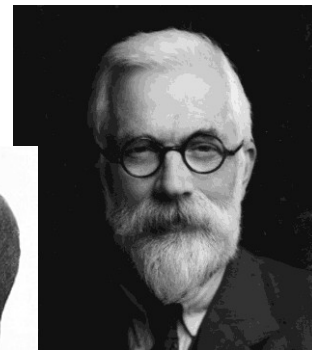
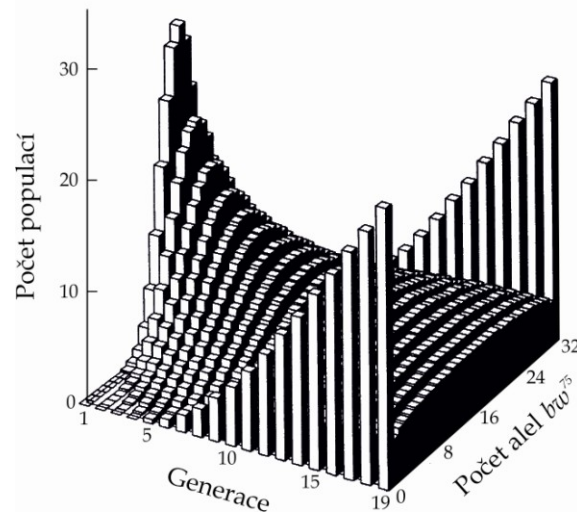
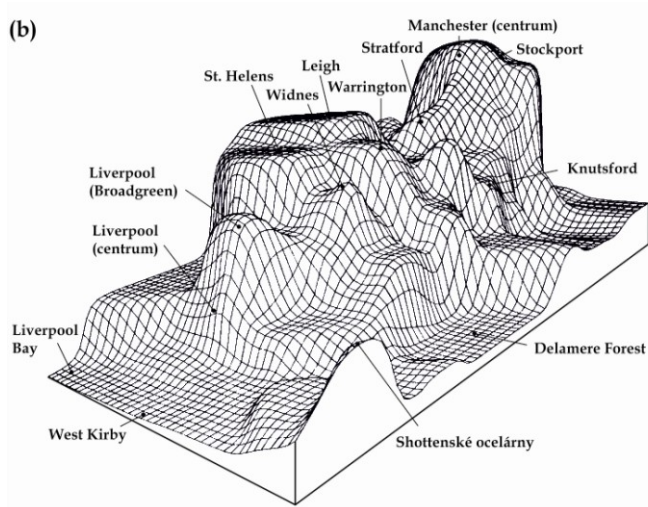


MECHANISMY MIKROEVOLUCE



**HARDYHO-
WEINBERGŮV
ZÁKON**

Mikroevoluce = vznik a osud genetické variability na druhové a nižší úrovni → děje a mechanismy v populacích

Evoluce = genetická změna populací v čase a prostoru

Evoluce = změna frekvence alel v populacích

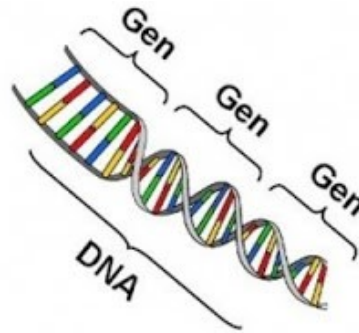
Populační genetik

idealizované predikce založené
na obecných principech (např.
mendelovská dědičnost)

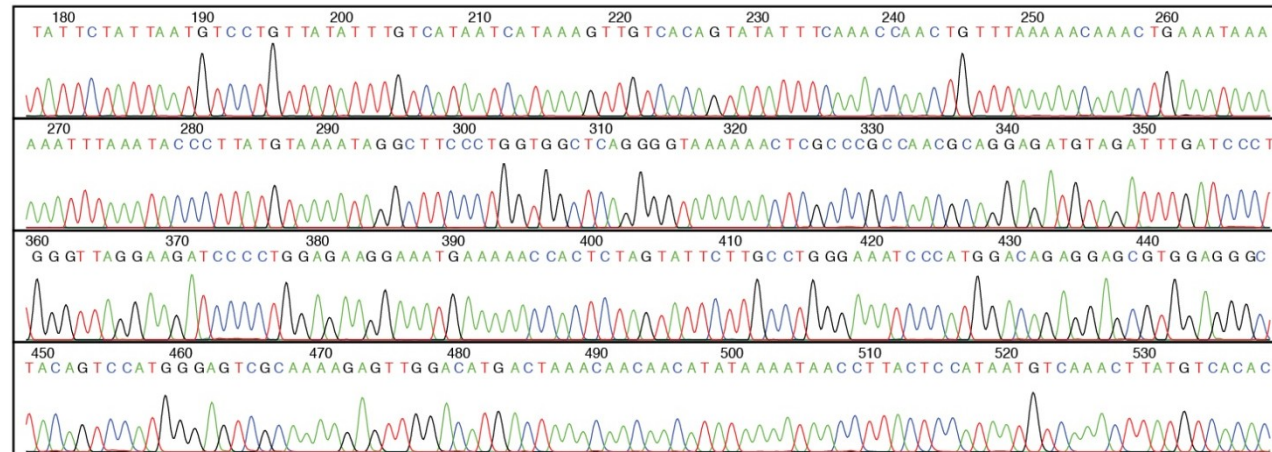
vysvětlení empirických výsledků
(mikroevoluční mechanismy)

3 základní premisy populační genetiky:

1. DNA se replikuje → dědičnost
2. DNA mutuje a rekombinuje → proměnlivost
3. Fenotypy vznikají interakcí DNA a prostředí → selekce



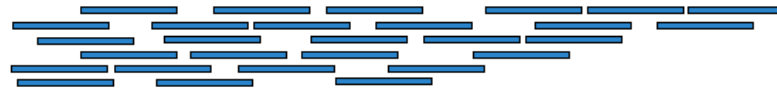
Gen, alela, lokus...



Multiple Copies of a Genome



Reads



High Coverage

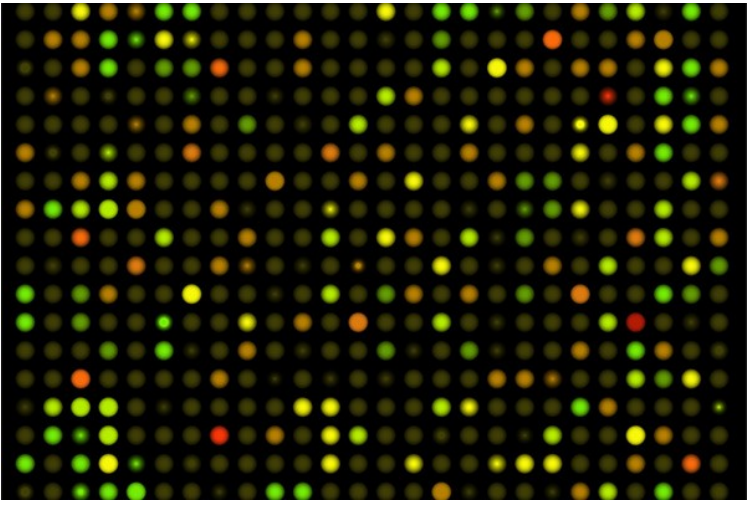
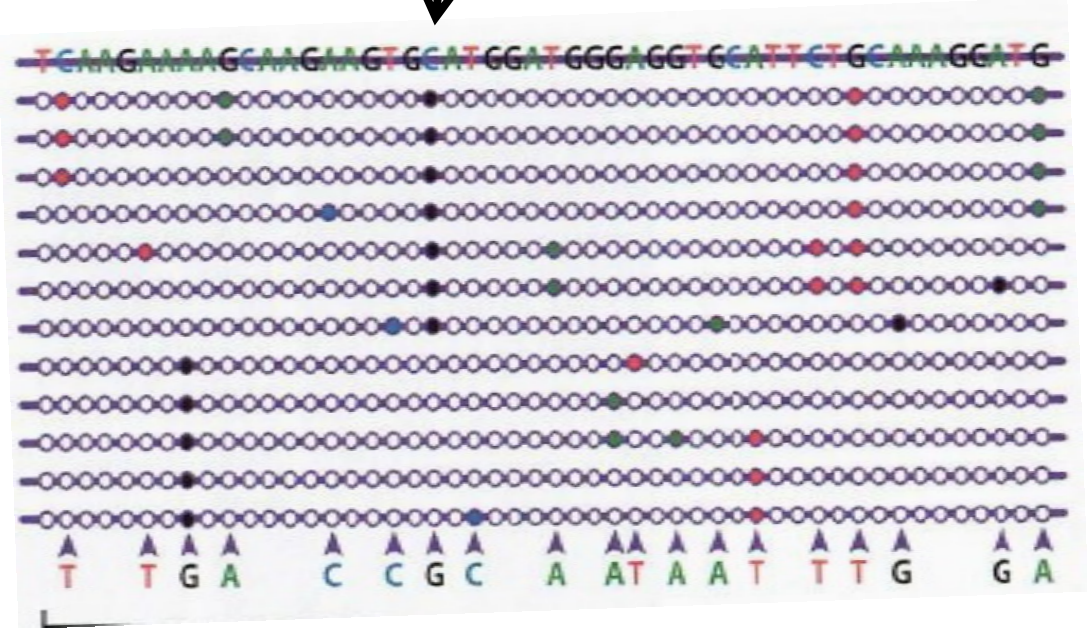
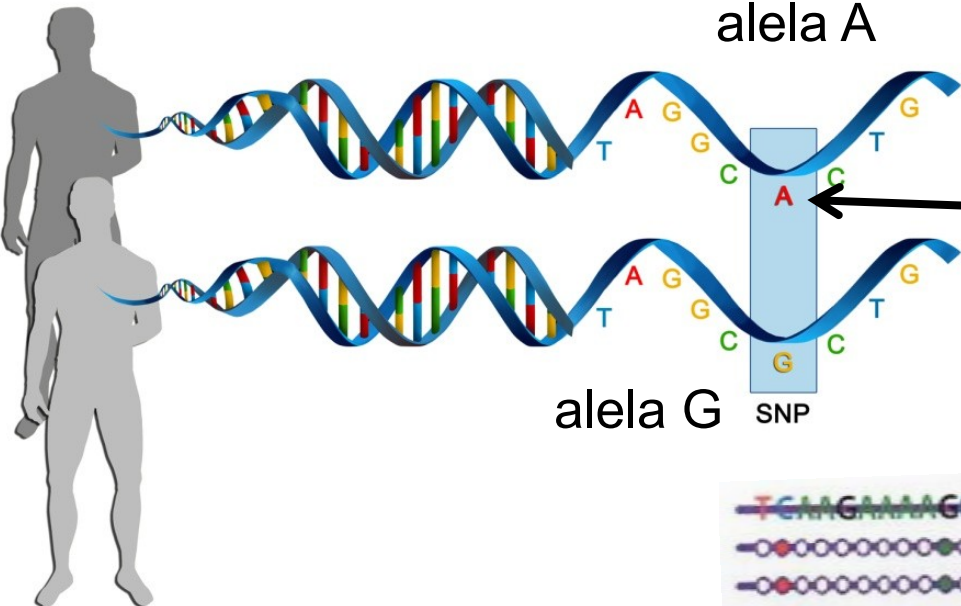
Low Coverage



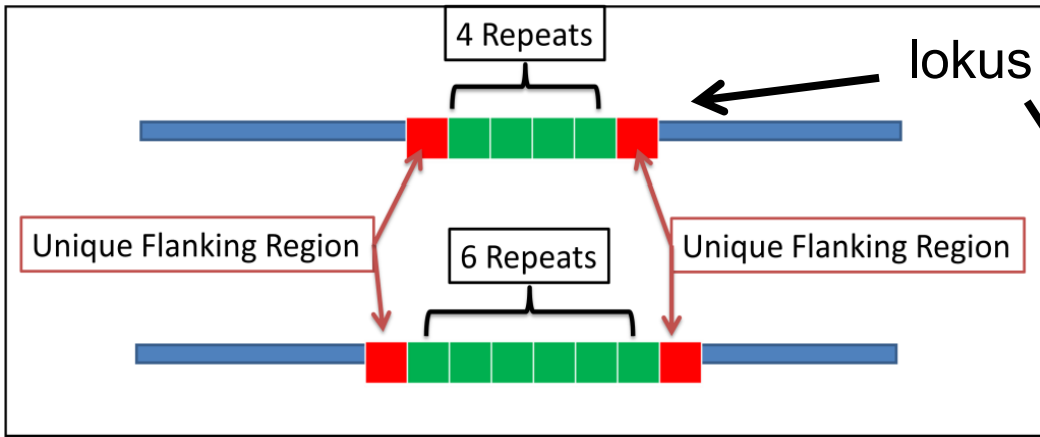
Consensus Sequence



Gen, alela, lokus...

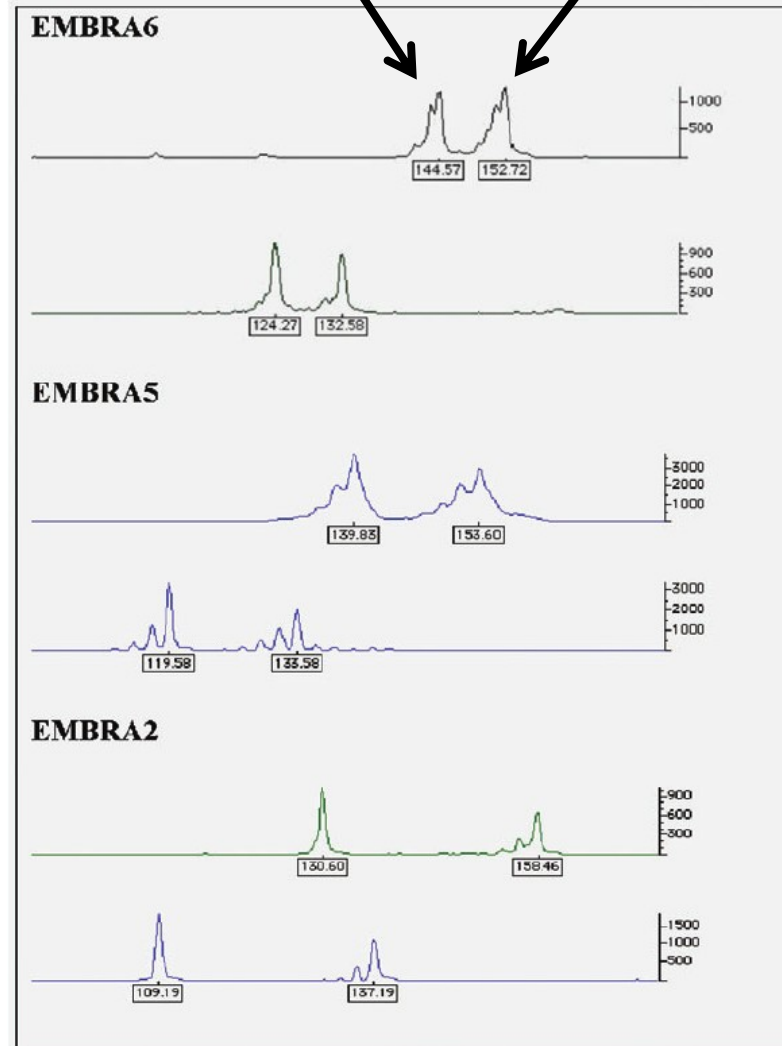


lokus



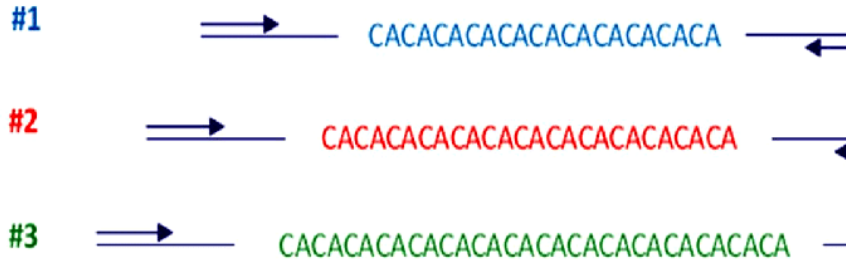
Genotyp 144/152

alela 144 alela 152

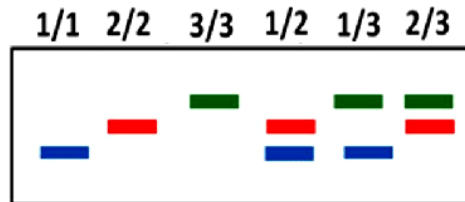


Alleles

(CA)_n Dinucleotide repeat



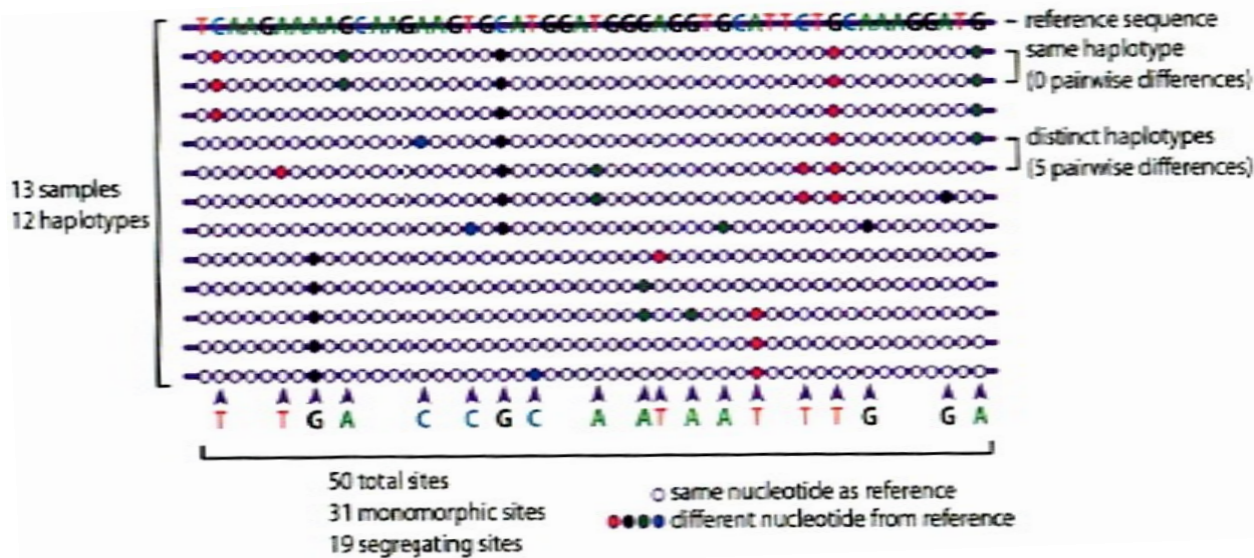
Genotypes



- Forward primer
- ← Reverse primer
- Flanking sequence

haplotyp (**haploidní genotyp**) = kombinace alel na různých částech sekvence DNA, které jsou přenášeny společně

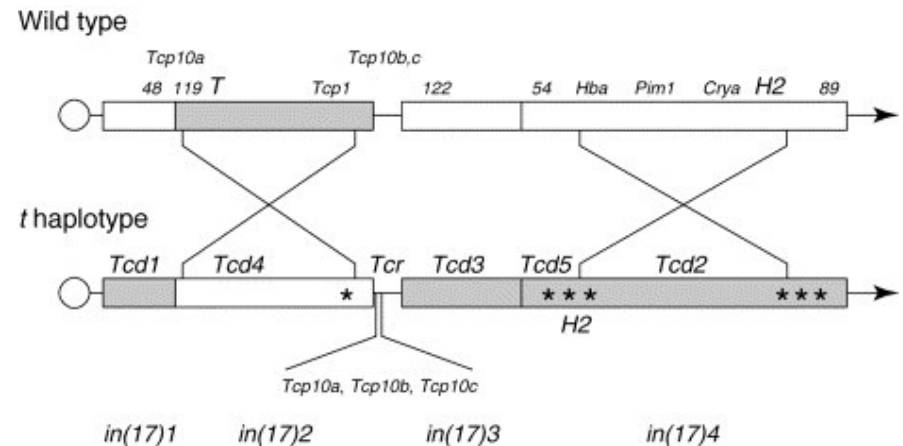
např. mtDNA, chr. Y



Haplotype	alela									lokus
	-11379	-11365	-4034	-3964	45	276	349	112	2019	
1	G	C	A	A	T	G	A	G	Ins	
2	G	C	A	A	T	T	A	A	Del	
3	G	C	A	A	G	G	G	A	Del	
4	G	C	C	A	T	G	A	A	Ins	
5	G	G	A	A	T	G	A	G	Ins	
6	A	C	C	A	T	T	A	A	Del	

haplotyp

t haplotyp:



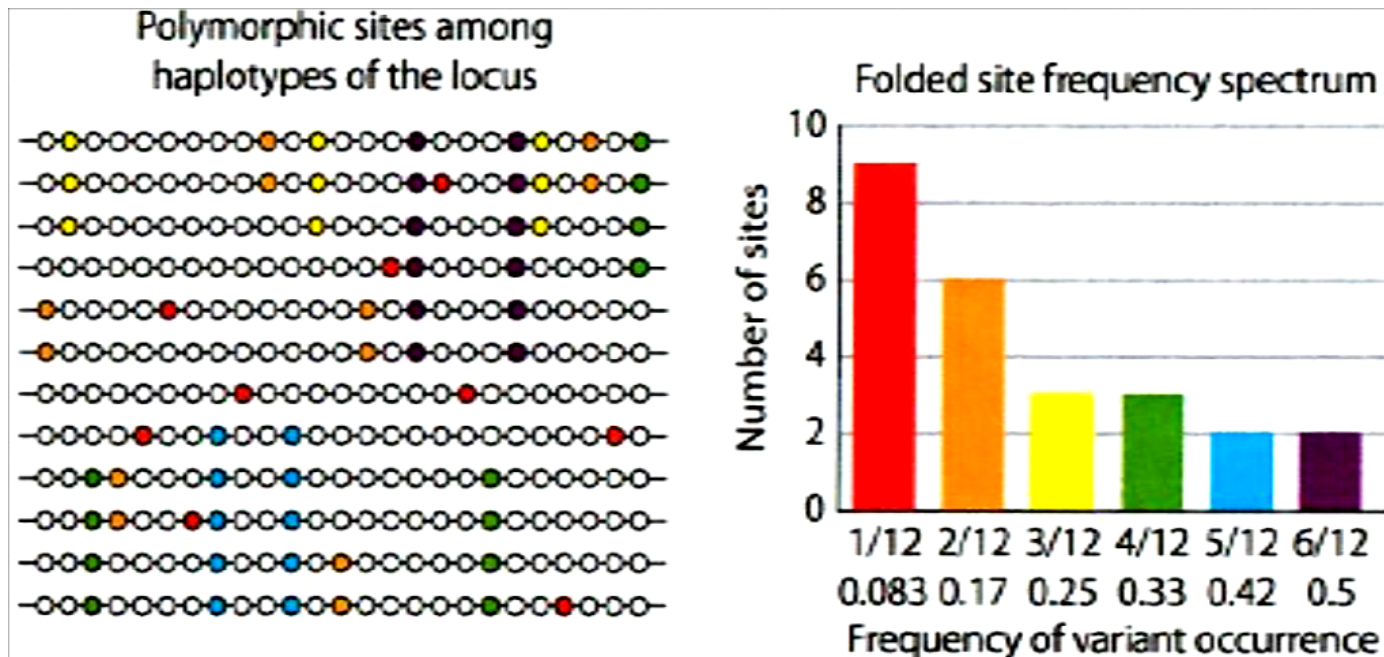
Spektrum alelových frekvencí (*site frequency spectrum, allele frequency spectrum*)

SNP

folded = jen vzácnější alely, považované za odvozené

unfolded = všechny alely, rozlišení ancestrálních a odvozených

umožňuje srovnání neutrální predikce a různých typů selekce



Genetická architektura:

počet lokusů (popř. jejich pozice v genomu)

počet alel na lokus

frekvence mutací (mutační rychlost)

způsob a pravidla dědičnosti

jeden lokus – např. autozom, X, Y, mtDNA

více lokusů – smíšený zp. dědičnosti, kombinace alel, rekombinace

Populace

T. Dobzhansky, E. Mayr:

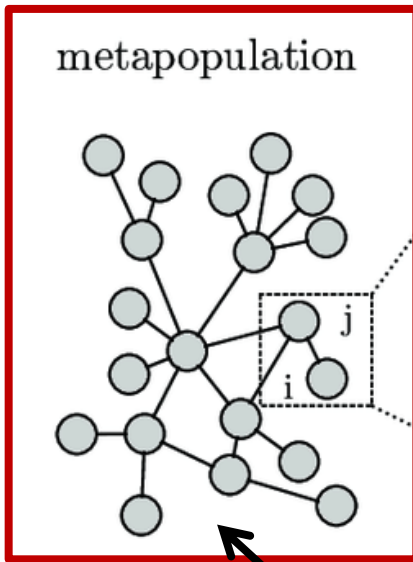
populace jako společný **genofond** (*gene pool*)

1. soubor sdílených genů/alel
2. soubor potenciálních gamet produkovaných všemi příslušníky populace

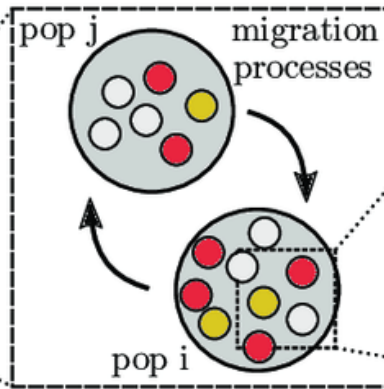
Pozor: ne vždy jsou oba koncepty totožné!



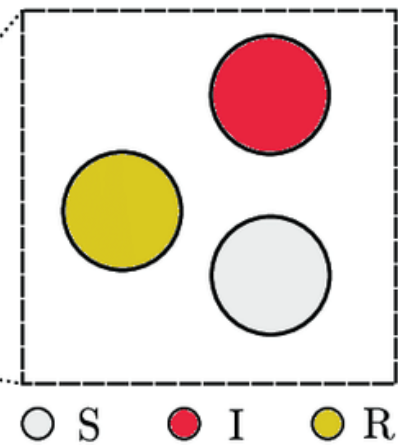
lokální populace (subpopulace, děmy)



subpopulation



individuals

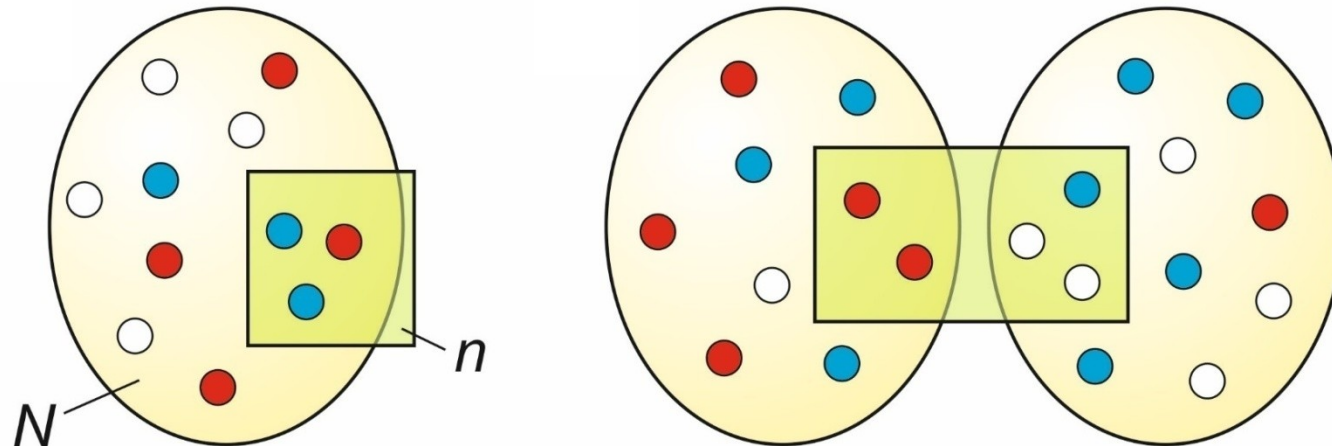


globální populace, metapopulace

populace přírodní, experimentální, zemědělské, modelové

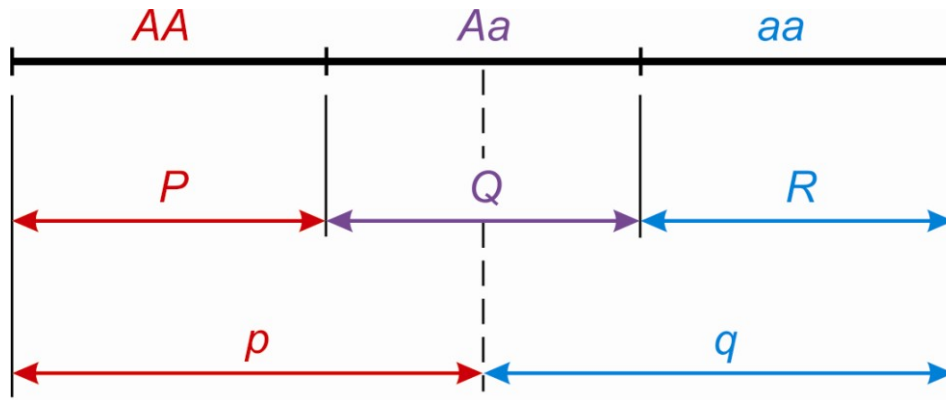
Kromě genů/gamet lokální populace sdílejí i
system páření (*system of mating*)

**Důsledky evoluce v subpopulacích a metapopulacích
můžou být zcela opačné!**



Pozor: populační vzorek \neq populace!

Genotypové a alelové frekvence

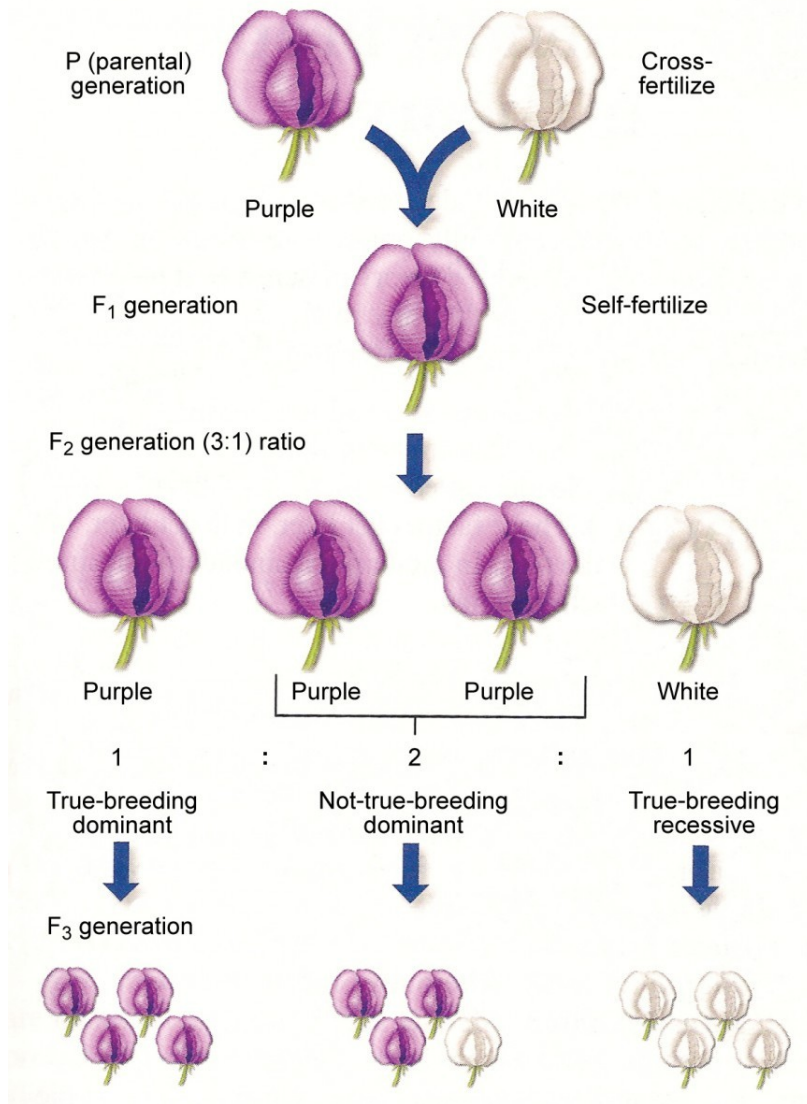


Relativní četnosti = frekvence: genotypové: $P (f_{AA})$, $Q (f_{Aa})$, $R (f_{aa})$
 alelové (genové): $p (A)$, $q (a)$

$$f_{AA} + f_{Aa} + f_{aa} = 1$$

$$p + q = 1$$

Genotyp	A_1A_1	A_1A_2	A_2A_2	Celkem
Četnost	n_1	n_2	n_3	N
Frekvence	$f_{AA} = n_1/N$	$f_{Aa} = n_2/N$	$f_{aa} = n_3/N$	
	$p = (2n_1 + n_2)/2N$		$q = (n_2 + 2n_3)/2N$	



Proč v přírodě nepozorujeme mendelovské poměry 3:1?

Proč v přírodě nepozorujeme mendelovské poměry 3:1?

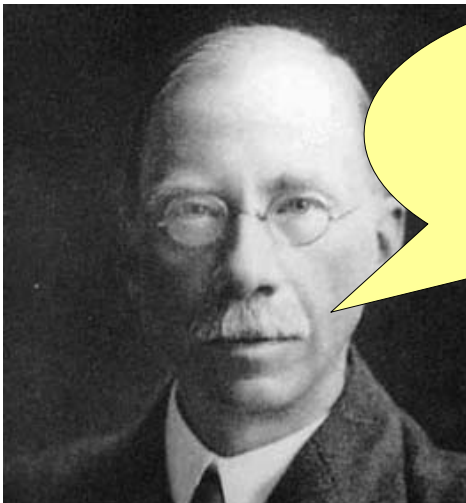


brachydaktylie



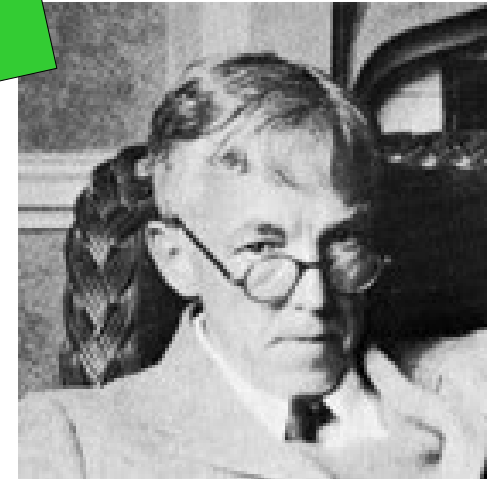
brachydaktylie =
dominantní
mendelovský znak

Reginald C. Punnett



Měli bychom
pozorovat 3× víc lidí
postižených než
nepostižených!

G. Udny Yule



G. H. Hardy

HARDYHO-WEINBERGŮV ZÁKON

1908



Godfrey Harold Hardy
(1877-1947)



Wilhelm Weinberg
(1862-1937)

Hardyho-Weinbergův model

Mechanismy produkce gamet (genetická architektura):

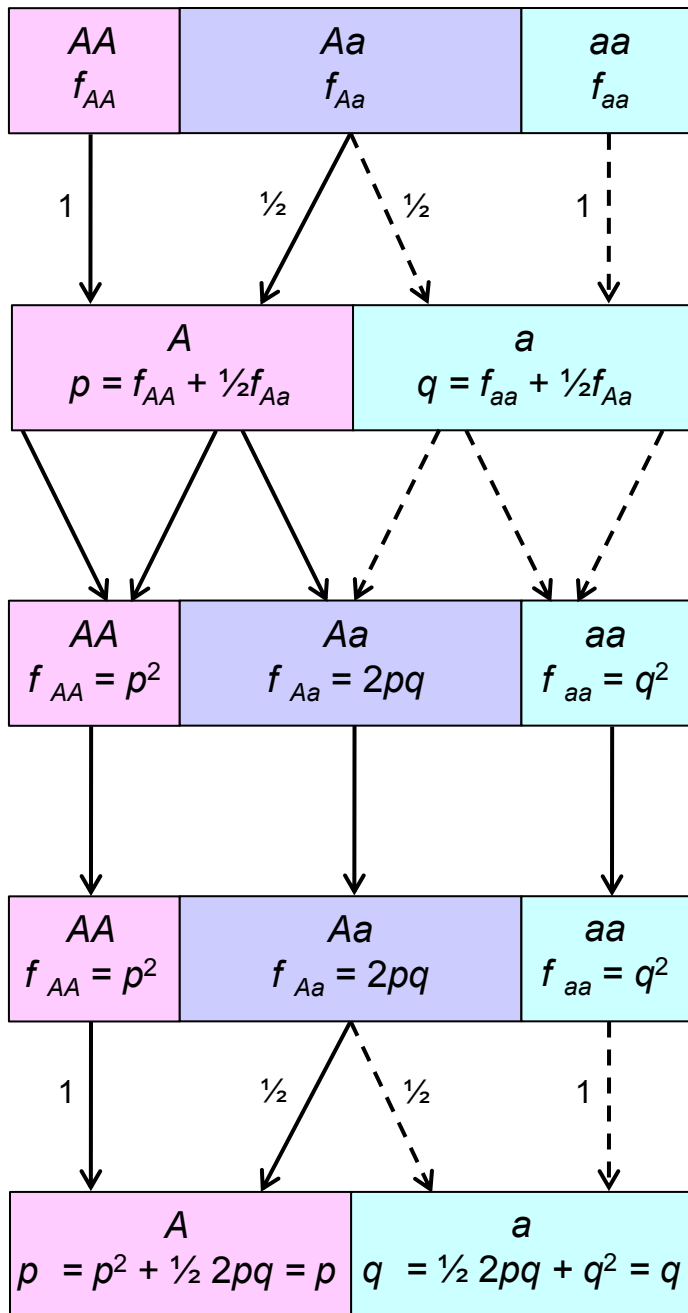
- diploidní organismy
- jeden autozomální lokus
- 2 alely
- žádné mutace
- mendelovská segregace (1:1)

Mechanismy spojování gamet (populační struktura):

- system páření → náhodné oplození (panmixie)
- velikost populace → nekonečná
- genetická výměna → žádná (absence toku genů)
- věková struktura → žádná (diskrétní generace)

Mechanismy vývoje fenotypů:

- všechny genotypy mají stejné fenotypy z hlediska schopnosti replikace DNA (absence selekce)



Adultní populace

mechanismus produkce gamet (z. o segregaci)

Genofond (populace gamet)

mechanismus spojení gamet (náhodné oplození)

Populace zygot

mechanismus vývoje fenotypu (žádný účinek na viabilitu, repr. úspěšnost nebo fertilitu)

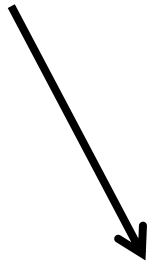
Adultní populace další generace

mechanismus produkce gamet (z. o segregaci)

Genofond příští generace

Frekvence gamety j v genofondu:

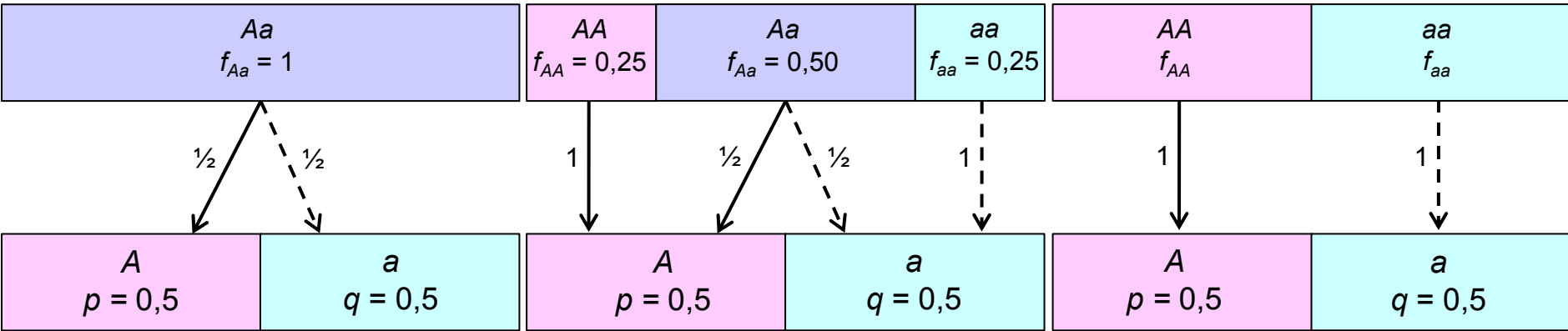
$$g_j = \sum_{\text{genotypy}} \text{Pr.}(\text{produkce gamety } j \text{ genotypem } k) \times \text{frekvence genotypu } k$$



Pr. produkce gamety A z $AA = 1$... z $Aa = \frac{1}{2}$... z $aa = 0$

Pr. produkce gamety a z $AA = 0$... z $Aa = \frac{1}{2}$... z $aa = 1$

známe-li zákony dědičnosti, z
frekvencí genotypů můžeme
vypočítat frekvence alel







různé adultní populace mohou produkovat stejný genofond gamet

⇒ genotypové frekvence *nejsou* jedinečně určeny frekvencemi gamet*)

⇒ musíme přidat další informaci → populační struktura (panmixie)!

*) matematicky: ze 2 proměnných nelze odvodit 3 proměnné!

Hardyho odvození:

		Samčí gamety	
		Alela	Frekvence
Samičí gamety	Alela	A	a
	Frekv.	p	q
A	p	 AA $p \times p = p^2$	 Aa $p \times q = pq$
	q	 aA $q \times p = qp$	 aa $q \times q = q^2$

Frekvence
v zygotech:

$$AA: G_{AA}' = p^2$$

$$Aa: G_{Aa}' = pq + qp = 2pq$$

$$aa: G_{aa}' = q^2$$

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

Pop. gamet efektivně nekonečná \Rightarrow výběr 1. gamety neovlivní výběr 2. gamety

V nekonečné pop. pravděpodobnost jevu = frekvence tohoto jevu

Weinbergovo odvození:

při náhodném párování Pr. 2 genotypů v páření = součin jejich frekvencí

Párování	Frekvence páru	Pravděpodobnost zygot		
		AA	Aa	aa
AA x AA	$f_{AA} \times f_{AA} = f_{AA}^2$	1	0	0
AA x Aa	$f_{AA} \times f_{Aa} = f_{AA}f_{Aa}$	1/2	1/2	0
Aa x AA	$f_{Aa} \times f_{AA} = f_{Aa}f_{AA}$	1/2	1/2	0
AA x aa	$f_{AA} \times f_{aa} = f_{AA}f_{aa}$	0	1	0
aa x AA	$f_{aa} \times f_{AA} = f_{aa}f_{AA}$	0	1	0
Aa x Aa	$f_{Aa} \times f_{Aa} = f_{Aa}^2$	1/4	1/2	1/4
Aa x aa	$f_{Aa} \times f_{aa} = f_{Aa}f_{aa}$	0	1/2	1/2
aa x Aa	$f_{aa} \times f_{Aa} = f_{aa}f_{Aa}$	0	1/2	1/2
aa x aa	$f_{aa} \times f_{aa} = f_{aa}^2$	0	0	1

$$f'_{AA} = f_{AA}^2 + 1/2f_{AA}f_{Aa} + 1/2f_{Aa}f_{AA} + 1/4f_{Aa}^2 = [f_{AA} + 1/2f_{Aa}]^2 = p^2$$

$$(a + b)^2 = a^2 + 2ab + b^2$$

$$f'_{Aa} = 1/2[2f_{AA}f_{Aa}] + 2f_{AA}f_{aa} + 1/2f_{Aa}^2 + 1/2[f_{Aa}f_{aa}] = 2[f_{AA} + 1/2f_{Aa}][f_{aa} + 1/2f_{Aa}] = 2pq$$

$$f'_{aa} = 1/4f_{Aa}^2 + 1/2[2f_{Aa}f_{aa}] + f_{aa}^2 = [f_{aa} + 1/2f_{Aa}]^2 = q^2$$

HARDYHO-WEINBERGŮV ZÁKON

1. Frekvence alel z generace na generaci stálé
= Hardyho-Weinbergova rovnováha

2. HW rovnováhy dosaženo už po 1 generaci náhodného křížení

Zobecnění:

geny vázané na X:

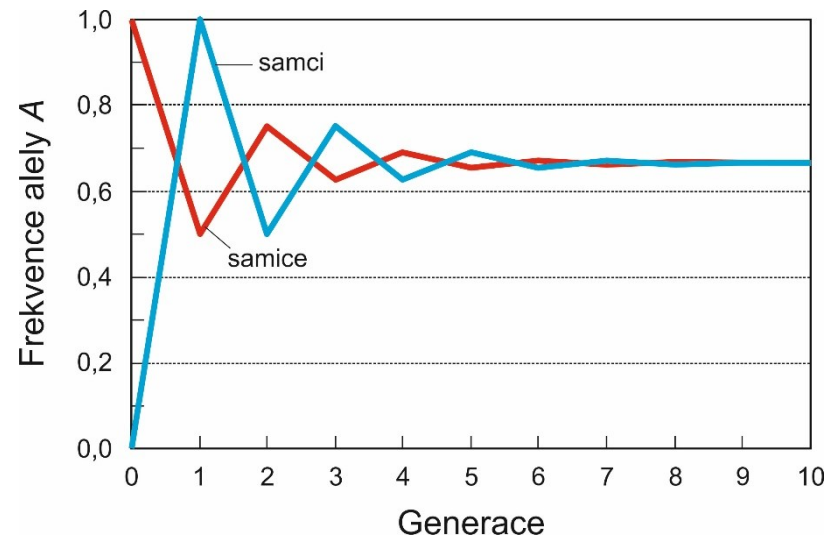
$$\text{samice: } p^2 + 2pq + q^2$$

$$\text{samci: } p + q$$

více alel:

$$3 \text{ alely: } p^2 + q^2 + r^2 + 2pq + 2pr + 2qr$$

$$\text{obecně } p_i^2 + 2p_{ij}$$



HARDYHO-WEINBERGŮV ZÁKON

Punnett



brachydaktylie

$p^2 + 2pq : q^2$
není důvod pro poměr
postižených a
nepostižených 3:1

Pokud letální alela *a* *recesivní* → 2 fenotypové kategorie:

AA + *Aa* → bez projevu choroby

aa → choroba se projeví

⇒ nemůžeme určit, kolik jedinců je nositelem choroby

ale předpokládáme-li HW, pak frekvence nemocných (můžeme spočítat) je q^2
a proto

$$\hat{q} = \sqrt{f_{aa}}$$

⇒ frekvence nositelů choroby =

$$2(1 - \hat{q})\hat{q}$$

Př.: fenylketonurie (PKU)

frekvence výskytu PKU v České republice je 1 : 9000, tj. zhruba 0,00011.

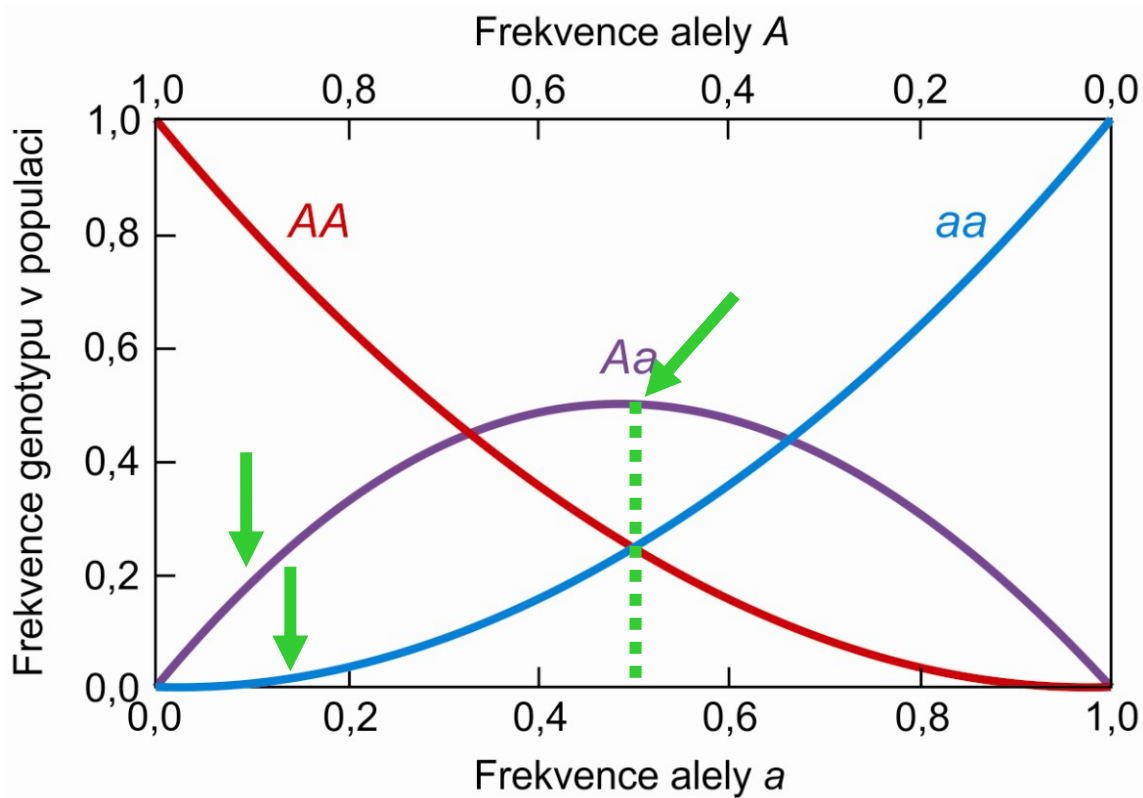
Jaké procento lidí jsou tzv. přenašeči?

Vyjdeme-li z HW zákona, je frekvence homozygotů $f_{aa} = q^2$, frekvence alely a je tudíž $q = \sqrt{0,00011} = 0,01049$.

$$**2pq = 2(1 - q)q = 2 \times (1 - 0,01049) \times 0,01049 = 0,02076.**$$

To znamená, že v České republice se vyskytují zhruba 2 % přenašečů fenylketonurie. K roku 2011 10 562 214 občanů \Rightarrow u nás přibližně 220 tisíc přenašečů PKU.

Frekvence vzácných alel:



heterozygoti nejfrekventovanější při $p = q = 0,5$

f_{Aa} se snižuje rychlostí $2pq$

f_{aa} rychlostí $q^2 \Rightarrow$ zvyšování $f_{Aa}/f_{aa} \rightarrow$ vzácná alela „schována“
v heterozygotním stavu

Příčiny neplatnosti H-W rovnováhy:

Metodické příčiny:

nulové alely (*allelic dropout*)

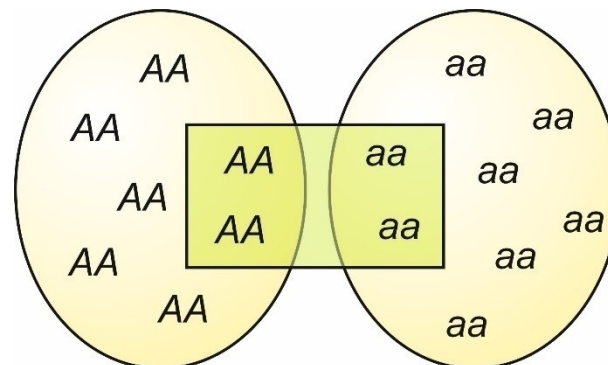
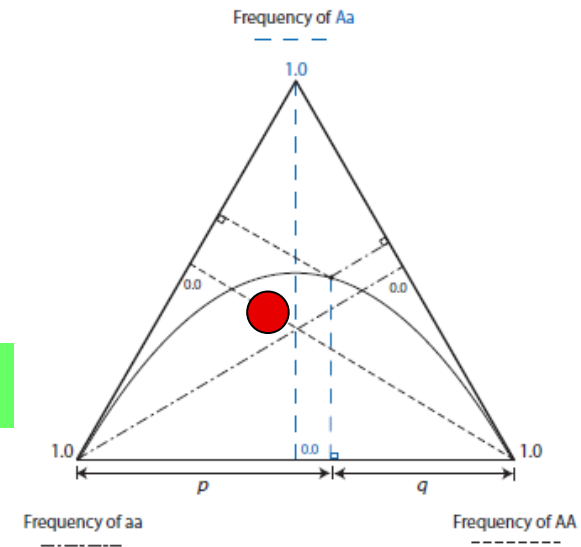
Neplatnost některého z předpokladů H-W modelu:

Snížení heterozygotnosti:

selektce proti heterozygotům

nenáhodné páření (inbreeding, asortativní páření)

strukturovanost populace (rozdílné frekvence alel, Wahlundův efekt)



Příčiny neplatnosti H-W rovnováhy:

Metodické příčiny:

nulové alely (*allelic dropout*)

Neplatnost některého z předpokladů H-W modelu:

Snížení heterozygotnosti:

selekce proti heterozygotům

nenáhodné páření (inbreeding, asortativní páření)

strukturovanost populace (rozdílné frekvence alel, Wahlundův efekt)

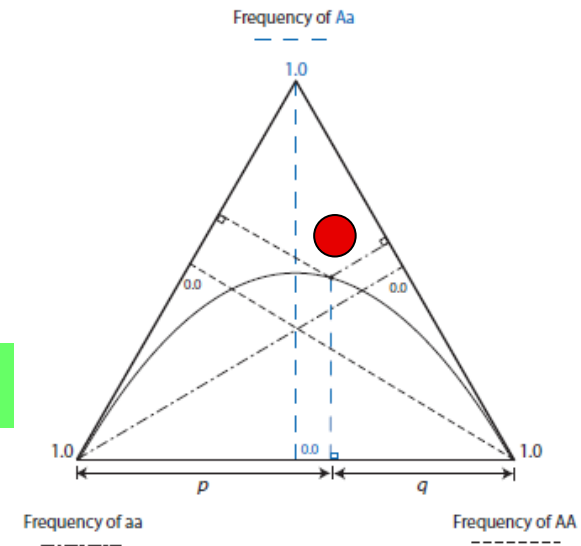
Zvýšení heterozygotnosti:

selekce podporující heterozygoty

nenáhodné páření (outbreeding, negativní asortativní páření)

migrace

mutace



H-W NA VÍCE LOKUSECH

blížkost lokusů = **vazba**

frekvence rekombinací r

místo 3 genotypových kategorií → 10

1 lokus → jedinec může předat jen gamety získané od rodičů

více lokusů → může předat i jiné, např:

rodiče $AB \times ab \rightarrow AB/ab$ gamety AB, ab s $\text{Pr.} = \frac{1}{2}(1 - r)$

ale i gamety Ab, aB (s $\text{Pr.} \frac{1}{2}r$), které se u rodičů nevyskytují

⇒ 2 definice genofondu (1. = soubor genů sdílených v populaci,

2. populace všech potenciálních gamet)

nejsou v případě rekombinace ekvivalentní

Může dojít k evoluci?

nutno znát, jak
je genofond
definován!

H-W NA VÍCE LOKUSECH

Př.: frekvence gamety AB v počátečním genofondu (g_{AB}) a následující generaci (g'_{AB}):

$$g'_{AB} = g_{AB} - rD \quad \text{kde } D = g_{AB}g_{ab} - g_{Ab}g_{aB} = \text{vazebná nerovnováha}$$

... podobně i pro ostatní typy gamet

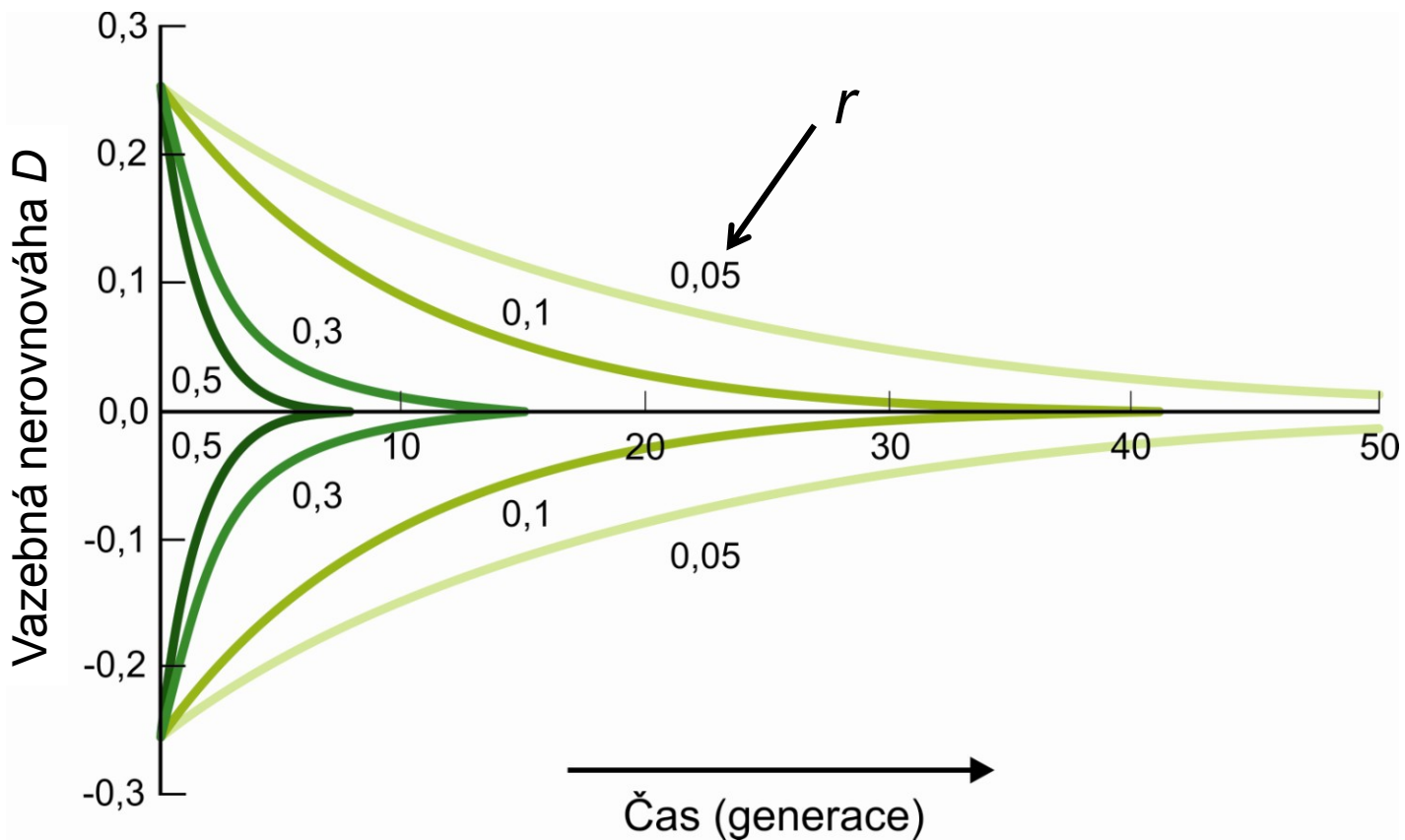
jestliže $r > 0$ (tj. dochází k rekombinaci) a $D \neq 0$ (existuje vazba)

pak platí $g_{xy} \neq g'_{xy} \Rightarrow$ **dochází k evoluci!**

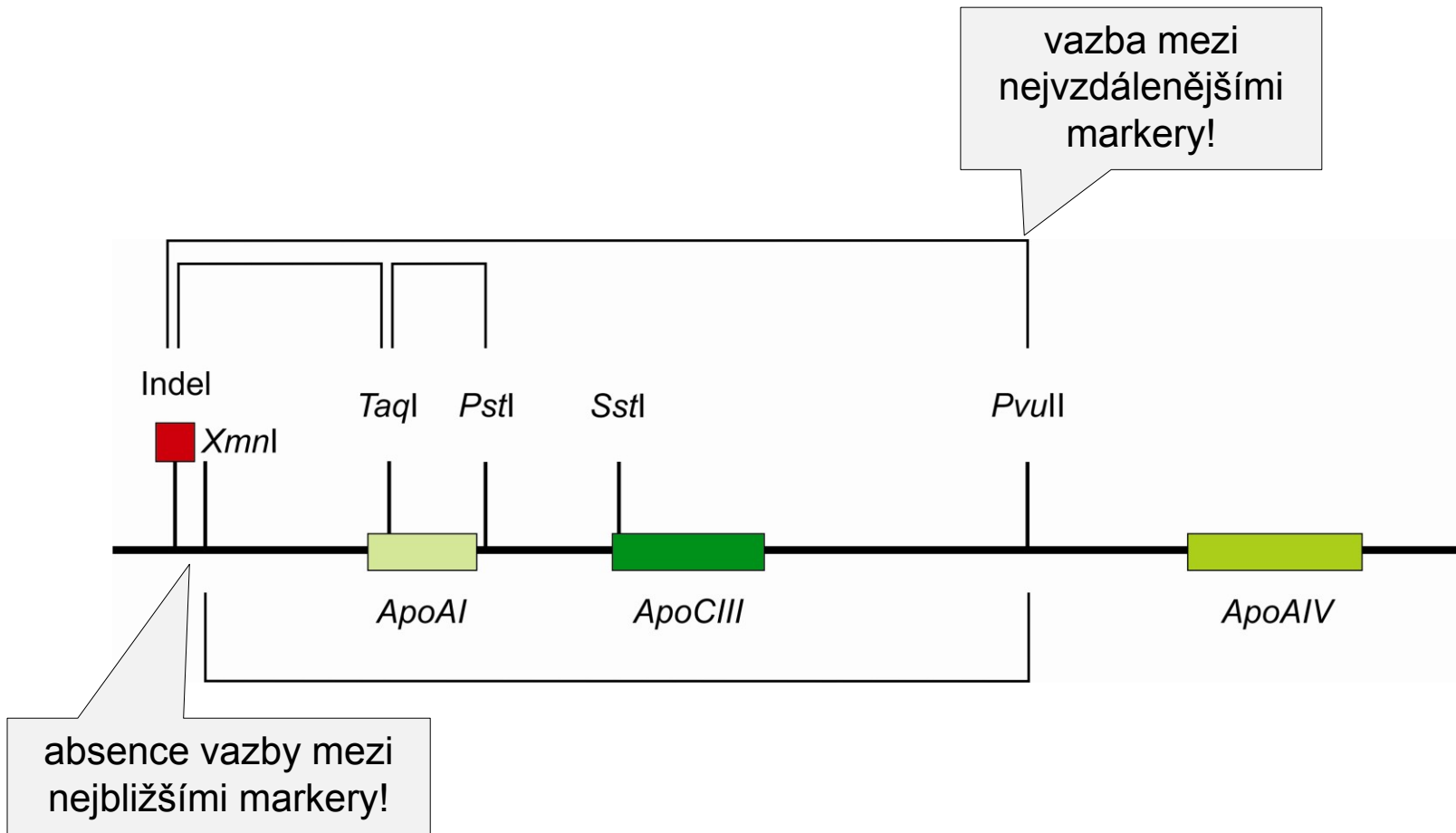
naopak při $r = 0$ multilokusový model \rightarrow jednolokusový (např. mtDNA)

k rekombinaci může dojít i *uvnitř* genu \sim jako mutace (tj. vznik nové alely)

Vztah D a rekombinace r :



Pozor! vazebná nerovnováha nemusí přesně odrazet fyzickou vzdálenost lokusů!!



Laurie et al. (2007): pokles D se vzdáleností u člověka a myši

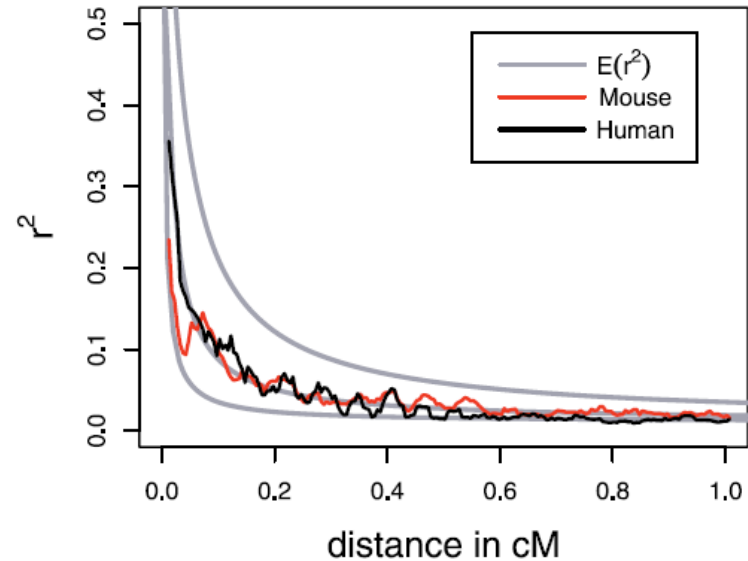
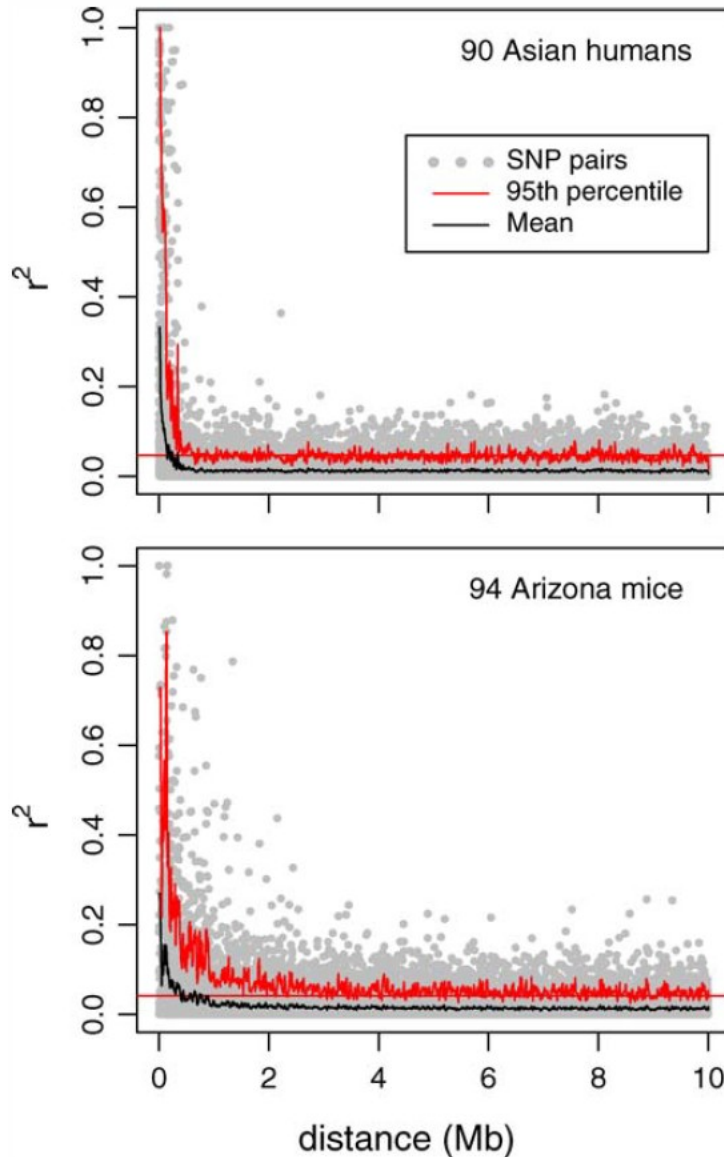


Figure 4. Mean LD in 94 Arizona Mice and 90 Asian Humans, Evaluated in a Sliding Window of 0.025 cM

The data are the same as in Figure 2, but physical distance has been converted to genetic distance using genome-wide averages of cM/Mb. The expected value of r^2 for a sample size of n diploids is calculated as $E(r^2) = (1 / (1 + 4N_e c)) + (1 / n)$, where N_e is effective population size and c is recombination rate (converted to cM for plotting by the Kosambi mapping function). In this case, $n = 92$ and $N_e = 1,000, 3,000,$ or $10,000$ (from top to bottom).

doi:10.1371/journal.pgen.0030144.g004

člověk: 1 cM ~ 1 Mb

myš: 1 cM ~ 2 Mb

Příčiny vazebné nerovnováhy:

vazebná nerovnováha
nemusí být mezi lokusy na
stejném chromozomu!

absence rekombinace (např. inverze)

nenáhodnost oplození

selekce

recentní mutace

vzorek směsí 2 druhů s různými frekvencemi

recentní splynutí 2 populací

omezená velikost populace (náhodné kolísání frekvencí alel)

u mnoha organismů crossing-over důležitý pro správný průběh meiózy
(aspoň 1 c-o na chromozom, jinak vznik aneuploidíí)

ženy s > c-o → > dětí

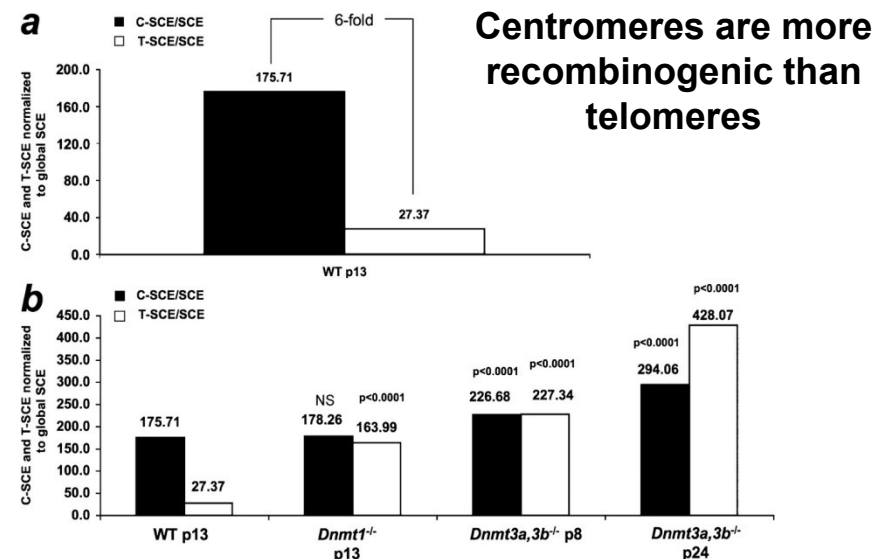
děti starších žen → > rekombinací

obecně častější u centromery, méně u telomer (neplatí pro všechny chromozomy,
druhové rozdíly – např. u cibule naopak)

malé chromozomy > frekvence rekombinací

rekombinační „hotspots“:

- u člověka ~25 000
- chybí u *Drosophila melanogaster* a *Caenorhabditis elegans*
- častý vznik a zánik
- zánik 1 místa často kompenzován zvýšenou aktivitou sousedního místa

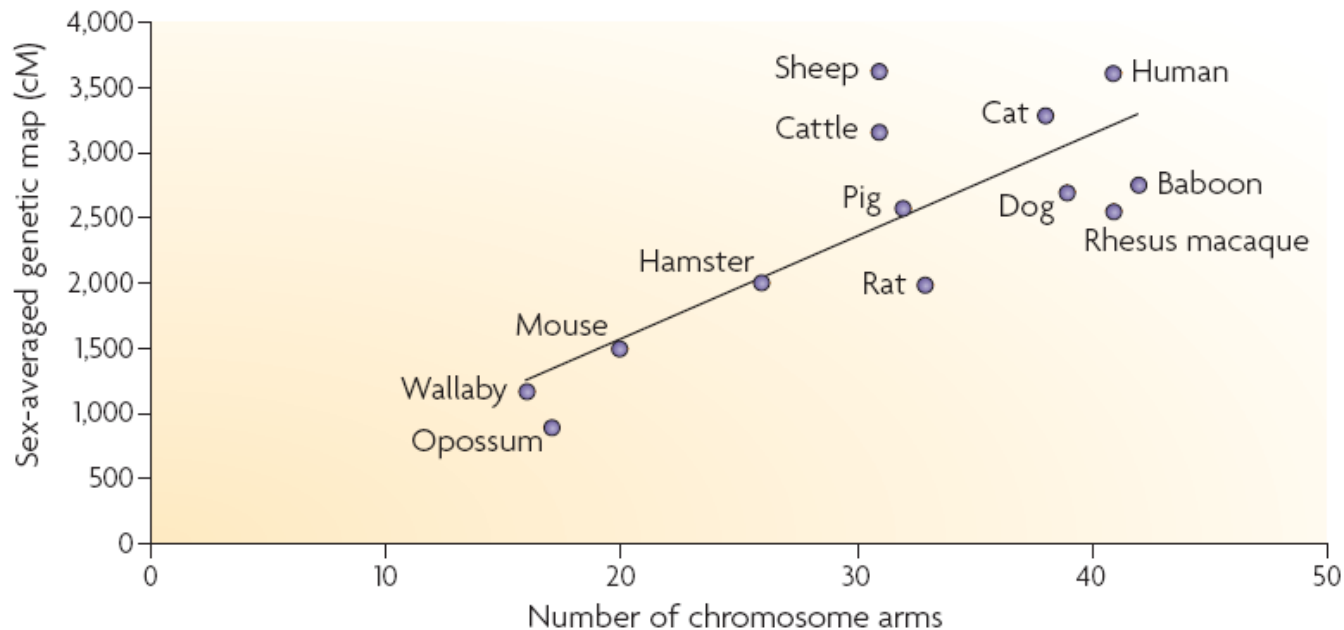


rozdíly v míře rekombinace mezi pohlavími:

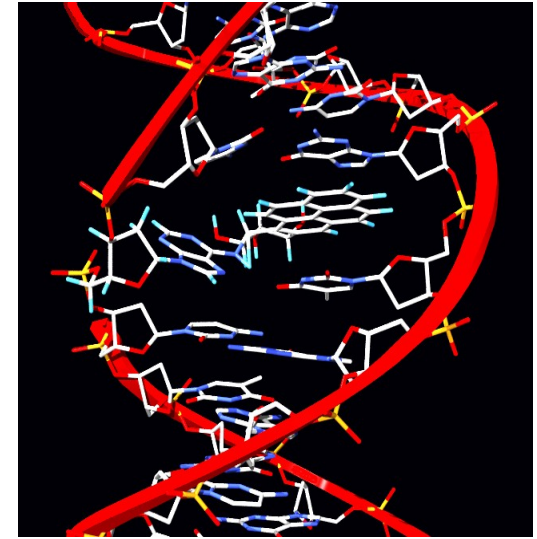
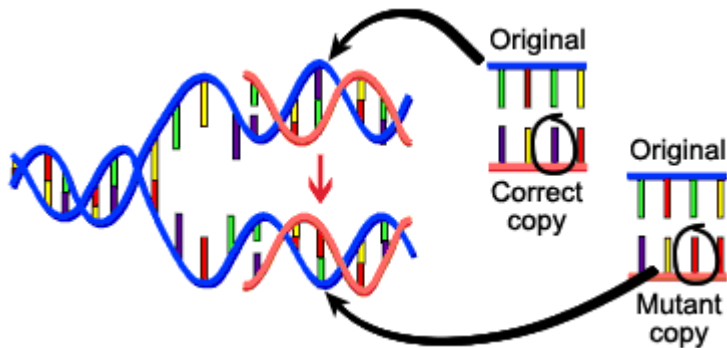
- **Haldaneovo-Huxleyovo pravidlo**: pokud jedno pohlaví nerekombinuje, jde o pohlaví heterogametické
- pokud rekombinují obě pohlaví, u samic většinou více rekombinací (člověk 1,7x, myš 1,3x)

rozdíly mezi druhy:

- druhy s více malými chromozomy → více rekombinací než druhy s menším počtem velkých chromozomů
- korelace s počtem ramen: více rekombinací v karyotypech s velkým množstvím chrom. ramen (aspoň 1 c-o/rameno, aby nedocházelo k aneuploidii?)



MUTACE



podle škodlivosti/prospěšnosti účinku:

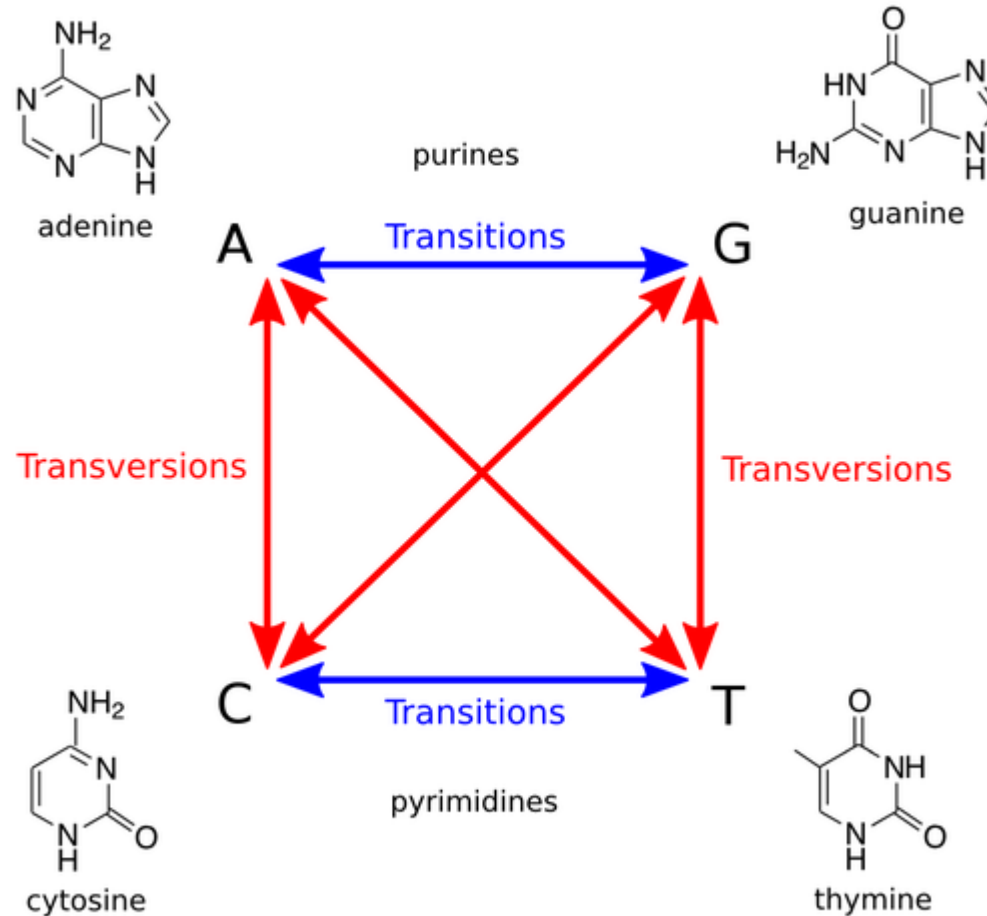
prospěšné
škodlivé
neutrální

podle rozsahu: genové/bodové
chromozomové
genomové

zpětné mutace: frekvence zpravidla 10× nižší

Bodové mutace

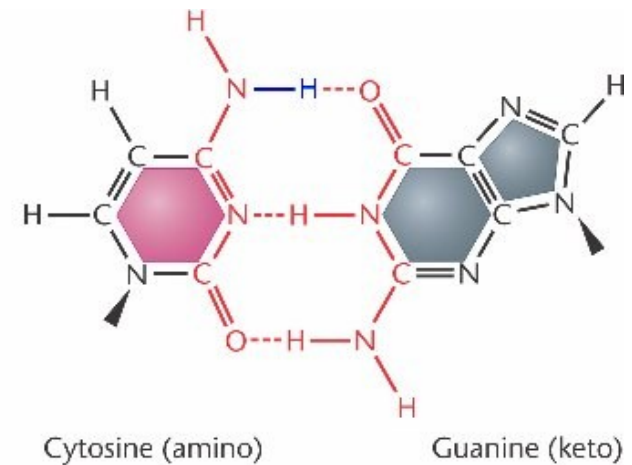
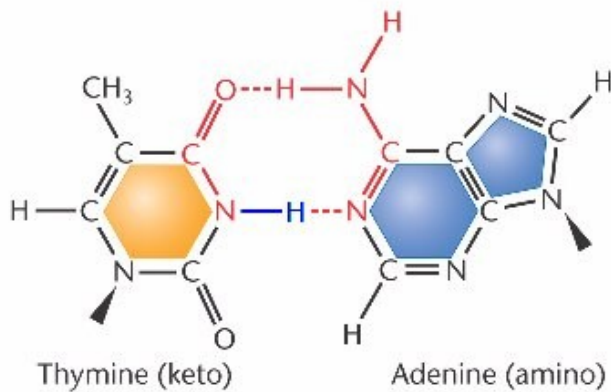
1. substituce (transice, transverze)



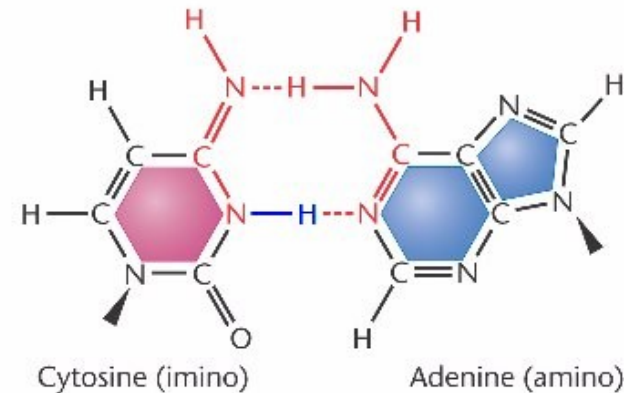
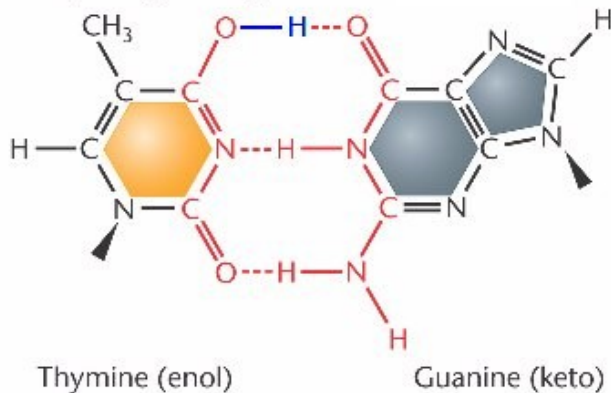
Bodové mutace

transverze → purin-purin, pyri-pyri ⇒ problém s párováním ⇒ oprava

(a) Standard base-pairing arrangements

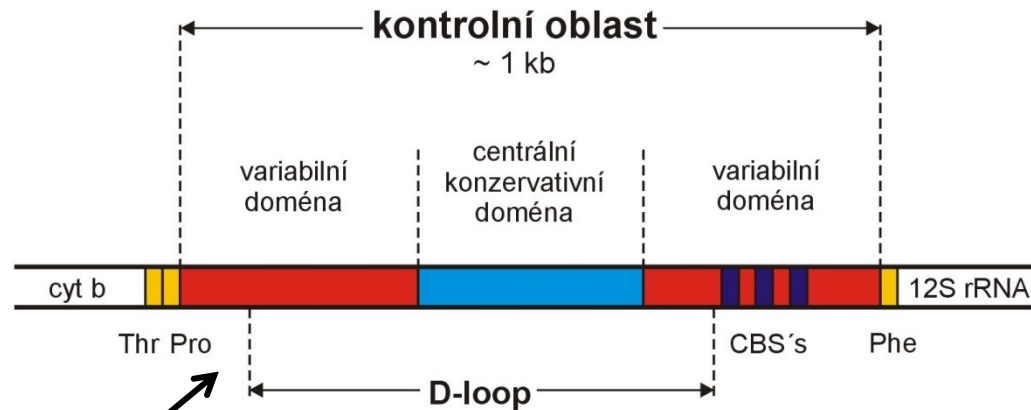


(b) Anomalous base-pairing arrangements



Bodové mutace

1. substituce (transice, transverze)



nekódující oblasti

synonymní

GTC → GTA
Val → Val

} = „tiché“ (*silent*) substituce

nesynonymní (záměnové)

měnící smysl (*missense*)

nesmyslné (*nonsense*)

GTC → TTC

Val → Phe

AAG → TAG

Lys → *ochre* (stop)

2.

inzerce

delece

ACGGT → ACAGGT

ACGGT → AGGT

} =

indels

→

posunutí čtecího
rámce

Chromozomové mutace (chr. přestavby)

inverze

pericentrické

paracentrické

translokace

fúze a disociace

(robertsonské translokace)

reciproké translokace celých ramen (WART)

delece

duplikace

inzerce

Genomové mutace

-ploidie (polyploidie)

především rostliny

u živočichů méně (bezobratlí, ryby, obojživelníci)

během evoluce obratlovců došlo ke 2 kolům duplikace celého genomu (2R-hypotéza)

polyploidní jedinci zpravidla větší (zvýšený objem buněk)

liché násobky genomu → problémy v meióze ⇒ reprodukční bariéra
(ne vždy – např. triploidní skokani)

autopolyploidie: kombinace dvou stejných genomů

fúze buněk

endoreplikace

abortivní buněčný cyklus

alopolyploidie: kombinace dvou různých genomů

fúze diploidních gamet

polyspermie

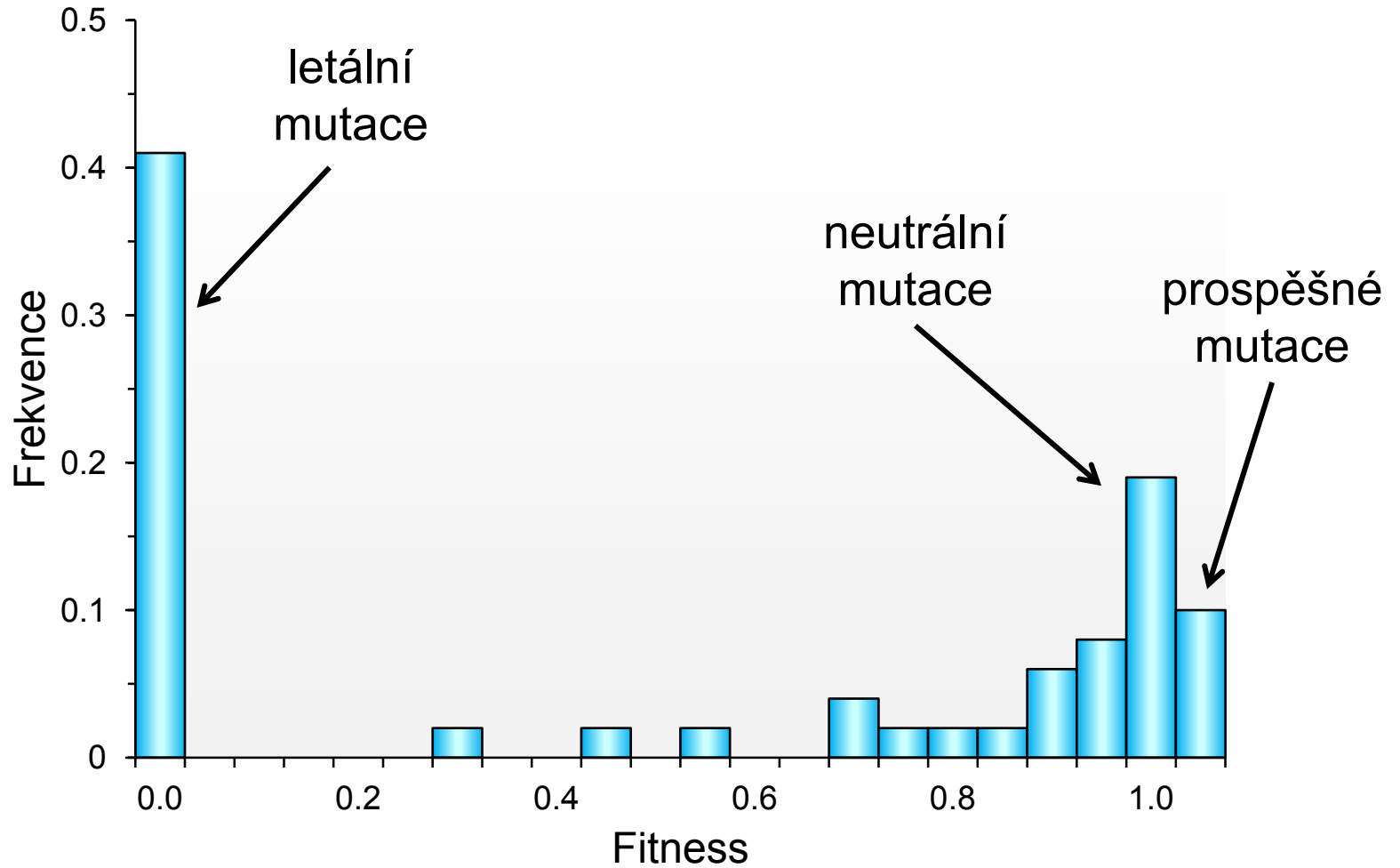
Genomové mutace

-zomie (monozomie, trizomie)

- většinou neslučitelné se životem
- **monozomie**: jediná životaschopná = X0 (Turnerův syndrom)
- **trizomie**: nerovnováha dávky genů (zvýšená exprese trizomického páru)
- životaschopné trizomie : XXY, XXX, XYY, Patau syndrom (chr. 13), Edwardsův s. (chr. 18), Downův s. (chr. 21)

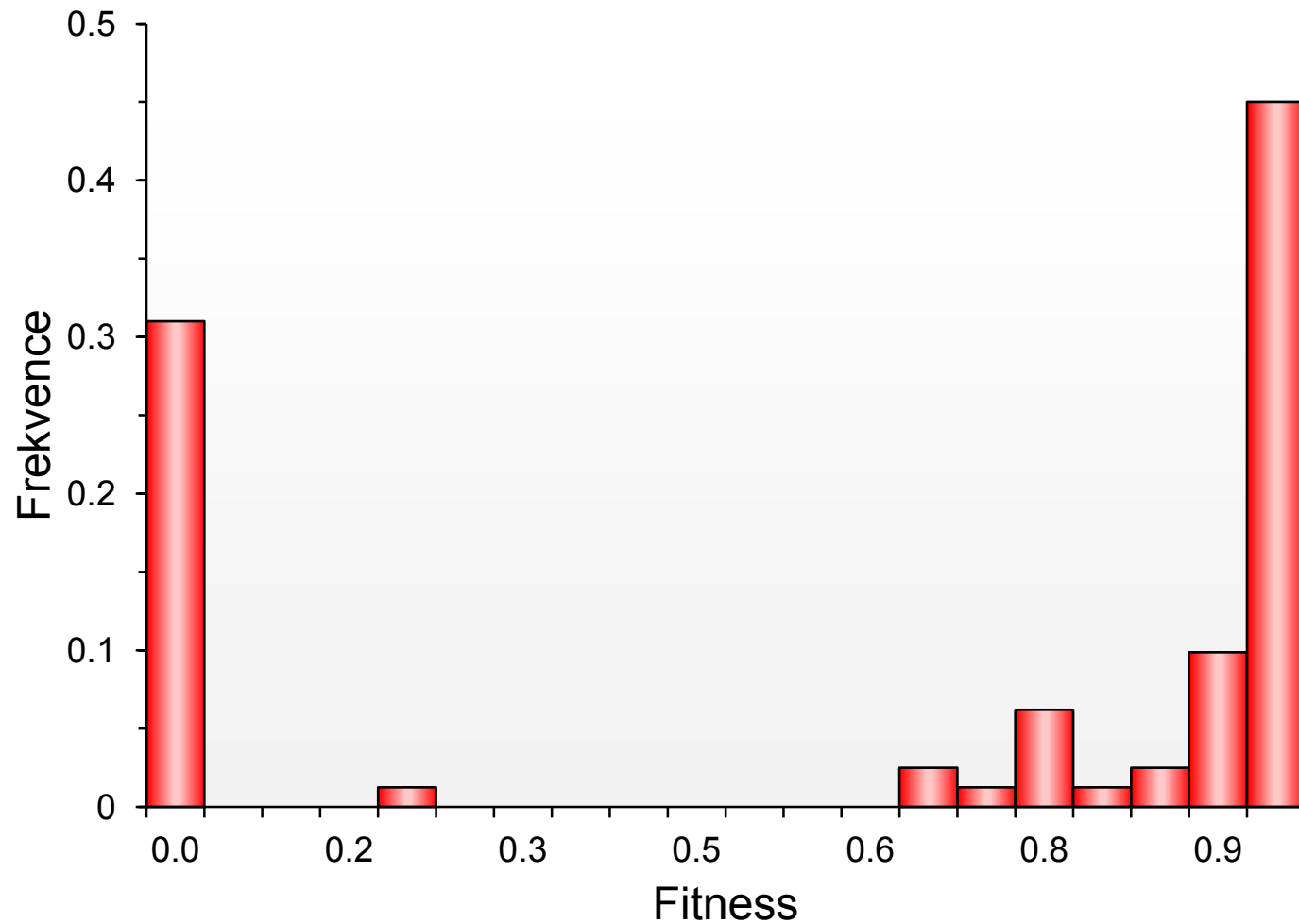
Distribution of fitness effects (DFE)

RNA virus (Sanjuan et al., PNAS 2004)



Distribution of fitness effects (DFE)

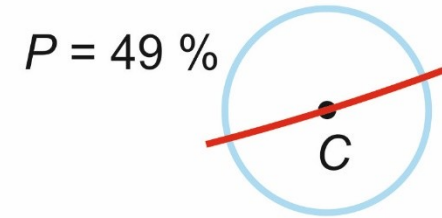
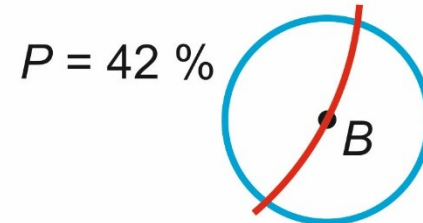
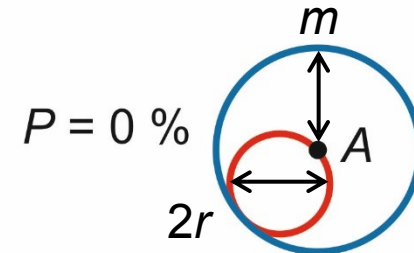
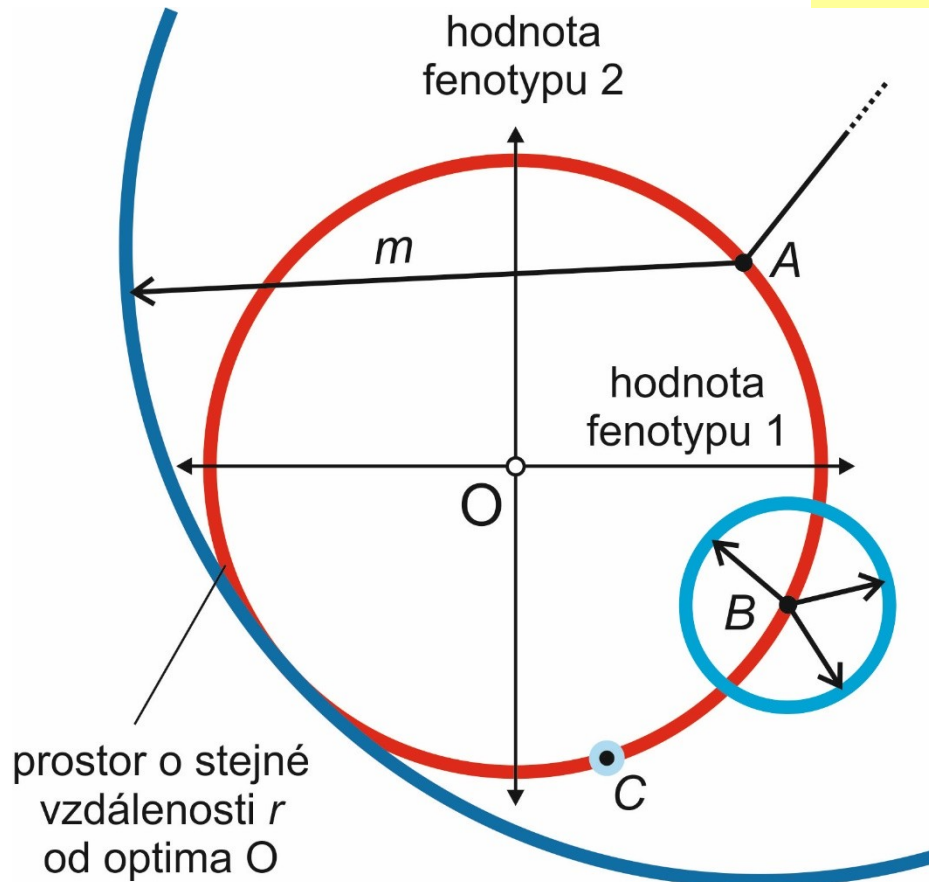
kvaskinka (Eyre-Walker and Keightley, Nature 2007)



Fisherův koncept hyperkoule:

čím větší dopad mutace na fitness, tím vyšší pravděpodobnost, že je škodlivá

$$\Pr(\text{mutace zvýší fitness}) = \frac{1}{2} \left(1 - \frac{m}{2r} \right)$$



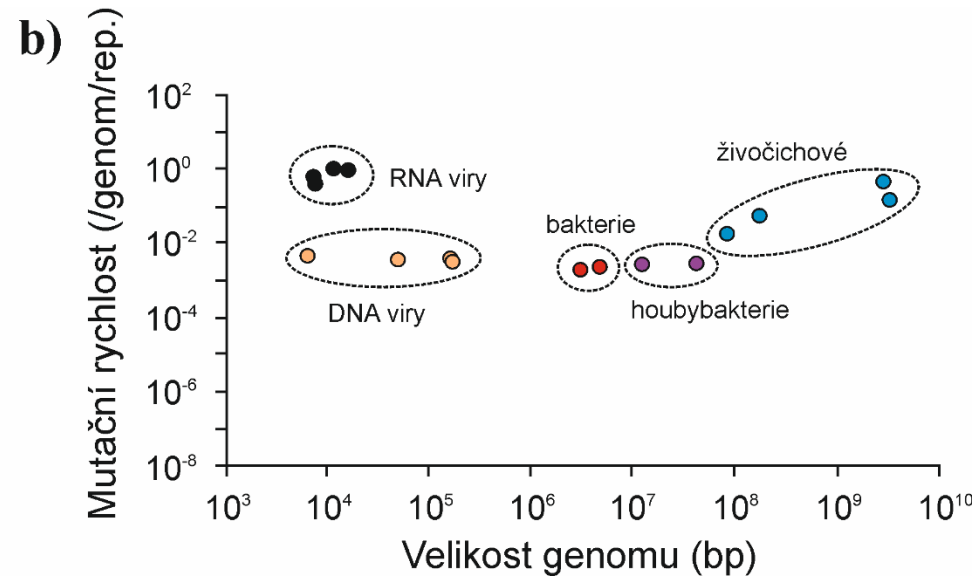
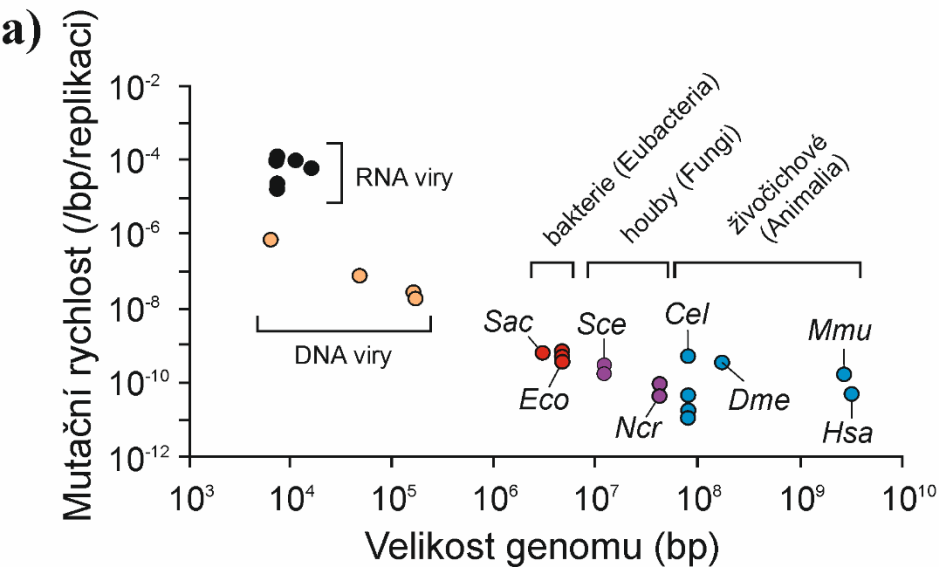
Mutační rychlost (μ)

mutace náhodné co do účinku, nenáhodné co do pozice a rychlosti

člověk: $1,1-2,3 \times 10^{-8}/\text{bp}/\text{generaci} \Rightarrow 10^{-10}/\text{bp}/\text{replikaci}$

E. coli: $\sim 10^{-9}/\text{bp}/\text{replikaci}$

negativní korelace mezi mutační rychlostí a velikostí genomu:



Mutační rychlost (μ)

různé typy mutací – různé rychlosti, např. transice > transverze

mutační „hotspots“: CpG u živočichů (metylovaný C \rightarrow T); TpT prokaryot
„SOS reakce“ bakterií (reakce na poškození DNA), „contingency loci“,
minisatelity (VNTR), mikrosatelity (STR)

mtDNA > jad. DNA (u člověka řádově 1000 \times)

pohlavní chromozomy > autozomy: Y > Z > A > X > W

vliv blízkosti počátku replikace, centromery, telomery, repetitivních sekvencí,
intenzity transkripce

studenokrevní živočichové: > t \Rightarrow > μ

RNA viry (HIV, chřipka) – absence opravných mechanismů \Rightarrow
 $\sim 1\,000\,000 \times$ vyšší frekvence mutací

paraziti, protilátky, imunoglobuliny

> μ somatických mutací

samci > samice: člověk 6 \times , hlodavci, liška: 2 \times ... více buněčných dělení
v zárodečných buňkách

rekurentní (opakované) mutace → mutační tlak:

počáteční frekvence $A = p_0 (= 1)$

mutace $A \rightarrow a$

frekvence A v čase $t = p_t$

mutační rychlost = μ (např. 10^{-4})

$$p_1 = p_0(1 - \mu)$$

$$p_2 = p_1(1 - \mu) = p_0(1 - \mu)(1 - \mu) = p_0(1 - \mu)^2 \dots p_t = p_0(1 - \mu)^t$$

$$p_t = p_0 (1 - \mu)^t$$

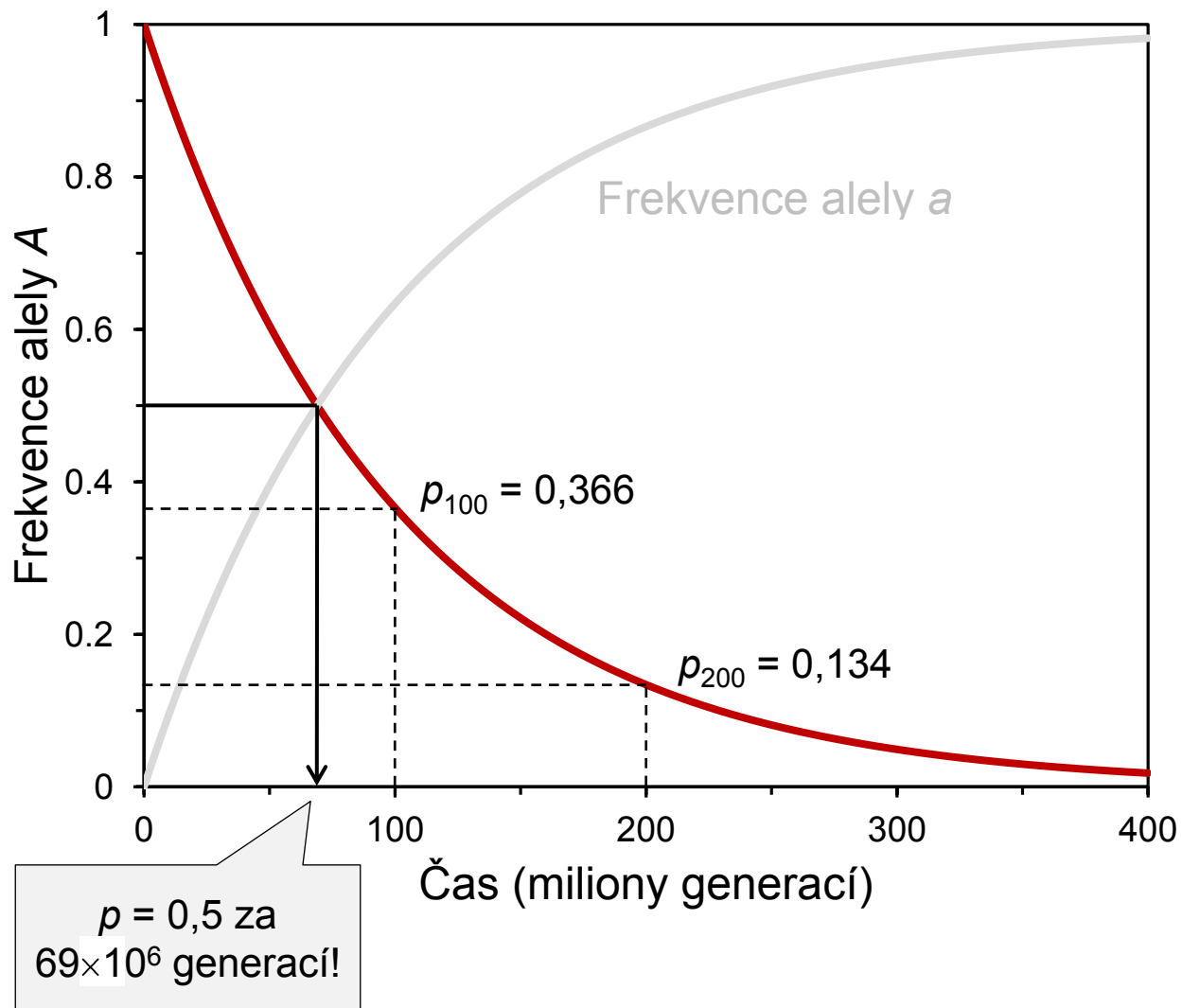
doba nutná ke snížení frekvence A na polovinu = $t_{1/2}$

$$t_{1/2} = \ln(1/2) / \ln(1 - \mu) \approx 0,6931 / \mu$$

\Rightarrow při $\mu = 10^{-4}$ $t_{1/2} \approx 6931$ generací; při $\mu = 10^{-5}$ $t_{1/2} \approx 69\,310$ generací;
při $\mu = 10^{-6}$ $t_{1/2} \approx 693\,100$ generací atd.

\Rightarrow změna frekvence alely mutací velmi pomalá

$10^{-8}/\text{bp/gen}$:



RNA viry: $10^{-4}/\text{bp/gen}$ → 6900 generací

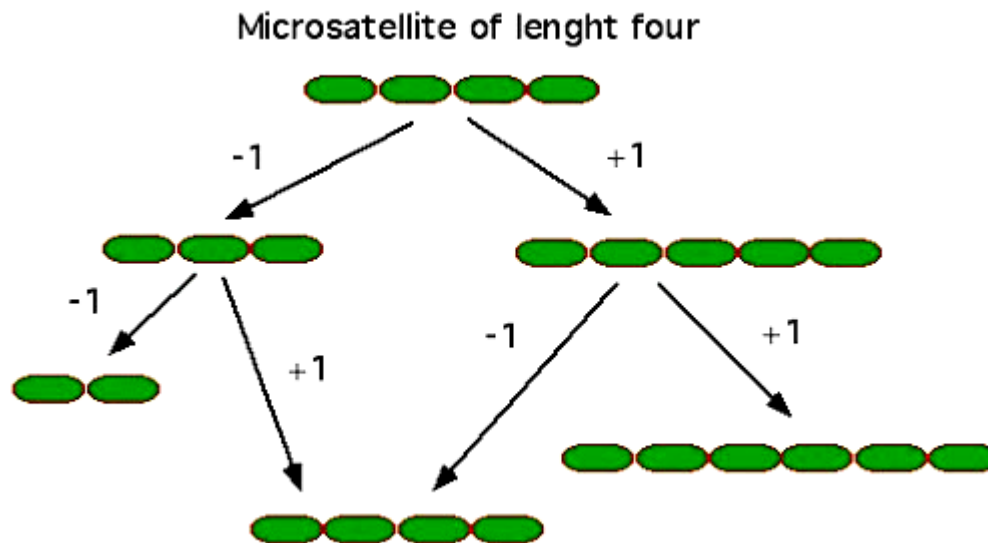
Mutační modely:

m. nekonečných alel (*infinite allele model*): každá mutace = nová alela;

m. nekonečných pozic (*infinite site model*): jen jedna mutace na pozici

m. konečných pozic (*finite site model*)

stepwise mutation model (SMM) – mikrosatelity



Mutační modely:

birth-death model: „narození“ = duplikace, „smrt“ = pseudogenizace
subfunkcionalizace, neofunkcionalizace

