

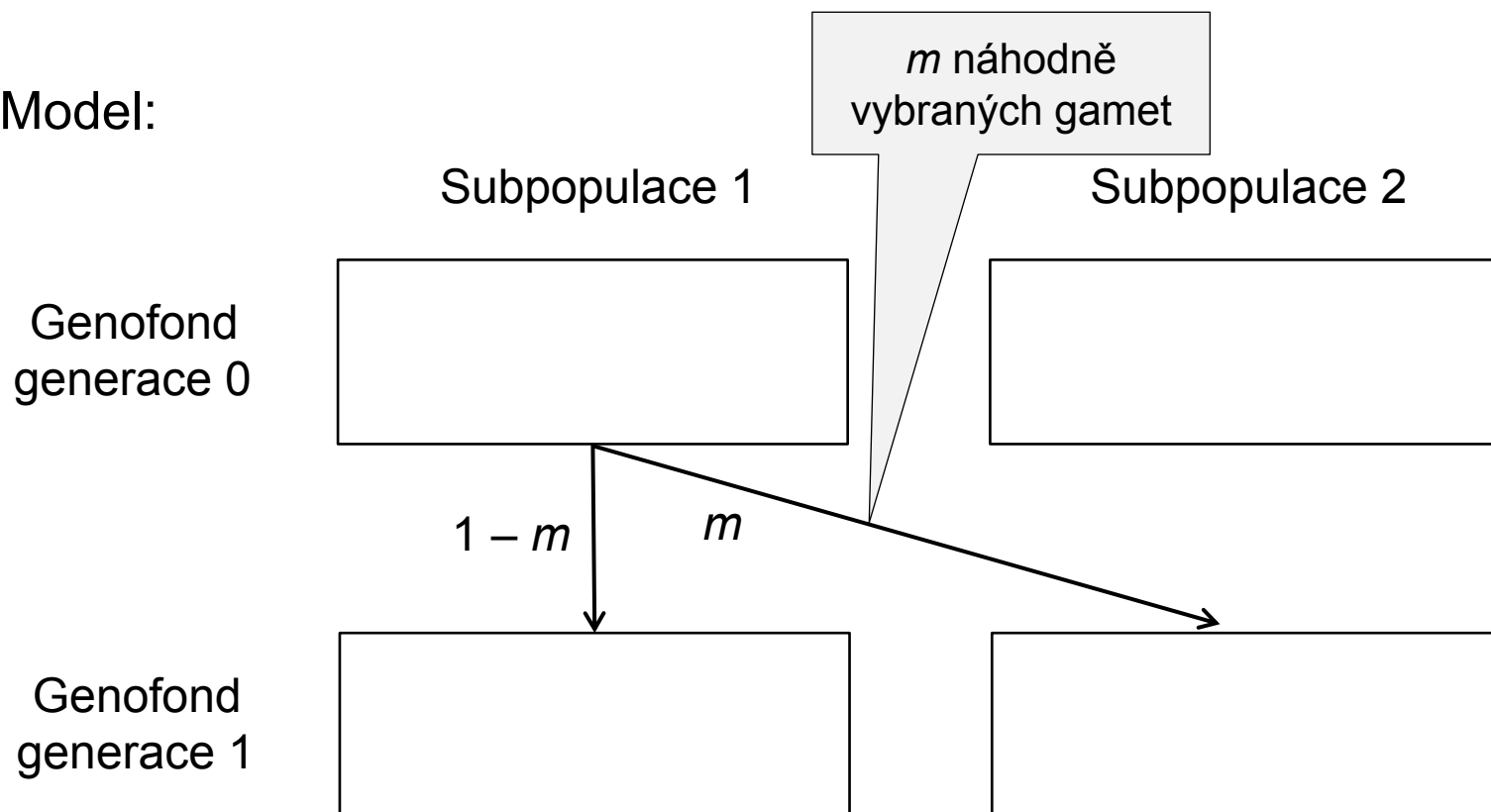
**TOK  
GENŮ**

HW model: jedna zcela izolovaná populace

× populace často rozděleny do subpopulací

genetická výměna mezi lokálními populacemi = tok genů (*gene flow*)

Model:



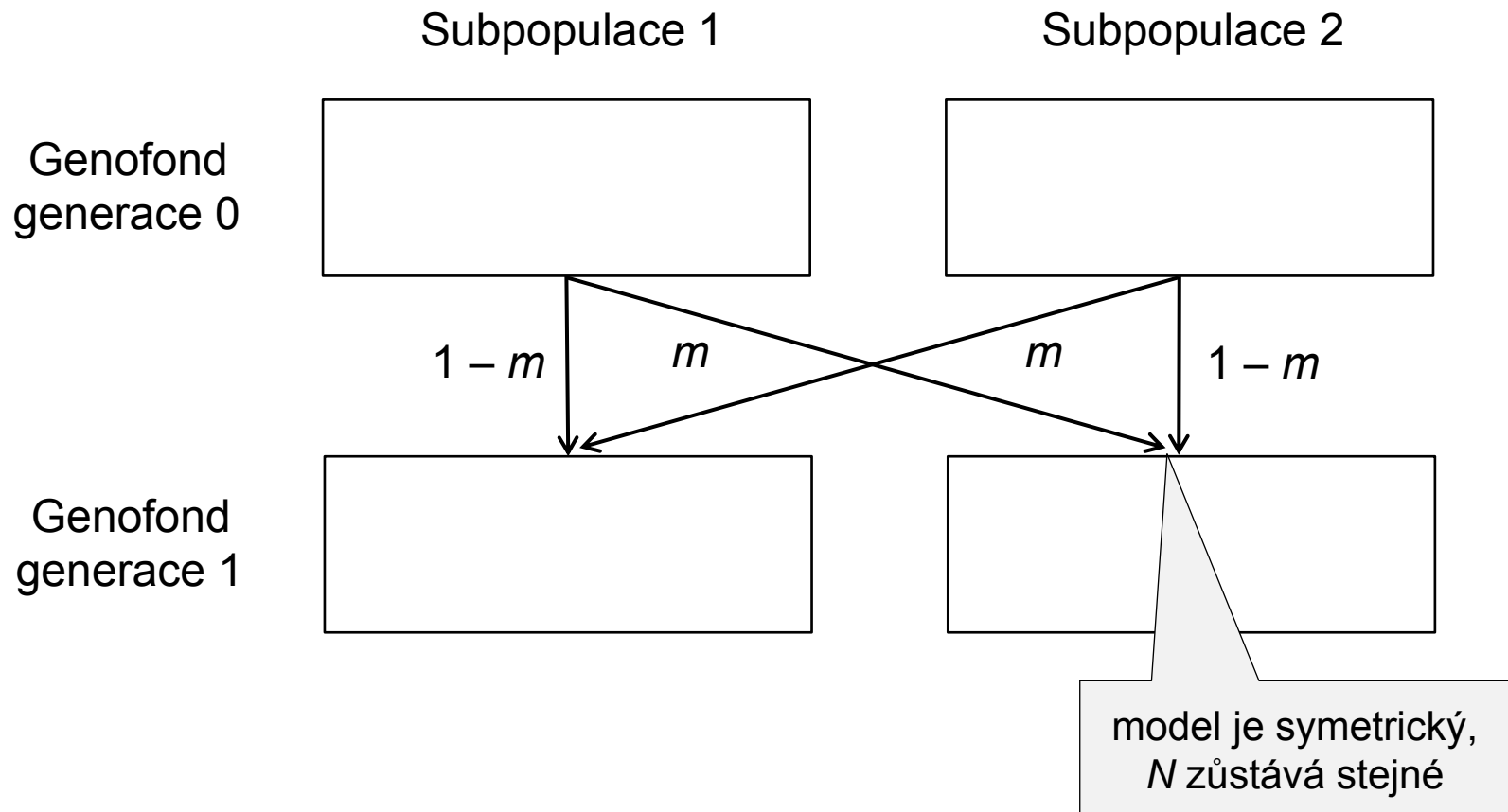
$m$  = míra genového toku (*migration rate*)

HW model: jedna zcela izolovaná populace

× populace často rozděleny do subpopulací

genetická výměna mezi lokálními populacemi = tok genů (*gene flow*)

Model:

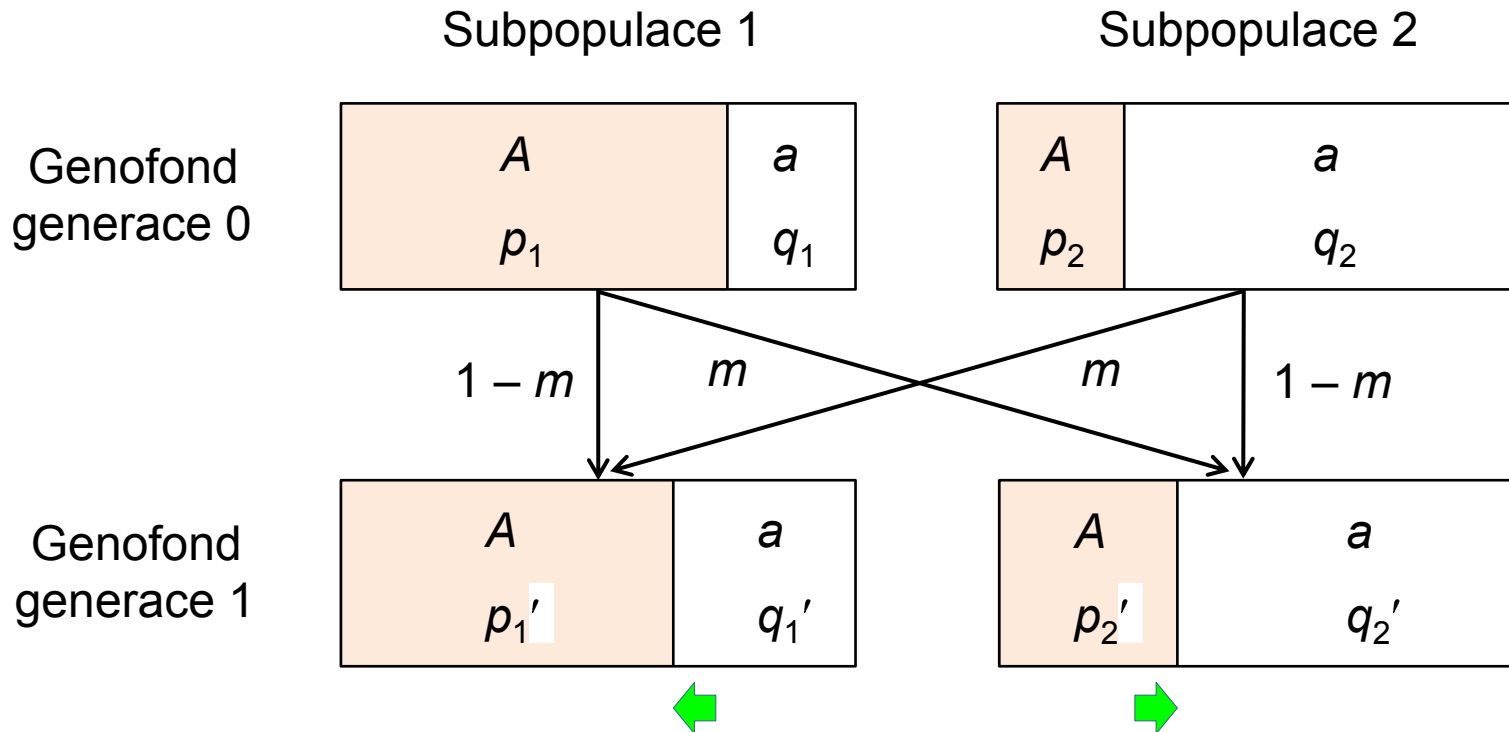


HW model: jedna zcela izolovaná populace

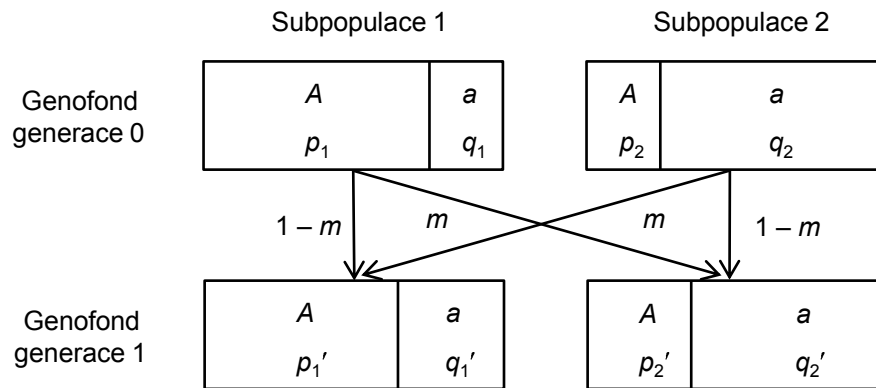
× populace často rozděleny do subpopulací

genetická výměna mezi lokálními populacemi = **tok genů** (*gene flow*)

Model:



$m$  = míra genového toku (*migration rate*)



Frekvence alely A:

počáteční generace:  $p_1$  (subpop. 1) a  $p_2$  (subpop. 2)

následující generace:  $p_1' = (1 - m)p_1 + mp_2$

**Evoluce?**

$$\Delta p_1 = p_1' - p_1 = (1 - m)p_1 + mp_2 - p_1 = -m(p_1 - p_2)$$

$$\Delta p_2 = -m(p_2 - p_1)$$

**EVOLUCE!**

lokální populace  
nejsou zcela  
izolované

$m > 0$   
a současně  
 $p_1 \neq p_2$

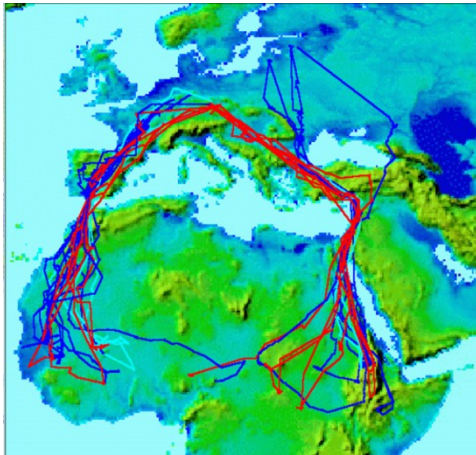
lokální populace  
jsou geneticky odlišné

# tok genů ≠ migrace!

$m$  = rozsah výměny  
gamet, ne jedinců



tok genů × žádná migrace



migrace na velké  
vzdálenosti ×  
žádný tok genů



Tok genů je evolučním mechanismem, který působí nenáhodným, predikovatelným způsobem:

rozdíl ve frekvencích  $d_0 = p_1 - p_2$

$$p_1' = (1 - m)p_1 + mp_2 = p_1 - md_0$$

analogicky  $p_2' = p_2 + md_0$

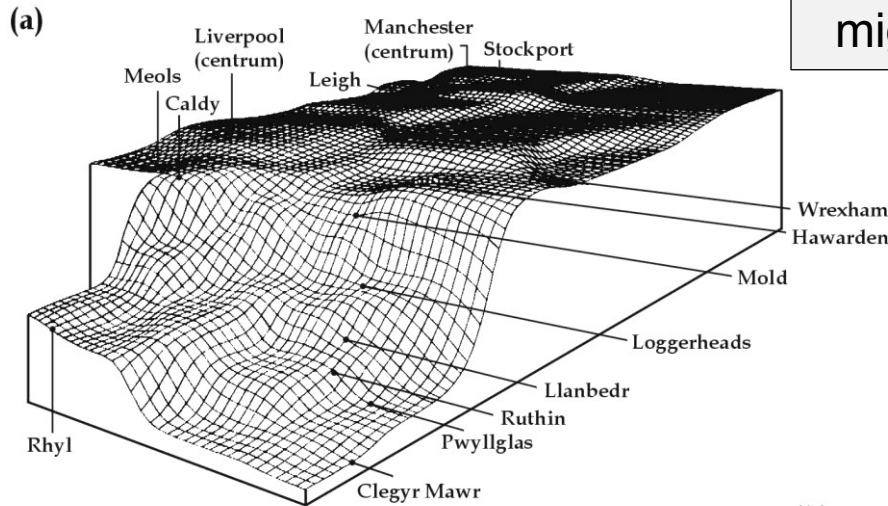
$$d_1 = p_1' - p_2' = p_1 - md_0 - p_2 - md_0 = d_0(1 - 2m)$$

$m = 0$  anebo  
 $d_0 = 0 \Rightarrow$   
žádná evoluce

po  $t$  generacích:  $d_t = d_0(1 - 2m)^t$ , tj. s tím, jak  $t \rightarrow \infty$   $d_t \rightarrow 0$

**Tok genů snižuje rozdíly ve frekvencích alel mezi lokálními populacemi**

# Př.: výskyt melanických forem můr v Anglii



vyšší  
migrace

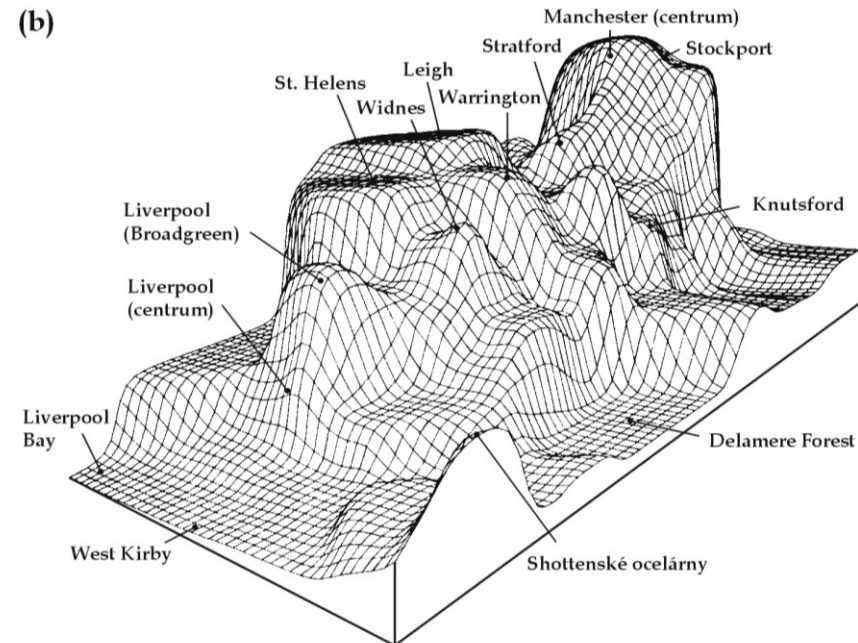


drsnokřídlec březový  
(*Biston betularia*)

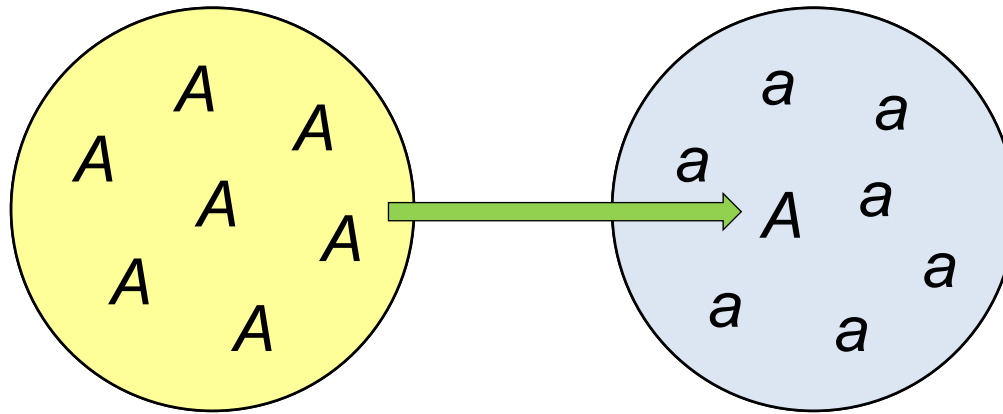
nižší  
migrace



zejkovce dvojzubý  
(*Odontoptera [Gonodontis] bidentata*)





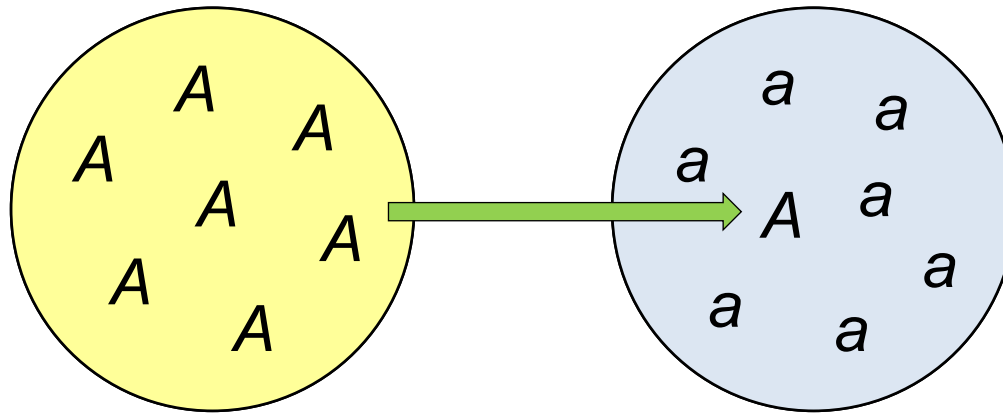


zanesení nové alely do lokální populace → tok genů ~ mutaci

ALE:

- 1) mutační rychlost ( $\mu$ ) velmi nízká × genový tok ( $m$ ) může být nízký i vysoký
- 2) tok genů se může týkat mnoha lokusů současně  
× mutace zasahuje jen 1 lokus

Mísením různých populací vzniká masivní **vazebná nerovnováha** rozprostřená po celém genomu (× mutace: LD jen mezi mutantním lokusem a jeho okolím, kde není rekombinace).



zanesení nové alely do lokální populace → tok genů ~ mutaci

ALE:

3) mutace většinou škodlivé (× migrací pozitivní i negativní alely)

4) mutace v době vzniku v nízké frekvenci ⇒ ztráta  
× migrace může vnést více kopií současně

⇒ tok genů může výrazně změnit genofond i během jediné generace

Tok genů snižuje variabilitu *mezi démy* × zvyšuje *uvnitř démů*

Tok genů a nenáhodné oplození:

**Inbreeding** a **asortativní páření** omezuje tok genů:

např. Tamilové v Indii: sňatky mezi bratranci a sestřenicemi ⇒ nízký tok genů s okolním obyvatelstvem

zavíječ kukuřičný (*Pyrausta nubilalis*):  
asortativní výběr stejné feromonové rasy



**Disasortativní páření** podporuje tok genů: reprodukční výhoda dispergujících jedinců

např. *D. melanogaster*: disasortativní výběr ⇒ jedna kosmopolitní populace

## Tok genů mezi bělochy a černochoy a vliv genetického i negenetického asortativního výběru:

S a J Amerika: největší příliv otroků ze Z a střední Afriky 1700-1808

USA: sociální klasifikace míšenců jako černochoů, přestože genotypově i fenotypově intermediární

asortativní páření podle barvy kůže (přičemž míšenec = „černoch“)  
+ větší počet bělochů ⇒ **asymetrický tok genů** běloši → černoši



## Tok genů mezi bělochy a černochoy a vliv genetického i negenetického asortativního výběru:

S a J Amerika: největší příliv otroků ze Z a střední Afriky 1700-1808

USA: sociální klasifikace míšenců jako černochoů, přestože genotypově i fenotypově intermediární

asortativní páření podle barvy kůže (příčemž míšenec = „černoch“)  
+ větší počet bělochů ⇒ **asymetrický tok genů** běloši → černoši

SZ Brazílie: více kategorií hybridů, „běloši“ spíše v kontextu kultury ⇒

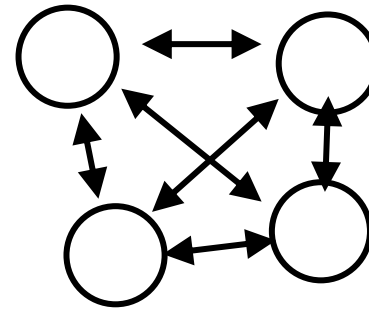
⇒ tok genů méně asymetrický a vyšší než v USA

→ Přes podobné počáteční podmínky dnes výrazné rozdíly

→ Tok genů je určován nejen migrací, ale i systémem páření ovlivněným jak genetickými, tak negenetickými faktory

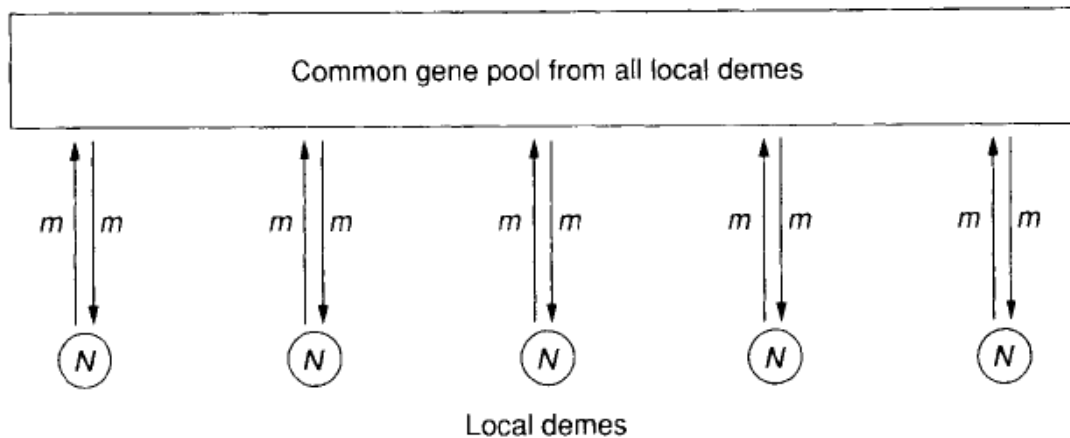
# Modely toku genů:

## 1. ostrovní (*island model*):



standardizovaný  
rozptyl frekvencí  
alel přes děmy

$$F_{ST} = \frac{1}{1 + 4Nm}$$

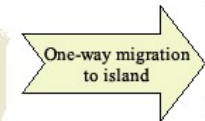


Continent-island model of Sewall Wright

### CONTINENT

A = Avirulence allele  $f(A) = p$   
a = virulence allele (mutant)

$Q = f(a)$  = frequency of  
virulence mutation on  
continent.



### ISLAND

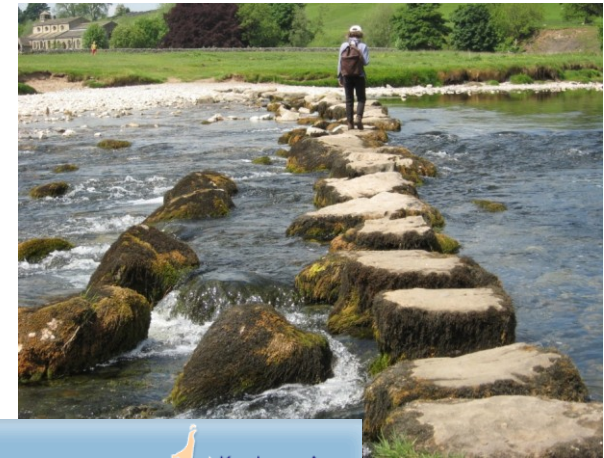
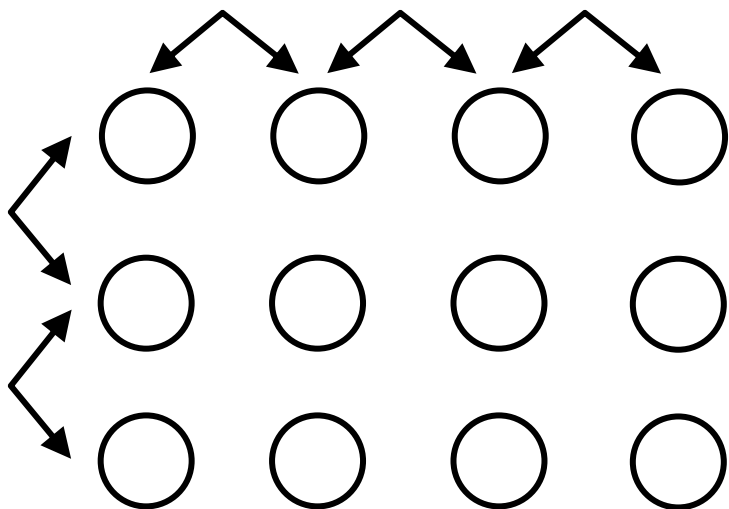
m immigrants  
1-m natives

asymetrický ostrovní model 2 populací = *continent-island*  
(= *mainland-island*) model

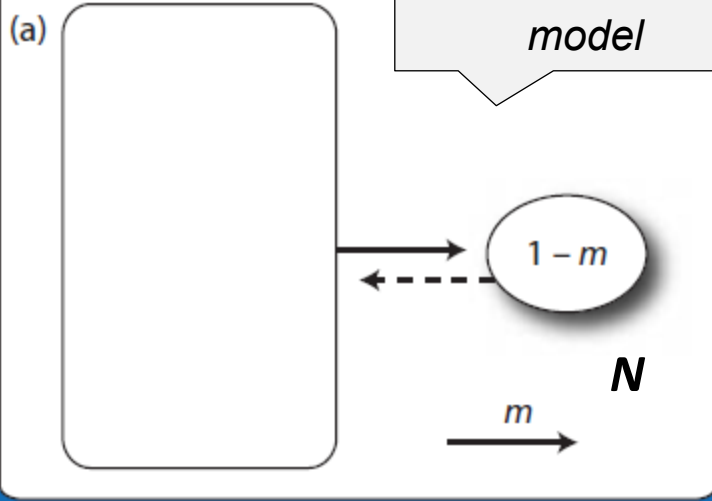
## 2. izolace vzdáleností (*isolation by distance*)

Př.: 2022 sňatků v údolí Ina (Japonsko, 1951):  
~ 50 % v rámci téže osady,  
> 2/3 ze stejné vesnice,  
< 1/3 ze vzdálenějších míst

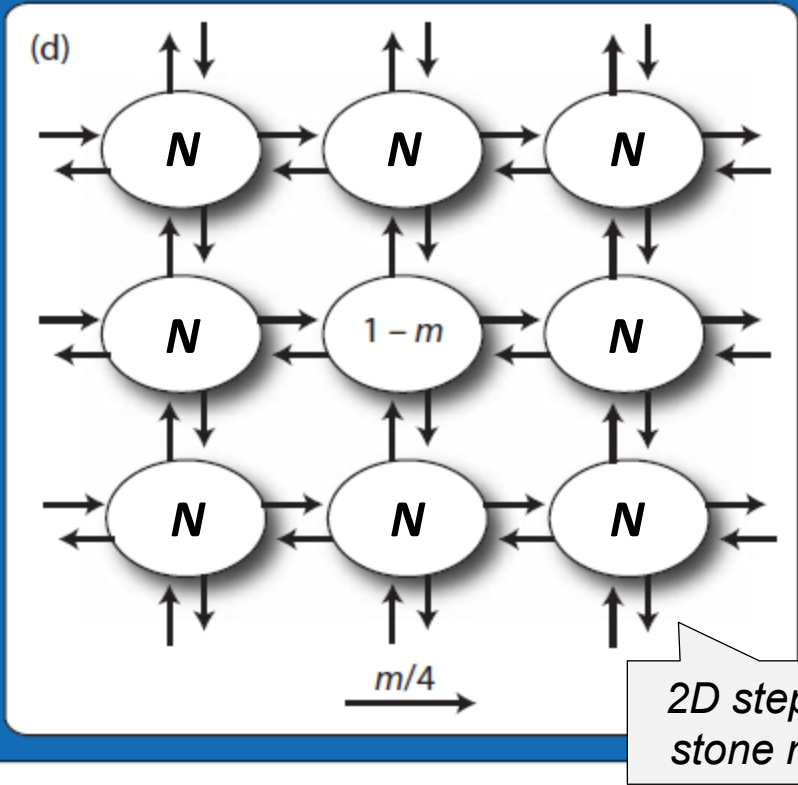
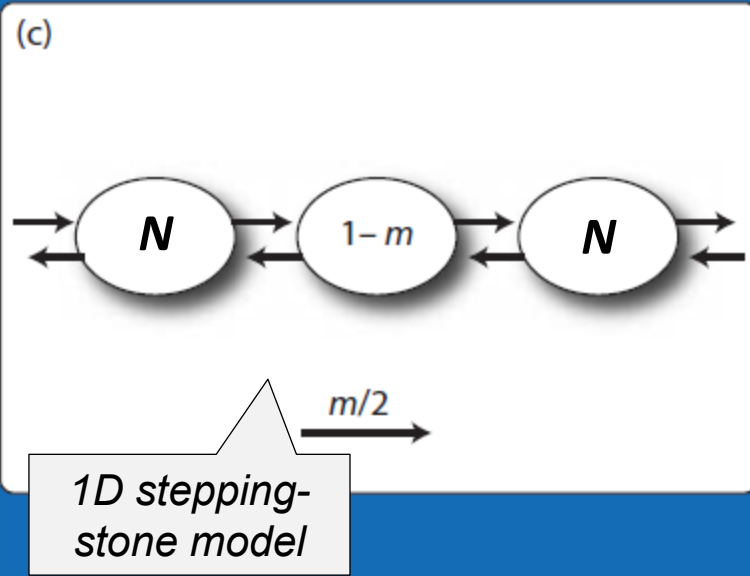
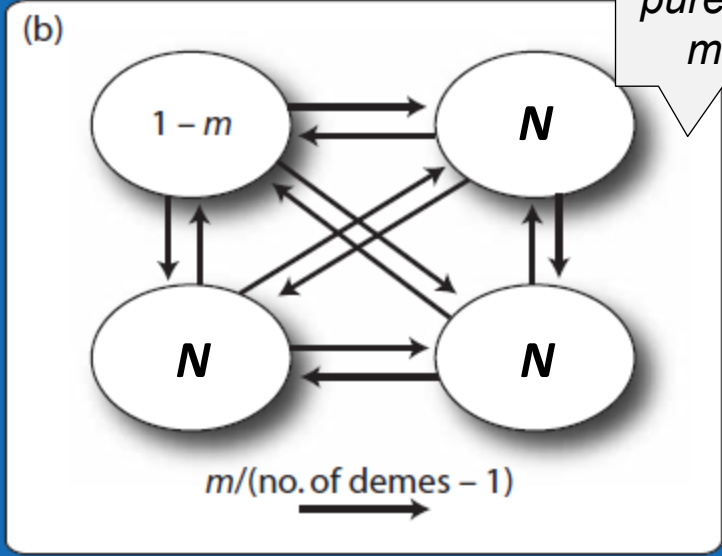
stepping stone: 1D, 2D



*continent-island model*



*pure island model*

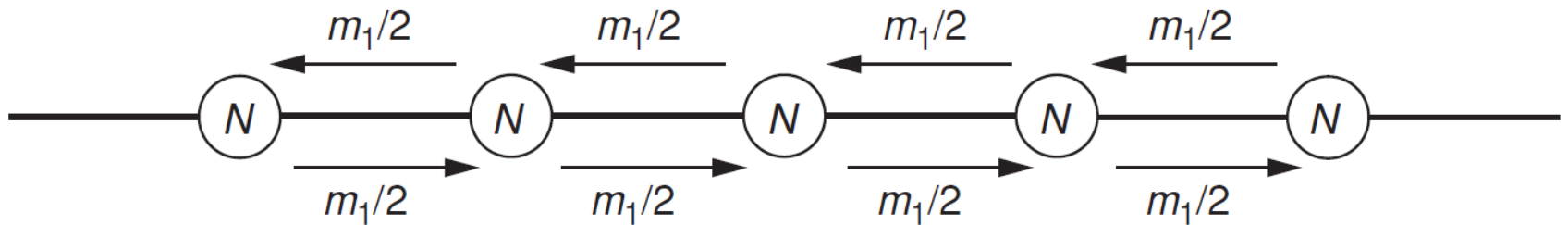




# 1D stepping stone model:

$m_1$  = frakce gamet dispergujících mezi sousedními démy

tok genů symetrický,  $m_1/2$  na jednu i druhou stranu

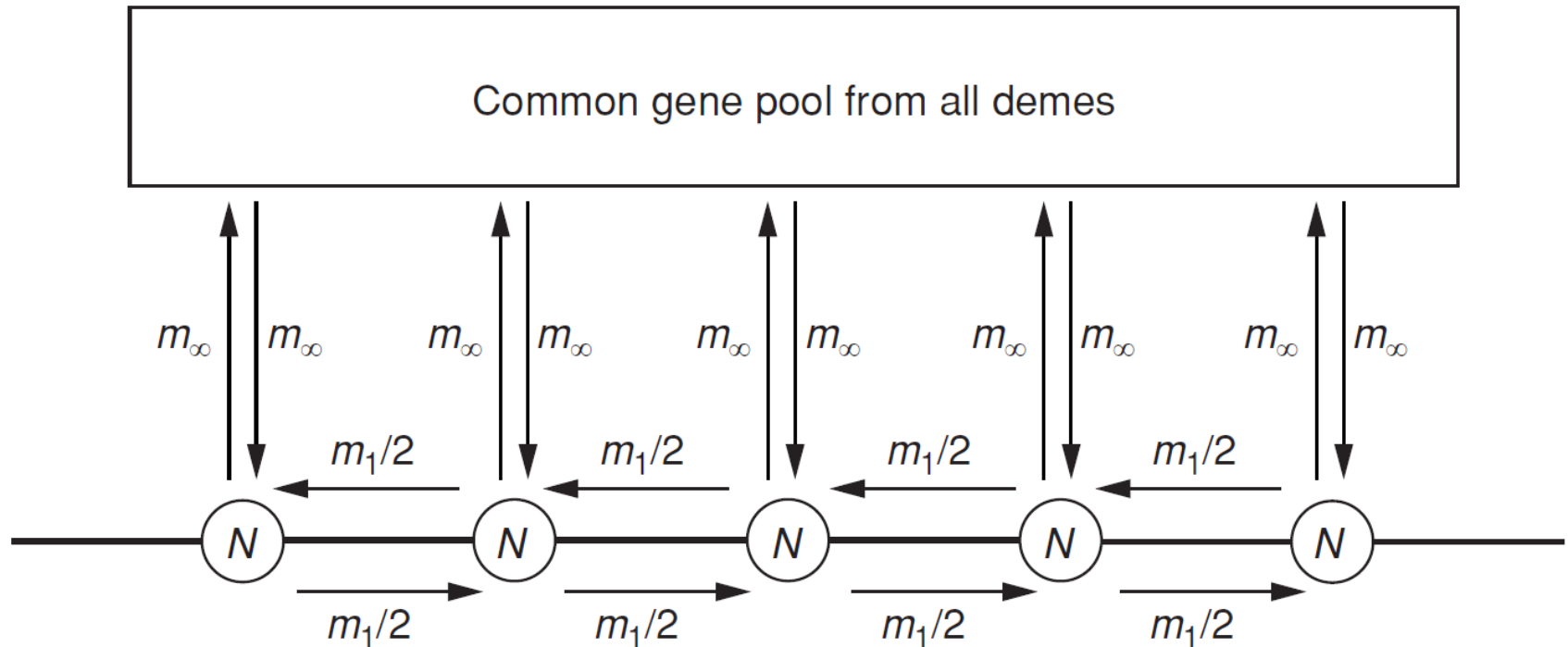


# 1D stepping stone model:

$m_1$  = frakce gamet dispergujících mezi sousedními démy

tok genů symetrický,  $m_1/2$  na jednu i druhou stranu

$m_\infty$  = frakce gamet dispergujících po celé populaci



většinou  $m_\infty \ll m_1 \Rightarrow F_{ST}$  můžeme aproximovat jako

$$F_{ST} \approx \frac{1}{1 + 4N\sqrt{2m_1m_\infty}}$$

rovnice je silně  
citlivá na migraci  
na velké  
vzdálenosti!

I když  $m_\infty \ll m_1$ , dopad toku genů dán součinem  $m_1m_\infty \Rightarrow$   
disperze na velké vzdálenosti má stále velký vliv na strukturovanost  
populace

např. jestliže  $N = 100$ ,  $m_1 = 0,1$  ...  $m_\infty = 0,01 \rightarrow F_{ST} = 0,053$   
...  $m_\infty = 0,001 \rightarrow F_{ST} = 0,276$

$\Rightarrow$  i když  $m_\infty = 10\text{--}100\times$  nižší než  $m_1$ , vliv na  $F_{ST}$  silný

důvodem větší rozdíl ve frekvencích mezi vzdálenými vs. sousedními  
démy

**Přesnější kvantifikace toku genů genetickými než  
přímými metodami (např. zpětným odchytém)!**

# izolace vzdáleností pro kontinuální populace:

*Linanthus parryae* z čeledi jirnicovitých (Polemoniaceae)  
z Mohavské pouště (Kalifornie) ... T. Dobzhansky, Sewall Wright

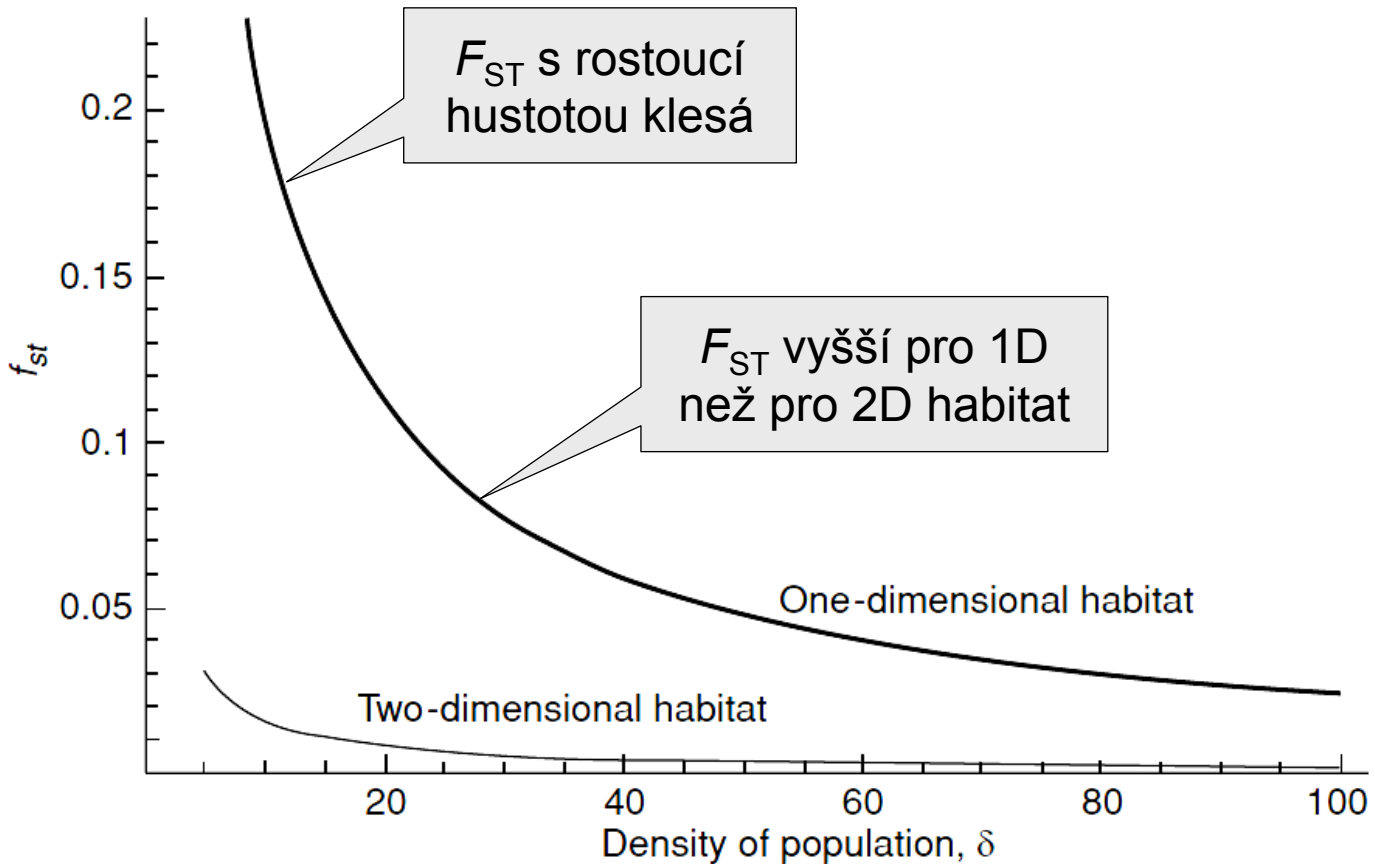
dém  $\rightarrow$  **čtvrť** (*neighborhood*)

$N \rightarrow \rho =$  **populační hustota** (*density*)

$\sqrt{m_1} \rightarrow \sigma =$  **disperze** (*dispersal*) = SD geografické  
vzdálenosti mezi místem narození rodičů a potomků



*L. parryae*



*L. parryae*

## izolace vzdáleností pro kontinuální populace:

velikost čtvrtě (*neighbourhood size*)  $\mathcal{N} \sim N_e$ :

$$1\text{D habitat: } \mathcal{N} = 2\rho\sigma\sqrt{\pi}$$

$$2\text{D habitat: } \mathcal{N} = 4\pi\rho\sigma^2$$

⇒ s rostoucí hustotou nebo disperzí roste *neighborhood size* a tedy i tok genů

S rostoucím počtem „kroků“ (stepping stone model) nebo geografickou vzdáleností (kontinuální model) by měla růst genetická diferenciace

→ (populačně)genetické distance ( $F_{ST}$ , Nei, atd.)

# Test izolace vzdáleností

**Mantelův test:** korelace mezi 2 symetrickými maticemi:  
matice geografických vzdáleností  
matice genetických vzdáleností,  $F_{ST}$ ,  $F_{ST}/(1 - F_{ST})$

## Mantel test

**X matrix**  
e.g. geographical  
distances

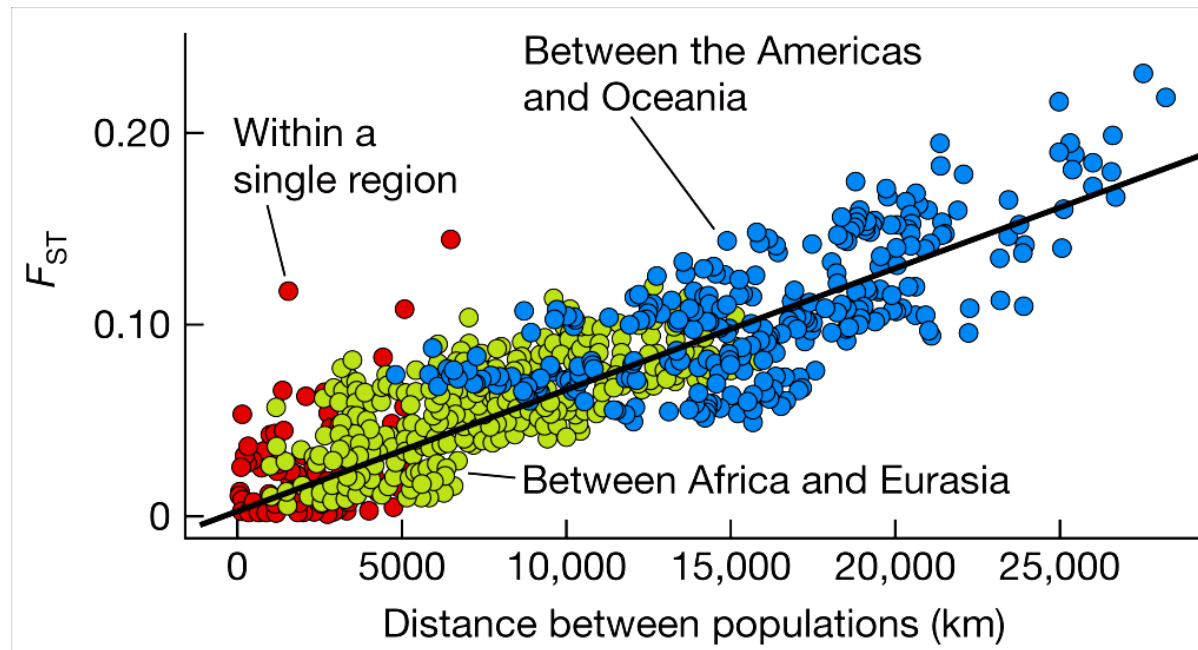
	A	B	C	D	E
A	0.0	1.2	2.6	1.8	3.2
B	1.2	0.0	3.1	0.5	2.7
C	2.6	3.1	0.0	1.1	4.2
D	1.8	0.5	1.1	0.0	3.4
E	3.2	2.7	4.2	3.4	0.0

**Y matrix**  
e.g. species  
dissimilarity

	A	B	C	D	E
A	0.00	0.29	0.56	0.45	0.49
B	0.01	0.00	0.48	0.06	0.12
C	0.21	0.17	0.00	0.27	0.59
D	0.07	0.04	0.16	0.00	0.02
E	0.45	0.34	0.78	0.21	0.00

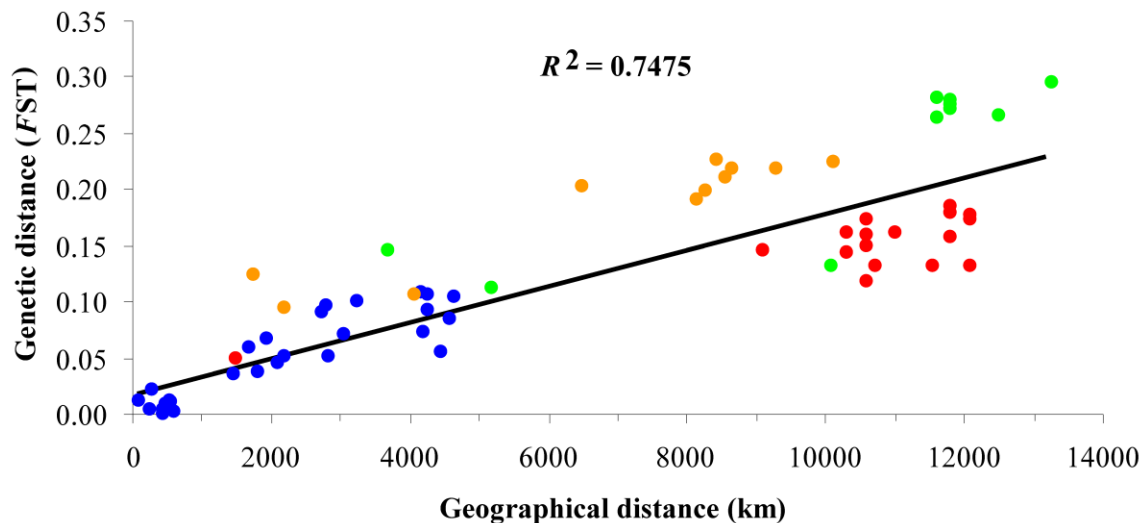
**Problem: are X and Y (in)dependent of each other?**

# Izolace vzdáleností (*isolation by distance*)



izolace vzdáleností u člověka ...

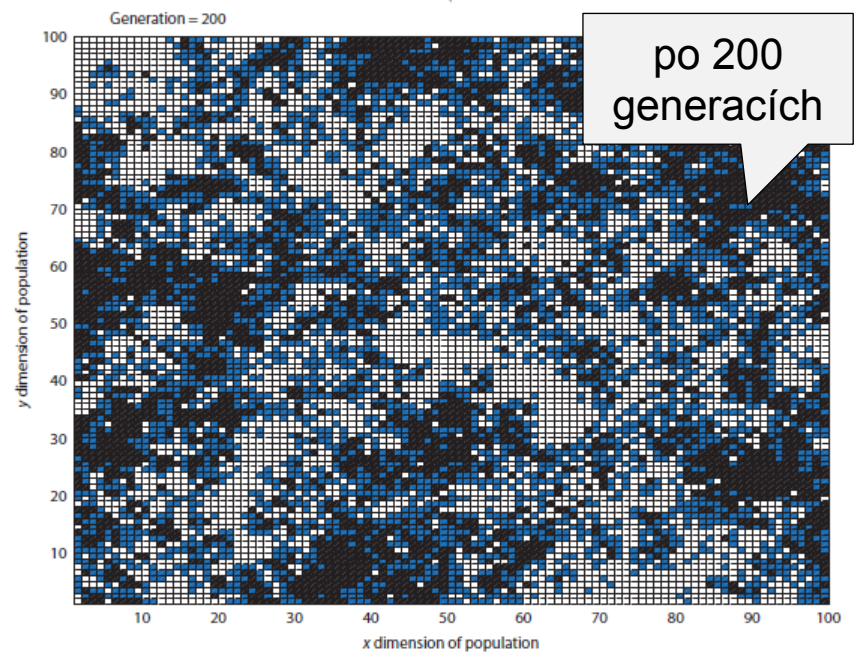
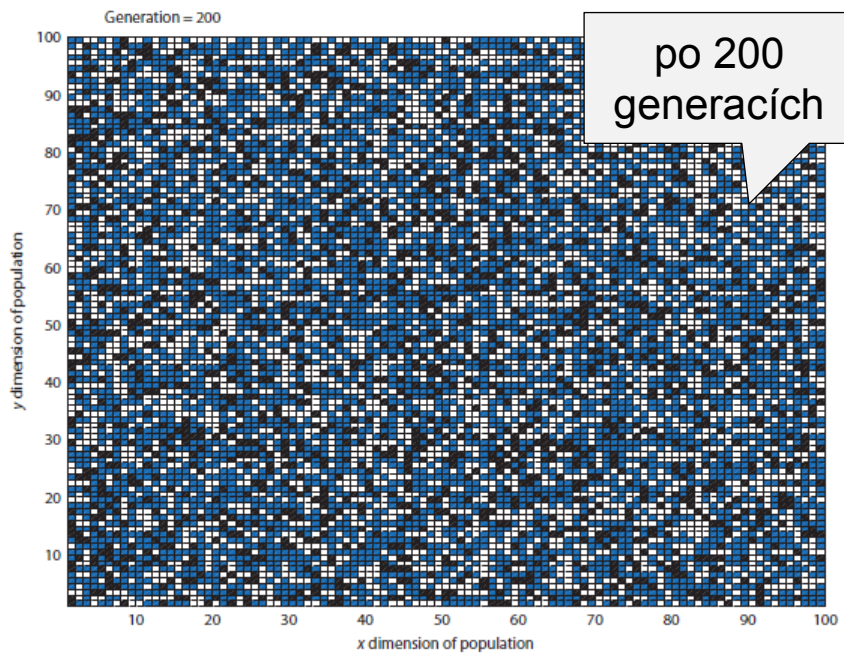
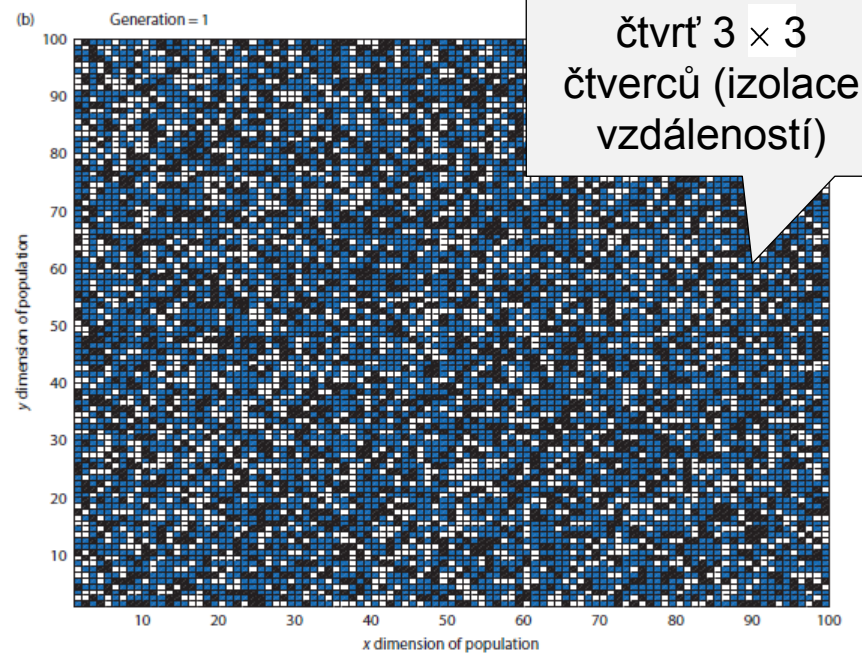
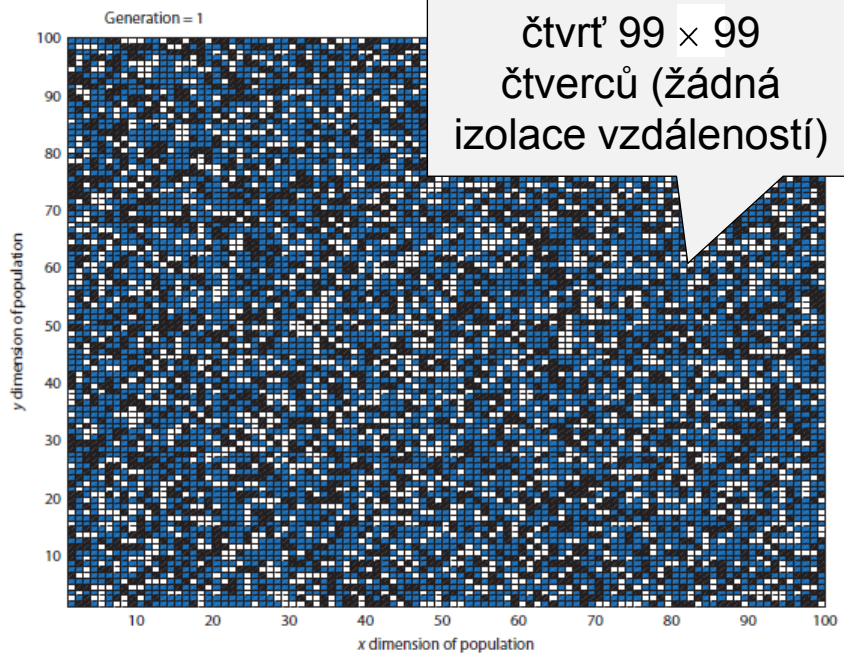
... a u *Helicobacter pylori*



Pairs of populations involving *H. pylori* in

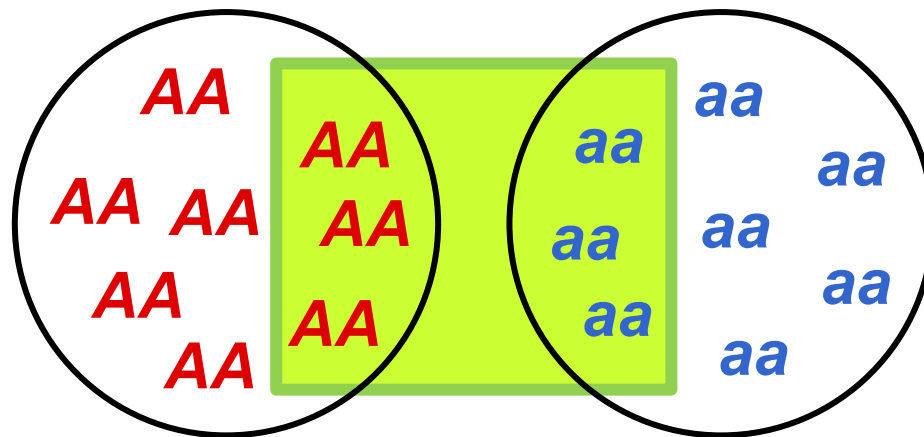
- West Africa only
- South Africa
- Madagascar
- San





# DRIFT V ROZDĚLENÉ POPULACI

Wahlundův efekt



absence  
heterozygotů

## Wahlundův efekt:

např. 2 stejně velké děmy,  $p_1 = 0,2$ ,  $p_2 = 0,7$ :

průměrné frekvence genotypů:

$$AA: (p_1^2 + p_2^2)/2 = (0,2^2 + 0,7^2)/2 = 0,265$$

$$Aa: (2p_1q_1 + 2p_2q_2)/2 = (2 \times 0,2 \times 0,8 + 2 \times 0,7 \times 0,3)/2 = 0,370$$

$$aa: (0,8^2 + 0,3^2)/2 = 0,365$$

## Wahlundův efekt:

např. 2 stejně velké démy,  $p_1 = 0,2$ ,  $p_2 = 0,7$ :

průměrné frekvence genotypů:

$$AA: (p_1^2 + p_2^2)/2 = (0,2^2 + 0,7^2)/2 = 0,265 \leftarrow$$

$$Aa: (2p_1q_1 + 2p_2q_2)/2 = (2 \times 0,2 \times 0,8 + 2 \times 0,7 \times 0,3)/2 = 0,370$$

$$aa: (0,8^2 + 0,3^2)/2 = 0,365$$

Pokud HW rovnováha, je frekvence  $A$  v celkové populaci:

$$p = (2f_{AA} + f_{Aa})/2 = (2 \times 0,265 + 0,370)/2 = 0,45$$

$\Rightarrow$  frekvence  $AA$  v celé populaci by měla teoreticky být

$$0,45^2 = \underline{0,2025} \text{ (ve skutečnosti } 0,265)$$

## Wahlundův efekt:

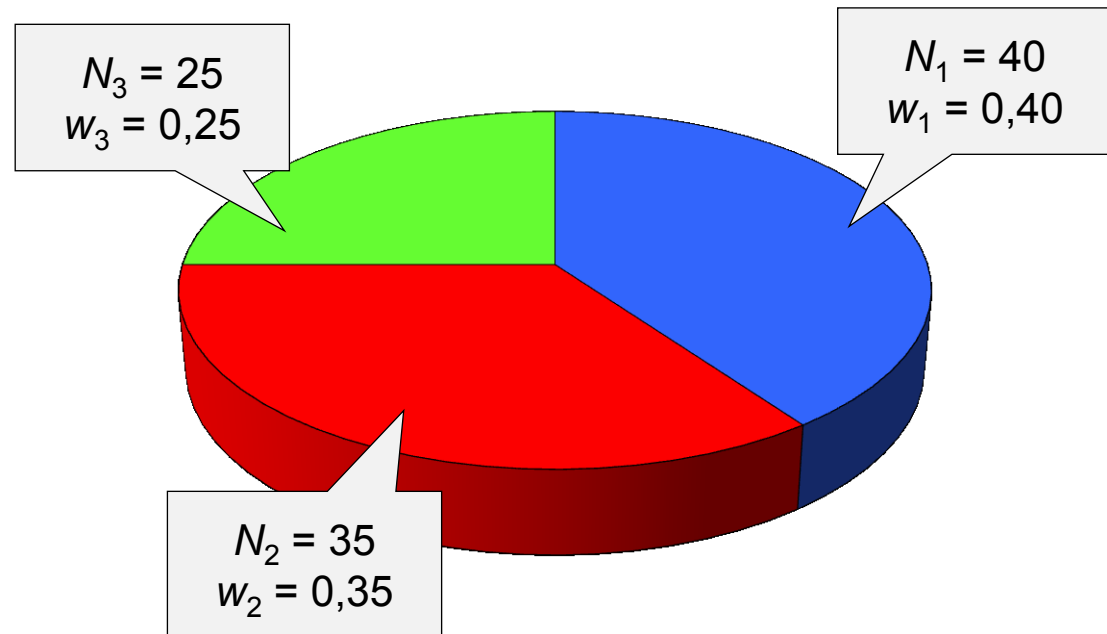
náhodné oplození, 1 lokus,  $A$ ,  $a$

$$N = \sum N_i, w_i = N_i/N, p_i$$

$$p = \sum w_i p_i, i\text{-tý dém: } AA: p_i^2, Aa: 2p_i q_i, aa: q_i^2$$

$$\text{celková populace: } f_{AA} = \sum w_i p_i^2, f_{Aa} = 2 \sum w_i p_i q_i, f_{aa} = \sum w_i q_i^2$$

např.:  $N = 100$



## Wahlundův efekt:

náhodné oplození, 1 lokus, A, a

$$N = \sum N_i, w_i = N_i/N, p_i$$

$$p = \sum w_i p_i, i\text{-tý dém: } AA: p_i^2, Aa: 2p_i q_i, aa: q_i^2$$

$$\text{celková populace: } f_{AA} = \sum w_i p_i^2, f_{Aa} = 2 \sum w_i p_i q_i, f_{aa} = \sum w_i q_i^2$$

rozptyl alelových frekvencí:

$$\text{Var}(p) = \sigma_p^2 = \sum_{i=1}^n w_i (p_i - \bar{p})^2 = \boxed{\sum_{i=1}^n w_i p_i^2} - \bar{p}^2$$

$= f_{AA}$

$$\sigma_p^2 = f_{AA} - \bar{p}^2 \Rightarrow f_{AA} = \bar{p}^2 + \sigma_p^2$$

$m = 0 \rightarrow$  nakonec náhodná fixace/extinkce ve všech démech

$\Rightarrow$  maximální rozptyl frekvencí:

$$Var(\max) = \bar{p}\bar{q} \text{ *)}$$

$\Rightarrow F_{ST}$  lze chápat jako standardizovaný rozptyl, tj. poměr skutečného rozptylu k teoretické maximální hodnotě:

$$F_{ST} = \frac{\sigma_p^2}{\bar{p}\bar{q}}$$

skutečný rozptyl

maximální rozptyl

\*)

$$Var(\max) = \bar{p}(1 - \bar{p})^2 + \bar{q}(0 - \bar{p})^2 = \bar{p}\bar{q}^2 + \bar{q}\bar{p}^2 = \bar{p}\bar{q}(\bar{p} + \bar{q}) = \bar{p}\bar{q}$$

$$f_{AA} = \bar{p}^2 + \sigma_p^2 \quad \text{a} \quad F_{ST} = \frac{\sigma_p^2}{\bar{p}\bar{q}} \Rightarrow \sigma_p^2 = \bar{p}\bar{q}F_{ST}$$

$$f_{AA} = \bar{p}^2 + \bar{p}\bar{q}F_{ST}$$

$$f_{Aa} = 2\bar{p}\bar{q}(1 - F_{ST})$$

$$f_{aa} = \bar{q}^2 + \bar{p}\bar{q}F_{ST}$$

měří odchylku od HW  
způsobenou rozdělením  
populace

$$f_{AA} = p^2 + pqF_{IS}$$

$$f_{Aa} = 2pq(1 - F_{IS})$$

$$f_{aa} = q^2 + pqF_{IS}$$

měří odchylku od HW  
způsobenou démovým  
inbreedingem



$F_{ST}$  i  $F_{IS}$  lze interpretovat různým způsobem, např.

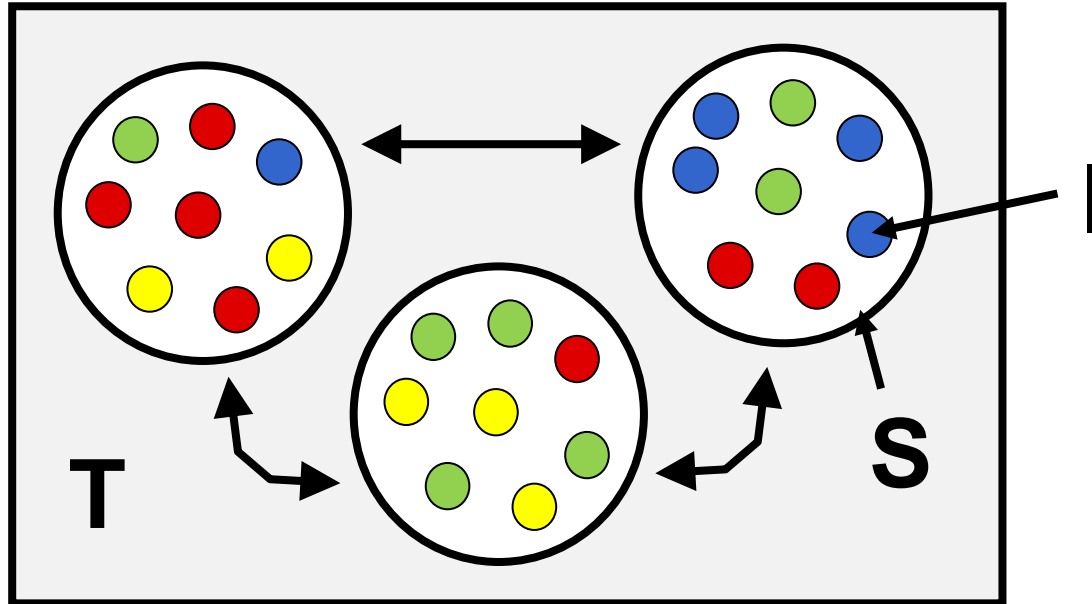
$F_{IS}$  = korelace mezi spojujícími se gametami

= vztah skutečné a očekávané heterozygotnosti v jediné nerozdělené populaci

$$F_{IS} = \frac{H_e - H_o}{H_e}$$

$$-1 \leq F_{IS} \leq +1$$

# Wrightova $F$ -statistika



S. Wright

koeficienty  $F_{IS}$ ,  $F_{ST}$ ,  $F_{IT}$

$$H_e \rightarrow H_S \quad H_o \rightarrow H_I$$

$$H_I = \frac{1}{n} \sum_{i=1}^n H_i \quad = \text{průměrná skutečná heterozygotnost v každé subpopulaci}$$

$$H_S = \frac{1}{n} \sum_{i=1}^n 2p_i q_i \quad = \text{průměrná očekávaná heterozygotnost v každém dému za předpokladu náhodného oplození}$$

$$H_T = 2\bar{p}\bar{q} \quad = \text{očekávaná heterozygotnost v celkové populaci za předpokladu náhodného oplození a žádné divergence frekvencí alel mezi subpopulacemi}$$

$$F_{IS} = \frac{H_S - H_I}{H_S}$$

průměrná skutečná heterozygotnost v démech

$$\sim F_{IS} = \frac{H_e - H_o}{H_e}$$

průměrná očekávaná heterozygotnost v démech

$$-1 \leq F_{IS} \leq +1$$

$F_{ST}$  = míra zvýšení Pr. identity stavem (*identity by state*, IBS) u alel vzorkovaných v téže subpopulaci (S) v porovnání se vzorkováním alel v celkové populaci (T):

$$F_{ST} = \frac{F_S - F_T}{1 - F_T}$$

pokud žádné rozdělení populace,  $F_S = F_T$

= odchylka od HW způsobenou rozdělením populace měřená pomocí rozptylu alelových frekvencí:


$$F_{ST} = \frac{\sigma_p^2}{\bar{p}\bar{q}}$$

nebo pomocí vztahu skutečné a očekávané heterozygotnosti:

$$f_{AA} = \bar{p}^2 + \bar{p}\bar{q}F_{ST}$$

$$f_{Aa} = 2\bar{p}\bar{q}(1 - F_{ST})$$

$$f_{aa} = \bar{q}^2 + \bar{p}\bar{q}F_{ST}$$


$$F_{ST} = 1 - \frac{f_{Aa}}{2\bar{p}\bar{q}} = \frac{2\bar{p}\bar{q} - f_{Aa}}{2\bar{p}\bar{q}} = \frac{H_T - H_S}{H_T}$$

pokud populace  
rozdělena, skutečná  
heterozygotnost je nižší,  
než teoretická

$H_T$  = očekávaná heterozygotnost v celkové populaci

$H_S$  = skutečná heterozygotnost v celkové populaci = průměrná HZ  
v subpopulacích

= **fixační index**

$$0 \leq F_{ST} \leq 1$$

$$F_{IT} = \frac{H_T - H_I}{H_T}$$

$$(1 - F_{IS})(1 - F_{ST}) = 1 - F_{IT}$$

v roce 1950 (původní Wrightův článek) pouze teoretické odvození  
pro 2 alely na lokus

v 70. letech elektroforéza → první skutečně genetická data

Weir and Cockerham (1984): odvození pro více alel

$$F_{ST} \rightarrow \theta$$

$$F_{IS} \rightarrow f$$

$$F_{IT} \rightarrow F$$

$$\hat{\theta} = \frac{s^2 - \frac{1}{2\bar{n} - 1} \left[ \bar{p}(1 - \bar{p}) - \frac{r - 1}{r} s^2 \right]}{\left[ 1 - \frac{2\bar{n}C^2}{(2\bar{n} - 1)r} \right] \bar{p}(1 - \bar{p}) + \left[ 1 + \frac{2\bar{n}(r - 1)C^2}{(2\bar{n} - 1)r} \right] \frac{s^2}{r}} .$$

## Odvození $F_{ST}$ ze sekvencí DNA:

$$F_{ST} = \frac{\pi_{\text{between}} - \pi_{\text{within}}}{\pi_{\text{between}}}$$

kde  $\pi$  = průměrný počet párových rozdílů mezi dvěma sekvencemi vybranými ze stejné subpopulace (*within*), nebo ze 2 různých subpopulací (*between*)

## Rovnováha driftu a toku genů:

Lze odvodit na základě inbreedingové efektivní velikosti populace v čase  $t$  generací:

$$\bar{F}(t) = \frac{1}{2N} + \left(1 - \frac{1}{2N}\right) \bar{F}(t-1)$$

v neidealizované populaci za  $N$  dosadíme  $N_{eF}$

Jaká je pravděpodobnost autozygotnosti v důsledku driftu a současně platí, že 2 náhodně vybrané gamety pocházejí ze stejného dému?

$$\bar{F}_t = \left[ \frac{1}{2N_{eF}} + \left(1 - \frac{1}{2N_{eF}}\right) \bar{F}(t-1) \right] (1-m)^2$$

platí pro malá  $m$   
(tj.  $m \gg m^2$ ,  
řádově  $\approx 1/N_{eF}$ )

Je-li populace v rovnováze, platí:

$$\hat{F} \approx \frac{1}{4N_{eF}m + 1}$$



## Rovnováha driftu a toku genů:

Lze odvodit na základě ostrovního modelu toku genů:

$$\sigma^2 = \frac{\bar{p}\bar{q}}{2N_{eV} - (2N_{eV} - 1)(1 - m)^2}$$

protože platí  $F_{ST} = \frac{\sigma_p^2}{\bar{p}\bar{q}}$  ,

pak pro malá  $m$ :  $\hat{F}_{ST} \approx \frac{1}{4N_{eV}m + 1}$

$$\hat{F}_{ST} \approx \frac{1}{4N_e m + 1}$$

Z toho plyne, že:

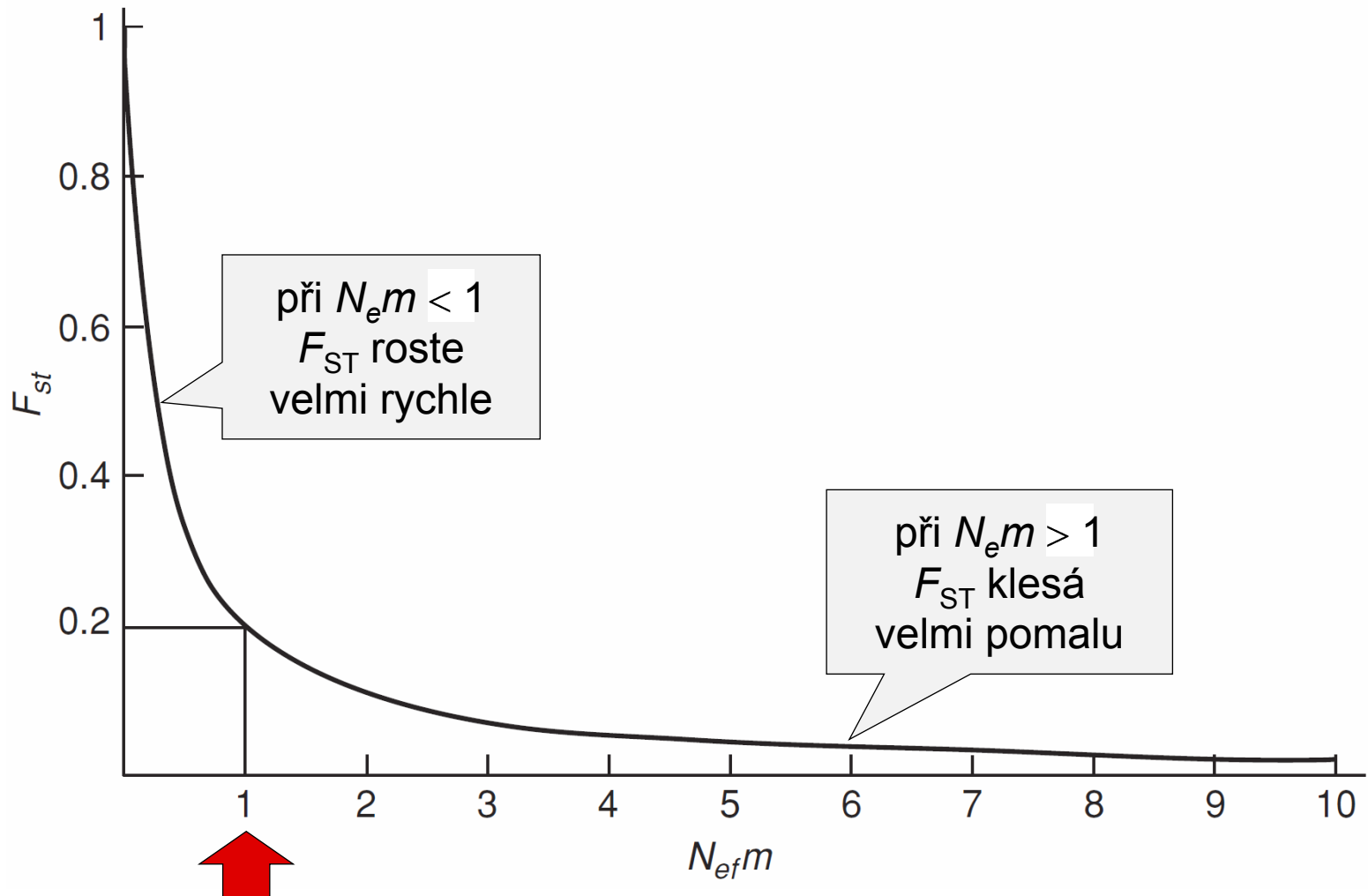
když větší tok genů ( $m$  roste)  $\Rightarrow F_{ST}$  klesá (větší variabilita v démech a menší rozdíly mezi nimi)

když  $1/N_e$  roste (větší drift)  $\Rightarrow F_{ST}$  roste (menší variabilita v démech a větší rozdíly mezi nimi)

i malý tok genů  $\Rightarrow$  2 subpopulace se efektivně chovají jako jedna evoluční linie

$\Rightarrow N_e m = 1$  (tj. 1 efektivní migrant za generaci) představuje významné rozhraní mezi relativním evolučním významem driftu a toku genů  
 $\rightarrow$  v průměru pouze jeden efektivní migrant za generaci stačí k tomu, aby tok genů převážil nad driftem a způsobil homogenizaci subpopulací

## Vztah počtu migrantů a míry izolace subpopulací:



## Míra homogenizace subpopulací nezávisí na míře toku genů, ale na efektivním počtu migrantů

např.  $N_e m = 1$  znamená, že dvě subpopulace sdílejí 80 % genů při  $N = 10^9$  i při  $N = 100$ , ale  $m$  v prvním případě =  $10^{-9}$ , ve druhém 0,01

Naopak při stejném  $m$  můžeme očekávat různé evoluční dopady

např.  $m = 0,01$ :  $N_e = 10^9 \Rightarrow F_{ST} \approx 0$ ;  $N_e = 100 \Rightarrow F_{ST} = 0,2$

$F_{ST}$  představuje rovnováhu mezi driftem a tokem genů

velké populace: pomalá divergence  $\Rightarrow$  stačí i slabá migrace

malé populace: rychlá divergence  $\Rightarrow$  tok genů musí být silnější

Z uvedeného také plyne, že pokud známe pouze  $m$  nebo jen  $N_e$ , nemůžeme nic usuzovat o populační struktuře. I druh s velkou lokální velikostí populace může vykazovat extrémní populační subrozdělení a naopak

Př.: plž *Rumina decollata* – početný, ale malá disperze  $\Rightarrow$   
 $F_{ST}$  v Montpellier (zahrady kolem Boulevard des Arceaux) = 0,294



naopak *Mus musculus* v Texasu – malé démy, ale větší disperze  $\Rightarrow$   
 $F_{ST} = 0,021$

Rozsah genetického rozdělení populací závisí  
na rovnováze mezi driftem a tokem genů,  
ne na jednom či druhém mechanismu samotném!

$F_{IS}$ ,  $F_{ST}$  a  $F_{IT}$  na příkladu indiánů kmene Janomamo:

tabu incestu  $\Rightarrow F_{IS} = -0,01$

malé  $N_e \Rightarrow F_{ST} = 0,073$

$$\begin{aligned} F_{IT} &= F_{ST} + F_{IS}(1 - F_{ST}) = \\ &= 0,073 - 0,01 \cdot 0,927 = 0,064 \end{aligned}$$



$\Rightarrow$  díky Wahlundově efektu a rozdělení populace v celém vzorku pozorujeme redukci heterozygotnosti, naznačující inbreeding, přestože ve skutečnosti se Janomamové vyhýbají příbuzenským svazkům

$F_{IT}$  směšuje lokální systém páření a rozdělení populace

## Vliv rozdělení populace na celkovou variabilitu - $N_{eF}$ :

pokud areál druhu fragmentován a mezi izoláty žádný tok genů,  
bude celková  $N_{eF}$  rovna lokální  $N_{eF}$ .

⇒ v této situaci se bude rodokmenový inbreeding akumulovat v důsledku driftu rychlostí, která závisí na velikosti fragmentů, ne na celkové populační velikosti druhu → důležité pro ochranu!

i když  $m > 0$ , pořád bude platit  $N_{eF} \ll N$

nicméně i nepatrný tok genů vede k tomu, že celková  $N_{eF}$  je o několik řádů větší než celková  $N_{eF}$  při úplné izolaci subpopulací!

## Vliv rozdělení populace na celkovou variabilitu - $N_{eV}$ :

$N_{eF}$  s rostoucí izolací dómů klesá  $\times$   $N_{eV}$  má opačný trend:

celková  $N_{eV}$  měří rozptyl frekvencí alel v důsledku driftu v populaci jako celku, ne pro každou lokální subpopulaci

Wright (1943): idealizovaná populace složená z  $n$  dómů o stejné velikosti  $N \Rightarrow$  velikost celé populace =  $nN$ ; tok genů =  $m$   
 $\Rightarrow$  pokud by populace byla panmiktická,  $N_{eV}$  by bylo =  $nN$

celková varianční efektivní velikost populace je:

$$N_{eV} = \frac{nN}{1 - F_{ST}}$$

tj. při  $F_{ST} = 0 \rightarrow N_{eV} =$  počet jedinců v celé populaci



Vliv rozdělení populace na celkovou variabilitu -  $N_{eV}$ :

$$N_{eV} = \frac{nN}{1 - F_{ST}}$$

ALE: při  $F_{ST} > 0$  (tj.  $1 - F_{ST} < 1$ ) bude  $N_{eV} > nN$

To znamená, že v populaci rozdělené na subpopulace je celková  $N_{eV}$  větší než celkový počet jedinců  $nN$ !!

S tím, jak  $F_{ST} \rightarrow 1$ ,  $N_{eV} \rightarrow \infty$ , i když počet jedinců v populaci je omezený!

Podobně pro kontinuální populaci s omezeným tokem genů (*isolation by distance*) platí, že ztráta alelické variability v ní je nižší než v panmiktické populaci.

Jak je možné, že rozdělení populace zvyšuje  $F$  a současně snižuje rychlost ztráty variability na úrovni celkové populace?

→ S postupnou fixací alel v lokálních démech se zvyšuje autozygotnost (tj. zvyšuje se  $N_{eF}$ )

ALE protože v démech jsou náhodně fixovány různé alely, celková variabilita zůstává zachována; v okamžiku fixace už žádný drift nemůže působit ⇒ s fixací alel v dalších démech se celková rychlost ztráty variability v démech zpomaluje

⇒ dilema pro ochranářskou praxi....

## ... dilema pro ochranářskou praxi:

prioritou zajistit tok genů mezi izoláty → zvýšení lokální  $N_{eF}$  i  $N_{eV}$ ,  
zvýšení genetické variability a tím lokální adaptivní flexibility, redukce  
celkového rodokmenového inbreedingu a tím snížení rizika  
inbrední deprese

ALE: úkolem také dlouhodobě zachovávat vysokou míru variability  
v celé populaci (druhu) – toho ovšem nejsnáze dosáhneme  
fragmentováním populace na izoláty ⇒ *trade-off*

→ otázka priorit: ztráta variability v izolátech × reintrodukce z jiných  
izolátů (nejen potenciální outbrední deprese)

## Rozdělení populací z hlediska sekvencí DNA – AMOVA:

Dosud 2 homologní geny

→ jsou nebo nejsou identické původem?

→ jsou nebo nejsou heterozygotní?

× se sekvencemi je to složitější: např. 2 sekvence 10 kb se můžou lišit v 1 nukleotidu, nebo v 50 nukleotidech

⇒ alternativy, např.:  $N_{ST}$ ,  $K_{ST}$

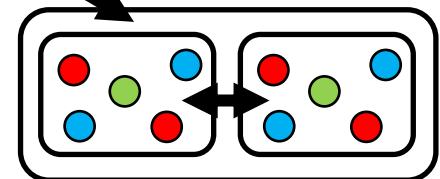
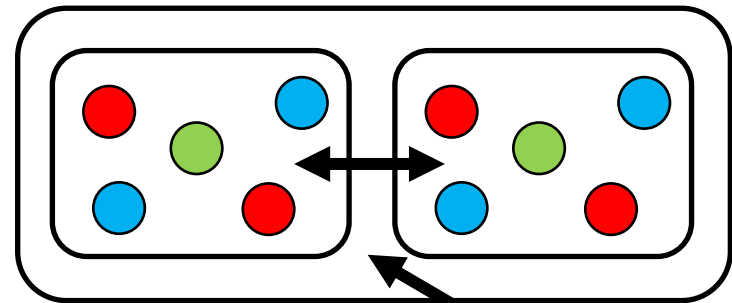
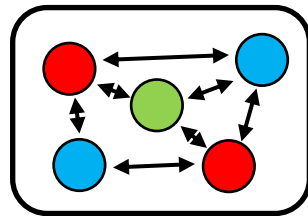
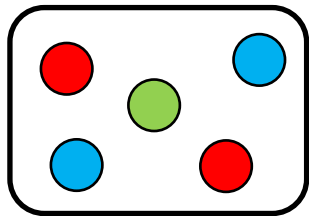
$\Phi_{ST}$  = používá molekulární genetické distance místo heterozygotnosti

→ analýza molekulárního rozptylu = AMOVA  
(*analysis of molecular variance*)

AMOVA (Excoffier et al. 1992): program Arlequin

několik hierarchických úrovní

variabilita v rámci jedince, variabilita mezi jedinci v populaci,  
mezi populacemi v rámci skupiny populací,  
variabilita mezi populacemi z různých skupin, celková variabilita



$$\Phi_{SC} = \frac{\sigma_b^2}{\sigma_b^2 + \sigma_c^2}$$

$$\Phi_{CT} = \frac{\sigma_a^2}{\sigma_T^2}$$

$$\Phi_{ST} = \frac{\sigma_a^2 + \sigma_b^2}{\sigma_T^2}$$

Pozor, někdy odlišné značení:  $\Phi_{SC} \sim F_{IS}$ ;  $\Phi_{CT} \sim F_{ST}$ ;  $\Phi_{ST} \sim F_{IT}$

## Populační struktura a různé typy dědičnosti:

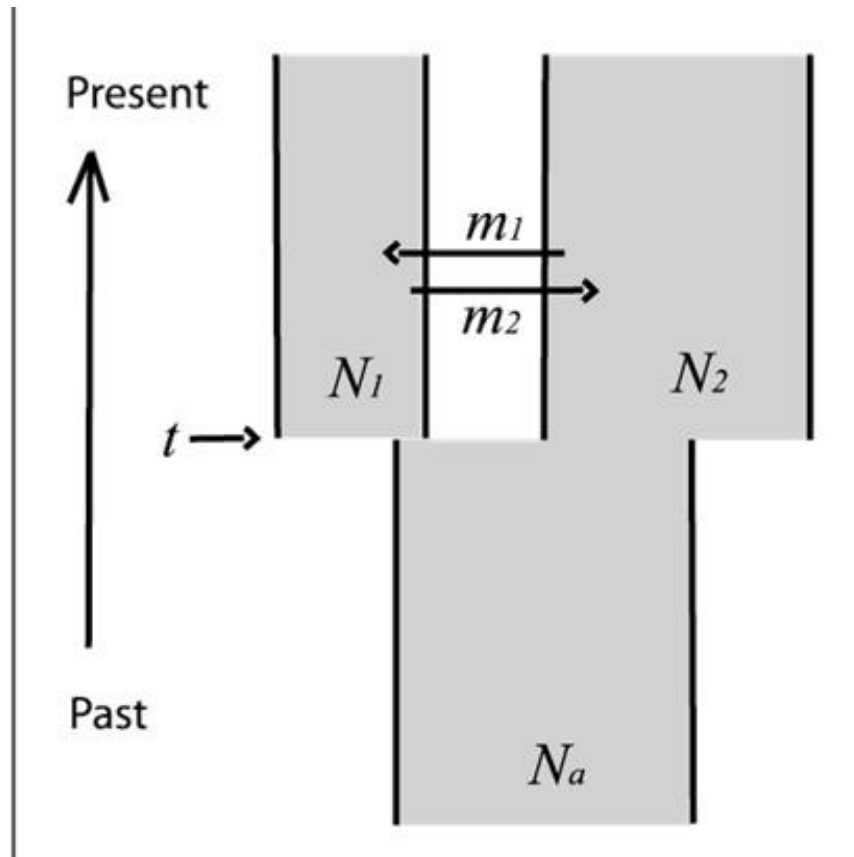
haploidní systém:  $F_{ST} = \frac{1}{1 + 2N_{eV}m}$

mtDNA, chr. Y:  $F_{ST} = \frac{1}{1 + N_{eV}m}$

chr. X při poměru pohlaví 1:1:  $F_{ST} = \frac{1}{1 + 3N_{eV}m}$

## Model izolace s migrací:

program IM (Hey and Nielsen 2004) a IMA (Hey and Nielsen 2007)  
... liší se způsobem odhadu parametrů



2 populace z 1 ancestrální  
celkem 6 parametrů:  
 $N_a, N_1, N_2$  ... pop. velikosti  
 $m_1, m_2$  ... tok genů  
 $t$  ... čas rozdělení populace

## Předpoklady IM a IMA modelu:

Historie populací reprezentována bifurkačním fylogenetickým stromem s kořenem

Topologie stromu je známa

Každá populace včetně ancestrální má konstantní velikost a je v souladu s předpoklady Wrightova-Fisherova modelu

Tok genů jednosměrný nebo obousměrný

Nedošlo k toku genů mezi nestudovanými (*unsampled*) a studovanými (*sampled*) populacemi (včetně ancestrální)



Extenze IMA modelu na  $\geq 2$  populace  
 Hey (2010):

pro  $k = 3$  populace 7 + 8 parametrů  
 počet migračních parametrů  
 rychle narůstá:

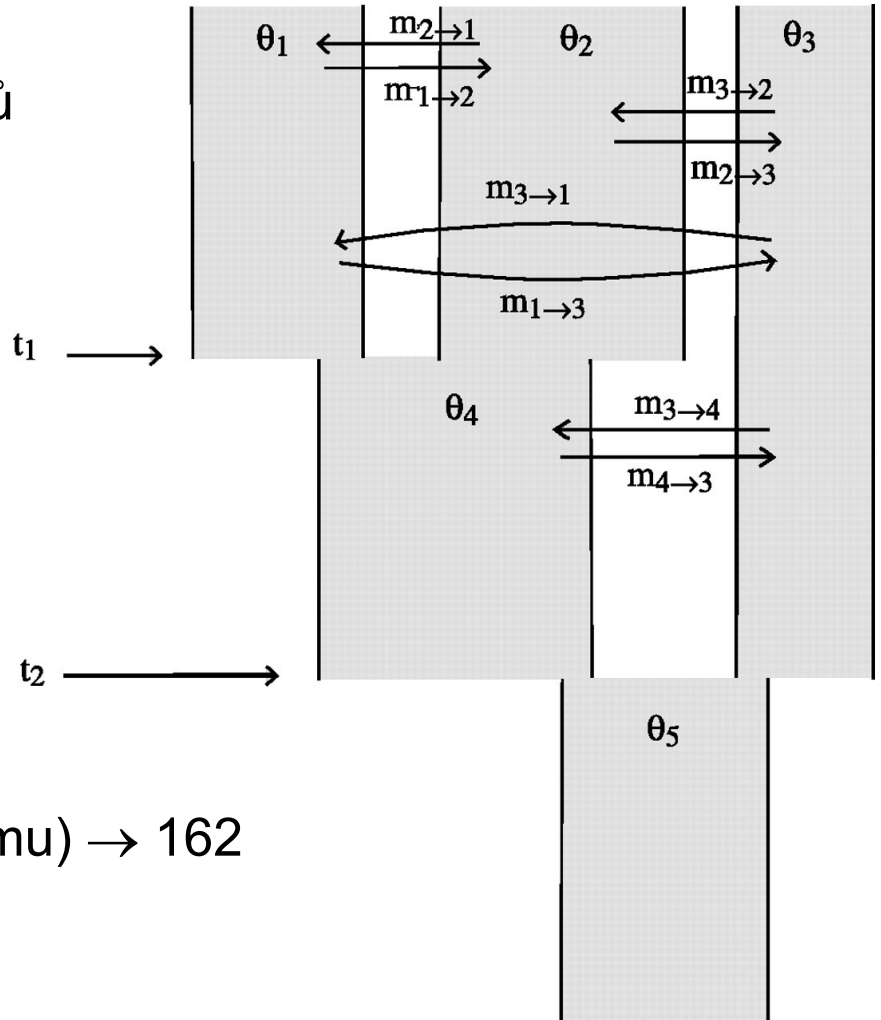
$$k(k - 1) + \sum_{i=2}^{k-1} 2(k - i) = 2(k - 1)^2$$

tj. pro  $k = 3 \rightarrow 8$

$k = 4 \rightarrow 18$

$k = 10$  (maximální hodnota programu)  $\rightarrow 162$   
 migračních parametrů

pro 3 populace  
 celkem 15  
 parametrů



# Pozor!!!

Všechny odhady populační struktury a toku genů jsou založeny na předpokladu, že drift a tok genů jsou v rovnováze!

Př. 1: populace recentně rozdělena na 2 velké subpopulace, mezi nimiž žádná migrace ( $m = 0$ )  $\rightarrow F_{ST} = 1$

× ve skutečnosti do dosažení rovnovážného stavu  $F_{ST} < 1 \Rightarrow$  chybně bychom dospěli k závěru, že  $m > 0$ , přestože ve skutečnosti obě subpopulace zcela izolované!

Př. 2: současná expanze areálu původně malé a homogenní populace  $\Rightarrow$  do dosažení rovnováhy budeme detekovat vysokou míru toku genů bez ohledu na skutečné hodnoty  $m$

# Pozor!!!

Př. 3: recentní tok genů mezi dlouhodobě izolovanými démy → protože  $F_{ST}$  klesá pomalu, dočasně budeme detekovat větší  $m$  než ve skutečnosti

Z uvedeného plyne, že  $F_{ST}$  je účinným indikátorem toku genů během evoluce pouze v případě, že populace jsou v rovnováze a nebyly ovlivněny recentními jevy!