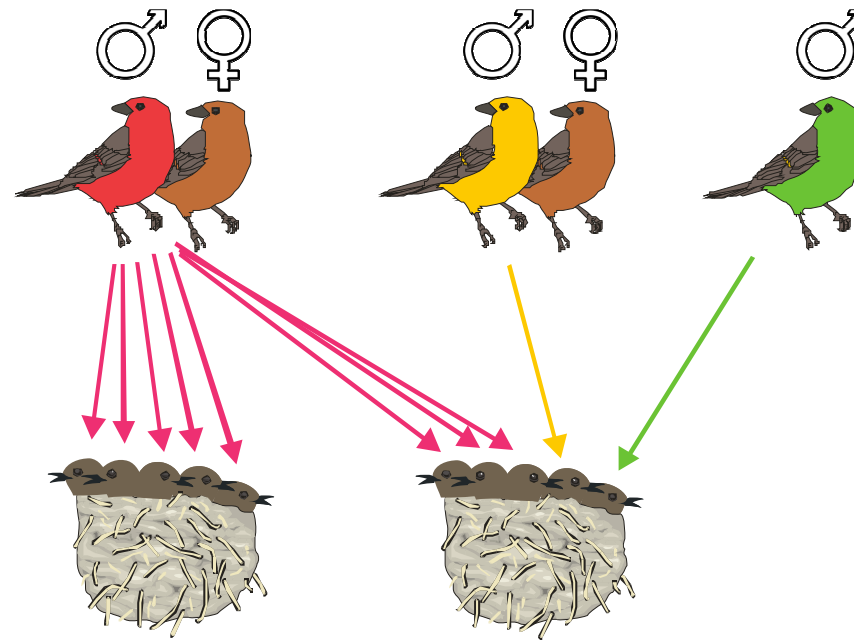


Molekulární identifikace

Druh, jedinec, pohlaví



Identifikace druhů

DNA barcoding

Definice druhu

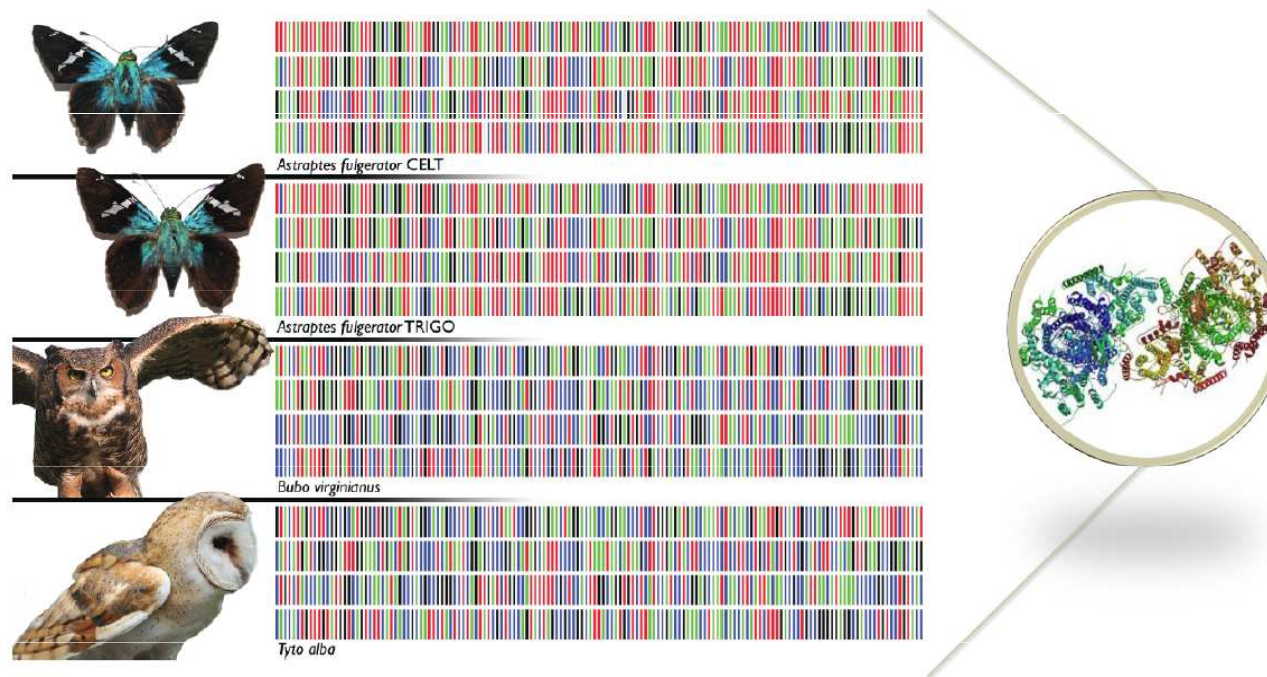
- Koncept **biologického druhu** (Mayr, 1942) – RIM = post- nebo prezygotické bariéry toku genů (ne vždy tak jednoduché – hybridní zóny, alopatická speciace, asexuální druhy atd.)
- Další koncepty (např. topotypický = morfologický, genetický, fylogenetický aj.)

Proč je vůbec potřeba druhy určovat (geneticky)?

- 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.

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

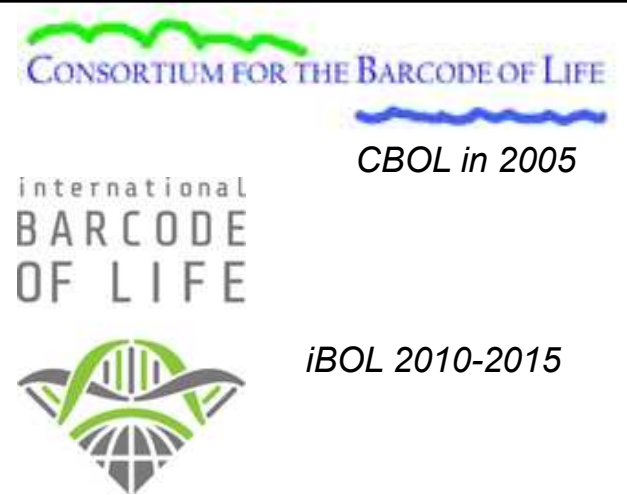
DNA Barcode: A short DNA sequence, from a uniform locality on the genome, used for identifying species



Biological identifications through DNA barcodes

Paul D. N. Hebert*, Alina Cywinska, Shelley L. Ball
and Jeremy R. deWaard

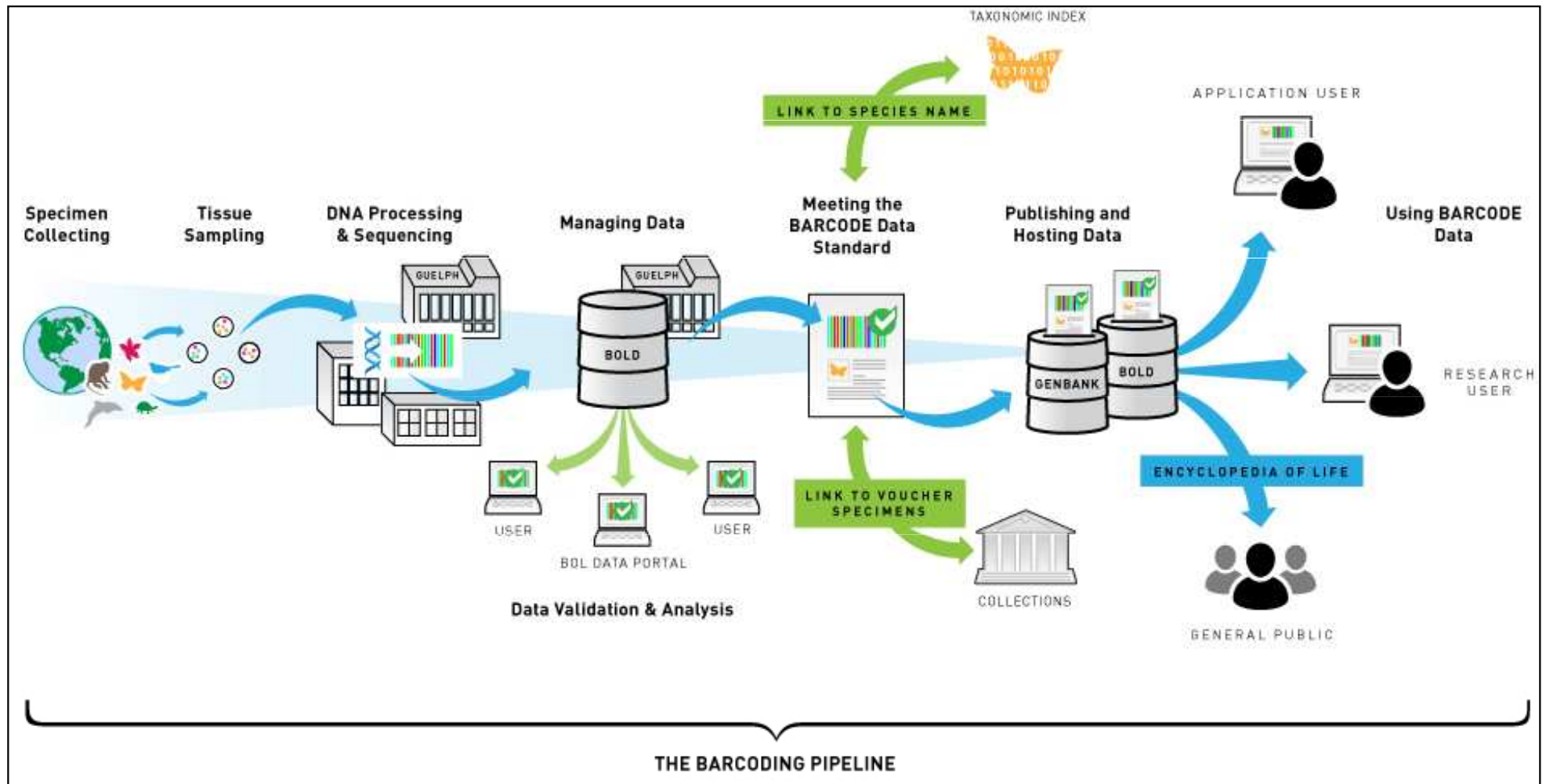
Department of Zoology, University of Guelph, Guelph, Ontario N1G 2W1, Canada



CBOL in 2005

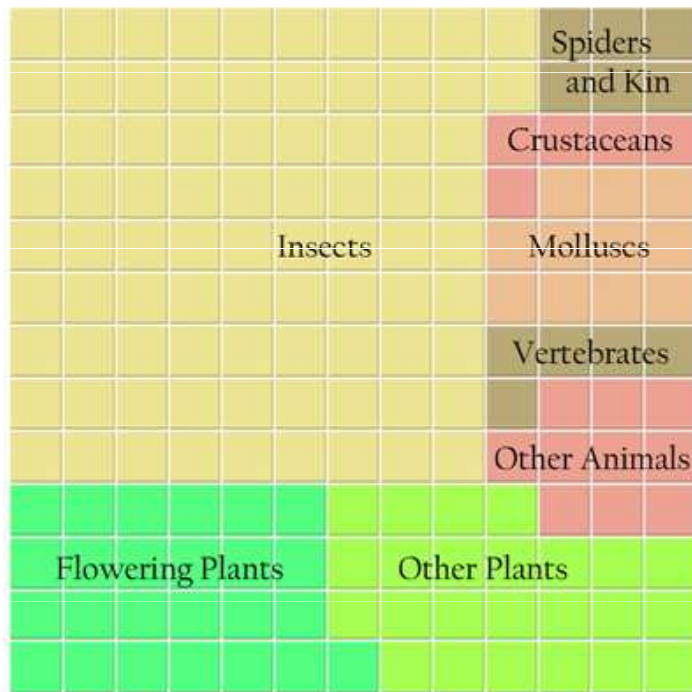
iBOL 2010-2015

first idea in 2003



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



Crisis of biodiversity and classical taxonomy



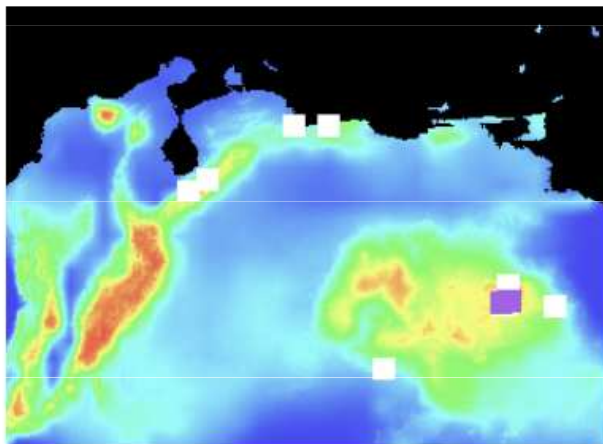
DNA barcoding is
important part of
„integrative taxonomy“

Integrative taxonomy

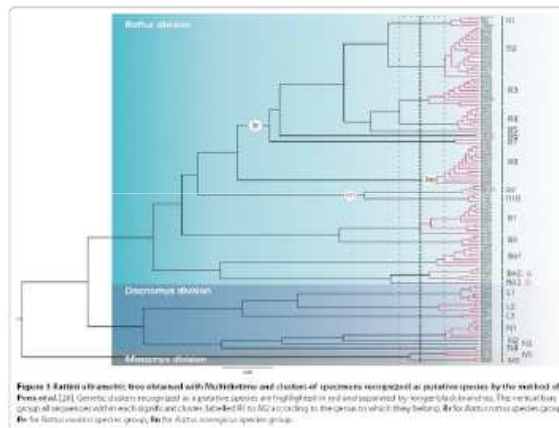
Integr. Taxonomy



Ecology



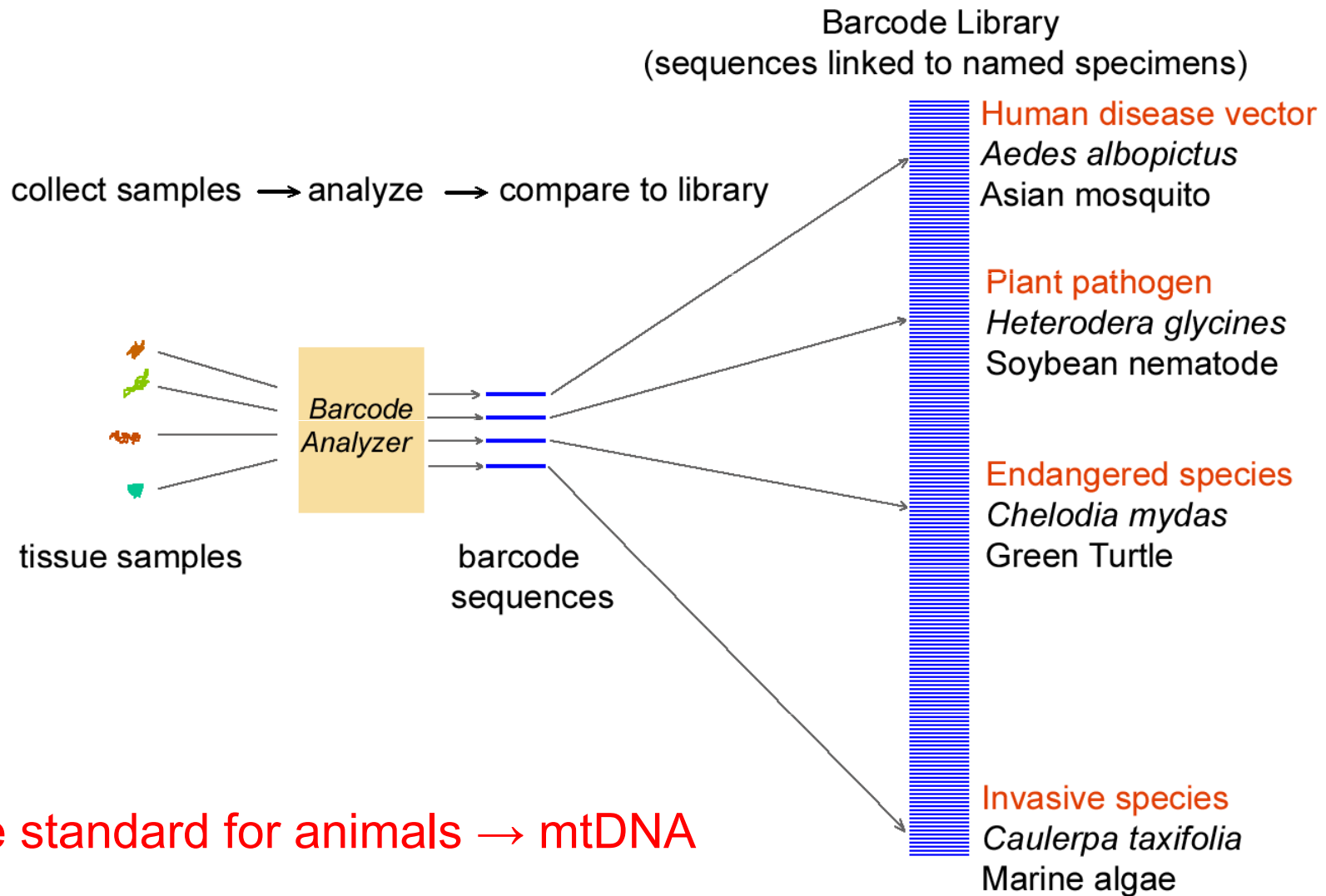
Genetics



Behavioural Biol.



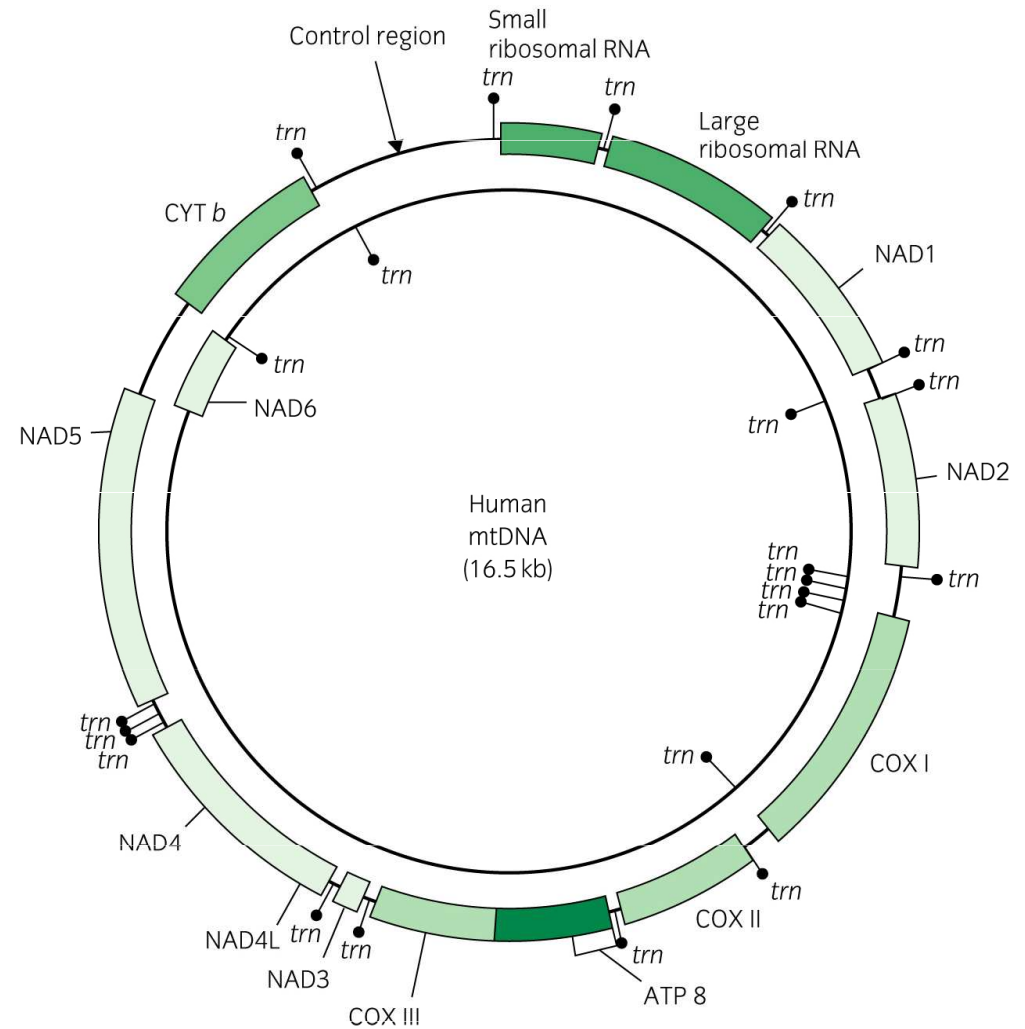
What are the benefits of standardization?



Suitable standard for animals → mtDNA

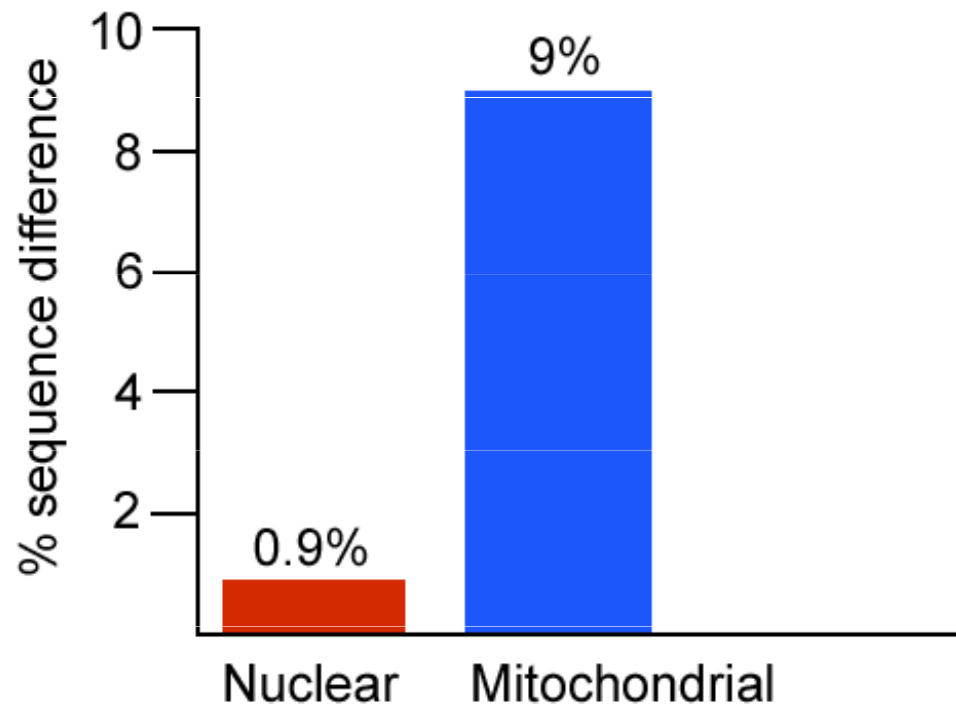
Why barcode animals with mitochondrial DNA?

Four properties make mitochondrial genomes especially suitable for identifying species



1. **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



- **2. 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.
- **3. 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.
- **4. 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.

Focus to date

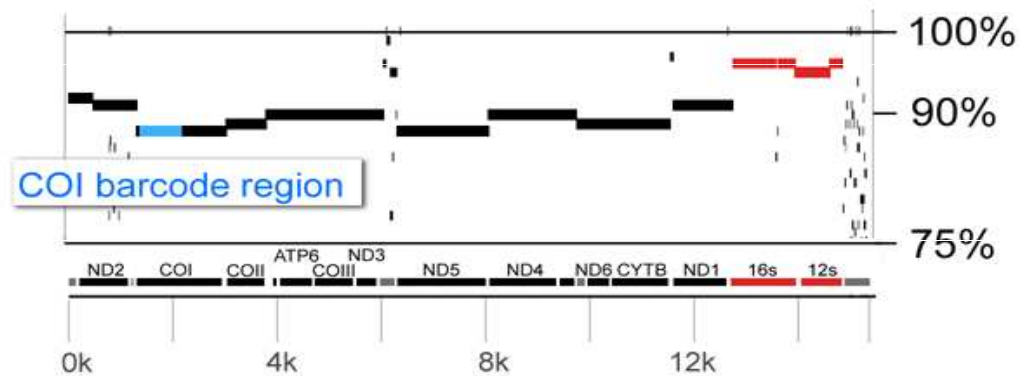
- **For animals**, a 658 base-pair fragment of the mitochondrial gene, **cytochrome oxidase subunit I** (mtCOI) – consensus for iBOL consortium
- **For plants**, mitochondrial genes do not differ sufficiently to distinguish among closely related species. Promising markers are genes on cpDNA: *matK* and *rbcL*
- **For bacteria**, a 16S-rDNA emerges as very useful marker (especially when using next-generation sequencing)
- For particular taxonomic groups, also other barcodes are widely used, e.g. cytochrom *b* for mammals

Cytochrome c oxidase I (COI or CoxI) 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

Barcoding, molecular taxonomy, and exploration of the diversity of shrews (Soricomorpha: Soricidae) on Mount Nimba (Guinea)

FRANÇOIS JACQUET^{1*}, VIOLAINE NICOLAS¹, CELINE BONILLO²,
CORINNE CRUAUD³ and CHRISTIANE DENYS¹

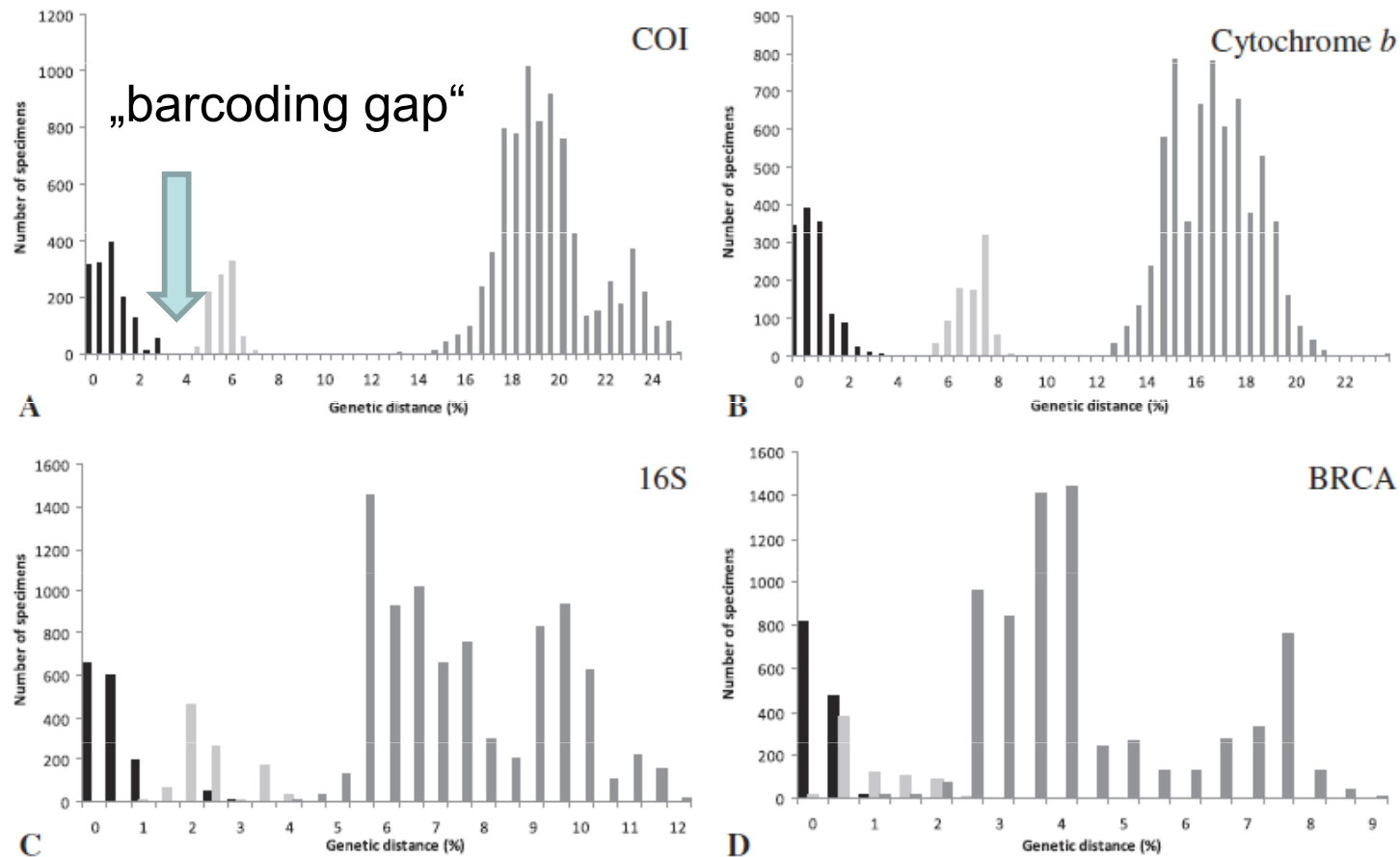


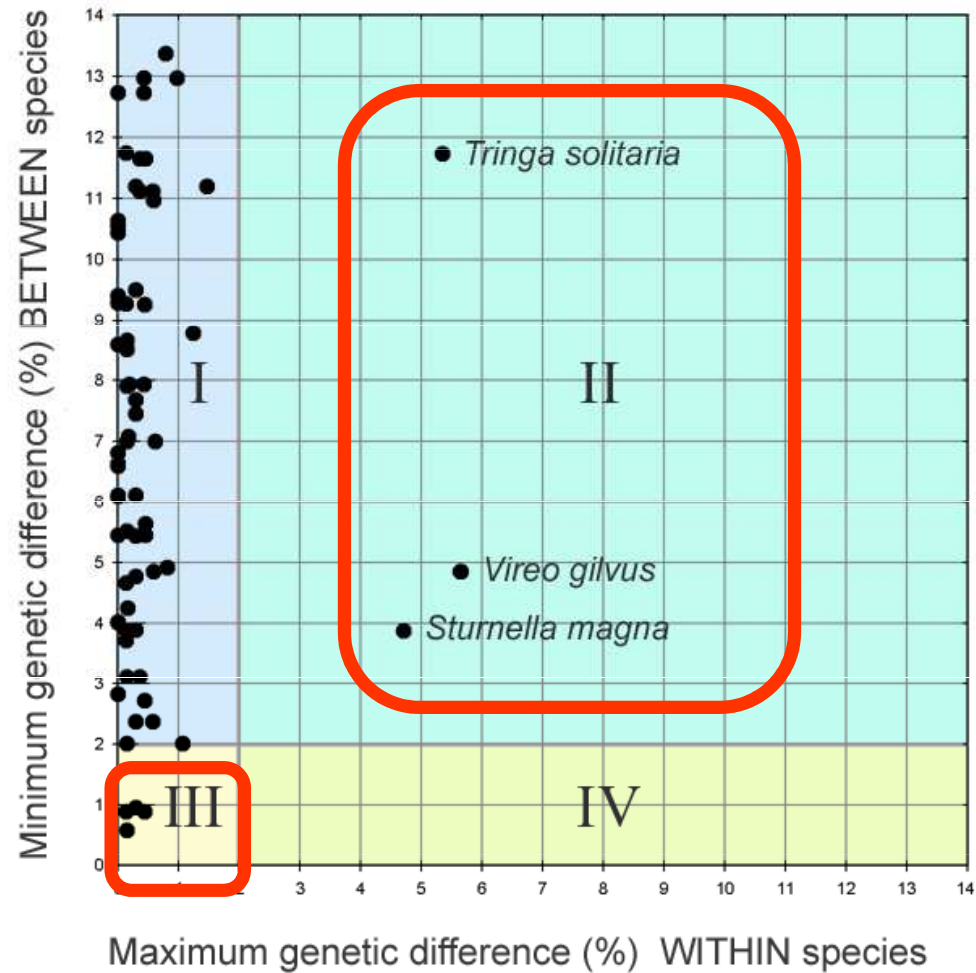
Figure 3. Histograms of divergence percentage values between all sequences. Black, intraspecific distances; light grey, interspecific distances amongst clade C; dark grey, interspecific distances excluding clade C. All sequences were corrected using the Kimura two-parameter substitution model. Sequences corresponding to nuclear copies of cytochrome *b* were excluded from the analysis. BRCA, breast cancer gene; COI, cytochrome oxidase I.

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 barcodes with closely-related species - many of them hybridize regularly
- in all groups studied so far, distinct barcode clusters with biologic co-variation suggest cryptic species

Barcoding North American birds

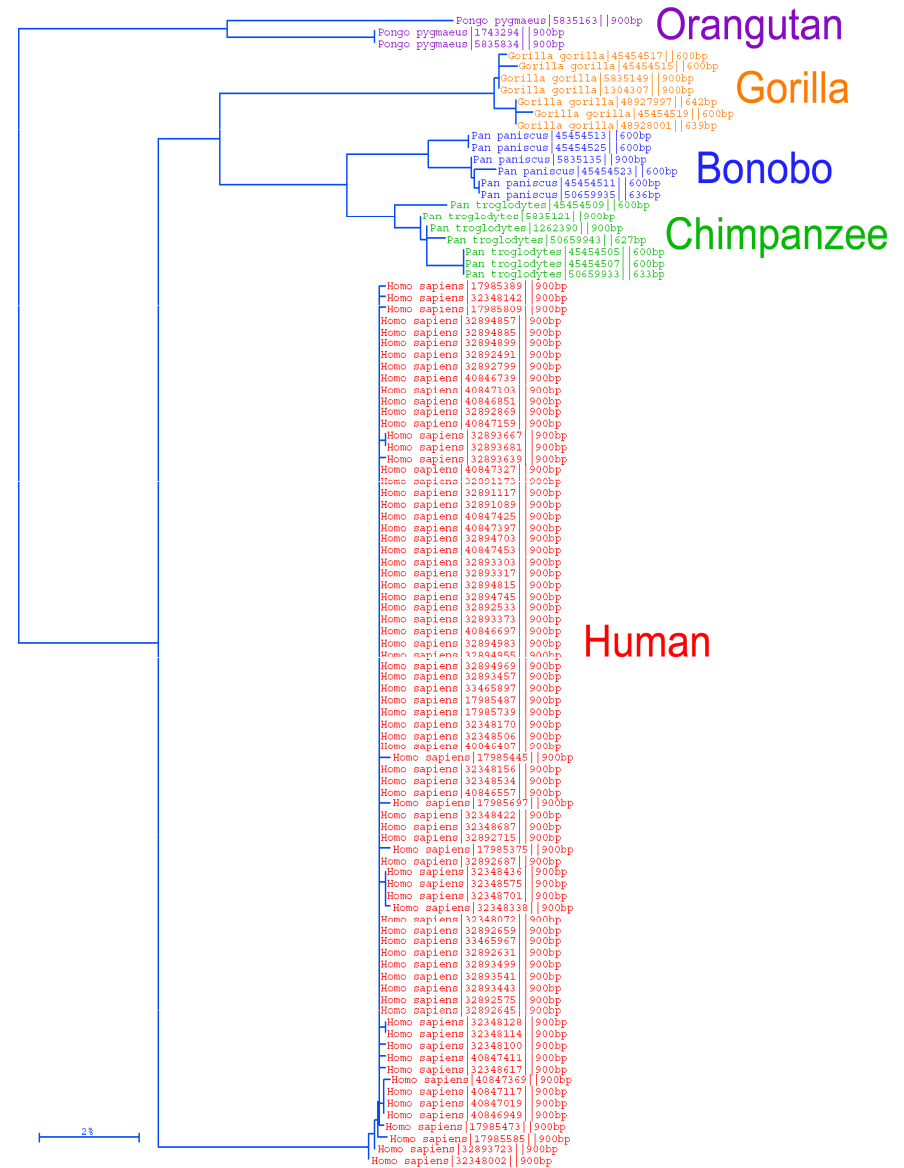
Interspecific vs. intraspecific COI barcode differences



Tak co třeba znamená toto?

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

Next generation sequencing of amplicons

1. Metagenomika/metatranskriptomika

- celé společenstvo půdních, vodních mikroorganismů, střevní mikroflóra
- PCR genu 16S (18S) rRNA
- lze i kvantifikovat

2. Složení potravy (COI barcoding)

- COI barcoding (masožravci)
- cpDNA (býložravci)

3. Analýza kontaminovaných vzorků

OPEN ACCESS Freely available online



Next-Generation Sequencing for Rodent Barcoding: Species Identification from Fresh, Degraded and Environmental Samples

Maxime Galan^{1*}, Marie Pagès^{1,2}, Jean-François Cosson¹

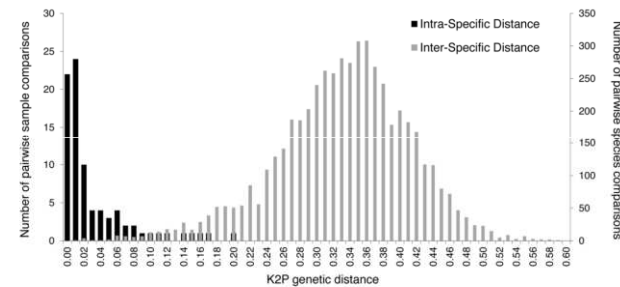


Figure 2. Distribution of pairwise K2P genetic distances within and between 103 rodent species (265 individuals) based on the 136 bp mini-barcode (cytb).

|SE|S|AM|E| BARCODE

SEquences Seeker & AMplicons Explorer for Barcoding

Genotyping and barcoding based on high-throughput multiplex amplicon sequencing

Illumina sequencers

Illumina MiSeq

4 millions reads/run
150 bp/read



Illumina GAIIx

300 millions reads/run
150 bp/read



Illumina HighSeq

1500 – 3000 millions reads/run
100 bp/read

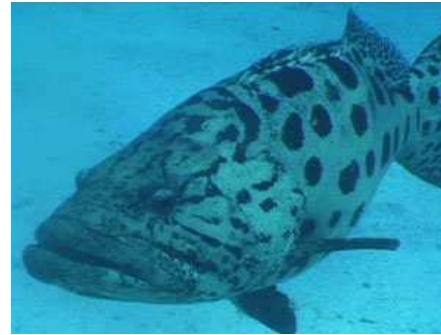


A barcoder? ... COMING SOON



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

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?

- horizontal gene transfer
- gene tree vs. organismal tree
- incomplete lineage sorting (usually recently diverged species)
- nuclear pseudogenes
- hybrids – mtDNA introgression

1. Horizontal gene transfer

DNA barcoding cannot reliably identify species of the blowfly genus *Protophthora* (Diptera: Calliphoridae)

T.L Whitworth, R.D Dawson, H Magalon and E Baudry

Proc. R. Soc. B 2007 **274**, doi: 10.1098/rspb.2007.0062, published 22 July 2007



Wolbachia within an insect cell
(25-70% druhů hmyzu)

Výsledky jaderné a mitochondriální DNA si neodpovídají

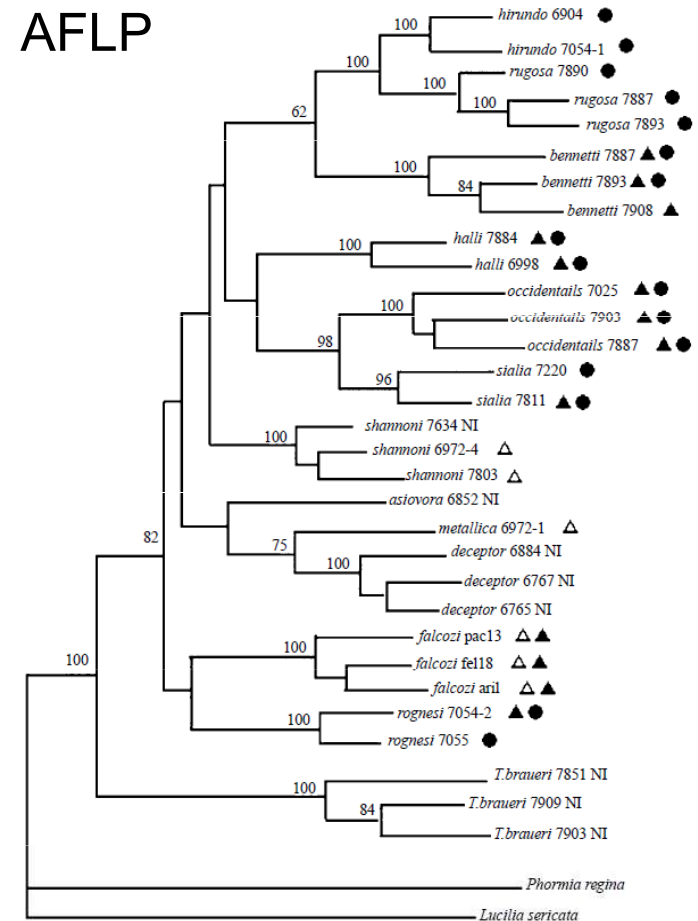
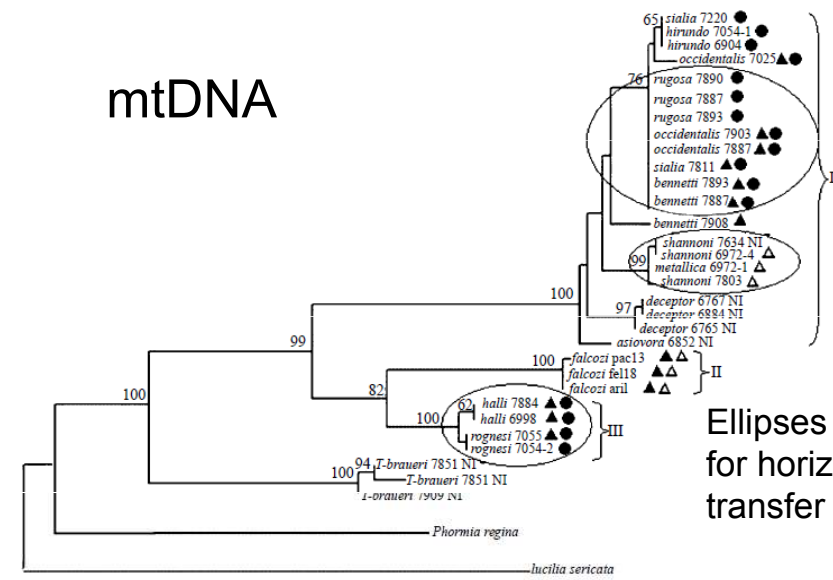


Figure 1. Phylogram of the *Protocalliphora* genus based upon AFLP data. The tree was generated by parsimony analysis using a heuristic search with tree bisection-reconnection. Bootstrap values are shown as percentage of 1000 replicates at each node only if they are 50% or greater. The *Wolbachia* infection status of each individual is shown on the tree. Individuals infected with *wA1*, *wA2* or *wB* *Wolbachia* strains are respectively represented by an open triangle, a solid triangle and a circle. Non-infected individuals are symbolized by NI.



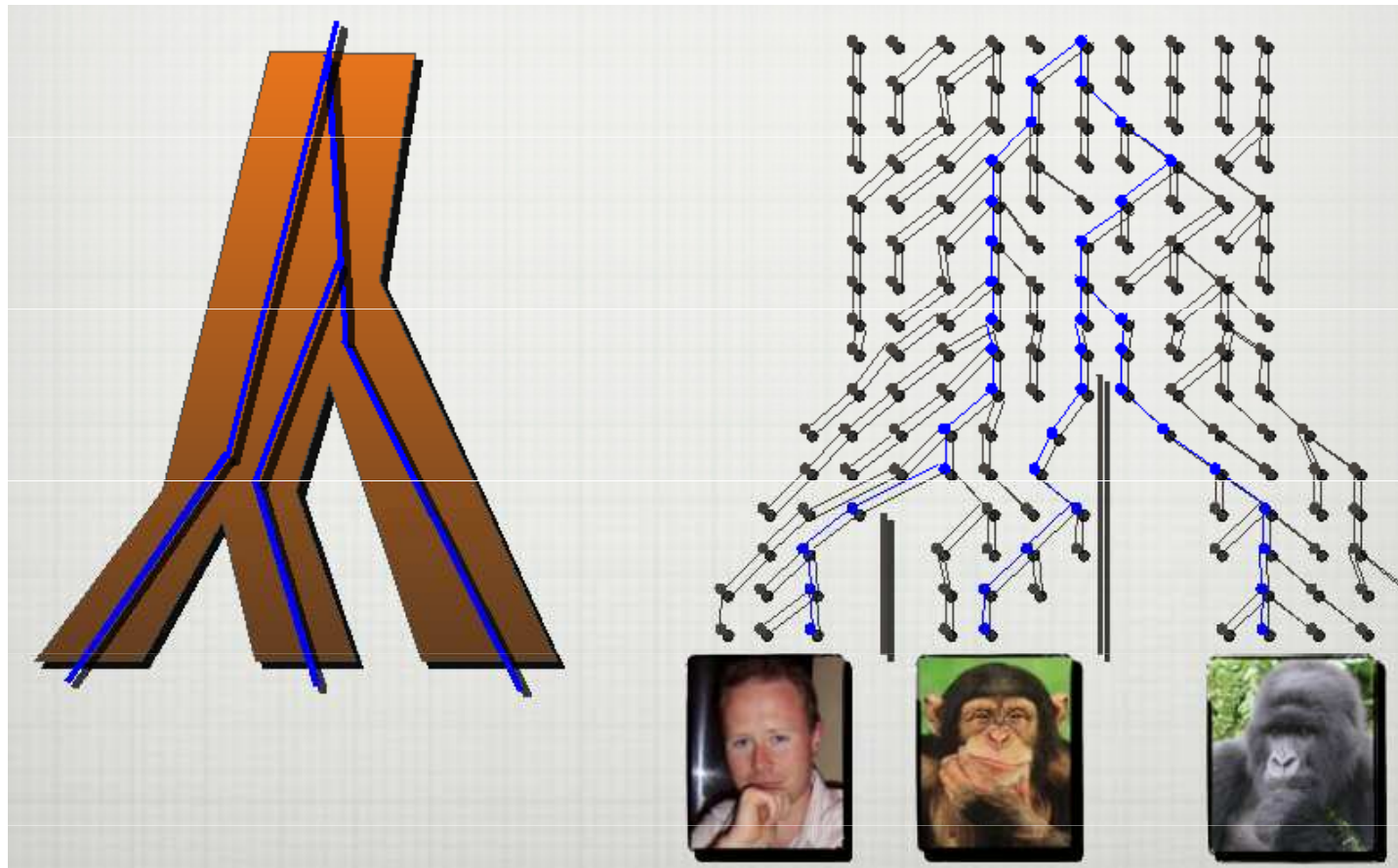
Ellipses = candidates for horizontal gene transfer

Figure 2. Phylogram of the *Protocalliphora* genus based upon COI and COII data (total of 953 bp). The tree was generated by maximum likelihood analysis using a heuristic search with tree bisection-reconnection. Bootstrap values are shown as percentage of 1000 replicates at each node only if they are 50% or greater. The *Wolbachia* infection status of each individual is shown on the tree. Individuals infected with *wA1*, *wA2* or *wB* *Wolbachia* strains are respectively represented by an open triangle, a solid triangle and a circle. Non-infected individuals are symbolized by NI. Three clusters defined using 3 or 1.8% divergence as threshold values (§3) are shown on the figure. The three ellipses indicate cases where horizontal transfer of *Wolbachia* between species seems probable (§3).

Horizontální přenos mtDNA přes *Wolbachia* (většinou mezi blízce příbuznými druhy, na úrovni rodů je barcoding OK)

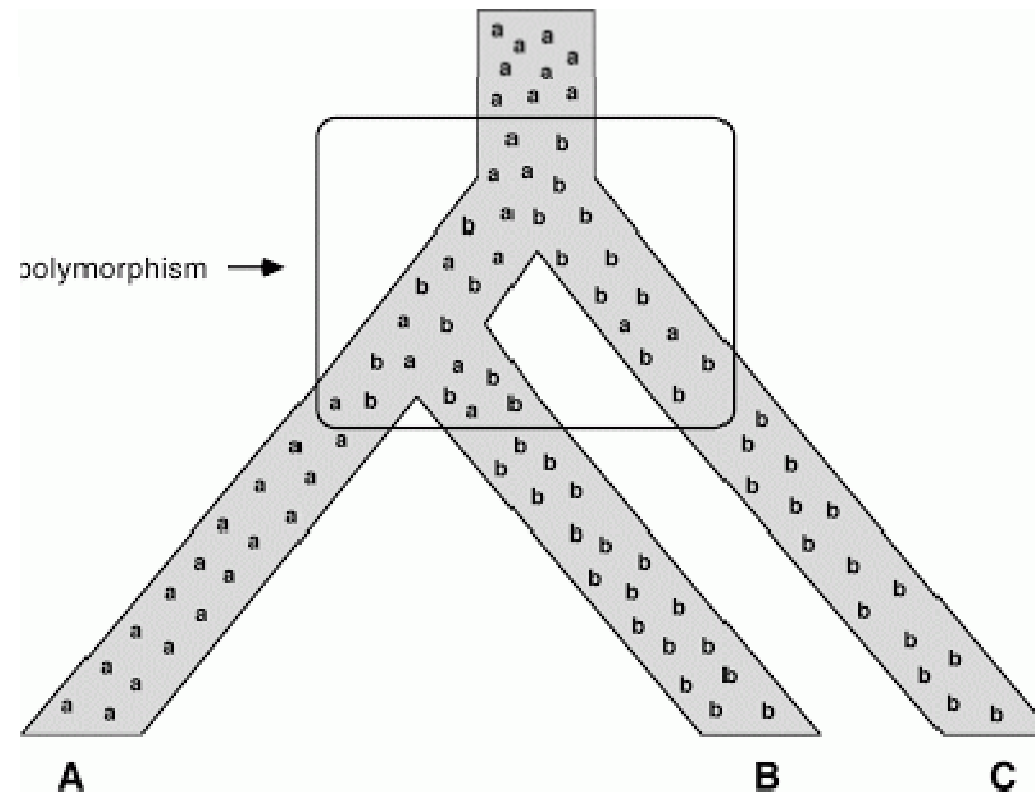
Symboly vpravo odpovídají typu *Wolbachia* infekce

2. Gene tree vs. organismal tree



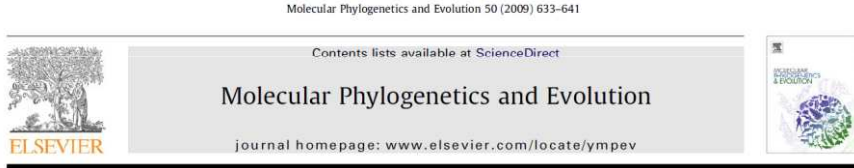
Jeden gen (navíc mitochondriální) není dostatečný pro definici druhu

3. Incomplete lineage sorting



Čas potřebný k fixaci polymorfismu = $2 N_e$ generací

4. Pseudogenes



False phylogenies on wood mice due to cryptic cytochrome-*b* pseudogene
 Sylvain Dubey^{a*}, Johan Michaux^b, Harald Br nner^c, Rainer Hutterer^d, Peter Vogel^e

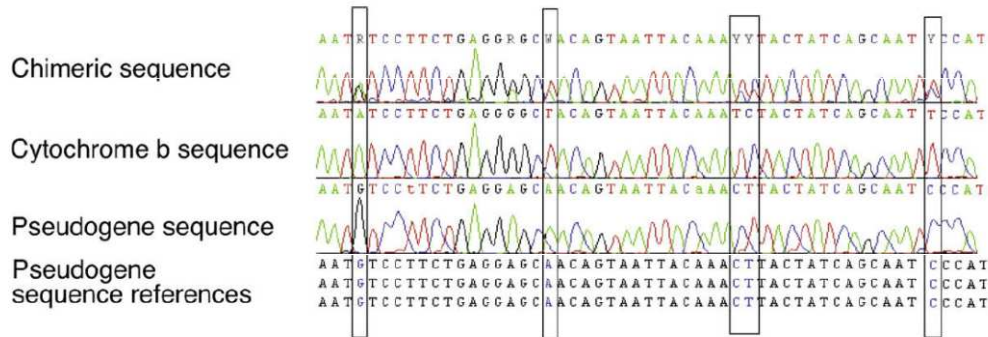
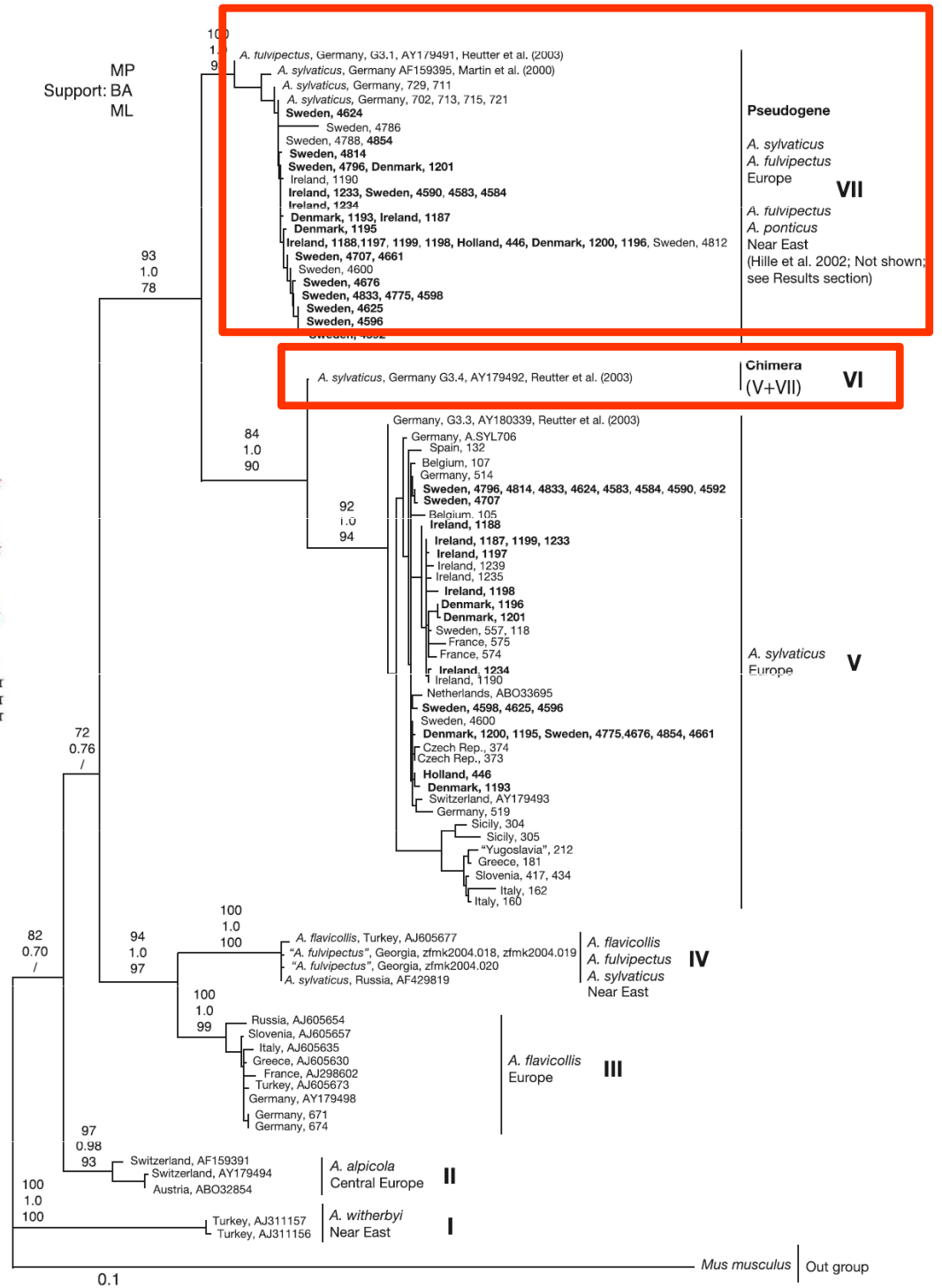


Fig. 2. Examples of electropherograms of a chimeric *cyt-b* sequence of lineage V/VII (VI) and of pure lineages V and VII.

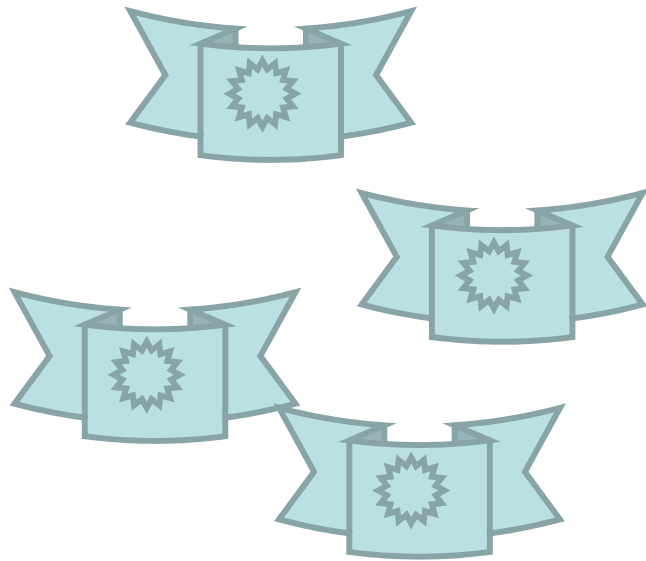
Heterozygotn  sekvenca v mtDNA
 → bud'te ostr ziti!

NUMTS = „nuclear copy of mtDNA sequences

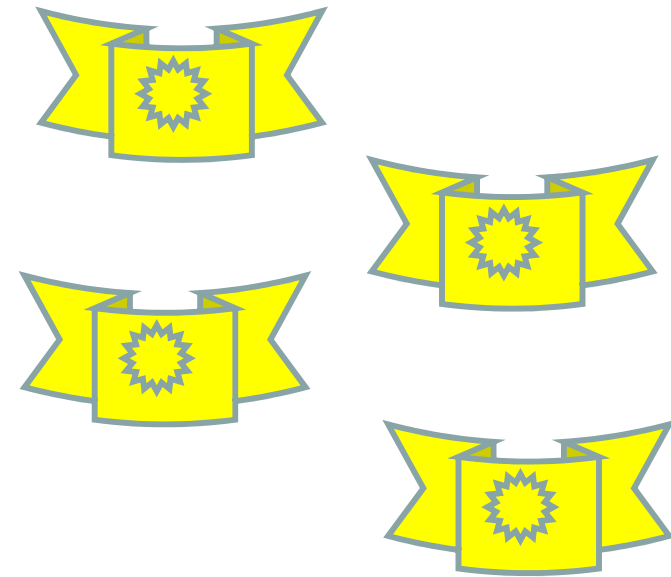
Dosty  asto gen pro cytochrom *b*



5. Introgression of mtDNA



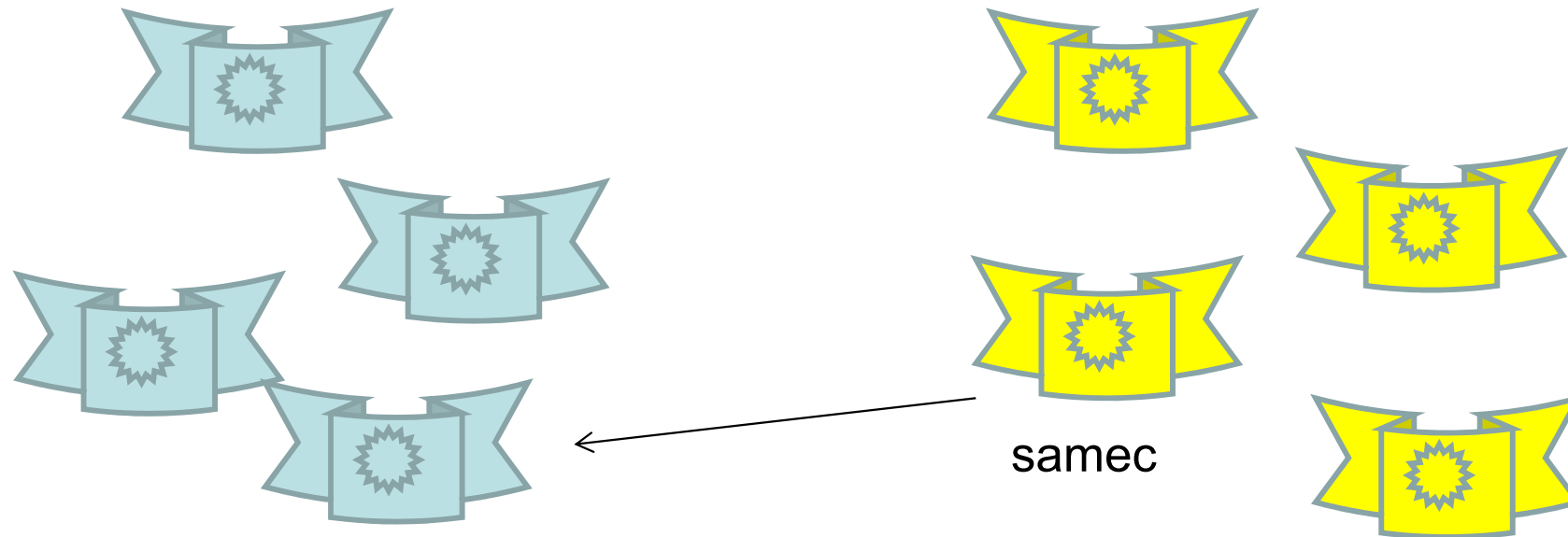
Myotis myotis - Evropa



Myotis blythii - Asie



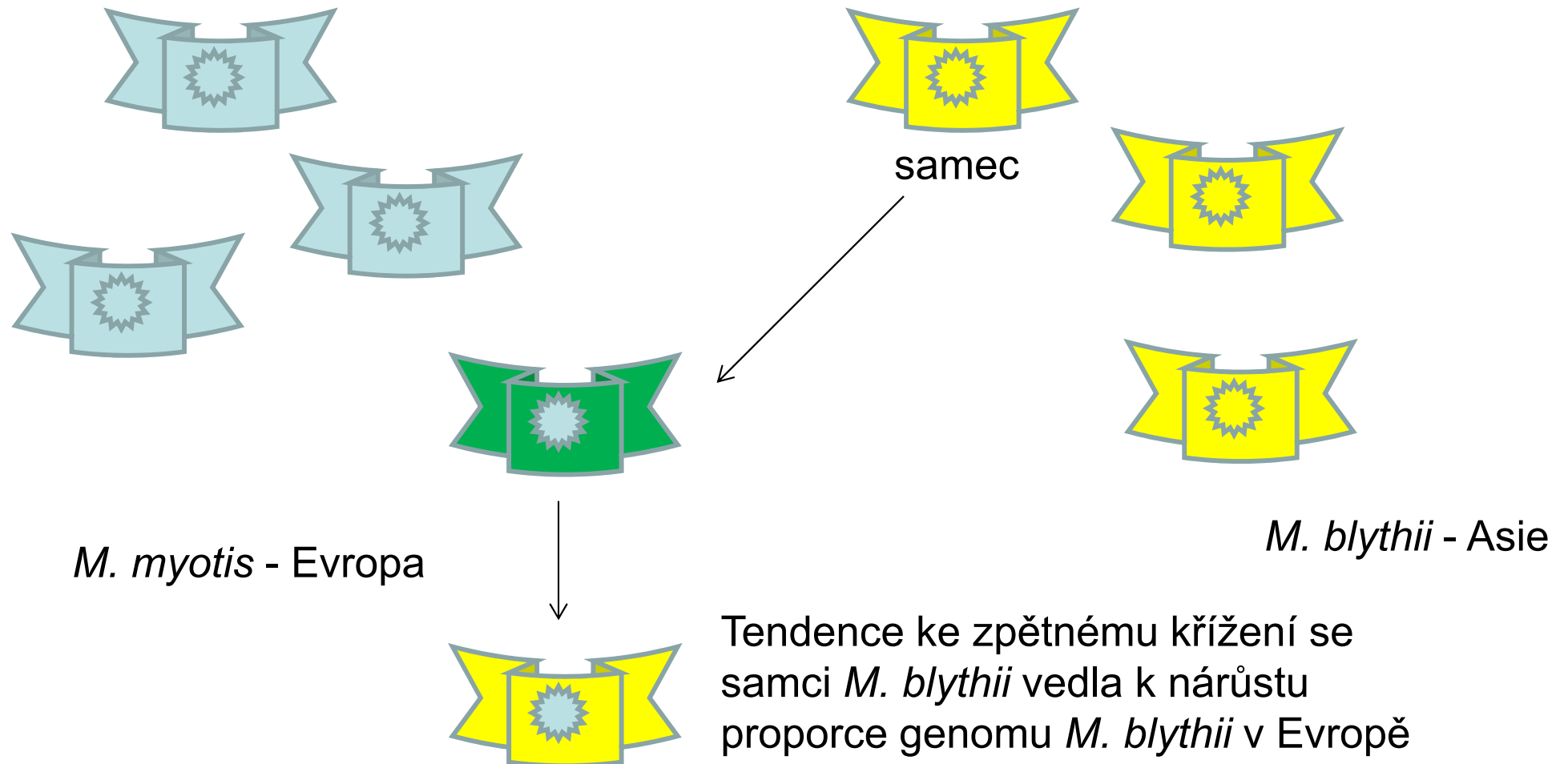
Příklad: *Myotis blythii* vs. *Myotis myotis* - introgrese mtDNA



M. myotis - Evropa

M. blythii - Asie

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



Kolonizující (invazní) druh „ukradne“ mtDNA původnímu druhu (Currat et al. 2008)

Plio-Pleistocene history of West African Sudanian savanna and the phylogeography of the *Praomys daltoni* complex (Rodentia): the environment/geography/genetic interplay

J. BRYJA,* L. GRANJON,^{†1} G. DOBIGNY,^{†2} H. PATZENHAUEROVÁ,* A. KONEČNÝ,^{**}
 J. M. DUPLANTIER,[†] P. GAUTHIER,[†] M. COLYN,[‡] L. DURNEZ,[§] A. LALIS[¶] and V. NICOLAS[¶]

Kolik je zde druhů?

Čtyři podle mtDNA genotypu – cca 7% divergence (cyt b)? -/97/-

Dva druhy podle fenotypu?

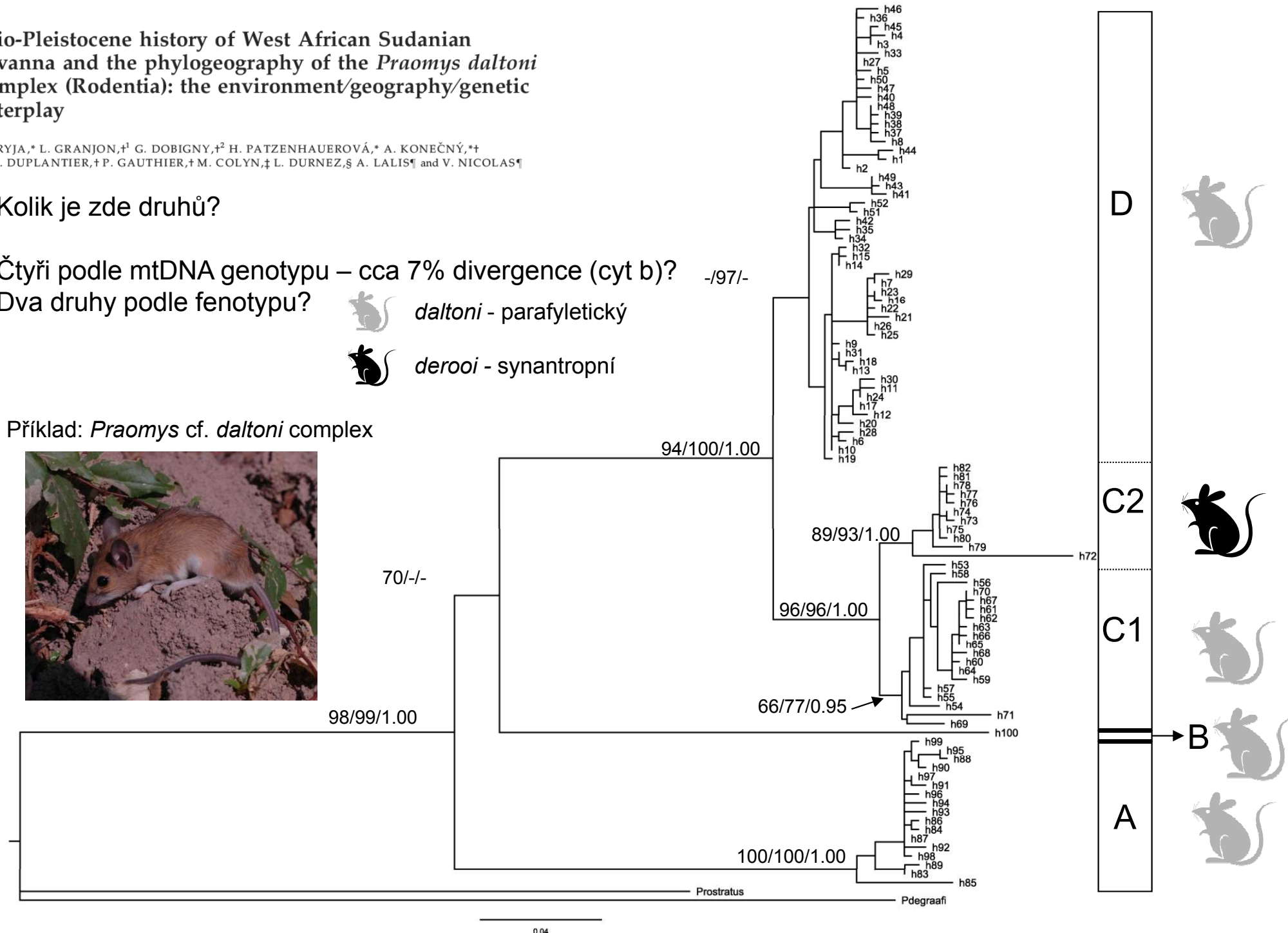


daltoni - parafyletický

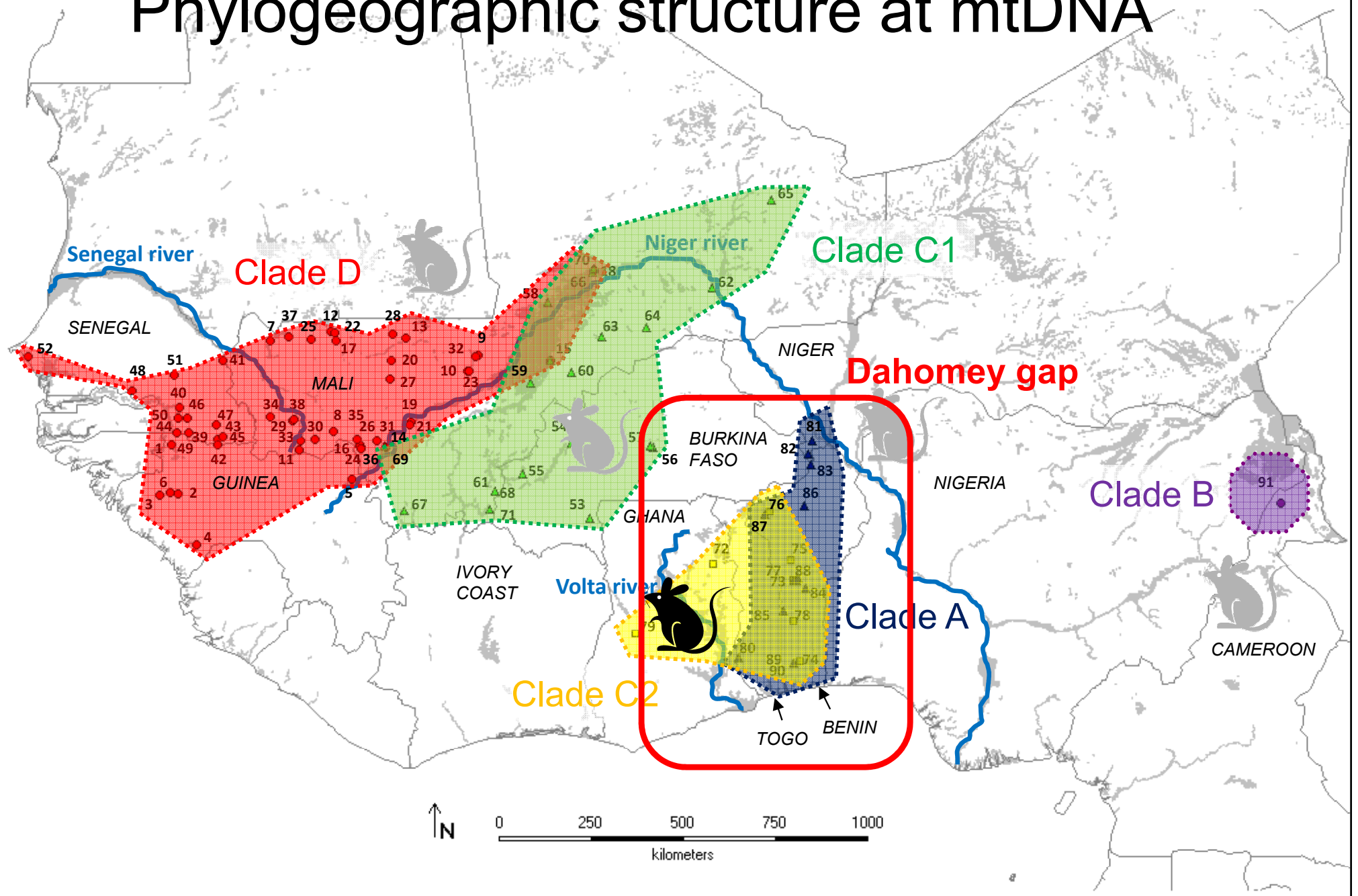


derooi - synantropní

Příklad: *Praomys* cf. *daltoni* complex



Phylogeographic structure at mtDNA



Partial mtDNA introgression in Dahomey gap

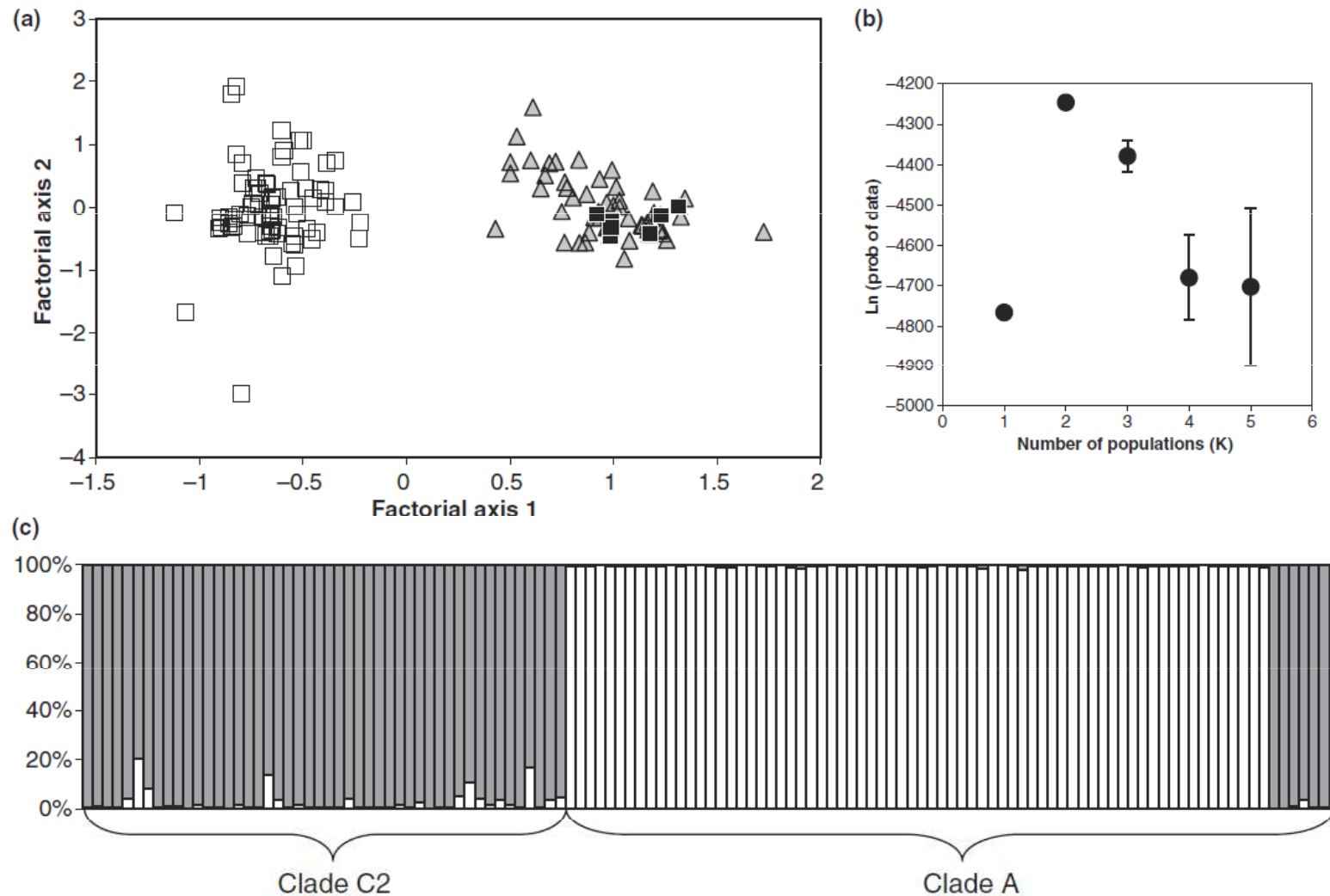
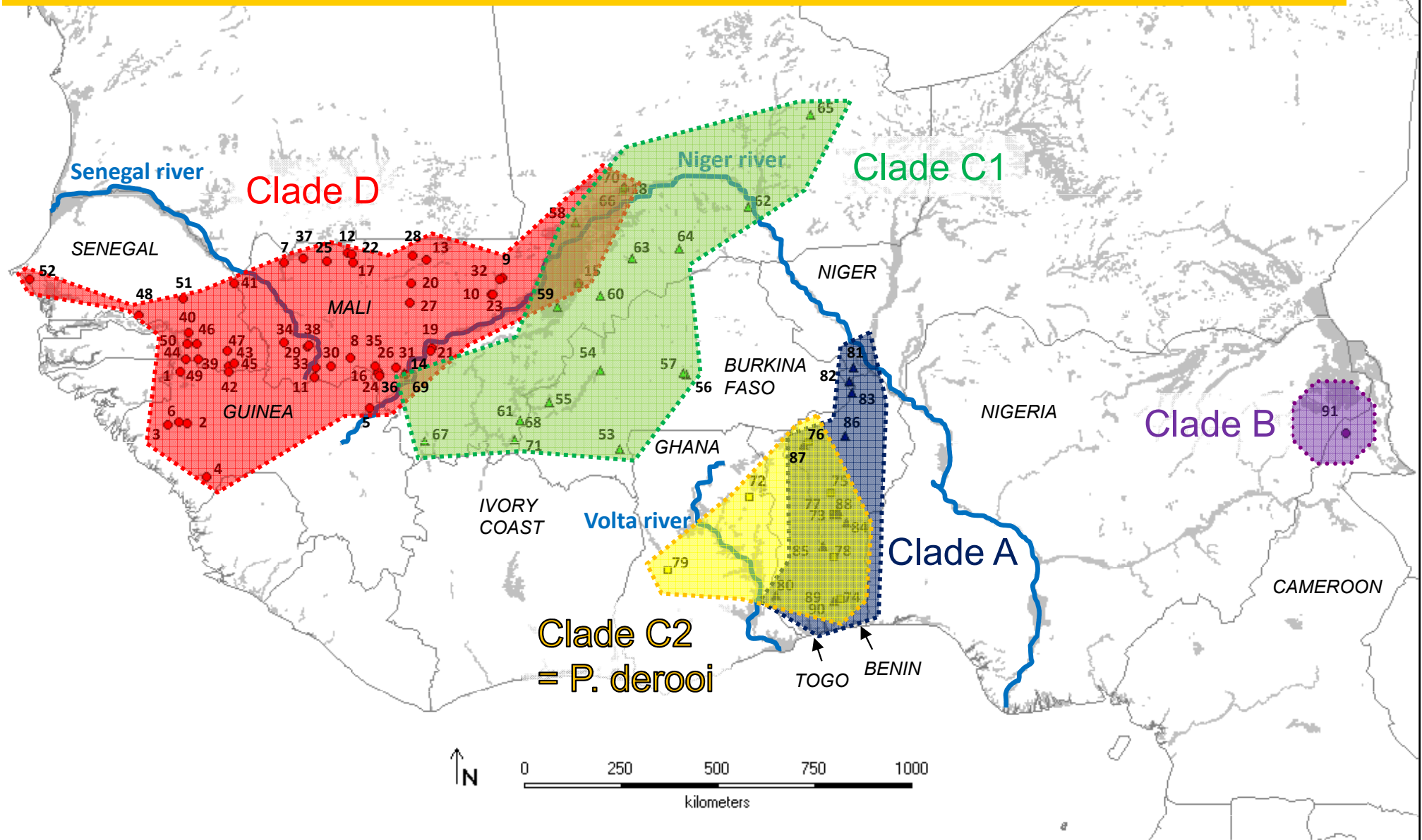
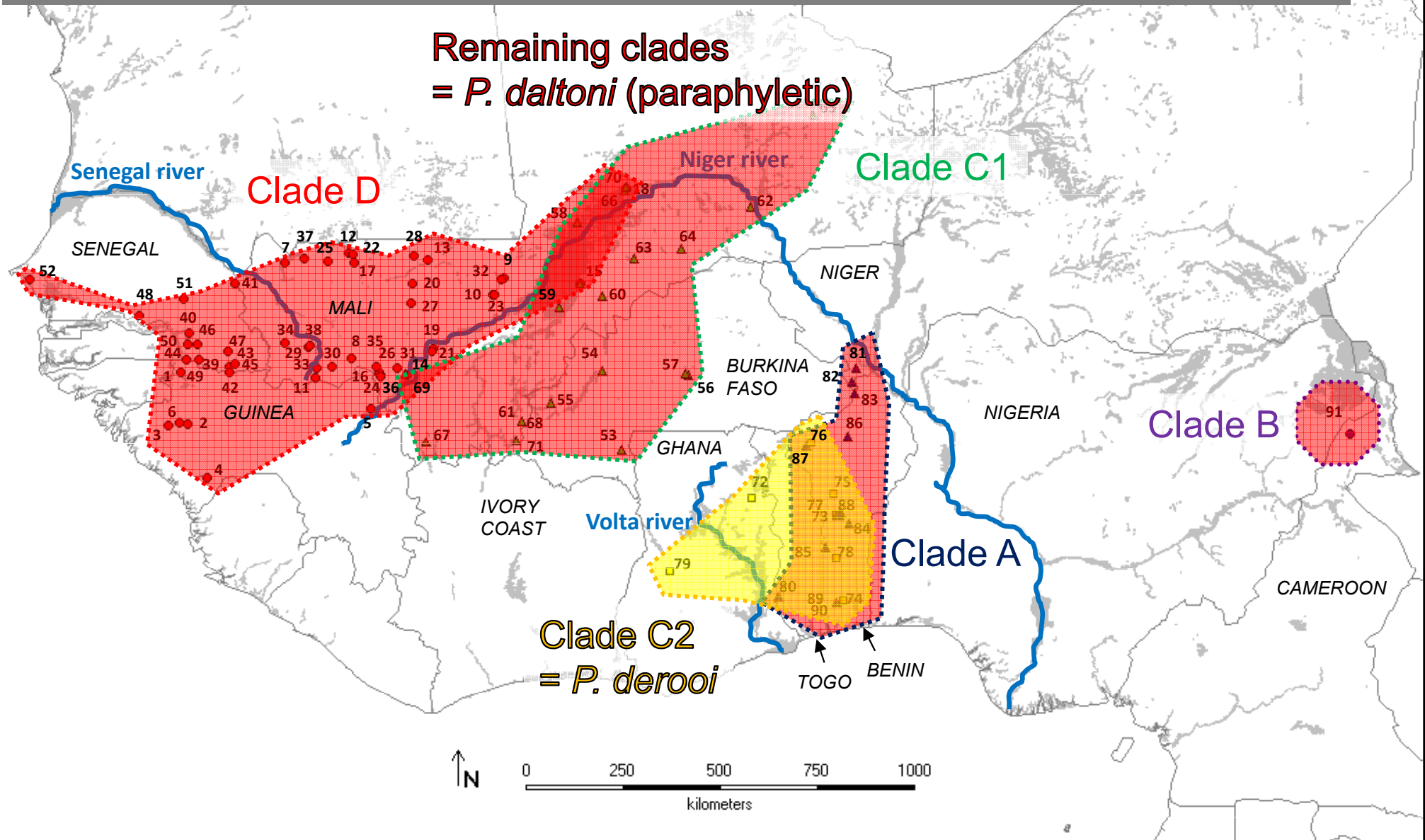


Fig. 6 Results of microsatellite analysis of individuals from Benin. (a) Factorial correspondence analysis in Genetix; *squares*—individuals with mtDNA from clade A, *triangles*—from clade C2. Note the seven individuals with mtDNA from clade A, but clustered within individuals with mtDNA from clade C2 (marked by *black squares*). (b) Likelihood of models in STRUCTURE for increasing number of populations (K); the highest likelihood is observed for $K = 2$. (c) Assignment of individuals to particular populations using the best model (i.e. $K = 2$; the run with the highest likelihood) in STRUCTURE.

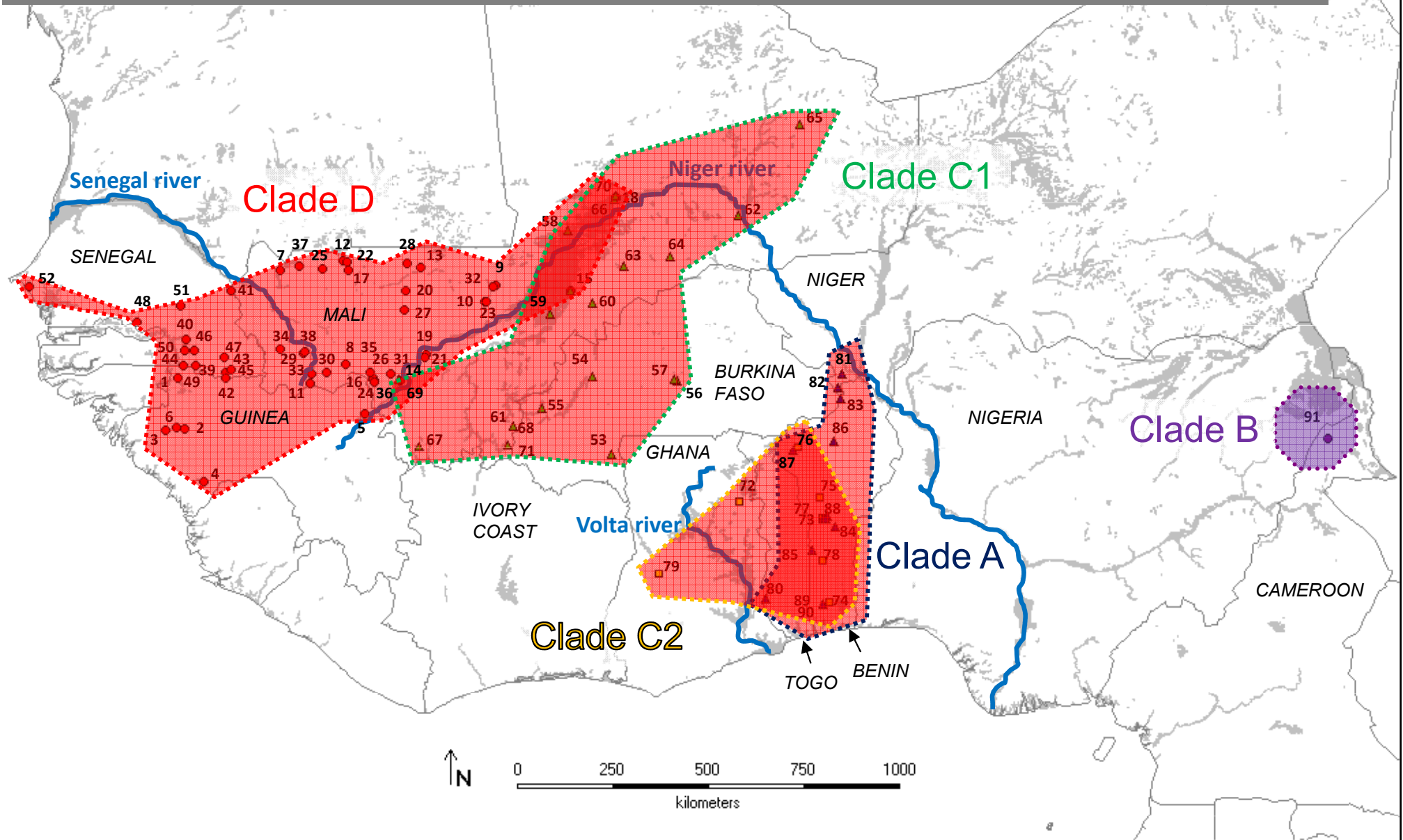
What is species???



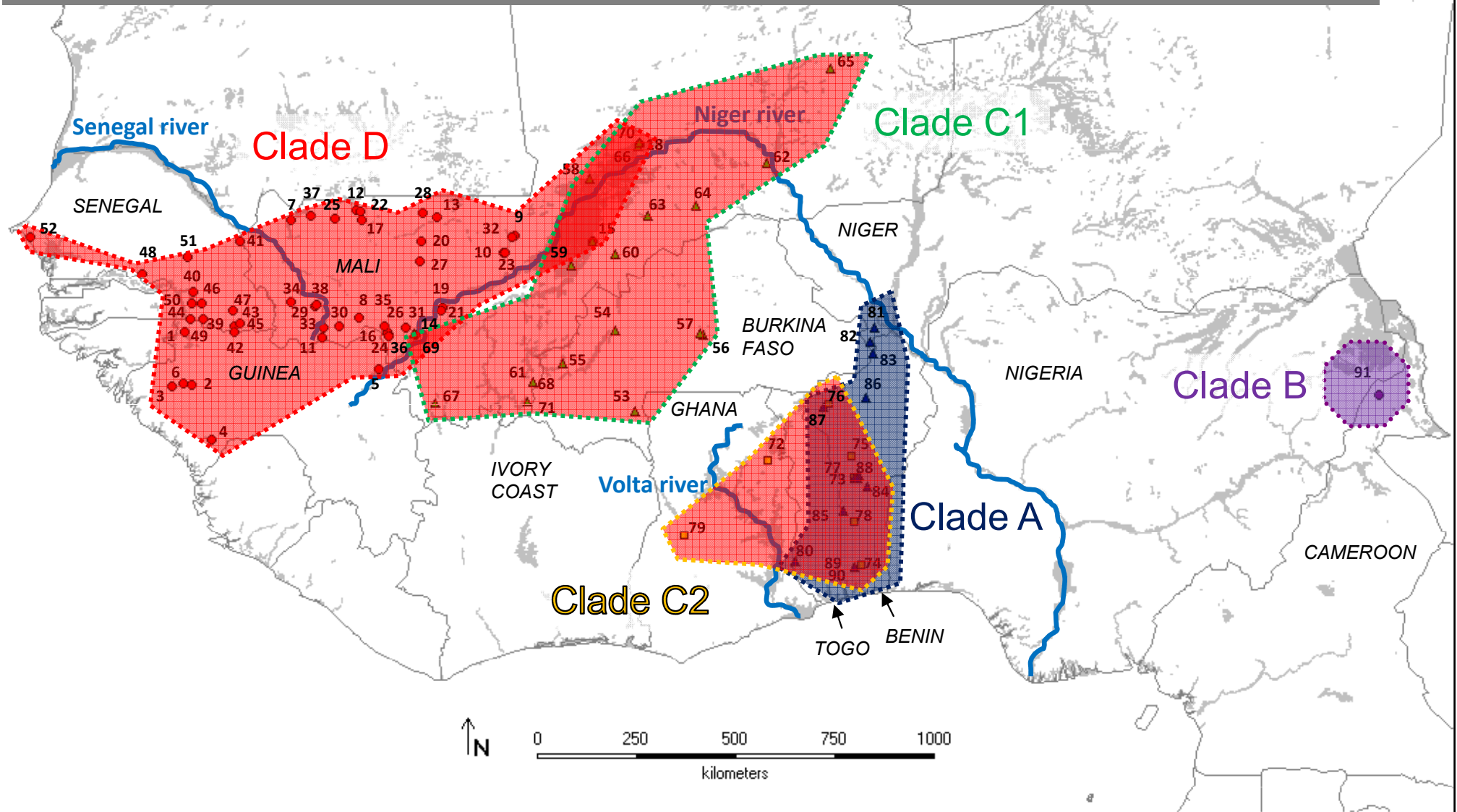
Morphology and ecology



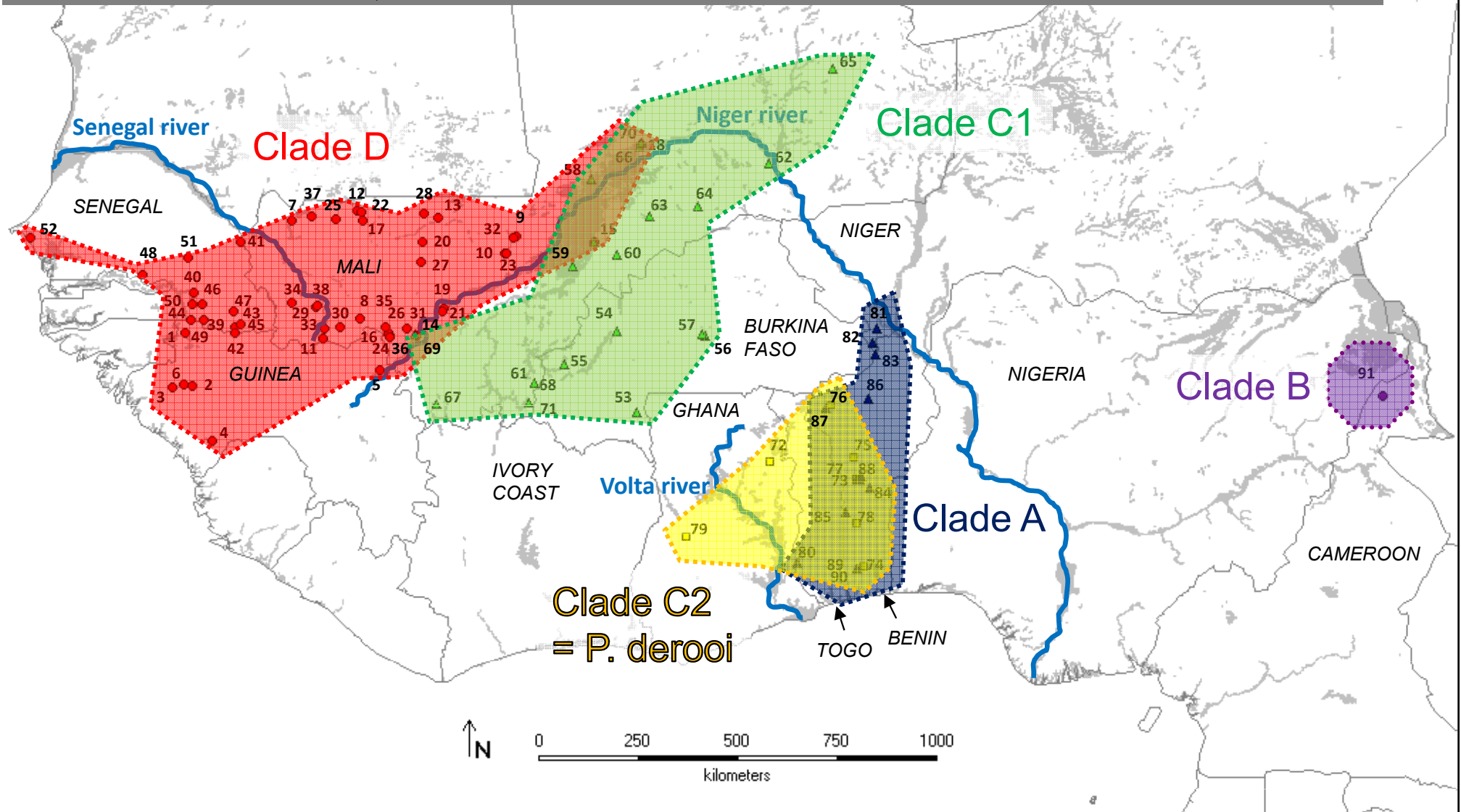
Karyotypes



Mitochondrial DNA + microsatellites in Benin + karyotypes



Splitting approach taking morphology and ecology into account – the reproductive barriers between clades A, B, C1, and D remains to be identified



Identifikace jedinců

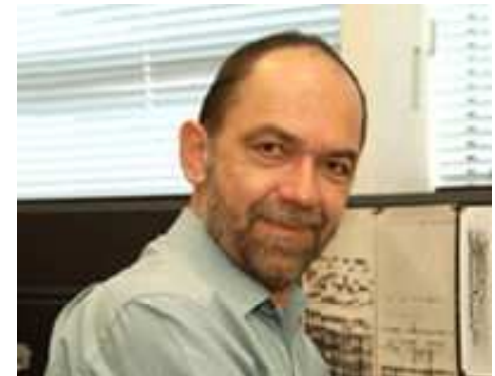
DNA fingerprinting

Identifikace jedinců – kdy?

- většinou když jedince nevidíme (příp. u zpětných odchyťů, když chybí značka)
- neinvazivní genetika – skrytá zvířata, vzorky trusu, moči, chlupů atd. – dá se spojit např. s individuální variabilitou potravy
- forenzní genetika – identifikace DNA v živočišných produktech, pytláctví, atd.
- druhová ochrana – např. sokolnictví (potvrzení rodičovství) apod.

Identifikace jedince – unikátní multilokusový genotyp

- minisatelitový fingerprinting
 - velké množství kvalitní DNA, technická náročnost
 - + univerzalita
- AFLP
 - kvalitní nedegradovaná DNA
 - + univerzalita



Alec Jeffreys, the pioneer of DNA profiling

- Mikrosatelity
 - + stačí malé množství nekvalitní DNA, **optimální pro neinvazivní přístupy**
 - je nutné znát konkrétní lokusy a sekvence specifických primerů
- SNPs
 - + velké množství znaků v genomu
 - malá variabilita, je potřeba hodně lokusů, nutno znát konkrétní lokusy

Mikrosatelity

- Tandemová opakování krátkých motivů

CTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT

- Izolace DNA

CTTTCTTTCTTTCTTTC

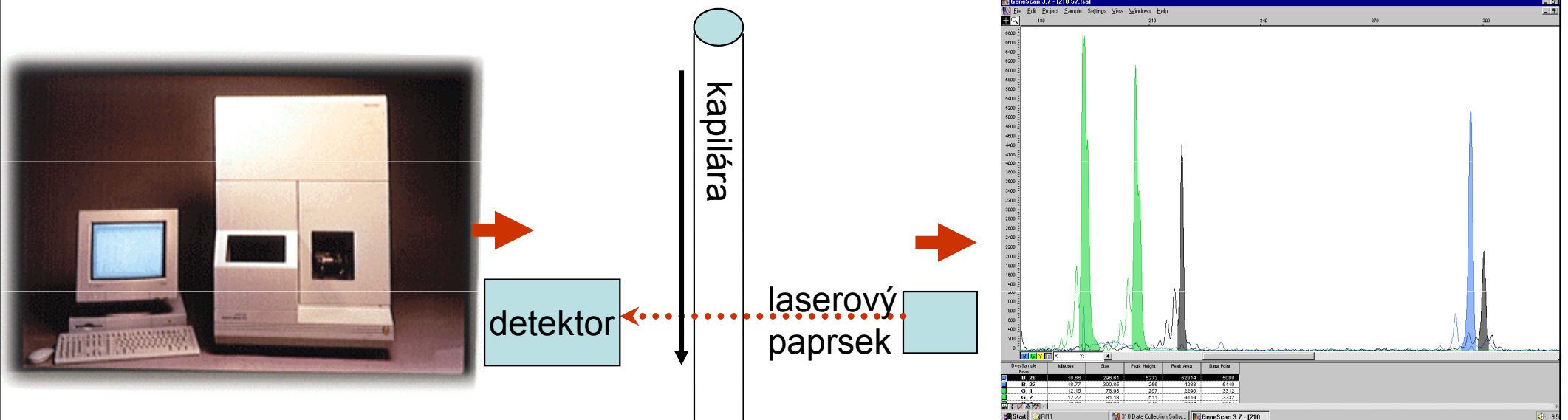
- PCR

CTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT

- Detekce

→ sekvenátor, fragmentační analýza

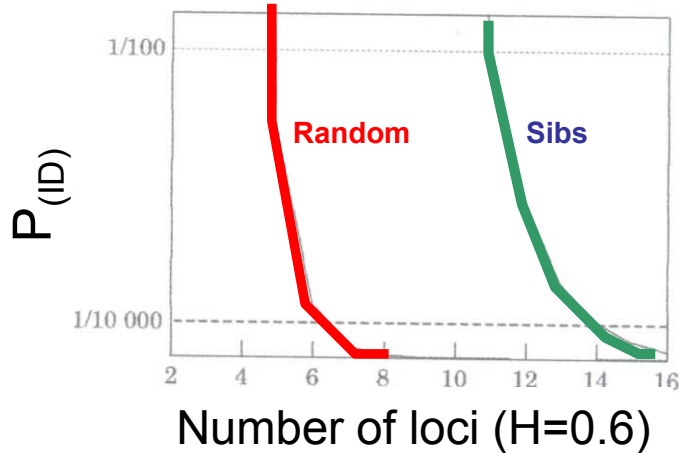
CTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT



Identifikace jedinců závisí na stupni polymorfismu

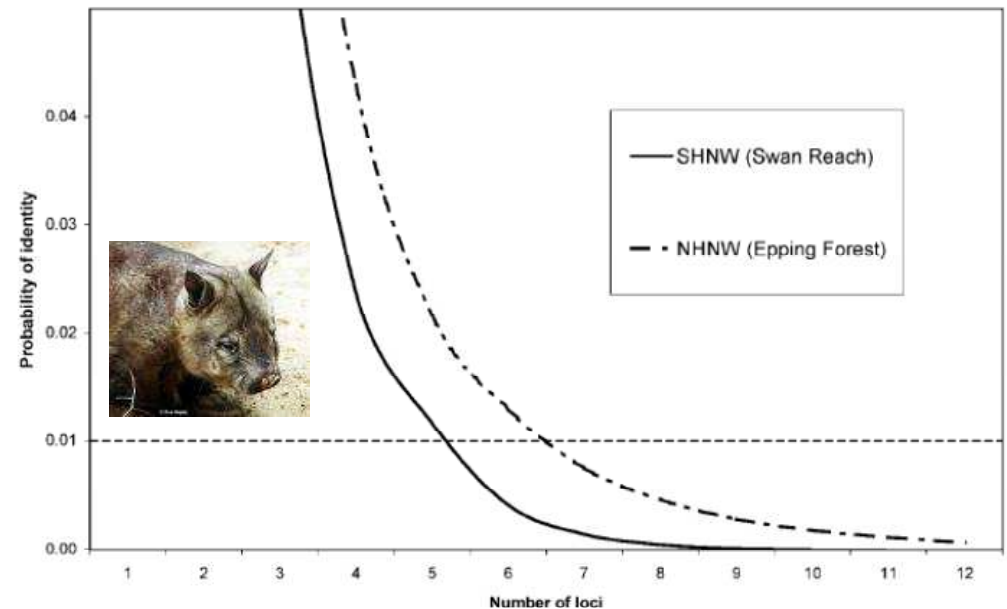
- multilocus microsatellite fingerprinting – power estimated as „probability of identity“ ($P_{(ID)}$) (Waits et al. 2001) – např. program GenAlex

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



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

- 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 s unikátními multilokusovými genotypy (o jednoho víc než podle stop a fotografií)
- Multiple-tube approach, mnohonásobné opakování PCR reakcí

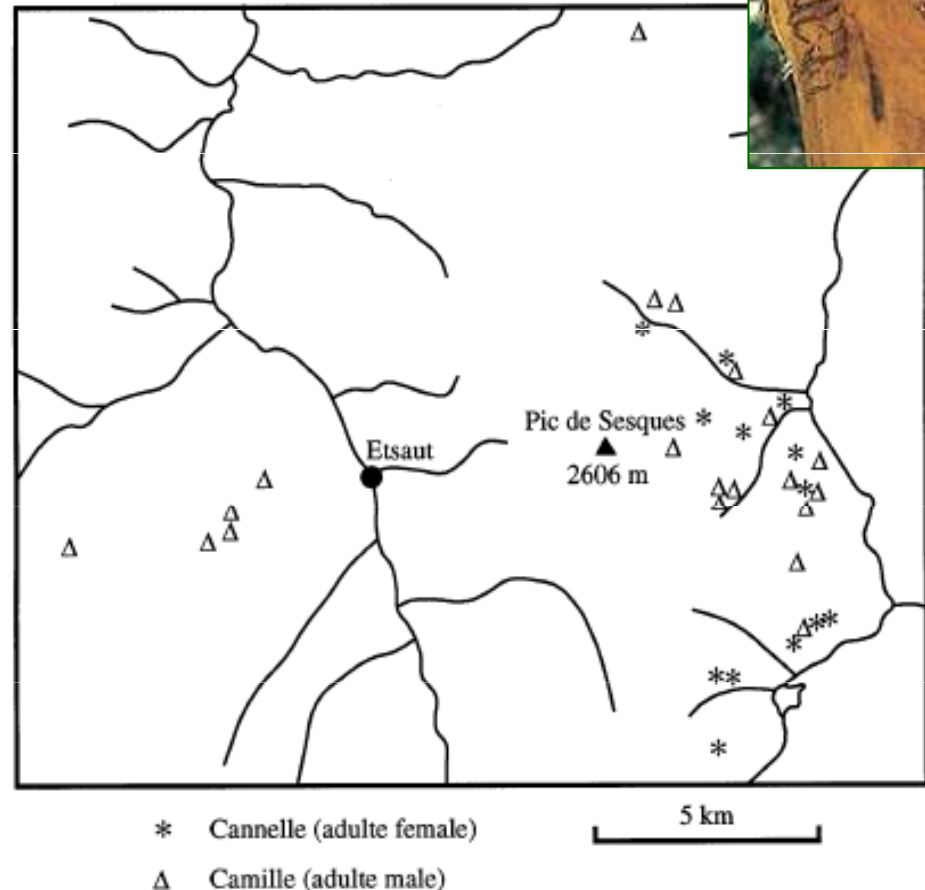


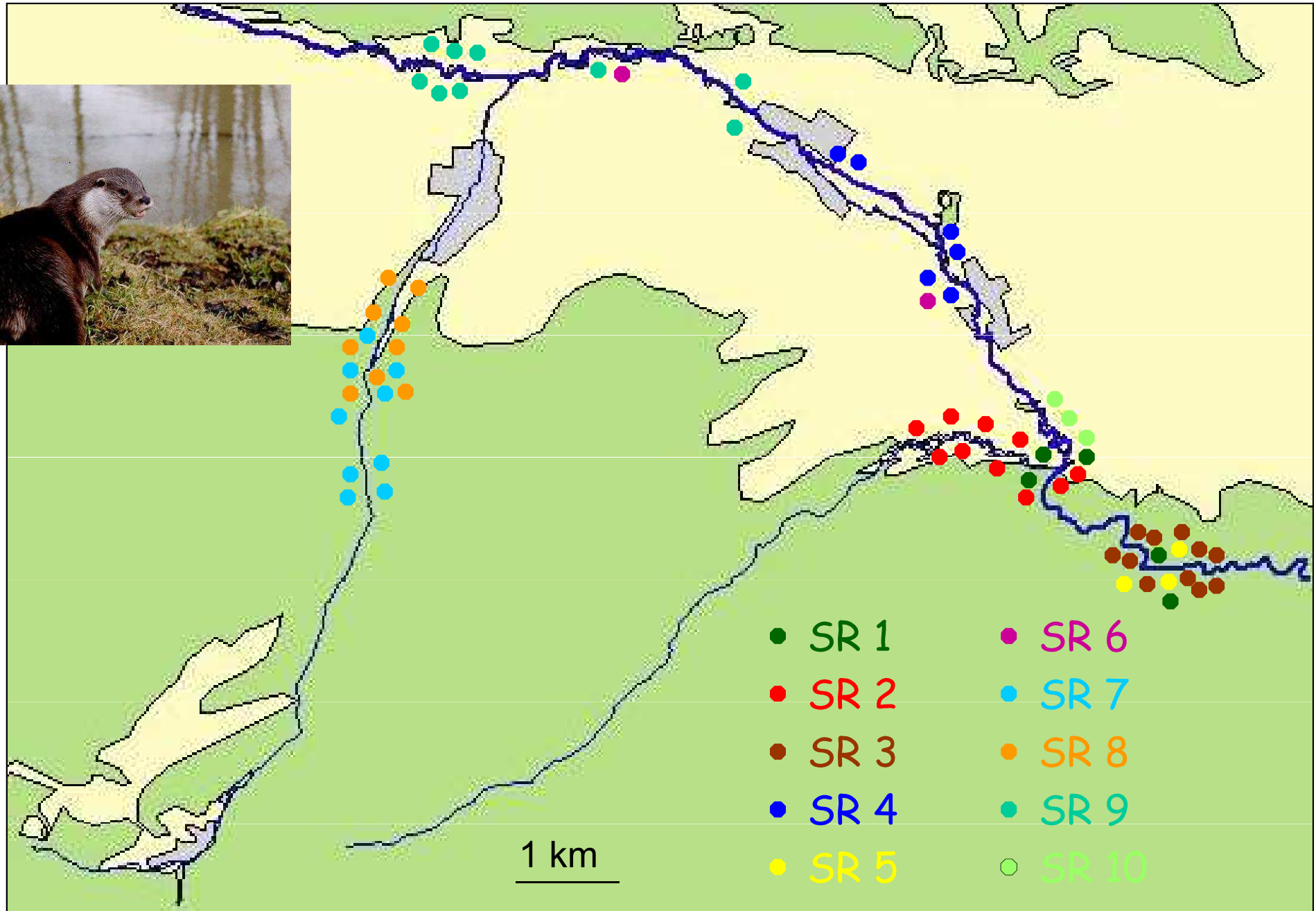
Fig. 3 Home range of two Pyrenean brown bears obtained by noninvasive genetic sampling and genotyping.

Analýza prostorové aktivity

- P. Hájková – PhD thesis

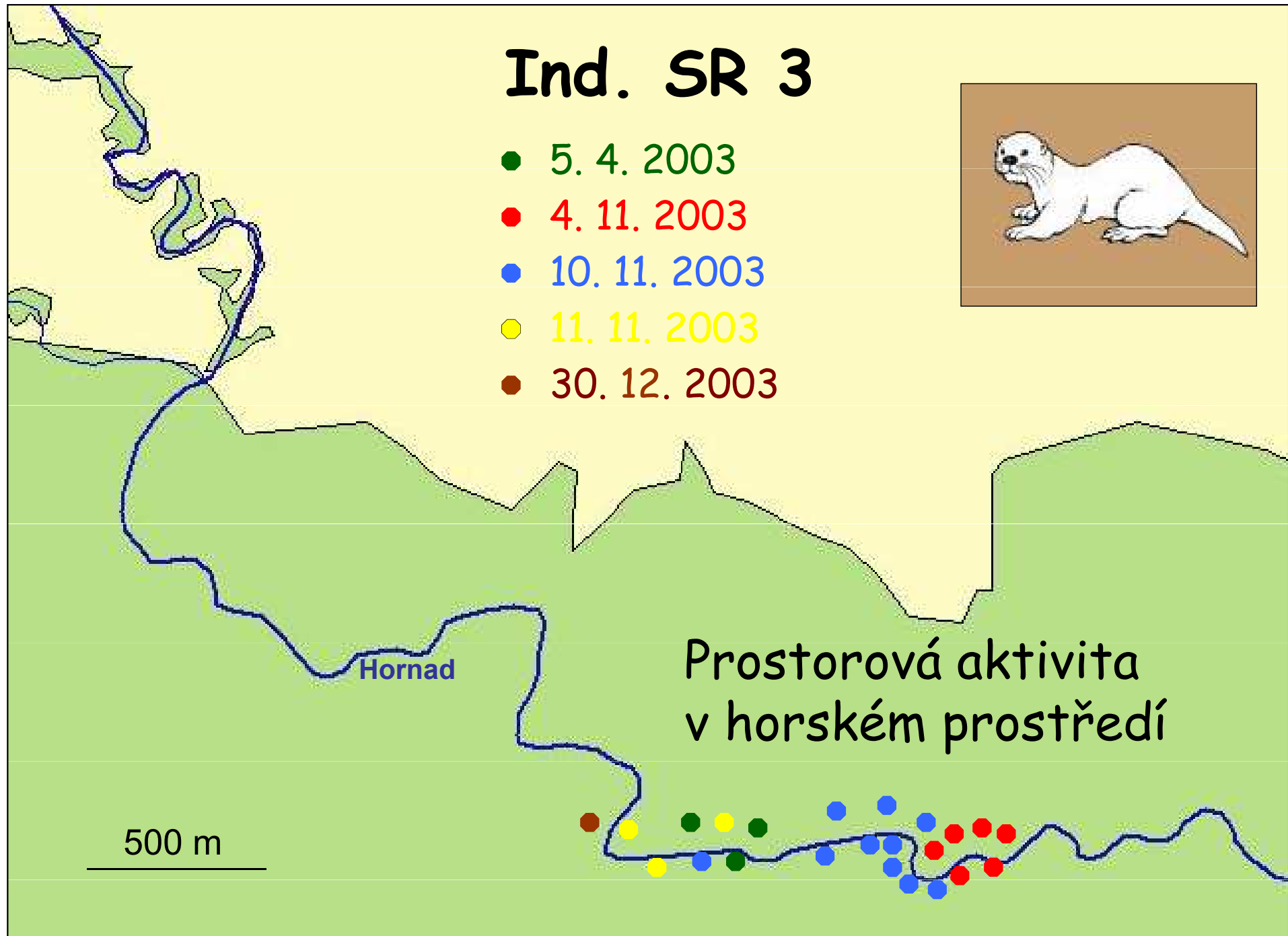
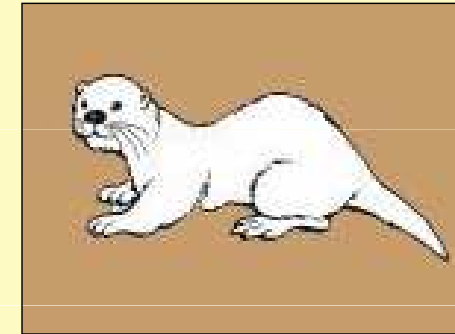


Identifikovaní jedinci - Hornád, NP Slovenský Raj



Ind. SR 3

- 5. 4. 2003
- 4. 11. 2003
- 10. 11. 2003
- 11. 11. 2003
- 30. 12. 2003



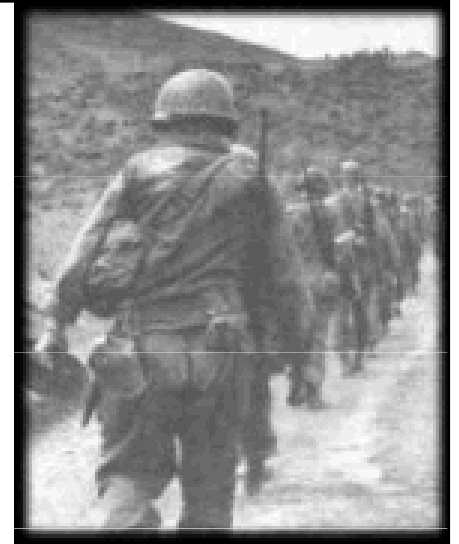
Lidská forenzní genetik

- **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**



Klony

Bambus *Sasa senanensis*

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



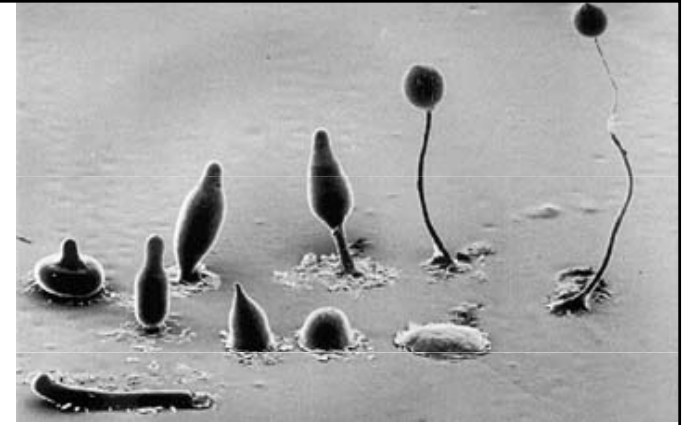
Slavní klonální bezobratlí

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

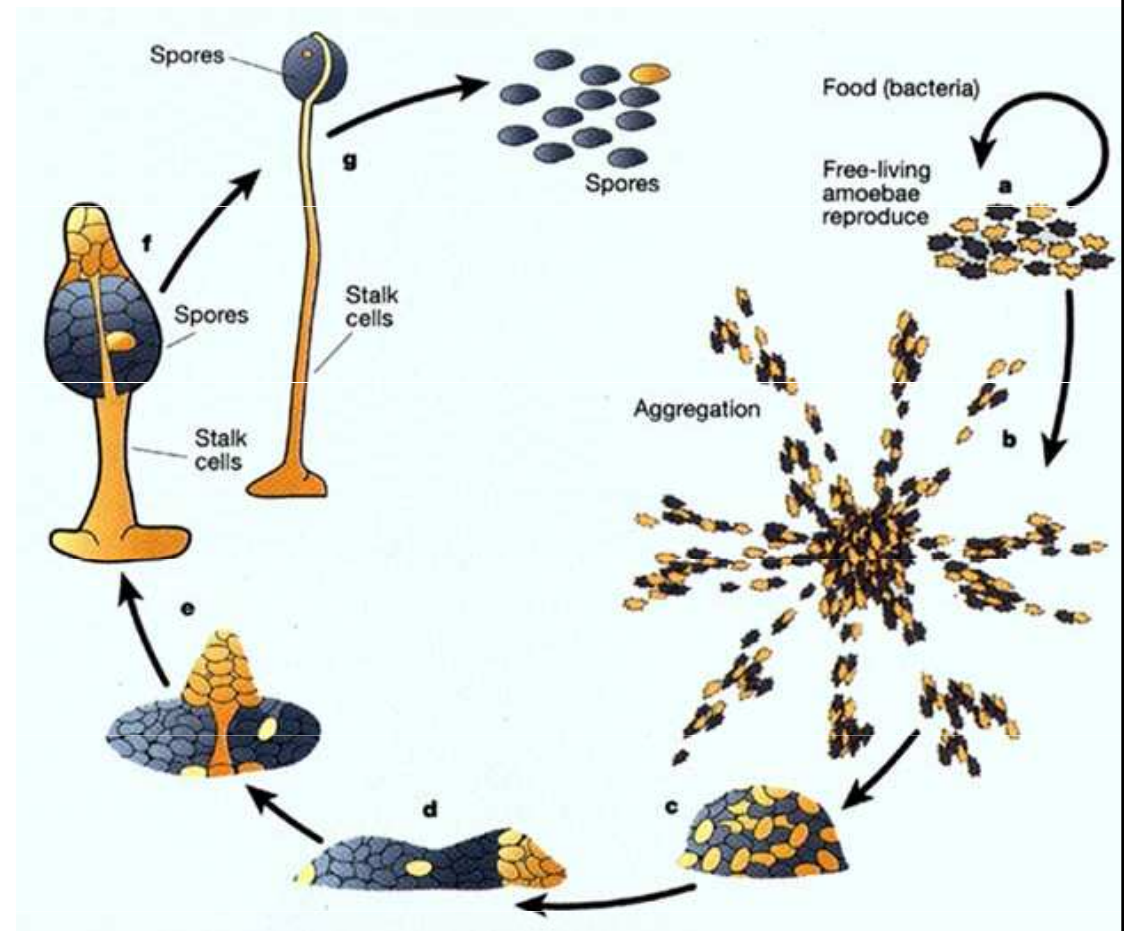


Darwinula stevensoni

Genetické chiméry

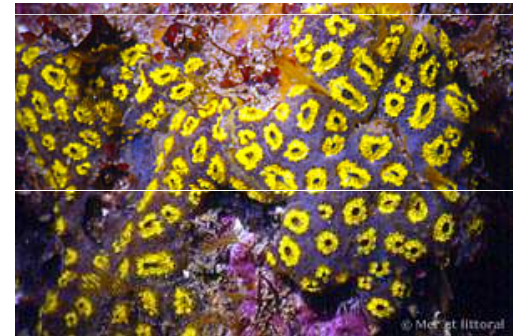


- organismy složené z buněk s různými genotypy
- *Dictyostelium discoideum*
chimérismus je pravidelná součást života



Genetické chiméry

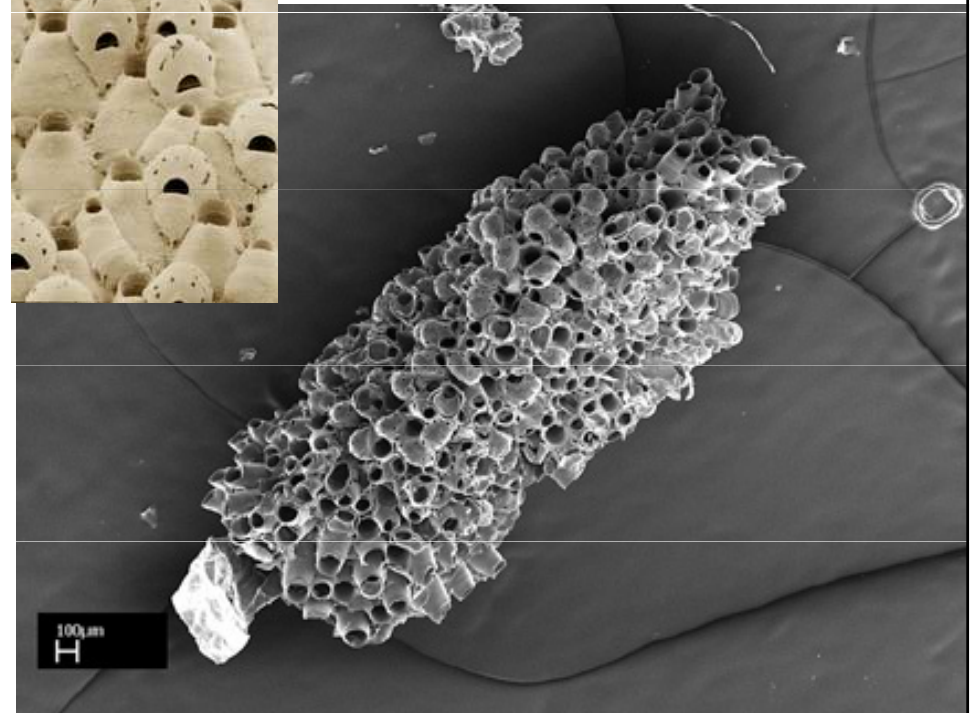
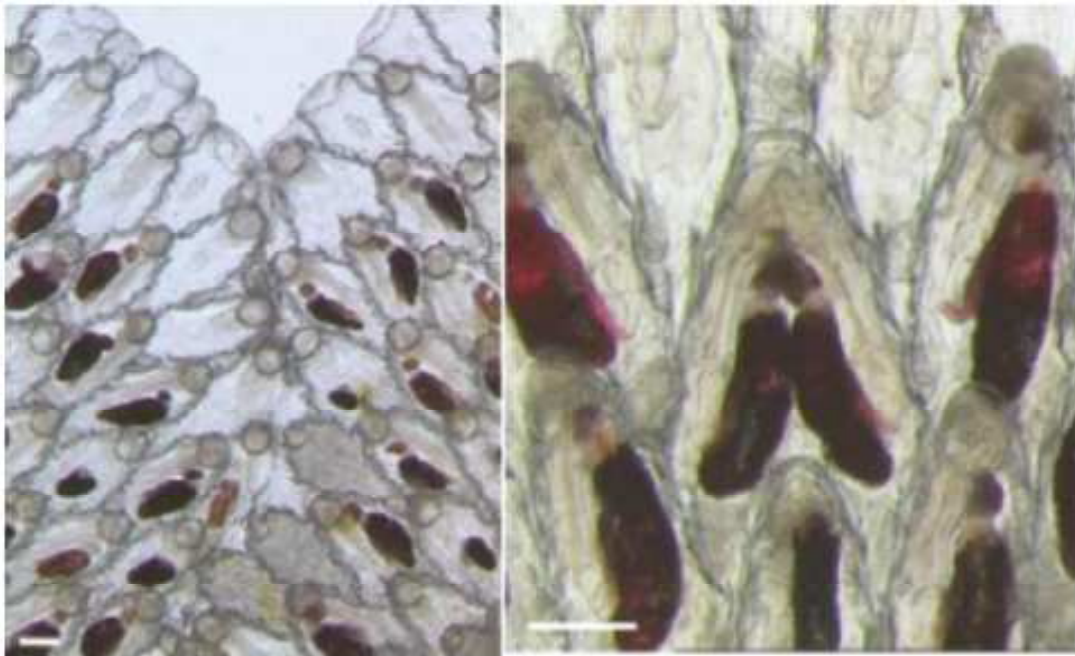
- *Ficus* srůst kořenů různých jedinců
- sumky *Botryllus schlosseri* chimérické kolonie příbuzní jedinci
- *Diplosoma listerianum* i nepřibuzní



Celleporella hyalina (Bryozoa)

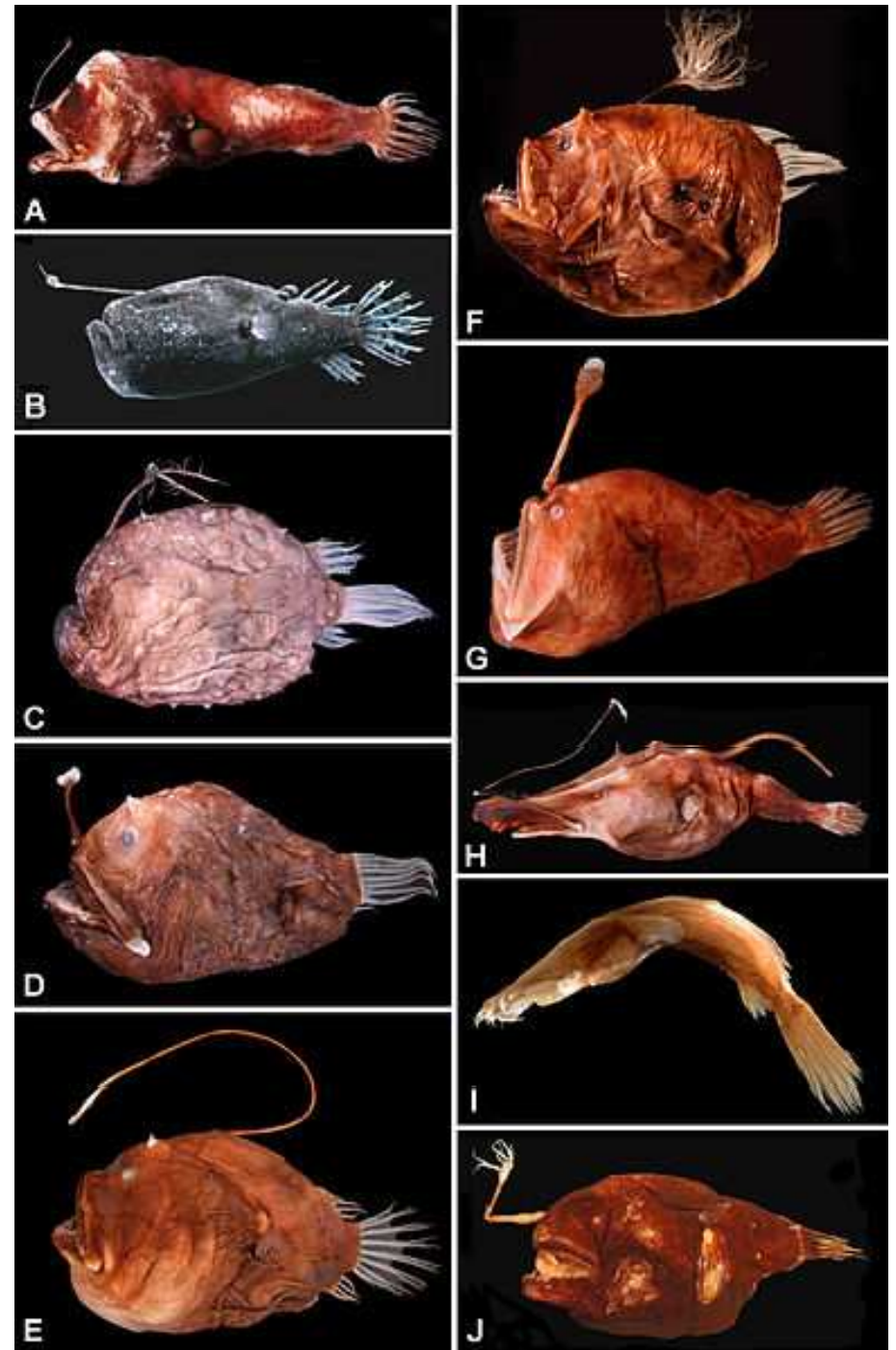
Hughes et al. 2004

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



Ceratioid anglerfish

- miniaturní samec po narození vyhledá samici, její kůže vyloučí hydrolytický enzym a samec přiroste
- vzniká hermafroditická chiméra



Genetické chiméry – „microchimerism“



- kosman bělovousý *Callithrix jacchus* (asi i rod *Saguinus*)
- dizygotická dvojčata
- DNA fingerprinting krve - 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 i u člověka
- (pohlaví dítěte před narozením lze určit i pomocí PCR sekvencí typických pro Chr Y, jako templát je periferní krev matky)

Identifikace pohlaví

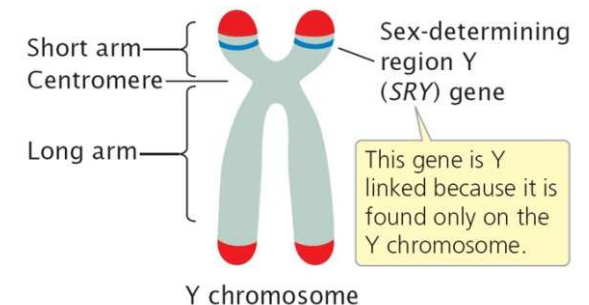
DNA sexing

Genetická identifikace pohlaví - proč

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

Genetická identifikace pohlaví

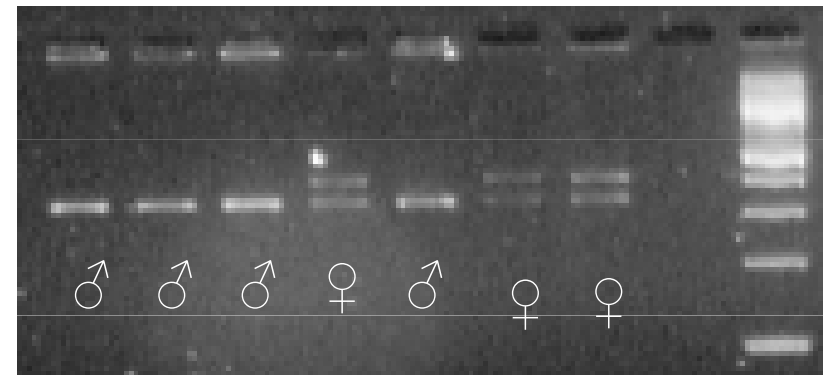
- druhy s genetickou determinací pohlaví (tj. nejčastěji pohlavní chromozómy)
- ptáci ($\text{♂} = ZZ$, $\text{♀} = ZW$)
- savci ($\text{♂} = XY$, $\text{♀} = XX$)
- amplifikace DNA oblasti specifické pro heterogametické pohlaví
- W, Y – malé chromozómy



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*

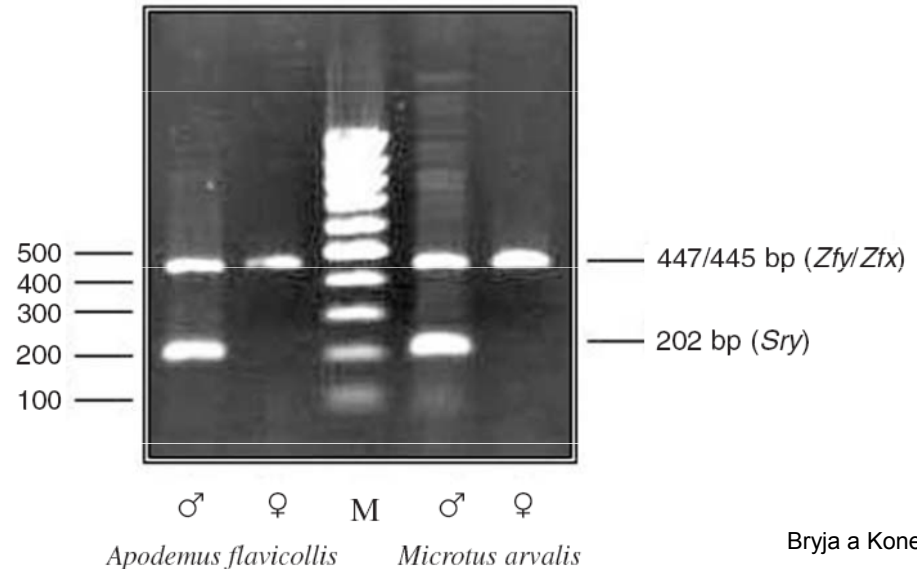
- Synové fungují jako pomocníci
- U adultů
2,31 samců na 1 samici
- Mláďata v hnízdě
poměr pohlaví 1:1 (57:57)
- První se líhnou samci
(v 17 hnízdech z 18)
Při opouštění hnízda jsou větší a těžší



medosavka hlučná

Určení pohlaví - savci

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



Bryja a Konečný 2003

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

M. cabrerae



Nannomys

Ellobius



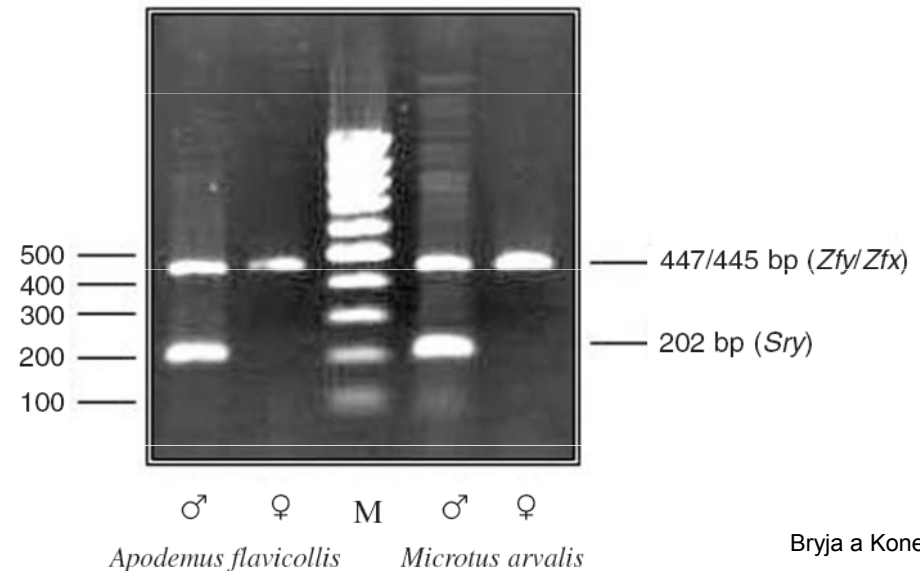
Tokudaia osimensis



© 奄美野生生物保護センター

Určení pohlaví - savci

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

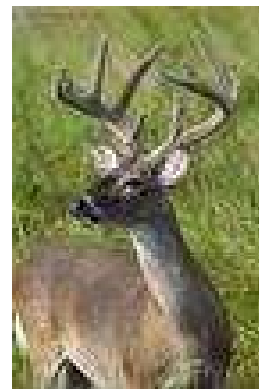


Bryja a Konečný 2003

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



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)

