

# Molecular approaches in behavioural ecology

## ■ Foraging

- Identifying prey, individual food preferences

## ■ Mating systems

- Extra-pair fertilizations
- Conspecific brood parasitism
- Mate choice (pre- and postcopulatory)
- Social breeding (relatedness)

## ■ Sex ratio manipulation

- Adaptive sex ratio
- Sex ratio conflicts

# „Foraging“ – study of diet preferences

- Identification of individuals (genetically) + their diet (other methods)
- Identification of individuals (other methods) + their diet (genetically)
- Identification of individuals (genetically) + their diet (genetically)

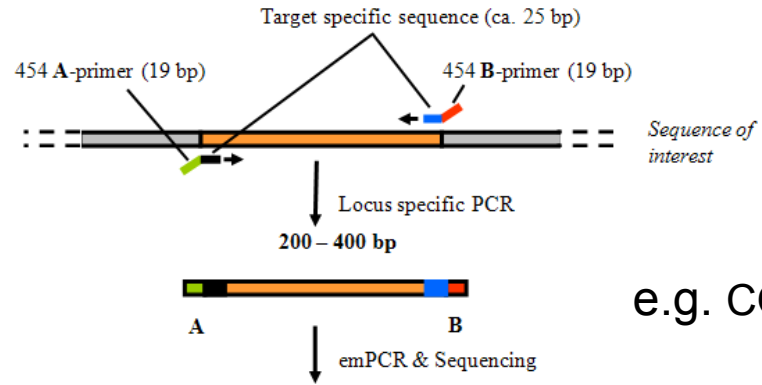
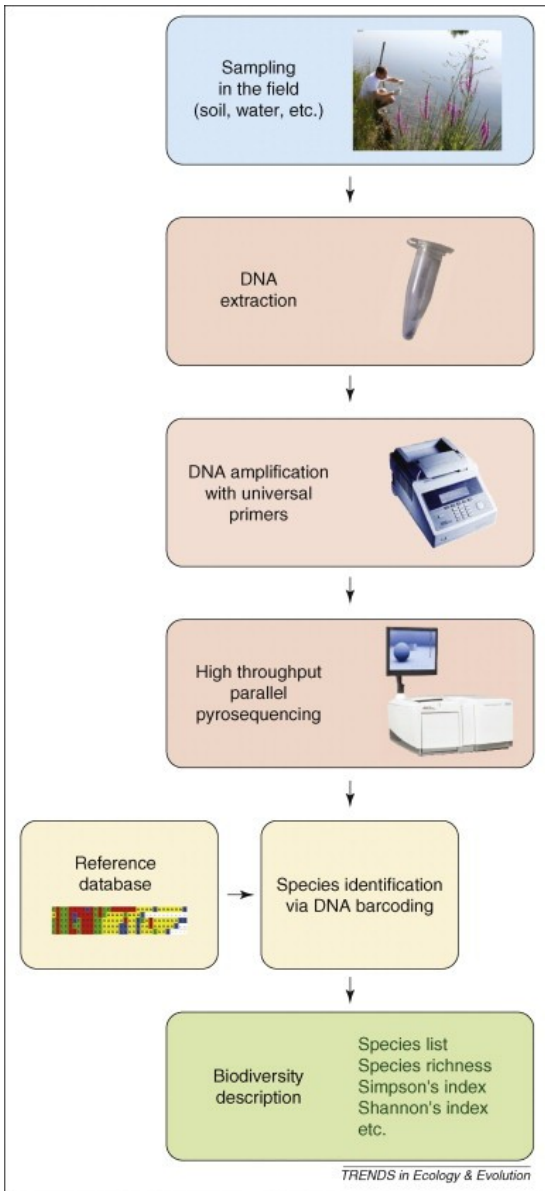
# „Foraging“ - non-invasive CMR studies

- **Population size – identification of individuals**
- „Capture-Mark-Recapture“ (review in Lukacs & Burnham 2005)
- Repeated sampling of the same individual (= multilocus genotype)
- Survival, population dynamics, etc.
- Closed population models, open population models, Robust design models
- Corrections for genotyping errors
- Faeces – analysis of **individual variability** in diet (e.g. coyotes - Fedriani & Kohn 2001)



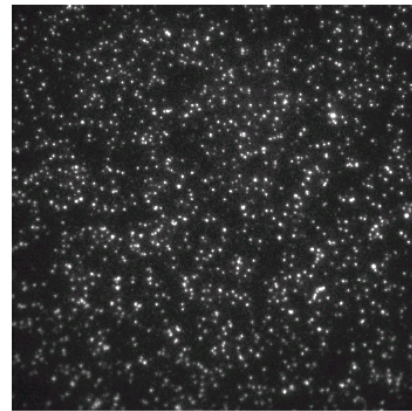
Population dynamics of coyotes (Prugh et al. 2005)

# Sequencing of mixed samples of diet (= PCR amplicons)

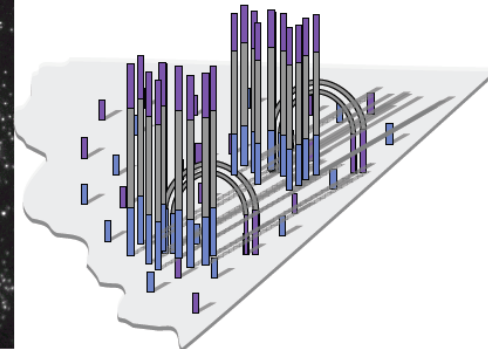


e.g. COI gene

parallel sequencing



Random array of clusters



~1000 molecules per ~ 1 um cluster  
 ~20-30,000 clusters per tile  
 ~40 M clusters per flowcell

# Example: Diet analysis of herbivores in Himalayas

Molecular Ecology Resources (2009) 9, 51–60

doi: 10.1111/j.1755-0998.2008.02352.x

## TECHNICAL ADVANCES

### New perspectives in diet analysis based on DNA barcoding and parallel pyrosequencing: the *trnL* approach

ALICE VALENTINI,\*† CHRISTIAN MIQUEL,\* MUHAMMAD ALI NAWAZ,‡§ EVA BELLEMAIN,\* ERIC COISSAC,\* FRANÇOIS POMPANON,\* LUDOVIC GIELLY,\* CORINNE CRUAUD,¶ GIUSEPPE NASCETTI,† PATRICK WINCKER,¶ JON E. SWENSON,‡\*\* and PIERRE TABERLET\*  
\*Laboratoire d'Ecologie Alpine, CNRS UMR 5553, Université Joseph Fourier, BP 53, F-38041 Grenoble cedex 9, France, †Dipartimento di Ecologia e Sviluppo Economico Sostenibile, Università degli Studi della Tuscia, via S. Giovanni Decollato 1, I-01100 Viterbo, Italy, ‡Department of Ecology and Natural Resource Management, Norwegian University of Life Sciences, Post Box 5003, NO-1432 Ås, Norway, §Himalayan Wildlife Foundation, 01, Park Road, Sector F-8/1 Islamabad 44000, Pakistan, ¶Genoscope – CNS, 2 rue Gaston Crémieux, BP 5706, F-91057 Evry cedex, France, \*\*Norwegian Institute for Nature Research, NO-7485 Trondheim, Norway



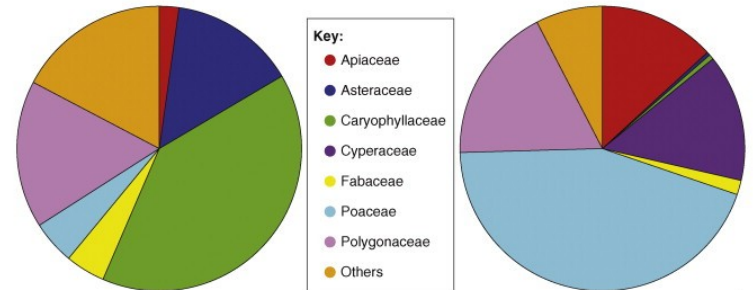
Deosai National Park, Pakistan



Golden marmot



Brown bear



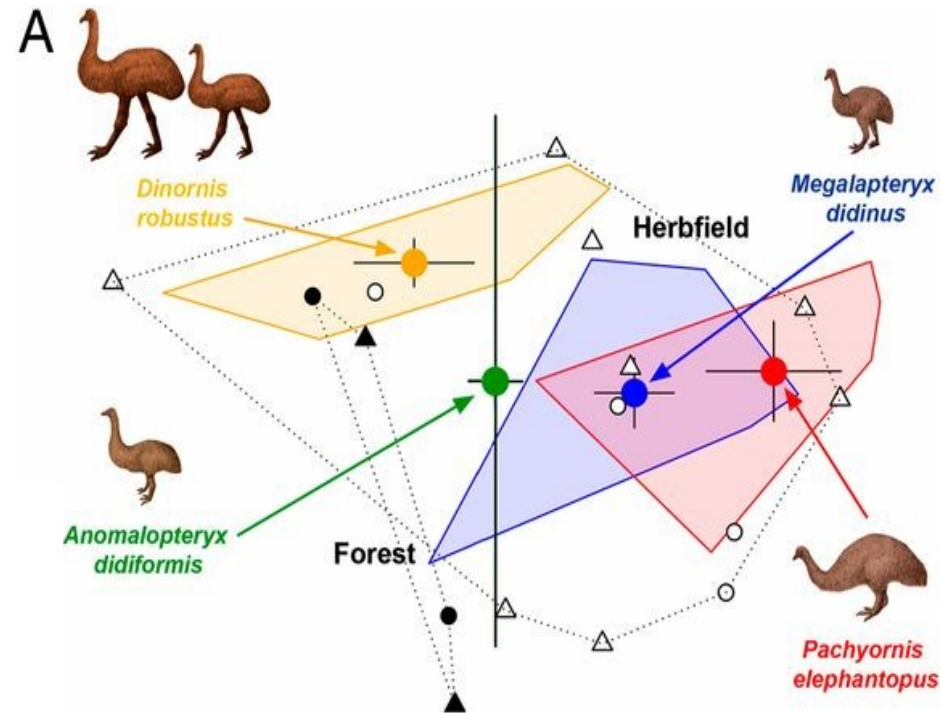
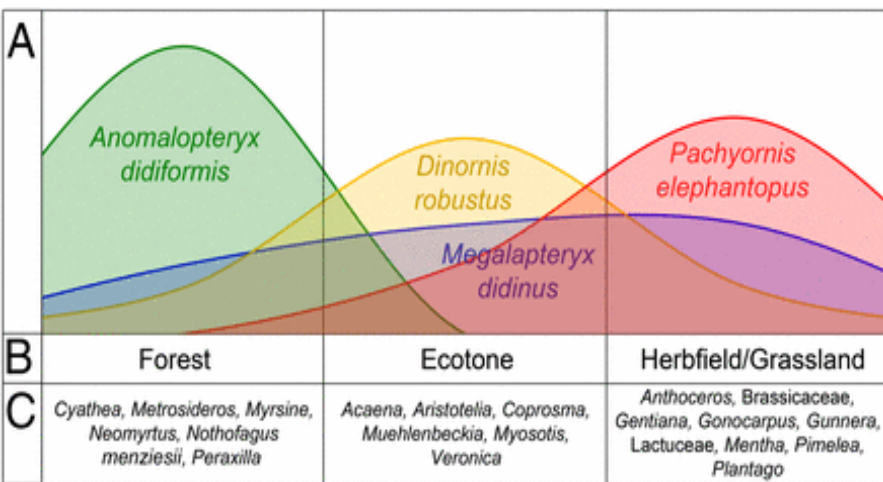
TRENDS in Ecology & Evolution

- 12 feces of each species were collected
- amplification with universal primers targeting a short fragment of the chloroplast *trnL* (UAA) intron
- the amplicons were analyzed on the 454 GS FLX sequencer
- more than 2000 DNA sequences were obtained per feces
- the plant taxa eaten were identified by comparison with available reference sequences

# Dietary niche of extinct taxa

## Ancient DNA from coprolites of moa (New Zealand)

Environmental requirements of individual species + separation of ecological niches



## Resolving lost herbivore community structure using coprolites of four sympatric moa species (Aves: Dinornithiformes)

Jamie R. Wood<sup>a,1</sup>, Janet M. Wilmshurst<sup>a</sup>, Sarah J. Richardson<sup>a</sup>, Nicolas J. Rawlence<sup>b,2</sup>, Steven J. Wagstaff<sup>a</sup>, Trevor H. Worthy<sup>c,3</sup>, and Alan Cooper<sup>b</sup>

<sup>a</sup>Landcare Research, Lincoln, Canterbury 7640, New Zealand; <sup>b</sup>Australian Centre for Ancient DNA, University of Adelaide, Adelaide, SA 5005, Australia;



# Mating systems

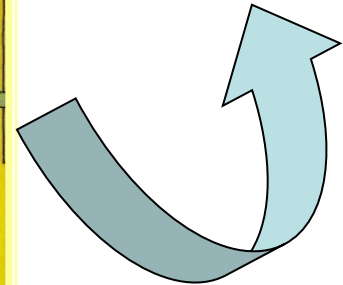
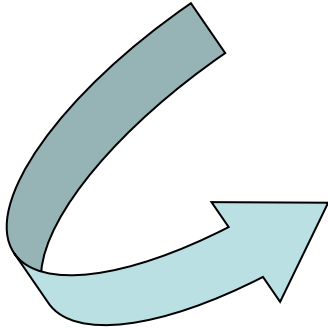
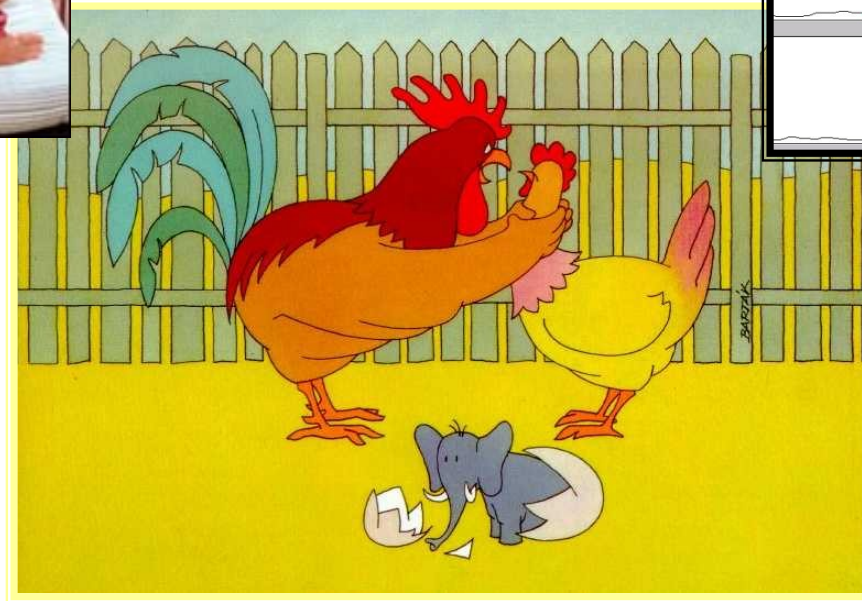
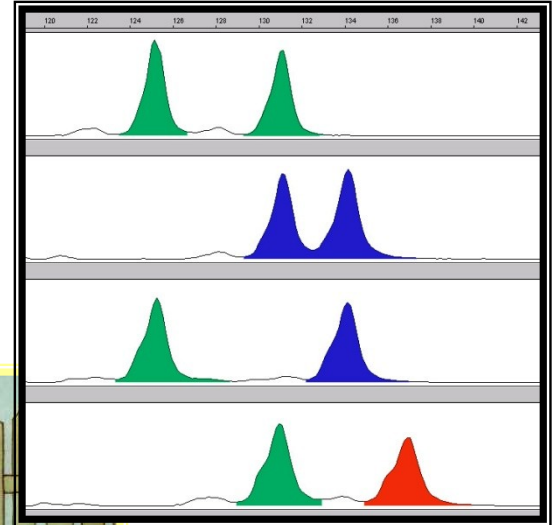
Mating system	No. of males	No. of females
Monogamy	1	1
Polygyny	1	Multiple
Polyandry	Multiple	1
Polygynandry	Multiple	Multiple
Promiscuity	Multiple	Multiple



Polygamy



# Parentage analysis





# Why parentage analyses?



- extra-pair fertilization – social vs. genetic father
- number of reproductively active individuals in a population
- study of factors affecting fitness (reproductive success)
- multiple paternity – analyses of mating systems (detection of promiscuity, etc.)



# Paternity x maternity x parentage

One parent known

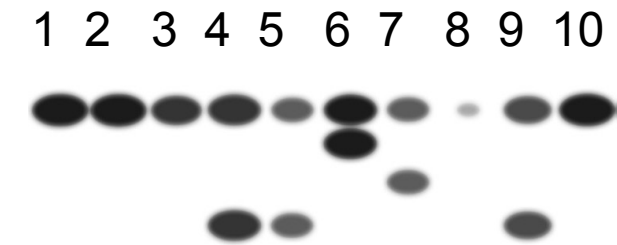


No parent known

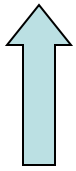
# History

- already first studies of genetic polymorphism → genetic techniques can solve parentage questions → one of the major topics in molecular ecology
- chromosomal polymorphism
- allozyme electrophoresis
- minisatellite DNA fingerprinting – overturn of existing paradigms in behavioural ecology (birds)
- statistical techniques for single-locus polymorphism (allozymes) – departure from practice (i.e. minisatellite DNA fingerprinting)

# History

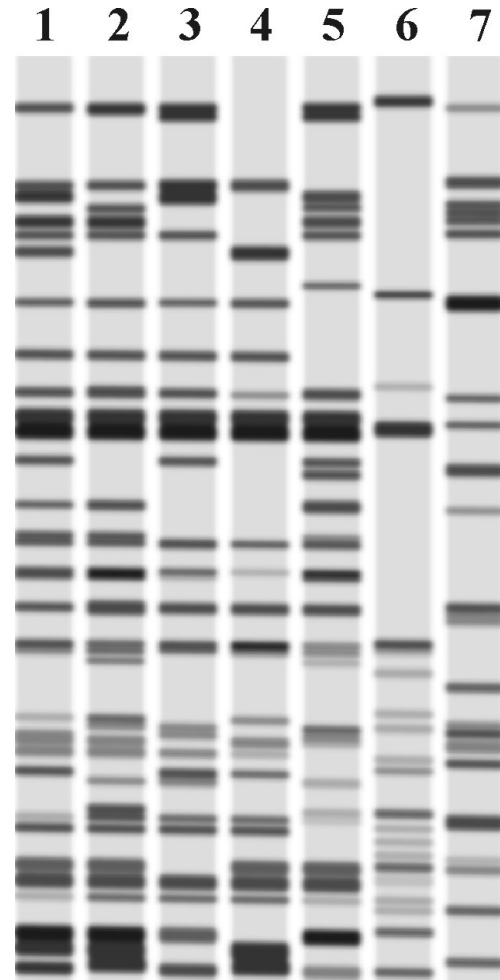


allozymes



low variation

multi-locus genotype



DNA fingerprinting, RAPD  
absence of theoretical models



## Protein fingerprinting

- Hohol severní  
*Andersson & Åhlund 2000*
- Proteins from egg white
- isoelectric focusing  
in immobilized pH gradients
- Intraspecific brood parasitism (more than half of nests)
- More frequent between related females

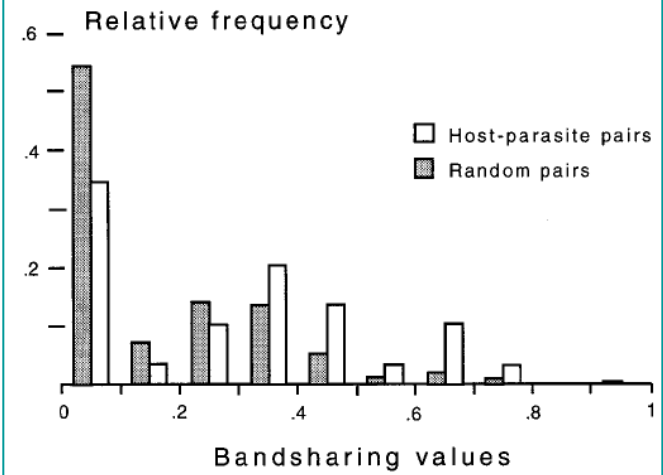


Fig. 2. Distribution of bandsharing values in the 29 pairs of host and primary parasite in 1986 and in the 861 different random pairs that can be drawn among the 42 other females not involved in host-parasite relationship with each other. Bandsharing values are significantly higher for the host-parasite pairs than for the randomly formed pairs (see *Host-Parasite Relatedness*).

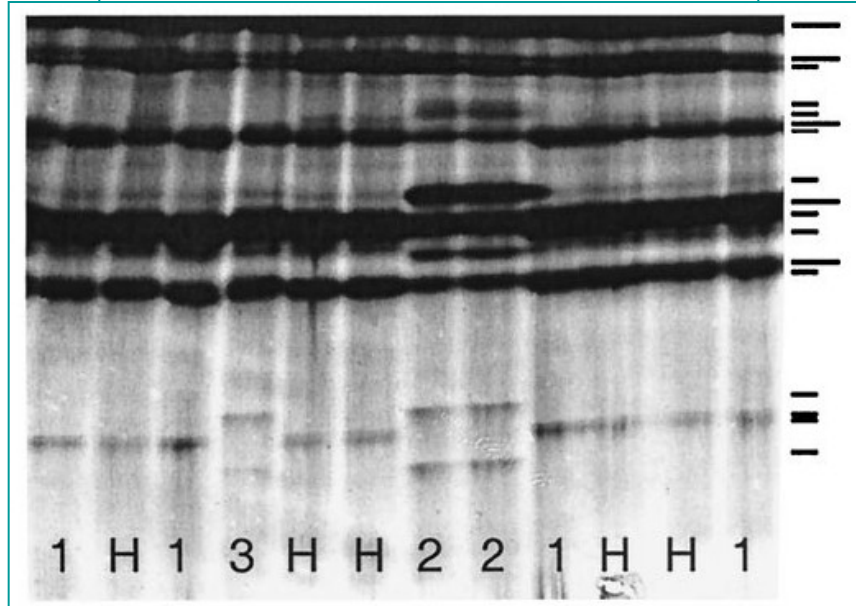


Fig. 1. Electrophoretic gel with albumen band patterns for the 12 eggs of a goldeneye clutch parasitized by three females. There are five host eggs (lanes marked H) and four, two, and one parasite eggs (lanes marked 1, 2, and 3, respectively). Bands that were scored for bandsharing analysis are marked with a short dash in the margin. Several different bands occur in all 902 eggs and can therefore be used as location references; they are marked with a longer dash.



# Mikrosatellites

(most frequent markers today in parentage analyses)

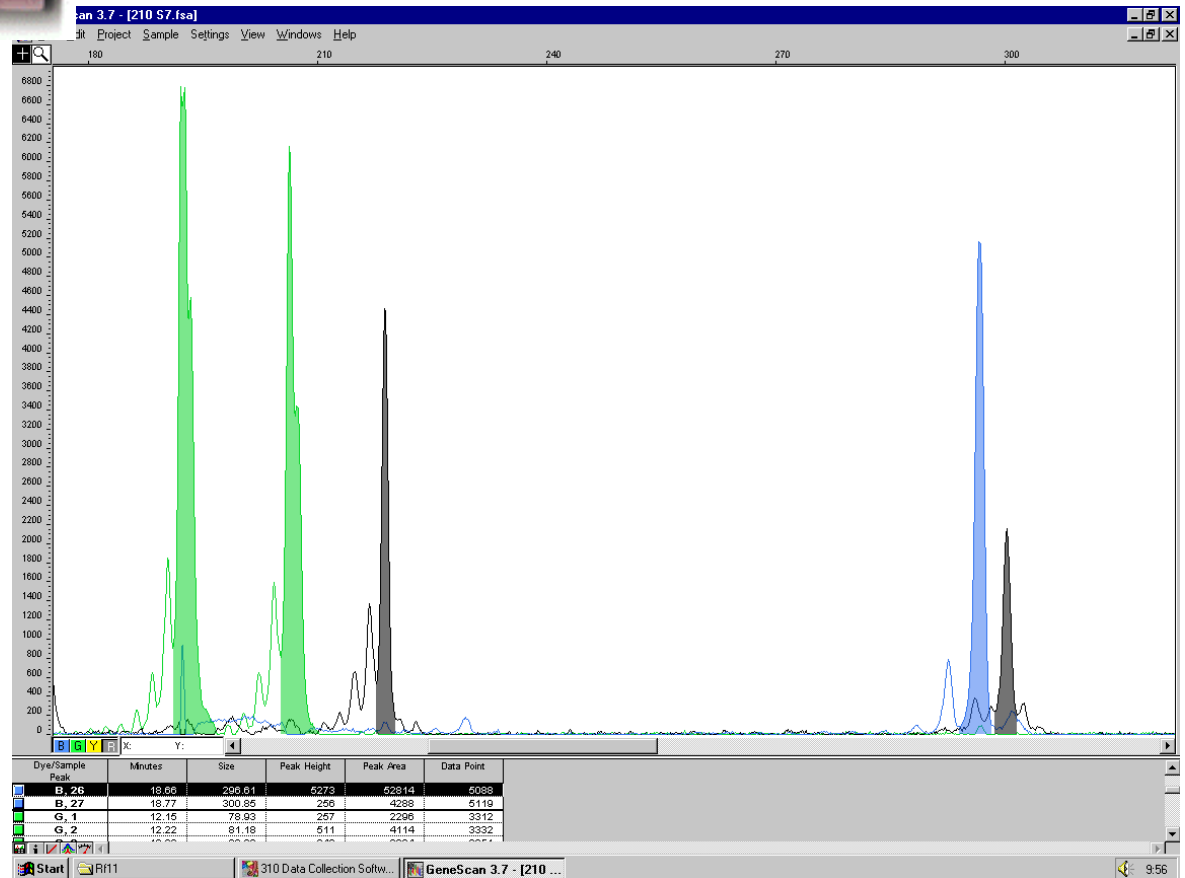
- Combination of theory (statistical models) and practice (highly polymorphic single-locus markers)
- Tandem repetitions of short motifs
- Ex.:  $(CTTT)_n$  or  $(CA)_n$
- Simple Mendelian inheritance – co-dominant



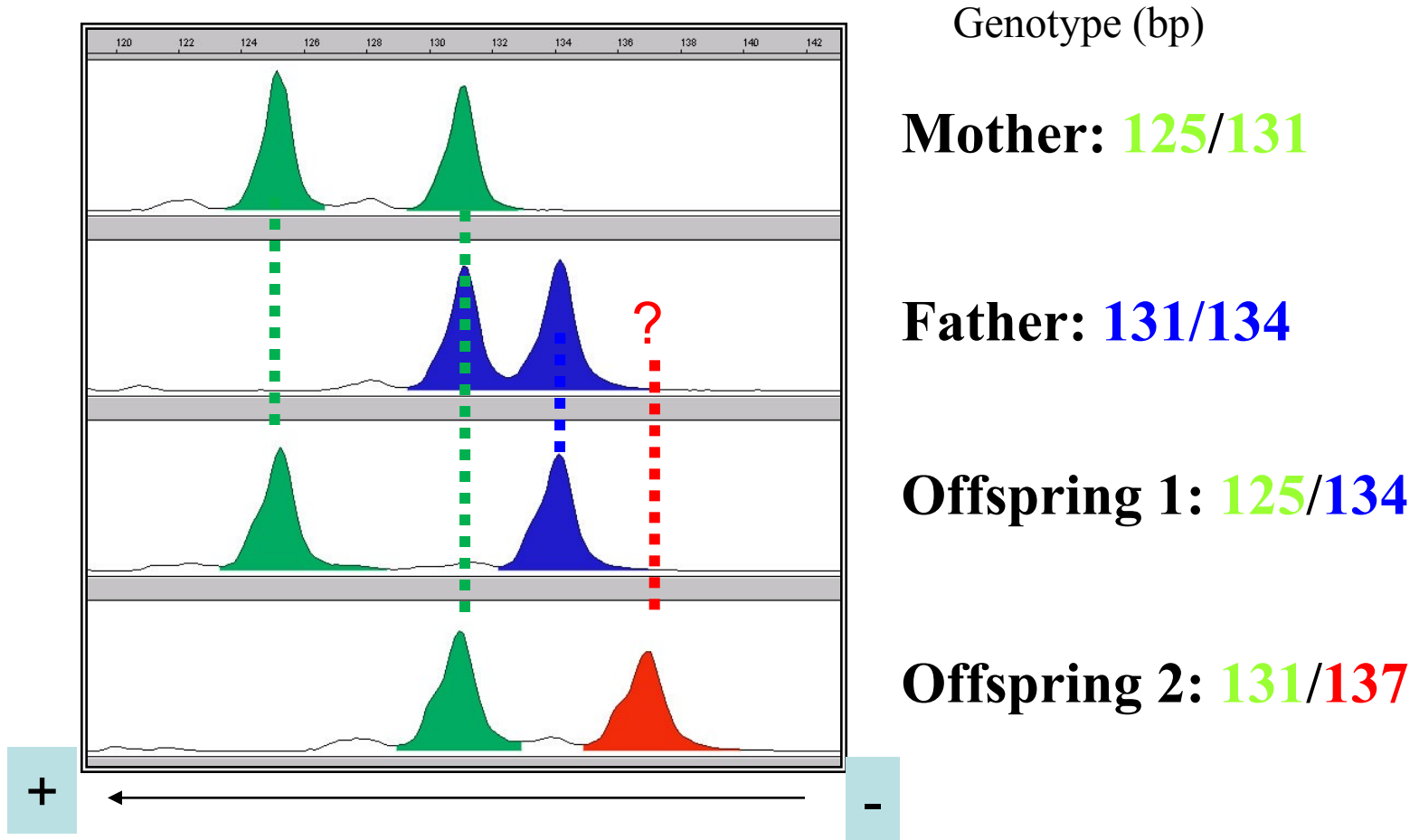
# Fragmentační analýza



Je možné analyzovat  
několik lokusů  
najednou.



# Example of one locus – fragment analysis of PCR products



Analysed male could sire offspring 1, but surely not offspring 2

# Spectrum of approaches

*Jones & Ardren 2003, Jones et al. 2010*

## ➤ Few candidate parents, all sampled, enough of variable markers

### → Exclusion

- incompatibility between parent and offspring → parentage rejected
- subsequent exclusion leads to identification of „true“ parents (→ biological father/mother)

## ➤ More non-excluded candidate parents

### → Methods based on maximum likelihood

Categorical and fractional allocation

(based on homozygosity → homozygotes have higher probability to transfer the allele to the next generation)

### → Full probability parentage analysis (Bayesian statistics)

(includes other parameters, e.g. distance between nests, behavioural data, etc.)

## ➤ Parents unknown and unsampled

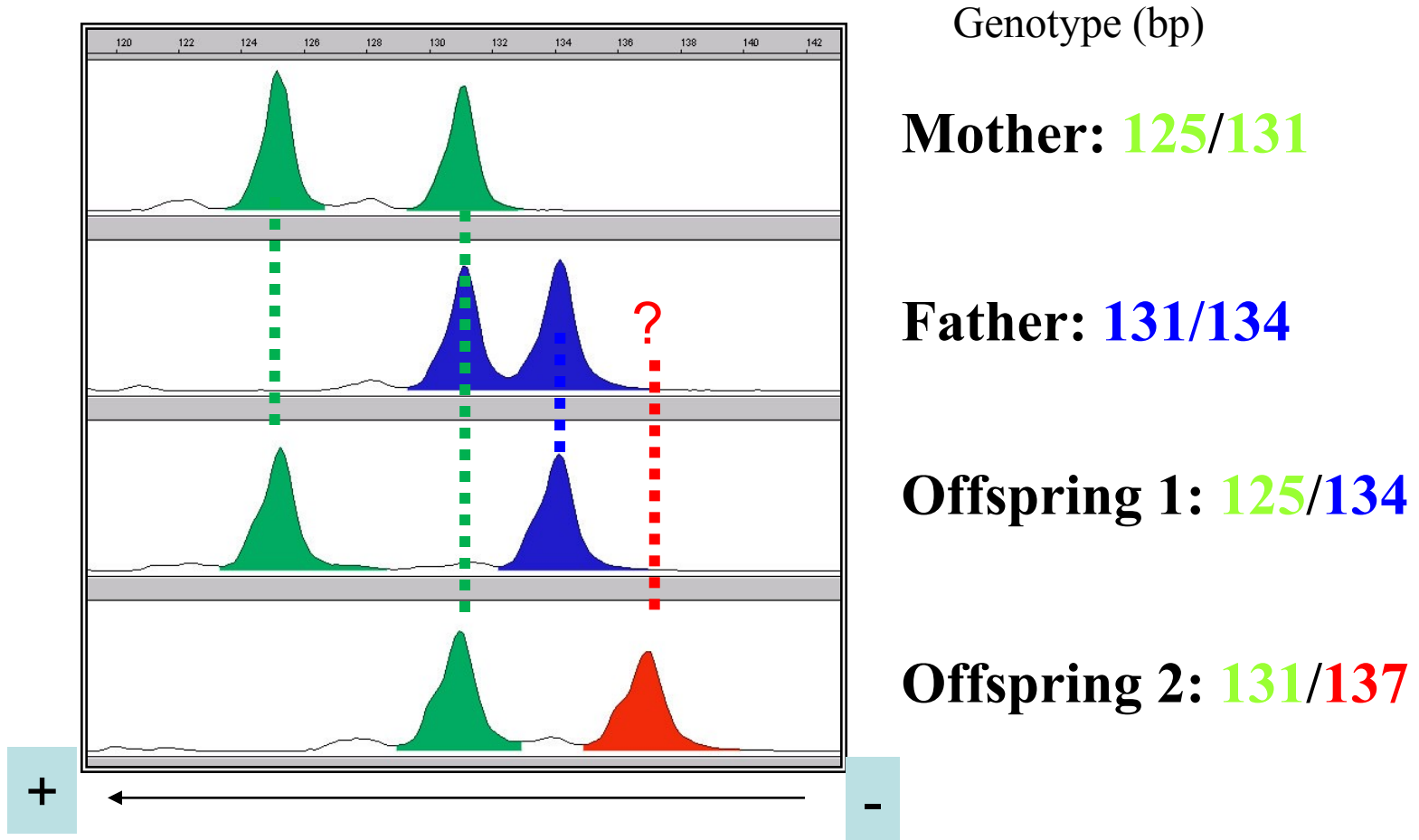
### → Parental reconstruction

Multiple genotypes from a single family → reconstruction of parental genotypes

E.g. identification of multiple paternity in fish



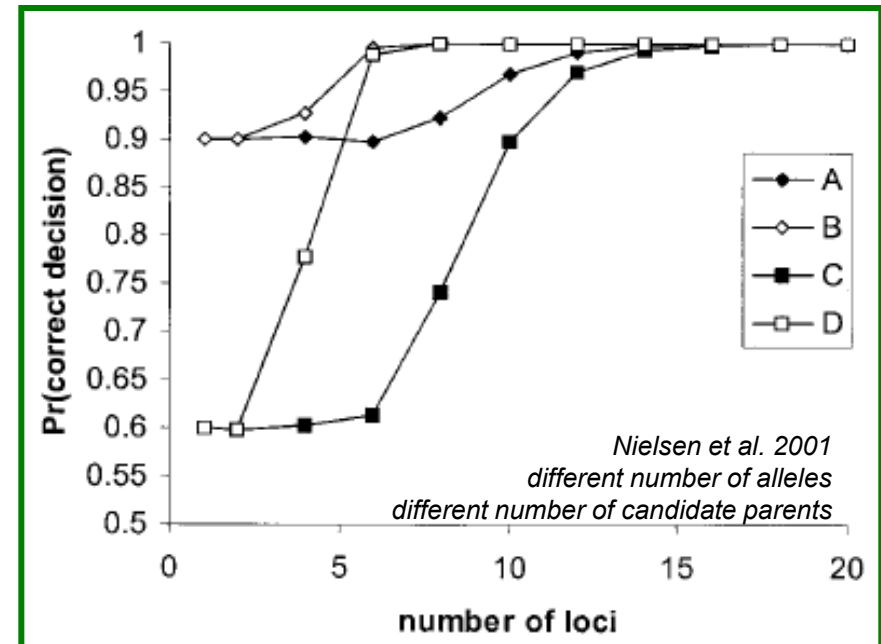
# Exclusion principle



Analysed male could sire offspring 1, but surely not offspring 2

# How many loci is necessary?

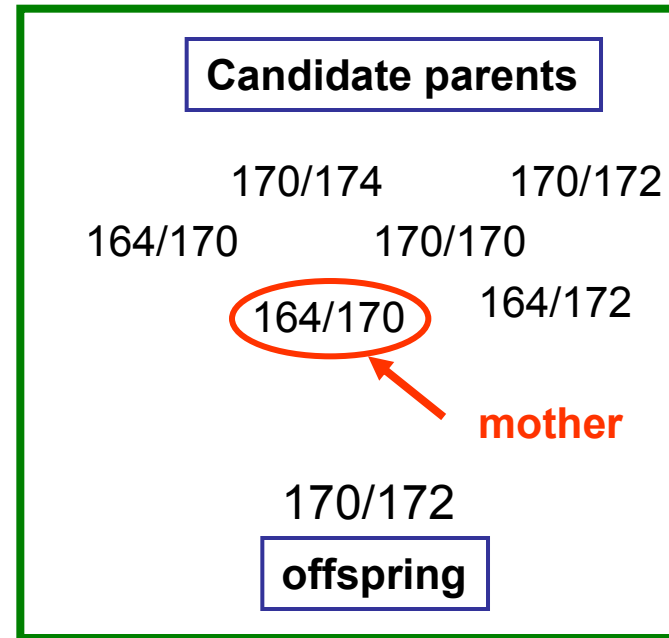
- No absolute value
- Depending on:
  - ✓ number of alleles
  - ✓ heterozygosity
  - ✓ allele frequencies
  - ✓ sample size
- ideal – hypervariable loci  
sometimes 3 are enough



- pilot study of adult part of population
- → **Exclusion probability**

# Exclusion probability

- Probability that randomly selected genotype in a population will not match the offspring,  
i.e. **probability**, that **randomly selected individual will be excluded** as a parent
- often as NonExclusion probability (NEP)  
(= 1 – Exclusion probability)
- it depends on number of loci and alleles
- good values of NEP are 0.01 and lower
- calculated e.g. in **Cervus**
- different values for the first and second parent



- Exclusion probability – rough estimate
- more suitable are usually simulations including also numbers and proportions of sampled individuals, genotyping errors and other problems (mutations, null alleles) – included in „likelihood approaches“ e.g. in Cervus

# Exclusion - complications

- Mendelian inheritance of microsatellites: genotyping errors, null alleles, mutations → false exclusions in strict conditions



# Null alleles and genotyping errors complicates analyses of relatedness (e.g. parentage)

	lokus 1 null alleles		lokus 2 genotyping error	
Mother	100	150	300	350
Father X	100	100	300	367
Offspring	150	150	350	365

**Father X** is always the „true“ father of the analysed offspring, but a simple exclusion analysis leads to erroneous results

# Mutations

*Ibarguchi et al. 2004*

- usually not considered
- however they can be relatively frequent  
( $4.5 \times 10^{-2}$  -  $5.1 \times 10^{-6}$  per locus)
- more frequent in males (2 až 6x)
- → exclusion based on a single locus can be erroneous

# Solution

- use of more loci
- Mutace → nejčastěji posun o jednu jednotku opakování  
například rodič (CTTT)<sub>7</sub> potomek (CTTT)<sub>8</sub>, délka fragmentu tedy například 120 u rodiče a 124 u potomka
- Probability of resemblance ( $P_R$ )  
Podívám se na ostatní lokusy a vypočítám s jakou pravděpodobností mohou alely sdílet dva nepříbuzní jedinci

$$P_{Ra} = (2p_a - p_a^2)^2$$

Pro jeden lokus

$$P_{RaCum} = \prod_r (P_{Ra_r})$$

Součin pravděpodobností přes lokusy



## *Uria lomvia*

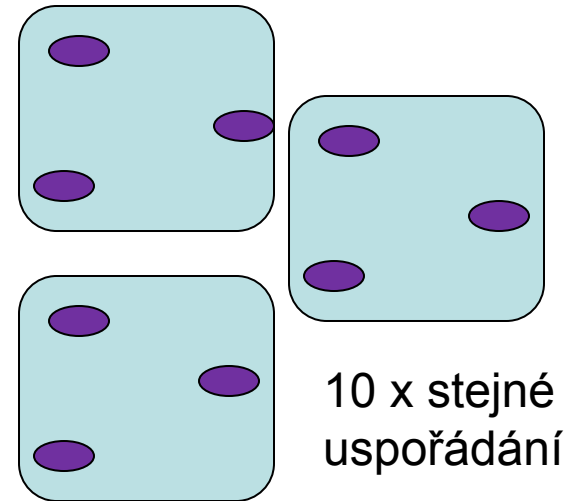
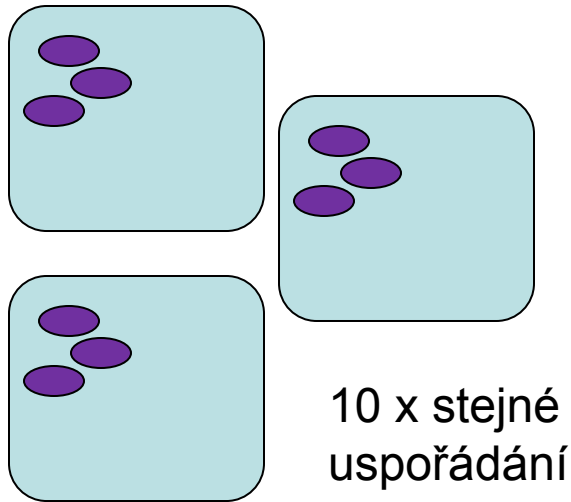


- **22%** mimopárová paternita bez korekce na mutace
- **7%** mimopárová paternita s korekcí soulad s jinou použitou technikou

# Exclusion - komplikace

- Mendelian rules of inheritance: genotyping errors, null alleles, mutations → false exclusions in strict conditions
- Extended family structure (příbuznost potenciálních rodičů) – nejhorší jsou sourozenci
- Linked loci – pokles variability
- Znaky na pohlavních chromozómech
- Problémy narůstají s rostoucím množstvím jedinců a lokusů
- „Exclusion“ je velmi užitečná metoda např. v experimentech, kde jsou všichni rodiče předem známi a zgenotypováni

# Př.: Faktory ovlivňující celoživotní fitness u hořavky duhové

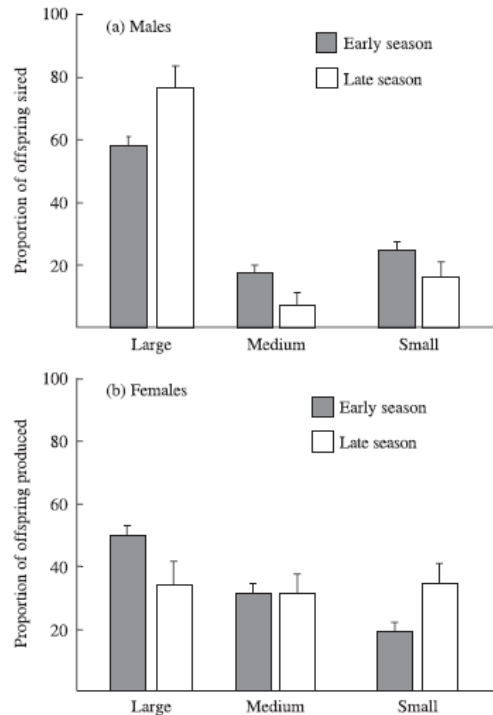
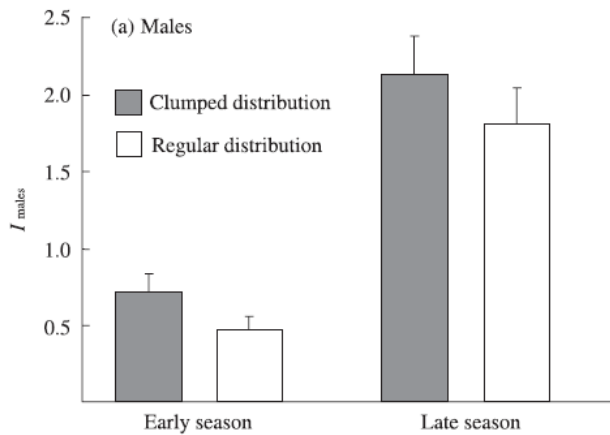


- 3 páry hořavek v každé nádrži
- přes 4000 embryí (odebírány v pravidelných intervalech, 5 mikrosatelitů)



- velikost těla
- kondice
- zbarvení duhovky
- velikost gonád
- parazitofauna

# Př.: Faktory ovlivňující celoživotní fitness u hořavky duhové



Nahloučená distribuce = selekce na velikost gonád

Rovnoměrná distribuce = selekce na velikost těla

Fig. 3 Mean (+ 1 SE) proportion (in percent) of offspring sired by individual (a) male and (b) female bitterling in relation to the body size categories at the start (dark columns) and end (open columns) of the reproductive season.

„velikost těla u samců rozhoduje“

„opportunity for sexual selection“  
(nezávisí na distribuci škeblí, ale na sezóně – na podzim je variabilita v reprodukčním úspěchu větší)



# Metody

*Jones & Ardren 2003, Jones et al. 2010*

➤ Málo potenciálních rodičů, všichni ovzorkováni, dostatek dobrých markerů

→ **Exclusion (prosté vyloučení)**

nekompatibilita mezi rodičem a potomkem → zamítnutí rodičovství

postupným vyřazováním nakonec zůstane jeden možný rodič (→ biologický otec nebo matka)

➤ **Nelze vyloučit více jedinců (zbyde více možných otců než 1)**

→ **Metody založené na maximální věrohodnosti**

Categorical and fractional allocation

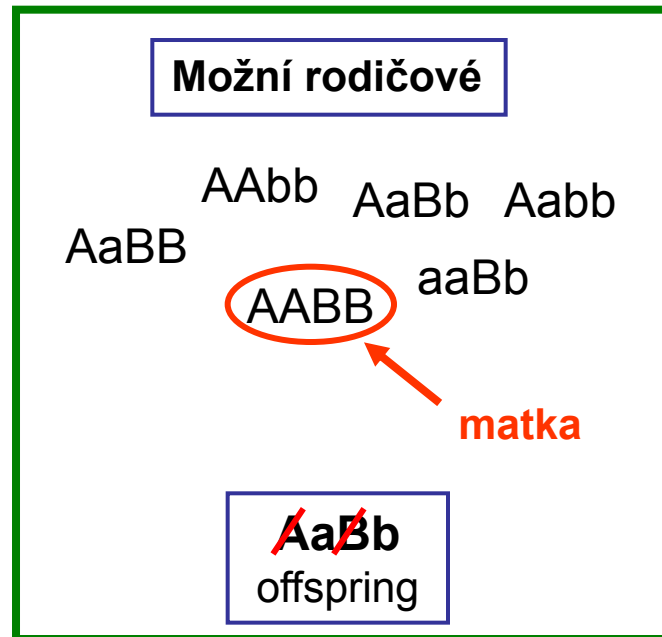
(založeno na homozygotnosti → homozygot má větší pravděpodobnost předat alelu)

→ **Full probability parentage analysis** (Bayesian statistics)

(zahrnutí i dalších dat jako vzdálenost hnízd, behaviorální data...)



# Categorical x fractional likelihood



-výpočet věrohodnosti („likelihood“) paternity pro jednotlivé kandidáty (na základě frekvencí alel)

# Categorical x fractional likelihood

- „likelihood“ skóre je vypočítáno na základě genotypů potomka a všech nevyložených rodičovských genotypů (na základě frekvencí alel v populaci)
- možnost zahrnout i další parametry (např. účinnost vzorkování či genotypizací)
- **Categorical I.:** potomek jako jednotka přiřazen otci, biologicky validní
- **Fractional I.:** potomek přiřazen všem kompatibilním otcům. Statisticky přesnější a proto někdy výhodnější (např. lepší pro srovnání reprodukčního úspěchu různých kategorií samců)

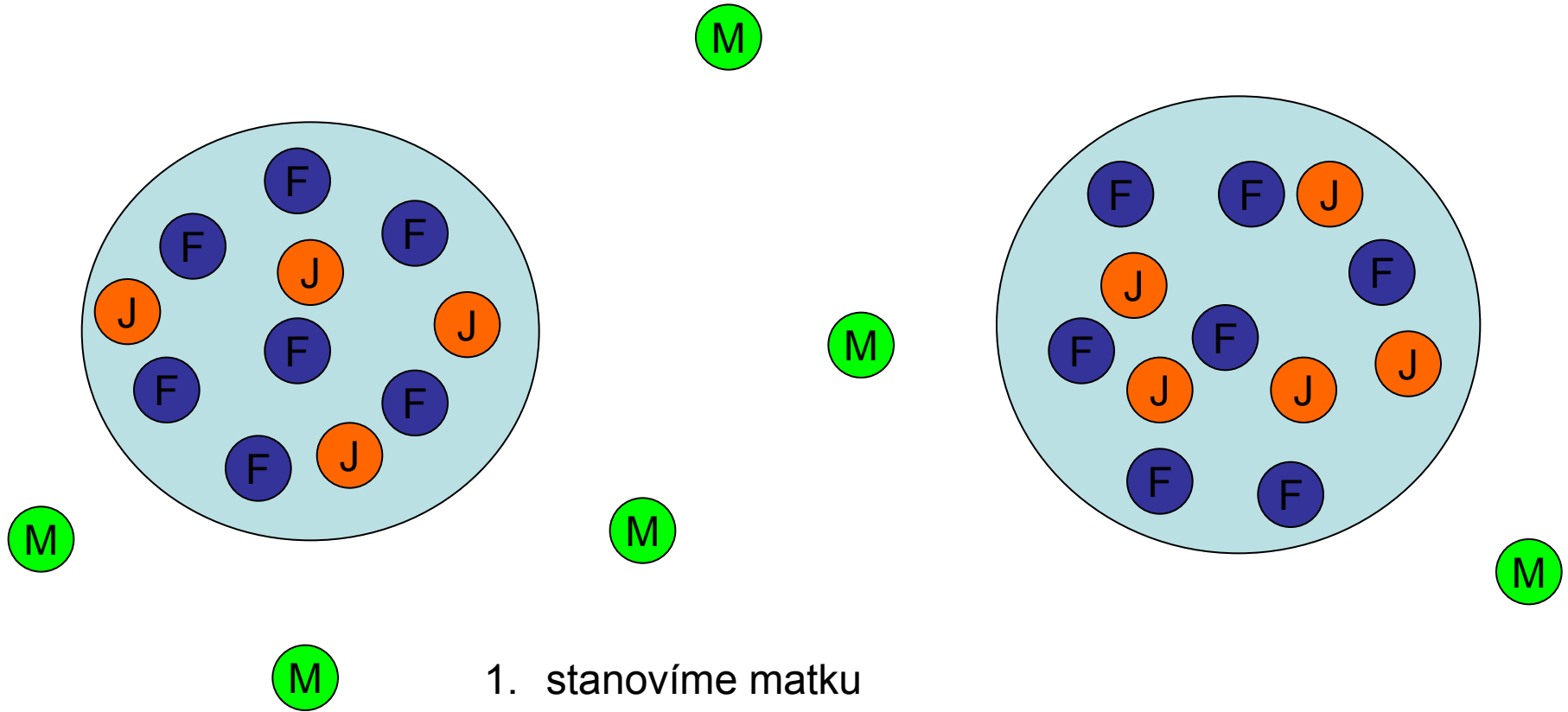


# Categorical allocation

- a) hledáme otce
- b) hledáme matku
- c) hledáme oba rodiče
- d) hledáme oba rodiče, ale neznáme pohlaví jednotlivých kandidátů

... program CERVUS

# Typický příklad: Maternita a paternita u netopýrů

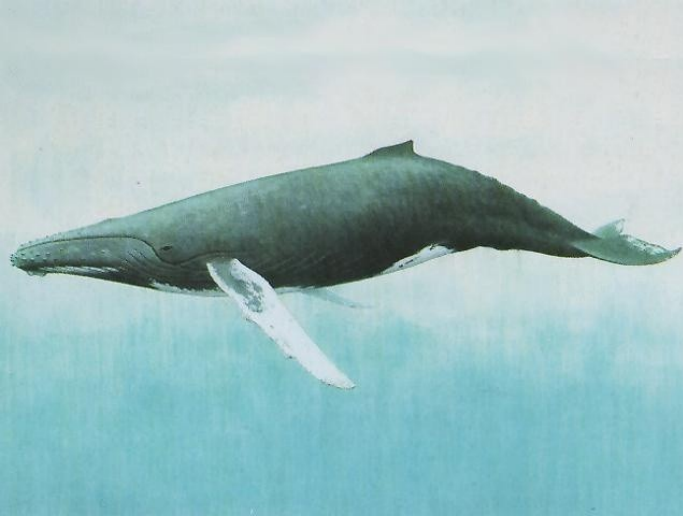


*Pipistrellus nathusii* – jižní Čechy



# Fractional allocation

- přiřadí určitou frakci (mezi 0 a 1) každého potomka všem nevyloženým kandidátním rodičům
- proporce každého potomka přiřazená danému rodiči odpovídá věrohodnosti („likelihood“) rodičovství – tj. je závislá na ostatních nevyložených rodičích
- statisticky přesnější a proto někdy výhodnější (např. lepší pro srovnání reprodukčního úspěchu různých kategorií samců)



*Megaptera  
novaeangliae*

*Nielsen et al. 2001*



- Až 25 samců se pokouší pářit se samicí
- Různé role samců - dominantní společník a vyzyvatel, sekundární společníci
- 6 mikrosatelitových lokusů
- Fractional likelihood paternity method
- Dominantní samci mají asi 3x více mláďat (ale rozdíly jsou pouze marginálně signifikantní)

# Full probability model

## Bayesiánská statistika

Molecular Ecology (2006) 15, 3715–3730



## Towards unbiased parentage assignment: combining genetic, behavioural and spatial data in a Bayesian framework

J. D. HADFIELD\*, D. S. RICHARDSON† and T. BURKE\*

*\*Department of Animal and Plant Sciences, University of Sheffield, Western Bank, Sheffield S10 2TN, UK, †Centre for Ecology, Evolution and Conservation, School of Biological Sciences, University of East Anglia, Norwich NR4 7TJ, UK*

- Složitější modely, zohledňují vzdálenost hnízd, frekvenci krmení, párování...

# Metody

*Jones & Ardren 2003, Jones et al. 2010*

## ➤ Málo potenciálních rodičů, všichni ovzorkováni, dostatek dobrých markerů

### → Exclusion (prosté vyloučení)

nekompatibilita mezi rodičem a potomkem → zamítnutí rodičovství

postupným vyřazováním nakonec zůstane jeden možný rodič (→ biologický otec nebo matka)

## ➤ Nelze vyloučit více jedinců (zůstane více možných otců než 1)

### → Metody založené na maximální věrohodnosti

Categorical and fractional allocation

(založeno na homozygotnosti → homozygot má větší pravděpodobnost předat alelu)

### → Full probability parentage analysis (Bayesian statistics)

(zahrnutí i dalších dat jako vzdálenost hnízd, behaviorální data...)

## ➤ Rodiče neznámi a neovzorkováni

### → Parental reconstruction

Z genotypů potomků z jedné rodiny → rekonstrukce genotypu rodičů

Například k určení „multiple paternity“



# Genotypic (parental) reconstruction

- rekonstrukce parentálních genotypů z velkého množství genotypů potomků (full- or halfsibs) – složitý algoritmus
- Často při analýzách paternity u ryb
- Například k určení „multiple paternity“, „rate of selfing“
- Zásadní je velký počet mlád'at ve snůšce/vrhu!
- Optimální je  $> 10$

# Počítačové programy

## Cervus

Newpat

Probmax

Kinship

Famoz

Pasos

Papa

Parente

Patri

a další viz Jones et al. 2010

(Některé programy počítají i s možnými chybami při určování genotypů  
*Cervus3, Kalinowski et al 2007, Newpat*)

# Co je nejdůležitější?

- Kvalitní data z terénu či experimentů!
- Nejlépe všichni dospělci z populace, behaviorální a jiná fenotypická data, mláďata přiřazena do rodin (matka + sociální partner)
- Špatná data = špatné výsledky

Stane se, že nemohu získat vzorky všech dospělých jedinců z populace (potenciálních rodičů) nebo nelze říci, která mláďata jsou z jedné rodiny

- Nemusí se zdařit najít rodiče všem potomkům
- I tak ale mohu zjistit ledacos zajímavého (multiple paternity) ...

# Příklad

## genotypy matky a embryí

	lokus 1		lokus 2	
Matka	100	150	300	350
1. embryo	100	115	300	320
2. embryo	150	120	350	310
3. embryo	120	100	350	365

# Alely od matky

	lokus 1		lokus 2	
Matka	100	150	300	350
1. embryo	100	115	300	320
2. embryo	150	120	350	310
3. embryo	120	100	350	365

Na druhém lokusu více **než dvě alely**,  
které nejsou od matky → více otců

... nebo mutace – data nutno korigovat – viz *Uria lomvia* („probability of resemblance“)

# *Apodemus agrarius* a *sylvaticus*

mláďata až od tří samců (*Bryja et al. 2008*)

Species	N litters	Minimum number of sires		
		1	2	3
<i>A. uralensis</i>	46 (40)	26 (22)	20 (18)	0 (0)
<i>A. flavicollis</i>	25 (16)	10 (7)	15 (9)	0 (0)
<i>A. sylvaticus</i>	22 (16)	7 (5)	13 (9)	2 (2)
<i>A. agrarius</i>	34 (26)	14 (8)	13 (11)	7 (7)



# Další příklady užitečnosti analýz paternity



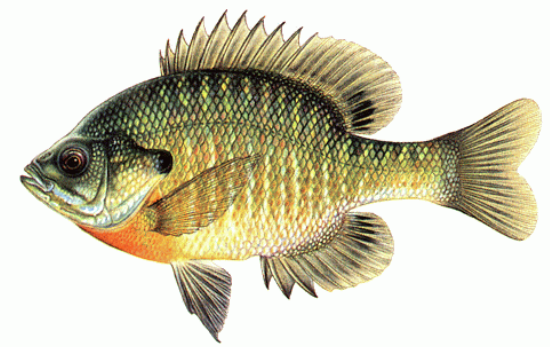


# Orangutan na Sumatře (paternity)

- Nápadný dimorfismus  
(vystouplé tváře a hrdelní vaky u samců)
- Předpoklad:  
Samci bez nápadných znaků hormonálně suprimováni a nemnoží se
- Mikrosatelity – 50 % mláďat je od samců bez nápadných znaků → **alternativní strategie**

*L. macrochirus*

# *Lepomis* (maternity)



- Samec (bourgeois male) hlídá hnízdo s jikrami
  - *L. marginatus*, *L. punctatus*, *L. auritus*  
samec má v hnízdě jikry od několika samic
  - *L. macrochirus*  
tři typy samců (alternativní reprodukční strategie):
    - bourgeois (> 7 let, staví hnízda)
    - mladí samci (vjíždějí do hnízd a vypouštějí spermie)  
– „sneakers“
    - nespárovaní staří samci (napodobují samice)
- 20% potomků není od bourgeois

# *Skladování spermií*



*Chrysemys picta*

- Dny u savců  
týdny u ptáků nebo hmyzu  
měsíce u mloků  
roky u hadů a želv
- *Chrysemys picta*  
mikrosatelity → po 3 roky mláďata od stejného otce, opakované páření se stejným otcem je nepravděpodobné

Varan – Zoo Liberec – mláďata po dvou letech od úmrtí samce



Photo Tavi Grepp

# *Halichoerus grypus*

## tuleň kuželozubý

*Wilmer et al. 1999*

- Kolonie, dimorfismus → polygynie
  - Dvě skotské kolonie North Rona a Isle of May
  - Vzoroky ze zhruba deseti let
  - 9 mikrosatelitových lokusů
  - IDENTITY, NEWPAT, CERVUS
- 
- Samci z centra kolonie úspěšnější (až 30x)
  - Samci reprodukčně aktivní 10 let i déle
  - Pro více než polovinu mláďat nenalezení otci → role páření ve vodě





# Ptáci

- Řada druhů považována za monogamní (pěvci)
- Ale u 75 % druhů ptáků mimopárová mláďata
- Skutečně monogamní pěvci – jen 14% (*Phylloscopus*)



*Sialia sialis*

8-35 % mimopárových mláďat



*Emberiza schoeniclus*

až 55% mimopárových mláďat



*Malurus cyaneus*

až 72% mimopárových mláďat



# Odhad % mimopárových mlád'at

- Dostatečný vzorek pro odhad
- Více než 200

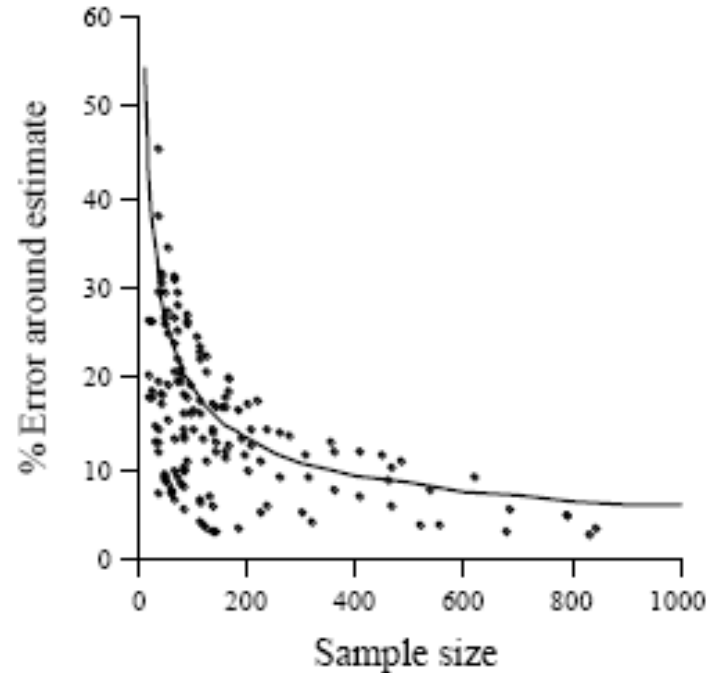


Fig. 2 The magnitude of error around actual estimates of EPP levels against the sample size of those studies. '% error' on the vertical axis refers to the magnitude of the difference between the upper and lower 95% confidence intervals around an estimate. The line plotted is this '% error' for a hypothetical population with a rate of 15% EPP across different sample sizes.

# Vzdálenější příbuznost jedinců

Sestry, bratři, sestřenice, bratřenci



Studium sociálních systémů

# „Social breeding“

- eusocialita = dělba práce – „hodně dělníků, málo plemeníků“



*Heterocephalus glaber*



*Synalpheus regalis*



termity a jiný hmyz

- „helpers“ – často se rozmnoží v následujících sezónách (cca 3% ptáků, hodně druhů savců a ryb)





# Příbuzenská selekce

- koncept inkluzivní fitness (Hamilton 1964)
- $r * b > c$  ( $r$  = relatedness,  $b$  = benefits,  $c$  = costs)
- za těchto podmínek je lepší pomáhat příbuzným než se sám množit

pomoc se sourozenci  
( $r = 0.5$ )

cena za pomoc  
(=sám se nemnoží)

$$(0.5) * (2) = 1$$

počet odchovaných  
sourozenců

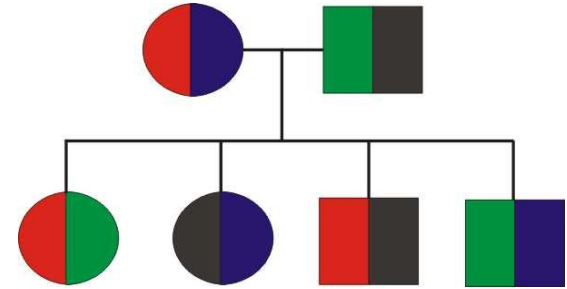
$$(0.5) * (3) > 1$$

Je výhodné se starat o sourozence (na úkor vlastního rozmnožování),  
jen pokud jsou 3 a více

# Příbuzenský koeficient $r$

- Celková „identity by descent“
- Diploidní organismy

- Jednovaječná dvojčata 1
- Rodič – potomek 0,5
- Sourozenci 0,5
- Nevlastní sourozenci 0,25
- Prarodiče a vnoučata 0,25
- Bratřenci a sestřence 0,125
- Nepříbuzní 0



- RELATEDNESS, KINSHIP (Mac), ML-RELATE (WinXP) odhad pomocí ML např. z mikrosatelitů
- BAYES

# Výpočet příbuznosti

$$r = \frac{\sum(p_y - p)}{\sum(p_x - p)}$$

Queller and Goodnight 1989

Jedinec	Lokus 1	Lokus 2
Helper – x	120/120	116/118
Matka rodu – y	120/122	118/118

Alela	$p_x$	$p_y$	$p$ = background allele frequency
120	1.0	0.5	0.65
116	0.5	0	0.20
118	0.5	1.0	0.35

$$r = \frac{[(0.5-0.65)+(0-0.20)+(1-0.35)]}{[(1-0.65)+(0.5-0.20)+(0.5-0.35)]} = \frac{0.30}{0.80} = \mathbf{0.375}$$

- ideální je cca 30-40 mikrosatelitů nebo > 100 SNPs



# *Cynopterus sphinx*

kaloň krátkonosý  
*Storz et al. 2001*



- Kolonie složené z harémů, v harému samec a 1 až 37 samic
- Příbuzenská struktura kolonie a harémů (kin structure)?
- 10 mikrosatelitových lokusů,  $r$ , KINSHIP
- $r$  blízké nule  $\rightarrow$  jedinci v kolonii jsou nepříbuzní
- Zásadní role disperze (mláďata z kolonie se v dospělosti nedrží pospolu)

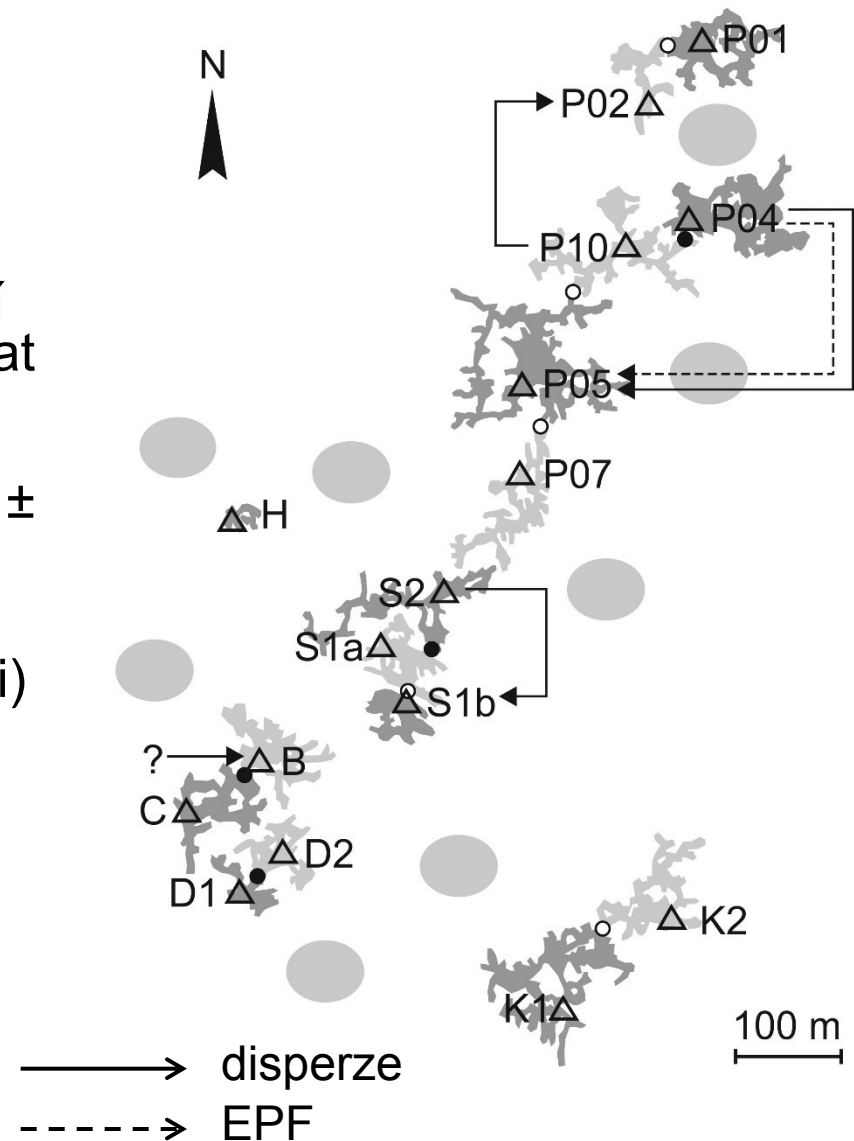
	Average pairwise $r$	95% confidence interval	No. pair- wise combin- ations
Mother-offspring pairs	0.494	0.478 to 0.511	185
Father-offspring pairs	0.508	0.485 to 0.530	118
Half-siblings	0.298	0.285 to 0.310	737
Pups (1997 cohort)	0.036	0.027 to 0.044	2211
Pups (1998 cohort)	-0.002	-0.006 to 0.003	6903
Adult females	-0.008	-0.010 to -0.006	21,736
Adult males	-0.001	-0.014 to 0.013	666

- Kolonie, max. cca 15 jedinců
- Množí se jen 1 samice (královna) a 1 samec
- 16 kolonií, 8 mikrosatelitů, CERVUS, ML-RELATE
- Královna má mláďata i se samci, kteří nejsou v kolonii - 11.8% mladých zvířat nejsou potomky rezidentního samce
- Příbuznost dominantního páru =  $0.04 \pm 0.06$
- V kolonii i nepříbuzní jedinci (imigranti)



# *Fukomys anelli*

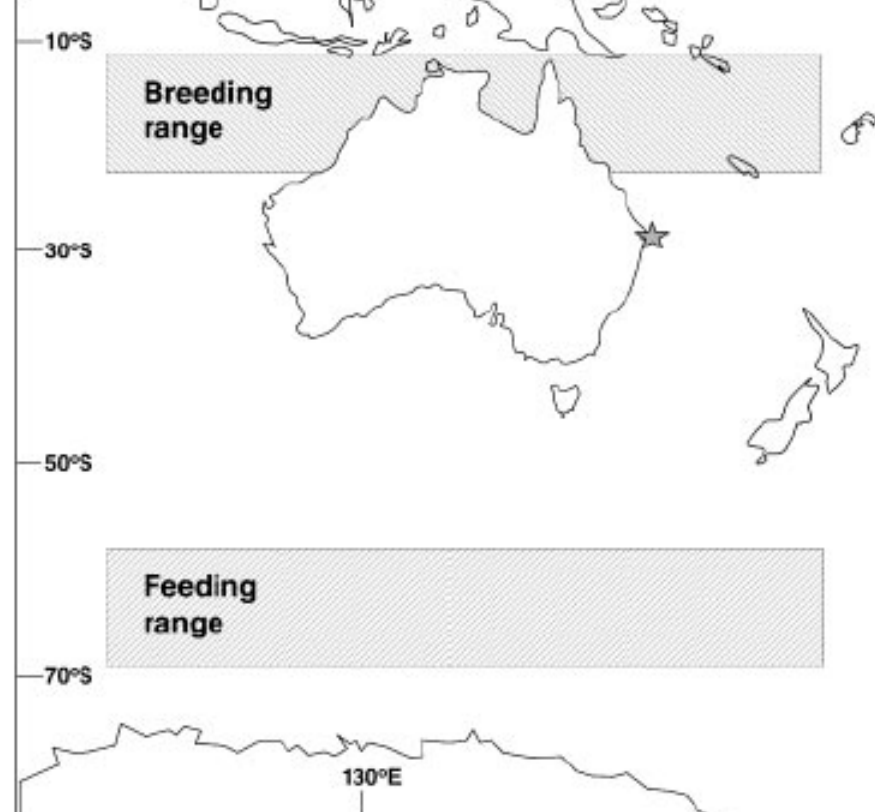
Patzenhauerová et al. 2013



# *Megaptera novaeangliae*

Valsecchi et al. 2002

- Cestují v malých skupinách
- Tvoří skupiny příbuzní?  
Kin selection?
- Mikrosatelity (8 lokusů), KINSHIP, NEWPAT
- Jediní příbuzní ve skupinách byly matky a jejich potomci.
- Kromě nich hodnoty  $r$  stejné jako při sloučení skupin dohromady
- Kin selection skupiny nevysvětluje





# Haplodiploidie

Hamilton (1972) - inkluzivní fitness

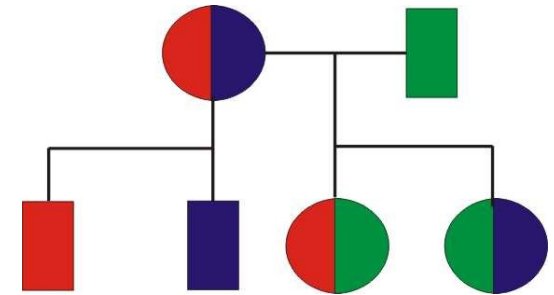
- haploidní, diploidní

- Jeden otec

- **Sestry:**
- **Matka – dcera**

0,75

0,5

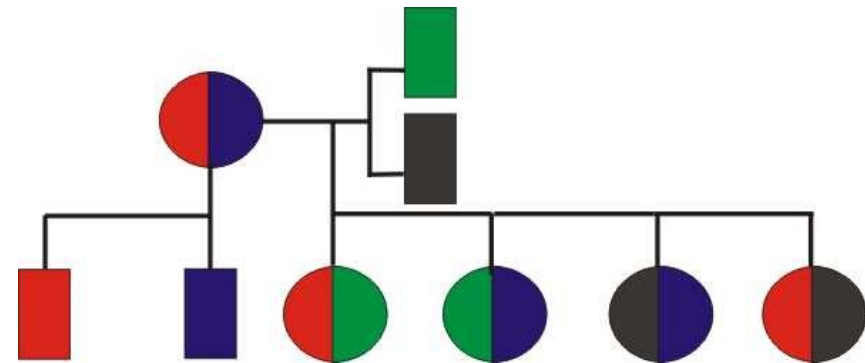


- Více otců

- **Sestry:**
- **Matka – dcera**

0,25 - 0,75

0,5



- Více matek i otců

$r$  různé

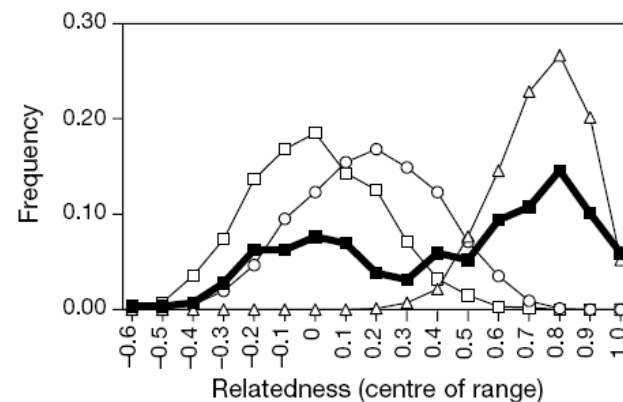


# *Polistes dominulus*

(dříve *P. gallicus*)

vosík francouzský *Queller et al. 2000*

- Hnízdo bez ochranného obalu
- Zakládá často více přezimovavších samic
- Dominantní samice klade vajíčka (>90%), subordinátní se starají o potravu
- Kin selection?  
(Jsou si samice příbuzné?)
- Ve třetině případů jsou samice nepřibuzné  
(ML 35% nepřibuzné  
7% sestřenice  
56% sestry)
- Jediná výhoda – nahrazení dominantní samice, pokud zahyne
- **Výjimka u sociálního hmyzu!**



**Figure 3** Observed relatedness distribution of *P. dominulus* foundresses (filled squares) and expected distributions for several relationships, grouped into intervals of width 0.1. The filled squares show the observed distribution for all relatedness estimates of foundress nestmate pairs, from the four foundress collections of Fig. 2. The other distributions, used in the likelihood analysis, show the distributions of relatedness estimates for simulated non-relatives (open squares, true  $r = 0$ ), cousins (open circles, true  $r = 3/16$ ) and full sisters (open triangles; true  $r = 3/4$ ).



# CERVUS

- využívá výpočtu „likelihood“
- přidává váhu jednotlivým alelám na základě jejich frekvence
- možnost nastavení míry genotypizační chyby, bere v úvahu nulové alely

# CERVUS

## PRVNÍ VERZE

- Marshall TC, Slate J, Kruuk LEB, Pemberton JM (1998) **Statistical confidence for likelihood-based paternity inference in natural populations.** *Molecular Ecology* 7 (5) , 639–655

## AKTUÁLNÍ VERZE 3.0

- Kalinowski ST, Taper ML, Marshall TC (2007) **Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment.** *Molecular Ecology* 16: 1099-1006.

DOWNLOAD - [www.fieldgenetics.com](http://www.fieldgenetics.com)

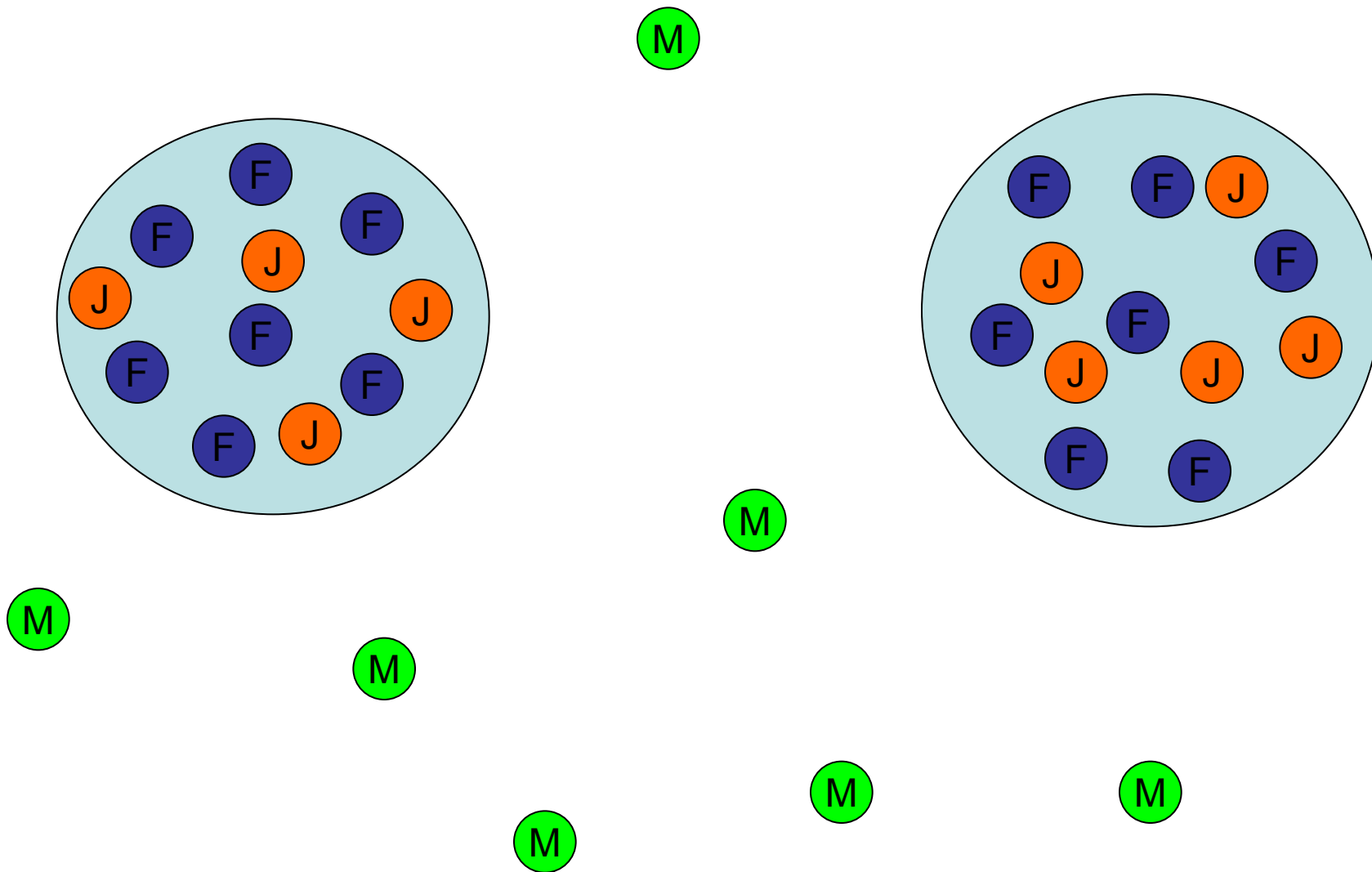
# CERVUS

Možnost použití

- a) hledáme otce
- b) hledáme matku
- c) hledáme oba rodiče
- d) hledáme oba rodiče, ale neznáme pohlaví jednotlivých kandidátů



# Typický příklad: Maternita a paternita u netopýrů



*Pipistrellus nathusii* – jižní Čechy (Jahelková, Hulva, Bryja)

# Omezení v programu CERVUS

- pouze diploidní data
- pouze kodominantní lokusy (např. mikrosatelity)
- lokusy nesmí být ve vazbě (test v Genepop 4.0)
- jen malé odchylky od HW rovnováhy

# Vstupní soubor



- 2 možnosti
  - vytvoření csv souboru v MS Excell
  - konverze z formátu jiného programu (Genepop, Kinship, Genetix)

# Vytvoření csv souboru

	A	B	C	D	E	F
1	ID	Loc1a	Loc1b	loc2a	loc2b	loc3a
2	1	155	155	128	130	256
3	2	155	157	124	128	258
4	3	159	152	124	124	254
5	4	155	157	126	124	254

Save As dialog box showing the file name 'cervus\_test' and the location 'D:\Cervus'.

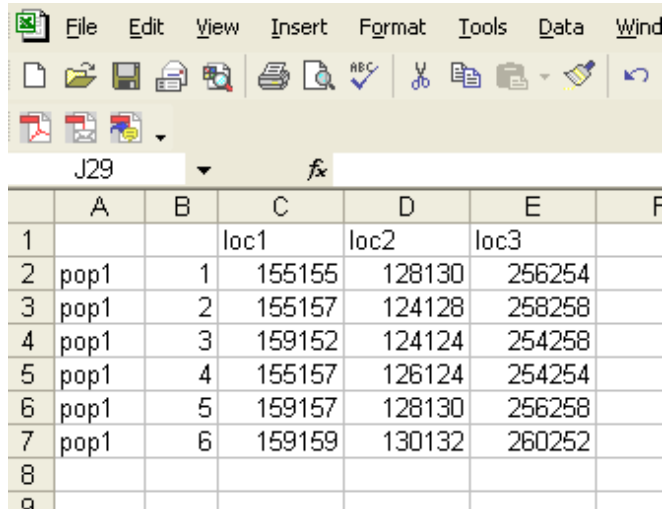
Windows Explorer showing the file 'cervus\_test.csv' selected in the D:\Cervus folder. A context menu is open, and 'Open With' is selected, showing 'Microsoft Excel' and 'WordPad' as options.

WordPad window showing the CSV data from the Excel table. A 'Replace' dialog box is open, with 'Find what' set to ';' and 'Replace with' set to ','.

```
ID;Loc1a;Loc1b;loc2a;loc2b;loc3a;loc3b
1;155;155;128;130;256;254
2;155;157;124;128;258;258
3;159;152;124;124;254;258
4;155;157;126;124;254;254
5;159;157;128;130;256;258
6;159;159;130;132;260;252
```

# Input file – konverze z jiného formátu

## Příklad – konverze z Genetix



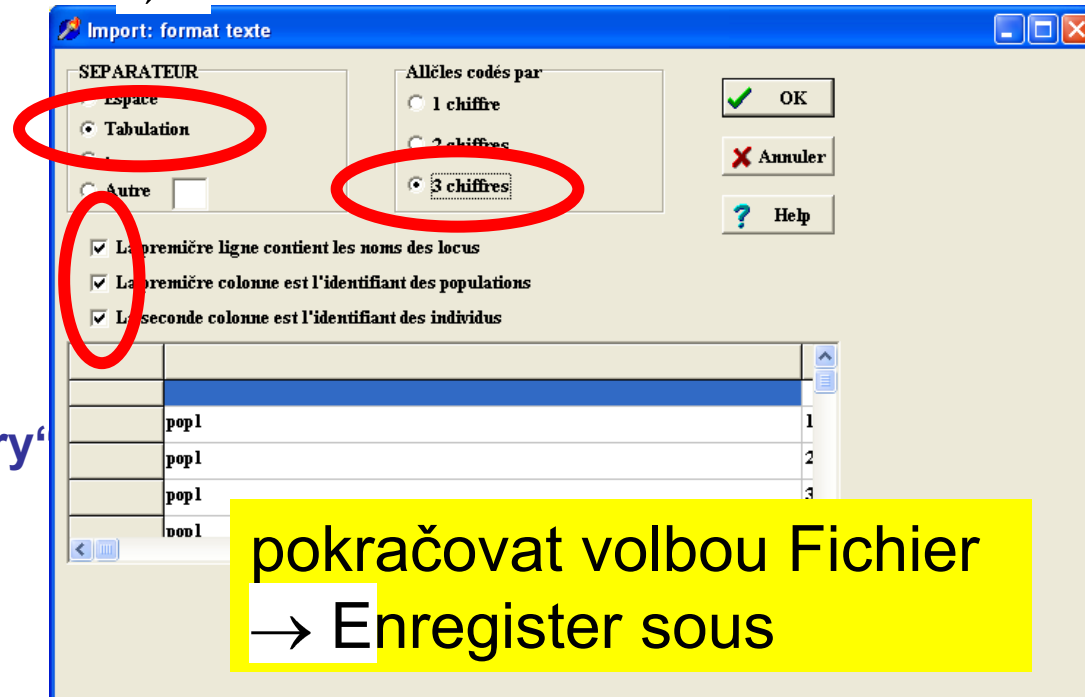
	A	B	C	D	E	F
1			loc1	loc2	loc3	
2	pop1	1	155155	128130	256254	
3	pop1	2	155157	124128	258258	
4	pop1	3	159152	124124	254258	
5	pop1	4	155157	126124	254254	
6	pop1	5	159157	128130	256258	
7	pop1	6	159159	130132	260252	
8						
9						

funkce CONCATENATE

Uložit jako „Text oddělený tabulátory“

Cesta v Genetix:

Fichier → Importer → vybrat soubor,  
označit Text avec séparateur



Import: format texte

SEPARATEUR

Espace

Tabulation

Autre

Allèles codés par

1 chiffre

2 chiffres

3 chiffres

La première ligne contient les noms des locus

La première colonne est l'identifiant des populations

La seconde colonne est l'identifiant des individus

OK

Annuler

Help

pop1 1

pop1 2

pop1 3

pop1

pokračovat volbou Fichier  
→ Enregistrer sous





# Načtení dat do Cervus

## Otevření souboru:

**.CSV**

File → View text file → vybrat soubor

## Z Genetix

Tools → Convert genotype file → Genetix to Cervus → vybrat soubor (Source file) a vybrat umístění a název konvertovaného souboru (Save as)

Cervus 3.0 - [Text file: cervus\_sk2.txt]

File Edit Analysis Tools Options Window Help

ID	Loc1a	Loc1b	loc2a	loc2b	loc3a	loc3b
1	155	155	128	130	256	254
2	155	157	124	128	258	258
3	159	152	124	124	254	258
4	155	157	126	124	254	254
5	159	157	128	130	256	258
6	159	159	130	132	260	252

Cervus 3.0 - [Genotype file conversion: cervus\_z\_genetix...]

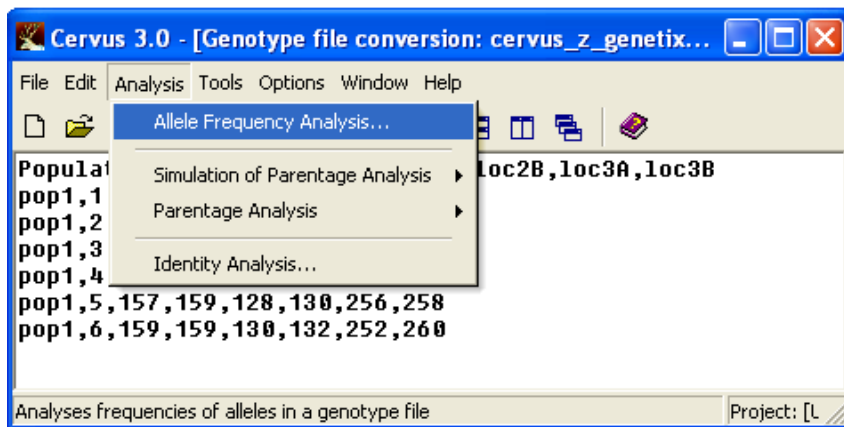
File Edit Analysis Tools Options Window Help

Population	ID	loc1A	loc1B	loc2A	loc2B	loc3A	loc3B
pop1	1	155	155	128	130	254	256
pop1	2	155	157	124	128	258	258
pop1	3	152	159	124	124	254	258
pop1	4	155	157	124	126	254	254
pop1	5	157	159	128	130	256	258
pop1	6	159	159	130	132	252	260

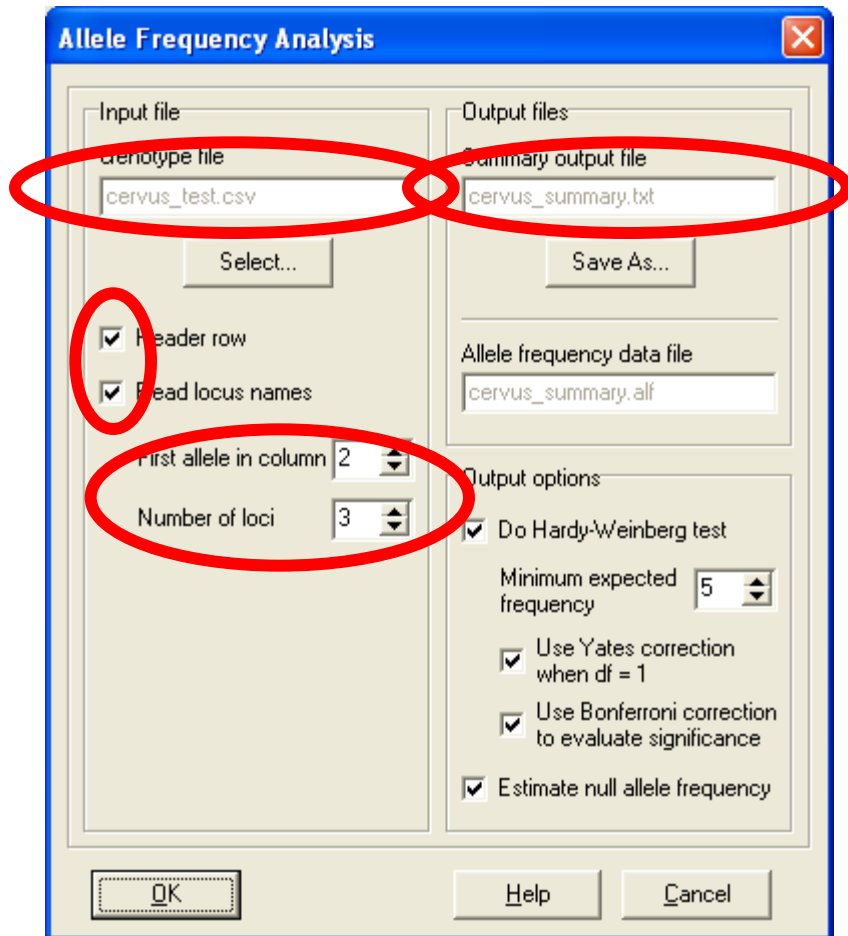
Project: [L

# Analýza dat

## 1) Frekvence alel



lépe počítat pouze pro  
dospělé, bez mláďat



# Analýza dat

Locus	k	N	HObs	HExp	PIC	NE-1P	NE-2P	NE-PP	NE-I	NE-SI	HW	F(Null)
loc1	4	6	0.667	0.773	0.652	0.721	0.555	0.385	0.141	0.431	NS	+0.0167
loc2	5	6	0.833	0.818	0.708	0.658	0.481	0.300	0.105	0.401	*	-0.0355
loc3	5	6	0.667	0.803	0.692	0.674	0.498	0.315	0.114	0.410	***	+0.0050

↑  
počet alel

↑  
počet zgenotypovaných  
jedinců

↑  
pozorovaná  
heterozygotnost

↑  
očekávaná  
heterozygotnost

↑  
míra informativnosti daného  
polymorfismu

↑   ↑   ↑   ↑   ↑  
non-exclusion probabilities

↑  
 $p$  pro test na nulové  
alely  
 $p$  pro test odchylky od HW  
rovnováhy

# Analýza dat

## 2) Simulace

Cervus 3.0 - [Allele frequency analysis: cervus\_summary.txt]

File Edit Analysis Tools Options Window Help

Allele Frequency Analysis...  
Simulation of Parentage Analysis ▶  
Parentage Analysis ▶  
Identity Analysis...

Maternity...  
Paternity...  
Parent Pair (Sexes Known)...  
Parent Pair (Sexes Unknown)...

\*\*\*\* Summary statistics \*\*\*\*

Locus	k	N	HObs	HExp	PIC	NE-1P	NE-2P	NE-PP	NE-I	NE-SI	HW	F(Null)
Loc1	3	5	0.600	0.733	0.586	0.782	0.635	0.486	0.189	0.467	ND	ND
loc2	5	6	0.833	0.818	0.708	0.658	0.481	0.300	0.105	0.401	ND	ND
loc3	5	6	0.667	0.803	0.692	0.674	0.498	0.315	0.114	0.410	ND	ND

Number of individuals: 6  
Number of loci: 3  
Mean number of alleles per locus: 4.33  
Mean proportion of individuals typed: 0.9444  
Mean expected heterozygosity: 0.7848  
Mean polymorphic information content (PIC): 0.6620

Project: [Unint...

Simulation of Parentage Analysis

Input file  
Allele frequency data  
cervus\_summary.alf  
Select...

Output files  
Summary output file  
cervus\_test\_sim.txt  
Save As...

Simulation data file  
cervus\_test\_sim.sim

Confidence  
Calculate confidence using:  
 LOD  
 Delta

Confidence levels:  
Relaxed (%) 80.00  
Strict (%) 95.00

Simulated genotypes  
Offspring 500  
Candidate fathers 20  
Prop. sampled 0.9000000  
Prop. loci typed 0.9444444  
Prop. heteromistyped 0.0100000  
Minimum typed loci 2

OK Options... Help Cancel

lépe počítat pouze pro  
dospělé, bez mláďat

# Analýza dat – výstup simulace

\*\*\*\* summary statistics \*\*\*\*

Critical values and success rates (one parent known):

Level	Confidence(%)	Delta Criterion	Tests	Success Rate
strict	95.00	0.51	8833	88%
Relaxed	80.00	0.00	9387	94%
Unresolved			613	6%

Critical values and success rates (neither parent known):

Level	Confidence(%)	Delta Criterion	Tests	Success Rate
strict	95.00	0.95	6815	68%
Relaxed	80.00	0.00	9471	95%
Unresolved			529	5%

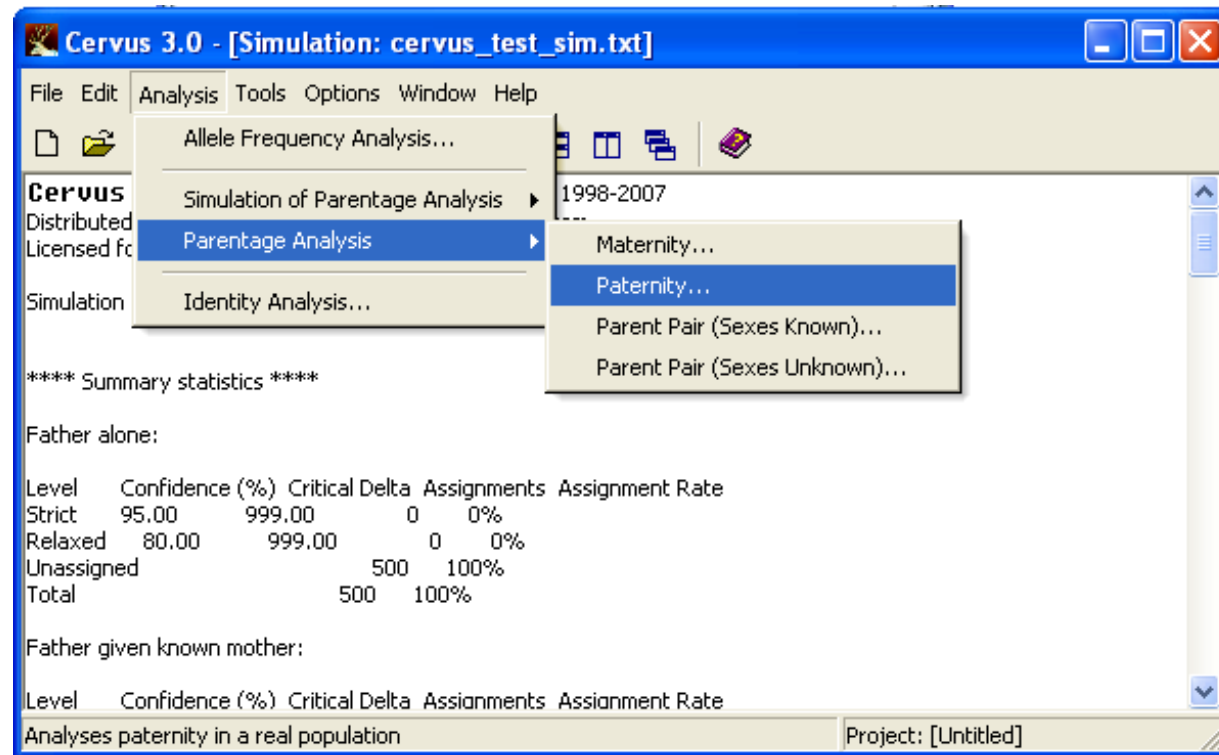
# Analýza dat

## 3) Vlastní určení paternity (a/nebo maternity)

### Kroky před

a) vytvoření souboru s mláďaty

b) vytvoření souboru s potenciálními otci (resp. matkami)



# Analýza dat

## 3) Vlastní určení paternity

Parentage Wizard - Step 1

Offspring data

Offspring file  
mladata.csv

Select...

Includes header row

Offspring ID in column 1

Includes known parents

ID in column 0

Includes candidate parents

First ID in column 0

OK

Cancel

<< Back

Next >>

Help



Parentage Wizard - Step 2

Candidate parent data

OK

Cancel

Candidate father file  
otcove.csv

Select...

<< Back

Next >>

Header row

Candidate parent IDs appear as:

One column for all offspring

One row for each offspring

One column for all offspring

Candidate ID in column 0

Help



# Analýza dat

Parentage Wizard - Step 3

Genetic data

Genotype file  
genotypy.csv  
Select...

Header row

ID in column 1

First allele in column 2

OK  
Cancel  
<< Back  
Next >>  
Help

Parentage Wizard - Step 4

Cervus files

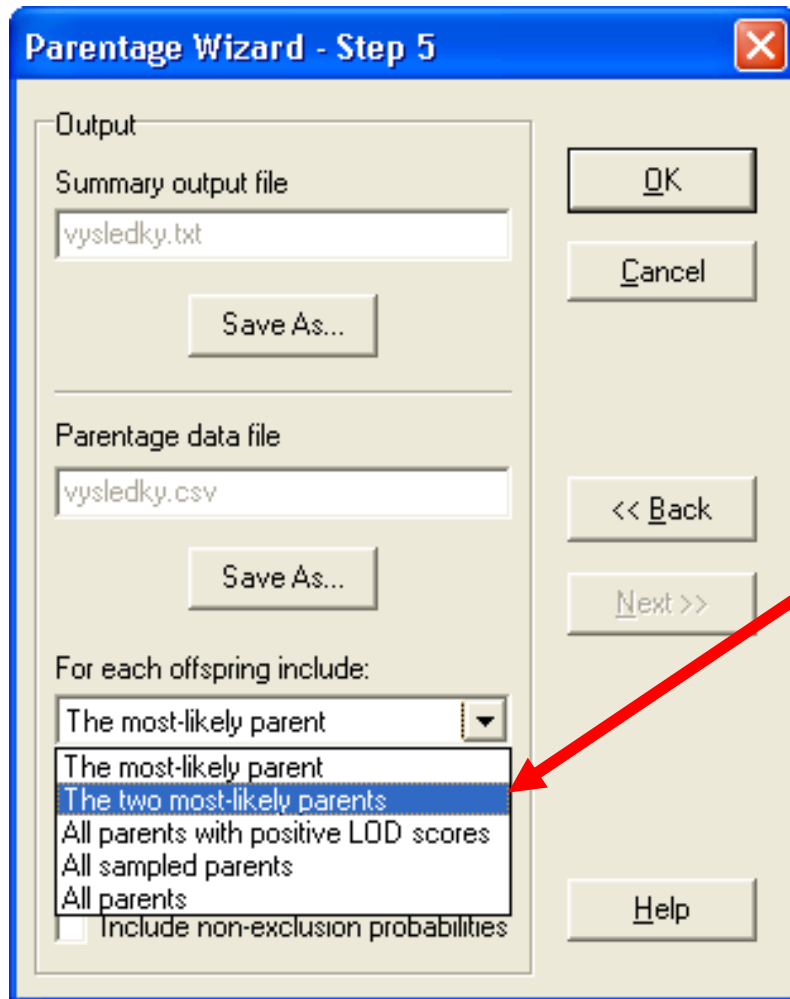
Allele frequency data  
cervus\_summary.alf  
Select...

Header row

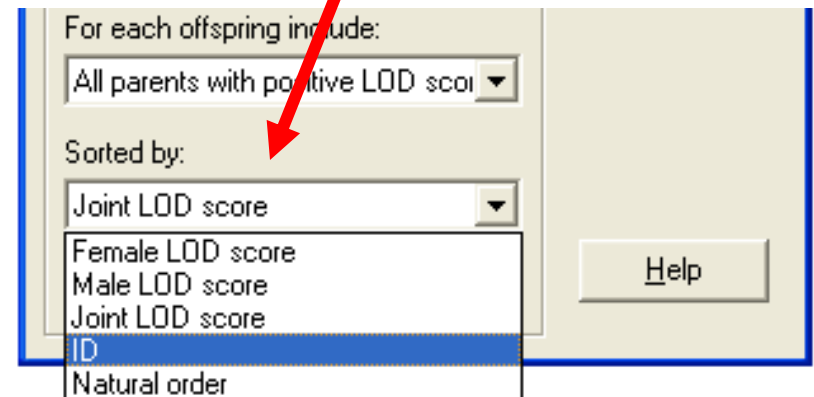
Simulation data  
cervus\_test\_sim.sim  
Select...

<< Back  
Next >>  
Help

# Analýza dat



výběr zobrazení  
výsledků analýzy



# Výstup

- před otevřením souboru s výstupem je třeba opět vyměnit čárky za středník
- mismatch – neseďící alely, nemožné zdědit po rodičích (může být ale způsoben nějakou chybou)
- non-exclusion probability – pravděpodobnost, že nepříbuzný kandidát není vyloučen jako rodič

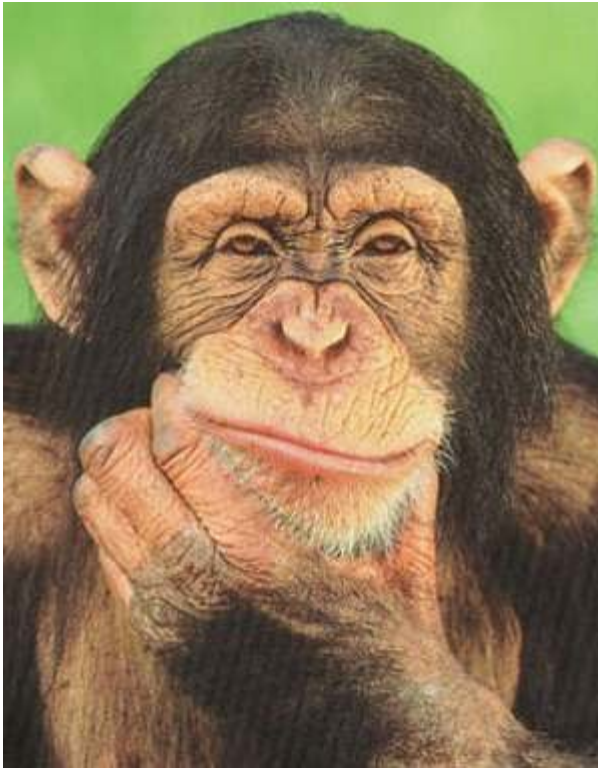
# Výstup

- LOD skóre – přirozený logaritmus pravděpodobnosti
  - menší než 0 – spíše není rodičem
  - rovný 0 – může být a nemusí být rodičem
  - větší než 0 – může být rodičem
- hodnota Delta – rozdíl mezi dvěma nejbližšími LOD skóre. Její kritická hodnota je vypočítávána při simulacích

# Výstup

- znak \* označuje otce/matky určené s 95% pravděpodobností
- znak + označuje otce/matky určené s 80% pravděpodobností
- znak – označuje nejpravděpodobnějšího otce/matku, který ale není určen jako rodič

# PŘÍKLADY



matka, otec, mládě,  
alela, LOD skóre....

# PŘÍKLAD 1

- máme 5 mlád'at, známe jejich matky
- existuje 10 potencionálních otců, každý z kandidátů může být otcem každého z mlád'at
- soubory: genotypy, mlád'ata, otcové

# PŘÍKLAD 2

- máme 5 mlád'at, neznáme jejich matky ani otce, zajímají nás jen matky
- existuje 10 potenciálních matek pro všechna mlád'ata
- soubory: jen genotypy2, ostatní soubory je třeba vytvořit



# PŘÍKLAD 3

- 10 mlád'at, neznáme ani otce, ani matku
- k dispozici genotypy 10 samců a 15 samic
- soubory – jen genotypy3, zbytek je třeba vytvořit

# PŘÍKLAD 4

- opět 5 mlád'at, známe jejich matky
- pro každé mládě existují 4 kandidátní otcové (označení kand, číslo mláděte, písmeno a-d)
- soubory – pouze genotypy<sup>4</sup>, zbytek třeba vytvořit