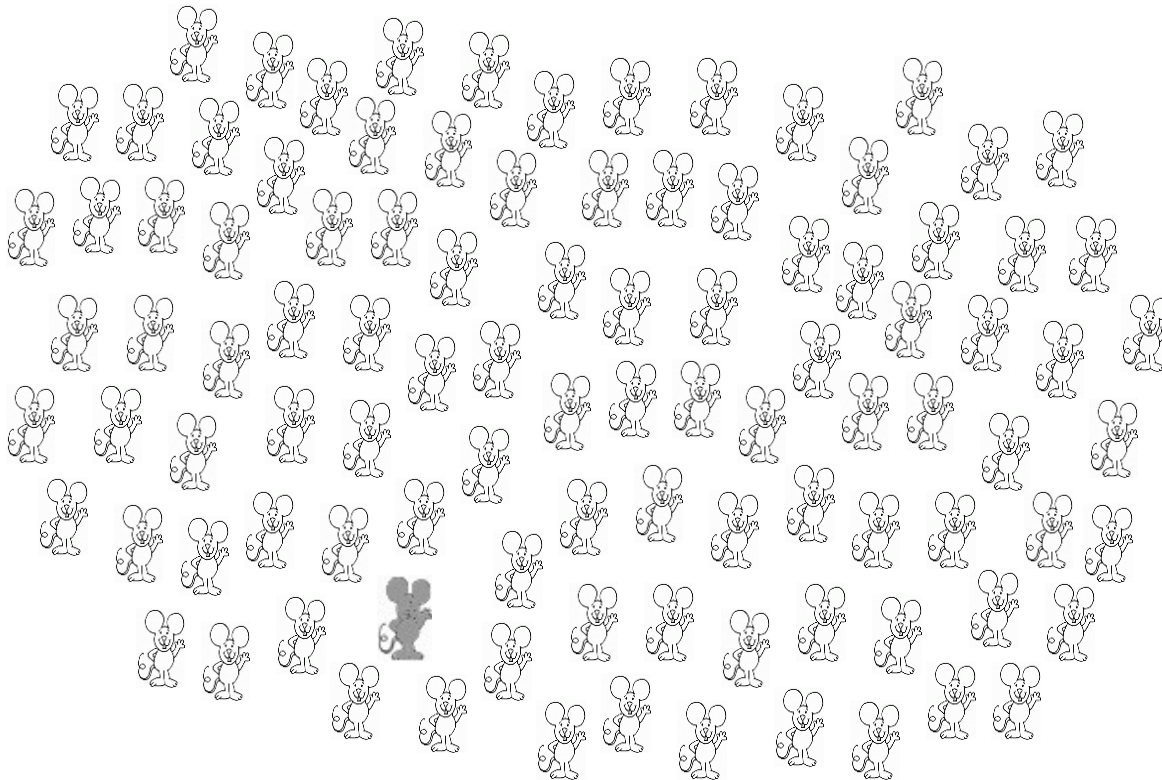


# Čím je daná mutační rychlost



# Mutační rychlost ( $\mu$ )

- frekvence nových mutací na generaci
- $\mu = 0.01$  (1 mutace na 100 jedinců na 1 generaci)

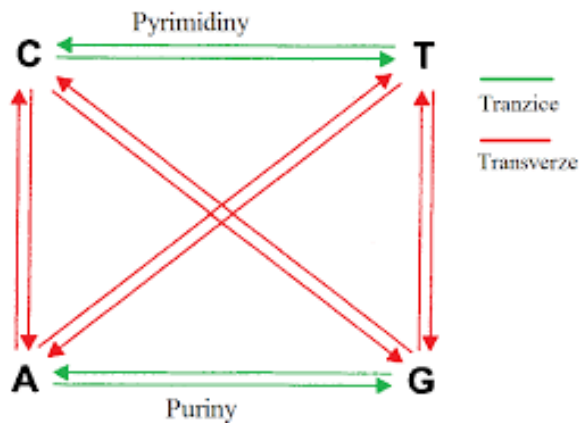
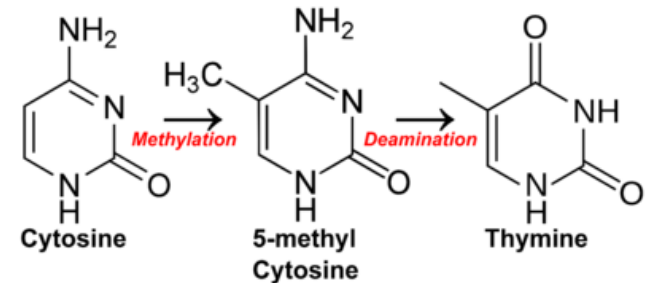


- Mutační rychlost (pro nukleotidové substituce) u lidí je cca  $1 \cdot 10^{-8}$ .  
(necelých 100 nových nukleotidových substitucí v celém genomu na jedince)

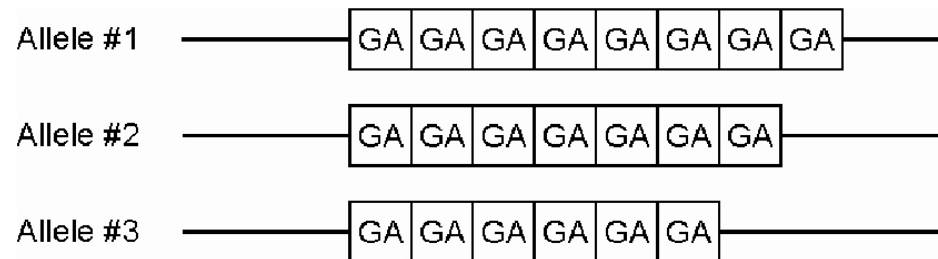
# Mutační rychlost

Je odlišná pro různé typy mutací

- rozsáhlejší mutace obecně méně časté
- mutační „hotspots“: CpG u živočichů, mikrosatelity.
- tranzice vznikají častěji než transverze



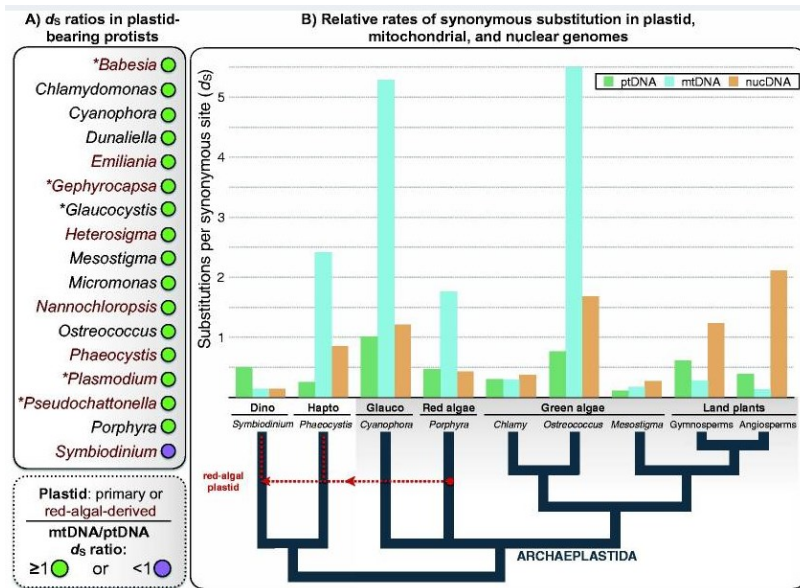
## mikrosatelit



# Mutační rychlost

## Je odlišná pro různé geny v rámci genomu

- Ovlivněna např. mírou transkripce dané sekvence, blízkostí k replikačnímu počátku, pozicí nukleozomu.
- Mutační rychlost u živočichů cca 10 x větší na mt DNA než v jaderné DNA



## Mutační rychlost mitochondriální a plastidové DNA u rostlin

GBE

### Mutation Rates in Plastid Genomes: They Are Lower than You Might Think

David Roy Smith\*

Department of Biology, University of Western Ontario, London, ON, Canada

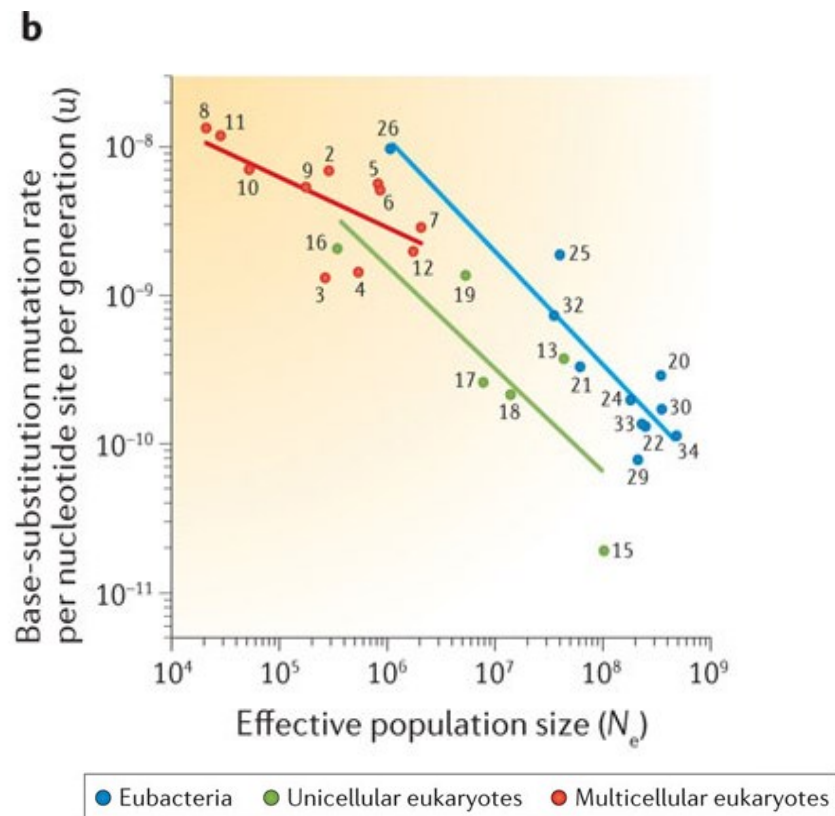
\*Corresponding author: E-mail: dsmit242@uwo.ca.

Accepted: April 9, 2015

# Mutační rychlost

## Liší se u různých organismů

- velikost genomu, efektivní velikost populace, velikost těla, intenzita metabolismu....



# Mutační rychlost v germinální linii somatických tkáních

Table 1. Mutation rates per nucleotide site ( $\times 10^{-9}$ ) in different tissues<sup>a</sup>

Species	Tissue	Cell divisions per generation <sup>a</sup>	Mutation rates <sup>b</sup>	
			Per generation	Per cell division
<i>Homo sapiens</i>	Germline	216	12.85	0.06
	Retina	55	54.45	0.99
	Intestinal epithelium	600	162.00	0.27
	Fibroblast (culture)			1.34
	Lymphocytes (culture)			1.47
<i>Mus musculus</i>	Male germline	39	38.00	0.97
	Brain		76.94	
	Colon		83.35	
	Epidermis		90.38	
	Intestine		117.69	
	Liver		237.88	
	Lung		166.83	
	Spleen		130.00	
<i>Rattus norvegicus</i>	Colon		178.38	
	Kidney		167.45	
	Liver		179.92	
	Lung		223.22	
	Mammary gland		57.70	
	Prostate		448.90	
	Spleen		101.62	
<i>Drosophila melanogaster</i>	Germline	36	4.65	0.13
	Whole body		380.92	
<i>Caenorhabditis elegans</i>	Germline	9	5.60	0.62
<i>Arabidopsis thaliana</i>	Germline	40	6.50	0.16
<i>Saccharomyces cerevisiae</i>		1	0.33	0.33
<i>Escherichia coli</i>		1	0.26	0.26

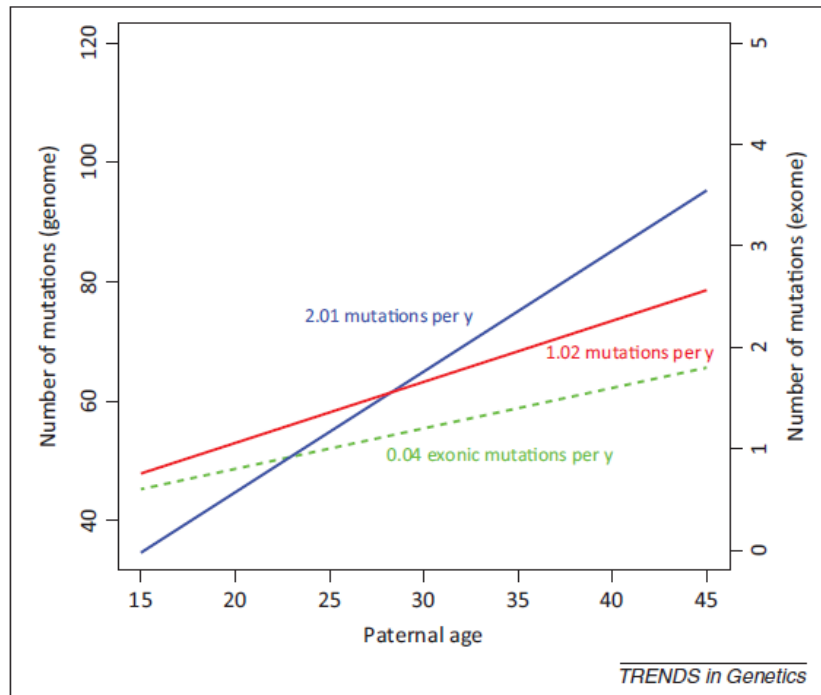
<sup>a</sup>References to data on numbers of germline cell divisions: human [Crow 2000]; *D. melanogaster* and mouse [57]; *C. elegans* [58]; and *A. thaliana* [59]. Numbers of cell divisions are unknown for the mouse and rat rates.

<sup>b</sup>Mammalian tissue-specific rates are given only for tissues in which at least two independent estimates have been acquired. All data on human mutation rates are taken from Lynch [36]. Data for somatic mutation rates in mouse and rat are derived from references contained within the supplementary material online. References to data on germline mutation rates are: *D. melanogaster* [5], *C. elegans* [4], *A. thaliana* [Ossowski *et al.*, 2009], *S. cerevisiae* [3], and *E. coli* [24].



# Většinu mutací dědíme od otců

- Většina nukleotidových substitucí vzniká v samčí germinální linii (více replikací DNA) (**male-driven evolution**).
- Počet nově vzniklých nukleotidových substitucí vzrůstá s věkem otce (paternal age effect).
- Rozdíl v mutační rychlosti mezi autosomy a pohlavními chromosomy.

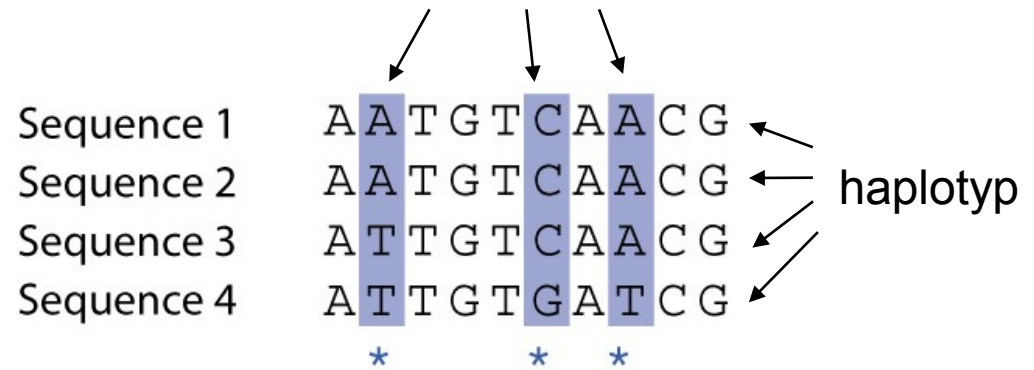


# Genetická diverzita - genetický polymorfismus ( $\theta$ )

Pro neutrální mutace platí:

$$\theta = 4N_e\mu$$

Polymorfní (segregující) místo



- Míra genetického polymorfismu v populaci závisí přímoúměrně na mutační rychlosti a velikosti populace.

# Empirické odhady genetické diverzity v populaci

## Nukleotidová diverzita (průměrná heterozygotita) ( $\pi$ )

Sequence 1	A	A	T	G	T	C	A	A	C	G
Sequence 2	A	A	T	G	T	C	A	A	C	G
Sequence 3	A	T	T	G	T	C	A	A	C	G
Sequence 4	A	T	T	G	T	G	A	T	C	G
Site number		*				*		*		
	1	2	3	4	5	6	7	8	9	10

Nucleotide diversity ( $\pi$ ):

1 A A T G T C A A C G  
2 A A T G T C A A C G  $d_{12} = 0$

1 A A T G T C A A C G  $d_{13} = 1$   
3 A T T G T C A A C G

2 A A T G T C A A C G  $d_{23} = 1$   
3 A T T G T C A A C G

1 A A T G T C A A C G  $d_{14} = 3$   
4 A T T G T G A T C G

2 A A T G T C A A C G  $d_{24} = 3$   
4 A T T G T G A T C G

3 A T T G T C A A C G  $d_{34} = 2$   
4 A T T G T G A T C G

$$\sum d_{ij} = 0 + 1 + 3 + 1 + 3 + 2 = 10$$

Number of pairs of sequences compared =  $[n(n - 1)]/2 = [4(3)]/2 = 6$

$\hat{\pi} = 10$  differences/6 pairs = 1.67 average pairwise differences

$\hat{\pi} = 1.67$  avg. differences/10 sites = 0.167 pairwise differences per site

## Proporce polymorfních míst ( $\theta_w$ )

Sequence 1	A	A	T	G	T	C	A	A	C	G
Sequence 2	A	A	T	G	T	C	A	A	C	G
Sequence 3	A	T	T	G	T	C	A	A	C	G
Sequence 4	A	T	T	G	T	G	A	T	C	G
		*				*		*		
Site number	1	2	3	4	5	6	7	8	9	10

Segregating sites ( $S$  and  $p_S$ ):

Sites 2, 6, and 8 have variable base pairs among the four sequences (columns marked with \*). These are segregating sites. Therefore, for these sequences  $S = 3$  segregating sites and  $p_S = 3/10 = 0.3$  segregating sites per nucleotide site examined.

$$\theta = S / n / H_{k-1}$$

$$H_{k-1} = 1 + \frac{1}{2} + \frac{1}{3}$$

$$\theta = 3 / 10 / 1,83 = 0,164$$

$S$  ... počet segregujících míst

$n$  ... počet nukleotidů v sekvenci

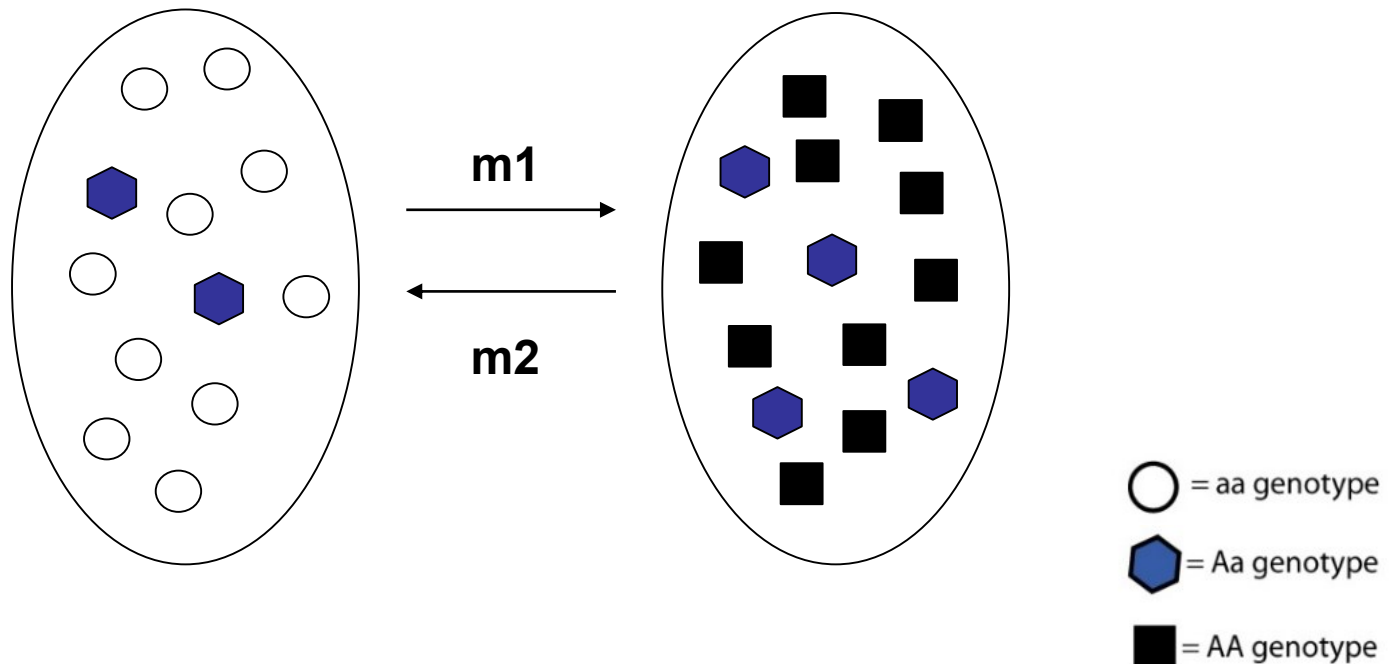
$k$  ... počet sekvencí

$H_{k-1}$  ... harmonické číslo

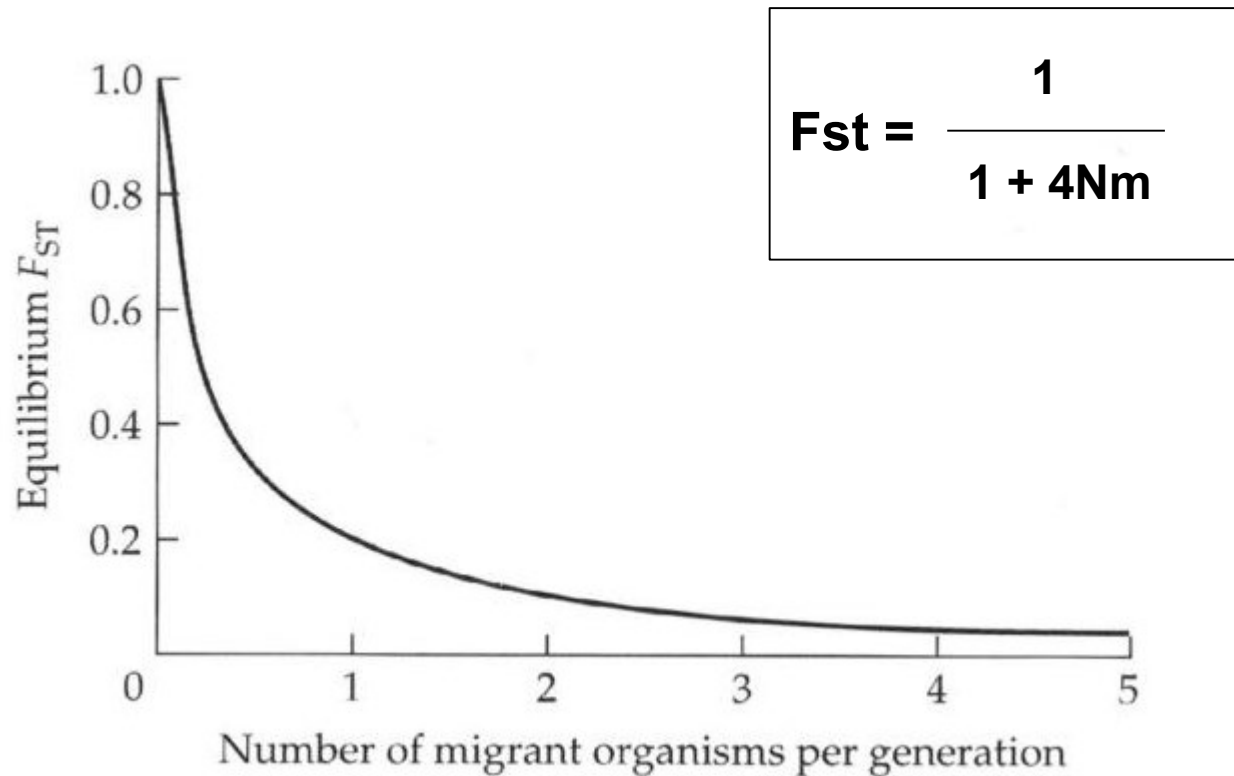
Za předpokladu neutrality by měly být oba odhady genetické diverzity stejné.

# Migrace (m)

- **Migrační rychlost (m):** pravděpodobnost, že jedinec z jedné populace nahradí jedince druhé populace (vztaženo na jedince a generaci).
- **Populační migrační rychlost (Nm):** počet migrantů za generaci



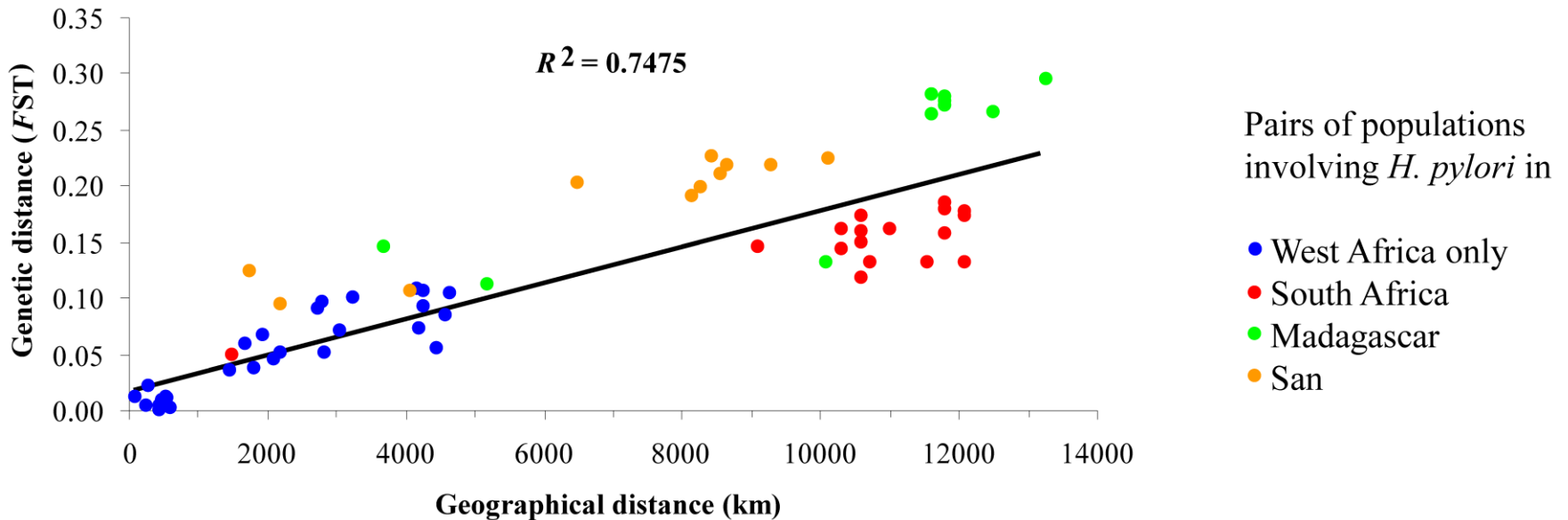
## Vztah mezi $F_{ST}$ a populační migrační rychlostí ( $Nm$ )



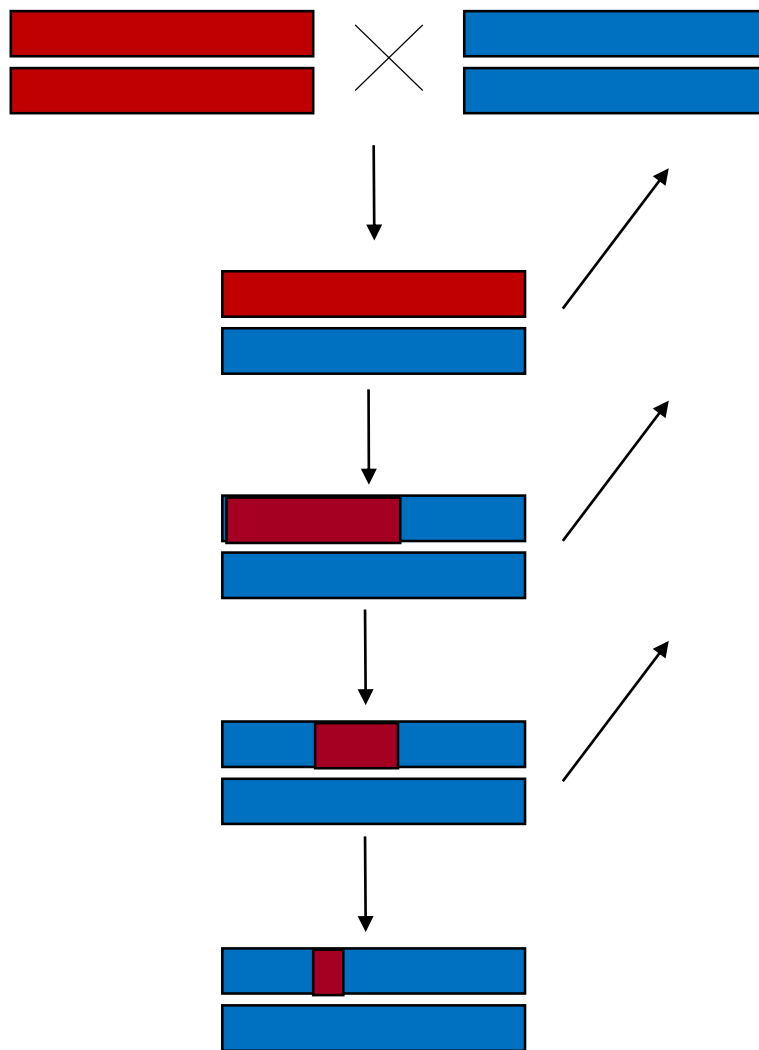
- Pokud  $Nm > 1$  je efekt migrace silnější než efekt genetického driftu, nedojde ke genetickému rozrůznění populací.

## Isolation by distance (izolace vzdáleností)

Míra genetické diferenciace ( $F_{ST}$ ) vzrůstá s geografickou vzdáleností.  
Vzniká když míra migrace je vyšší mezi geograficky bližšími populacemi.



# Genová introgrese (genový tok)



Přenos jednotlivých genů z  
jedné populace do druhé  
pomocí zpětného křížení.



# Adaptivní introgrese



Myš domácí (*Mus musculus*) získala alely k rezistenci vůči warfarinu z myši středozevní (*Mus spretus*).

## Adaptivní introgrese: Co jsme zdědili od neandrtálců?

- Neafrická lidská populace 2-4% genomu původem od neandrtálců.
- V Tibetské populaci mutace v genu EPAS1 usnadňuje život ve vysokých nadmořských výškách. Původem z neandrtálců.
- U inuitů fixované alely genů WARS2 a TBX15 původem z Denisovanů. Hrají roli v adaptaci na chladné podnebí (vytváření adipocytů a tepla).
- Adaptivní introgrese se také týká genů ovlivňujících kvalitu a zbarvení kůže a srsti a dále genů imunitního systému.
- Některé alely získané od neandrtálců způsobují choroby (např. obezitu, deprese).

