Odhad heritability v přírodních/divokých populacích

prof. Ing. Tomáš Urban, Ph.D.
urban@mendelu.cz

Odhad heritability v přírodních populací

Kovariance mezi příbuznými pro kvantitativní znaky je základem pro odhad dědičnosti znaků a jejich genetických korelací (Falconer, 1989).

Klasicky se úroveň příbuznosti mezi příbuznými vypočítává ze známých rodokmenů.

K odhadu genetických a environmentálních komponent rozptylu kvantitativních znaků na základě porovnání úplných sourozenců a polosourozenců, rodičů a jejich potomků nebo podle reakce na selekci lze použít řadu statistických metod.

Dva důsledky omezení, která vyplývají z tradičních kvantitativních genetických postupů, jsou:

1. organismy s dlouhou generační dobou (např. stromy, velcí obratlovci) je velmi obtížné studovat a
2. neznámý vztah mezi odhady heritability a genetickými korelacemi vytvořenými v laboratoři a těmi, které se projevují ve volně žijících populacích v přirozených podmínkách prostředí.

Problém přírodních populací

- zkoumání kvantitativní genetické variability v přírodních populacích, které nevyžadují experimentální manipulaci, a umožnit genetické studie organismů, které nelze studovat v laboratoři buď z důvodu dlouhého života, nebo obtížného chovu v zajetí.

Ve většině přírodních populací není možné vytvořit křížence a obvykle není ani možné sledovat přirozené vzorce rozmnožování. A i když si organismy vytvářejí teritoria a zdánlivě se rozmnožují monogamně, takže je možné identifikovat pravděpodobné rodiče (např. mnoho ptáků), mimopárové rozmnožování a jejich vliv na vztahy v rámci rodiny není snadné zjistit.

'Natural' heritability

Vytváření umělých populací ve volné přírodě, **cross-fostering design** nebo **vytváření populací přizpůsobených laboratornímu prostředí** (např. linie Rose u Drosophila melanogaster; Chippindale et al., 1996).

U drozofily se odhadovala heritabilita v přírodě pomocí regrese laboratorně vypěstovaných potomků na rodiče nasbírané v terénu.

h^2 - dvojnásobek regresního koeficientu z hodnot samic odchycených ve volné přírodě s hodnotami průměrů jejich laboratorně chovaných dcer.

Table 2. Heritability and standard error per month at the PNI population of *Drosophila mediopunctata*

	Sep/86 ($n = 28$)	Jun/87 ($n = 47$)	Aug/87 ($n = 120$)	Nov/87 ($n = 45$)	Mar/88 ($n = 18$)	May/88 ($n = 84$)	Total ($n = 342$)
PC1-TOT	0.13 ± 0.14^{ns}	-0.01 ± 0.13^{ns}	$0.51 \pm 0.11^{*}$	-0.10 ± 0.12^{ns}	0.10 ± 0.15^{ns}	0.07 ± 0.08^{ns}	$0.18 \pm 0.06^{*}$
PC2-TOT	0.09 ± 0.17^{ns}	$0.52 \pm 0.20^{*}$	$0.57 \pm 0.13^{*}$	0.51 ± 0.20^{ns}	-0.11 ± 0.32^{ns}	0.13 ± 0.18^{ns}	$0.50 \pm 0.07^{*}$
PC3-TOT	0.41 ± 0.21^{ns}	0.16 ± 0.16^{ns}	$0.57 \pm 0.16^{*}$	$0.60 \pm 0.19^{*}$	0.55 ± 0.35^{ns}	0.40 ± 0.17^{ns}	$0.47 \pm 0.08^{*}$

$^{ns}P > 0.01$.
$^{*}P < 0.01$; n = number of families.

Cross-fostered – „křížen ě“ odchovávaná mláďata

Smith, James N. M.; Dhondt, André A. (1980). *EXPERIMENTAL CONFIRMATION OF HERITABLE MORPHOLOGICAL VARIATION IN A NATURAL POPULATION OF SONG SPARROWS. Evolution, 34(6), 1155–1158.* doi:10.1111/j.1558-5646.1980.tb04058.x

- regrese skóre středu rodičů vůči průměrnému skóre potomků pro experimentální rodiny s pravými i náhradními (foster, pěstouni) rodiči a pro kontrolní rodiny.

TABLE 2. *Estimates of single-parent/offspring heritability[1] of four morphological characters in cross-fostered and control broods of song sparrows.*

	Young in experimental families				Young in control families			
	Female-offspring heritability		Male-offspring heritability		Female-offspring heritability		Male-offspring heritability	
Character	h^2	SE	h^2	SE	h^2	SE	h^2	SE
Beak length	0.16	0.34	0.71*	0.27	0.29	0.43	−0.04	0.29
Beak depth	1.33**	0.40	1.23**	0.37	1.54**	0.43	0.50	0.50
Beak width	0.22	0.34	0.59*	0.24	0.47	0.25	0.30	0.27
Tarsus length	0.92*	0.37	0.71	0.57	1.12**	0.37	0.90**	0.33
Sample size (families)	26		23		32		30	
Average	0.66		0.81		0.86		0.42	

* = $P < .05$; ** = $P < .01$.
[1] Heritability (h^2) = 2 × slope.

Více markerů (tj. celkový počet alel) by zvýšilo přesnost odhadu příbuznosti a zvýšení celkového počtu jedinců výrazně zvyšuje přesnost a správnost odhadů heritability a genetických korelací.

Nyní: odhad heritability na základě podílu fenotypové variance vysvětlené genotypovanými SNP.

Yang et al. (2010) hodnotili svou metodu odhadem podílu variance vysvětlené ~290 k SNP u 3925 lidí a po korekci na SNP, které nejsou genotypizovány, a na SNP s nižší frekvencí minoritních alel se poměrně blíží heritabilitě výšky odhadované ze sourozeneckých modelů, ~0,8.

Software GCTA byl také použit k odhadu heritability pro hmotnost u lidí (Yang et al. 2011), inteligenci (Davies et al. 2011), náchylnost k nemocem (Lee et al. 2012) a osobnost (Verweij et al. 2012).

Podobné metody byly použity k rozdělení genetické variability v délce křídel (Robinson et al. 2013), velikosti snůšky a hmotnosti vajec populace volně žijících ptáků (Santure et al. 2013).

Heritabilita odhadnutá pomocí genetických markerů

- metoda umožňující kvantitativní genetickou analýzu volně žijících populací s přirozenými vzorci rozmnožování.
- metody byly založeny na lineárních vztazích mezi odhady příbuznosti na základě markerů a fenotypy. Kvůli nejistotě v odhadu příbuznosti a záměně příbuznosti s prostředím však odhady h^2 z těchto metod nebyly přesné.

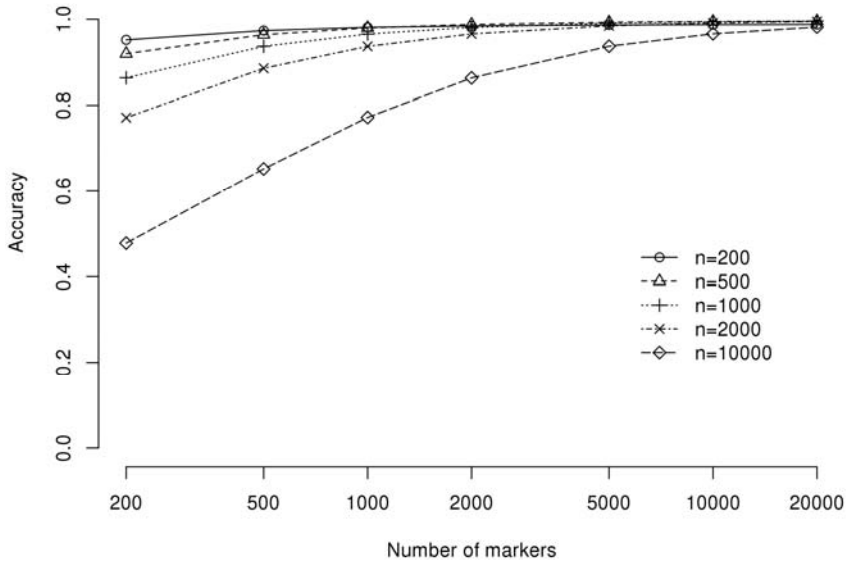
1) jedinci jsou ve studované populaci genotypizováni pro lokusy molekulárních markerů a vyhodnoceni z hlediska kvantitativních znaků.
2) na základě údajů z markerů postupem maximální pravděpodobnosti (ML) odvodí příbuznosti mezi dvojicemi jedinců, přičemž se předpokládá směs nepříbuzných a úplných sourozeneckých párů.
3) odhady příbuznosti jsou kombinovány s údaji o kvantitativních znacích ve smíšeném modelu, aby bylo možné odvodit h^2 a genetické korelace.

Odhad podílu fenotypové variance vysvětlené genotypovanými SNP

Analýza GCTA se skládá ze dvou kroků. Nejprve se všechny SNP použijí k výpočtu matice genetické příbuznosti vztahů (GRM) mezi vzorky.

GRM se pak použije jako prediktor ve smíšeném lineárním modelu s rysem jako odpovědí k odhadu h^2.

Vliv velikosti vzorku a hustoty markerů na korelaci mezi skutečnou a teoretickou genetickou příbuzností



			Realized h^2 ($\hat{h}^2$)		SDS		REML	
h^2	n	m	$\overline{\hat{h}^2}$ ± SD	$\hat{h}^2$Range of	$\overline{\hat{h}^2}$ ± SD[b]	$\hat{h}^2$Range of	$\overline{\hat{h}^2}$ ± SD	$\hat{h}^2$Range of
0.2	500	200	0.187±0.030	0.146–0.243	0.193±0.078	0.081–0.313	0.187±0.047*	0.121–0.249
0.2	500	2000	0.190±0.032	0.140–0.244	0.197±0.077	0.082–0.342	0.199±0.065	0.095–0.303
0.2	500	20000	0.186±0.031	0.136–0.238	0.184±0.088	0.054–0.319	0.192±0.069	0.095–0.300
0.2	1000	200	0.194±0.029	0.153–0.243	0.198±0.069	0.106–0.337	0.197±0.039	0.134–0.261
0.2	1000	2000	0.189±0.022	0.157–0.223	0.192±0.045	0.118–0.262	0.193±0.040	0.132–0.263
0.2	1000	20000	0.191±0.024	0.155–0.228	0.201±0.086	0.091–0.342	0.194±0.048	0.113–0.273
0.2	5000	200	0.199±0.011	0.181–0.216	0.196±0.041	0.136–0.275	0.198±0.011	0.179–0.216
0.2	5000	2000	0.199±0.013	0.181–0.218	0.198±0.039	0.149–0.269	0.199±0.016	0.173–0.223
0.2	5000	20000	0.200±0.011	0.184–0.217	0.204±0.050	0.124–0.297	0.201±0.019	0.174–0.233
0.5	500	200	0.476±0.049	0.407–0.562	0.492±0.138	0.296–0.725	0.482±0.048*	0.415–0.561
0.5	500	2000	0.480±0.052	0.394–0.564	0.495±0.119	0.300–0.702	0.499±0.070	0.384–0.602
0.5	500	20000	0.474±0.051	0.386–0.556	0.480±0.150	0.271–0.694	0.480±0.085*	0.337–0.602
0.5	1000	200	0.487±0.046	0.419–0.562	0.494±0.114	0.344–0.683	0.495±0.035	0.435–0.548
0.5	1000	2000	0.481±0.036	0.428–0.535	0.482±0.072*	0.374–0.613	0.491±0.047	0.417–0.565
0.5	1000	20000	0.485±0.038	0.423–0.542	0.511±0.144	0.343–0.774	0.490±0.050	0.408–0.563
0.5	5000	200	0.498±0.017	0.470–0.525	0.497±0.063	0.408–0.615	0.497±0.010*	0.479–0.513
0.5	5000	2000	0.497±0.020	0.468–0.528	0.498±0.052	0.428–0.595	0.498±0.016	0.472–0.521
0.5	5000	20000	0.499±0.017	0.474–0.526	0.507±0.069	0.411–0.609	0.501±0.022	0.464–0.534
0.8	500	200	0.782±0.034	0.733–0.837	0.802±0.177	0.538–1.096	0.790±0.022*	0.753–0.828
0.8	500	2000	0.785±0.035	0.722–0.838	0.801±0.135	0.592–1.022	0.797±0.044	0.725–0.860
0.8	500	20000	0.781±0.035	0.715–0.833	0.797±0.202	0.536–1.114	0.783±0.064*	0.670–0.874
0.8	1000	200	0.790±0.031	0.742–0.837	0.797±0.137	0.612–1.008	0.797±0.014	0.775–0.819
0.8	1000	2000	0.786±0.024	0.749–0.821	0.785±0.084	0.652–0.953	0.795±0.027	0.754–0.842
0.8	1000	20000	0.789±0.026	0.746–0.825	0.829±0.191	0.573–1.189	0.796±0.032	0.736–0.841
0.8	5000	200	0.799±0.011	0.780–0.815	0.799±0.081	0.687–0.955	0.798±0.004*	0.792–0.806
0.8	5000	2000	0.798±0.013	0.779–0.817	0.800±0.057	0.715–0.899	0.799±0.007	0.787–0.809
0.8	5000	20000	0.799±0.011	0.783–0.816	0.809±0.088	0.665–0.937	0.801±0.013	0.778–0.819

*Indicates a significant deviation from h^2 according to a t-test: $t = (\overline{\hat{h}^2} - h^2)/(SD/\sqrt{100})$.

[a]The 27 simulation trials consist of three levels of true narrow-sense heritability (h^2), three sample sizes (n) and three marker densities (m). In each simulation sample, h^2 is estimated by the symmetric difference squared (SDS) method implemented in our R package, SDS/R and by residual maximum likelihood (REML) method implemented in the GCTA software.

[b]SD = standard deviation.

doi:10.1371/journal.pone.0102715.t001

Odhad heritability u člověka pomocí analýz dvojčat

prof. Ing. Tomáš Urban, Ph.D.
urban@mendelu.cz

Dvojčata

Jedinci s identickým genotypem ~ podobně jako inbrední linie

Studie dvojčat (lidé, skot)

Zohlednění společného prostředí – od zygoty po narození, a v dalším období, kdy jsou spolu -> variance mezi páry dvojčat obsahuje varianci společného prostředí (V_{Ec}), spojená s V_G

Tento problém lze řešit analýzou dvou typů dvojčat:

monozygotní (MZ) a dizygotní(DZ)

- MZ ~ genetická identita
- DZ ~ úplní sourozenci

- ve společném prostředí mají přibližně stejný rozměr jako MZ

Odhad množství genetické variance, ptáme se nakolik méně jsou si podobní DZ než MZ?

Pokud jsou si MZ dvojčata fenotypově podobnější než DZ dvojčata, pak to naznačuje, že genetická variabilita více ovlivňuje fenotypové rozdíly.

Klasický design dvojčat je také známý jako model ACE.

Model ACE odhaduje heritabilitu v úzkém smyslu za předpokladu, že genetická variance dominance je nulová

Zavedením **modelu ADE** [v němž jsou složky variance modelovány jako aditivní (A), dominantní a epistázové (D) a nesdílené vlivy prostředí] lze odhadovanou heritabilitu rozdělit na aditivní a dominantní složky variance. Pokud se u dvojčat MZ a DZ projevují nepřímé genetické účinky rodičů podobně, budou tyto účinky zahrnuty do složky společného prostředí (C).

Předpoklady studia dvojčat

- Společné prostředí přispívá k danému fenotypu stejnou měrou u obou párů dvojčat MZ i DZ, což se označuje jako "předpoklad stejného prostředí" (equal environment assumption - EAA)
  - Pokud by EEA neplatila, odhady heritability by byly pravděpodobně nadsazené, protože rozdílná prostředí by byla mylně přisuzována rozdílům v genetické variabilitě
- Dvojčata jsou zobecnitelná na obecnou populaci, pokud jde o fenotyp, který je předmětem zájmu. Platnost tohoto předpokladu byla prokázána v mnoha studiích
- V rámci populace dochází k „náhodnému páření“.
- „nejasná role podobnosti prostředí“ na genetické faktory a výsledky je omezená.

Pokud všechny tyto předpoklady platí, pak porovnáním fenotypových korelací MZ a DZ dvojčat lze spolehlivě odhadnout heritability.

Složení komponent variance mezi páry dvojčat

	Mezi páry, σ_b^2	Uvnitř párů, σ_w^2
MZ dvojčata	$V_A + V_D + V_{Ec}$	V_{Ew}
DZ dvojčata	$½ V_A + ¼ V_D + V_{Ec}$	$½ V_A + ¾ V_D + V_{Ew}$
Rozdíl (MS_{MZ} - MS_{DZ})	$½ V_A + ¾ V_D$	$½ V_A + ¾ V_D$

Za předpokladu, že obě složky variance prostředí, V_{Ec} a V_{Ew}, jsou stejné pro MZ a DZ, se V_I neuvažuje; **celková genetická variance je stejná u MZ a u DZ**

Rozdíly mezi MZ a DZ dvojčaty v obou komponentách odhadují ½ V_A společně se ¾ V_D
$(\sigma_{bMZ}^2 - \sigma_{bDZ}^2)$+$(\sigma_{wMZ}^2 - \sigma_{wDZ}^2)$

Korelace mezi páry dvojčat je podílem komponenty variance mezi páry fenotypovou variancí ~ dvojnásobek rozdílu mezi MZ korelace a DZ korelace:
„Heritabilita“ $= \frac{\left(V_A + 1\frac{1}{2}V_D\right)}{V_P}$
Výpočet se více blíží **heritabilitě v širším smyslu**, než v úzkém smyslu.

Odhady heritability – genomické metody

- U nepříbuzných jedinců
- genomické metody se týkají statistických metod aplikovaných na molekulárně genetická (včetně celogenomových) data, která byla buď přímo změřena, nebo imputována pomocí referenčních panelů.
- Některé genomové metody se zaměřují na odhad „SNP heritability“ - což je zvláštní případ heritability v úzkém smyslu odhadované z měřených SNP.
- **LD score regression**
- **Genomic relatedness restricted maximum likelihood**

Genomické metody: příbuzní jedinci

Společným omezením genomických metod aplikovaných na nepříbuzné jedince je nemožnost zohlednit vliv prostředí.

Byly vyvinuty různé metody, které umožňují implementaci genomických metod ve velkých vzorcích příbuzných jedinců.

Sibling regression

Může poskytnout odhad heritability v úzkém slova smyslu; žádné další předpoklady o distribuci účinků SNP; zahrnuje účinky vzácných SNP do odhadů heritability; odolný vůči chybám genotypování a určitému chybění; částečně zohledňuje stratifikaci populace.

Pro přesné výsledky je nutná velká velikost vzorku vzhledem k malé směrodatné odchylce IBD sdílených mezi sourozenci, např. Visscher et al. stanovili v rámci svého vzorku průměrný podíl IBD sdílených genomem v rámci sourozeneckých párů se směrodatnou odchylkou 0,036; odhad je vztažen ke zvolené referenční populaci.

Další genetické parametry

- koeficient opakovatelnosti
- genetické korelace

Koeficient opakovatelnosti

Opakované měření stejné vlastnosti na stejném jedinci v průběhu jeho života (nejlépe za stejných podmínek) → podobnost měření je závislá na velikosti genetické determinace.

Koeficient opakovatelnosti udává podíl proměnlivosti zapříčiněný genetickými rozdíly z celkové fenotypové proměnlivosti.

- Na stejném místě (prostorové opakování, topografická)
- Opakování v časovém období (s věkem)

Koeficient opakovatelnosti - r_{op}

$$S_P^2 = S_G^2 + S_{E_P}^2 + S_{GE_P}^2 + S_{E_T}^2 + S_{GE_T}^2$$

S_g^2 S_e^2

stálé vlivy: $s = S_G^2 + S_{E_P}^2 + S_{GE_P}^2 = \boxed{S_g^2}$

podobnost

dočasné vlivy: $e = S_{E_T}^2 + S_{GE_T}^2 = \boxed{S_e^2}$

rozdílnost

Měření jsou **rozdílná** → vlivy dočasné jsou různé

Měření jsou **podobná** → vlivy stálé jsou stejné

Zdroj proměnlivosti mezi jedinci

Zdroj proměnlivosti mezi měřeními u jedince

Korelace mezi fenotypovými hodnotami různých měření u všech jedinců populace je OPAKOVATELNOST:

koeficient opakovatelnosti: r_{op}

$$r_{op} = r = \frac{S_g^2}{S_P^2} = \frac{S_G^2 + S_{E_P}^2 + S_{GE_P}^2}{S_P^2} = h_š^2 + \frac{S_{E_P}^2 + S_{GE_P}^2}{S_P^2}$$

$$h_u^2 < h_š^2 < r_{op}$$

$$h^2 = r_{op} - \frac{\sigma_{Ep}^2 + \sigma_{GEp}^2}{\sigma_P^2}$$

r_{op} je horní hranicí h^2

Př.: U 250 prasnic ve velkochovu byly sledovány počty všech narozených selat za jejich první čtyři vrhy. Vypočítejte odhad ***koeficientu opakovatelnosti*** této užitkové vlastnosti včetně jeho střední chyby.

Proměnlivost	SS	df	MS	složení MS
Mezi skupinami jedinců mezi prasnicemi (a)	1681,99	$df_a = p - 1 = 249$	6,777	$\sigma_e^2 + n_0\sigma_g^2$
Uvnitř skupin (e)	3044,25	$df_e = n - p = 750$	4,059	σ_e^2
Celková (c)	4756,24	$df_c = n - 1 = 999$	-	

$MS_a = \sigma_e^2 + n_0\sigma_g^2$

$MS_e = \sigma_e^2$

p = 250 počet prasnic
n = 1000 počet sledovaných vrhů
k = 4 počet opakování u jedné prasnice, zde platí: $k = n_0$ vážený počet potomků

Odhad variance genetické: $\sigma_g^2 = \frac{MS_a - MS_e}{n_0} = 0{,}6740$

Odhad variance fenotypové: $\sigma_P^2 = (\sigma_G^2 + \sigma_{Ep}^2 + \sigma_{GEp}^2) + \sigma_{Et}^2 + \sigma_{GEt}^2 = \sigma_g^2 + \sigma_e^2 = 4{,}7330$

Výpočet odhadu intraklasního koeficientu korelace $\rho = r_{op}$ - **koeficientu opakovatelnosti**:

$$r_{op} = \frac{\sigma_g^2}{\sigma_P^2} = 0{,}1424$$

$$se_{r_{op}} = \sqrt{\frac{2.(1-\rho)^2\left(1+(n_0-1)\rho\right)^2}{n_0(n_0-1)(p-1)}} = 0{,}0317$$

$$Y_{ijkl} = HY_i + M_j + a_k + p_k + b_1 x_{ijkl} + b_2 (x_{ijkl})^2 + e_{ijkl}$$

To estimate variances with the **repeatability model**, **REML**, was used.
$\mathbf{Y_{ijkl}}$ is milk, fat, or protein yield with twice daily milking and complete records of length 240 to 305 d for cow k in herd-year i, calving month j, and lactation l;
$\mathbf{HY_i}$ is effect of herd-year i;
Mj is effect of calving month j ;
$\mathbf{x_{ijkl}}$ is the cow age at calving;
$\mathbf{a_k}$ is additive genetic effect of cow k;
$\mathbf{p_k}$ is the permanent environmental effect of cow k;
$\mathbf{b_1}$ and $\mathbf{b_2}$ are the partial regression coefficients for linear and quadratic effects of age at calving;
$\mathbf{e_{ijkl}}$ is random residual associated with each record.

Heritability and Repeatability for Milk Production Traits of Japanese Holsteins from an Animal Model. Suzuki M., Van Flack L.D., 1994, J Dairy Sci 77:583-588

TABLE 3. Estimated variance components,[1] heritabilities, and repeatabilities (R) of milk yield for 10 data files with means and empirical standard errors.

Subsets	$\hat{\sigma}^2_a$	$\hat{\sigma}^2_p$	$\hat{\sigma}^2_e$	h^2	R
1	428,434	342,251	705,362	.29	.52
2	392,230	308,267	652,472	.29	.52
3	376,293	261,408	608,725	.30	.51
4	378,277	305,932	703,141	.27	.49
5	452,460	291,879	708,161	.31	.51
6	376,338	292,226	666,288	.28	.50
7	489,450	345,570	597,150	.34	.58
8	435,667	411,703	545,908	.34	.61
9	288,459	288,008	543,053	.26	.51
10	480,144	444,486	522,103	.33	.64
$\bar{X}$	409,775	329,173	625,236	.30	.54
SE	17,999	17,484	21,573	.01	.01

[1] $\hat{\sigma}^2_a$ = Estimated variance component for additive genetic effects (kilograms squared), $\hat{\sigma}^2_p$ = estimated variance component for permanent environmental effects (kilograms squared), and $\hat{\sigma}^2_e$ = estimated error variance (kilograms squared).

TABLE 5. Estimated variance components,[1] heritabilities, and repeatabilities (R) of protein yield for 10 data files with means and empirical standard errors.

Subsets	$\hat{\sigma}^2_a$	$\hat{\sigma}^2_p$	$\hat{\sigma}^2_e$	h^2	R
1	317.0	321.4	616.3	.25	.51
2	298.9	276.2	569.7	.26	.50
3	280.3	237.2	527.3	.27	.50
4	268.3	296.8	604.4	.23	.48
5	354.7	271.5	604.1	.29	.51
6	273.3	277.3	573.1	.24	.49
7	358.5	403.7	496.6	.28	.61
8	427.7	371.6	456.0	.34	.64
9	227.0	241.3	461.8	.24	.50
10	175.6	247.2	451.2	.20	.48
$\bar{X}$	298.1	294.4	536.1	.26	.52
SE	21.4	16.8	19.7	.01	.02

[1] $\hat{\sigma}^2_a$ = Estimated variance component for additive genetic effects (kilograms squares), $\hat{\sigma}^2_p$ = estimated variance component for permanent environmental effects (kilograms squared), and $\hat{\sigma}^2_e$ = estimated error variance (kilograms squared).

TABLE 4. Estimated variance components,[1] heritabilities, and repeatabilities (R) of fat yield for 10 data files with means and empirical standard errors.

Subsets	$\hat{\sigma}^2_a$	$\hat{\sigma}^2_p$	$\hat{\sigma}^2_e$	h^2	R
1	662.2	451.3	1077.0	.29	.51
2	576.8	393.5	943.3	.30	.51
3	552.0	339.7	890.3	.31	.50
4	511.2	450.4	973.2	.26	.50
5	678.3	372.1	982.2	.33	.52
6	463.7	446.9	926.9	.25	.50
7	811.9	476.4	833.4	.38	.61
8	633.4	485.7	768.9	.34	.59
9	423.1	394.0	787.6	.26	.51
10	345.3	390.8	786.8	.23	.48
$\bar{X}$	565.8	420.1	897.0	.30	.52
SE	41.2	14.5	30.5	.01	.01

[1] $\hat{\sigma}^2_a$ = Estimated variance component for additive genetic effects (kilograms squared), $\hat{\sigma}^2_p$ = estimated variance component for permanent environmental effects (kilograms squared), and $\hat{\sigma}^2_e$ = estimated error variance (kilograms squared).

Význam r_{op}

Určuje horní hranici koeficientu heritability (obsahuje i vlivy dominance a interakce genů, navíc i složku stálých vlivů prostředí a interakci genů a prostředí.

Pro odhad není nutné mít skupiny příbuzných jedinců.

Upřesňuje stanovení skutečně geneticky podmíněné užitkové hodnoty (odhad genotypové odchylky).

Pro výpočet odhadu koeficientu heritability.

Zpřesnění účinnosti selekce.

Fenotypové, genetické, genotypové a paratypové korelace

Vztahy mezi vlastnostmi

- **pozitivní** – zvýšením jednoho znaku se zvyšuje i druhý
- **negativní** – zvýšením jednoho znaku se snižuje druhý

Fenotypová korelace

- závislost mezi pozorovanými hodnotami $\mathbf{P_x}$ a $\mathbf{P_y}$ je daná kombinací závislostí genotypových hodnot a účinků prostředí;
- korelační páry = dvojice měření znaku **x** a znaku **y** u téhož jedince;

r_P

r_{PxPy}

Genetická korelace

genotypová = závislost mezi genotypovými hodnotami znaků **x** a **y**;

r_G;

r_{GxGy};

genetická = závislost aditivních hodnot obou znaků – častější;

- vyjadřuje rozsah, ve kterém dvě měření odrážejí, co je geneticky stejná vlastnost

(délka křídel : délka těla r = 0,75)

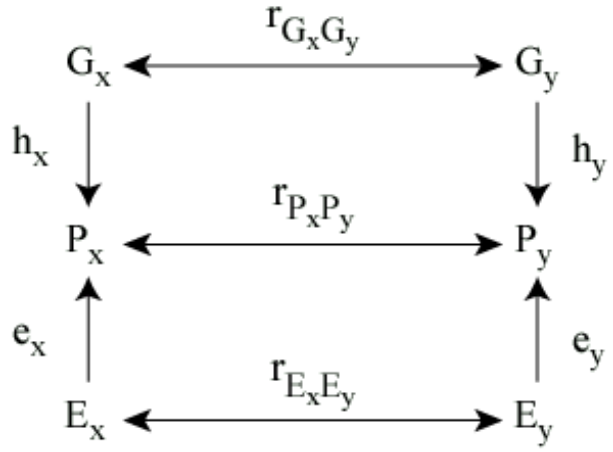
Paratypová korelace

Prostřeďová korelace

Závislost mezi efekty prostředí

r_E

r_{ExEy}



Ze schématu vyplývá **rozklad fenotypové korelace:**

$$r_{P_xP_y} = h_x \,.\, h_y \,.\, r_{G_xG_y} + e_x \,.\, e_y \,.\, r_{E_xE_y}$$

$$r_{P_xP_y} = h_x \,.\, h_y \,.\, r_{G_xG_y} + \sqrt{\left(1-h_x^2\right)\left(1-h_y^2\right)} \,.\, r_{E_xE_y}$$

Výpočet korelace – obecně

Zohledňuje variance (rozptyly) obou vlastností a jejich vzájemnou kovarianci:

$$r_{PxPy} = \frac{\sum[(x-\bar{x})\cdot(y-\bar{y})]}{\sqrt{\sum(x-\bar{x})^2 \cdot \sum(y-\bar{y})^2}}$$

Zjednodušeně: $r_{PxPy} = \dfrac{\mathrm{cov}_{xy}}{\sqrt{\mathrm{var}(x)\cdot\mathrm{var}(y)}}$

Vznik genetických korelací

Pleiotropní působení genů

= gen ovlivňuje více znaků ve stejném čase

Vazba genů

= geny lokalizovány v jedné vazbové skupině

- čím blíže, tím je vazba silnější
- není stálá – crossing over
- vazba mezi skupinami genů

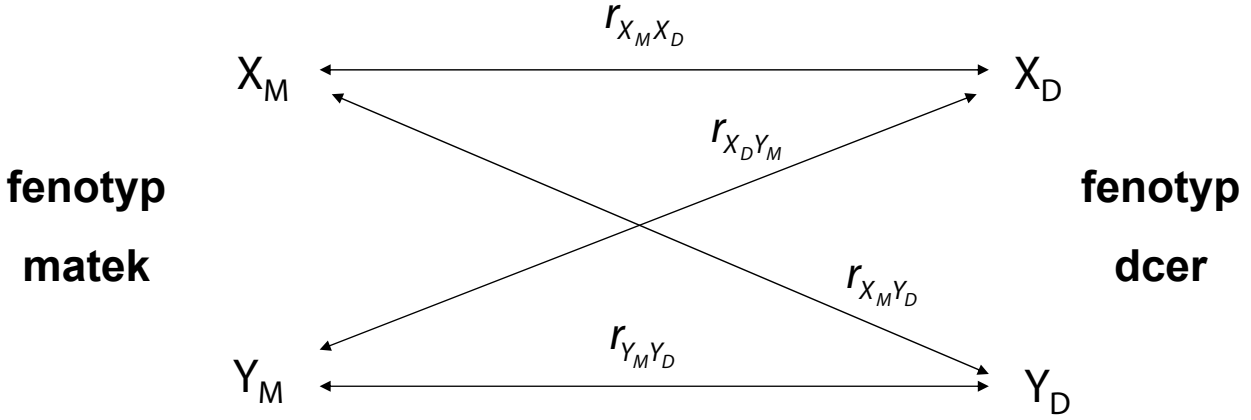
- **zavedení genů do populace**
  - ~ dočasné korelace
  - dané intenzivním využíváním vybraných jedinců
  - Při změně intenzity rozmnožování určitých jedinců genetické korelace zanikají
  - narušení náhodným pářením

Metody odhadu genetické korelace

- metoda korelace křížem
- analýza variance a kovariance
- realizovaná genetická korelace na základě výsledku selekce

1. Korelace křížem

- známe fenotypovou hodnotu znaku x, y u rodiče a potomka



$$r_{G_X G_Y} = \frac{r_{X_D Y_M} + r_{X_M Y_D}}{2 \cdot \sqrt{r_{X_M X_D} \cdot r_{Y_M Y_D}}}$$

Křížová korelace různých vlastností u matek a dcer

Křížová korelace stejných vlastností u matek a dcer

2. Analýza variance a kovariance

- Jednofaktorová
- Dvoufaktorová
- u skupin příbuzných jedinců
- Cíl: určit **variance** a **kovariance** genetické a prostředové
- Počítáme 2 ANOVY pro 2 vlastnosti a navíc analýzu kovariance současně pro obě vlastnosti

$$r_{G_X G_Y} = \frac{\mathrm{cov}(G_{xy})}{\sqrt{\mathrm{var}(G_x)\cdot \mathrm{var}(G_y)}}$$

$$r_{E_X E_Y} = \frac{\mathrm{cov}(E_{xy})}{\sqrt{\mathrm{var}(E_x)\cdot \mathrm{var}(E_y)}}$$

$$r_{P_X P_Y} = \frac{\mathrm{cov}(P_{xy})}{\sqrt{\mathrm{var}(P_x)\cdot \mathrm{var}(P_y)}}$$

Vztah mezi obsahem bílkovin (x) a tuku (y) v mléce dojnic byl sledován v osmi chovech (k), u skupin polosester (b). Výpočet odhadu byl proveden u 487 dojnic (n), dcer po 69 otců (b) bez ohledu na pořadí laktace.

% bílkovin $x_{ijk} = \mu + a_i + b_{ij} + e_{ijk}$
% tuku $y_{ijk} = \mu + a_i + b_{ij} + e_{ijk}$

	součet čtverců		stupně volnosti	součet součinů	**střední čtverec (variance)**		**střední produkt (kovariance)**
Zdroj proměnlivosti	SS_x	SS_y	F	SP_{xy}	MS_x	MS_y	MP_{xy}
- mezi chovy	$SS_{a\,x}$ = 4,7959	$SS_{a\,y}$ = 3,6014	f_a = k -1 = 7	$SP_{a\,xy}$ = 2,2283	$MS_{a\,x}$ = 0,685128	$MS_{a\,y}$ = 0,514485	$MP_{a\,xy}$ = 0,318328
- mezi otci uvnitř chovů	$SS_{o\,x}$ = 6,1125	$SS_{o\,y}$ = 12,6068	f_o = b – k = 61	$SP_{o\,xy}$ = 4,4461	**$MS_{o\,x}$ = 0,100204**	**$MS_{o\,y}$ = 0,206668**	**$MP_{o\,xy}$ = 0,072887**
- mezi dojnicemi uvnitř skupin otců	$SS_{e\,x}$ = 18,1950	$SS_{e\,y}$ = 32,5841	f_e = n – b = 418	$SP_{e\,xy}$ = 10,1466	**$MS_{e\,x}$ = 0,043528**	**$MS_{e\,y}$ = 0,077952**	$MP_{e\,xy}$ = 0,024274
- celková	$SS_{c\,x}$ = 29,1034	$SS_{c\,y}$ = 48,7923	f_c = n – 1 = 486	$SP_{c\,xy}$ = 16,8210	-	-	-

Protože ve skupinách byly různé počty pozorování, tak byl vypočítán vážený počet pozorování n_0 = 6,6494.

$$\sigma^2_{g_x} = \frac{MS_{Ox} - MS_{ex}}{n_0} = 0{,}00852$$

$$\sigma^2_{g_y} = \frac{MS_{Oy} - MS_{ey}}{n_0} = 0{,}01936$$

$$\mathrm{cov}_{g_{xy}} = \frac{\mathrm{cov}_{Oxy} - \mathrm{cov}_{exy}}{n_0} = \frac{MP_{Oxy} - MP_{exy}}{n_0} = 0{,}00731$$

$$\hat{r}_{G_xG_y} = \frac{\mathrm{cov}_{g\,xy}}{\sqrt{\sigma^2_{g_x}.\sigma^2_{g_y}}} = 0{,}569$$

$$\hat{r}_{E_xE_y} = \frac{\mathrm{cov}_{e\,xy} - 3\mathrm{cov}_{g\,xy}}{\sqrt{\sigma^2_{e_x} - 3.\sigma^2_{g_x}}\sqrt{\sigma^2_{e_y} - 3.\sigma^2_{g_y}}} = \frac{MP_{e\,xy} - 3\mathrm{cov}_{g\,xy}}{\sqrt{MS_{e_x} - 3.\sigma^2_{g_x}}\sqrt{MS_{e_y} - 3.\sigma^2_{g_y}}} = 0{,}1239$$

$$\hat{r}_{P_xP_y} = \frac{\mathrm{cov}_{c\,xy}}{\sqrt{\sigma^2_{c_x}.\sigma^2_{c_y}}} = \frac{SP_{c\,xy}}{\sqrt{SS_{c_x}.SS_{c_y}}} = 0{,}4464$$

Správnost výpočtu odhadů korelací je možno **ověřit rozkladem fenotypové korelace**

$$h^2_x = 4.\rho = 4\frac{\sigma^2_{g_x}}{\sigma^2_{g_x} + \sigma^2_{e_x}} = 4\frac{\sigma^2_{g_x}}{\sigma^2_{P_x}} = 0{,}655 \qquad h_x = \sqrt{h^2_x} = 0{,}8093 \qquad e_x = \sqrt{e^2_x} = \sqrt{1 - h^2_x} = 0{,}5874$$

$$h^2_y = 4.\rho = 4\frac{\sigma^2_{g_y}}{\sigma^2_{g_y} + \sigma^2_{e_y}} = 4\frac{\sigma^2_{g_Y}}{\sigma^2_{P_y}} = 0{,}7957 \qquad h_y = \sqrt{h^2_y} = 0{,}8921 \qquad e_y = \sqrt{e^2_y} = \sqrt{1 - h^2_y} = 0{,}4520$$

$$\hat{r}_{P_xP_y} = h_x\,.\,h_y\,.\,r_{G_xG_y} + e_x\,.\,e_y\,.\,r_{E_xE_y} = 0{,}4438$$

3. Z výsledku selekce

- Realizovaná genetická korelace 2 vlastností na základě výsledků selekce, podle výše genetického zisku
  - selekční pokus
  - selekce v běžných podmínkách

a) Selekční pokus

- $\Delta G_{X/YS}$ genetický zisk v selektované části populace
- $\Delta G_{X/Y}$ genetický zisk v neselektované části populace

$$r_{G_X G_Y} = \frac{\Delta G_{XS}}{\Delta G_Y} \cdot \frac{\Delta G_{YS}}{\Delta G_X}$$

b) Selekce v běžných podmínkách

- provádíme selekci dle jednoho znaku
- stanovíme dosažený genetický zisk pro oba znaky
- známe-li heritabilitu obou znaků, pak při selekci podle X

$$r_{G_X G_Y} = \frac{\Delta G_Y}{i_x h_X \sigma_{G_Y}} = \frac{\Delta G_Y}{i_x h_X h_Y \sigma_{P_Y}}$$

Korelovaný selekční efekt

- o kolik se v generaci po selekci na znak X změní znak Y, který nebyl předmětem selekce
- vlastnosti ve vazbě (skot: mléčné složky)

a) *přímá selekce* – selekční zisk za jednu generaci lze získat za pomoci vztahu:

$$\Delta G_x = i.h_x^2.\sigma_{Px} \qquad \Delta G_y = i.h_y^2.\sigma_{Py}$$

b) ***nepřímá selekce*** – selekční zisk za jednu generaci lze získat pomocí vztahu **korelovaného selekčního zisku (efektu).** Při selekci na jeden znak, lze ze znalosti genetické korelace s druhým znakem predikovat o kolik se změní druhý znak, na který neprobíhala selekce.

$$_{kor.}\Delta G_y = \Delta G_x \,.\, b_{G\,xy} \qquad b_{G\,yx} = r_{GxGy}.\frac{\sigma_{Gy}}{\sigma_{Gx}}$$

$$_{kor.}\Delta G_x = i.h_x.h_y.r_{GxGy}.\sigma_{Px}$$

$$_{kor.}\Delta G_x = \Delta G_y \,.\, b_{G\,xy} \qquad b_{G\,xy} = r_{GxGy}.\frac{\sigma_{Gx}}{\sigma_{Gy}}$$

$$_{kor.}\Delta G_y = i.h_x.h_y.r_{GxGy}.\sigma_{Py}$$

Příklady genetických korelací

skot	r_P	r_G	r_E
produkce mléka : % tuku	-0,26	-0,38	-0,18
prod. mléka v 1 : 2 laktaci	0,40	0,75	0,26
prasata			
přírůstek : hřbetní tuk	0,00	0,13	-0,18
přírůstek : výkrmnost	0,66	0,69	0,64
drůbež			
hm. těla : hm. vejce	0,33	0,42	0,23
hm. těla : snáška	0,01	-0,17	0,08