

Statistika & STR typing



INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

STR lokus

Mám typický multialelický autozomální STR lokus „A“

- známo je celkem 9 alel (1-9)
- každá z těchto alel má nějakou frekvenci ve studované populaci

1	2	3	4	5	6	7	8	9
0,02	0,08	0,11	0,24	0,09	0,21	0,15	0,09	0,01

Jak ale vlastně zjistím frekvenci alel v populaci?

→ musím udělat populační studii



Populační studie

„Vezmi reprezentativní vzorek dostatečně velký, aby tvá data byla relevantní, a genotypuj ho.“

Co to je „REPREZENTATIVNÍ VZOREK (VÝBĚR)“?

→ jen sladké mámení... nikdy nemáme plně reprezentativní vzorek dané populace, můžeme se mu jen blížit

Jaké budou hlavní požadavky na reprezentativnost?

Co to je „DOSTATEČNĚ VELKÝ“?

→ záleží na nás, s jakou přesností chceme získat údaj o alelických frekvencích – čím větší vzorek, tím přesnější data (tzn. menší SD) a tím menší problém s řídkými tabulkami (viz později)

Lze dosáhnout 100% přesnosti?



Populační studie

každou osobu z výběru genotypuji – tzn. stanovím sestavu jejích alel

stanovím četnost výskytu každé alely

- **POZOR!** – heterozygot 2/3 – započítávám 1x alelu **2** a 1x alelu **3**
- homozygot 3/3 – započítávám 2x alelu **3**

pak stanovím frekvence jednotlivých alel – pro alelu X platí:

$$P(X) = n(X)/N$$

1	2	3	4	5	6	7	8	9
8	32	44	48	36	42	60	36	4
400	400	400	400	400	400	400	400	400
0,02	0,08	0,11	0,24	0,09	0,21	0,15	0,09	0,01



Populační studie

...tolik stanovení frekvence alel. Toto mohu udělat u jakéhokoli polymorfismu, mohu to udělat u libovolného genetického markeru, biochemického markeru, morfologického markeru, mohu to udělat i pro polymorfismus barvy auta atd.

Ale to je jen začátek!

Už známe frekvence, ale teď musíme ověřit, jak se alely chovají, abychom mohli dále hodnotit

Co je Svatým Grálem populační genetiky?

HW - equilibrium



HW - equilibrium

Co jsou základní předpoklady HWE?

- nekonečná velikost populace
 - žádné mutace
 - žádná selekce
- náhodný pohlavní výběr



HW - equilibrium

NEKONEČNÁ VELIKOST POPULACE

Nekonečnost populace je nezbytným předpokladem, který eliminuje náhodné změny frekvencí alel

Jak takovým náhodným změnám říkáme?

genetický drift

Každá populace je konečná = je tvořena konečným počtem jedinců, nicméně má-li alespoň rozumnou velikost, je vliv driftu zanedbatelný a HWE poskytuje dobré odhady



HW - equilibrium

NEEXISTENCE MUTACÍ

Neexistence mutací je dalším nezbytným předpokladem, který eliminuje náhodné změny frekvencí alel

Mohou existovat polymorfní STR lokusy, které nemutují a nikdy nemutovaly?

Opět platí, že dochází-li k mutacím pouze v rozumné míře, je jejich vliv zanedbatelný a HWE poskytuje dobré odhady



HW - equilibrium

NULOVÝ SELEKČNÍ TLAK

Neexistence selekce je nezbytným předpokladem, který eliminuje nenáhodné změny frekvencí alel

Jednotlivé alely samy o sobě nesmějí přinášet nějakou selekční výhodu respektive nevýhodu

Existují STR lokusy, které selekční (ne)výhodu přinášejí?

Pokud samy alely nejsou předmětem selekce, nesmí se tímto předmětem selekce stát ani nepřímo – nesmí existovat genová vazba mezi daným STR lokusem a nějakým funkčním genem, který je selektován

Jak nazýváme tuto „nepřímou selekci“?

genetický draft = selekce na pozadí



HW - equilibrium

NÁHODNÝ POHLAVNÍ VÝBĚR

Neexistence závislého výběru partnera je dalším nezbytným předpokladem, který eliminuje nenáhodné změny frekvencí alel

Jednotlivé alely samy o sobě nesmějí být předmětem pohlavního výběru

Pokud samy alely nejsou předmětem PV, nesmí se tímto předmětem PV stát ani nepřímo – nesmí existovat genová vazba mezi daným STR lokusem a nějakým funkčním genem, který je PV

*Blondýna, 30 let, příslušnice mitotypu Uršula, by se ráda seznámila se sportovně založeným nositelem Y-haplotypu J2.
Zn. Společná evoluce.*

HW - equilibrium

VÝSLEDEK

V populaci, která je z hlediska sledovaného lokusu v HW rovnováze, se vše řídí jednoduchou kombinatorikou a teorií pravděpodobnosti – jako „házení kostkami“

Frekvence jednotlivých genotypů lze vidět jako polynom n-tého stupně, kde n je počet alel

u 2 alel

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

u 3 alel

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



HW - equilibrium

OBECNĚ BUDE PLATIT, ŽE:

frekvence homozygotů P/P je p^2

frekvence heterozygotů P/Q je $2pq$

počet osob, které nemají žádnou alelu P, je $(1-p)^2$

počet osob, které mají alespoň jednu alelu P, je $1-(1-p)^2$

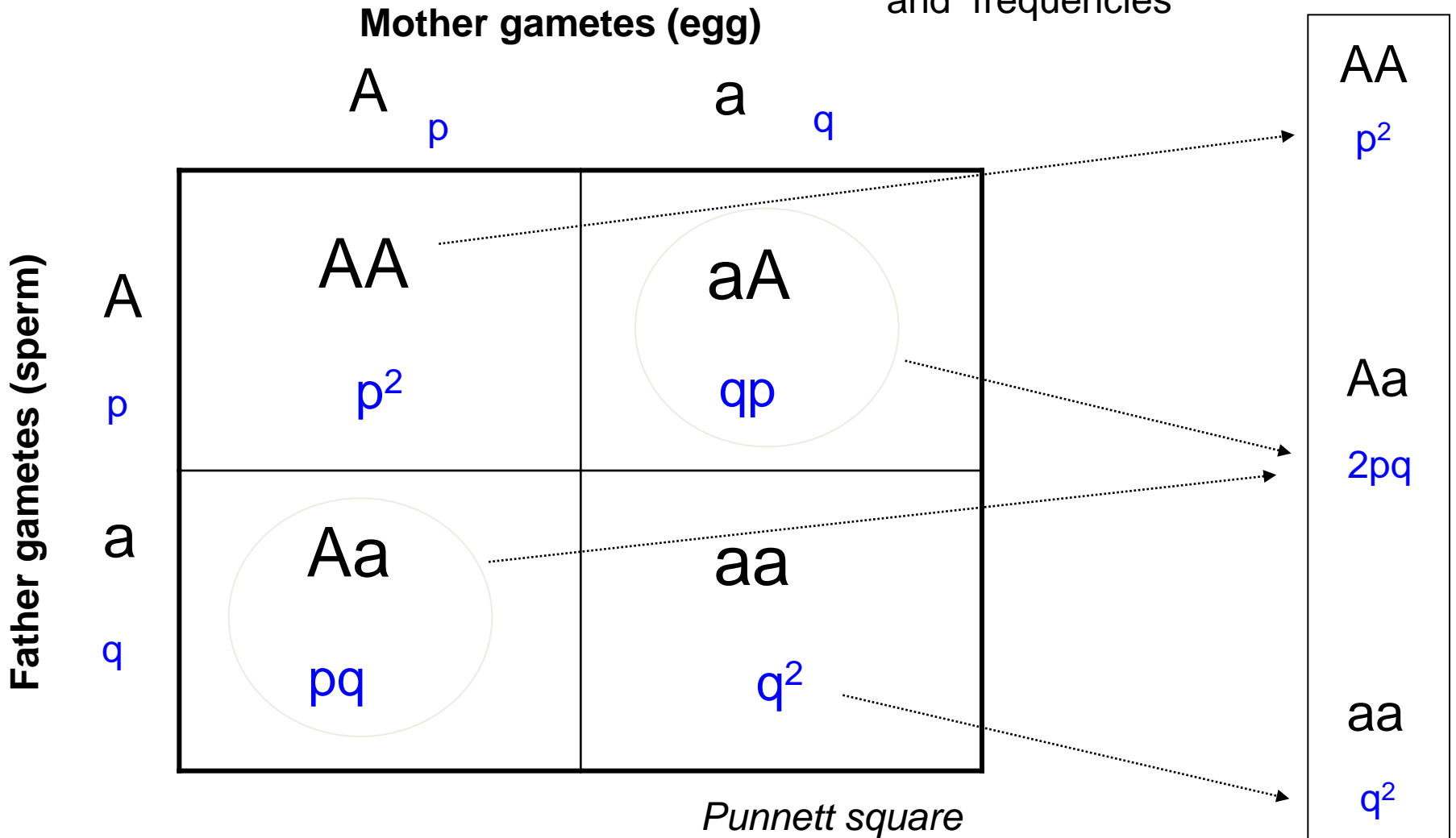
počet osob, které nemají žádnou alelu P ani Q, je $(1-p-q)^2$

počet osob, které mají alespoň jednu alelu P nebo Q, je $1-(1-p-q)^2$

**K odvození slouží
MAGICKÝ PUNNETTŮV ČTVEREC**



Resulting genotype combinations and frequencies



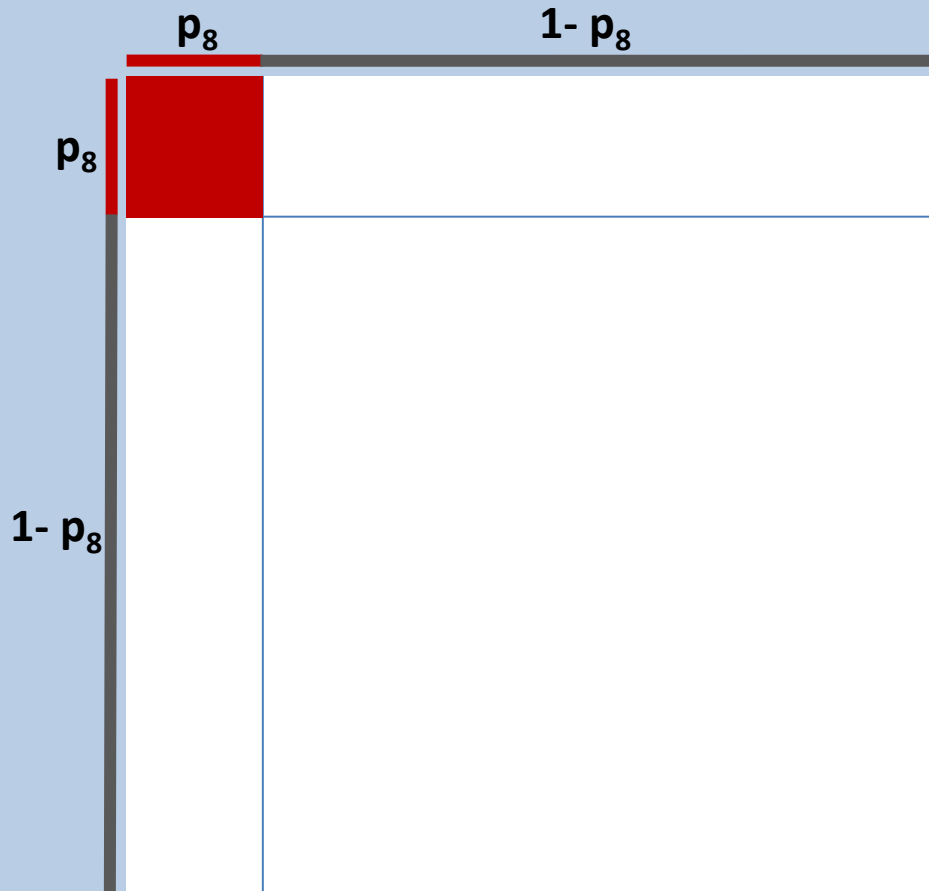
Freq (A) = p

Freq (a) = q

$p + q = 1$

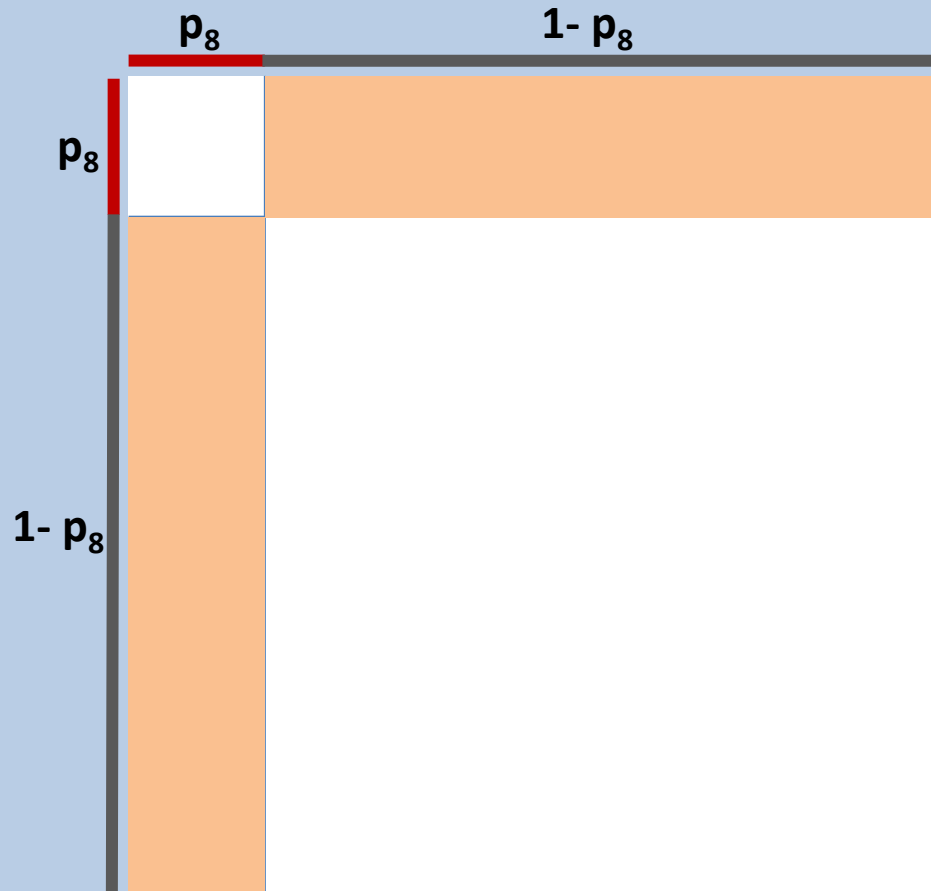
$(p + q)^2 = p^2 + 2pq + q^2$

Maternální alela

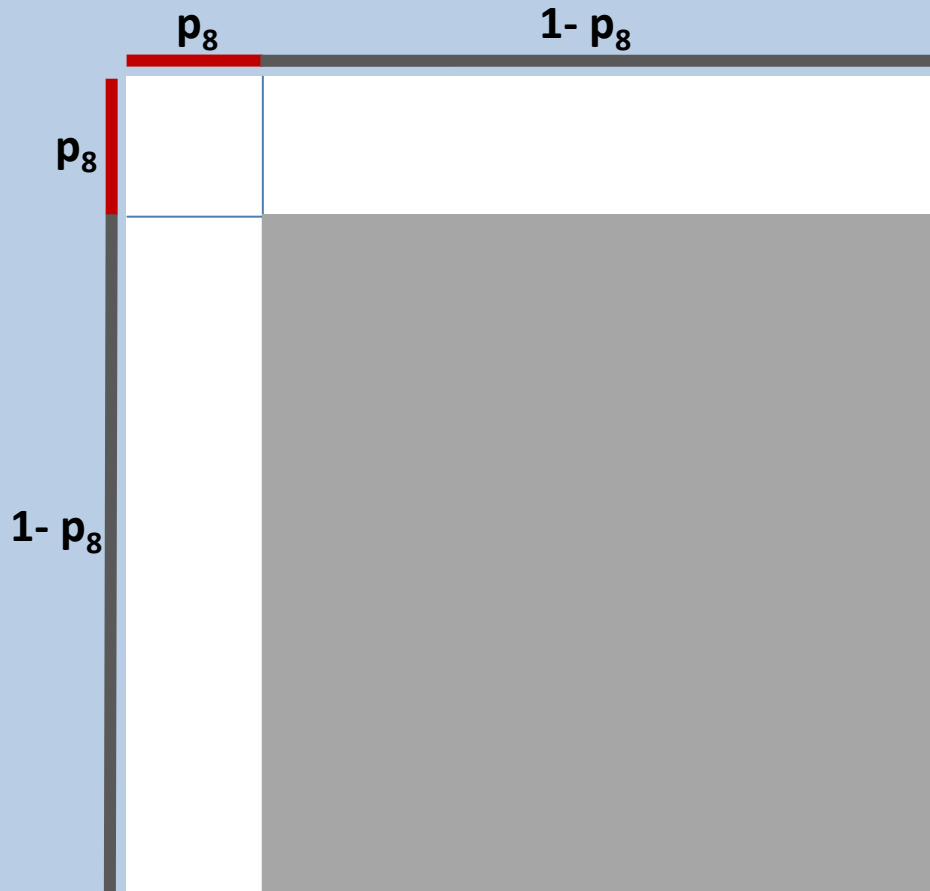


Maternální alela

Paternální alela



Maternální alela

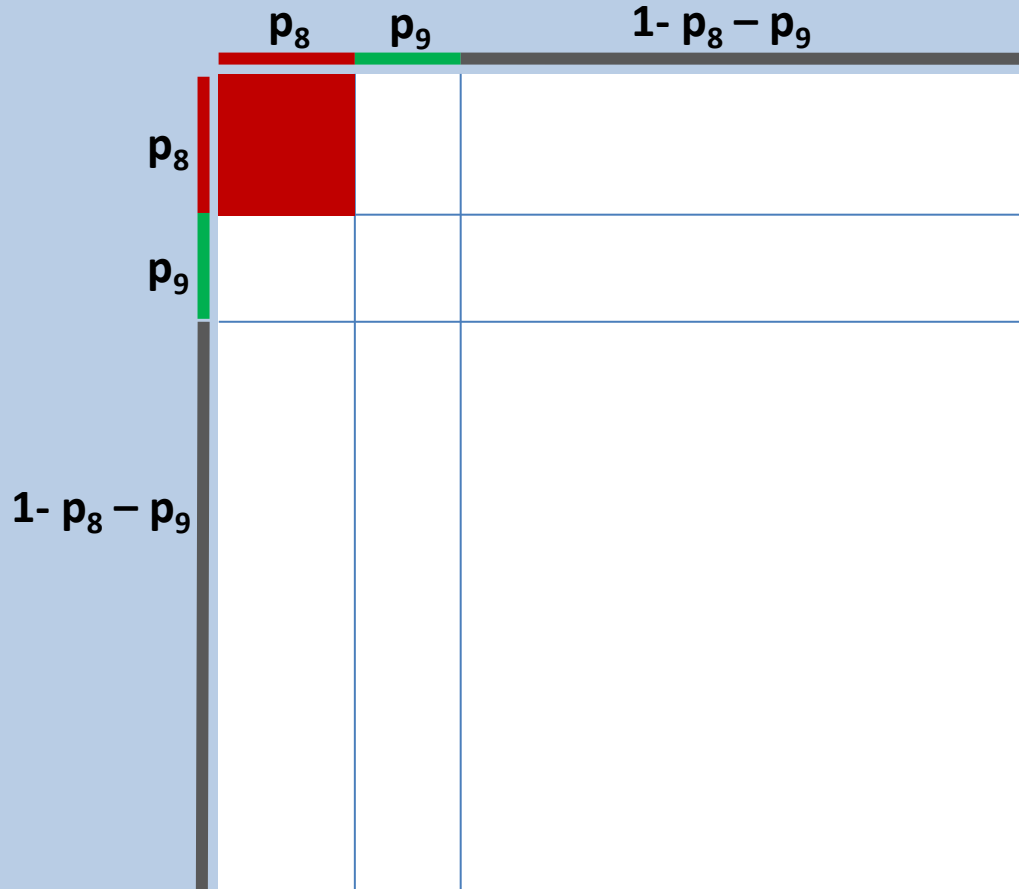


8

9

Maternální alela

Paternální alela

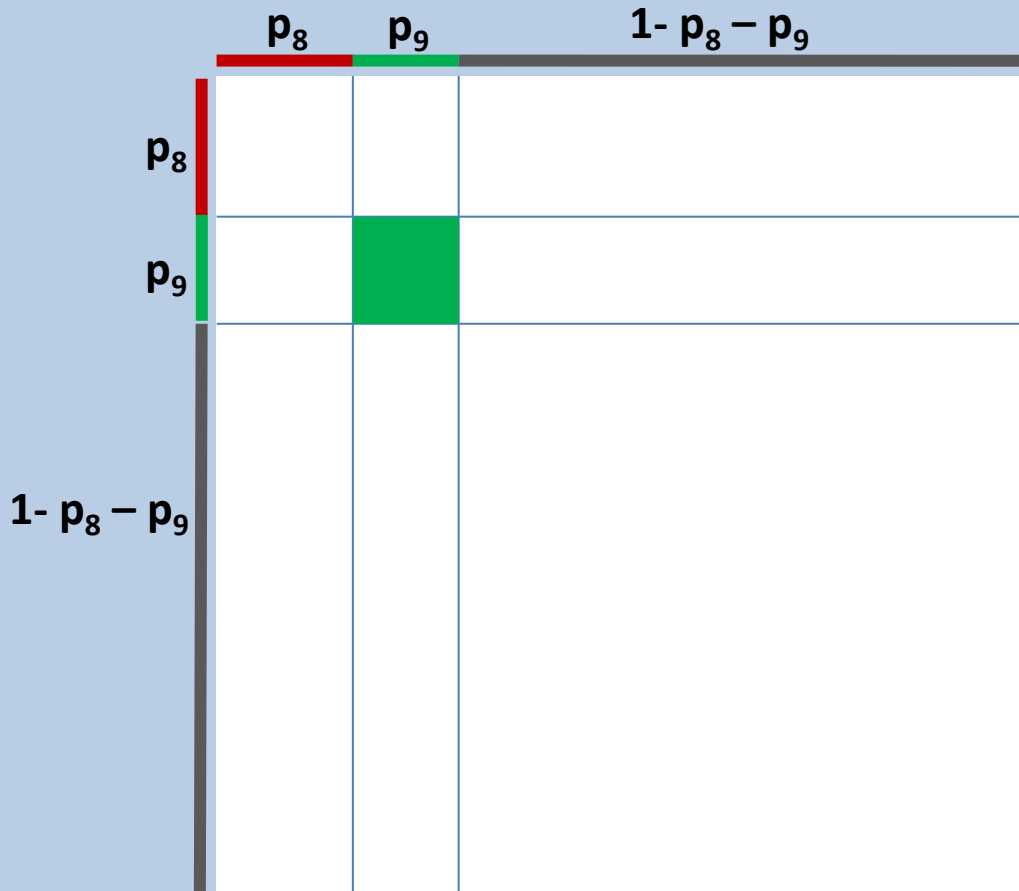


8

9

Maternální alela

Paternální alela

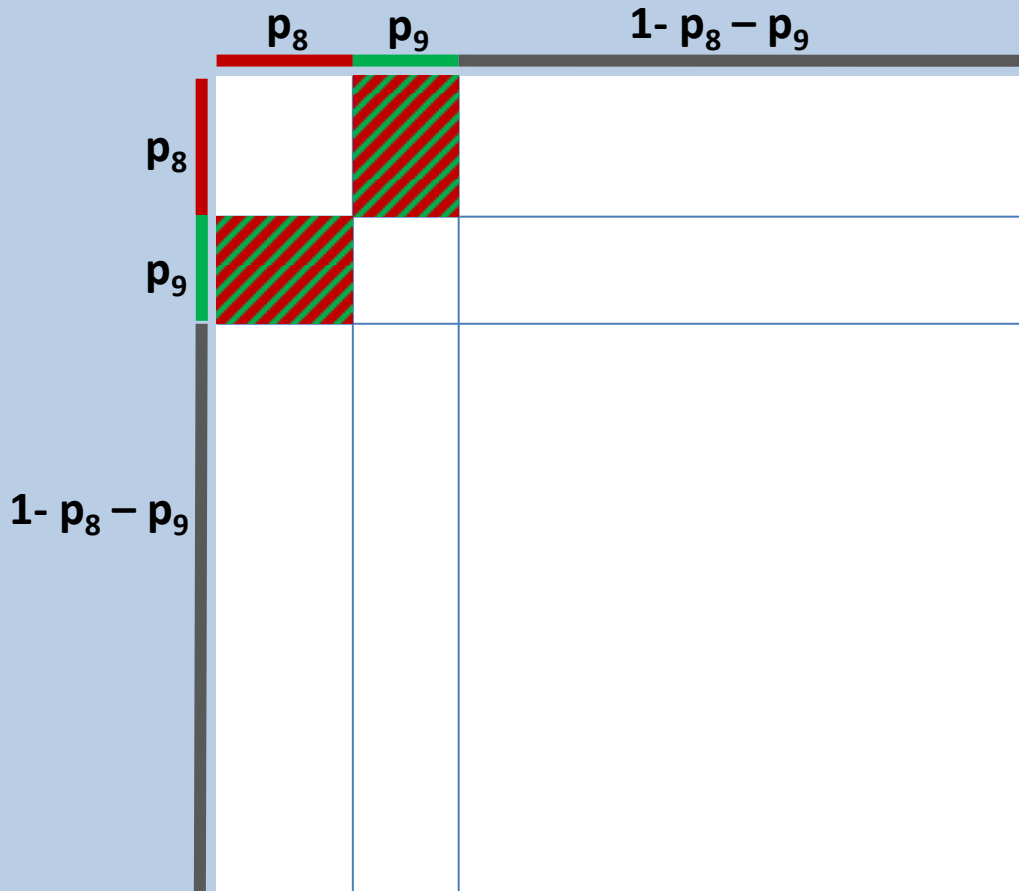


8

9

Maternální alela

Paternální alela

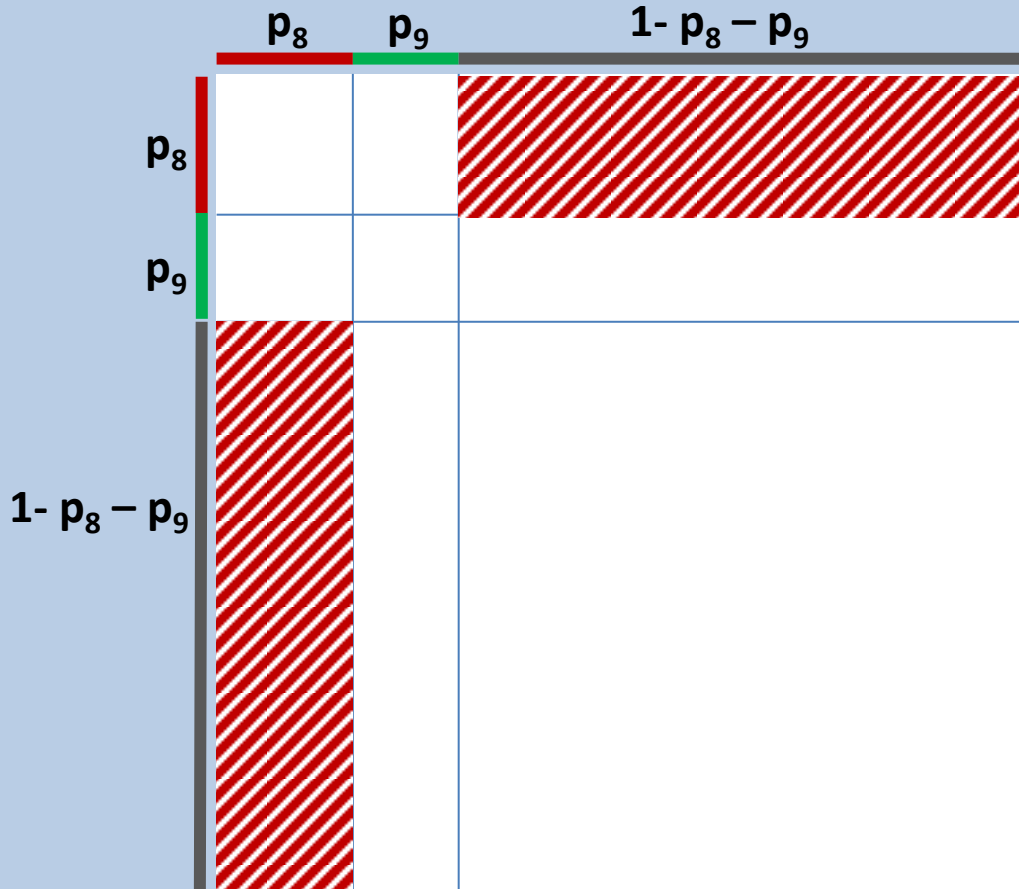


8

9

Maternální alela

Paternální alela

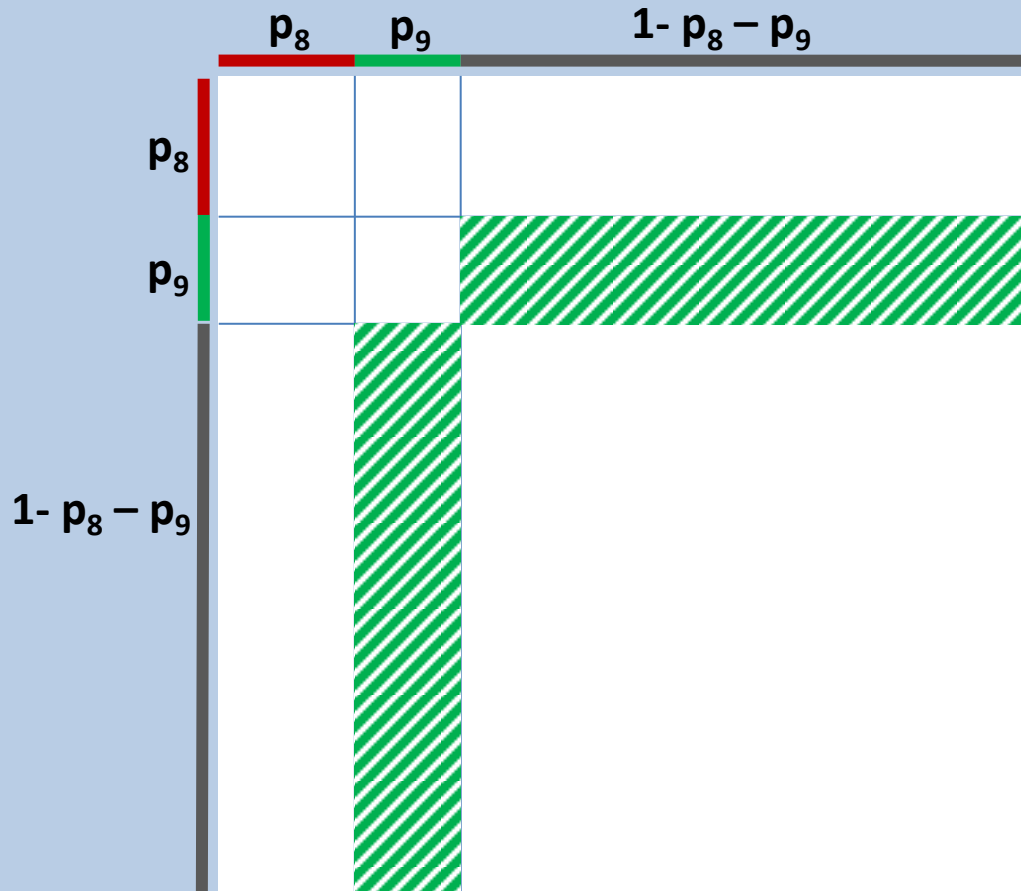


8

9

Maternální alela

Paternální alela

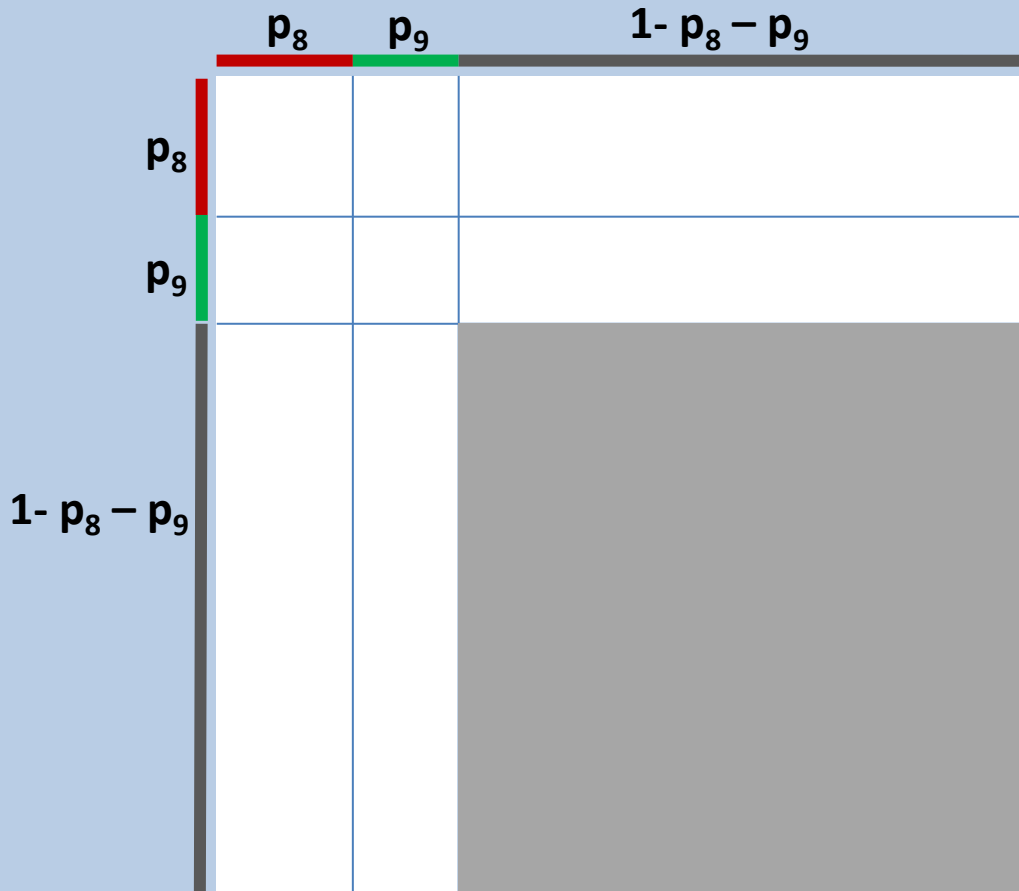


8

9

Maternální alela

Paternální alela



HW - equilibrium

Multialelický autozomální STR lokus - několik jednoduchých úloh:

1	2	3	4	5	6	7	8	9
0,02	0,08	0,11	0,24	0,09	0,21	0,15	0,09	0,01

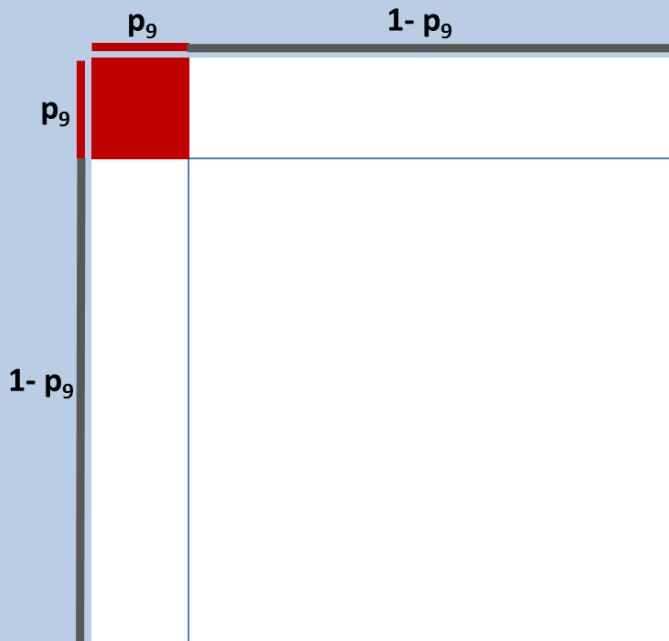
1. kolik % je v populaci osob s genotypem 9/9 ?
2. kolik % osob má alespoň jednu alelu 4 ?
3. kolik % osob má právě jednu alelu 8 ?
4. kolik % spermií v celé populaci nese alelu 7 ?



HW - equilibrium

1. kolik % je v populaci osob s genotypem 9/9 ?

1	2	3	4	5	6	7	8	9
0,02	0,08	0,11	0,24	0,09	0,21	0,15	0,09	0,01



$$P = p_9 \times p_9 = p_9^2$$

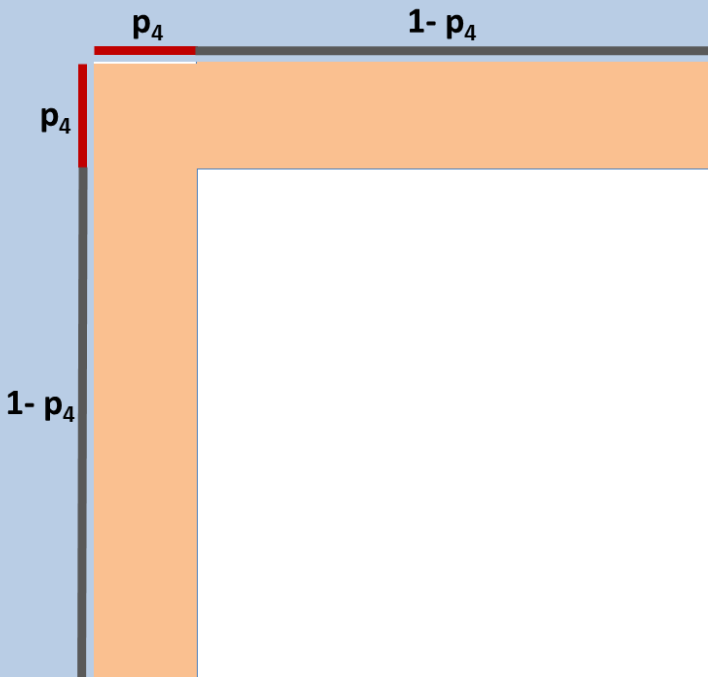
$$P = 0,01 \times 0,01 = 0,0001$$

0,01%

HW - equilibrium

2. kolik % osob má alespoň jednu alelu 4 ?

1	2	3	4	5	6	7	8	9
0,02	0,08	0,11	0,24	0,09	0,21	0,15	0,09	0,01



$$P = (p_4 \times p_4) + 2(p_4 \times 1-p_4)$$

$$P = 2p_4 - (p_4 \times p_4)$$

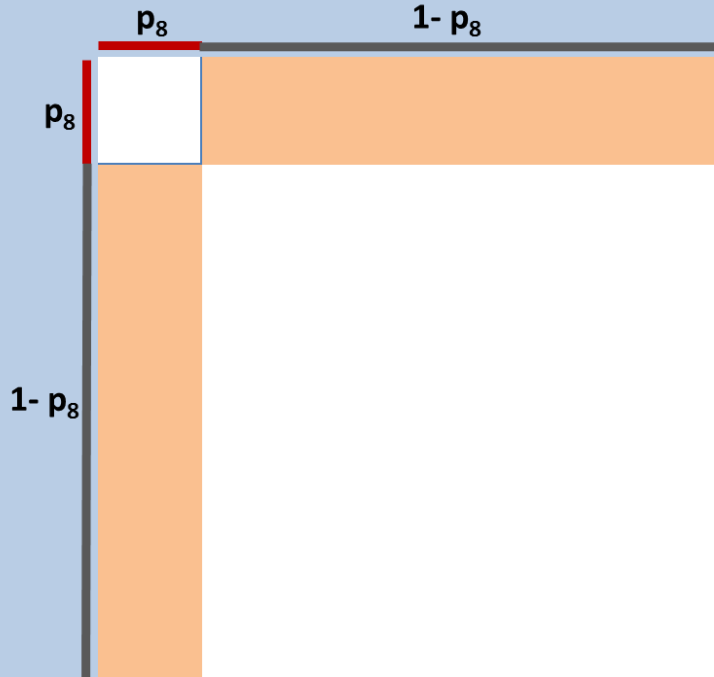
$$P = 1 - (1 - p_4)^2$$

$$P = 0,4224 \quad 42,24\%$$

HW - equilibrium

3. kolik % osob má právě jednu alelu 8 ?

1	2	3	4	5	6	7	8	9
0,02	0,08	0,11	0,24	0,09	0,21	0,15	0,09	0,01



$$P = 2 \times (1 - p_8) \times p_8 = 2(p_8 - p_8^2)$$

$$P = 0,1638$$

16,38%

HW - equilibrium

4. kolik % spermií v celé populaci nese alelu 7 ?

1	2	3	4	5	6	7	8	9
0,02	0,08	0,11	0,24	0,09	0,21	0,15	0,09	0,01

$$P = p_7 = 0,15 \quad 15\%$$



HW - equilibrium

OVĚŘENÍ

PLATÍ-LI V POPULACI HWE, MUSÍ PLATIT:

frekvence homozygotů P/P je p^2

frekvence heterozygotů P/Q je $2pq$

pro všechny genotypy dané populace

Porovnáám, jak moc se tyto očekávané frekvence blíží
frekvencím zjištěným v populační studii



HW - equilibrium

OVĚŘENÍ

mám lokus A s alelami **1** ($p_1=0,1$), **2** ($p_2=0,3$) a **3** ($p_3=0,6$)

genotyp	1/1	2/2	3/3	1/2	1/3	2/3
Očekávaná četnost	0,01	0,09	0,36	0,06	0,12	0,36
Pozorovaná četnost	0,02	0,06	0,39	0,06	0,13	0,34
rozdíl	0,01	0,03	0,03	0	0,01	0,02

musím posoudit, zda tyto rozdíly mezi očekávaným a pozorovaným mohou být pouze dílem náhody, či zda jsou signifikantní a ukazují, že HWE neplatí

jak? - χ^2



HW - equilibrium

Důsledky pro forenzní výpočty

pokud zjistím, že pro lokus HWE neplatí, nemohu počítat frekvence jednotlivých genotypů prostou kombinatorikou; mohu

- a) užívat výrazně komplikovanější výpočty s koeficienty
- b) vyloučit lokus z testování

pokud zjistím, že pro lokus HWE platí, pracuji při hodnocení důkazů s frekvencemi tak, jak je počítá HWE

