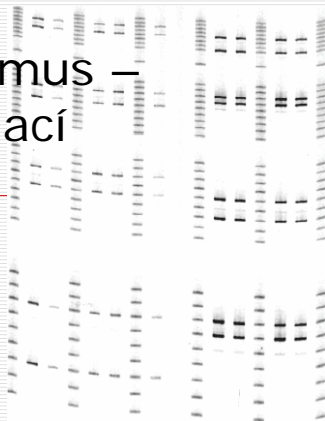


Genetický polymorfismus – popis struktury populací

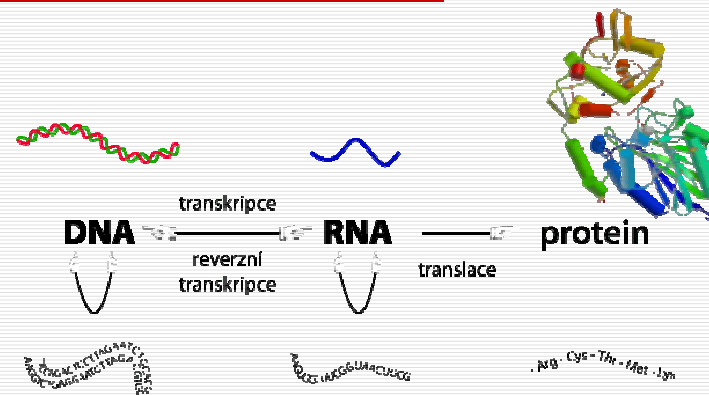


Dr. Ing. Urban Tomáš
 ÚSTAV GENETIKY MZLU Brno

urban@mendelu.cz
<http://www.mendelu.cz/af/genetika/>

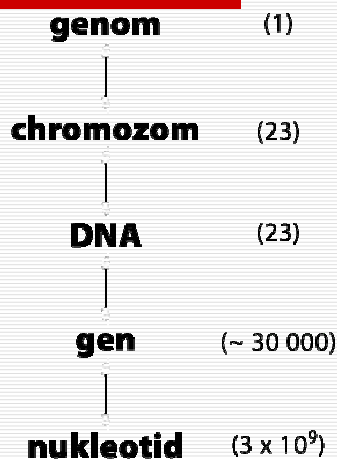
Seminář doktorského grantu 523/03/H076 :
Molekulární data v genetice populací a hodnocení diverzity
 ZF JU České Budějovice
 25. – 26. 10. 2004

Tok GI v buňce



Dědičnost – tok informací mezi buňkami (během buněčného dělení) a přes generace (pomocí meiózy)

Organizace genetického materiálu



Molekulární technologie a automatizace spolu se sofistikovanými výpočetními programy umožňují určit sekvence nukleotidů v genomu.

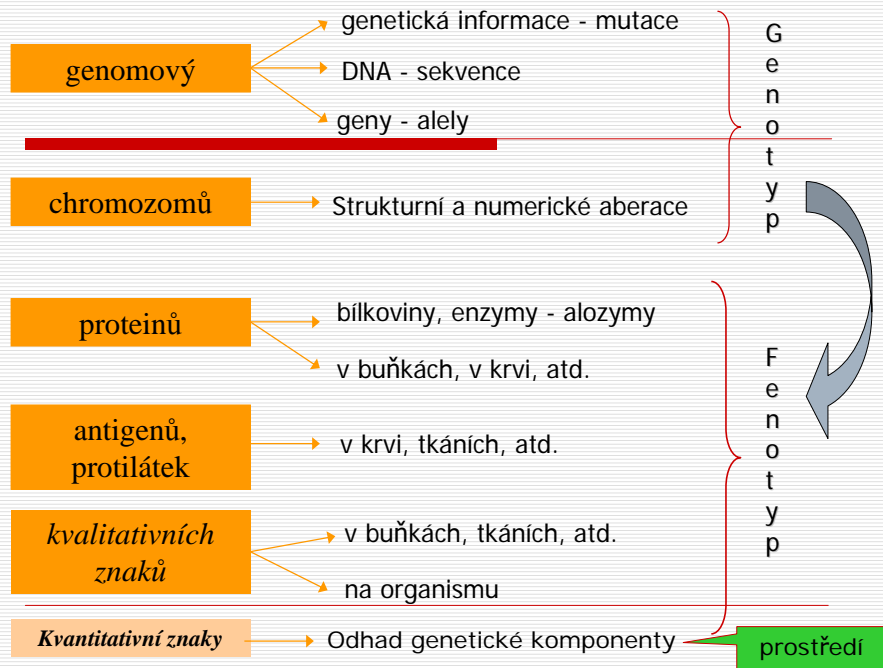
Definice polymorfismu

- poly = mnoho; morfé = tvar
- *Polymorfismus* = výskyt dvou nebo více variant (tvarů) určitého znaku v populaci daného druhu

příklad polymorfismu znaku – postavení uší u prasete



znak kvalitativní

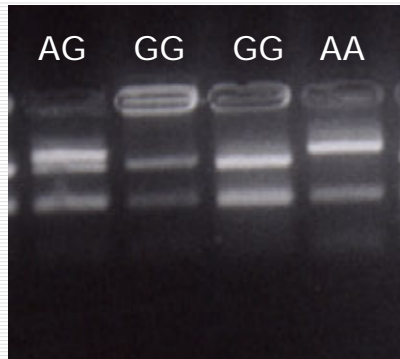


Proč využívat molekulární genetické markery ?

- Molekulární data poskytují genetickou informaci (nejvíce informativní)
 - Možno testovat všechny biologické druhy
 - Molekulární metody zpřístupňují téměř neomezený zdroj genetické variability (3 mld bp)
 - Kompletní DNA sekvence je již u mnoha modelových druhů (ale pro populační a evoluční studie je taková informace nadbytečná)
 - 2003 – 1,4 mil SNP v databázi lidského genomu
 - Mol. data poskytují společné měřítko pro měření divergence mezi druhy, plemeny, ...
 - Molekulární přístupy se mění a rozvíjejí
-
- Proč nevyužívat molekulárně genetické markery?
 - lidí, náklady

PCR - RFLP přístup studia polymorfizmu

- Restriction fragment length polymorphism of the M307 polymorphic site in the *FUT1* gene



Horák, Urban, Dvořák (2004)

Genomový polymorfismus

podstata: - bodové mutace (záměna, delece, atd. bazí)
- repetice sekvencí (kratších nebo delších)

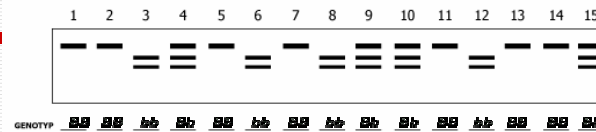
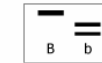
zkoumá: - molekulární genetika
- genomika

metody zjišťování: PCR, RFLP, DGGE, TGGE, SSCP, sekvenování

- krátký přehled Mgr. Cívánová

Genetický polymorfismus jako míra genetické variance v populacích

- Lokus je polymorfní
 - Četnost nejrozšířenější alely není větší než 0,95
 - Nebo 0,99



? absolutní frekvence genotypů! $BB = 7$ $Bb = 4$ $bb = 4$

? relativní frekvence genotypů!

$$d = \frac{D}{N} \qquad h = \frac{H}{N} \qquad r = \frac{R}{N}$$

$$f(BB) = 0,4667 \qquad f(Bb) = 0,2667 \qquad f(bb) = 0,2667$$

? relativní frekvence alel!

$$p = d + \frac{1}{2}h \qquad f(B) = 0,600 \qquad q = r + \frac{1}{2}h \qquad f(b) = 0,400$$

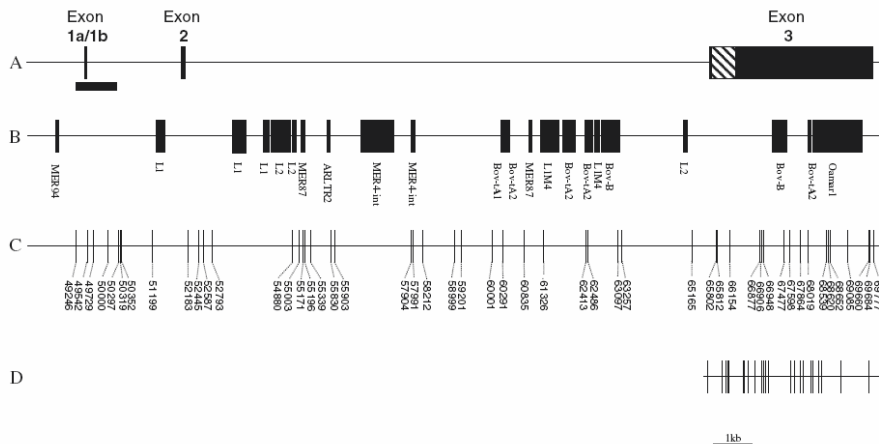


Figure 1 Diagram illustrating the genomic architecture of the bovine *PRNP* gene. In (A) the three exons are depicted by solid black boxes and the ORF is shown as a hatched box within exon 3. The solid line underneath exon 1 marks the promoter region identified by Inoue *et al.* (1997). (B) Indicates location of repeat elements. (C) Shows the positions of the variants identified here within GenBank Accession number AJ298878. (D) Indicates the positions of the variants within the exon 3 region of ovine *PRNP* but the positions are not given for clarity.

Sequence variation in the bovine and ovine *PRNP* genes

D. Hills, J. Schlaepfer, S. Comincini, I. MacLean, G. Dolf, L. Ferretti, I. Olsaker, J. L. Williams
Animal Genetics 2003 34:3 p. 183

Formy variability v bovinním PRNP genu

- 82 % jsou SNP
 - Substitute
 - 79 % tranzice (A > G; C > T)
 - 21 % transverze (A, G > C, T)

Table 4 Nucleotide diversity data in cattle.

Gene	Length (bp)	Number of polymorphisms
LEP	1788	20
APP	458	24
GRO3	829	3
GRO1	456	2
GROX	912	6
IL8R	523	4
SDF1	605	6
IFNG	892	7
TNFA	639	3
IL8	662	7
TRA	2110	12
PRNP	21287	51
PRNP (exon 3)	4517	19

Table 2 Sequence variants identified in bovine PRNP.

Position ¹	Variant	Frequency ²			
49246	A.. > G	0.50	60291	T.. > C	0.88
49542	G.. > A	0.92	60835	A.. > G	0.73
49729	ins GGGGGCCCGCGGC	0.54	61326	A.. > G	0.92
50000	G.. > A	0.58	62413	G.. > A	0.65
50297	G.. > A	0.69	62486	G.. > A	0.81
50319	G.. > A	0.69	63097	G.. > T	0.92
50352	G.. > A	0.65	63257	A.. > G	0.92
51199	G.. > A	0.53	65165	T.. > C	0.57
52183	G.. > A	0.65	65802	del GCTGGGGCCAGCCTCATGGAGGTG	0.88
52445	C.. > A	0.65	65812	G.. > A	0.60
52587	G.. > A	0.88	66154	C.. > T	0.85
52793	T.. > C	0.88	66877	C.. > T	0.50
54863	TTGTACATT.. > GTAACITTTGTCTCCATTTCTTTTATGATATATGGTCAAGTTTTCTATTTAA AAATTG	0.62	66906	A.. > G	0.92
55003	T.. > C	0.65	66948	T.. > C	0.92
55171	G.. > A	0.76	67477	T.. > G	0.92
55196	A.. > G	0.84	67598	G.. > A	0.73
55339	T.. > A	0.92	67864	G.. > A	0.62
55830	G.. > A	0.92	68019	(TTTTTGTAAAGGTAC) ₂ ..>(TTTTTGTAAAGGTAC) ₃	0.92
55903	CAC.. > GAACTTGTTA	0.65	68408	del AAGAA	
57904	A.. > G	0.62	68539	T.. > A	0.77
57991	del CGA	0.73	68620	T.. > G	1
58212	T.. > A	0.69	68652	C.. > T	0.77
58999	G.. > A	0.54	69085	G.. > A	1
59201	G.. > C	0.65	69660	T.. > G	0.73
60001	(A) ₅ .. > (A) ₇	0.88	69684	(T) ₄ .. > (T) ₅	0.96
			69774	G.. > A	0.50

¹Position in reference sequence AJ298878.
²Frequency of most common allele in Holstein-Friesian samples only.

Hardy-Weinbergova rovnováha a vazbová rovnováha

- V populacích existuje genetická rovnováha – pro studium změn frekvencí
- 2 typy rovnováhy
 - Hardy –Weinbergova
 - vazbová

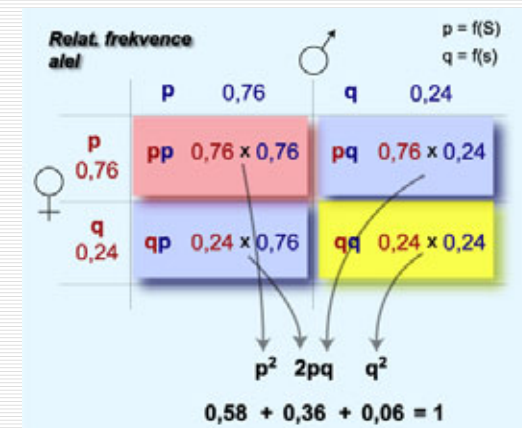
Odvození H.-W. principu

	Vajíčka ♀	
♂ Spermie	p ~	q ~
p ~	p ² ~ AA	pq ~ Aa
q ~	pq ~ Aa	q ² ~ aa

$$(A + a)^2 = 1AA + 2Aa + 1aa = N$$

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

Tvorba nové generace



Jedna frekvence alel – více frekvencí genotypů

☐ Pouze v 1 případě je populace v genetické rovnováze!

$$p = d + \frac{1}{2}h$$

$$p = \sqrt{p^2}$$

p^2 (AA)	$2pq$ (Aa)	q^2 (aa)	p (A)	q (a)
0,60	0,40	0,00	0,80	0,20
0,61	0,38	0,01	0,80	0,20
0,64	0,32	0,04	0,80	0,20
0,70	0,20	0,10	0,80	0,20
0,75	0,10	0,15	0,80	0,20
0,80	0,00	0,20	0,80	0,20

➤ Pouze je-li populace v genetické rovnováze, lze odvodit frekvenci genotypů z frekvencí alel dle

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

Testování HWE

☐ χ^2 test (Pearson)

Vícealelové lokusy

☐ exact test (Fisher)

☐ likelihood-ratio test (χ^2)

χ^2 test dobré shody

H_0 –
lokus je
v HWE

H_A –
lokus
není v
HWE

Testování genetické rovnováhy		χ^2 test (test dobré shody)		
	SS (p^2)	FS ($2pq$)	FF (q^2)	Σ
Pozorované frekvence (absolutní)	66	20	14	100
Očekávané frekvence (relativní)	0,5776	0,3648	0,0576	1,00
Očekávané frekvence (absolutní)	57,76	36,48	5,76	100
$d = (P - O)$	8,24	16,48	8,24	
$\chi^2 = \frac{(P - O)^2}{O}$	1,176	7,445	11,788	20,408

$$df = 3 - 1 - 1 = 1$$

$$\chi^2_{(1, 0,05)} = 3,84$$

$$\chi^2_{(1, 0,01)} = 6,64$$

Internetové aplikace - HWE

- ☐ http://wbiomed.curtin.edu.au/genepop/genepop_op1.html
- ☐ Tests for deviation from Hardy-Weinberg equilibrium and tests for association (SNP) <http://ihg.gsf.de/cgi-bin/hw/hwa1.pl>
- ☐ HARDY (for large numbers of alleles (microsatellites) <http://www.stat.washington.edu/thompson/Genepi/Hardy.shtml>
- ☐ <http://pritch.bsd.uchicago.edu/software.html>
- ☐ ...

Měření vazbové nerovnováhy

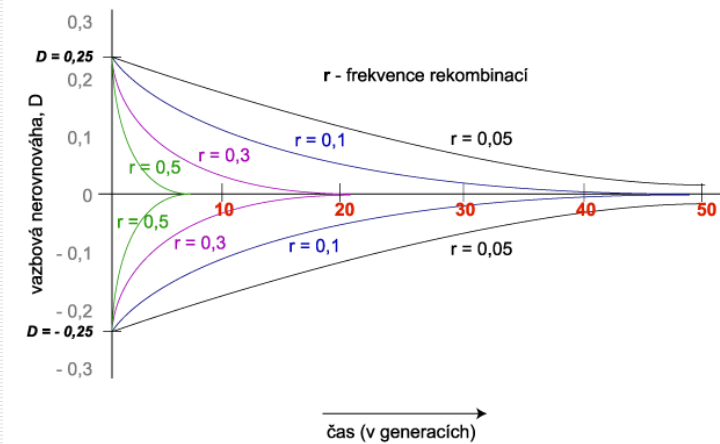
- Nenáhodné asociace mezi chromozomálními lokusy v populaci
- Alely jednoho lokusu nemusí být v asociaci s alelami druhého lokusu
- Vazbová rovnováha: náhodné alelické asociace mezi alelami na jakémkoliv lokusu
- Dosažení rovnováhy je však pomalé, závisí na rekombinaci – r (pravděpodobnost, že gameta přenášená jedincem je rekombinantní)

	B (p_1)	b (p_2)
A (q_1)	p_{AB} ($p_1 q_1$)	p_{Ab} ($p_2 q_1$)
a (q_2)	p_{aB} ($p_1 q_2$)	p_{ab} ($p_2 q_2$)

$$D = P_{11}P_{22} - P_{12}P_{21}$$

Cis *Trans*
Frekvence haplotypů

$$\text{Nerovnováha v generaci } t: D_t = (1 - r)^t D_0$$



Testování vazbové nerovnováhy

- χ^2 test – 2 lokusy s 2 alelami

$$\chi_{ij}^2 = \frac{n[d(ij)]^2}{f(i)[1 \pm f(i)] f(j)[1 \pm f(j)]}$$

$d(ij) = f(ij) - f(i)f(j)$

- nerovnováha alely i v lokusu 1 a alely j v lokusu 2

$f(i), f(j)$ - frekvence alely i v lokusu 1, alely j v lokusu 2

$f(ij)$ - frekvence haplotypu ij (lokus 1 a 2)

Ab	20	0,40	$p(L1)$	0,5	$f(A)$
ab	10	0,20	$q(L1)$	0,5	$f(a)$
AB	5	0,10	$p(L2)$	0,4	$f(B)$
aB	15	0,30	$q(L2)$	0,6	$f(b)$
	50	1,00			

$d(Ab)$	0,10	$d(ij) = f(ij) - f(i)f(j)$
$d(ab)$	-0,10	
$d(AB)$	-0,10	
$d(aB)$	0,10	

Ab	$\chi^2 =$	8,333333	0,5
			0,06
ab	$\chi^2 =$	8,333333	0,5
			0,06
AB	$\chi^2 =$	8,333333	0,5
			0,06
aB	$\chi^2 =$	8,333333	0,5
			0,06

$$\chi_{ij}^2 = \frac{n[d(ij)]^2}{f(i)[1 \pm f(i)] f(j)[1 \pm f(j)]}$$

Příklad v excelu

PROGRAM**DOWNLOAD**

Arlequin	http://lqb.unige.ch/arlequin/
CONVERT	http://www.agriculture.purdue.edu/fnr/html/faculty/Rhodes/Students%20and%20Staff/glaubitz/software.htm
DISPAN	http://www.bio.psu.edu/People/Faculty/Nei/Lab/software.htm
FSTAT	http://www2.unil.ch/izea/software/fstat.html
GDA (Genetic Data Analysis)	http://lewis.eeb.uconn.edu/lewishome/software.html
Genetix4.04	http://www.univ-montp2.fr/~genetix/genetix/genetix.htm
KINSHIP:	http://www.gssoftnet.us/GSoft.html
Microsat	http://hpgl.stanford.edu/projects/microsat/
PAUP 4.0b	http://paup.csit.fsu.edu/
PCAGen	http://www2.unil.ch/izea/software/pcagen.html
Phylip	http://evolution.genetics.washington.edu/phylip.html
Popgene	http://www.ualberta.ca/~fyeh/
Structure	http://pritch.bsd.uchicago.edu/software.html
TreeView	http://taxonomy.zoology.gla.ac.uk/rod/treeview.html