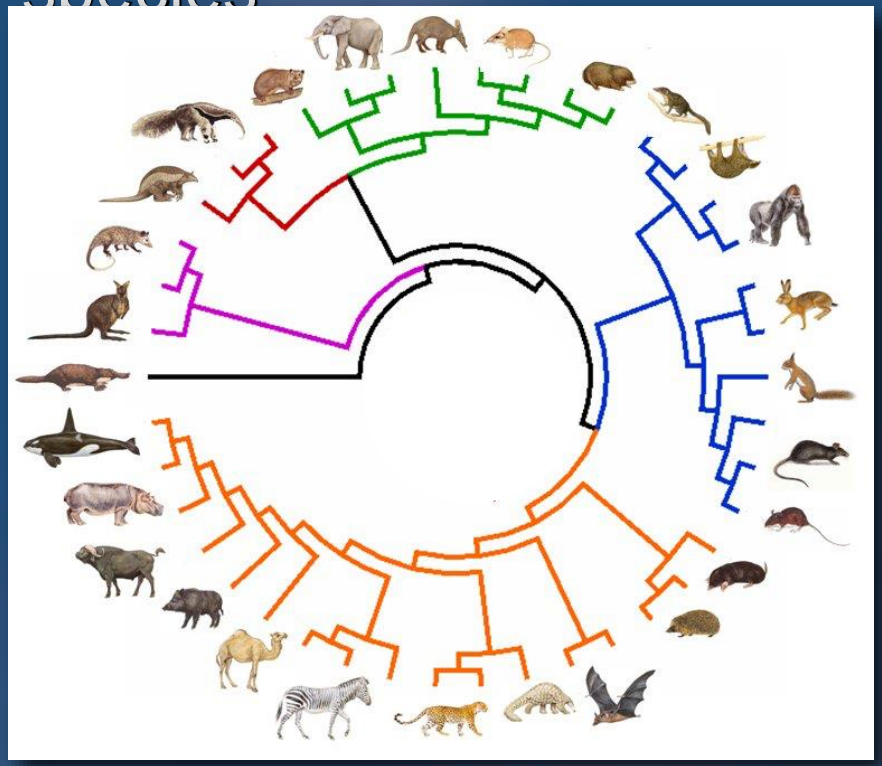
The use of mtDNA analysis to identify animal species

Promega / CURML Workshop
Lausanne - 24.11.2015

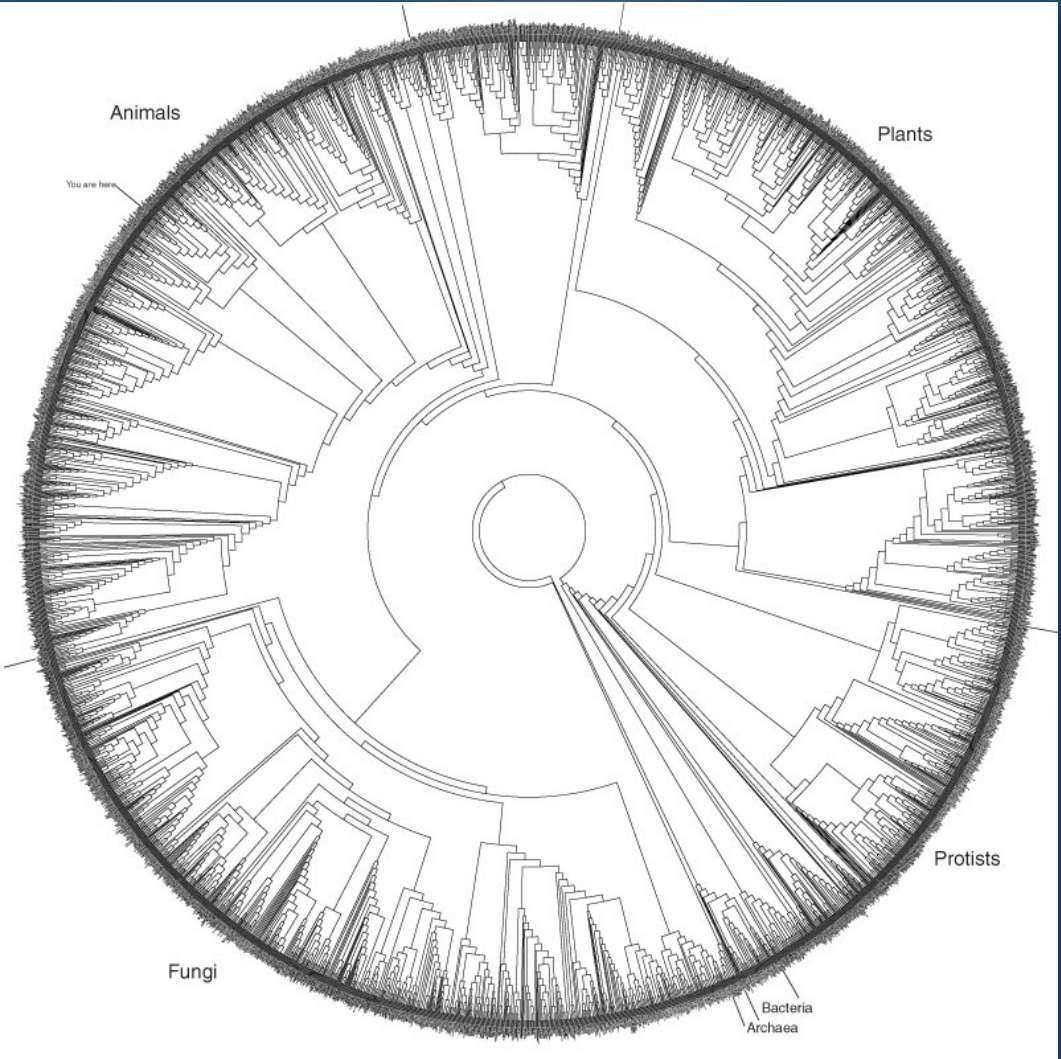


Dr Luca Fumagalli

Centre Universitaire Romand de Médecine Légale
Ch. de la Vulliette 4
CH-1000 Lausanne 25 (Suisse)

Laboratoire de Biologie de la Conservation
Département d'Ecologie et Evolution
Biophore, Université de Lausanne
CH-1015 Lausanne (Suisse)
www.unil.ch/lbc

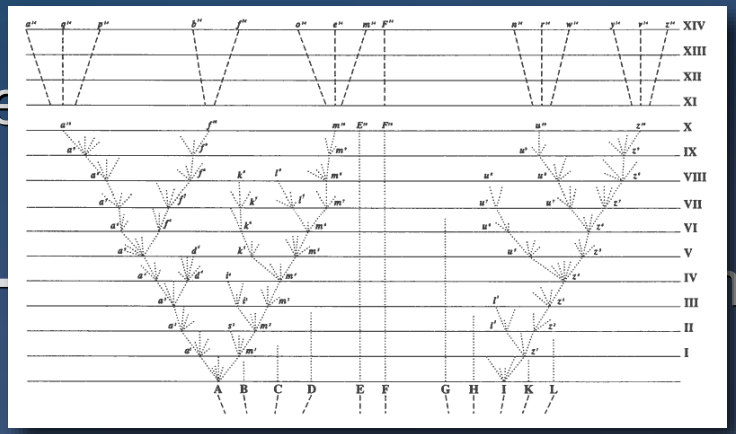
The use of mtDNA analysis to identify animal species



Introduction to molecular phylogeny

All life forms on earth share a common origin, and their ancestries can be traced back to one or a few organisms that lived approx. 4 billion years ago.

→ all animals, plants and bacteria are related by descent to e



→ ms are descended from more recent common ancestors than are distantly related ones

C. Darwin – *Origin of Species* (1859)
Illustration of 'Descent with modification'

Phylogenetic hypothesis

- Phylogenetic hypothesis

all living beings share a common ancestor, from which they diverged by accumulating DNA substitutions.

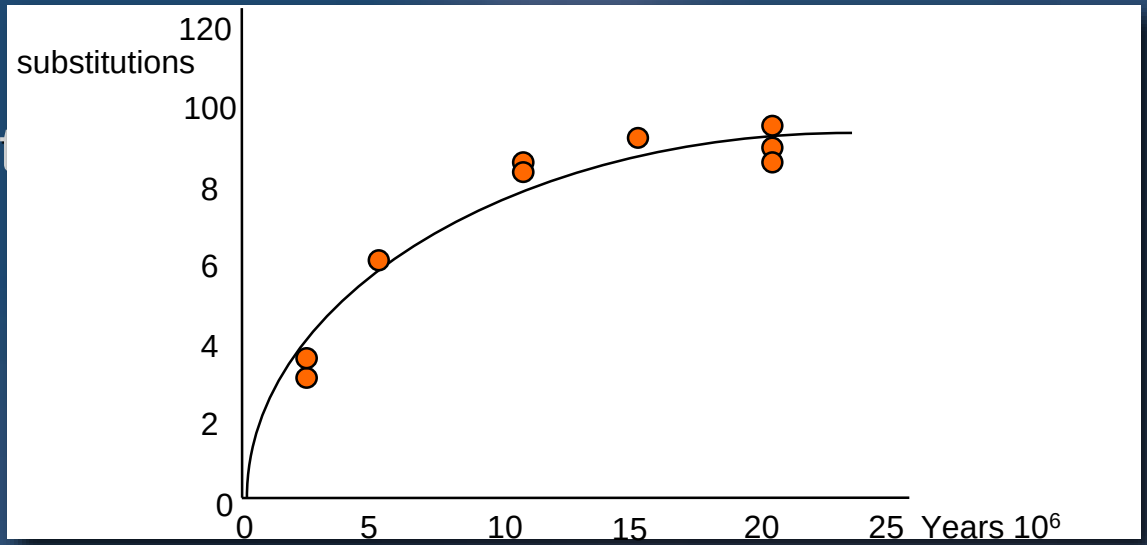
→ two homologous DNA sequences will be more similar if their common ancestor is recent than if the common ancestor

Phylogenetic hypothesis

Distance measures for nucleotide sequences

- *Observed differences*: count the nr. of nucleotide sites

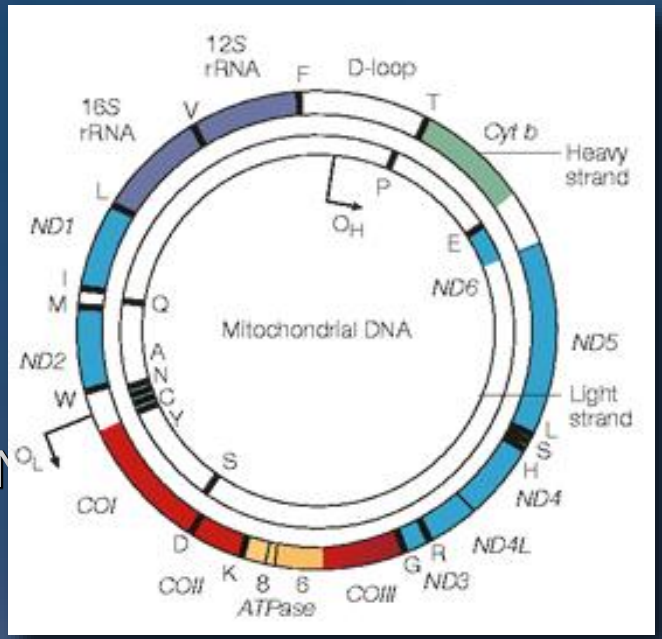
at which t



- Sequences separated by 20 Myr are no more different than sequences that diverged 15 Myr
- Relationship not linear, but deflected downwards due to multiple hits (most of the sites changing have already changed before)

Animal mtDNA

- Non-recombining
- Fast-evolving
- Uniparental inheritance
- High copy number
- Ubiquitous
- Simple structure
- 37 genes coding for 13 proteins, 22 tRNAs, 2 rRN
- Gene order ± conserved
- Small size [human: 16.5kb]
- Large database available



A milestone in evolutionary genetics, mtDNA-RFLP analysis inaugurated and
dominated the field of phylogeography in the 1980s.

Still widely employed, direct sequencing of individual genes, regular or chip-
based sequencing of whole mtDNA genomes (*mitogenomics*).

Animal mtDNA

Why using the mitochondrial genome for species ID?

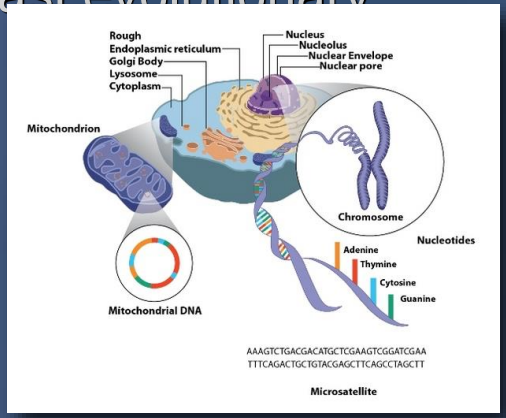
Rapid evolution: mtDNA is *polymorphic* at the intraspecific level, but less than at the interspecific level.

One quarter the effective population size (haploid and uniparental) of nuclear
loci means populations *diverge quickly* with respect to nDNA.

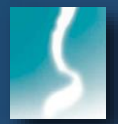
Non-recombining, *maternal* inheritance so haplotypes can be ordered phylogenetically into a *gene genealogy* interpretable as the matriarchal component of an *organismal pedigree* (good record of past evolutionary events).

Easy to amplify (many mitochondria per cell).

Large database available.



Animal mtDNA



Proc. Natl. Acad. Sci. USA
Vol. 86, pp. 6196–6200, August 1989
Evolution

Dynamics of mitochondrial DNA evolution in animals: Amplification and sequencing with conserved primers

(cytochrome *b*/12S ribosomal DNA/control region/evolutionary genetics/molecular phylogenies)

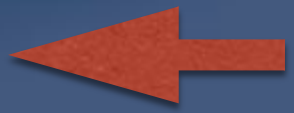
T. D. KOCHER*, W. K. THOMAS*, A. MEYER*†‡, S. V. EDWARDS*†‡, S. PÄÄBO*, F. X. VILLABLANCA†‡, AND A. C. WILSON*

Departments of *Biochemistry and †Zoology, and ‡Museum of Vertebrate Zoology, University of California, Berkeley, CA 94720

Universal mtDNA animal primers:

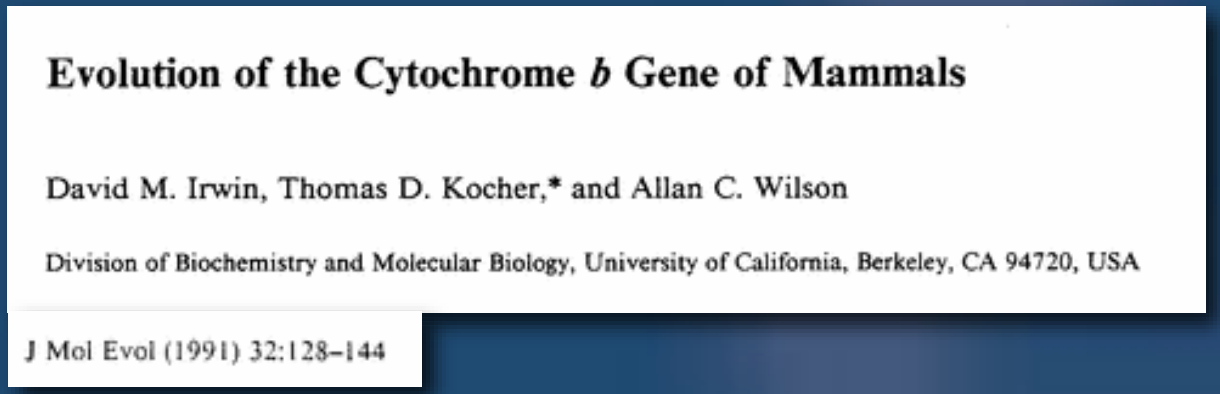
Cytochrome *b*, 12S rRNA and control region

Phylogenetics
Systematics
Population genetics
Forensics
(…)



Species ID

Animal mtDNA



Evolution of the Cytochrome *b* Gene of Mammals

David M. Irwin, Thomas D. Kocher,* and Allan C. Wilson

Division of Biochemistry and Molecular Biology, University of California, Berkeley, CA 94720, USA

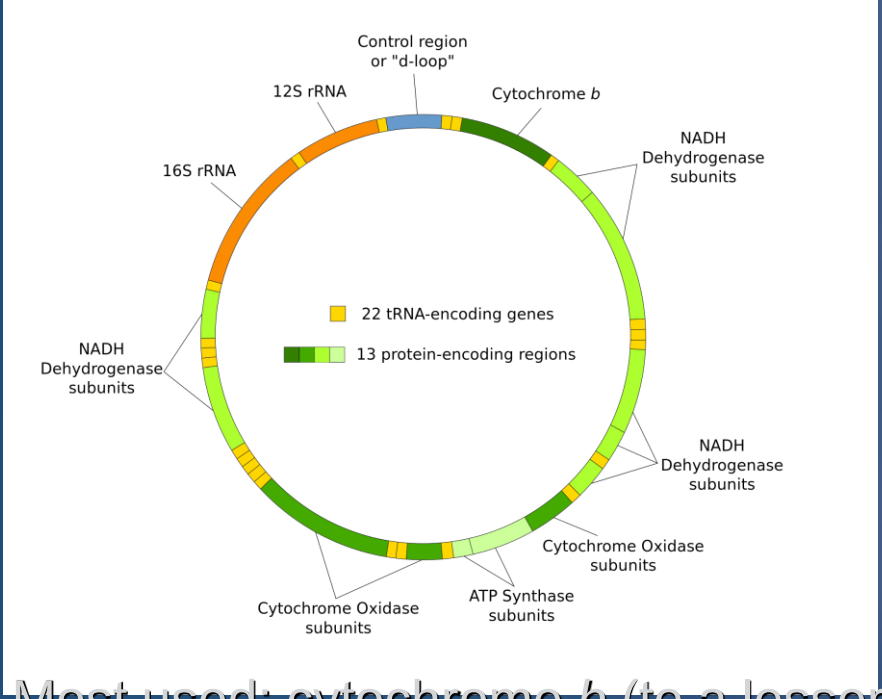
J Mol Evol (1991) 32:128–144

mtDNA genetic code

Substitution rates depend on:

- codon position: 1st, 2nd and 3rd positions ratio 10:1:35
- structure/function model: residues in transmembrane region evolve faster

Animal mtDNA

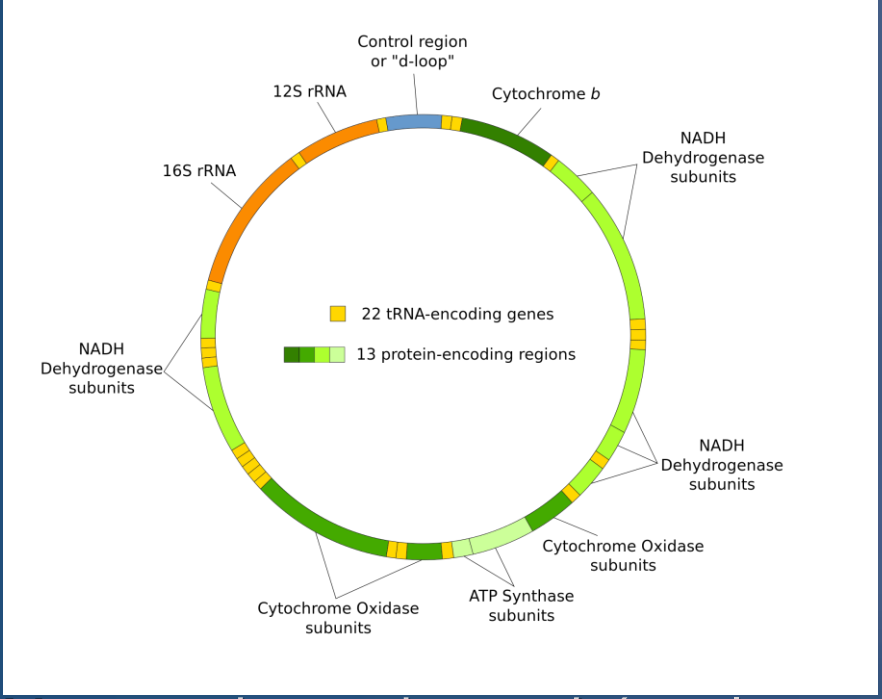


Most used: cytochrome *b* (to a lesser amount 12S and 16S rRNA, COI*)

higher degree of variability among than within species,
availability of universal primer pairs amplifying virtually any vertebrate spp,
substantial amount of DNA sequences deposited in DNA databases

**DNA barcoding*

Animal mtDNA



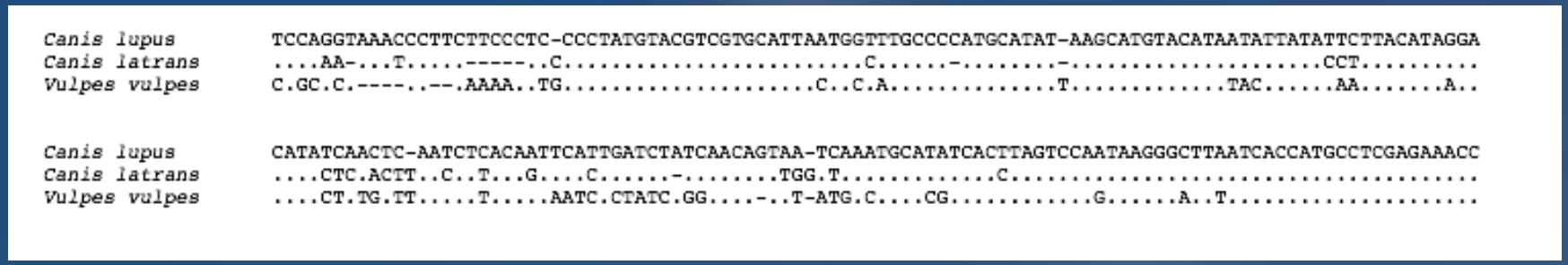
Most used: cytochrome *b* (to a lesser amount 12S and 16S rRNA, COI*)

RESTRICTION: limited variability as a coding region, therefore:

inability to differentiate closely related species;
identical within all vertebrate lineages, impeding the discrimination by direct DNA sequencing of different spp. in mixed samples when using a single universal primer pair

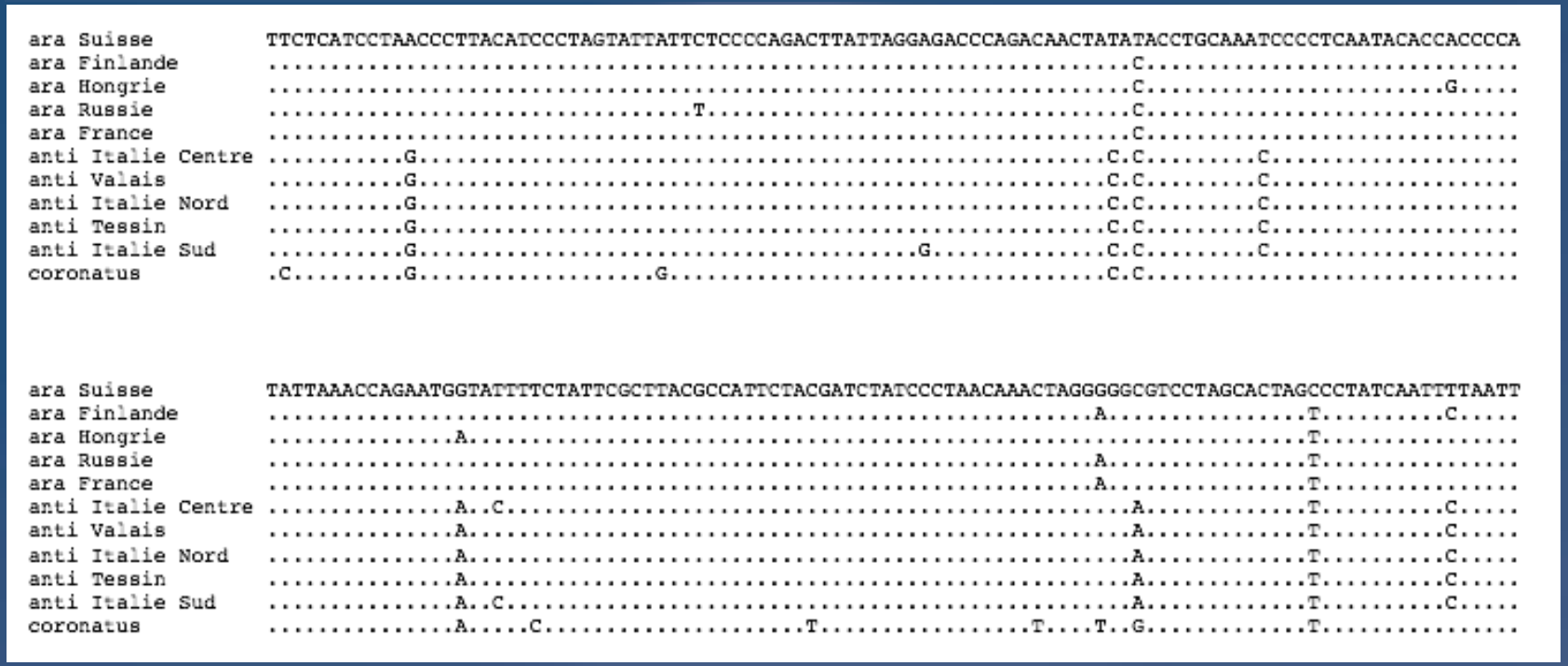
The use of mtDNA analysis to identify animal species

Alignment of 200 bp of the mtDNA control region of three canid species (*Canis lupus*: *C. latrans*: coyote; *Vulpes vulpes*: red fox



The use of mtDNA analysis to identify animal species

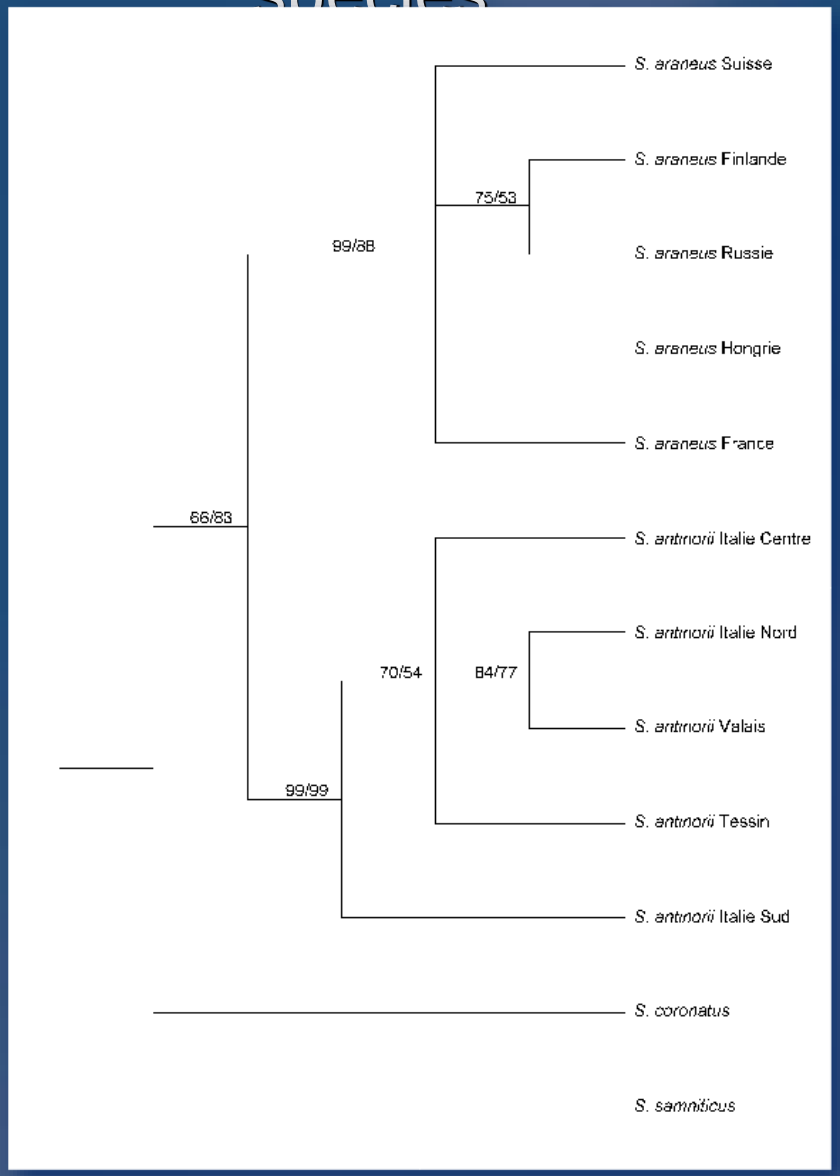
Alignment of 200 bp of the mtDNA cyt *b* gene in 11 specimens of shrews belonging to the genus *Sorex*



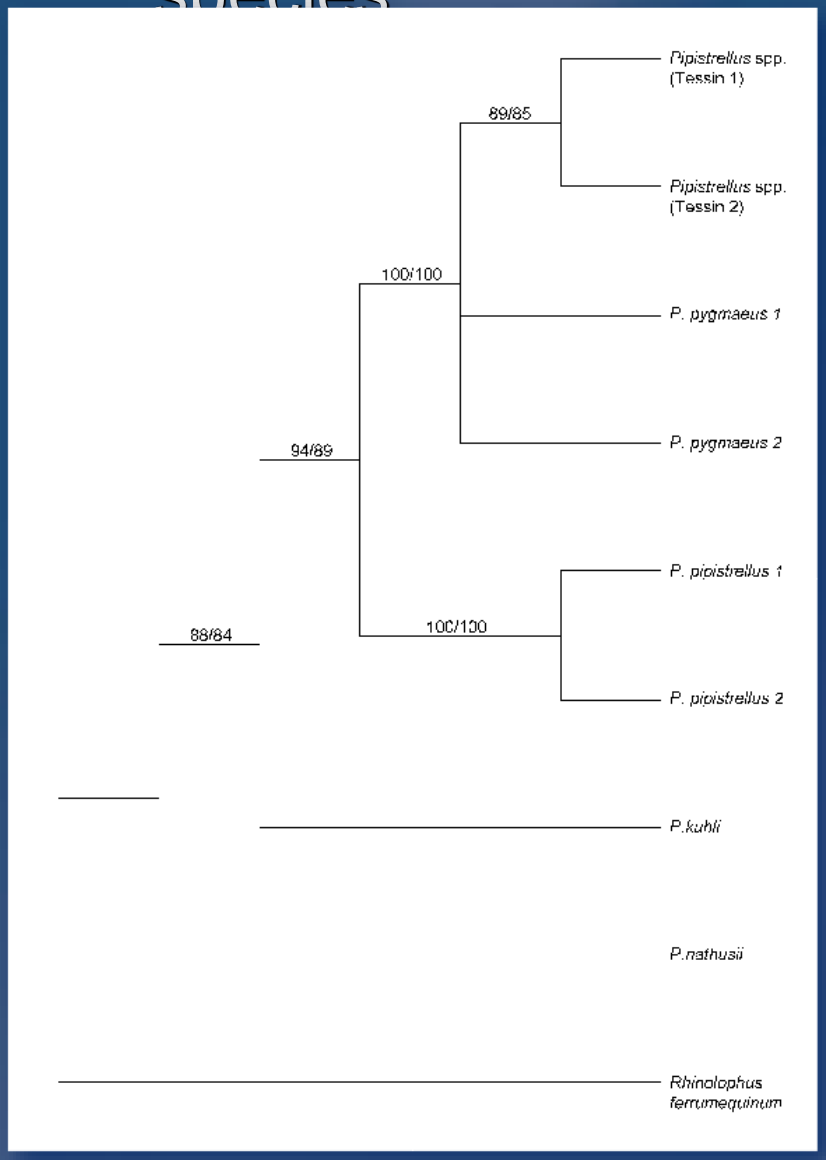
```
ara Suisse          TTCTCATCCTAACCCTTACATCCCTAGTATTATTCTCCCCAGACTTATTAGGAGACCCAGACAACTATATACCTGCAAATCCCCTCAATACACCACCCCA
ara Finlande        ..........................................................C.....................................
ara Hongrie         ..........................................................C.....................G.....
ara Russie          ................................T.........................C.....................................
ara France          ..........................................................C.....................................
anti Italie Centre ...........G..................................................C.C.........C.....................
anti Valais         ...........G..................................................C.C.........C.....................
anti Italie Nord    ...........G..................................................C.C.........C.....................
anti Tessin         ...........G..................................................C.C.........C.....................
anti Italie Sud     ...........G.....................................G............C.C.........C.....................
coronatus           .C.........G.....................G..............................C.C.....................................


ara Suisse          TATTAAACCAGAATGGTATTTTCTATTCGCTTACGCCATTCTACGATCTATCCCTAACAAACTAGGGGGCGTCCTAGCACTAGCCCTATCAATTTTAATT
ara Finlande        ..........................................................A.................T..........C.....
ara Hongrie         ...............A..........................................................T.....................
ara Russie          ..........................................................A.................T.....................
ara France          ..........................................................A.................T.....................
anti Italie Centre ...............A..C.....................................................A.............T..........C.....
anti Valais         ...............A..........................................................A.............T..........C.....
anti Italie Nord    ...............A..........................................................A.............T..........C.....
anti Tessin         ...............A..........................................................A.............T..........C.....
anti Italie Sud     ...............A..C.....................................................A.............T..........C.....
coronatus           ...............A.....C.....................T.................T....T..G.............T.....................
```

The use of mtDNA analysis to identify animal species



The use of mtDNA analysis to identify animal species

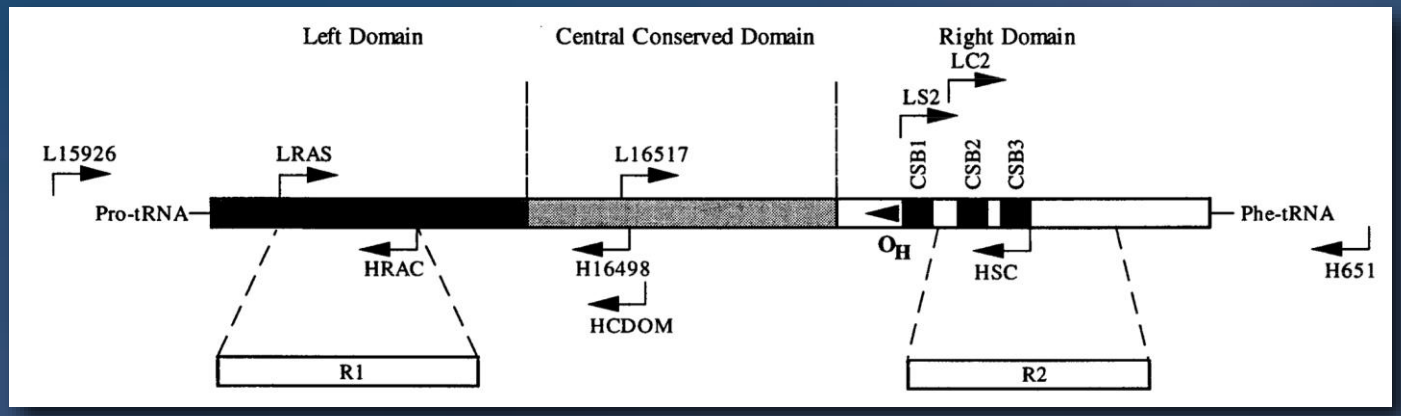


Animal mtDNA

A single non-coding region: control region

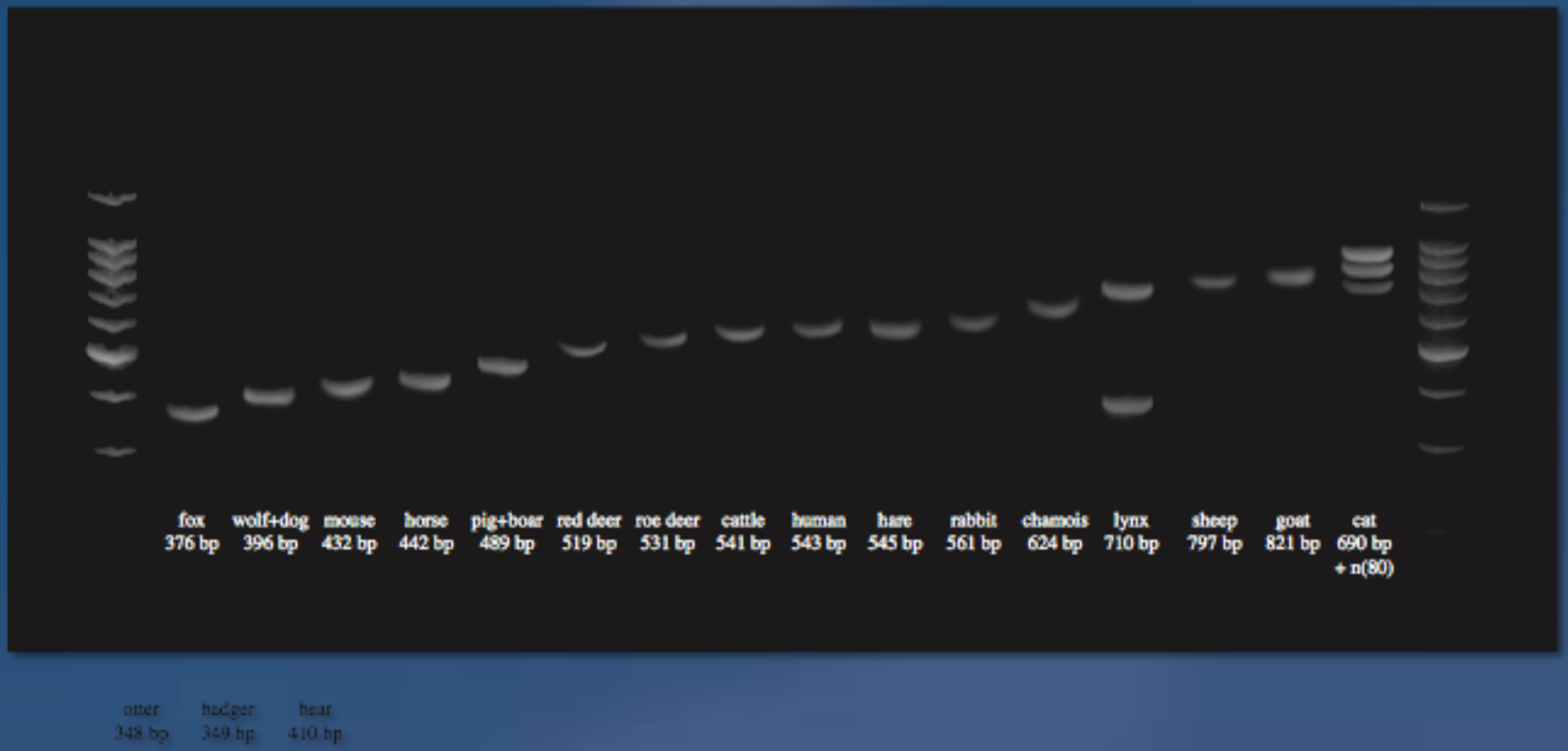
RAPID EVOLUTION, therefore more effective than cyt *b* and other mtDNA coding regions for the identification of closely related species

Presence of tandem repeats: sequence length polymorphism, useful for DNA mixtures



(Fumagalli et al. 1996)

- Species ID



(Pun *et al.* 2009)

• Species ID

Animal attacks on animal

- Dead pelicans in a zoo
- Predator from the zoo?

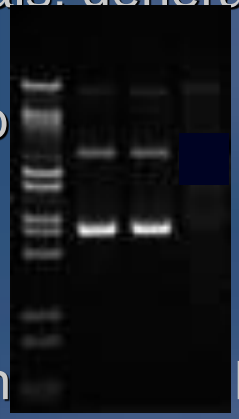
Dama dama
Canis familiaris

- swabs around wounds
- mtDNA control region: 2 amplicons of 396 bp et de 518 bp, ID by sequencing
- Dog probable predator, fallow deer contamination during storing

• Species ID

Animal attacks on animal

- 40+ cases in summer 2005 in NW Switzerland
- Mutilations on domestic animals, generally found dead
- cows, horses, donkeys, sheep
- sexual organs, face, tail, ears
- alleged human mutilator, but n DNA found
  - Swabs around two dead cows
  - mtDNA control region: 2 amplicons of 541 bp et 376 bp, ID by sequencing
  - Probable *post-mortem* interference



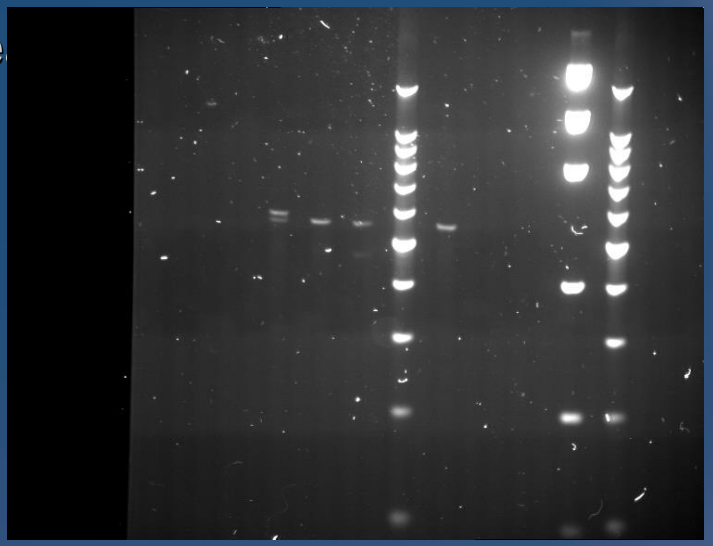
Bos taurus

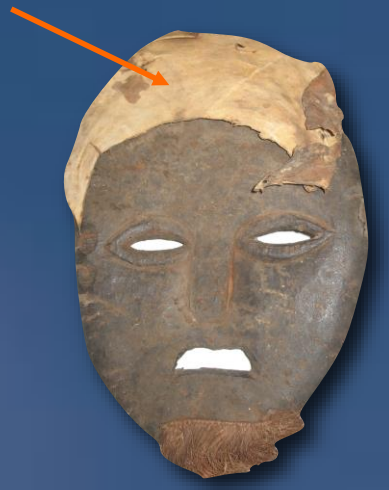
Vulpes vulpes

• Species ID

Expertise for exportation permit

- Ancient african mask (ca. 100 years), value ca. 25’000 $
- Le





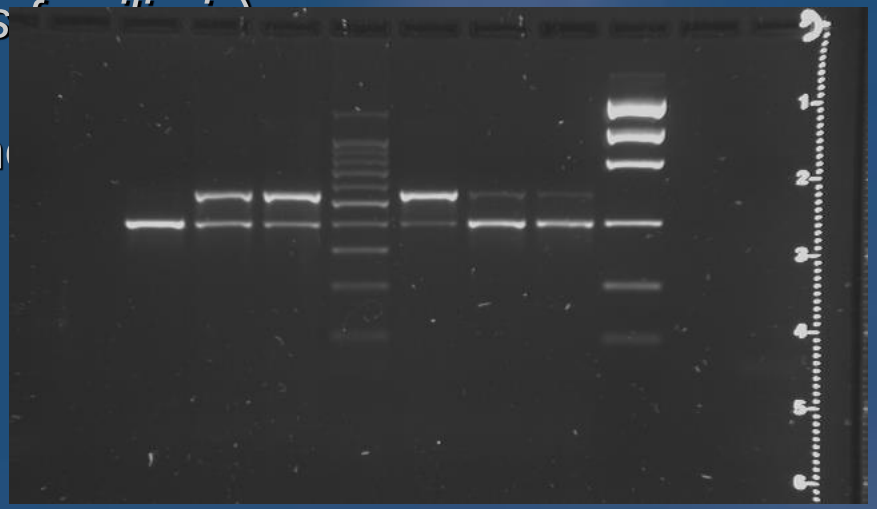
- mtDNA control region: amplicons of. 550 bp, ID by sequencing
- *Genetta maculata* ou *G. tigrina*

• Species ID

Animal attack on human

• dog bite *(Canis*

• swabs on cloth



Homo sapiens
Canis familiaris

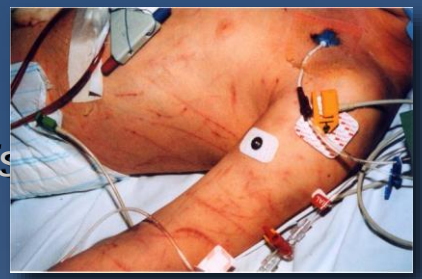
• mtDNA control region: 2 amplicons of 543 bp et 396 bp, ID by sequencing
• DNA profiling: 8 STR loci
• Two reference dogs

→ 6 samples same DNA profile, identical to one of the references,
→ allelic *f* in a reference population of 63 dogs of various breeds, correction factor theta of 3%
→ $LR > 1 \times 10^9$

• Species ID

Alleged human attack on human

• Child laying unconscious in the snow, most of his clothes removed or torn, and several wounds on his body. Presence of the family dog (German sheperd).



• No human DNA on child except it's own

• 58 samplings on 7 samples (our sampling), 8 STR loci

• Saliva from the family dog (reference)

→ DNA profiles 7 samples identical, and identical to reference

→ allelic *f* in a reference population of 95 dogs of various breeds,

→ $P_{(ID)random} < 8/10^{13}$

→ $P_{(ID)sib} < 1.9/10^{4}$

• Wounds compatible with dog attack (controversial)

• mtDNA control region: 30 bands of expected size for *Canis familiaris*

→ Confirmation *C. familiaris* by DNA sequencing on 13 bands

• 58 samplings on 7 samples (our sampling), 8 STR loci

• Saliva from the family dog (reference)

→ DNA profiles 7 samples identical, and identical to reference

→ allelic *f* in a reference population of 95 dogs of various breeds,

→ $P_{(ID)random} < 8/10^{13}$

→ $P_{(ID)sib} < 1.9/10^{4}$

• 58 samplings on 7 samples (our sampling), 8 STR loci (body and clothes), and by us 2 years later (clothes)

• Species ID

Accident caused by animals

- bird strike
- swabs over spots on engine

- mtDNA cyt *b* bird specific universal primers

→ 3 swabs, species identified: Common kestrel (*Falco tinnunculus*)

- Species ID
  - Illegal animal traffic (poaching)

roe deer (*Capreolus capreolus*)

red deer (*Cervus elaphus*)

wild boar (*Sus scrofa*)

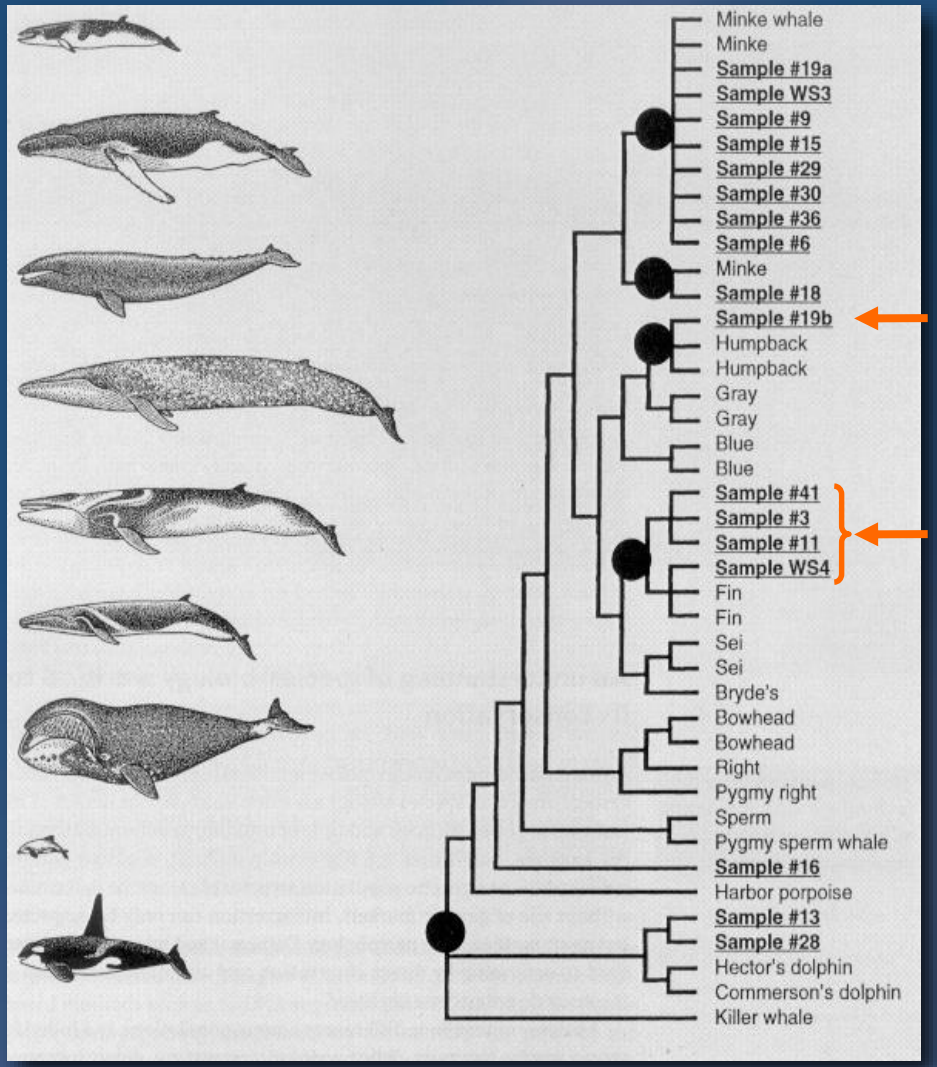
wolf (*Canis lupus*)

lynx (*Lynx lynx*)

(…)

→ mtDNA: Non-invasive genetic sampling and aDNA

- Species ID



(after Baker & Palumbi 1996)

- Species ID

DNA microarrays

DNA mixtures and multiple species

Journal of Applied Ecology 2008, **45**, 967–975 doi: 10.1111/j.1365-2664.2007.01415.x

Molecular identification of vertebrate species by oligonucleotide microarray in food and forensic samples

Fabrice Teletchea[1]†, Jacques Bernillon[2], Marilyne Duffraisse[1,3], Vincent Laudet[3] and Catherine Hänni[1,3]*

mtDNA cyt *b* gene
373 bp (268+124 bp)
71/77 vertebrate spp. unambig. identified (mammals, birds, fish)

OPEN ACCESS Freely available online PLoS one

Identifying Fishes through DNA Barcodes and Microarrays

Marc Kochzius[1,2]*, Christian Seidel[1¤a], Aglaia Antoniou[3], Sandeep Kumar Botla[1¤b], Daniel Campo[4¤c], Alessia Cariani[5], Eva Garcia Vazquez[4], Janet Hauschild[1¤d], Caroline Hervet[6¤e], Sigridur Hjörleifsdottir[7], Gudmundur Hreggvidsson[7,8], Kristina Kappel[1], Monica Landi[5¤f], Antonios Magoulas[3], Viggo Marteinsson[7], Manfred Nölte[9], Serge Planes[6¤g], Fausto Tinti[5], Cemal Turan[10], Moleyur N. Venugopal[11], Hannes Weber[1], Dietmar Blohm[1]

mtDNA cyt *b*, 16S, COI
30 out of 50 fish spp. identified

RESTRICTIONS: development time, costs, no flexibility (species

etermined)

- Species ID

Next generation sequencing

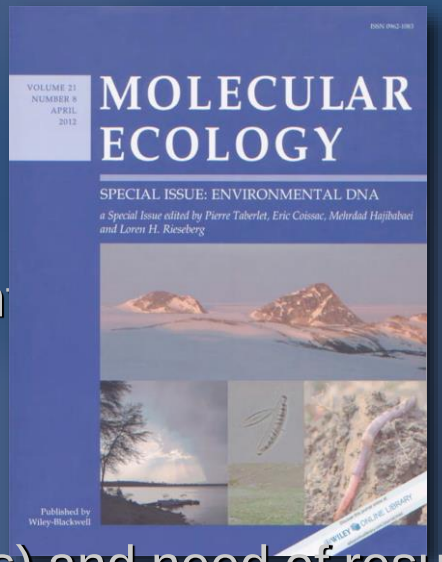
A Universal Method for Species Identification of Mammals Utilizing Next Generation Sequencing for the Analysis of DNA Mixtures

Andreas O. Tillmar, Barbara Dell'Amico, Jenny Welander, Gunilla Holmlund

cies

2013. PLoS ONE 8(12).

eDNA



Biodiversity assessmen

Diet analysis

Ecology

(…)

RESTRICTIONS: costs (few samples) and need of results on a week-by-

k basis

- Species ID

Next generation sequencing

Population assignment (SNPs)

ARTICLE

Received 23 Dec 2011 | Accepted 12 Apr 2012 | Published 22 May 2012

DOI:10.1038/ncomms1845

Gene-associated markers provide tools for tackling illegal fishing and false eco-certification

Einar E. Nielsen[1], Alessia Cariani[2,3], Eoin Mac Aoidh[4], Gregory E. Maes[3], Ilaria Milano[2,5], Rob Ogden[6], Martin Taylor[7], Jakob Hemmer-Hansen[1], Massimiliano Babbucci[5], Luca Bargelloni[5], Dorte Bekkevold[1], Eveline Diopere[3], Leonie Grenfell[6], Sarah Helyar[8], Morten T. Limborg[1], Jann T. Martinsohn[4], Ross McEwing[6], Frank Panitz[9], Tomaso Patarnello[5], Fausto Tinti[2], Jeroen K.J. Van Houdt[3], Filip A.M. Volckaert[3], Robin S. Waples[10], FishPopTrace Consortium* & Gary R. Carvalho[7]

• Difficulty and constraints

- Very high number of different species

Development of methods

- no commercial kits

- research and development only exclusively performed in « non-forensic » research laboratories

- From the university to the forensic laboratory

Extent of techniques and species

Research and development

Restricted budgets

Very variable caseworks

Intermediate solutions

Flexibility and adaptation

Set up of quality systems

Avoid prohibitive analysis costs