

**DRIFT II**

# EFEKTIVNÍ VELIKOST POPULACE

HW – nekonečná populace

× omezená velikost populace

⇒ vyšší IBD (autozygotnost)

⇒ zvyšující se rozptyl frekvencí alel mezi démy v čase

⇒ fixace/extinkce alel

všechny tyto jevy nepřímo úměrné velikosti populace  $N$

Použitím idealizované populace můžeme odvodit přesný kvantitativní vztah driftu a velikosti populace → **Wrightův-Fisherův model**

# Wrightova-Fisherova populace:

diploidní, hermafrodit

velikost omezená, žádné fluktuační  $N$

náhodné oplození

kompletní izolace (žádný tok genů)

diskrétní generace

žádná věková struktura

žádná selekce

rozptyl výběru gamet do další generace<sup>\*)</sup>

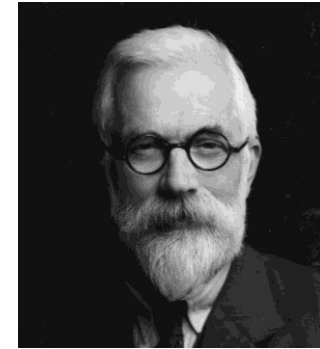
→ Poissonovo rozdělení

<sup>\*)</sup> tj. každý jedinec může přispět

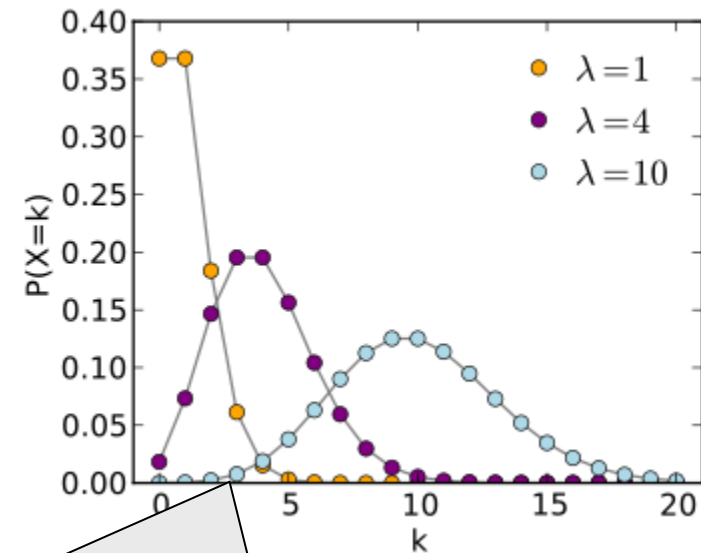
0, 1, 2, 3, ... gametami



S. Wright



R.A. Fisher



**Poissonovo rozdělení**  $\approx$  biomické pro velký počet pokusů ( $n \rightarrow \infty$ ) a malou pravděpodobnost jevu ( $p \rightarrow 0$ ) (v praxi  $n > 30$ ;  $p \leq 1/10$ , pak  $\lambda = np$ ) např. mutace

**Efektivní velikost populace  $N_e$**  nám umožňuje měřit sílu driftu v neideální populaci

Př.: Tristan da Cunha: porušení všech předpokladů W-F modelu  
předpokládáme prům.  $F = 0$  v zakladatelské generaci  
můžeme vypočítat prům. generační dobu a prům.  $F$  v každé gen.

Jak velké  $N$  potřebujeme použít pro idealizovanou populaci, abychom dostali stejné prům.  $F$  při stejných startovních podmínkách a po stejném počtu generací?

Výsledné  $N = N_e$

**Efektivní velikost populace = počet jedinců idealizované populace, která vykazuje hodnotu dané populačně genetické veličiny\*) rovnou hodnotě této veličiny ve studované neidealizované populaci**

\*) zde buď zvýšení průměrné pravděpodobnosti autozygotnosti na autozomálním lokusu ( $N_{eF}$ ), nebo zvýšení rozptylu frekvencí alel přes generace/přes subpopulace ( $N_{eV}$ )

## Podobně jako u koeficientu inbreedingu neexistuje jediná efektivní velikost!!

Díky  $N_e$  můžeme měřit sílu driftu v reálných populacích, které mohou porušovat různé předpoklady WF modelu v různé míře

Např. efekt zakladatele:

- 1) 20 gonochoristů (ne-hermafroditů) – 10 samic, 10 samců  
jinak ~ W-F model
- 2) 25 jednodomých rostlin – 50 % samosprašnost, 50 % náhodné opylení  
jinak ~ W-F model

Ve kterém případě je vliv driftu silnější? →  $N_e$

# Inbreedingová efektivní velikost $N_{eF}$

**Generace 0:**  $F(0) = 0$ ; velikost  $N \Rightarrow$  v každé generaci  $2N$  gamet

**Generace 1:** žádná dvojice gamet nemůže být IBD ( $F(0) = 0$ ), ale protože jde o hermafrodity (def. W-F), můžou pocházet od stejného rodiče

Jaká je Pr., že 2 gamety pocházejí od stejného rodiče?  
protože předpokládáme náhodné oplození,  $Pr = 1/(N)$   
jestliže došlo k samooplození, Pr. že obě alely totožné =  $1/2$

$\Rightarrow$  celková Pr (IBD v Gen.1) =  $1/(N) \times 1/2 = 1/(2N) = F(1)$   
tj. Prům. Pr. IBD se v 1. generaci zvýšila o  $1/(2N)$

**Generace 2:**

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

Pr. autogamie  
v Gen1

IBD v důsledku spojení gamet stejného prarodiče, vážená Pr., že tyto gamety už nejsou IBD díky autogamii

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

po několika matematických tricích: 
$$\bar{F}(t) = 1 - \left(1 - \frac{1}{2N_{eF}}\right)^t$$

a tedy

$$N_{eF} = \frac{1}{2\{1 - [1 - \bar{F}(t)]^{1/t}\}}$$

$N_{eF}$  závisí jen na prům.  
 $F$  a generačním čase  $t$ ,  
nezávisí na  $N$ !

*G. spekei*, 1979,  $N = 19$ :

$F = 0,1283$ ; prům.  $t = 1,7 \Rightarrow N_{eF} = 6,4$  ( $N_{eF} \approx 1/3 N$ )

kdybychom ale vzali rodiče 19 jedinců z 1979 jako referenční generaci  $\rightarrow N_{eF}$  pro 15 potomků = 48,1 ( $N_{eF} \approx 2,5 N$ )





*G. spekei*, 1979,  $N = 19$ :

$F = 0,1283$ ; prům.  $t = 1,7 \Rightarrow N_{eF} = 6,4$  ( $N_{eF} \approx 1/3 N$ )

kdybychom ale vzali rodiče 19 jedinců z 1979 jako referenční generaci  $\rightarrow N_{eF}$  pro 15 potomků = 48,1 ( $N_{eF} \approx 2,5 N$ )

problém v tom, že referenční generace 19 jedinců z definice definována jako  $F = 0$

$\Rightarrow$  v prvním případě  $N_{eF}$  ukazuje rychlost akumulace inbreedingu po založení populace, ve druhém  $N_{eF}$  kvantifikuje účinek nového – outbredního – chovného programu

**Efektivní velikost populace může být i vyšší než  $N$  (*census population size*)!**

**Pro stejnou populaci můžeme dostat velmi rozdílné efektivní velikosti v závislosti na volbě referenční populace!**

## Varianční efektivní velikost $N_{eV}$

Genetický drift způsobuje 1) náhodné odchylky od frekvencí alel ( $p$ ) v předchozí generaci a 2) rozdíly v  $p$  mezi subpopulacemi → obojí měřeno pomocí rozptylu (variance)

Idealizovaná populace: 1 lokus; alely  $A$ ,  $a$ ; velikost  $N$ ,  
 $p$  = frekvence alely  $A$ ,  $q = 1 - p$  (frekvence  $a$ );  $x$  = počet alel ve vzorku  
sampling alel v souladu s binomickým rozdělením

protože drift nemá směr, očekávaná frekvence  $A$  v další generaci =  $p$

$$p = x/(2N)$$

$$\text{Rozptyl alelových frekvencí} = \text{Var}\left(\frac{x}{2N}\right) = \frac{\text{Var}(x)}{(2N)^2} = \frac{2Npq}{(2N)^2} = \frac{pq}{2N}$$

v neidealizované populaci  $N \rightarrow N_{eV}$

Rozptyl v čase roste  $\Rightarrow$  po  $t$  generacích:

$$\text{Var}(p_t) = pq \left[ 1 - \left( 1 - \frac{1}{2N} \right)^t \right]$$

s  $t \rightarrow \infty$   $\text{Var} \rightarrow \text{Var}(\text{max}) = pq$   
(= stav, kdy všechny alely buď  
fixovány, nebo ztraceny)

v neidealizované populaci  $N \rightarrow N_{eV}$ ;  $\text{Var}(p_t) \rightarrow \sigma_t^2$

a **varianční efektivní velikost**

$$N_{eV} = \frac{1}{2 \left\{ 1 - \left[ 1 - \frac{\sigma_t^2}{pq} \right]^{1/t} \right\}}$$

závisí pouze na  $\text{Var}$ ,  $p$ ,  
 $q$  a  $t$ , nezávisí na  $N$ !

frekvence  $A$  v  
počáteční generaci

$= 1 - p$

*G. spekei*,  $N = 15$ :

rodiče 19 jedinců z 1979 jako referenční generace  
→  $N_{eV} = 20,1$  ( $N_{eF} = 48,1 \approx 239 \% N_{eV}$ )



### Inbreeding, drift a ztráta alelické diverzity:

Ani děmový, ani rodokmenový inbreeding nemá přímý dopad na frekvence alel  $\Rightarrow$  sám o sobě ani nezrychluje, ani nezpomaluje ztrátu genetické variability

V mnoha reálných, početně omezených populacích existuje negativní korelace mezi inbreedingem a variabilitou

$\Rightarrow$  drift, ne inbreeding je skutečnou příčinou ztráty genetické variability

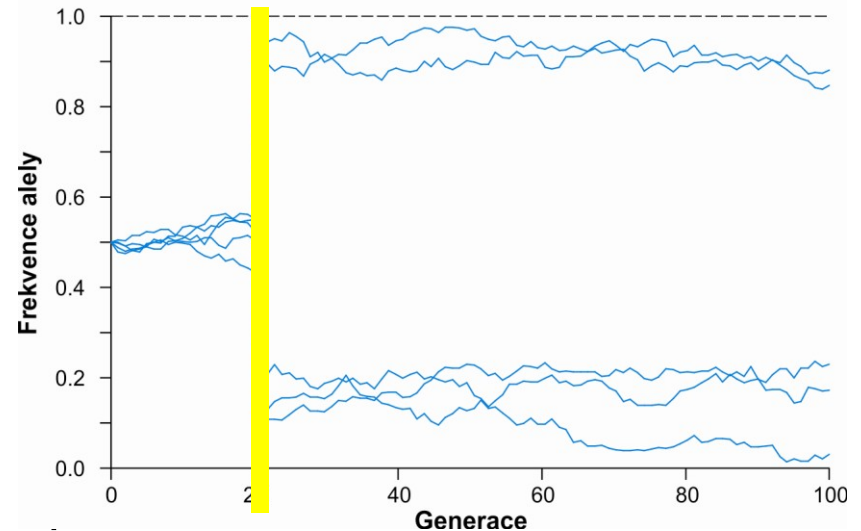
$\Rightarrow$  inbreeding nezpůsobuje ztrátu alelické diverzity v malých populacích, její příčinou je samotné omezení velikosti populace

## Srovnání $N_{eF}$ a $N_{eV}$ :

$N_{eF}$  je citlivá k počtu rodičů

$N_{eV}$  je citlivá k počtu potomků

Př.: redukce  $N$  ve 21. generaci  
z 1000 na 4  
a zvýšení na 1000 ve 22. generaci:



$N_{eF}$ : gamety náhodně vybrány z 1000 rodičů  $\rightarrow$  4 jedinci v gen. 21

pravděpodobnost, že dvě vybrané gamety od stejného rodiče =  $1/1000$

pravděpodobnost, že se tyto gamety účastnily samooplození a jsou

$$\text{IBD} = \frac{1}{2} \Rightarrow \text{prům. } F \text{ ve 21. generaci} = \frac{1}{2000} \Rightarrow N_{eF} = \underline{1000} \quad (N = 4)$$

$N_{eV}$ : vybráno pouze 8 gamet  $\Rightarrow p(21) = pq/8 \Rightarrow N_{eV}(21) = \underline{4} \quad (N = 4)$

## $N_e$ a děmový inbreeding $F_{IS}$ :

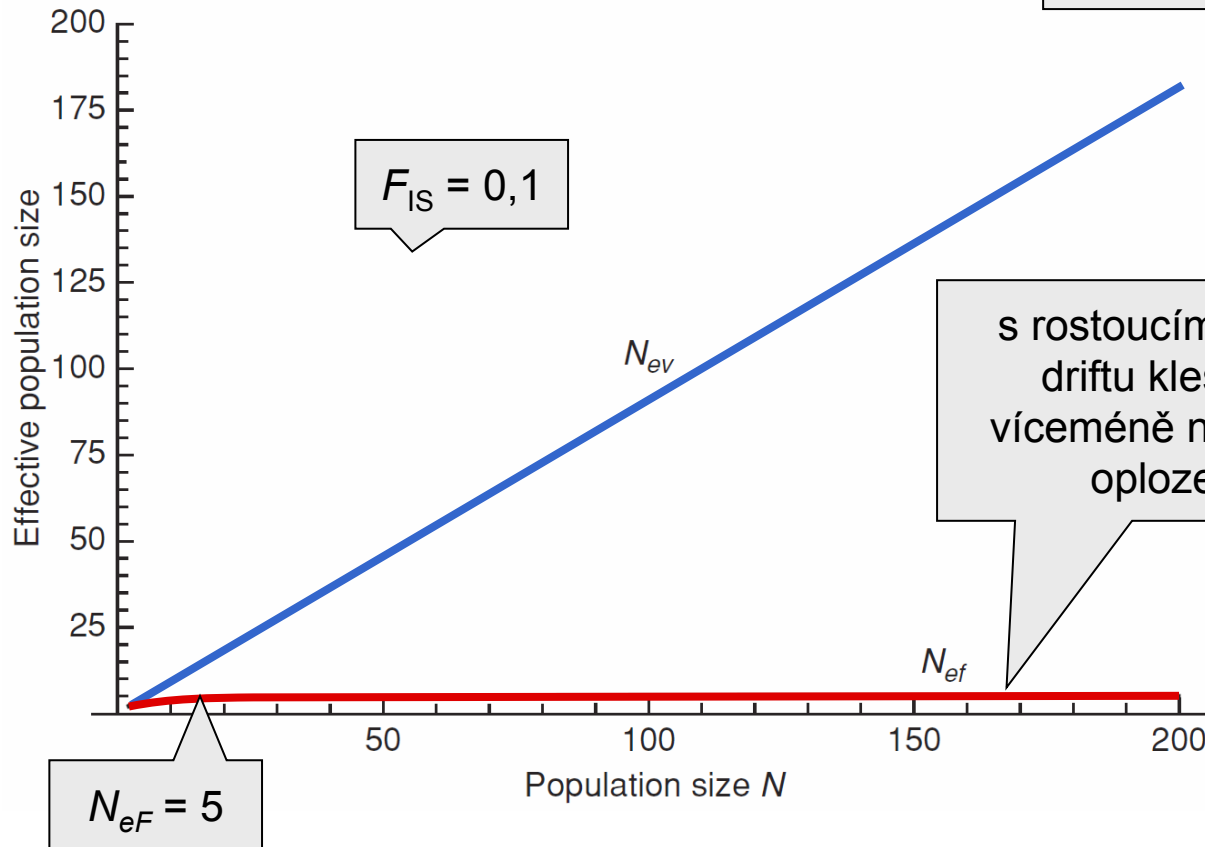
s rostoucím  $N$  se  $N_{eF}$  moc nezvětšuje

$$N_{eF} = \frac{N}{1 + F_{IS}(2N - 1)}$$

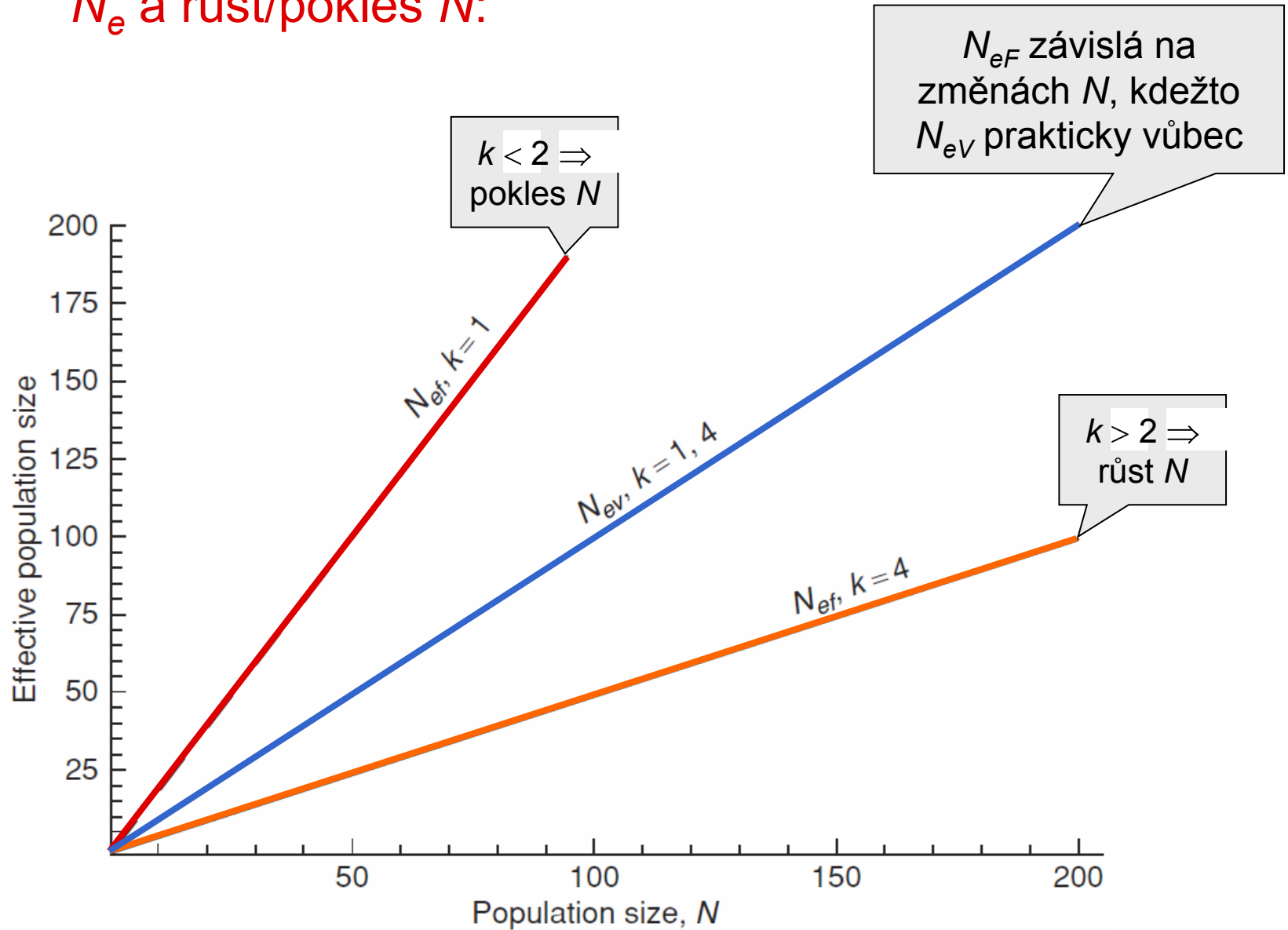
$$N_{eV} = \frac{N}{1 + F_{IS}}$$

s rostoucím  $N$  se  $N_{eF}$  zvětšuje

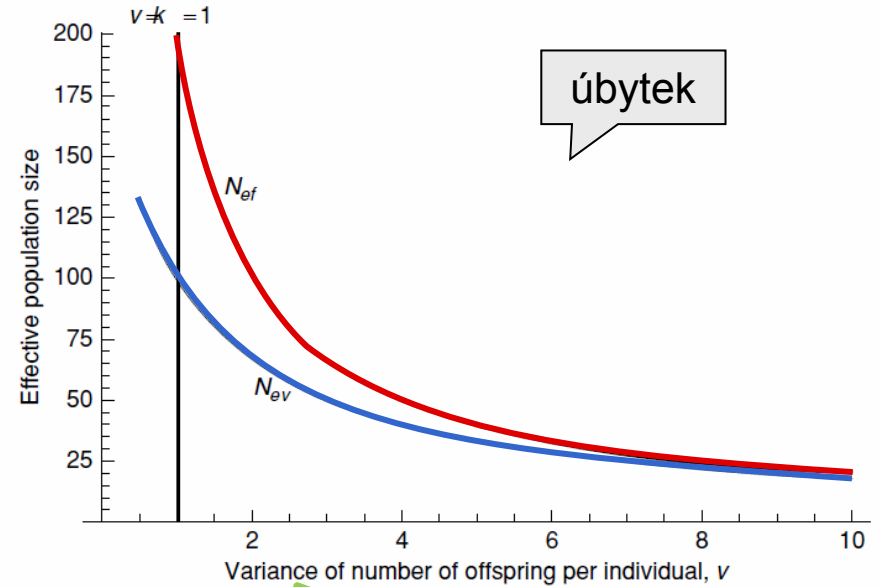
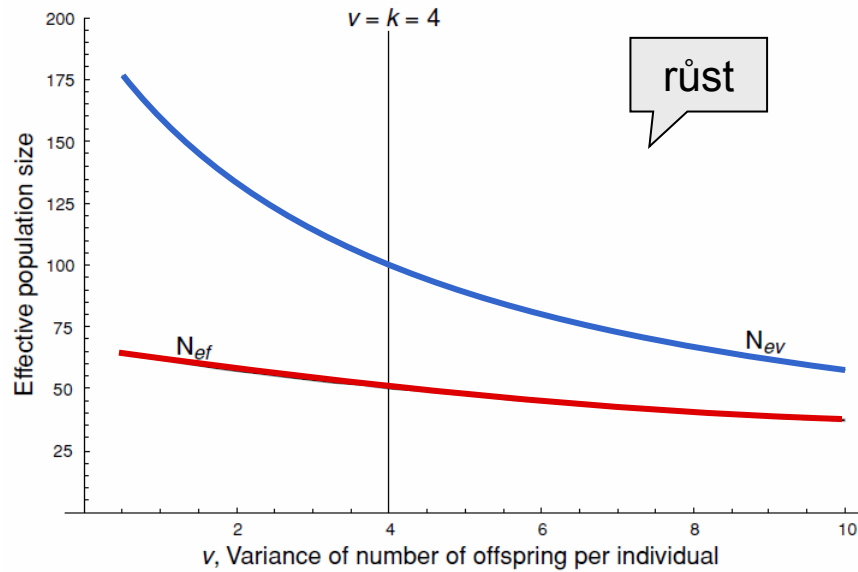
při  $F_{IS} = 0$   $N_{eF} = N_{eV}$



# $N_e$ a růst/pokles $N$ :



# $N_e$ a rozdíly v počtu potomků (rozpor s Poissonovým rozdělením):



čím větší rozptyl počtu potomků, tím nižší  $N_{eF}$  i  $N_{eV}$

pokud populace roste ( $k > 2$ ),  $N_{eV}$  citlivější k rozptylu počtu potomků než  $N_{eF}$

× pokud populace klesá ( $k < 2$ ),  $N_{eV}$  méně citlivá k rozptylu počtu potomků než  $N_{eF}$



## Každý genetický znak vyžaduje vlastní $N_e$

Pro geny na autozomech, pohlavních chromozomech a mtDNA existuje odlišná efektivní velikost populace:

autozomy:  $N_e$  nebo  $4N_e$

X, Z:  $\frac{3}{4}N_e$  nebo  $3N_e$

Y, W, mtDNA:  $\frac{1}{4}N_e$  nebo  $1N_e$

## Vliv kolísání populační velikosti:

efektivní velikost lze aproximovat jako harmonický průměr  $\Rightarrow$  velký vliv malých  $N$ !!

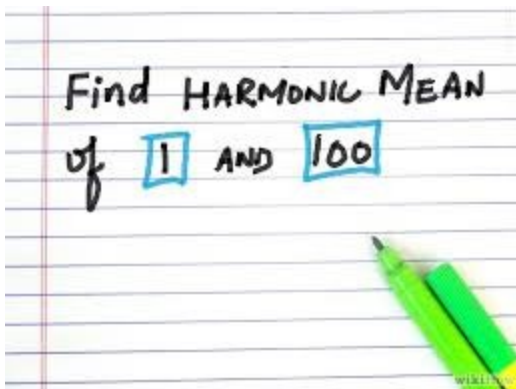
$$N_{eV} = \frac{t}{\frac{1}{N_1} + \frac{1}{N_2} + \dots + \frac{1}{N_t}}$$

harmonický průměr

$$N_{eF} = \frac{t}{\frac{1}{N_0} + \frac{1}{N_1} + \dots + \frac{1}{N_{t-1}}}$$

na rozdíl od  $N_{eV}$  závislost na rodičovské generaci

průměr mnohem blíže nižší hodnotě!



Handwritten formula on lined paper:  $H = \frac{2}{\frac{1}{a_1} + \frac{1}{a_2}}$   
 $= \frac{2}{\frac{1}{1} + \frac{1}{100}}$ . A green highlighter is visible at the bottom right.

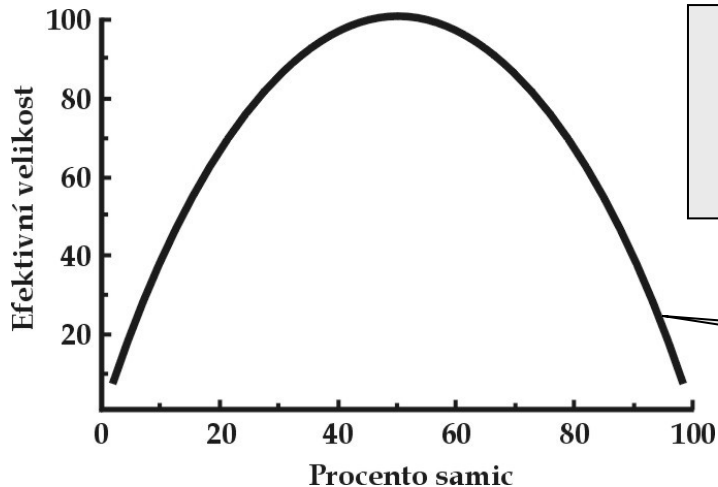
Handwritten calculation on lined paper:  $= \frac{2}{1.01}$   
 $= 1.9802$ . The result 1.9802 is enclosed in a blue box. Below it, it says "Harmonic Mean of 1 and 100". A green highlighter is visible at the bottom right.

## Vliv vychýleného poměru pohlaví:

předchozí výpočty a aproximace předpokládaly stejný počet samců a samic přispívajících svými geny do dalších generací

$N_m$  = počet rozmnožujících se samců,  $N_f$  = počet samic

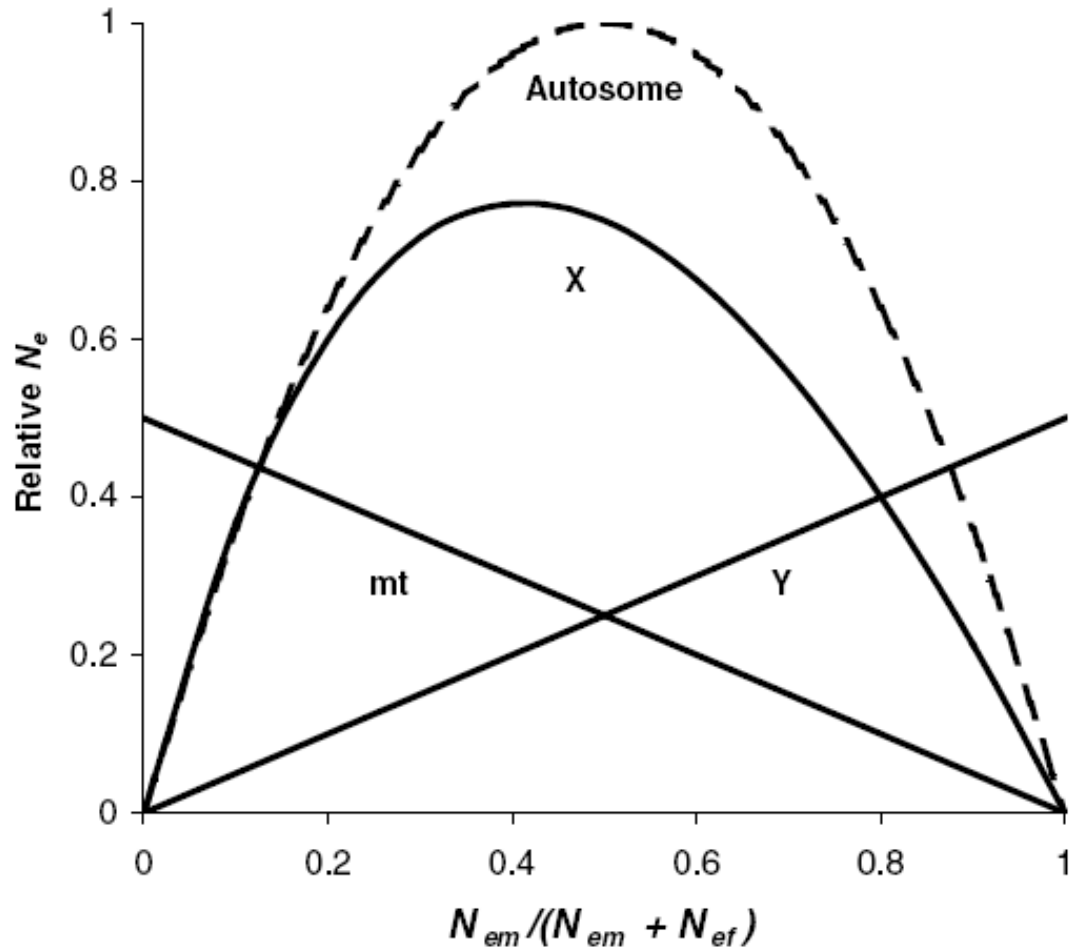
$$N_e = \frac{4N_m N_f}{n_{\pm}} \rightarrow N_e = \frac{4N_f}{n_{\pm}} \approx$$



z uvedeného vztahu vyplývá, že pokud se v populaci bude rozmnožovat pouze jeden samec, bude  $N_e \approx 4$  bez ohledu na celkový počet jedinců

čím větší odchylky od vyrovnaného poměru pohlaví, tím nižší  $N_e$

vliv poměru pohlaví na  $N_e$  odlišný pro různé genetické znaky:

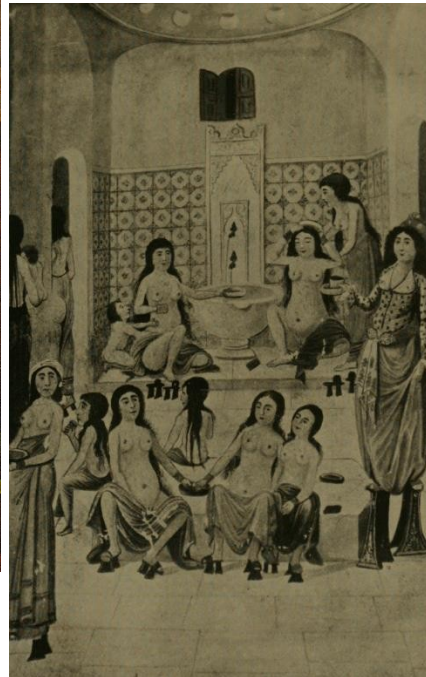


## Vliv vychýleného poměru pohlaví:

Př.: rypouš sloní (*Mirounga leonina*):

v harému poměr pohlaví 1:40

ale efektivní poměr 1:4-5 díky nevěrám a krátké době dominance samce (1-2 roky)



## Geny pod selekcí:

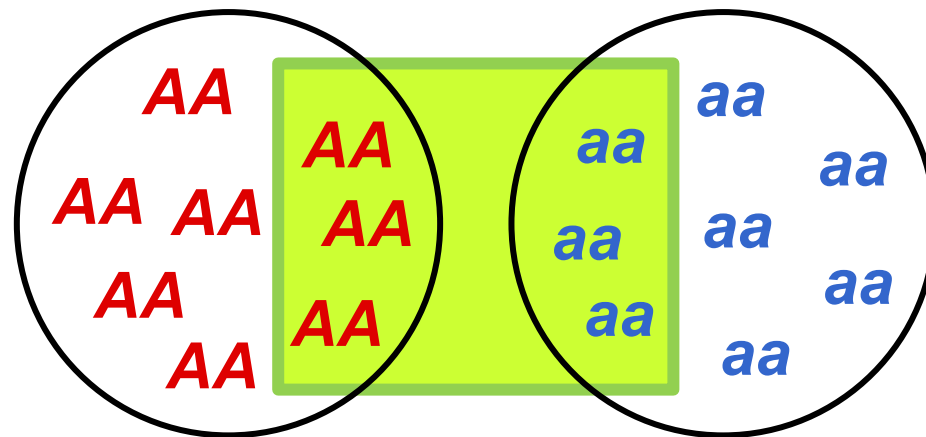
Jestliže na gen působí selekce, je rozptyl v počtu potomků mezi jedinci v populaci vysoký (jedinci s výhodnou alelou zanechají více potomstva)

⇒  $N_e$  pro tento gen nižší než pro gen selektivně neutrální

Platí i pro části genů!

# DRIFT V ROZDĚLENÉ POPULACI

Wahlundův efekt



absence  
heterozygotů

## Wahlundův efekt:

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

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

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

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

frekvence  $A$  v celkové populaci:  $(2 \times 0,265 + 0,370)/2 = 0,45$

$\Rightarrow$  frekvence  $AA$  v celé populaci =  $0,45^2 = 0,2025 < \underline{0,370}$



## Wahlundův efekt:

náhodné oplození, 1 lokus, A, a  
 $N, N_i, w_i = N_i/N, p_i$

$p = \sum w_i p_i$ ,  $i$ -tý dům: AA:  $p_i^2$ , Aa:  $2p_i q_i$ , aa:  $q_i^2$

celková populace: AA:  $\sum w_i p_i^2$ , Aa:  $2 \sum w_i p_i q_i$ , aa:  $\sum w_i q_i^2$

rozptyl alelových frekvencí:  $\text{Var}(p) = \sigma_p^2 = \sum_{i=1}^n w_i (p_i - \bar{p})^2 = \sum_{i=1}^n w_i p_i^2 - \bar{p}^2$   
 $= G_{AA}$

standardizovaný rozptyl:

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

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

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

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

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

měří odchylku od HW  
způsobenou rozdělením  
populace

~

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

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

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

měří odchylku od HW  
způsobenou děmovým  
inbreedingem

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

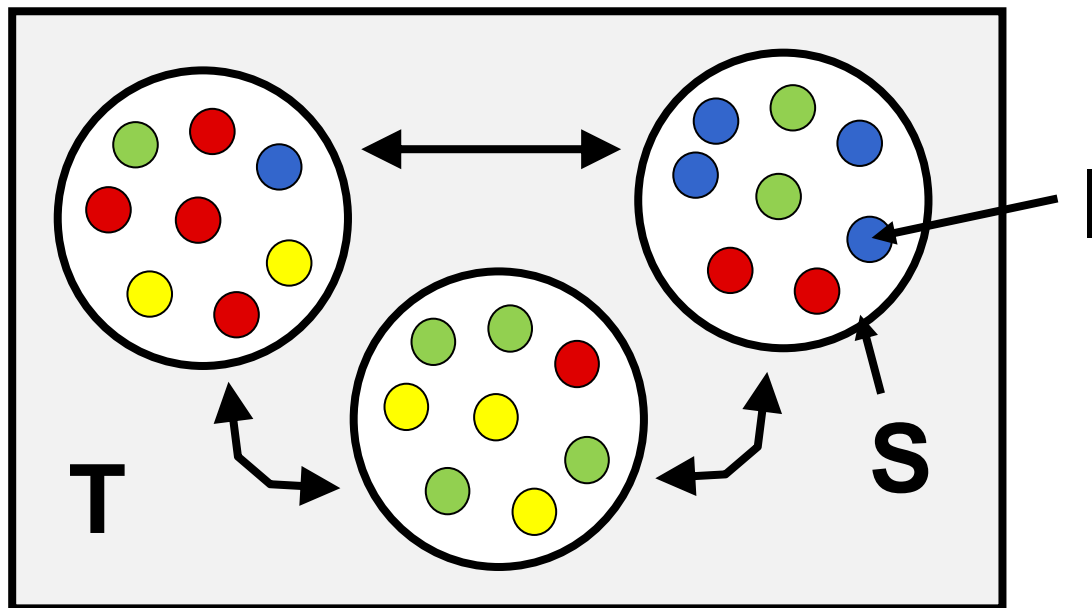
$\Rightarrow$  maximální rozptyl frekvencí:

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

$\Rightarrow F_{ST}$  lze chápat jako poměr skutečného rozptylu k teoretické maximální hodnotě:

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

## S. Wright: $F$ -statistika



S. Wright

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

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

$F_{IS}$  = korelace mezi splývajícími gametami

= vztah skutečné a očekávané heterozygotnosti v jediné nerozdělené populaci

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

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

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

průměrná očekávaná  
heterozygotnost v  
démoch

průměrná skutečná  
heterozygotnost v  
démoch

$F_{ST}$  = Pr., že 2 alely náhodně vybrané z celkové populace (T) a z lokální subpopulace (S) jsou stejné (*identity by state*, IBS):

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

= 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:

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



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

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

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

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

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

## $F_{ST}$ a koalescence .... příště!

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 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 děmu?

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

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)

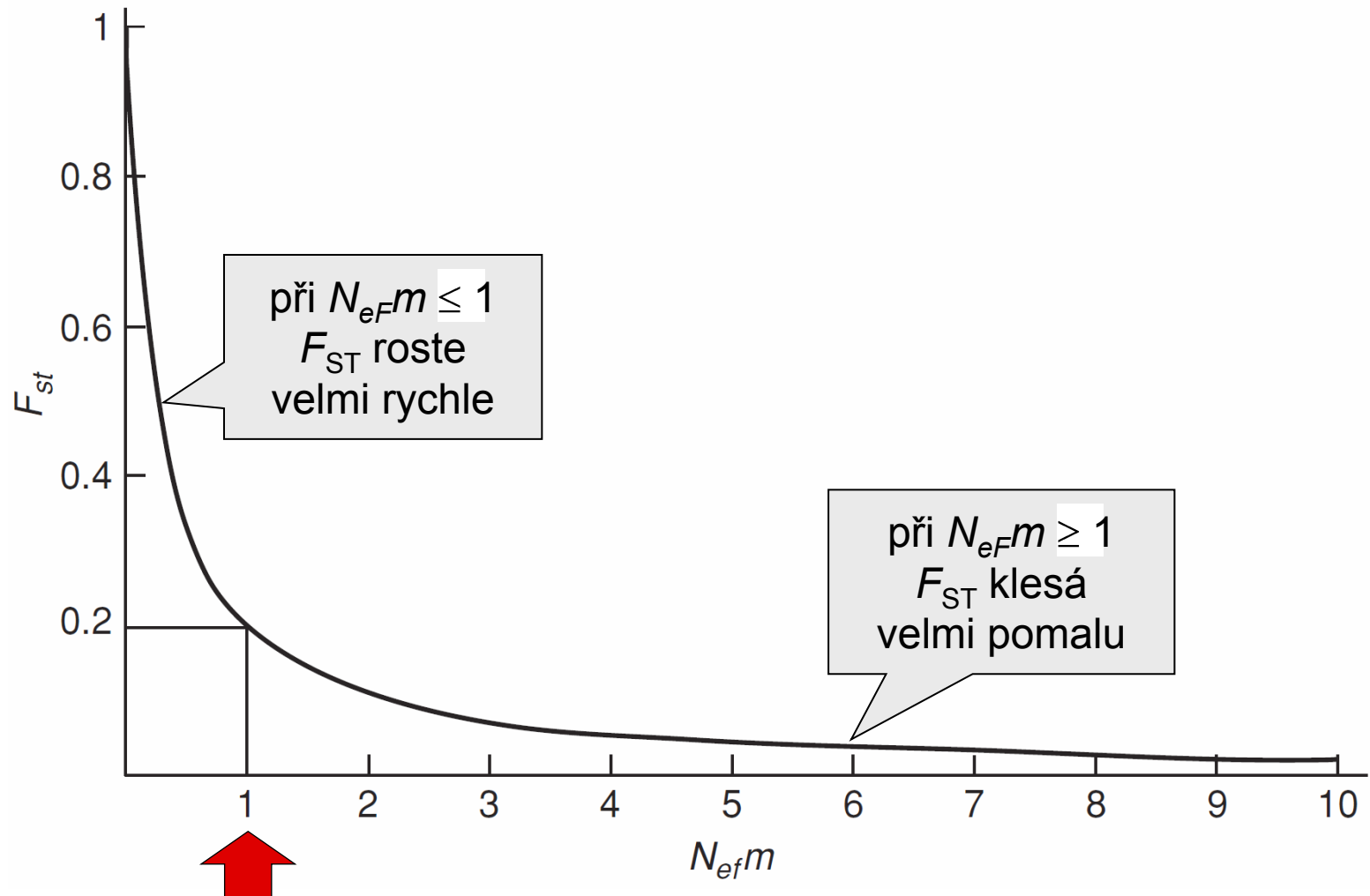
když  $1/N_e$  roste (větší drift)  $\Rightarrow F_{ST}$  roste (menší variabilita v démech a větší rozdíly mezi nimi)

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

\*) tj. nezávisí na skutečném počtu migrantů, ale na lokální efektivní velikosti pop.

# Vztah počtu migrantů a míry izolace subpopulací:



Míra homogenizace subpopulací nezávisí na migrační frekvenci, ale na 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.  $m = 0,01 \rightarrow N_e = 10^9 \Rightarrow F_{ST} \approx 0$ ;  $N_e = 100 \Rightarrow F_{ST} = 0,2$

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

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



$\Rightarrow$  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 démů 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  $n$  démů 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 počet jedinců  $N$ !!

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 ochránářskou 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

## 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

2–4 hierarchické úrovně

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



$$\Phi_{SC} = \frac{\sigma_b^2}{\sigma_b^2 + \sigma_c^2} \quad \Phi_{CT} = \frac{\sigma_a^2}{\sigma_T^2} \quad \Phi_{ST} = \frac{\sigma_a^2 + \sigma_b^2}{\sigma_T^2}$$

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

## Populační struktura a různé typy dědičnosti:

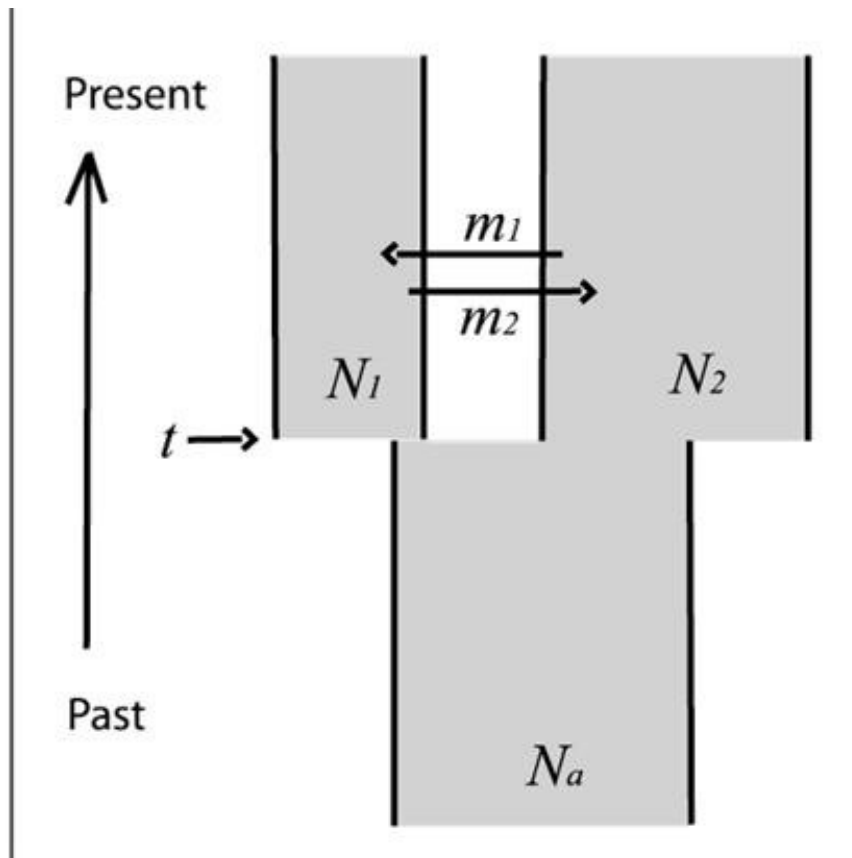
haploidní systém:  $F_{ST} = \frac{1}{1 + 2N_e V m}$

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

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

## 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:

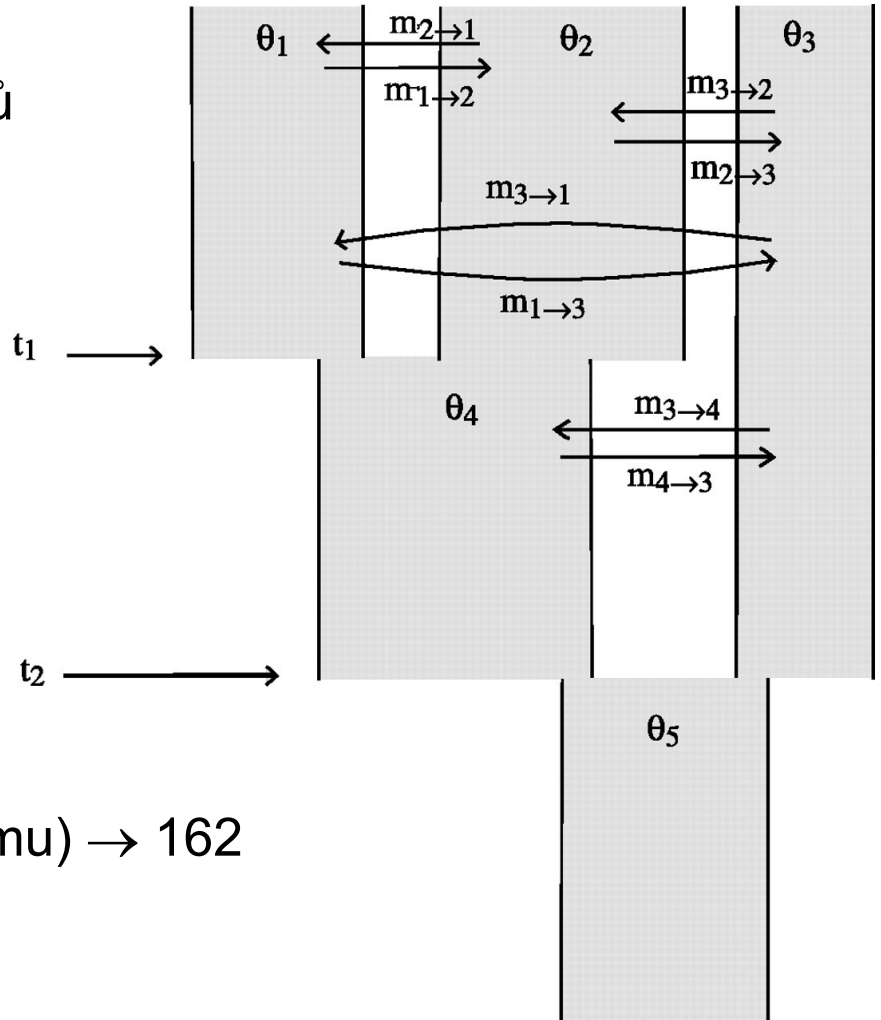
$$k(k - 1) + \sum_{i=2}^{k-1} 2(k - i) = 2(k - 1)^2$$

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!