

# Molekulární identifikace

Druh, jedinec, pohlaví



evropský  
sociální  
fond v ČR



EVROPSKÁ UNIE



MINISTERSTVO ŠKOLSTVÍ,  
MLÁDEŽE A TĚLOVÝCHOVY



OP Vzdělávání  
pro konkurenceschopnost



INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

# Definice druhu

- Koncept biologického druhu (Mayr, 1942) – RIM = post- nebo prezygotické bariéry toku genů (ne vždy tak jednoduché – hybridní zóny, alopatická speciace, asexuální druhy atd.)
- Ochranařská biologie – je nutno rozhodnout o taxonomické jednotce (druhu), která vyžaduje pozornost
- Forezní genetika, vývojová stadia bez determinačních znaků, identifikace kořisti v trávicím traktu predátorů, atd.

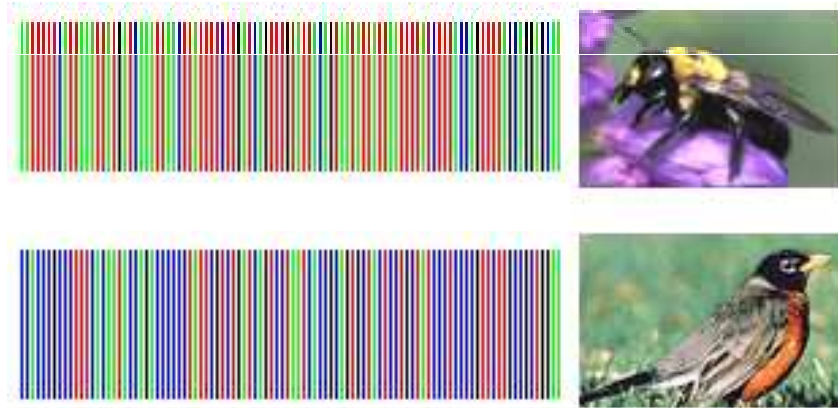
# What *is* DNA Barcoding?



- **Definition:** Derivation of short DNA sequence(s) that enables species identification or recognition in a particular domain of life (e.g., eucaryotes).
- **Barcode of Life Initiative (BOLI)** will resolve barcodes for named species and use a barcoding approach to assess undescribed biological diversity.

# BARCODING LIFE, ILLUSTRATED

Goals, Rationale, Results



Mark Stoeckle, The Rockefeller University

Paul E. Waggoner, Connecticut Agricultural Experiment Station

Jesse H. Ausubel, Alfred P. Sloan Foundation

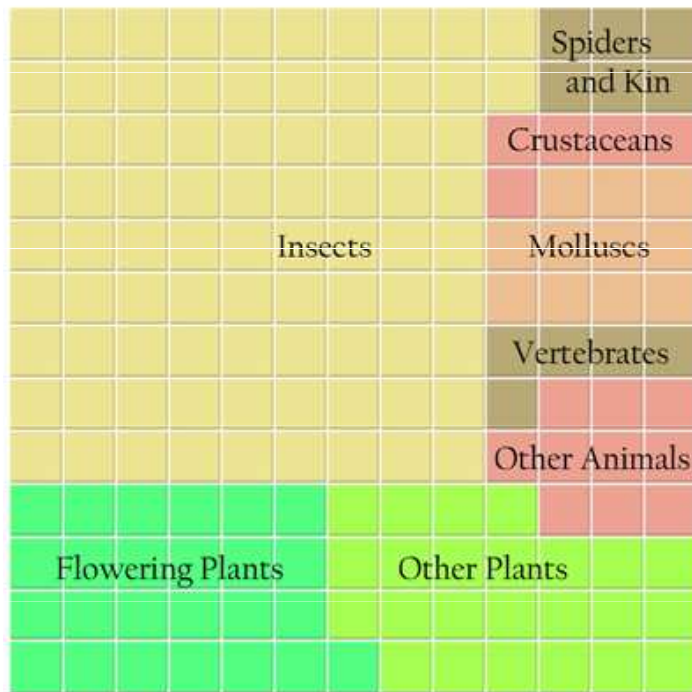


Barcoding is a standardized approach to identifying plants and animals by minimal sequences of DNA, called DNA barcodes.

**DNA Barcode:** A short DNA sequence, from a uniform locality on the genome, used for identifying species.

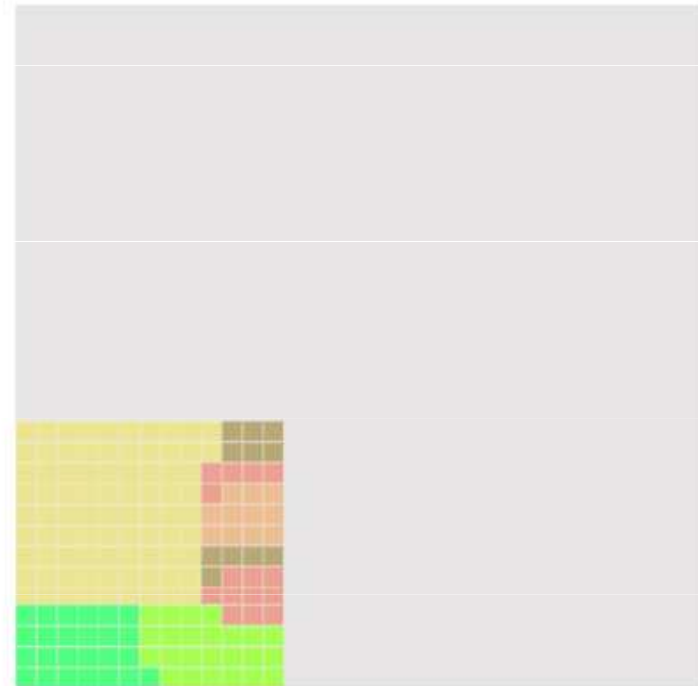
# Why barcode animal and plant species?

**Known Biodiversity**  
*Approximately 1.7 million named species of plants and animals.*

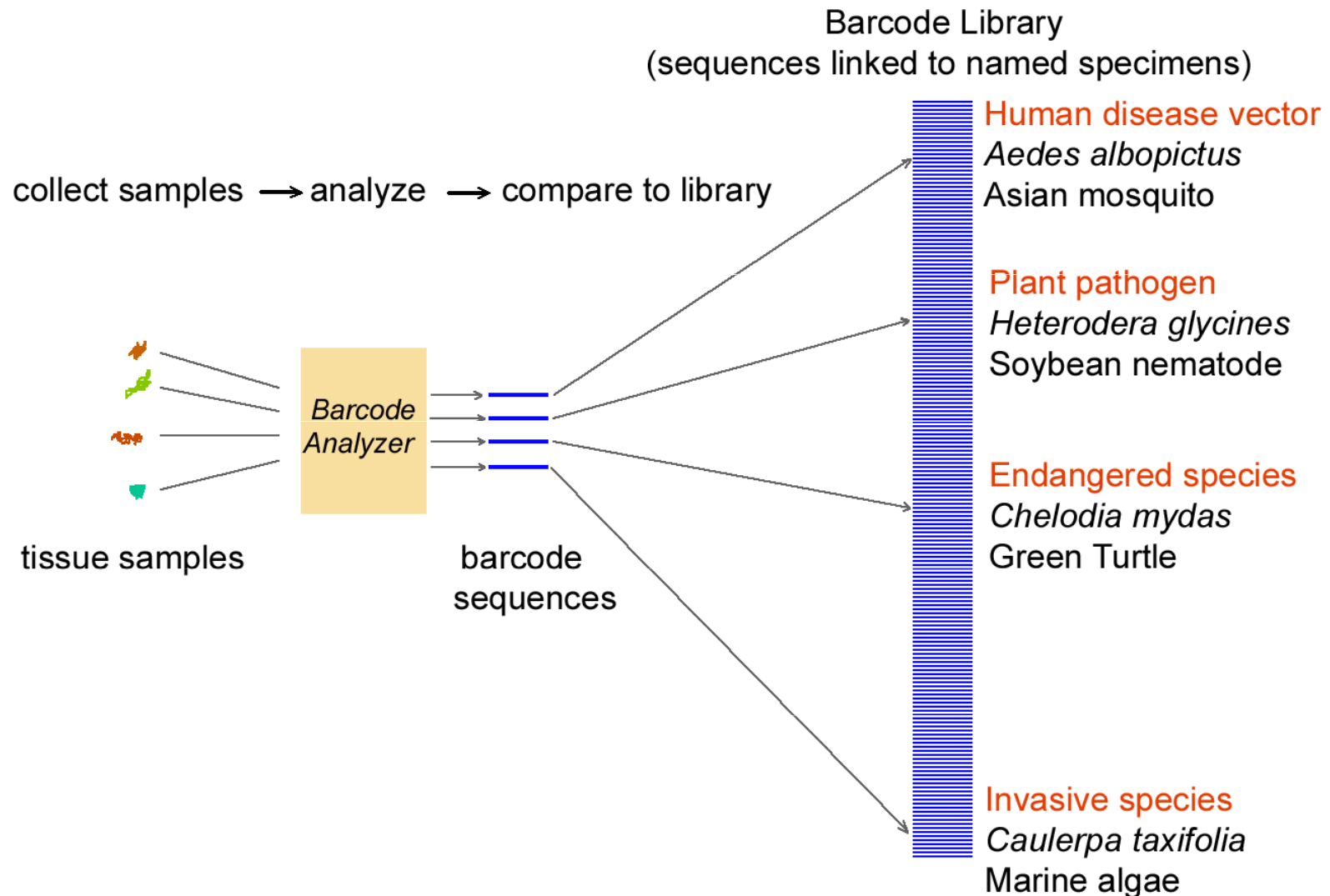


1 square = 10,000 species

**Estimated Biodiversity**  
*10 million species*



## 2. What are the benefits of standardization?



Researchers have developed numerous ways to identify species by DNA, typically tailoring the approach to answer a specific question in a limited set of species (e.g. phylogeny).

Like convergence on one or a few railroad gauges, **barcoding aims to capture the benefits of standardization**, which typically lowers costs and lifts reliability, and thus speeds diffusion and use.

- Results so far suggest that a **mitochondrial gene will enable identification of most animal species.**
- **Focus to date:** For animals, a 658 base-pair fragment of the mitochondrial gene, cytochrome oxidase subunit I (mtCOI).
- For plants, mitochondrial genes do not differ sufficiently to distinguish among closely related species. Promising approaches to standardize plant identification use one or possibly more barcode regions are under development.
- For bacteria, a 16S-rDNA emerges as very useful marker (especially in using next-generation sequencing)

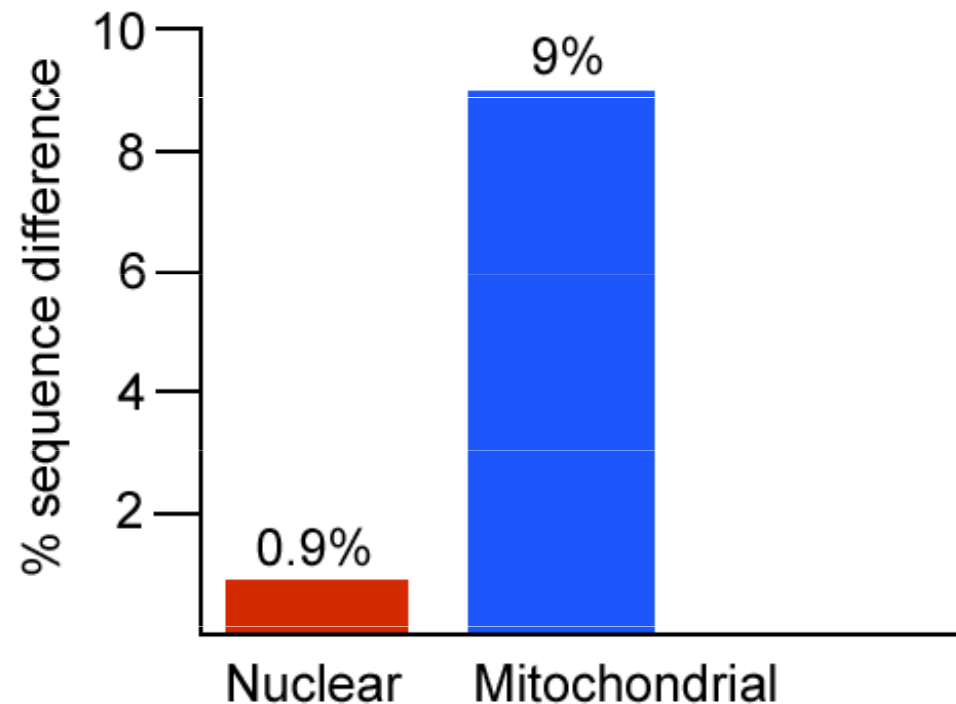
# Why barcode animals with mitochondrial DNA?

Mitochondria, energy-producing organelles in plant and animal cells, have their own genome. Twenty years of research have established the utility of mitochondrial DNA sequences in differentiating among closely-related animal species.

Four properties make mitochondrial genomes especially suitable for identifying species:

Greater differences among species, on average 5- to 10-fold higher in mitochondrial than in nuclear genes. Thus shorter segments distinguish among species, and because shorter, less expensively.

Average sequence differences in nuclear and mitochondrial DNA between human and chimp



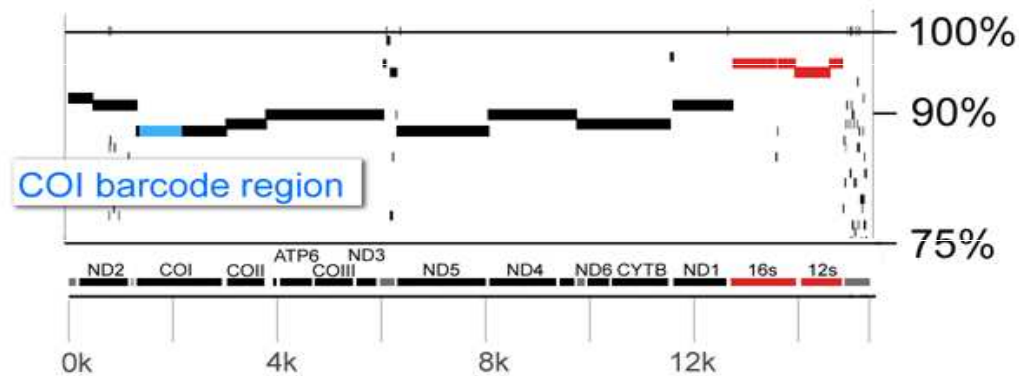


- **Copy number.** There are 100-10,000 more copies of mitochondrial than nuclear DNA per cell, making recovery, especially from small or partially degraded samples, easier and cheaper.
- **Relatively few differences within species** in most cases. Small intraspecific and large interspecific differences signal distinct genetic boundaries between most species, enabling precise identification with a barcode.
- **Introns, which are non-coding regions interspersed between coding regions of a gene, are absent from mitochondrial DNA** of most animal species, making amplification straightforward. Nuclear genes are often interrupted by introns, making amplification difficult or unpredictable.

Cytochrome c oxidase I (COI) contains differences representative of those in other protein-coding genes. Possible gains in accuracy or cost using a different protein-coding gene would likely be small.

Percent identity plot (PIP) analysis of complete mitochondrial genomes

*Anopheles gambiae*  
vs.  
*A. quadrimaculatus*



*Homo sapiens*  
vs.  
*Pan troglodytes*



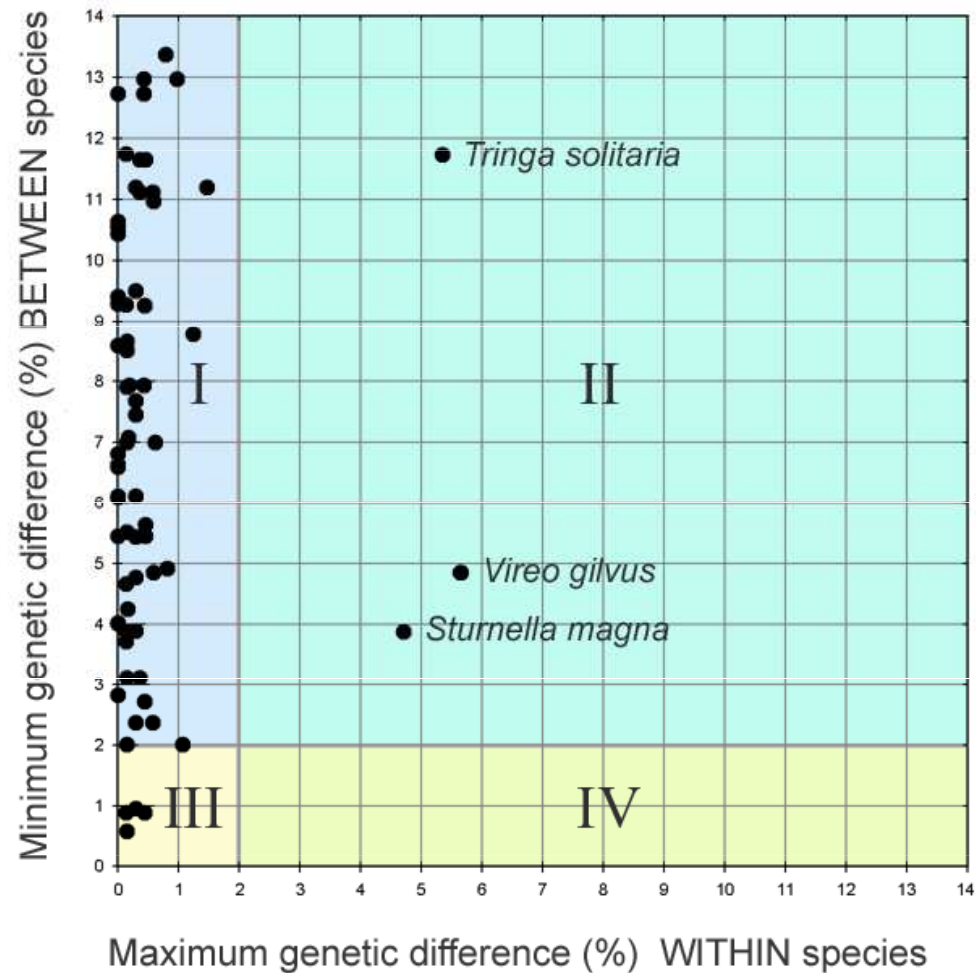
— protein coding gene  
— ribosomal RNA gene

## What do barcode differences among and within animal species studied so far suggest?

- Barcodes identify most animal species unambiguously.
- Approximately 2-5% of recognized species have shared or overlapping barcodes with closely-related species. Many of the species with overlapping barcodes hybridize regularly.
- In all groups studied so far, distinct barcode clusters with biologic co-variation suggest cryptic species.

Barcoding  
North  
American birds  
highlights  
probable  
cryptic species

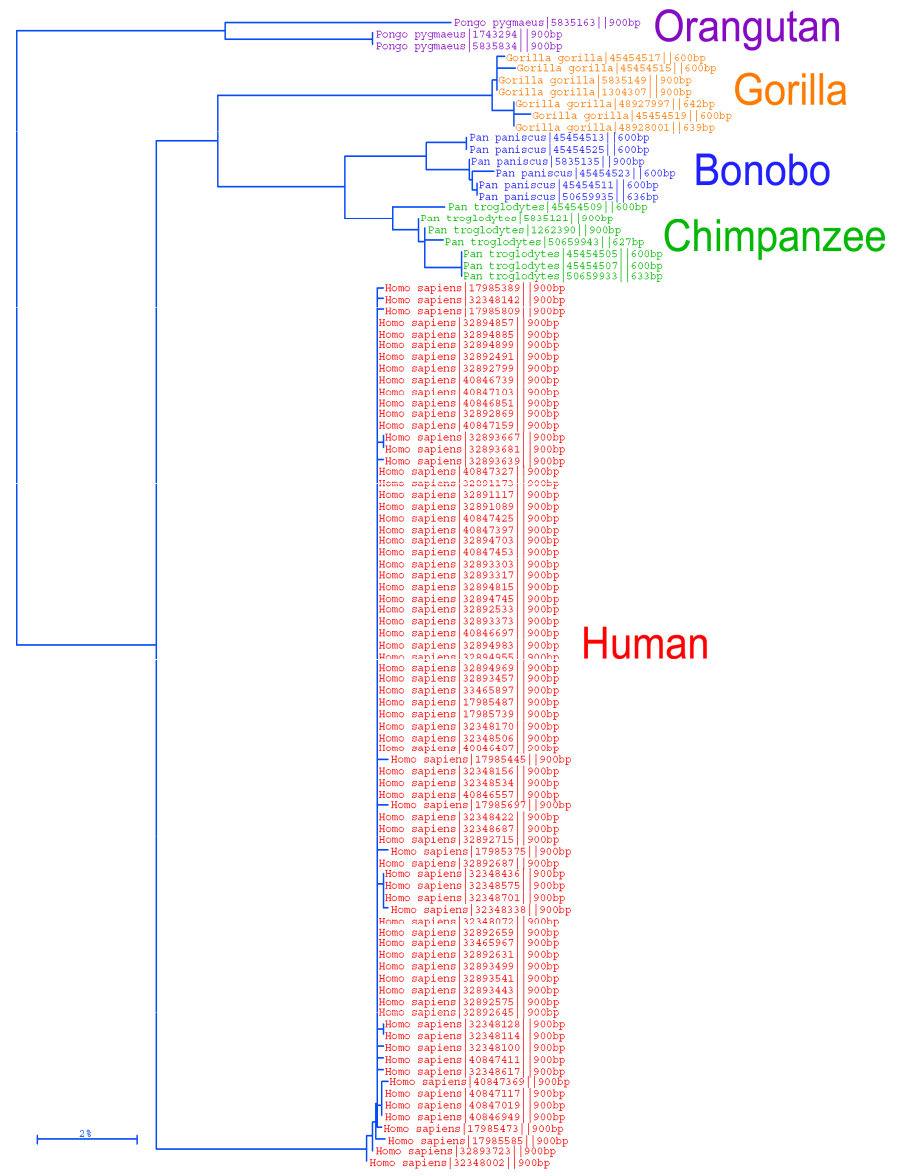
Interspecific vs. intraspecific COI barcode differences



Results for 73 species of North American are shown. Quadrants represent different categories of species:

- I. consistent with current taxonomy
- II. probable lumped species (candidate for taxonomic split)
- III. recent divergence, hybridization, or synonymy.
- IV. probable taxonomic misidentification

Barcodes affirm the unity of the species *Homo sapiens*. Comparisons show we differ from one another by only 1 or 2 nucleotides out of 648, while we differ from chimpanzees at 60 locations and gorillas at 70 locations.



A barcoder?



Mark Stoeckle The Rockefeller University  
Paul Waggoner Connecticut Agricultural Experiment Station  
Jesse Ausubel Alfred P. Sloan Foundation

# **Maintaining Credibility Amongst Taxonomic Community**

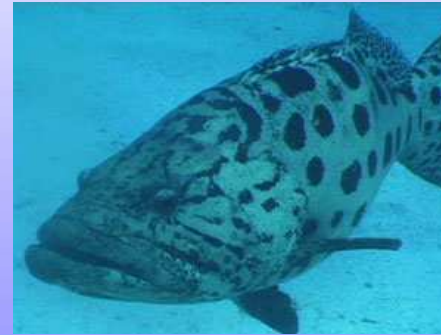
Recognizing the intellectual content of taxonomy

Demonstrating the utility of DNA barcoding in classical taxonomy problems

Integrating DNA barcoding approaches with classical taxonomy and taxonomic methods



# What *isn't* DNA Barcoding?

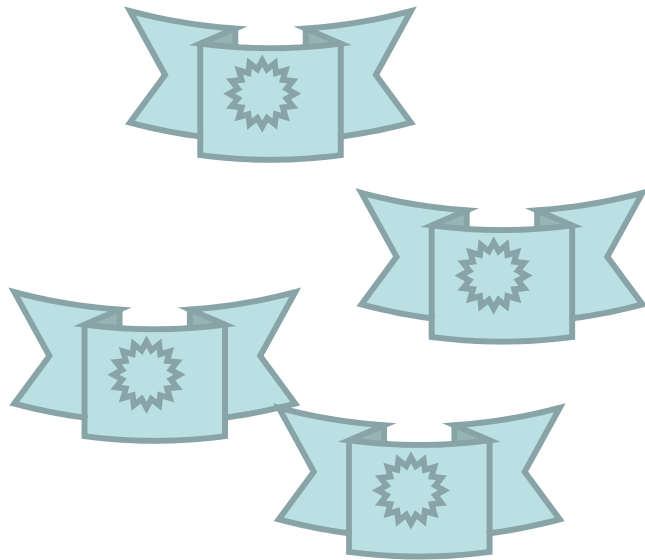


- It is not intended to, in any way, supplant or invalidate existing taxonomic practice.
- It is not DNA taxonomy; it does not equate species identity, formally or informally, with a particular DNA sequence.
- It is not intended to duplicate or compete with efforts to resolve deep phylogeny (e.g., *Assembling the Tree of Life*, ATOL).

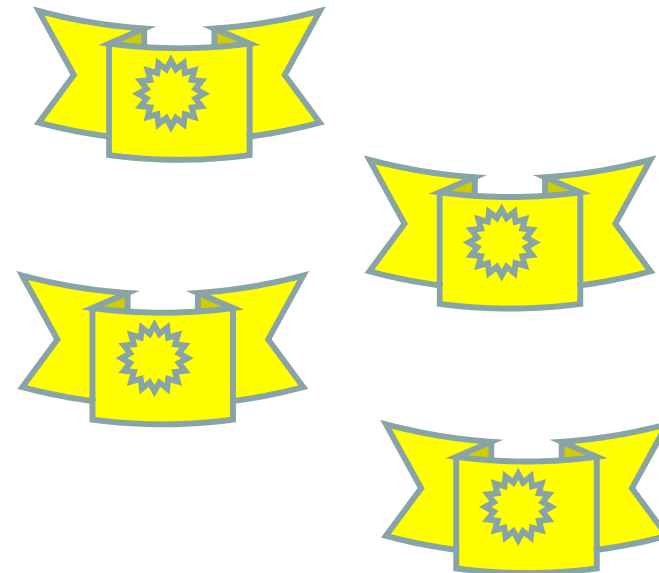
# What are the main limits to barcoding encountered so far?

- Groups with little sequence diversity
- Incomplete lineage sorting
- Resolution of recently diverged species
- Hybrids
- Nuclear pseudogenes

# Příklad: *Myotis blythii* vs. *Myotis myotis* - introgrese mtDNA

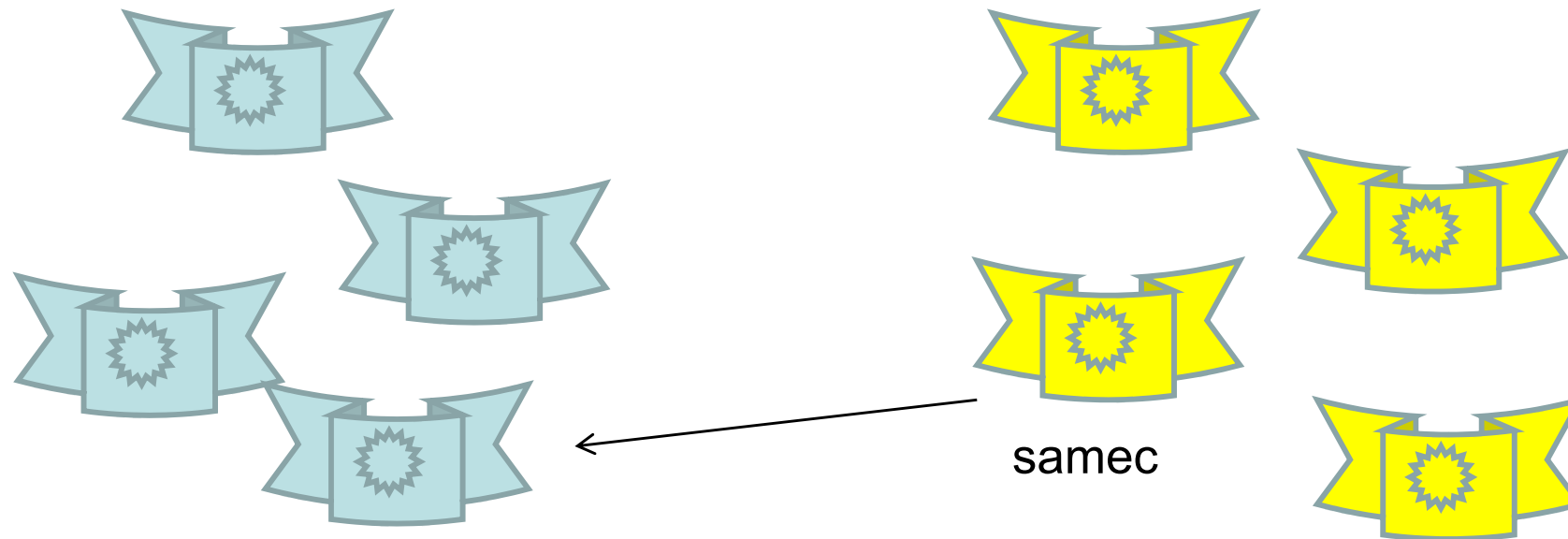


M. myotis - Evropa



M. blythii - Asie

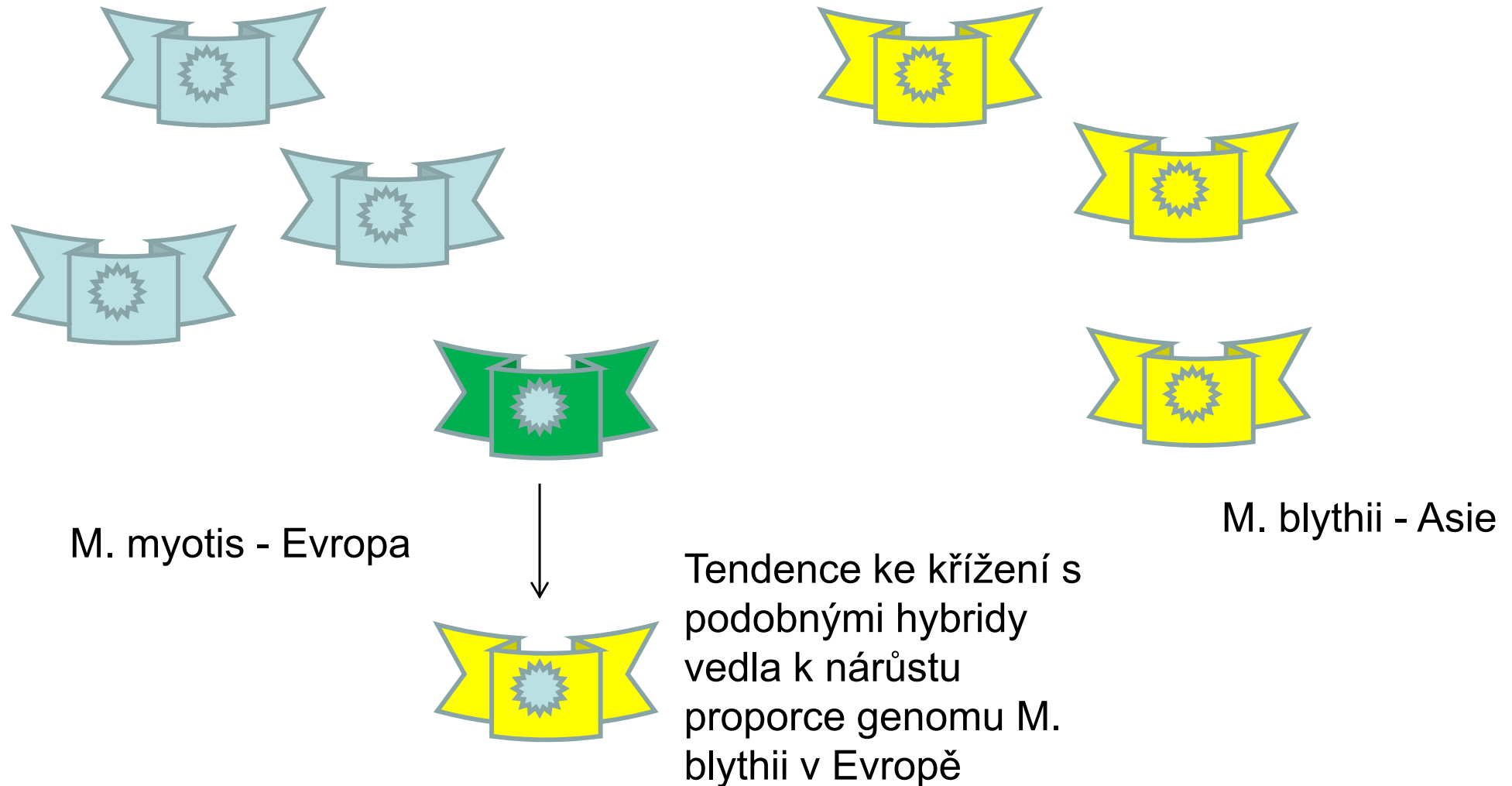
# Příklad: *Myotis blythii* vs. *Myotis myotis* - introgrese mtDNA



*M. myotis* - Evropa

*M. blythii* - Asie

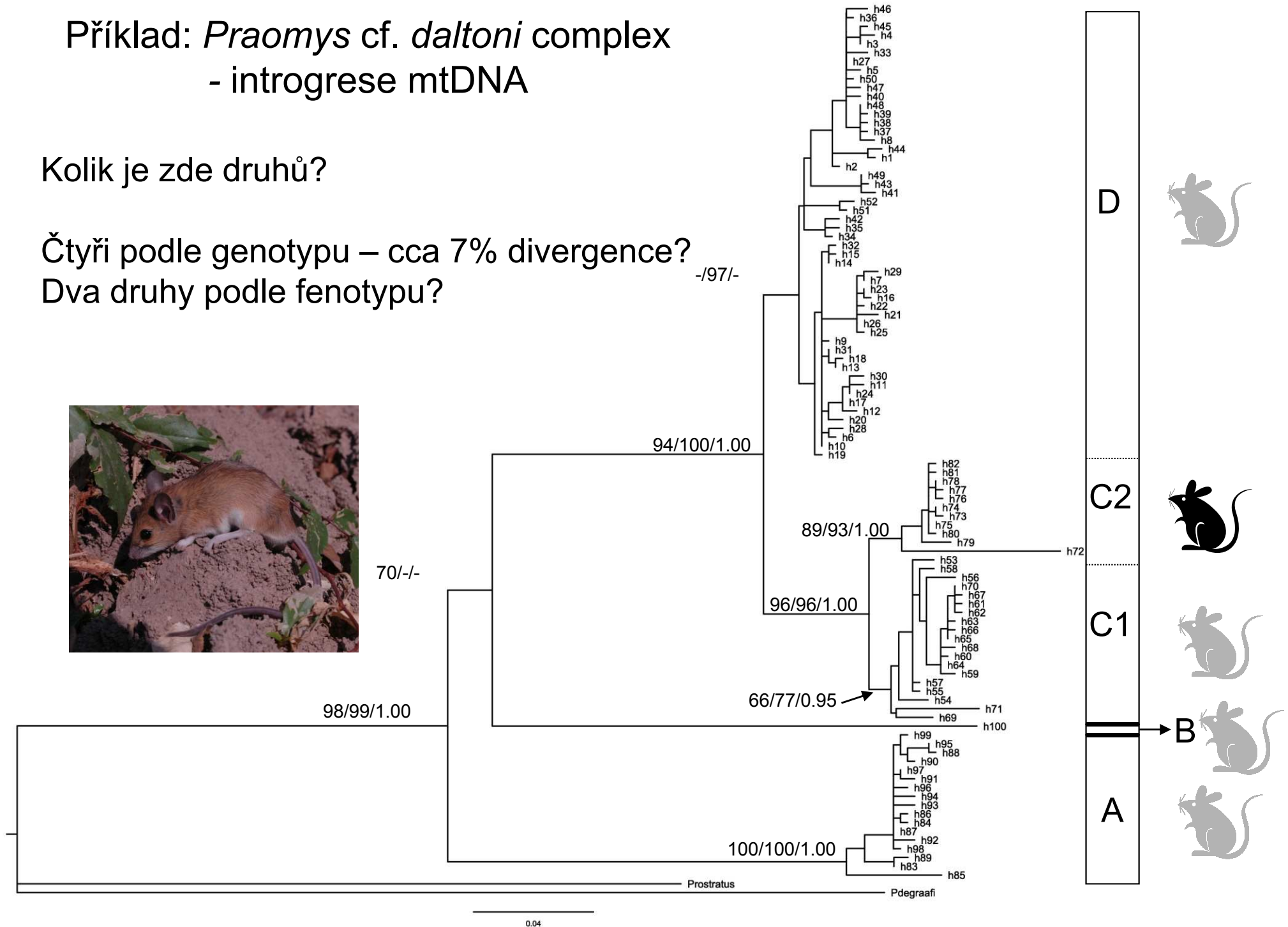
# Příklad: *Myotis blythii* vs. *Myotis myotis* - introgrese mtDNA

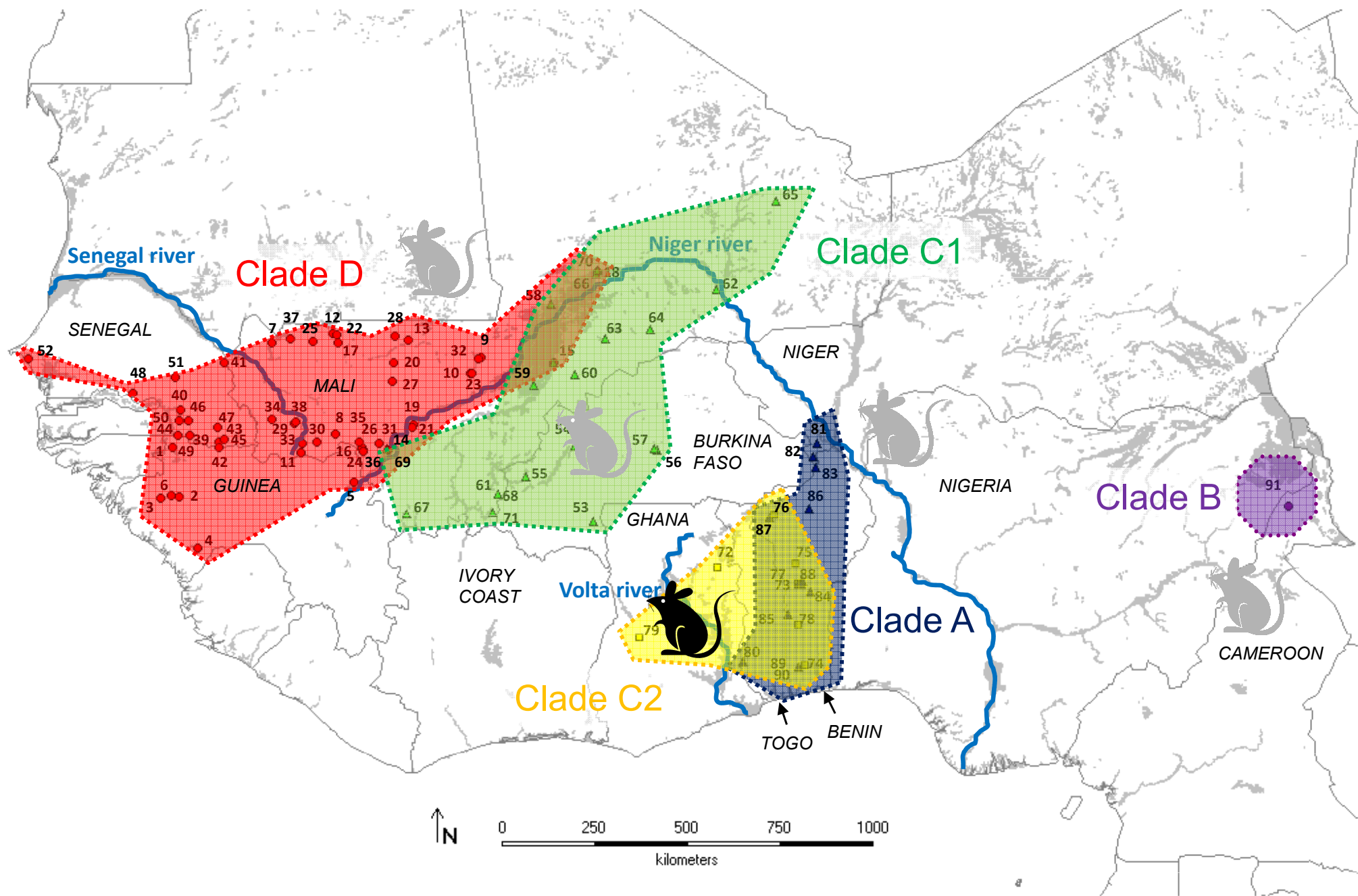


# Příklad: *Praomys cf. daltoni* complex - introgrese mtDNA

Kolik je zde druhů?

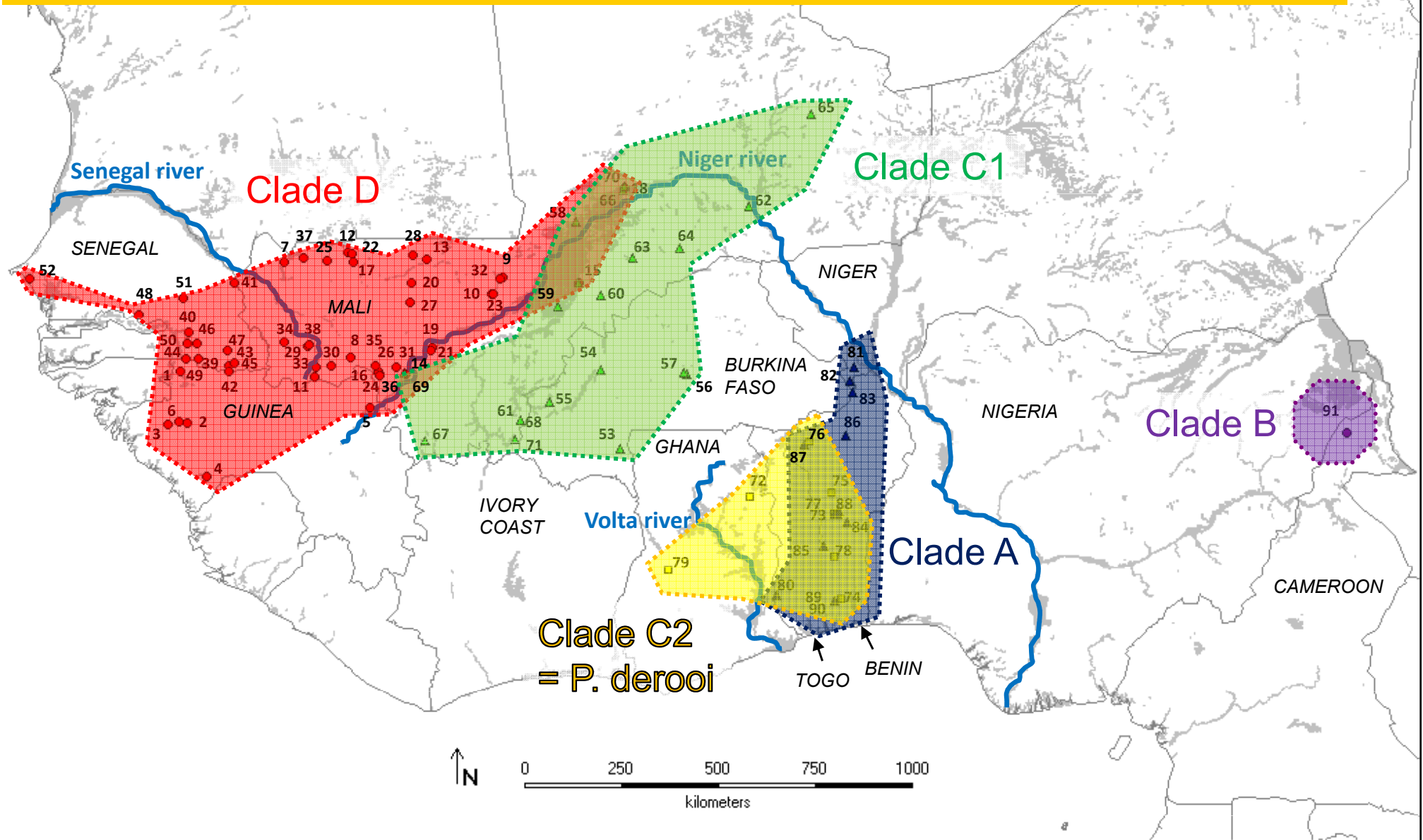
Čtyři podle genotypu – cca 7% divergence?  
Dva druhy podle fenotypu?



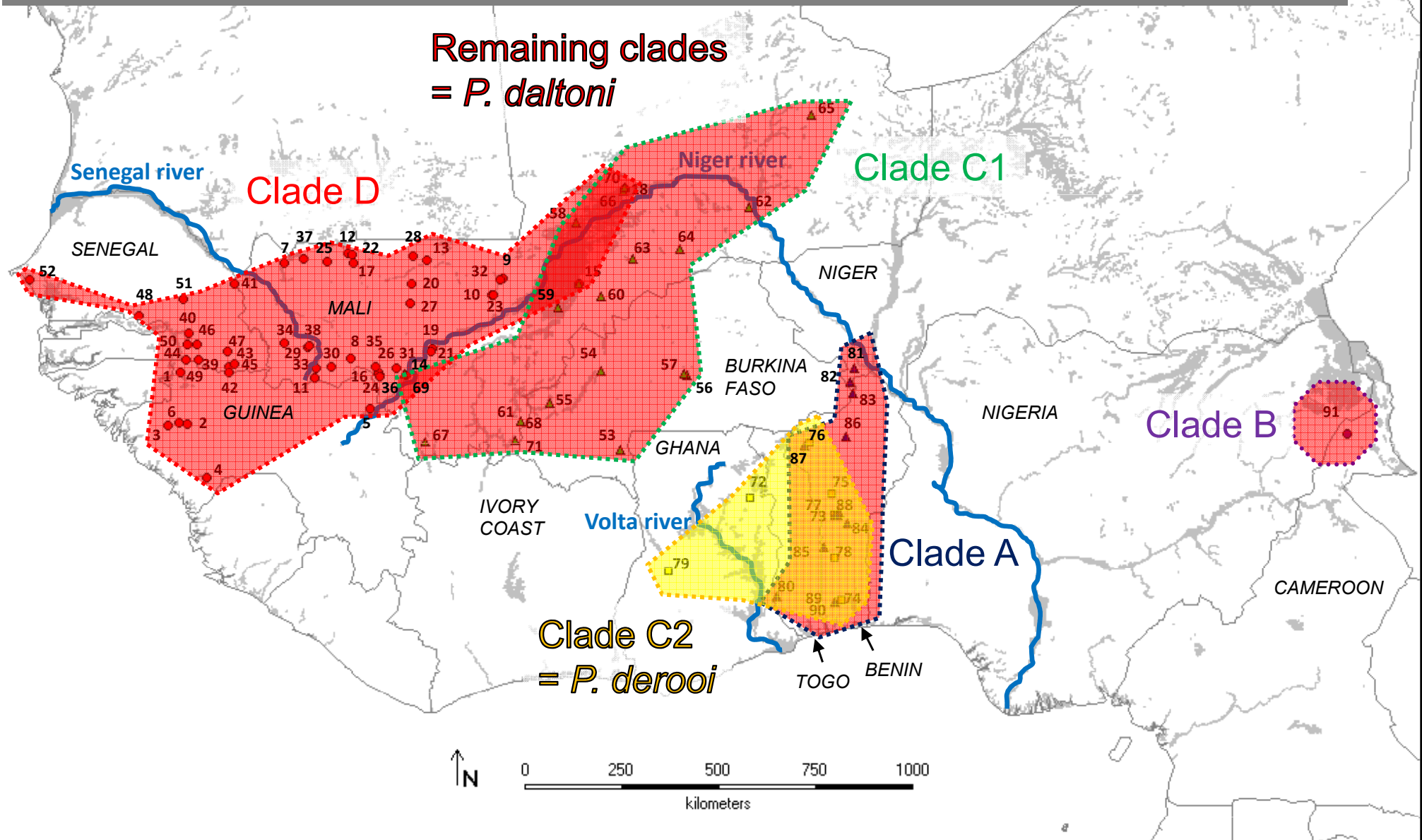




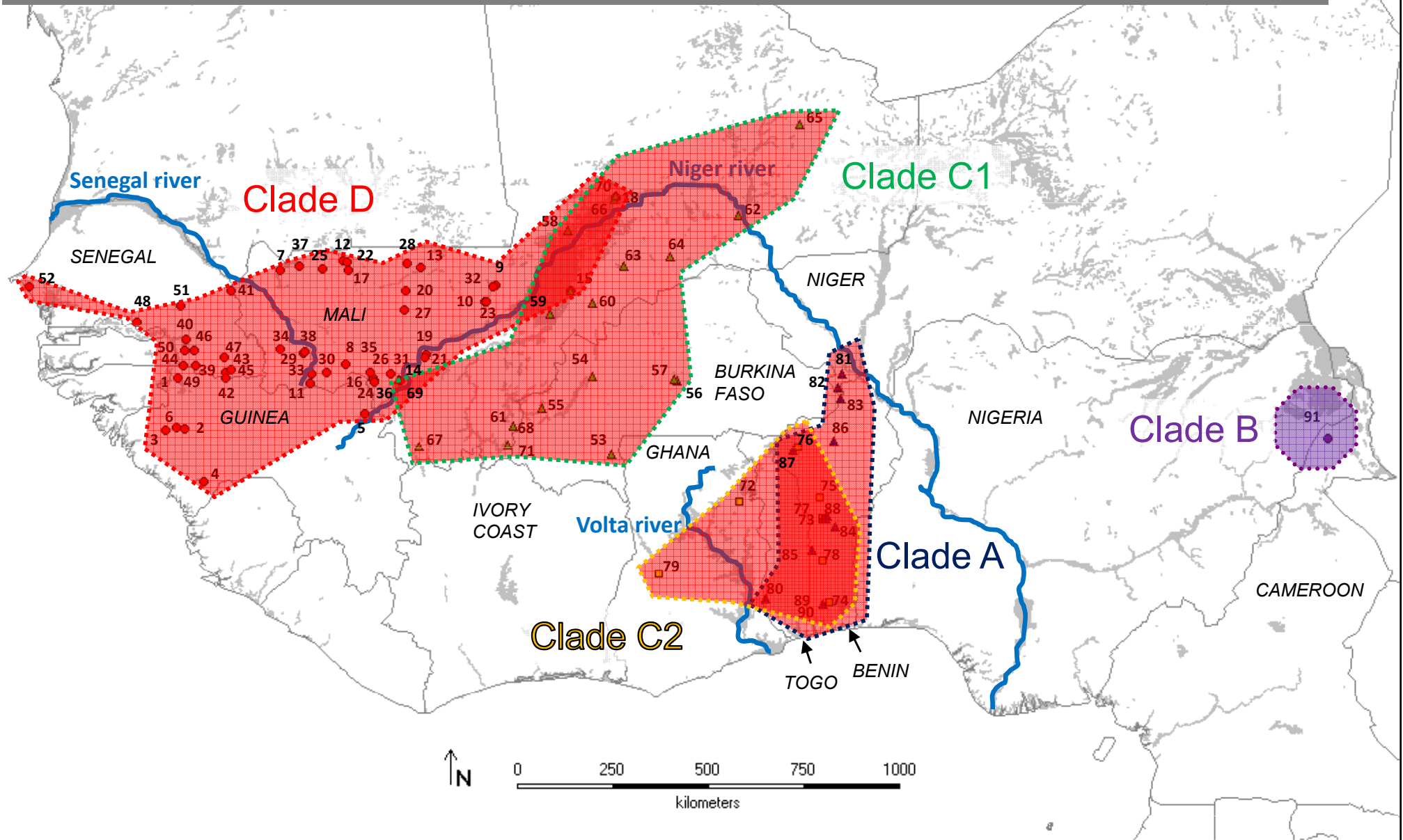
# What is species???



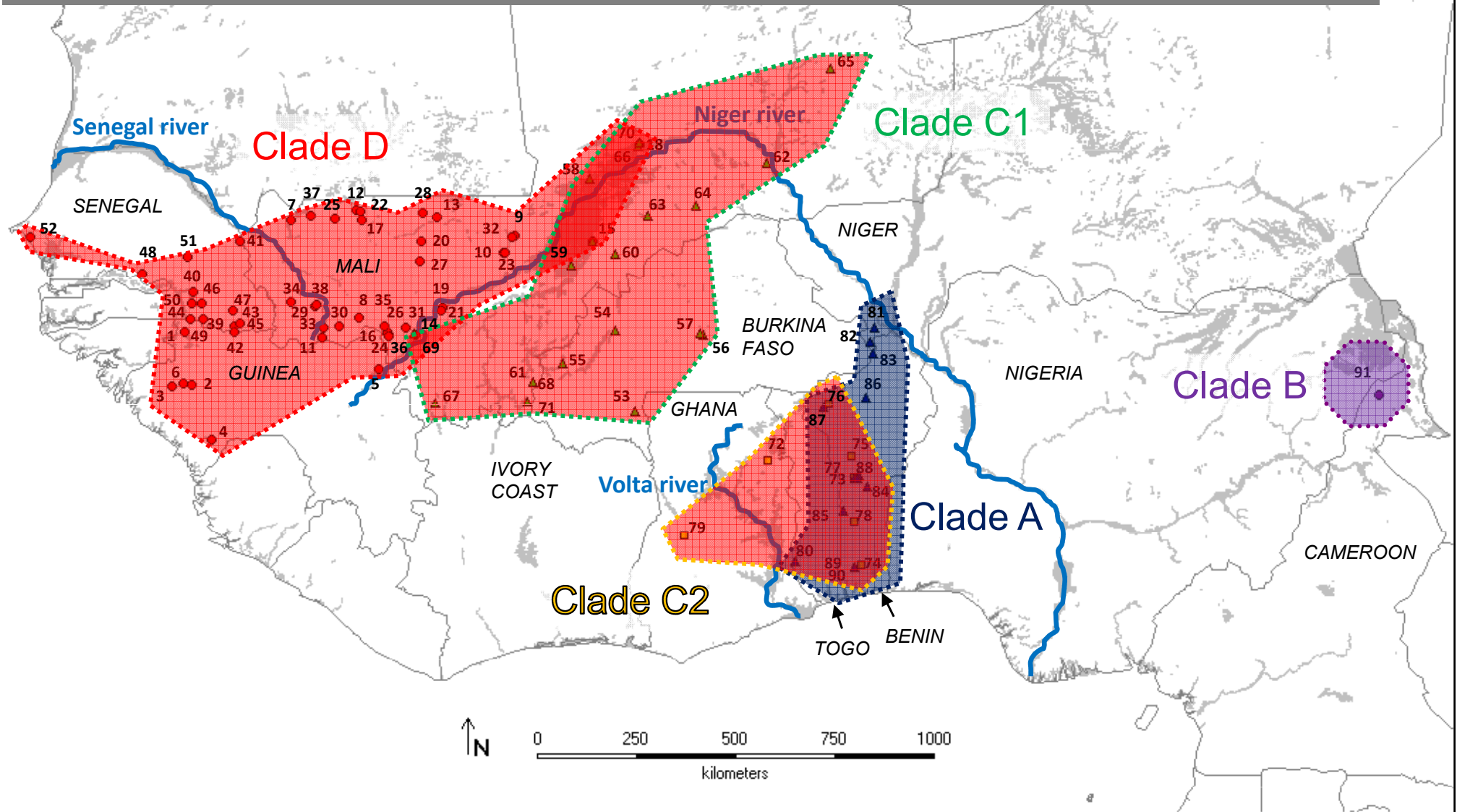
# Morphology and ecology



# Karyotypes

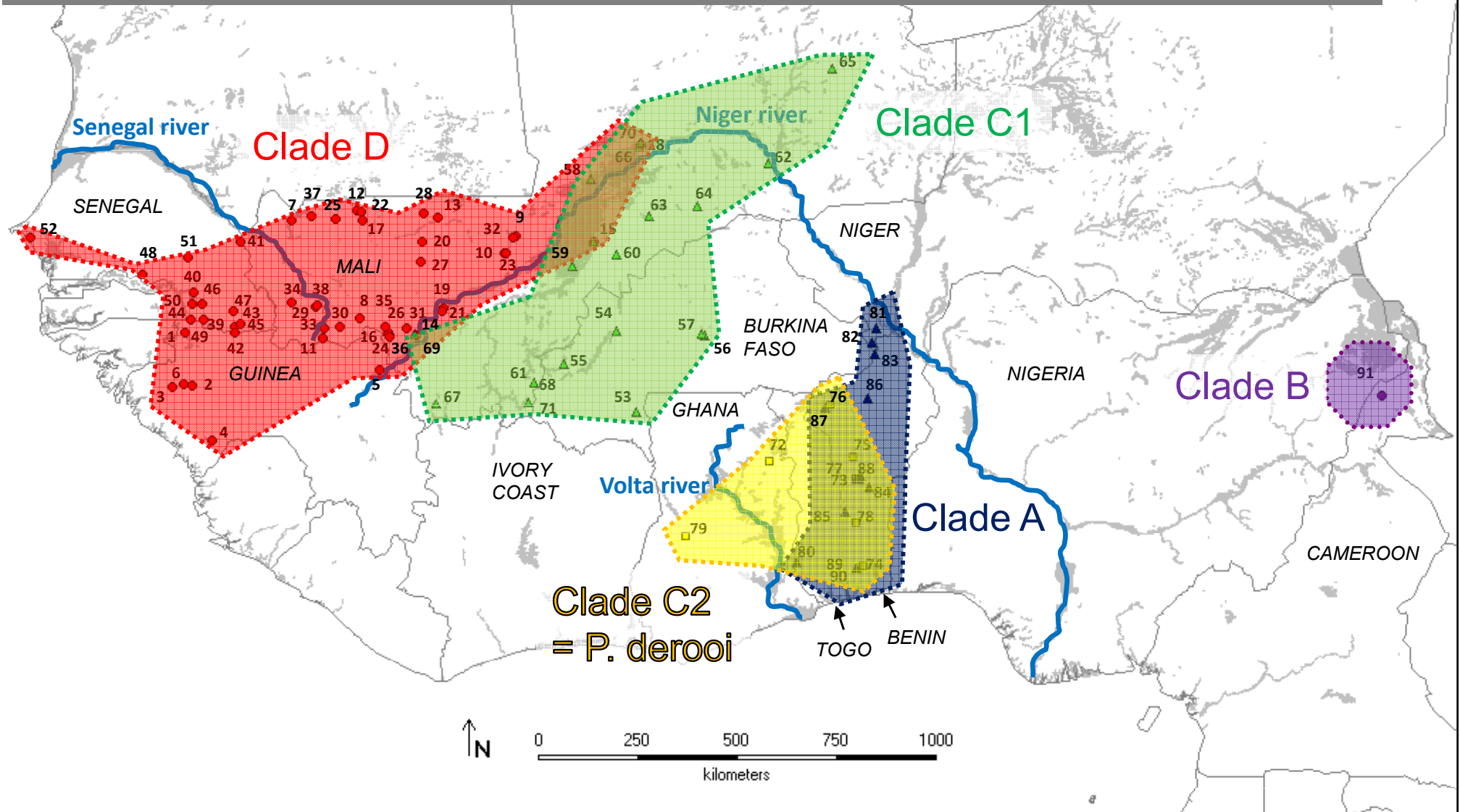


# Mitochondrial DNA + microsatellites in Benin + karyotypes





Splitting approach taking morphology and ecology into account – the reproductive barrier between Clade C1 and Clade D remains to be identified



# Who is advancing barcoding?



CONSORTIUM FOR THE BARCODE OF LIFE



**CONSORTIUM FOR THE BARCODE OF LIFE (CBOL)** is an international initiative devoted to developing DNA barcoding as a global standard in taxonomy.

**CBOL** is a collaboration of natural history museums, herbaria, biological repositories, and biodiversity inventory sites, together with academic and commercial experts in genomics, taxonomy, electronics, and computer science. CBOL has more than 100 institutional members in 40 countries.

The inaugural meeting for **CBOL** was held at The National Museum of Natural History, Washington DC in May 2004. The initial organizational support for **CBOL** was provided by a 2.5 year grant from the The Alfred P. Sloan Foundation, renewed in 2006 for 3 years.

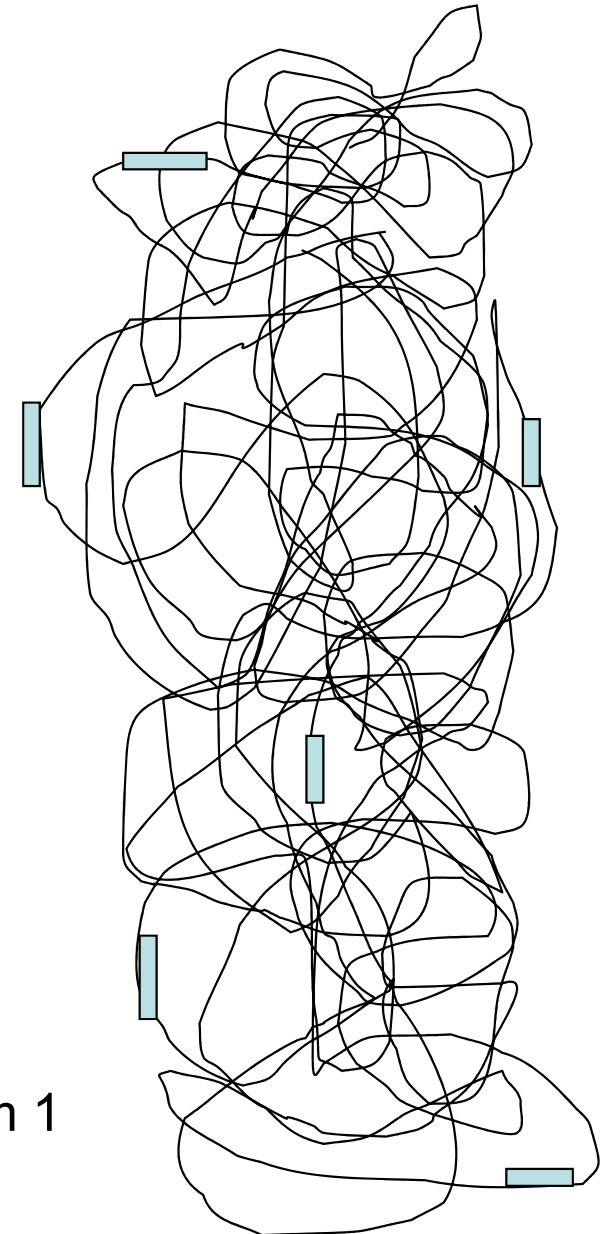
# Identifikace jedince - metody

- **DNA fingerprinting (název dnes používán pro různé metody)**
  - velké množství kvalitní DNA
  - technická náročnost
  - +univerzalita
- **AFLP**
  - kvalitní nedegradovaná DNA
  - +univerzalita
- **Sekvenování, alozomy**
  - nemusí rozlišit jedince
- **Mikrosatelity**
  - + stačí malé množství nekvalitní DNA, **optimální pro neinvazivní přístupy**
  - je nutné znát konkrétní lokusy a sekvence specifických primerů



# Multi-locus genetic markers

- Mnoho znaků náhodně rozmístěných v genomu
  - minisatellite DNA fingerprinting (celogenomová RFLP)
  - AFLP (amplified fragment length polymorphism)
- presence vs. absence = dominantní znaky (neodliší heterozygota)



Př.: chromozóm 1

## RFLP (restriction fragments length polymorphism) a AFLP (amplified fragments length polymorphism)

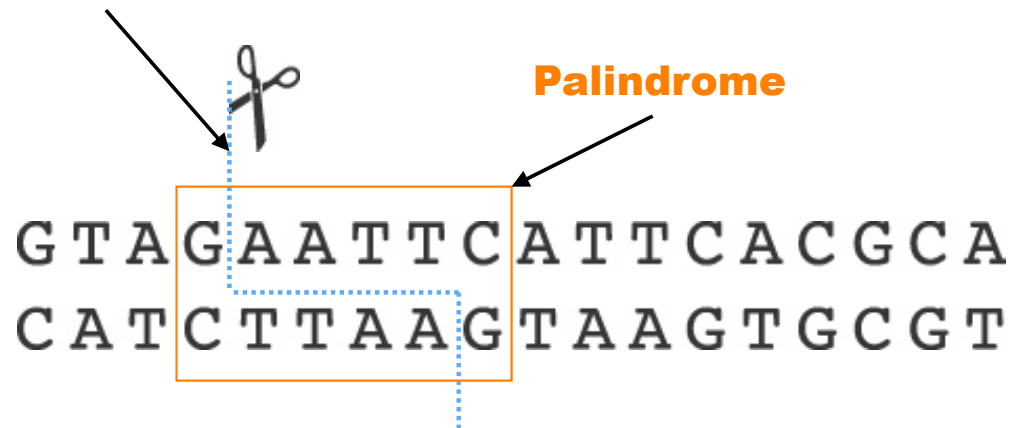
- Základem je naštípání DNA restrikčními enzymy
- → milióny fragmentů různé délky
- Rozdíly díky:
  - přítomnosti/absenci restrikčních míst
  - tandemovým opakováním (mikrosatelity, minisatelity)
  - delece a inzerce
- Zviditelnění jen určité části fragmentů

# Restrikční endonukleázy

## Enzyme Site Recognition

- Each enzyme digests (cuts) DNA at a specific sequence = restriction site
- Enzymes recognize 4- or 6- base pair, palindromic sequences (eg GAATTC)

Restriction site



Palindromic



Fragment 1

Fragment 2

# Common Restriction Enzymes



**EcoRI**

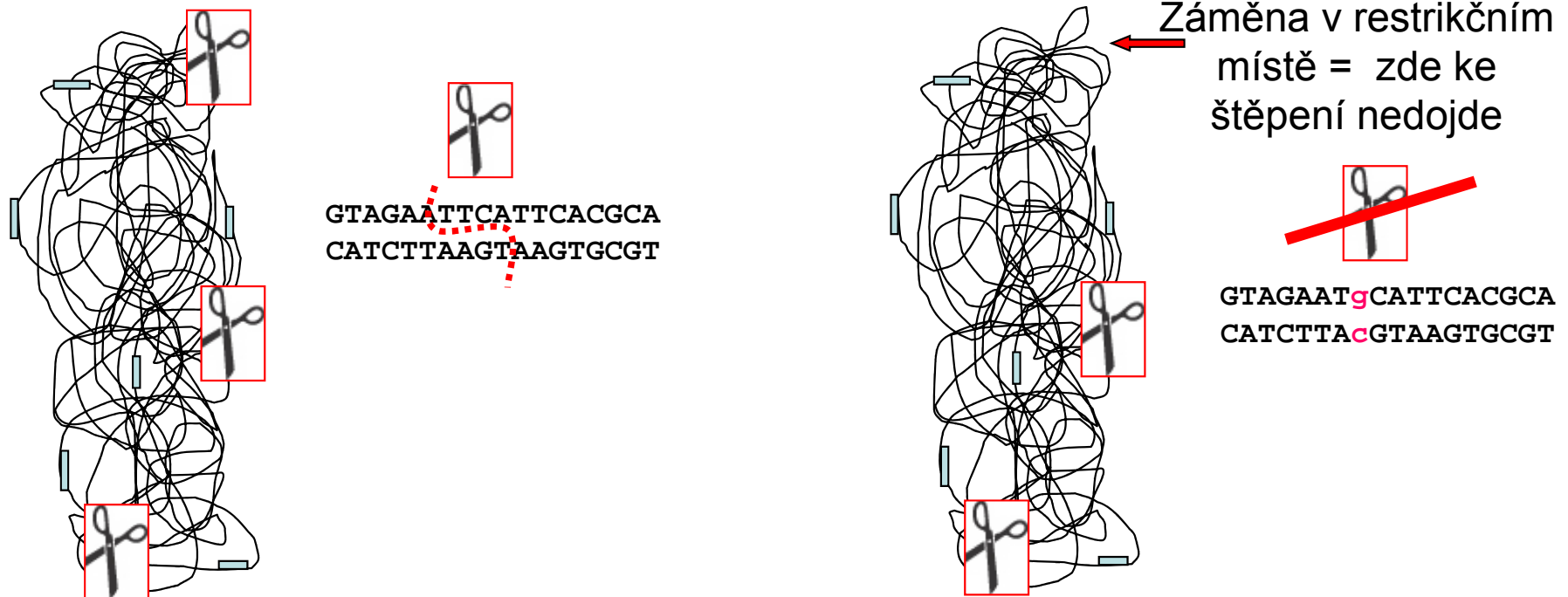
- Escherichia coli
- 5 prime overhang



**PstI**

- Providencia stuartii
- 3 prime overhang

# Každý jedinec má jedinečný genom



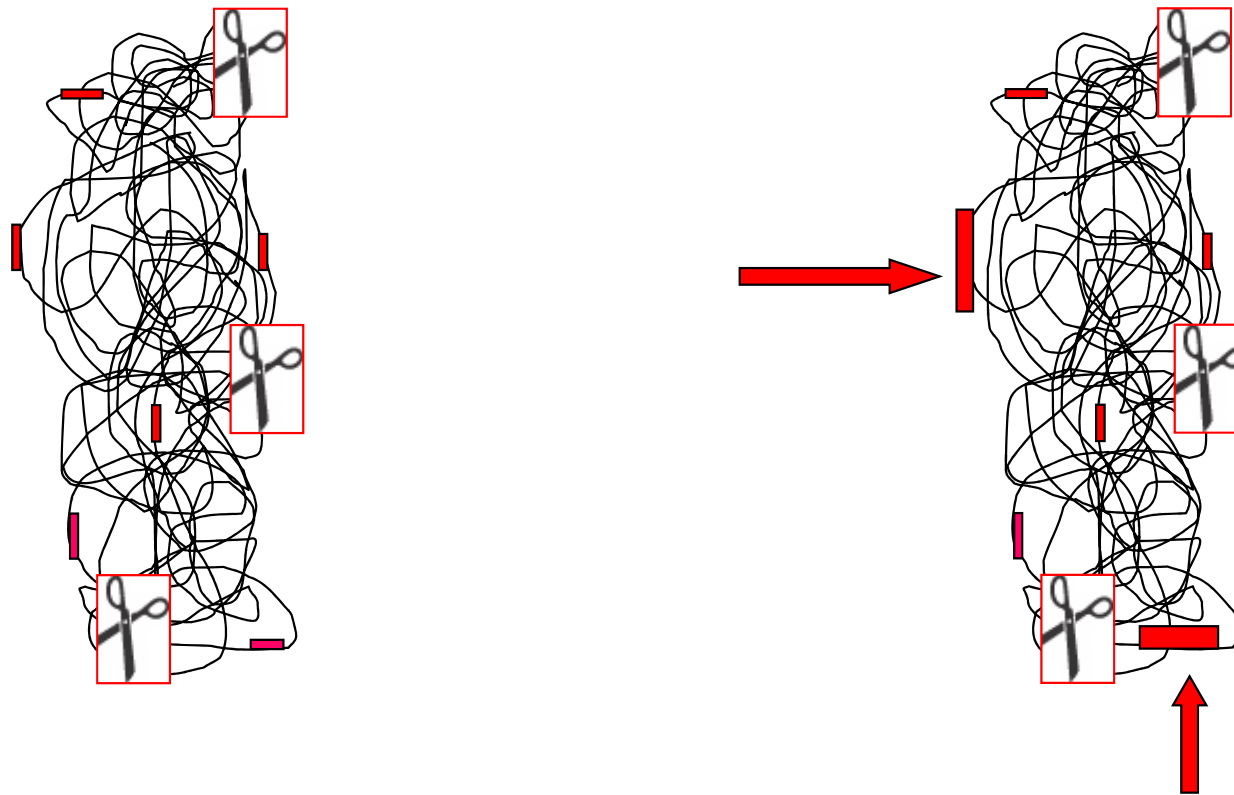
1. Ztráta nebo nabytí restričního místa

# Každý jedinec má jedinečný genom



2. Ztráta nebo nabytí SINE (např. [Alu](#) sekvence) nebo LINE

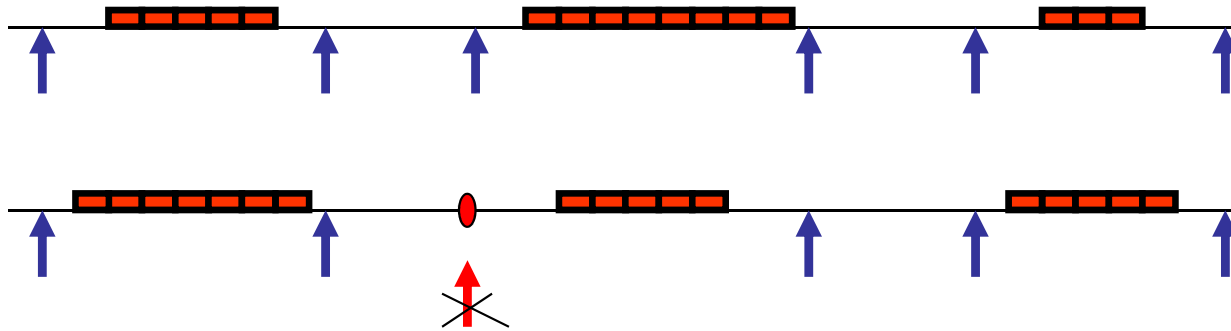
# Každý jedinec má jedinečný genom



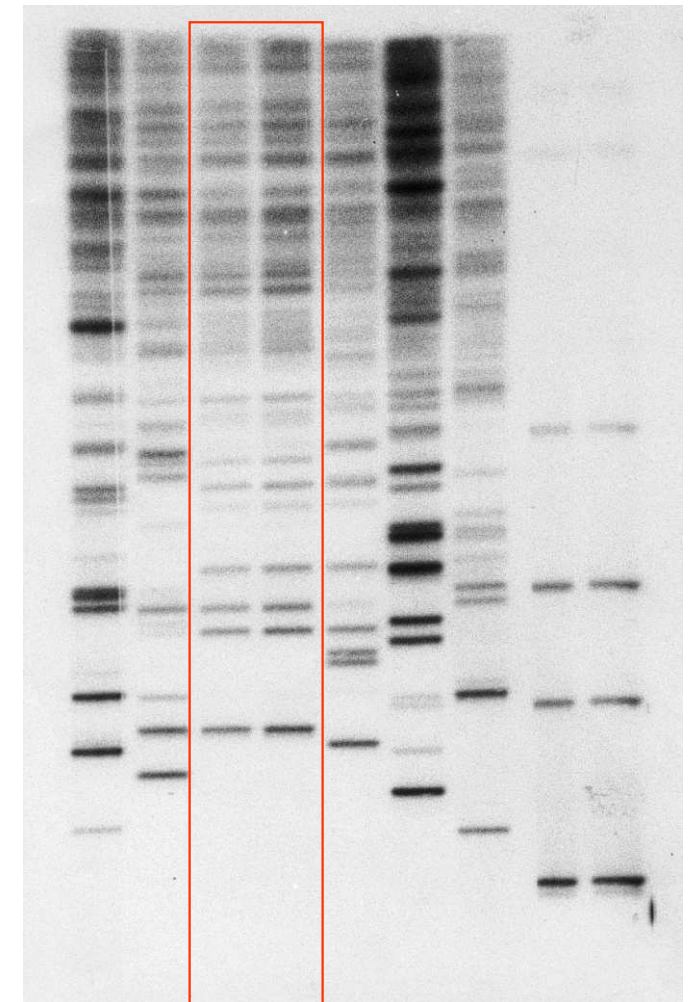
3. Vysoká mutační rychlost **minisatelitů** – rozdíly v počtu repeticí, tj. v délce daného úseku

# DNA fingerprinting

Sir  
Alec Jeffreys



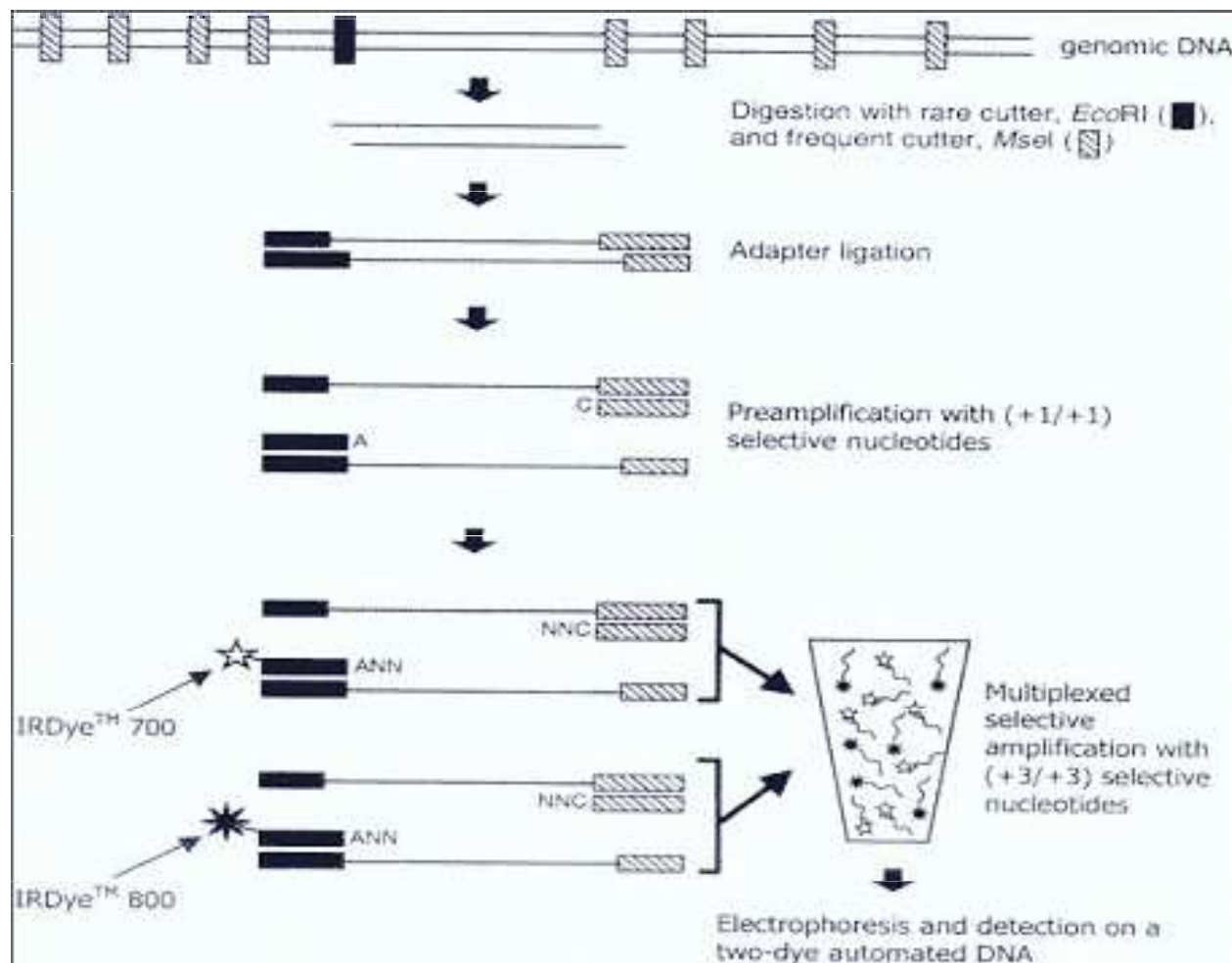
- Jeffreys et al. 1985
  - „naštípání“ DNA pomocí restričního enzymu
  - Rozdělení fragmentů pomocí elektroforézy
  - Hybridizace (jako sonda minisatelit, např. Alu)
- Dostatek kvalitní DNA
- Vše na jednom gelu
- Laboratorní náročnost





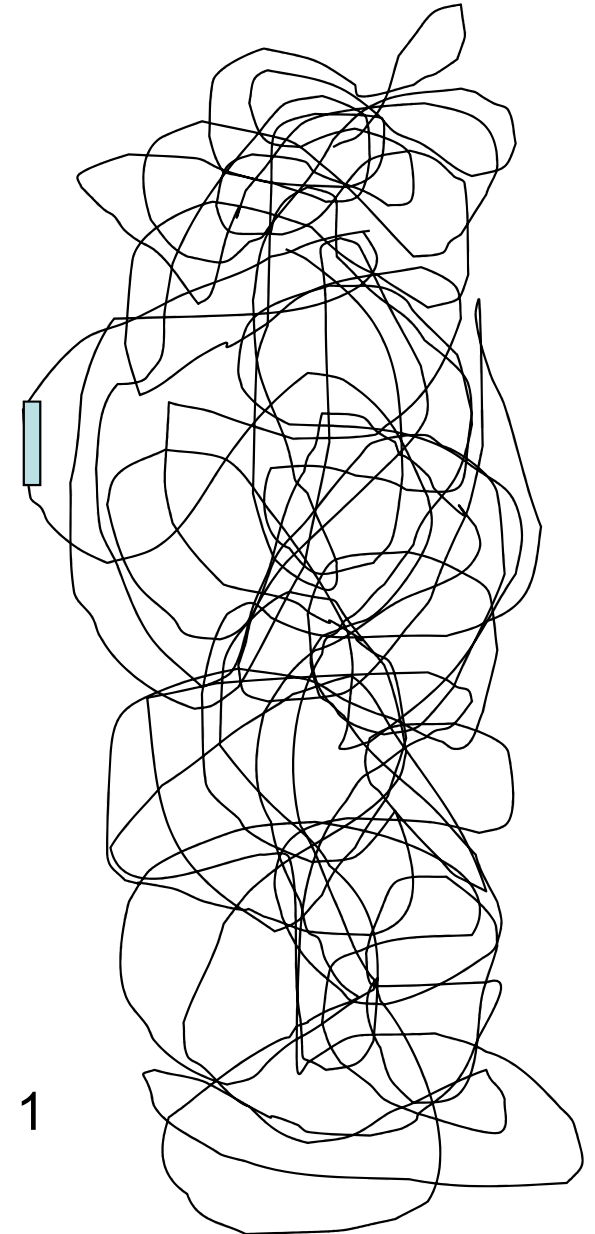
# AFLP

- Naštípání DNA restrikcími enzymy
- Navázání adaptérů
- Amplifikace primery o nukleotid delšími než je adaptér (jinak sekvence komplementární k adaptéru)
- Amplifikace primery o nukleotid delšími než je adaptér (jinak sekvence komplementární k adaptéru)
- Amplifikace značenými primery o 3 nukleotidy delšími než je adaptér
- Detekce – gely, sekvenátor



# Single-locus genetic markers

- kodominantní – možno stanovovat frekvence alel
- allozomy a jiné funkční geny - **MM**
- mikrosatelity – délkový polymorfismus
- SNPs (single nucleotide polymorphisms) – sekvenční polymorfismus
- SINE, LINE



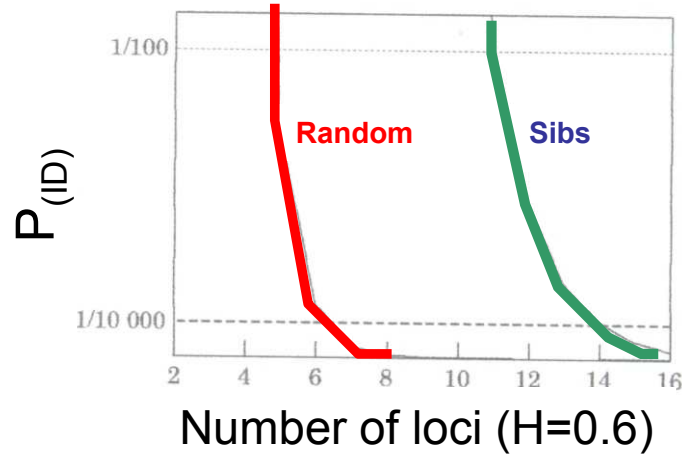
Př.: chromozóm 1



# Identifikace jedinců závisí na stupni polymorfismu

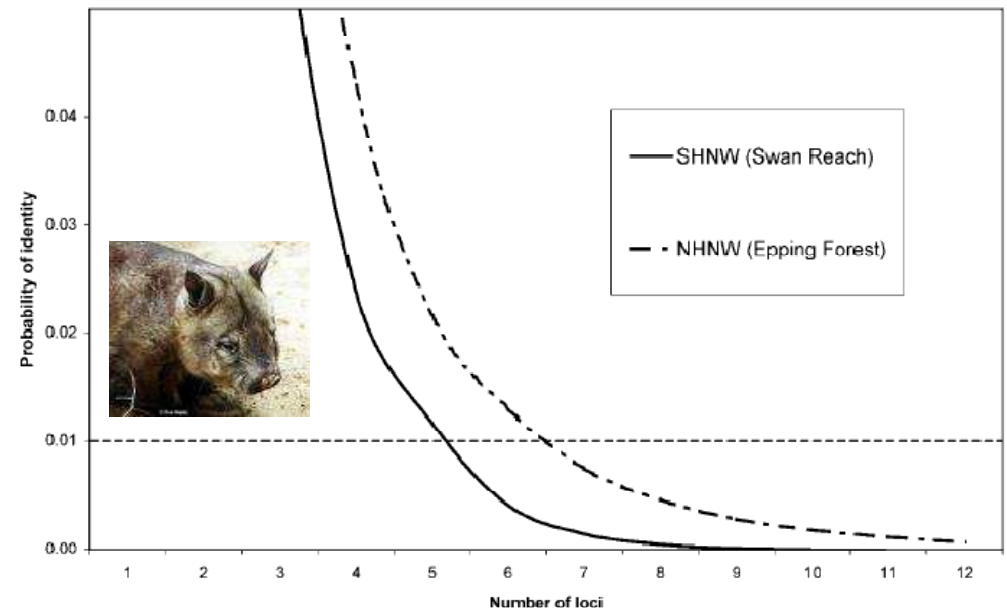
- multilocus microsatellite fingerprinting – power estimated as „probability of identity“ ( $P_{(ID)}$ ) (Waits et al. 2001)

$$P_{(ID)} = \sum p_i^4 + \sum \sum (2 p_i p_j)^2$$



$$P_{(ID)sib} = 0.25 + (0.5 \sum p_i^2) + [0.5(\sum p_i^2)^2] - (0.25 \sum p_i^4)$$

- pilot studies with tissue samples are required to identify  $P_{(ID)}$  in a population studied by e.g. Non-invasive methods





# Medvědi v Pyrenejích

*Taberlet et al. 1997*



- Trus a chlupy
- 24 mikrosatelitových lokusů
- 4 samci a jedna samice  
(o jednoho víc než podle stop a fotografií)
- Multiple-tube approach,  
mnohonásobné opakování PCR  
reakcí

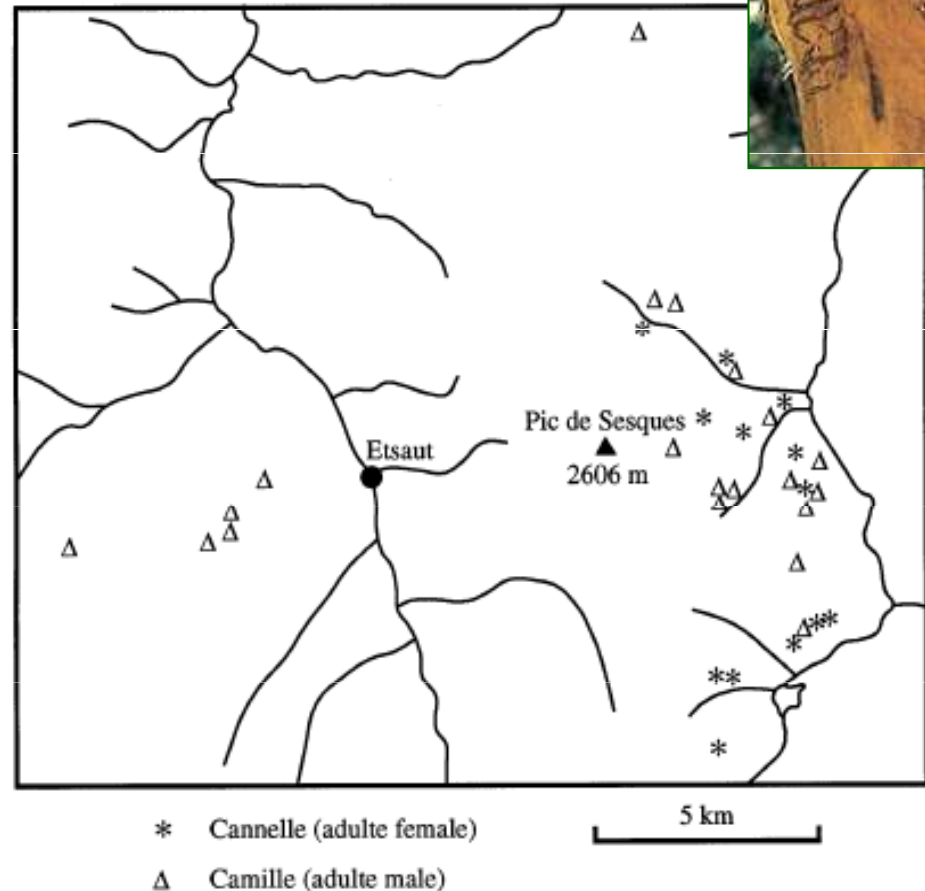


Fig. 3 Home range of two Pyrenean brown bears obtained by noninvasive genetic sampling and genotyping.

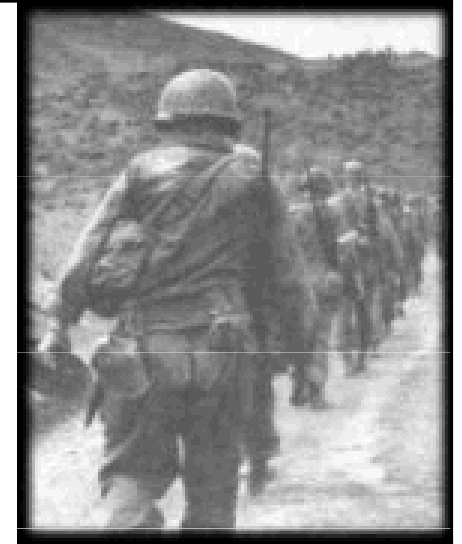
# Lidská forenzní genetik

- **Pozůstatky vojáků z války** Vietnam a Korea

Identifikace na základě mtDNA příbuzných osob  
(Ize jen někdy)

V současnosti: vzorek DNA (krve) při odvodu, jiné markery

Armed Forces Repository of Specimen Samples for the Identification of Remains



- **Soudní pře**  
Clinton-Lewinská  
Pozůstatky ruského cara Nikolaje II

- **Kriminalistika**

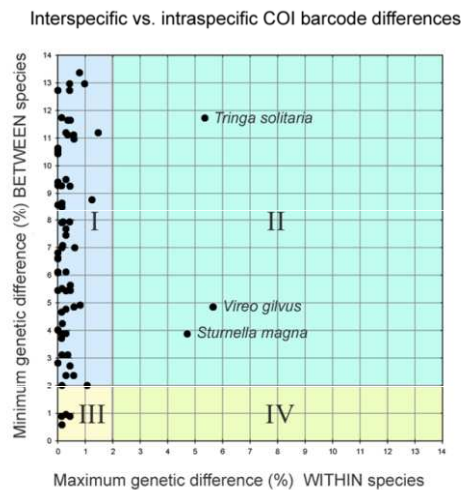
- **Oběti tragických událostí**





# Genetická identifikace

Barcoding  
North  
American birds  
highlights  
probable  
cryptic species

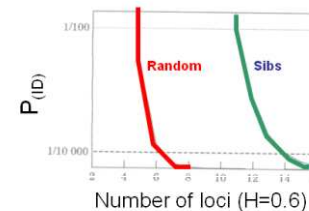


Results for 73 species of North American are shown. Quadrants represent different categories of species:  
I. consistent with current taxonomy  
II. probable lumped species (candidate for taxonomic split)  
III. recent divergence, hybridization, or synonymy.  
IV. probable taxonomic misidentification

## Identifikace jedinců závisí na stupni polymorfismu

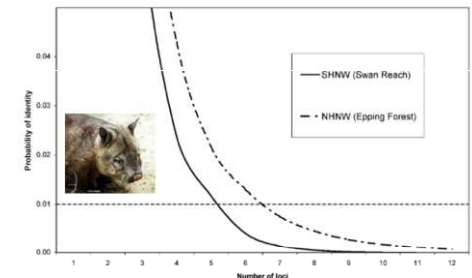
- multilocus microsatellite fingerprinting – power estimated as „probability of identity“ ( $P_{(ID)}$ ) (Waits et al. 2001)

$$P_{(ID)} = \sum p_i^4 + \sum \sum (2p_i p_j)^2$$



$$P_{(ID)_{sib}} = 0.25 + (0.5 \sum p_i^2) + [0.5(\sum p_i^2)^2] - (0.25 \sum p_i^4)$$

- pilot studies with tissue samples are required to identify  $P_{(ID)}$  in a population studied by e.g. Non-invasive methods



Identifikace druhu  
(„DNA barcoding“)

Identifikace jedince  
(„DNA fingerprinting“)

# Klony

## Bambus *Sasa senanensis*

- Suyama et al. 2000
- Plocha 10 hektarů
- AFLP
- 22 klonů
- Klon na ploše 300 m v průměru





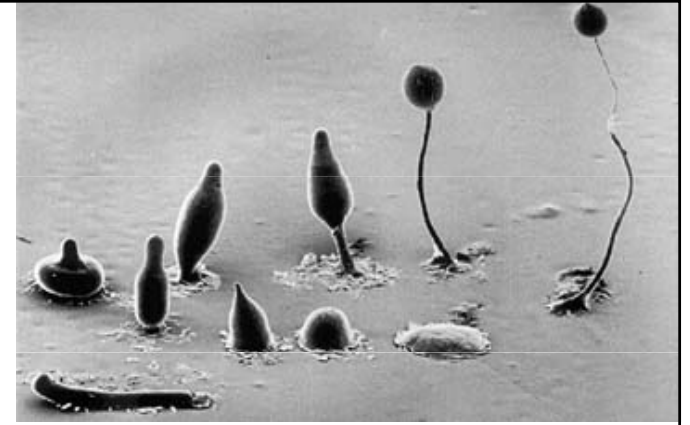
# Slavní klonální bezobratlí

- Rotifera – Bdelloidea
- Ostracoda  
(*Darwinula*)
- Partenogenetické  
klony vysokého stáří  
(milióny let)

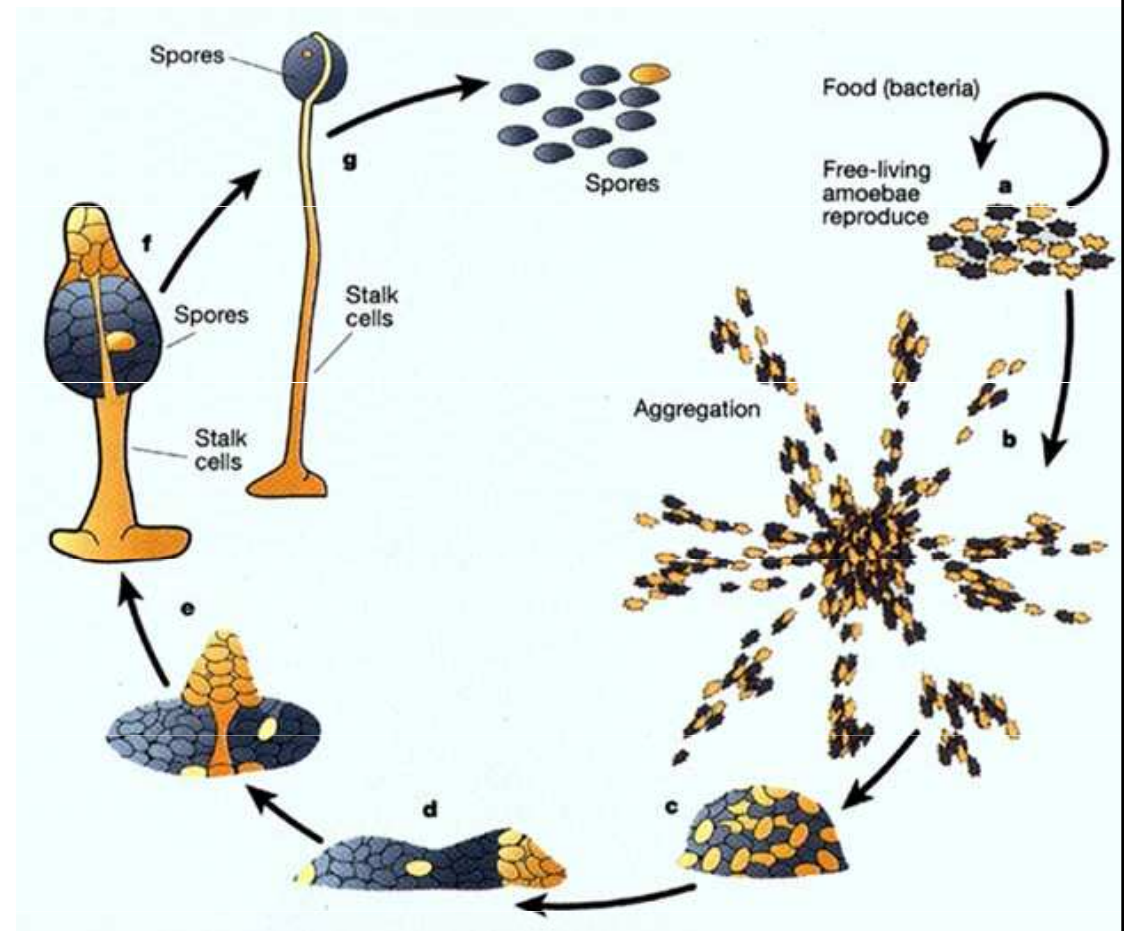


*Darwinula stevensoni*

# Genetické chiméry

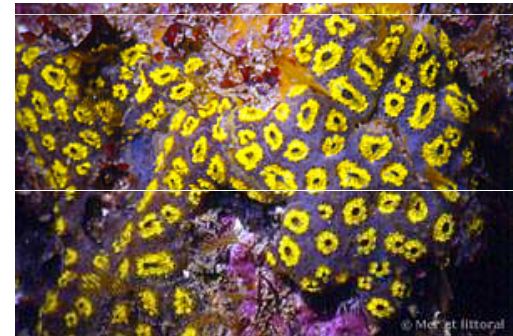


- organismy složené z buněk s různými genotypy
- *Dictyostelium discoideum*  
chimérismus je pravidelná součást života



# Genetické chiméry

- *Ficus* srůst kořenů různých jedinců
- sumky *Botryllus schlosseri* chimérické kolonie příbuzní jedinci
- *Diplosoma listerianum* i nepřibuzní

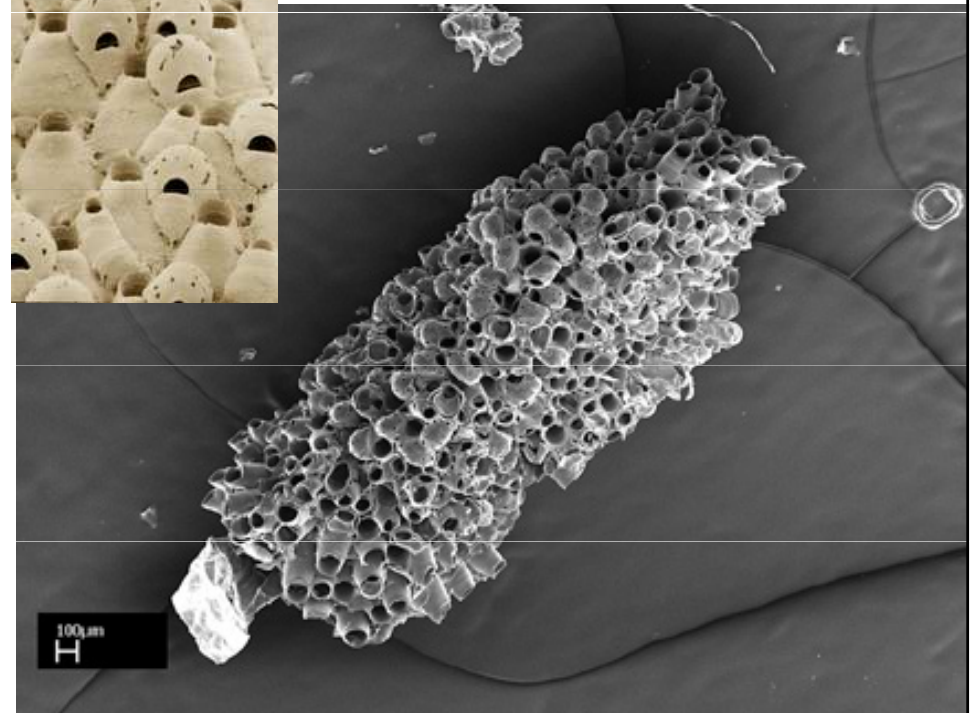
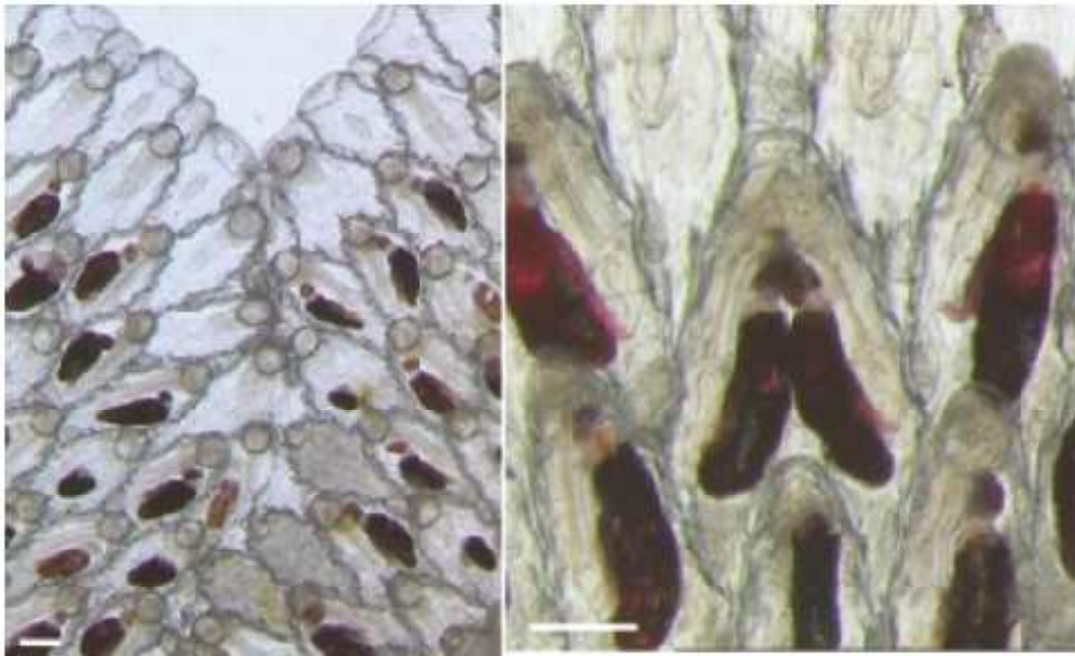




# *Celleporella hyalina* (Bryozoa)

Hughes et al. 2004

- Pravděpodobnost fúze koreluje s příbuzností
- Histokompatibilita
- Lepší rozpoznávání v pokročilejších fázích  
→ dozrávání imunokompetence
- Speciální proteiny (spongikany...)



# Genetické chiméry



- kosman bělovousý *Callithrix jacchus* (asi i *Saguinus*)
- Dizygotická dvojčata
- DNA fingerprinting krve
- Jsou to hematopoietické chiméry
- Během embryonálního vývoje vzájemná výměna buněk kostní dřeně
- Týká se to asi jen krve (neinvazivní metody – chlupy, trus →jeden genotyp)
- Průnik embryonálních erytroblastů a volné DNA přes placentu u člověka
- (pohlaví dítěte před narozením lze určit i pomocí PCR sekvencí typických pro Chr Y, jako templát periferní krev matky)

# Genetická identifikace pohlaví

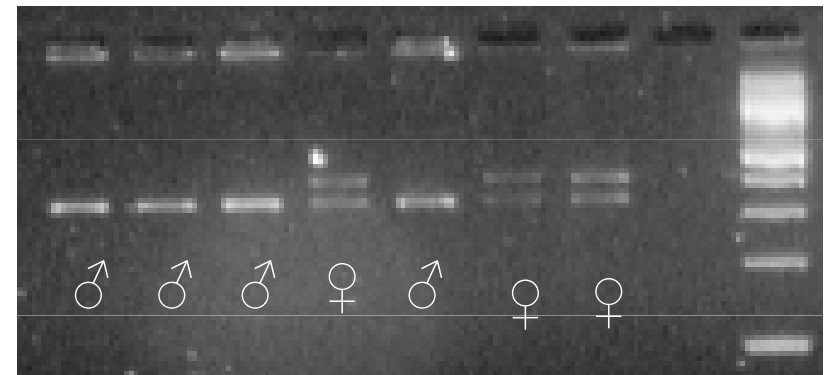
- 1) druhy s nevýrazným pohlavním dimorfismem (ptáci)
- 2) zárodky v ranném stádiu ontogeneze (embrya)
- 3) neinvazivní metody (trus, skořápky, šupiny)

# Genetická identifikace pohlaví

- druhy s genetickou determinací pohlaví (tj. nejčastěji pohlavní chromozómy)
- ptáci ( $\text{♂}=\text{ZZ}$ ,  $\text{♀}=\text{ZW}$ )
- savci ( $\text{♂}=\text{XY}$ ,  $\text{♀}=\text{XX}$ )
- amplifikace DNA oblasti specifické pro heterogametické pohlaví
- W, Y – malé chromozómy

# Určení pohlaví – ptáci

*Griffith et al. 1998*



- *CHD1W* a *CHD1Z*, geny na pohlavních chromosomech (chromobox-helicase-DNA-binding gene (CHD) – Griffiths & Tiwari 1995)
- Primery amplifikují introny obou genů
- Introny se mohou lišit délkou
- Existují už tři možnosti běžně používaných primerů
- Problematické druhy Struthioniformes





# *Manorina melanocephala*

(Meliphagidae) *Arnold et al. 2001*

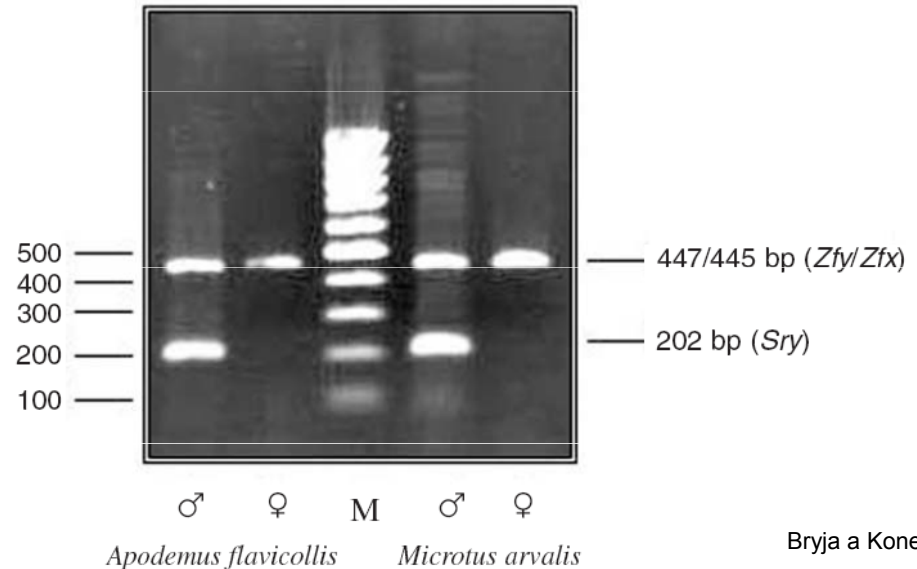
- Synové fungují jako pomocníci
- U adultů  
2,31 samců na 1 samici
- Mláďata v hnízdě  
poměr pohlaví 1:1 (57:57)
- První se líhnou samci  
(v 17 hnízdech z 18)  
Při opouštění hnízda jsou větší a těžší



medosavka hlučná

# Určení pohlaví - savci

- Amplifikace genu na Chr Y (*Sry*)  
(nejlépe duplex PCR s genem na X nebo autosomech)



Bryja a Konečný 2003

- Sry* DNA-binding motif (HMG box)  
velmi podobný u vačnatců i placentálů

- Microtus cabrerae*  
Sry na Chr X  
*Ellobius*, *Tokudaia*  
Sry zcela chybí
- Nannomys*  
Velká variabilita

*M. cabrerae*



*Ellobius*



*Nannomys*



*Tokudaia osimensis*

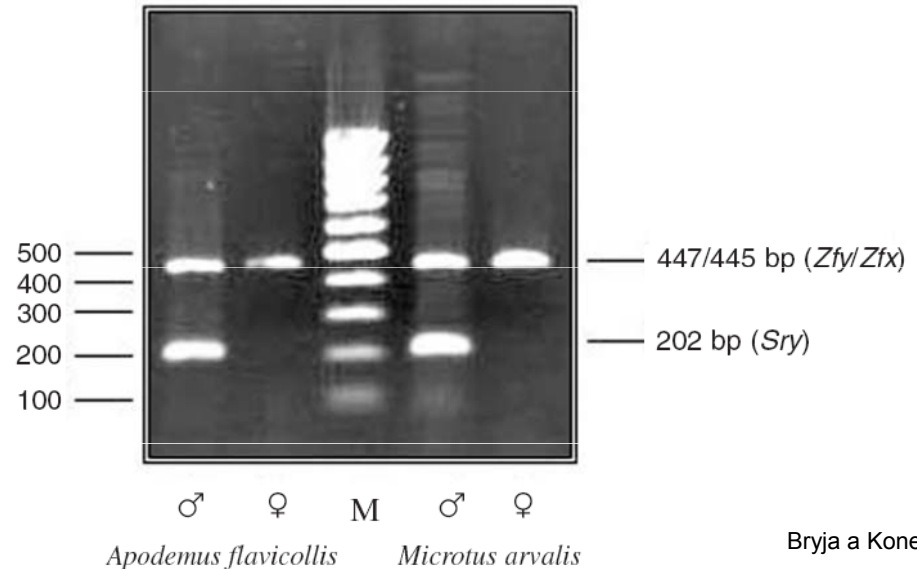


# Určení pohlaví - savci

- Amplifikace genu na Chr Y (*Sry*)  
(nejlépe duplex PCR s genem na X nebo autosomech)

- *Sry* DNA-binding motif (HMG box)  
velmi podobný u vačnatců i placentálů

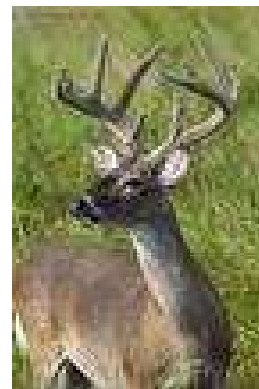
- Analýzy z trusu: nutno používat druhově specifické markery  
(jinak cross-amplification s druhy tvořícími potravu)



Bryja a Konečný 2003



X



Murphy et al. 2003

# Určení pohlaví – jiné skupiny

- Chr Y občas i u rostlin  
*Rumex*
- Plazi  
*Calotes versicolor*  
Sry má i 50% samic!
- Hledání markerů pomocí nescifických metod (RAPD, AFLP)

