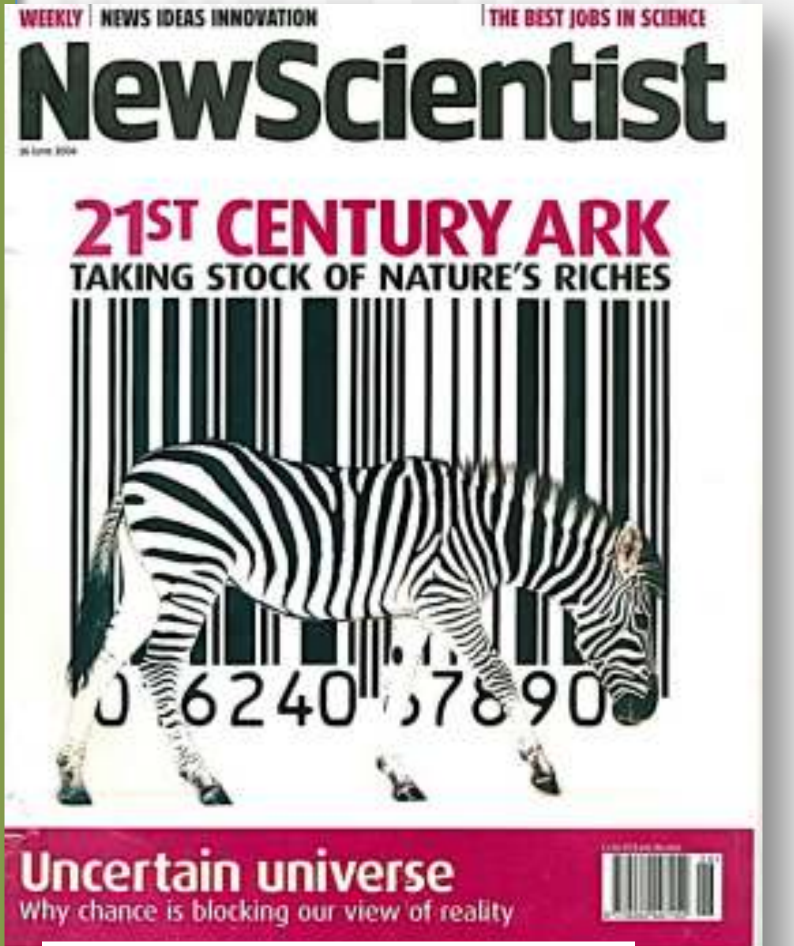
DNA Barcoding


WEEKLY | NEWS IDEAS INNOVATION | THE BEST JOBS IN SCIENCE

NewScientist

21ST CENTURY ARK

TAKING STOCK OF NATURE'S RICHES

Uncertain universe

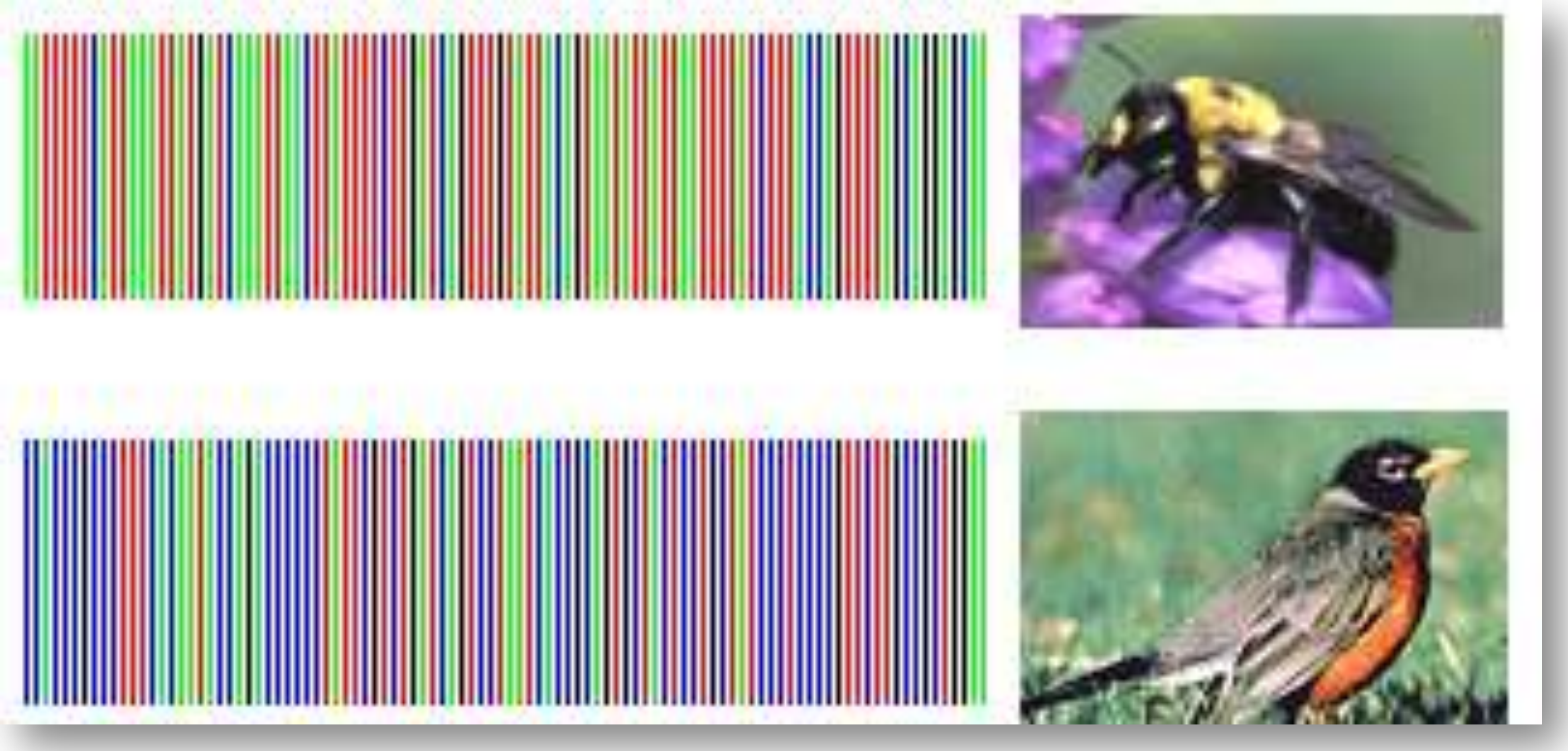
Why chance is blocking our view of reality

New Scientist, 26 June, 2004

A new diagnostic tool for rapid species recognition identification, and discovery

DNA Barcoding

Approccio **standardizzato** per **l'identificazione di specie** animali e vegetali tramite **corte sequenze di DNA**, chiamate DNA barcode.



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

DNA Barcoding

Cos'è?

- ✓ Nasce dall'idea di trovare una singola regione del DNA utile per l'identificazione di tutti i taxa
- ✓ Marcatore **universale** per l'identificazione di specie e di individui all'interno della stessa specie (**CODICE A BARRE**)
- ✓ Una raccolta di dati (*reference databases*) fruibili anche da non-specialisti
- ✓ Un salto nel passato (possibilità di analizzare campioni museali ed erbari)

Il codice a barre deve riconoscere le specie sulla base del rapporto

Variabilità interspecie / **Variabilità intraspecie**

DNA Barcoding

Cos'è?

- ✓ Nasce dall'idea di trovare una singola regione del DNA utile per l'identificazione di tutti i taxa
- ✓ Marcatore **universale** per l'identificazione di specie e di individui all'interno della stessa specie (**CODICE A BARRE**)
- ✓ Una raccolta di dati (reference databases) fruibili anche da non-specialisti
- ✓ Un salto nel passato (possibilità di analizzare campioni museali ed erbari)

Cosa NON è?

- ✓ Non sostituisce le altre discipline (morfologia) né le altre tecniche/marcatori
- ✓ **Universale**

DNA Barcoding

ELSEVIER

Update

TRENDS in Parasitology Vol.19 No.12 December 2003

545

Research Focus

DNA barcoding of parasites and invertebrate disease vectors: what you don't know can hurt you

Nora J. Besansky, David W. Severson and Michael T. Ferdig

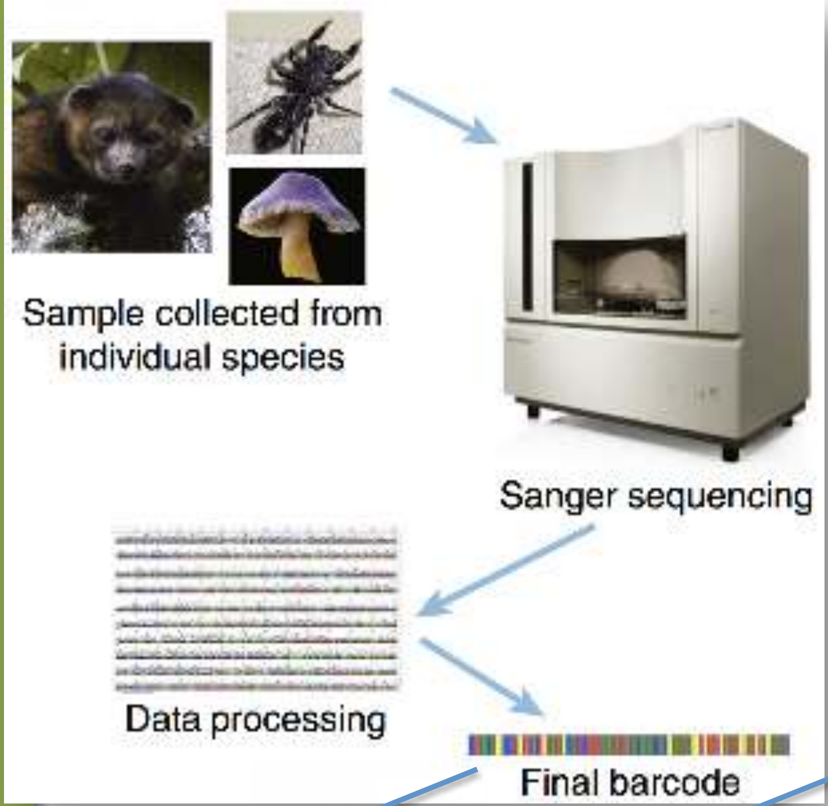
Center for Tropical Disease Research and Training, Department of Biological Sciences, University of Notre Dame, Notre Dame, IN 46556, USA

"the role of any molecular diagnostic is to aid research, not to serve as an end in itself. Barcoding … is independent of questions as to whether individual taxa are species, what species are (or should be), and where they fit in a unified tree of life…. Barcoding is not an end in itself, but will boost the rate of discovery. The unique contribution of DNA barcoding to … taxonomy and systematics is a compressed timeline for the exploration and analysis of biodiversity."

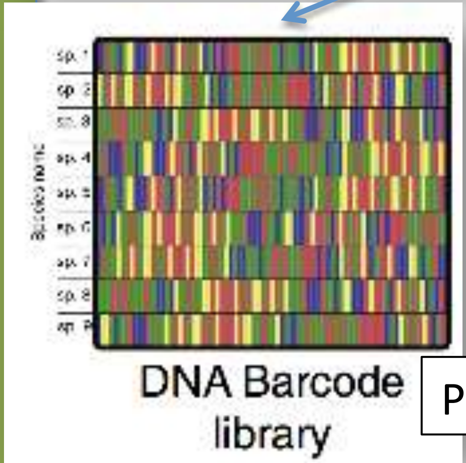
Ma di cosa necessita il progetto DNA Barcode?

In pratica …

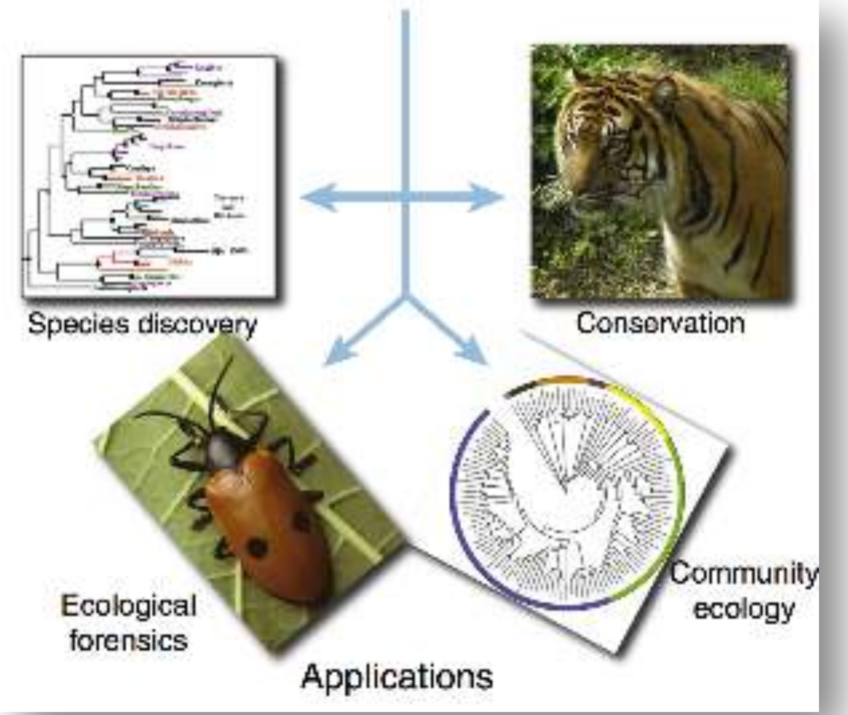
Basic Workflow to generate DNA barcodes



- ✓ Sequenziamento veloce **Next Generation Sequencing (NGS)**
- ✓ Creazione di banche dati **BOLD**
- ✓ Tecniche informatiche per le procedure di comparazione



Public databases



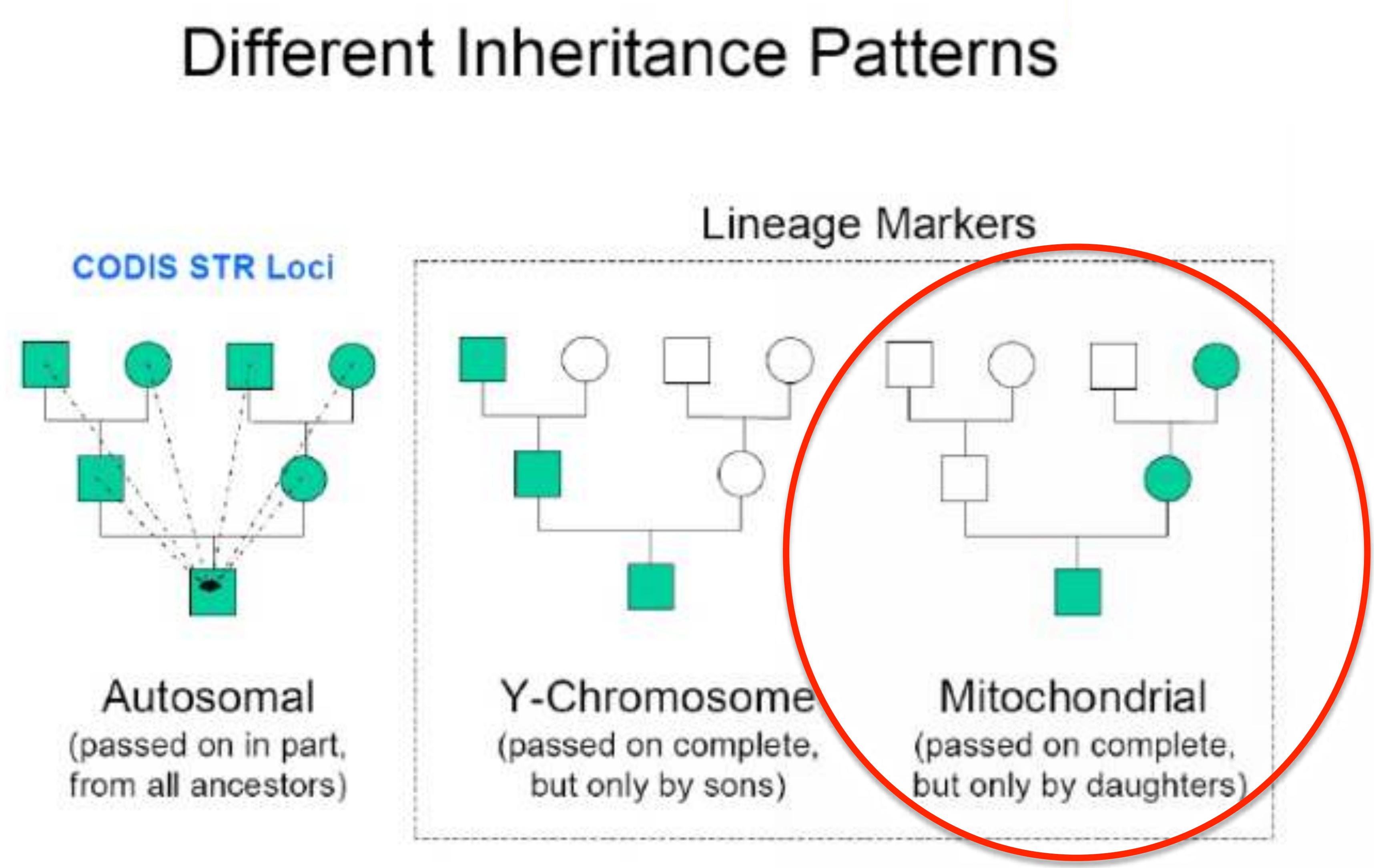
DNA Barcoding

A selected and standardized **set of BI-PARENTALLY INHERITED autosomal markers**, so called MICROSATELLITES, has become **the GOLDEN STANDARD in human identity testing**, including Forensic DNA analysis.

“DNA profiles” in criminal databases provide powerful tools for more effective law enforcement

But . . .

DNA Barcoding



Different Inheritance Patterns

Lineage Markers

CODIS STR Loci

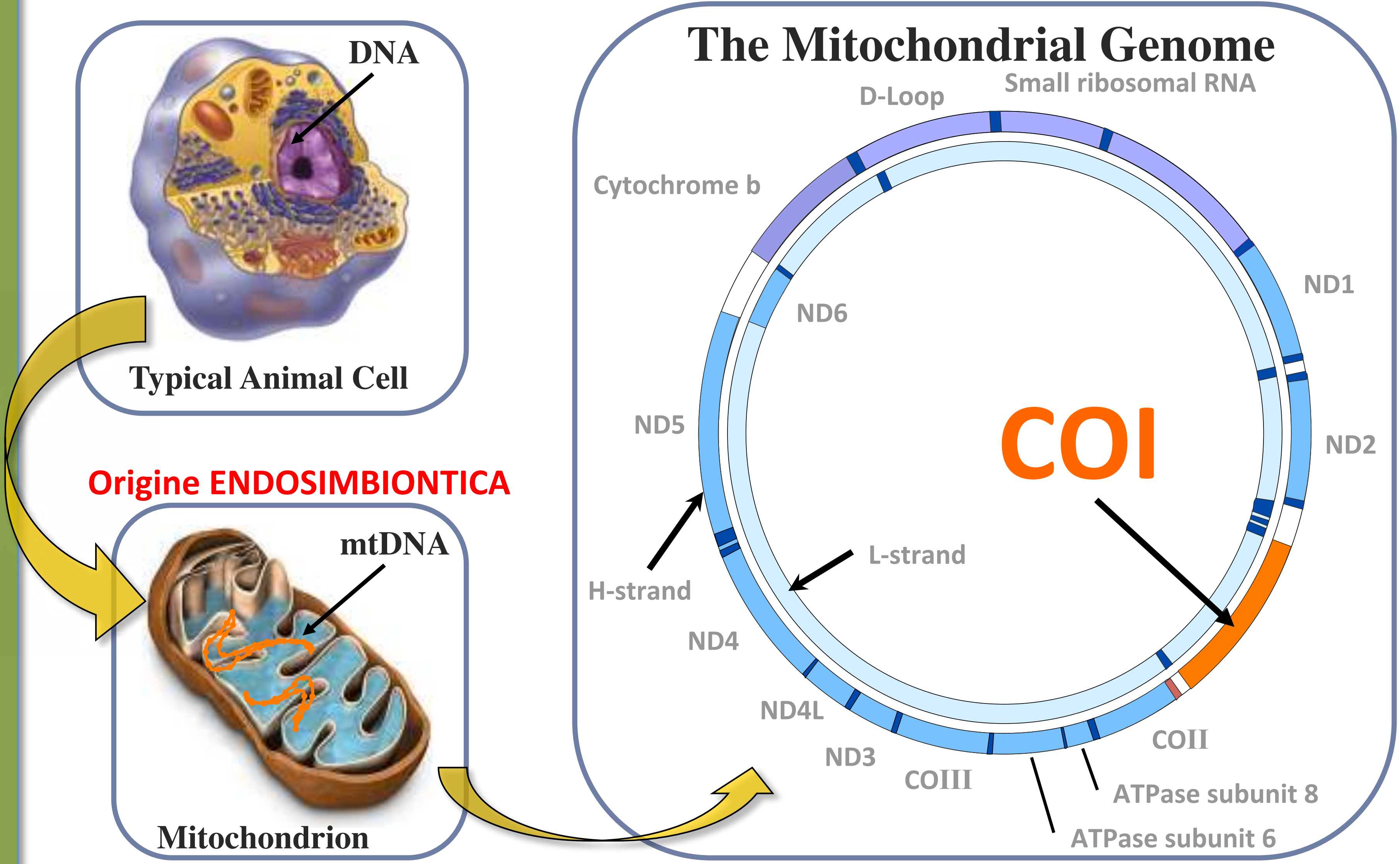
Autosomal
(passed on in part, from all ancestors)

Y-Chromosome
(passed on complete, but only by sons)

Mitochondrial
(passed on complete, but only by daughters)

Sex-specific inherited markers

An Internal ID System for All Animals



Typical Animal Cell

Origine ENDOSIMBIONTICA

Mitochondrion

The Mitochondrial Genome

Mitochondrial genome as barcode – Why?

- ✓ Piccolo (uomo 16.569 bp), circolare e semplice
- ✓ Aploide e NON-ricombinante
- ✓ Multi-copy

Per cellula ci sono 1.000-10.000 copie in più di mtDNA rispetto a nDNA →

- Facilità di estrazione da piccole porzioni di tessuto
- Facilita di amplificazione da DNA degradato

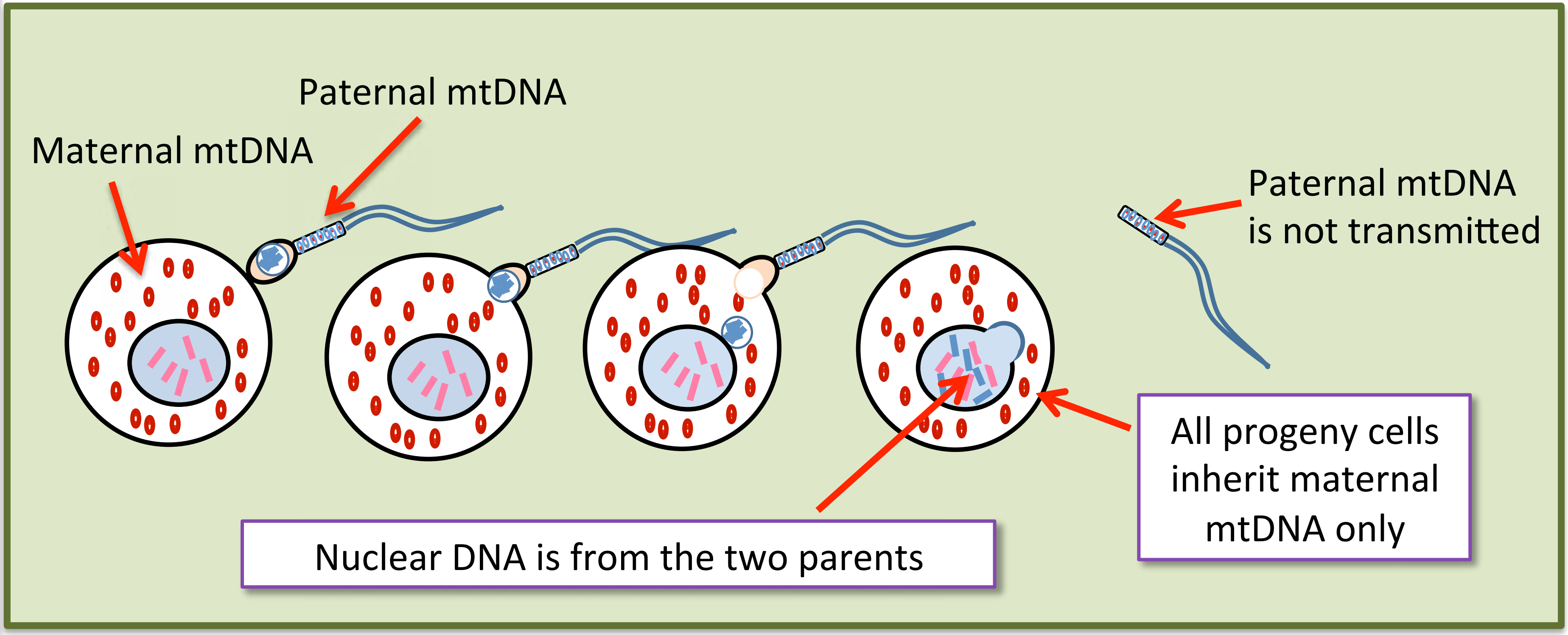
- ✓ Compatto – Assenza di INTRONI

Il 93%-95% del mtDNA è codificante

I geni nucleari sono spesso interrotti da lunghe sequenze non codificanti che complicano l'amplificazione

Mitochondrial genome as barcode – Why?

- ✓ Ereditato per via **UNIPARENTALE**



Paternal mtDNA

Maternal mtDNA

Paternal mtDNA is not transmitted

All progeny cells inherit maternal mtDNA only

Nuclear DNA is from the two parents

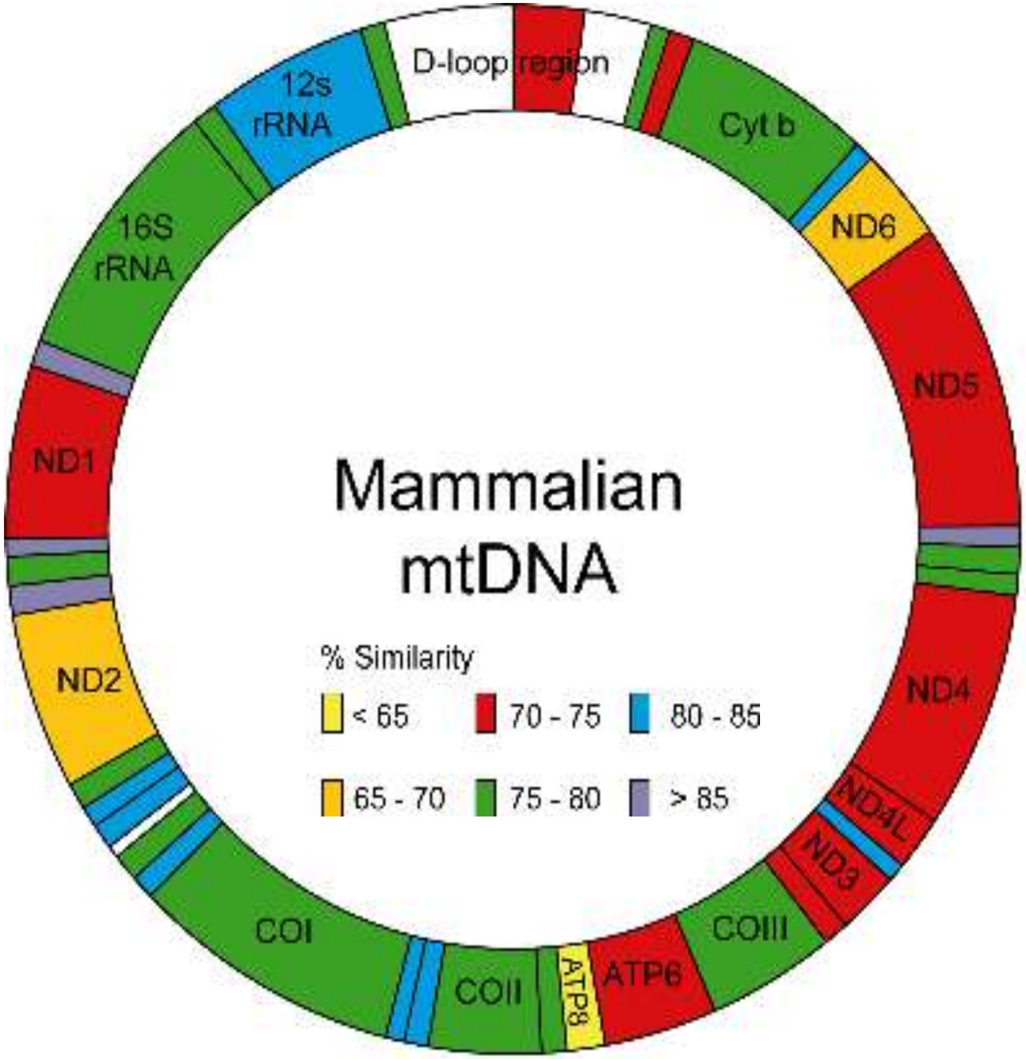
Le sue differenze dipendono dall'accumulo di mutazioni e non da ricombinazione

Mitochondrial genome as barcode – Why?

✓ **ETEROGENEO**

Some regions are very **CONSERVED** and show a high similarity across mammal species but also across all Vertebrates and other phyla.

Other are **HYPERVARIABLE** (Es. dLoop Control Region)



Mitochondrial genome as barcode – Why?

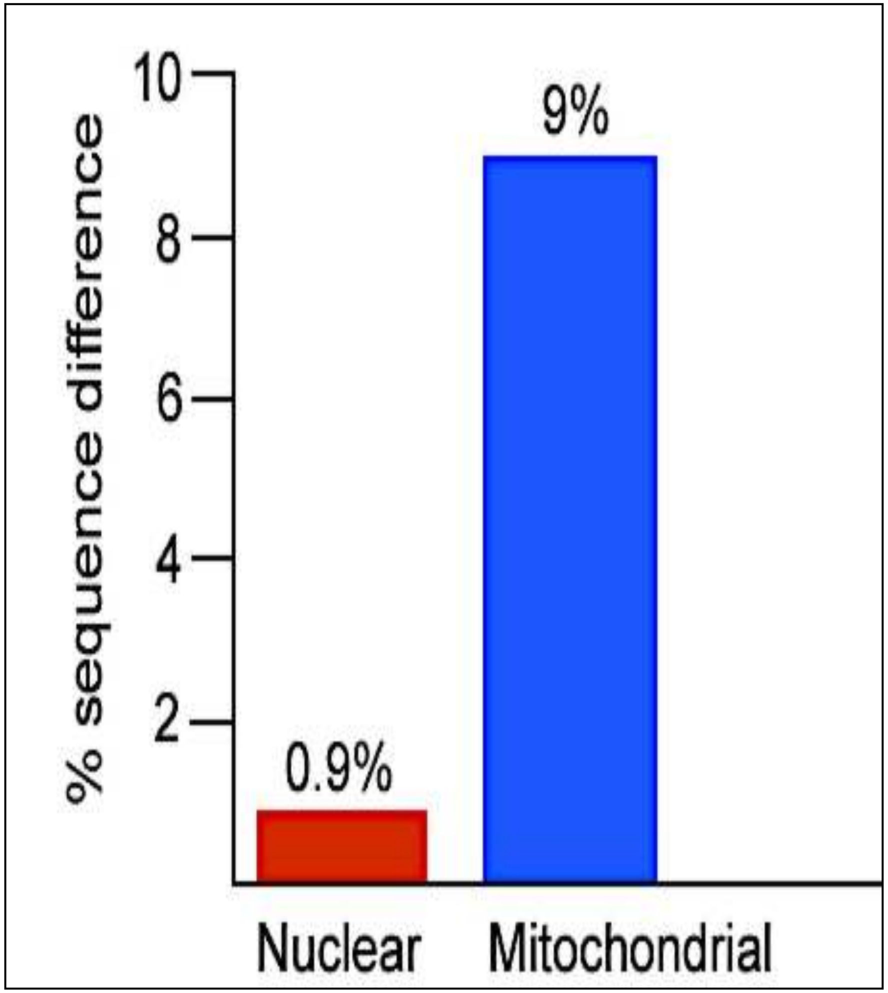
✓ **ETEROGENEO**

Grandi differenze interspecifiche

5-10 volte maggiori nel mtDNA contro nDNA

⇩

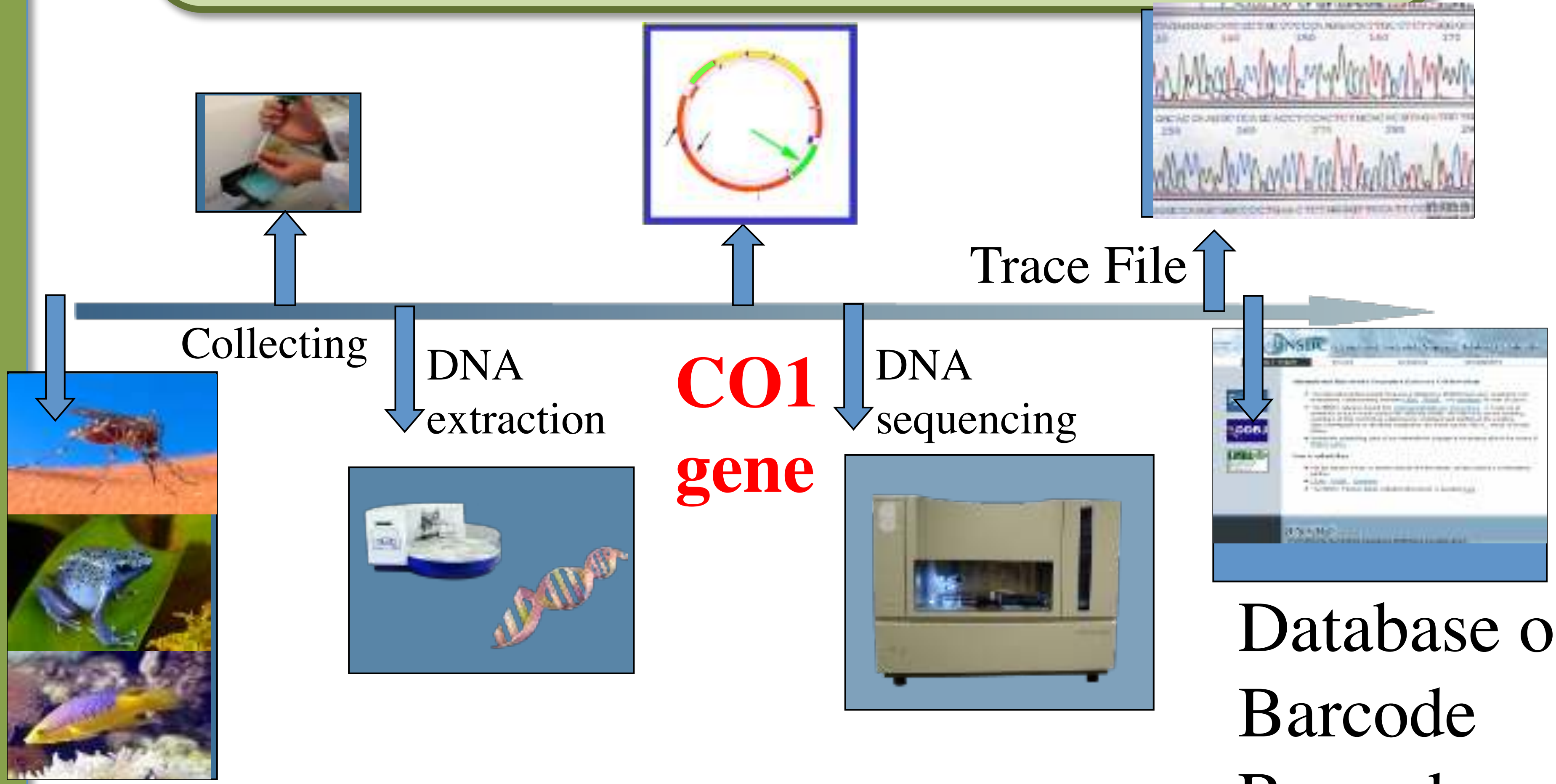
Garantisce una precisa distinzione di specie



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

How Barcoding is Done

From specimen to sequence to species



Voucher Specimen

Collecting

DNA extraction

CO1 gene

DNA sequencing

Trace File

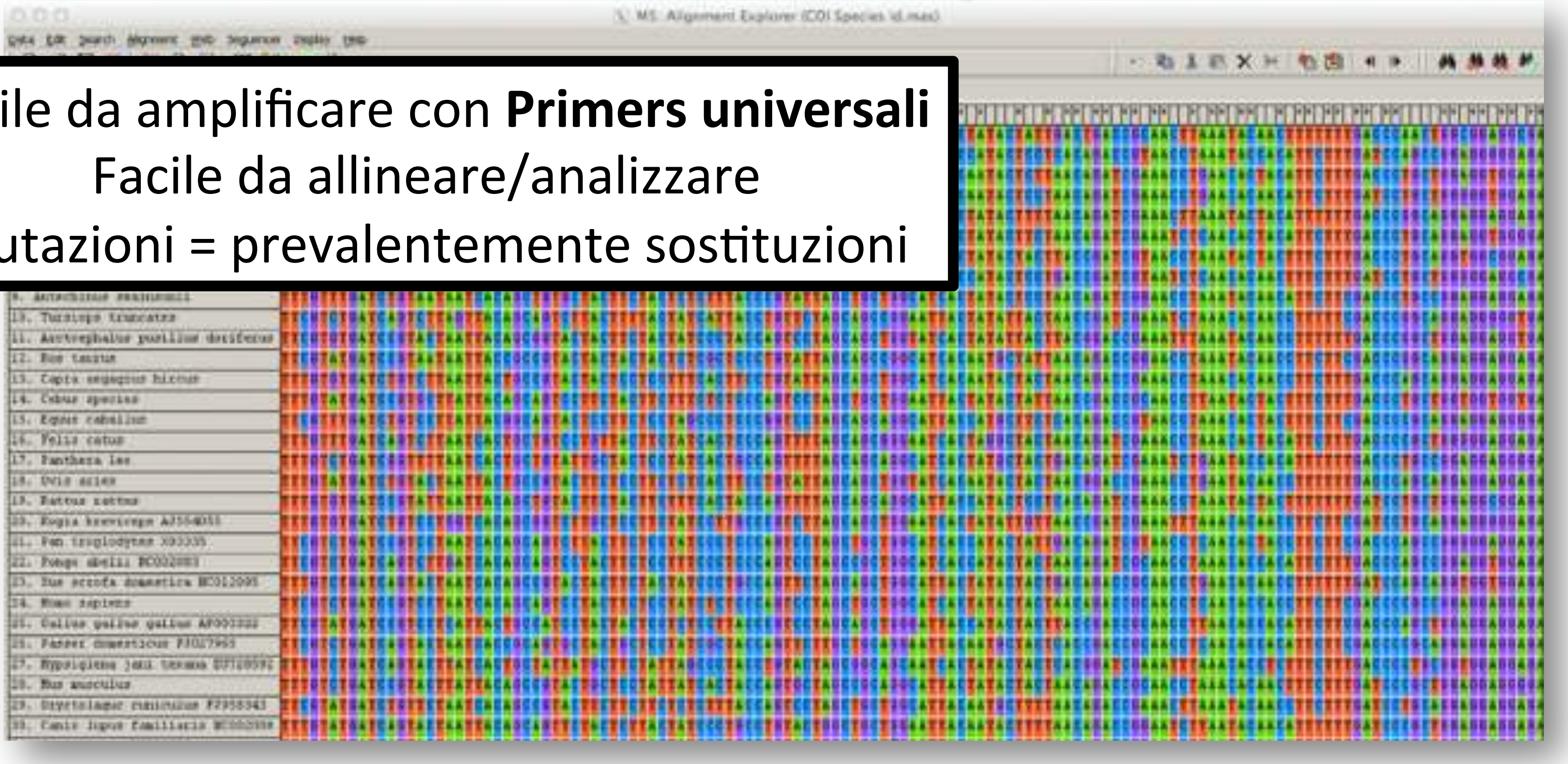
Database of Barcode Records

Cytochrome c oxidase (COI)

- ✓ Gene mitocondriale codificante
- ✓ Componente della catena respiratoria
- ✓ Catalizza la reazione di riduzione dell’ossigeno in acqua

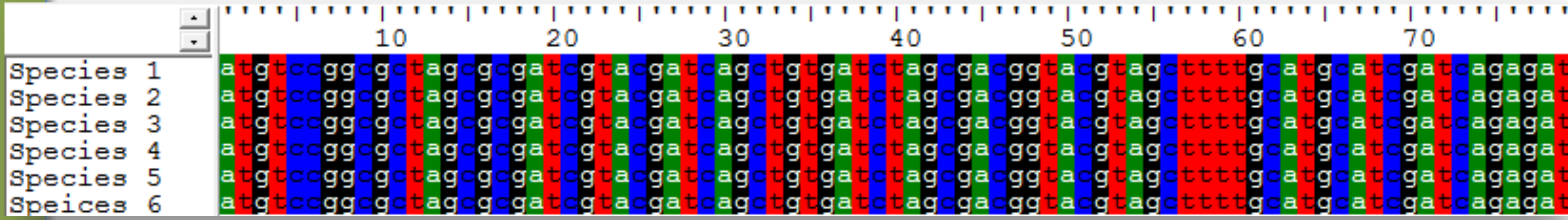
CONSERVATA
Basso tasso di mutazione

Facile da amplificare con **Primers universali**
Facile da allineare/analizzare
Mutazioni = prevalentemente sostituzioni



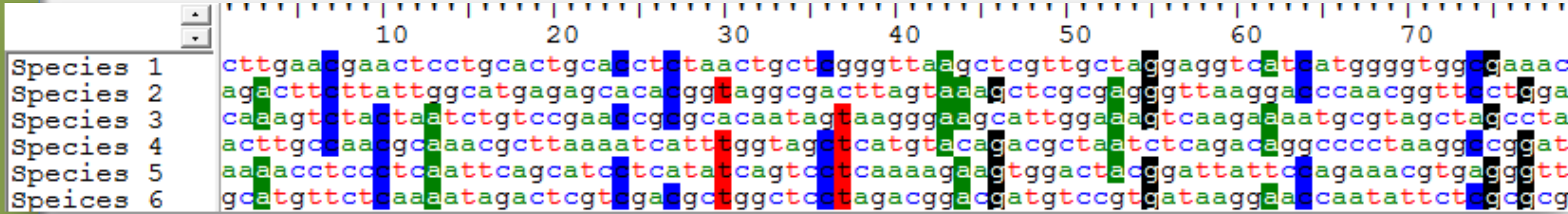
Esempi

Fail: Sequence is completely conserved, good for PCR, but uninformative as barcode



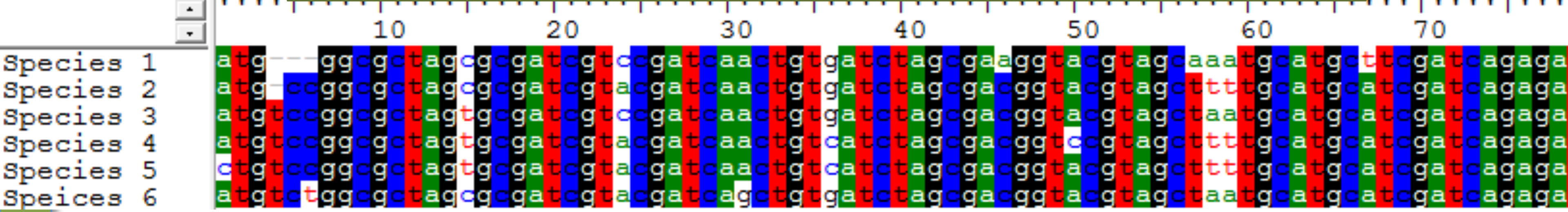
```
Species 1   atgtccggcgctagcgcgatcgtacgatcagctgtgatctagcgacggtacgtagcttttgcatgcatcgatcagagat
Species 2   atgtccggcgctagcgcgatcgtacgatcagctgtgatctagcgacggtacgtagcttttgcatgcatcgatcagagat
Species 3   atgtccggcgctagcgcgatcgtacgatcagctgtgatctagcgacggtacgtagcttttgcatgcatcgatcagagat
Species 4   atgtccggcgctagcgcgatcgtacgatcagctgtgatctagcgacggtacgtagcttttgcatgcatcgatcagagat
Species 5   atgtccggcgctagcgcgatcgtacgatcagctgtgatctagcgacggtacgtagcttttgcatgcatcgatcagagat
Speices 6   atgtccggcgctagcgcgatcgtacgatcagctgtgatctagcgacggtacgtagcttttgcatgcatcgatcagagat
```

Fail: Sequence shows no conservation, impossible for PCR, but good as barcode



```
Species 1   cttgaacgaactcctgcactgcacctctaactgctcgggttaagctcgttgctaggaggtcatcatggggtggcgaaac
Species 2   agacttcttattggcatgagagcacacggtaggcgacttagtaaagctcgcgagggttaaggacccaacggttcctgga
Species 3   caaagtctactaatctgtccgaaccgcgcacaatagtaagggaagcattggaaagtcaagaaaatgcgtagctagccta
Species 4   acttgccaacgcaaacgcttaaaatcatttggtagctcatgtacagacgctaatctcagacaggcccctaaggccggat
Species 5   aaaacctccctcaattcagcatcctcatatcagtcctcaaaagaagtggactacggattattccagaaacgtgagggtt
Speices 6   gcatgttctcaaaatagactcgtcgacgctggctcctagacggacgatgtccgtgataaggaaccaatattctcgccgcg
```

Win: Sequence shows some (ideally ~70%) conservation, good for PCR, good as barcode

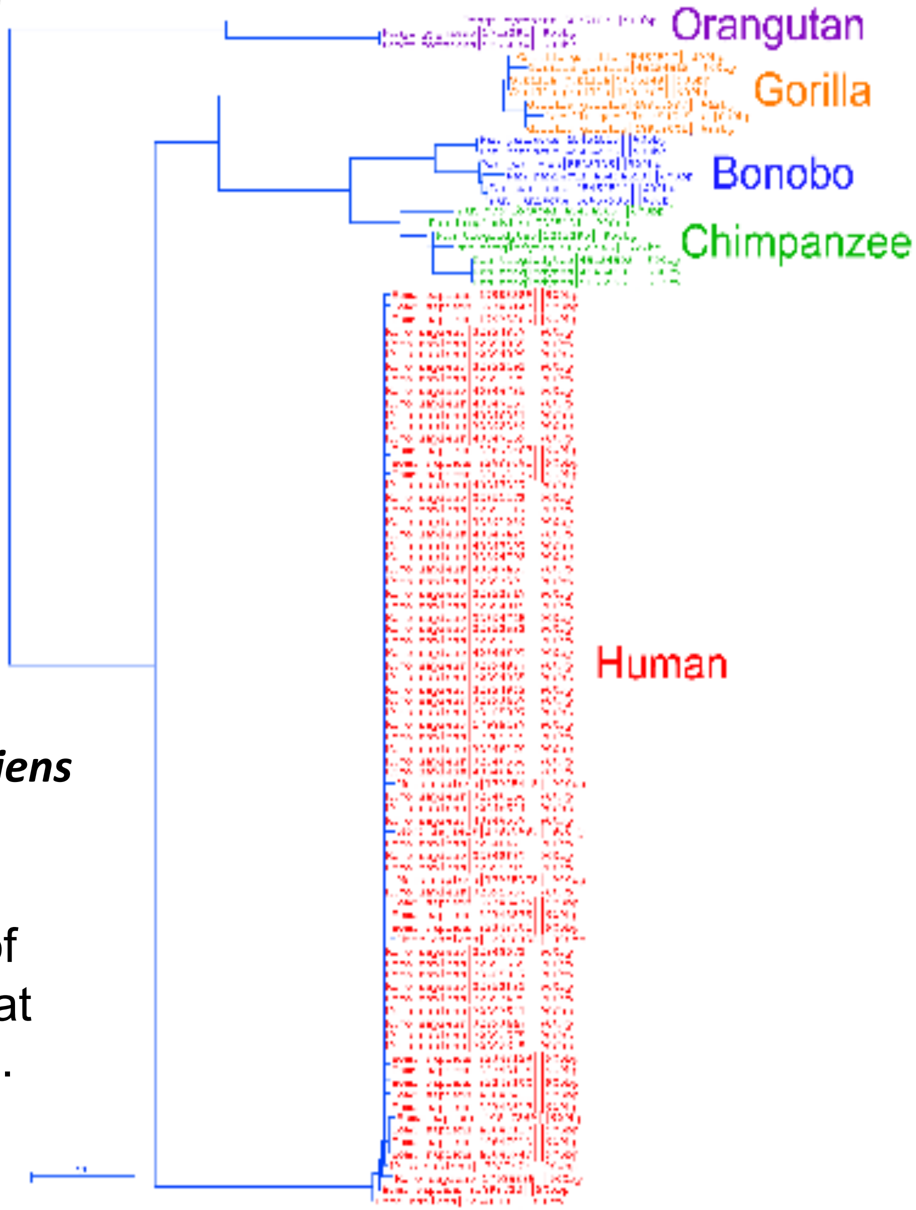


```
Species 1   atg---ggcgctagcgcgatcgtccgatcaactgtgatctagcgaaggtacgtagcaaatgcatgcttcgatcagaga
Species 2   atg-ccggcgctagcgcgatcgtacgatcaactgtgatctagcgacggtacgtagctttgcatgcatcgatcagaga
Species 3   atgtccggcgctagtgcgatcgtccgatcaactgtgatctagcgacggtacgtagctaatgcatgcatcgatcagaga
Species 4   atgtccggcgctagtgcgatcgtacgatcaactgtcatctagcgacggtccgtagctttgcatgcatcgatcagaga
Species 5   ctgtccggcgctagtgcgatcgtacgatcaactgtcatctagcgacggtacgtagctttgcatgcatcgatcagaga
Speices 6   atgtctggcgctagcgcgatcgtacgatcagctgtgatctagcgacggtacgtagctaatgcatgcatcgatcagaga
```

Cytochrome c oxidase (COI)

- ✓ Studi di comparazione di 13.320 specie congeneriche hanno dato esiti comparabili con la classificazione e migliorato le ambiguità in specie criptiche

- ✓ **Media di errore con utilizzo di COI è < 5%**

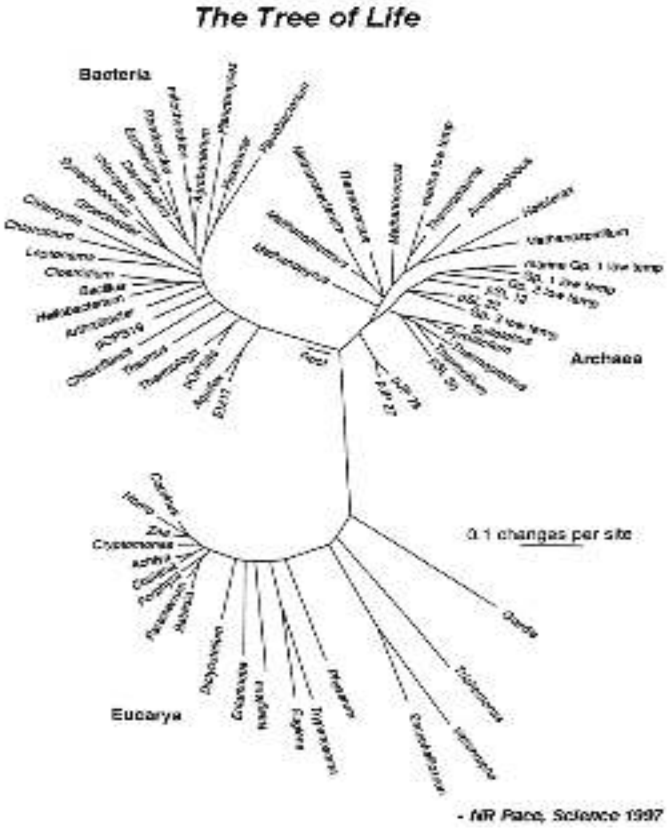


Confermata l'unità della specie *Homo sapiens*

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.

DNA Barcoding

Is it anything really new?



- NR Pace, Science 1997

………NO ………. BUT!

Scale & Standardization

‘One gene fits all’

Cytochrome c oxidase (COI)

La COI diventa il **DNA Barcode per eccellenza permettendo il riconoscimento del 98% delle specie ANIMALI**

Cytochrome c oxidase (COI) - **Limiti**

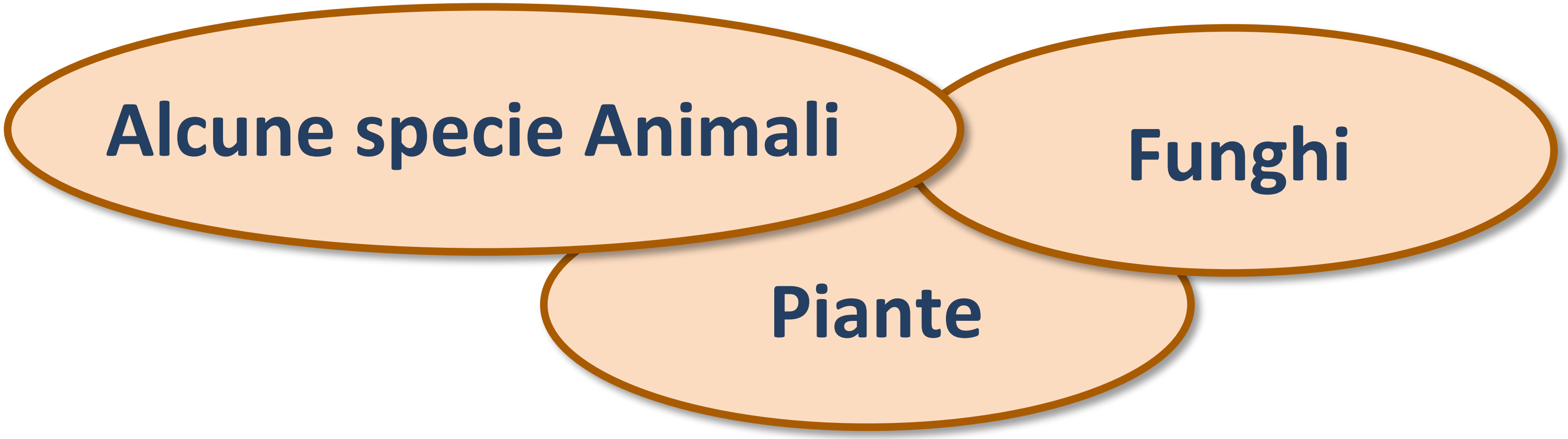
✓ Regole sull’eredità del mtDNA non universali

- Casi di doppia eredità parentale per mtDNA
- Casi di ricombinazione (e.g. molluschi)
- Casi di perdita del mtDNA

} Animali

✓ Impossibilità di identificazione gli **IBRIDI e specie che hanno subito passate introgressioni**

✓ **Variabilità interspecifica troppo bassa** in alcuni taxa



DNA Barcoding

- DNA barcode è diventato uno strumento standard per scoprire, descrivere, e comprendere la biodiversità.
- DNA barcode assieme ai dati morfologici, biochimici e ecologici ha permesso la caratterizzazione di un'enorme quantità di specie prima non riconoscibili.
- Solo nei primi 8 mesi del 2014 si contano 310 pubblicazioni nelle quali il DNA barcode viene usato per l'identificazione di nuove specie.

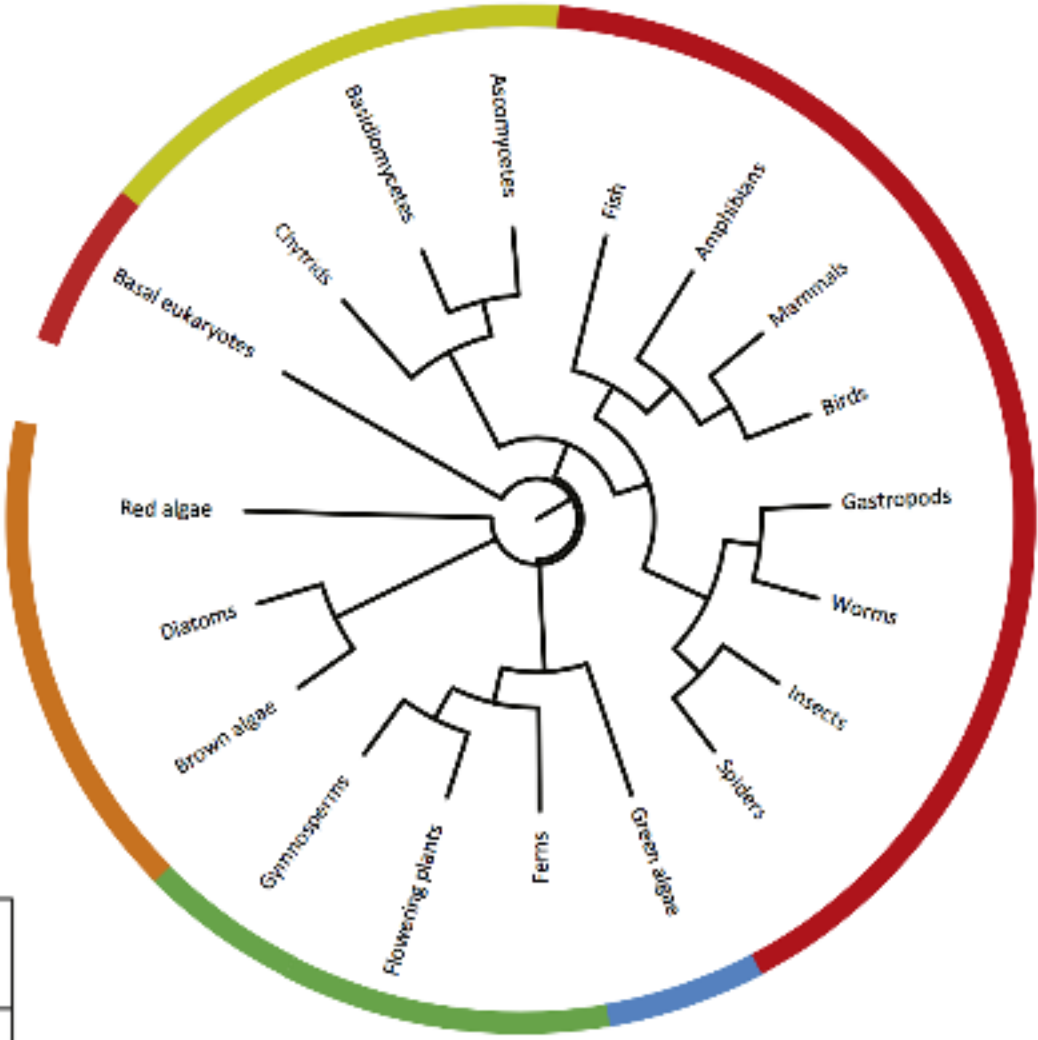
- MA . . . Se all'inizio si pensava di poter utilizzare un unico marcatore come **barcode universale** per completare il Tree of Life degli organismi Eucarioti.

DNA Barcoding

In pratica . . .

Servono diverse regioni regioni per un'adeguata identificazione tra diversi *lineages*

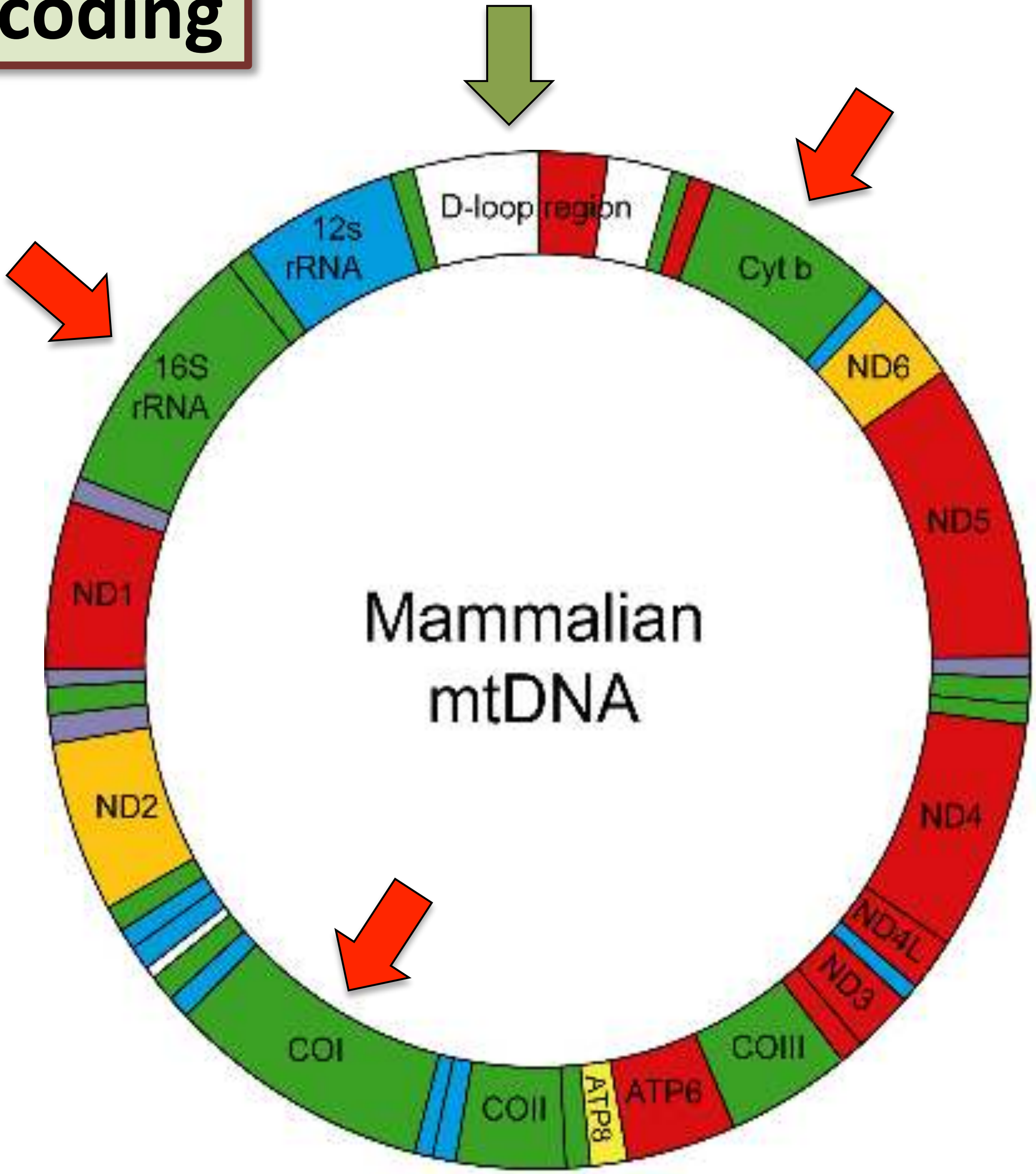
Tree of life



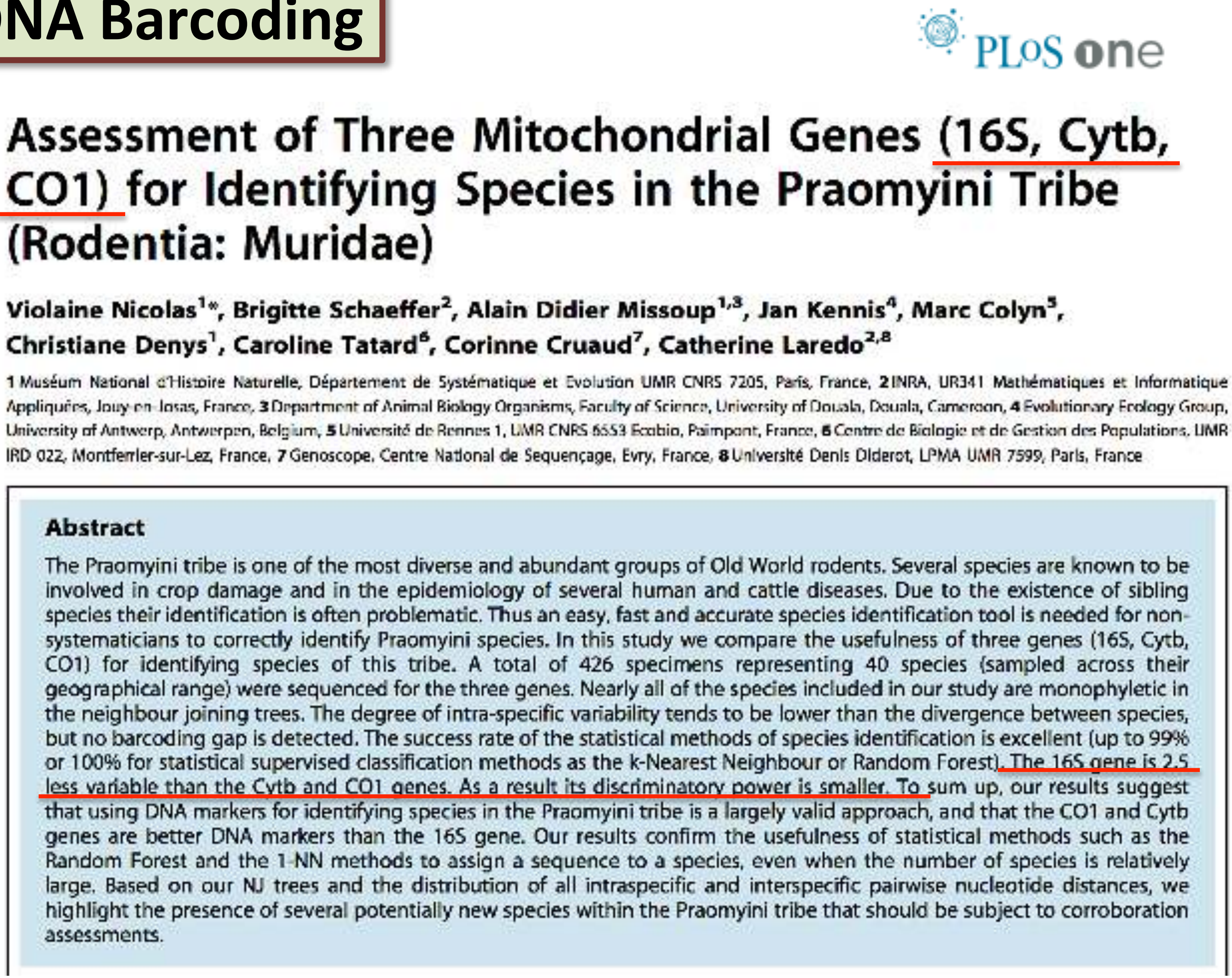
Key:	Color	Clade	Primary barcode(s)	Secondary barcode(s)
	🟥	Animals	CO1	CO1, 16S
	🟨	Fungi	ITS	LSU D1/D2
	🟦	Green algae	*tuf*A	LSU D2/D3
	🟩	Land plants	*rbc*L/*mat*K	*psb*A-*trn*H/ITS
	🟧	Algae	CO1-5P	LSU D2/D3

TRENDS in Ecology & Evolution

DNA Barcoding



DNA Barcoding



PLoS one

Assessment of Three Mitochondrial Genes (16S, Cytb, CO1) for Identifying Species in the Praomyini Tribe (Rodentia: Muridae)

Violaine Nicolas[1]*, Brigitte Schaeffer[2], Alain Didier Missoup[1,3], Jan Kennis[4], Marc Colyn[5], Christiane Denys[1], Caroline Tatard[6], Corinne Cruaud[7], Catherine Laredo[2,8]

1 Muséum National d'Histoire Naturelle, Département de Systématique et Evolution UMR CNRS 7205, Paris, France, 2 INRA, UR341 Mathématiques et Informatique Appliquées, Jouy-en-Josas, France, 3 Department of Animal Biology Organisms, Faculty of Science, University of Douala, Douala, Cameroon, 4 Evolutionary Ecology Group, University of Antwerp, Antwerpen, Belgium, 5 Université de Rennes 1, UMR CNRS 6553 Ecobio, Paimpont, France, 6 Centre de Biologie et de Gestion des Populations, UMR IRD 022, Montferrier-sur-Lez, France, 7 Genoscope, Centre National de Sequençage, Evry, France, 8 Université Denis Diderot, LPMA UMR 7599, Paris, France

Abstract

The Praomyini tribe is one of the most diverse and abundant groups of Old World rodents. Several species are known to be involved in crop damage and in the epidemiology of several human and cattle diseases. Due to the existence of sibling species their identification is often problematic. Thus an easy, fast and accurate species identification tool is needed for non-systematicians to correctly identify Praomyini species. In this study we compare the usefulness of three genes (16S, Cytb, CO1) for identifying species of this tribe. A total of 426 specimens representing 40 species (sampled across their geographical range) were sequenced for the three genes. Nearly all of the species included in our study are monophyletic in the neighbour joining trees. The degree of intra-specific variability tends to be lower than the divergence between species, but no barcoding gap is detected. The success rate of the statistical methods of species identification is excellent (up to 99% or 100% for statistical supervised classification methods as the k-Nearest Neighbour or Random Forest). The 16S gene is 2.5 less variable than the Cytb and CO1 genes. As a result its discriminatory power is smaller. To sum up, our results suggest that using DNA markers for identifying species in the Praomyini tribe is a largely valid approach, and that the CO1 and Cytb genes are better DNA markers than the 16S gene. Our results confirm the usefulness of statistical methods such as the Random Forest and the 1-NN methods to assign a sequence to a species, even when the number of species is relatively large. Based on our NJ trees and the distribution of all intraspecific and interspecific pairwise nucleotide distances, we highlight the presence of several potentially new species within the Praomyini tribe that should be subject to corroboration assessments.

DNA Barcoding - Plants

- Low heterogenety in the mithocondrial COI gene.

Low rates of nucleotide substitution in mtDNA compared with animals

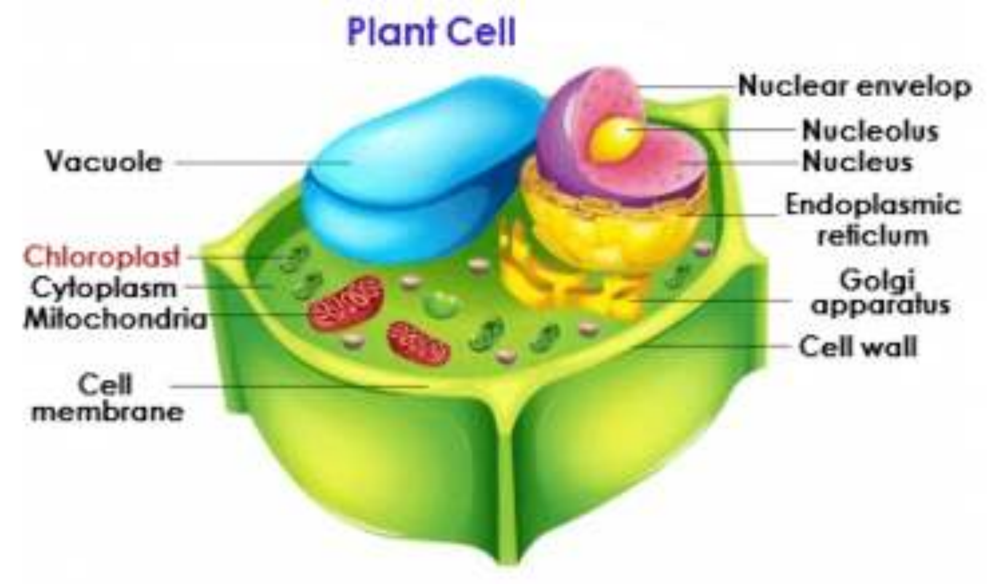
PLANT CHLOROPLAST GENOME

as an alternative to the animal mtDNA

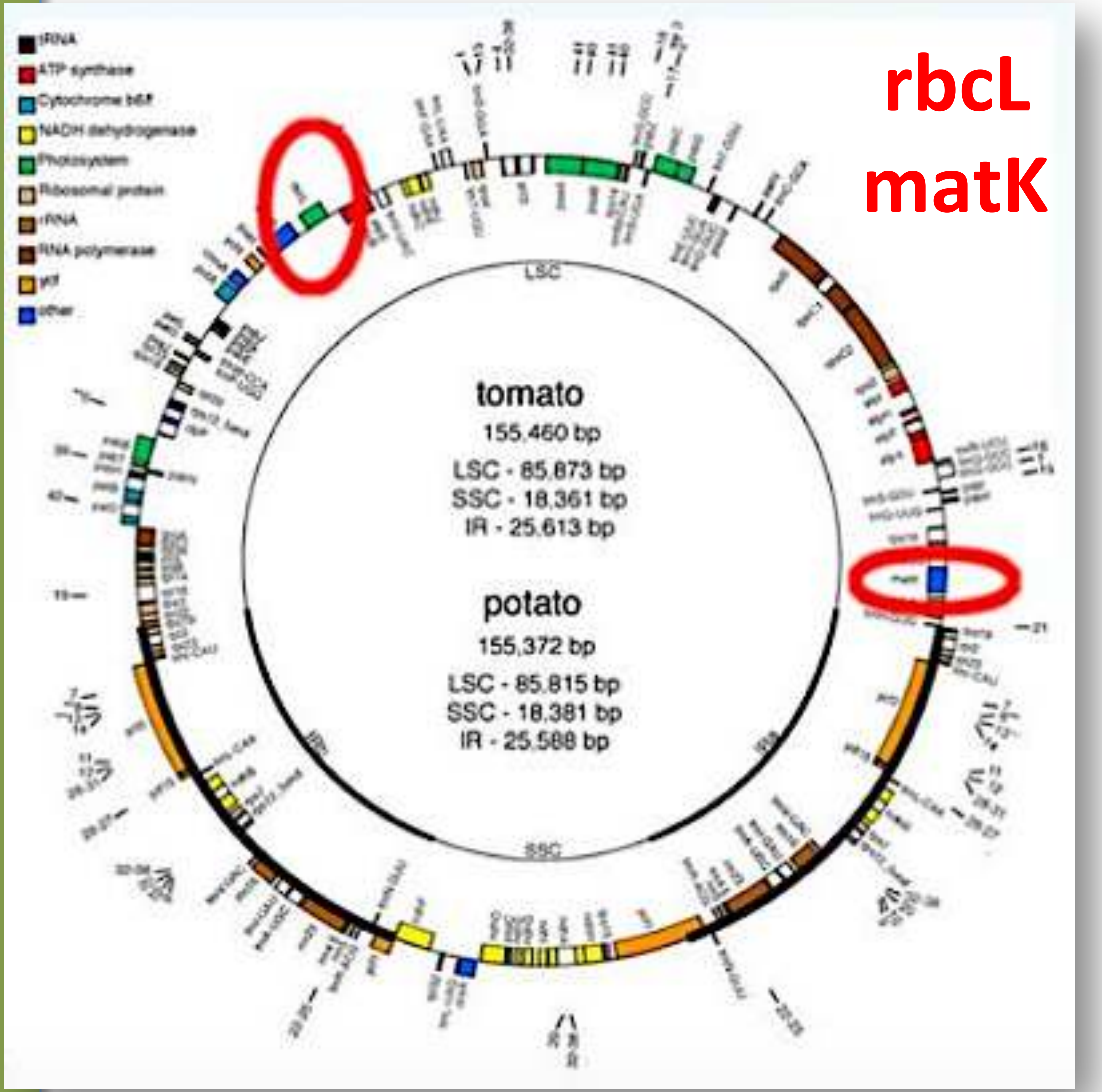
- Presente in tutte le cellule
- Elevato numero di copie
- Regioni codificanti e non

MA . . .

In generale, presenta un basso tasso evolutivo



DNA Barcoding - Plants



rbcL

matK

Core plant barcoding region

Si è arrivati alla conclusione di ricercare la sequenza barcode per le piante considerando delle **combinazioni di più geni**

rbcL + matK

- Identificano il **genere** nel **100%** dei casi
- Identificano la **specie** nel **72%** dei casi

- - - - - - - - - -

psbA + trnH

Markers supplementari per casi di difficile interpretazione

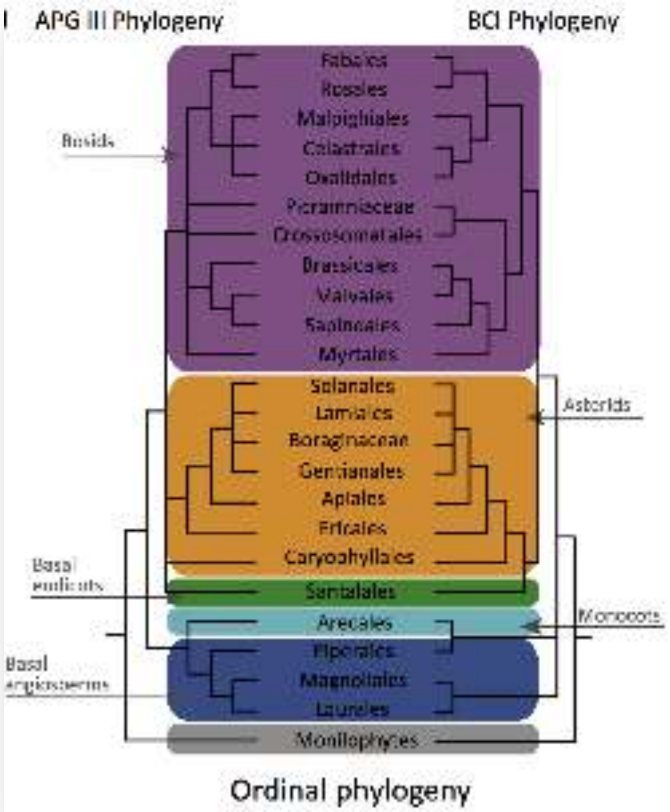
DNA Barcoding - **Plants**

Prima applicazione fu la ricostruzione filogenetica della comunità di piante della foresta di Barro Colorado Island (Panama).

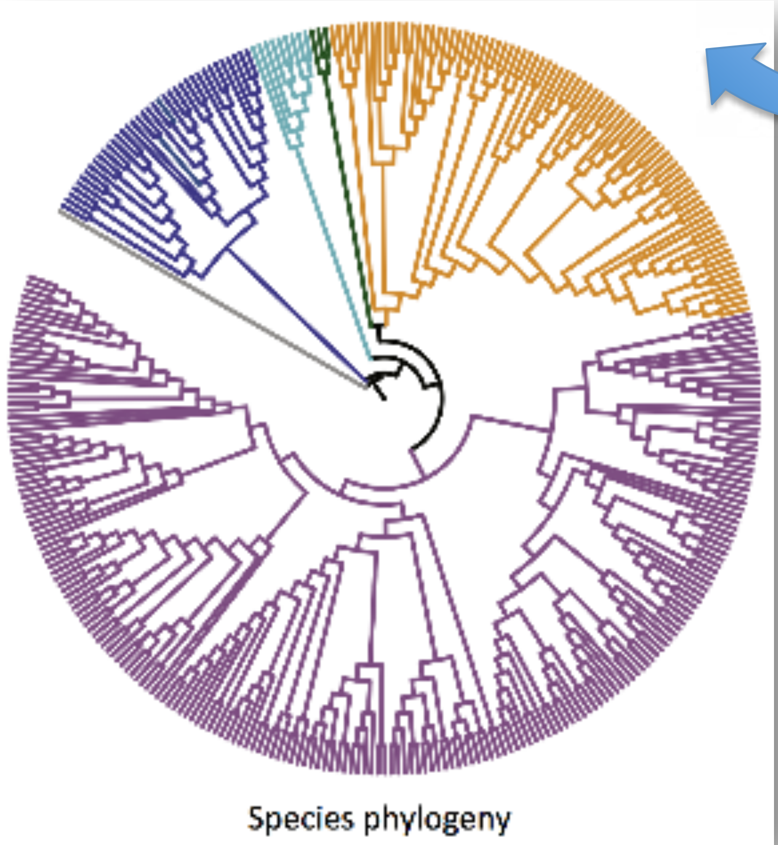
Stimate 281 specie vegetali

DNA Barcoding - Plants

- Robusta ricostruzione delle relazioni filogenetiche con **l’approccio multilocus**
- Sensibile **miglioramento della topologia dell’albero all’estremità dei rami**



Ordinal phylogeny



Species phylogeny

rbcL + matK

Miglior stima dei nodi profondi

trnH + psbA

Miglior stima dei nodi alle estremità dei rami

TRENDS in Ecology & Evolution

DNA Barcoding - Fungi

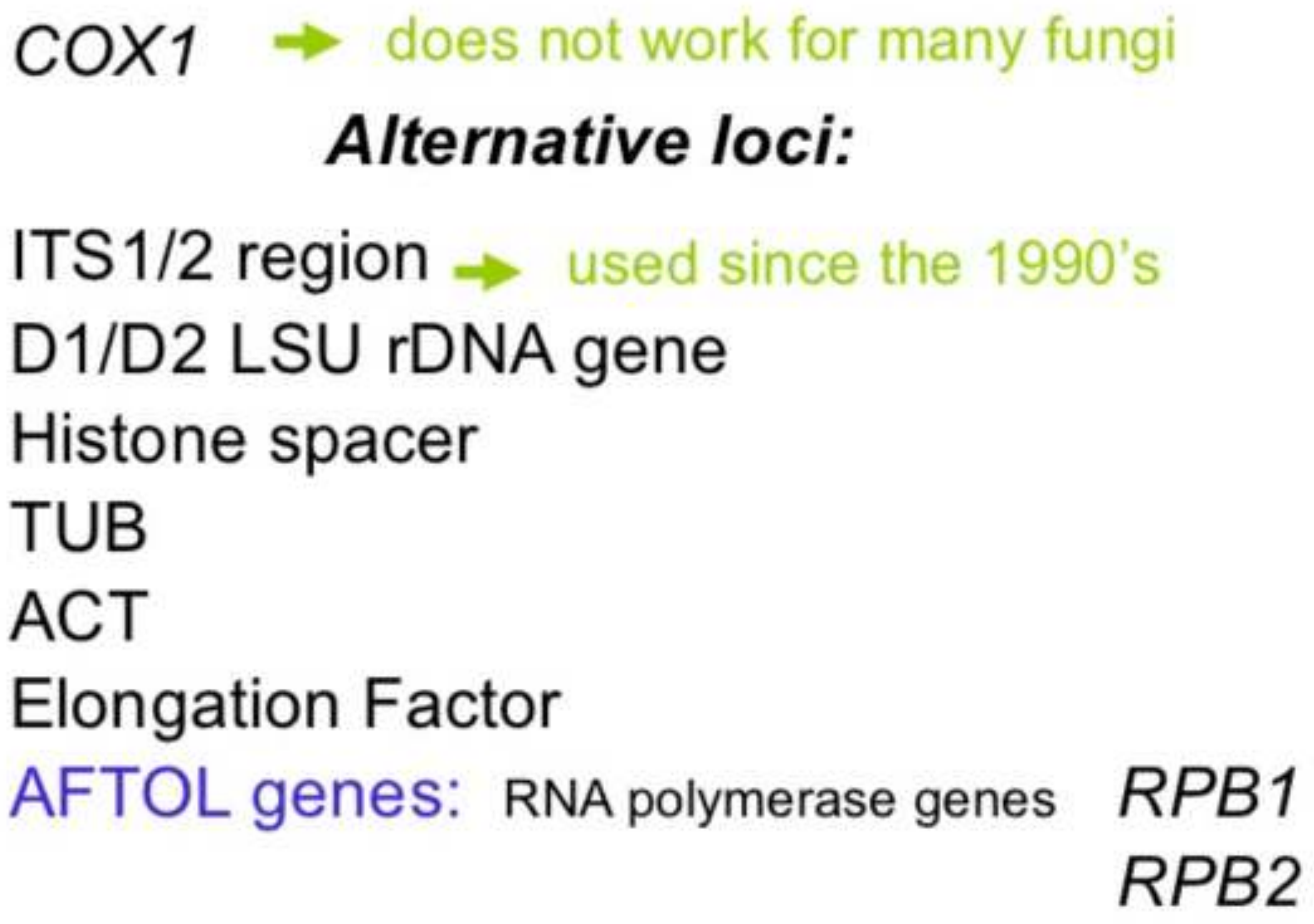
- Rivestono ruoli importanti negli ecosistemi terrestri
- Racchiudono gruppi molto diversi di organismi (es. da unicellulari a macroorganismi, ricco di specie criptiche)
- Incremento della frequenza di infezioni fungine con cure limitate dovute ad una bassa risoluzione delle tecniche di identificazione morfo/fisiologiche.

Fungal Barcoding Database managed by the

International Fungal Working group

(http://www.fungalbarcoding.org)

DNA Barcoding - Fungi



COX1 → does not work for many fungi

Alternative loci:

ITS1/2 region → used since the 1990's

D1/D2 LSU rDNA gene

Histone spacer

TUB

ACT

Elongation Factor

AFTOL genes: RNA polymerase genes *RPB1*

RPB2

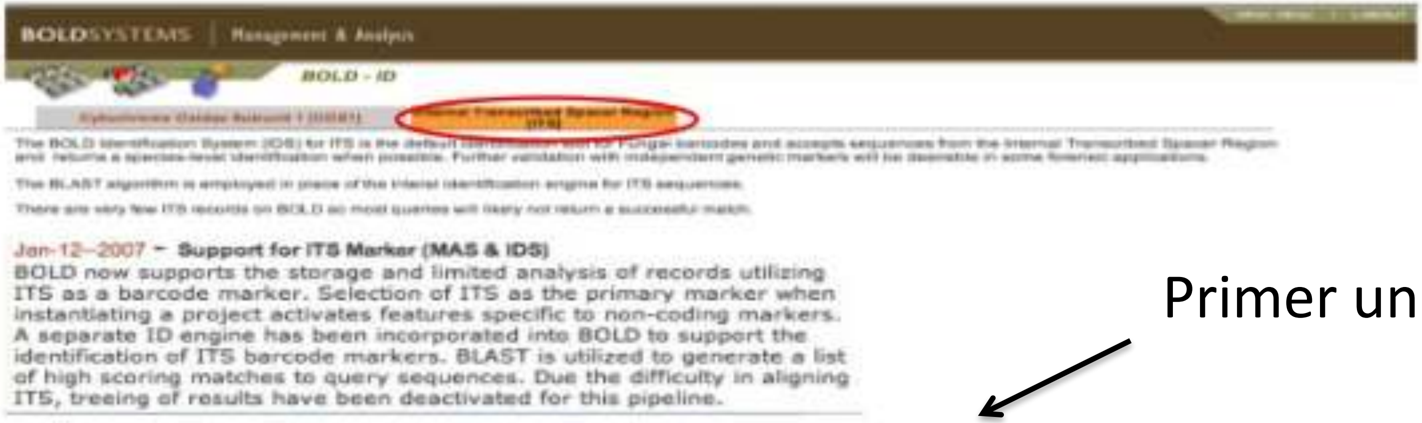
Non riconoscono fino a livello di specie!

DNA Barcoding - Fungi

BARCODE OF LIFE DATA SYSTEMS v2.5

Advancing species identification and discovery through the analysis of short, standardized gene regions

The International Sub-commission on Fungal Barcoding has proposed the ITS region as the prime fungal barcode or the default region for species identification (http://www.allfungi.com/its-barcode.php).



Jan-12-2007 – **Support for ITS Marker (MAS & IDS)**
BOLD now supports the storage and limited analysis of records utilizing ITS as a barcode marker. Selection of ITS as the primary marker when instantiating a project activates features specific to non-coding markers. A separate ID engine has been incorporated into BOLD to support the identification of ITS barcode markers. BLAST is utilized to generate a list of high scoring matches to query sequences. Due the difficulty in aligning ITS, treeing of results have been deactivated for this pipeline.

Primer universali

Fungal specific primers: **SR6R: 5' AAGTATAAGTCGTAACAAGG 3'**
LR1: 5' GGTTGGTTTCTTTTCCT 3'

[Vilgalys & Hesters (1990) J. of Bacteriol. 20: 4238-4246]

© WMeyer, USYD, Australia, BOL4 Adelaide Nov 2011

DNA Barcoding (COI)– Punti di forza

- ✓ Possibilità di discriminare *sibling or cryptic species*
- ✓ Chiarire le relazioni tra variabili morfologiche entrospecie (es. individuare diversi stadi di sviluppo)
- ✓ Può aiutare nell'identificazione di interazioni tra specie (es. specificità ospite parassita)
- ✓ Applicato all'entomologia e all'entomologia forense
- ✓ Primo passo verso l'ideazione di un sistema palmare
- ✓ Valorizzazione delle collezioni museali
- ✓ Frodi commerciali e genetica della conservazione
- ✓ Identificazione delle specie/individuo sulla base di parti di individui

DNA Barcoding (COI)– Punti di forza

Guanacaste Conservation Area
(Costa Rica).
Una sola specie di *common skipper butterfly* definita da caratteri morfologici degli individui adulti.

Analisi del DNA barcodes ha identificato un complesso di 10 specie diverse.
Successivamente confermate da incongruenze nella dieta e nelle caratteristiche delle larve.

Discovery of cryptic species

"Astraptes fulgerator" complex
(Lepidoptera: Hesperiidae)

TRENDS in Ecology & Evolution

DNA Barcoding (COI)– Punti di forza

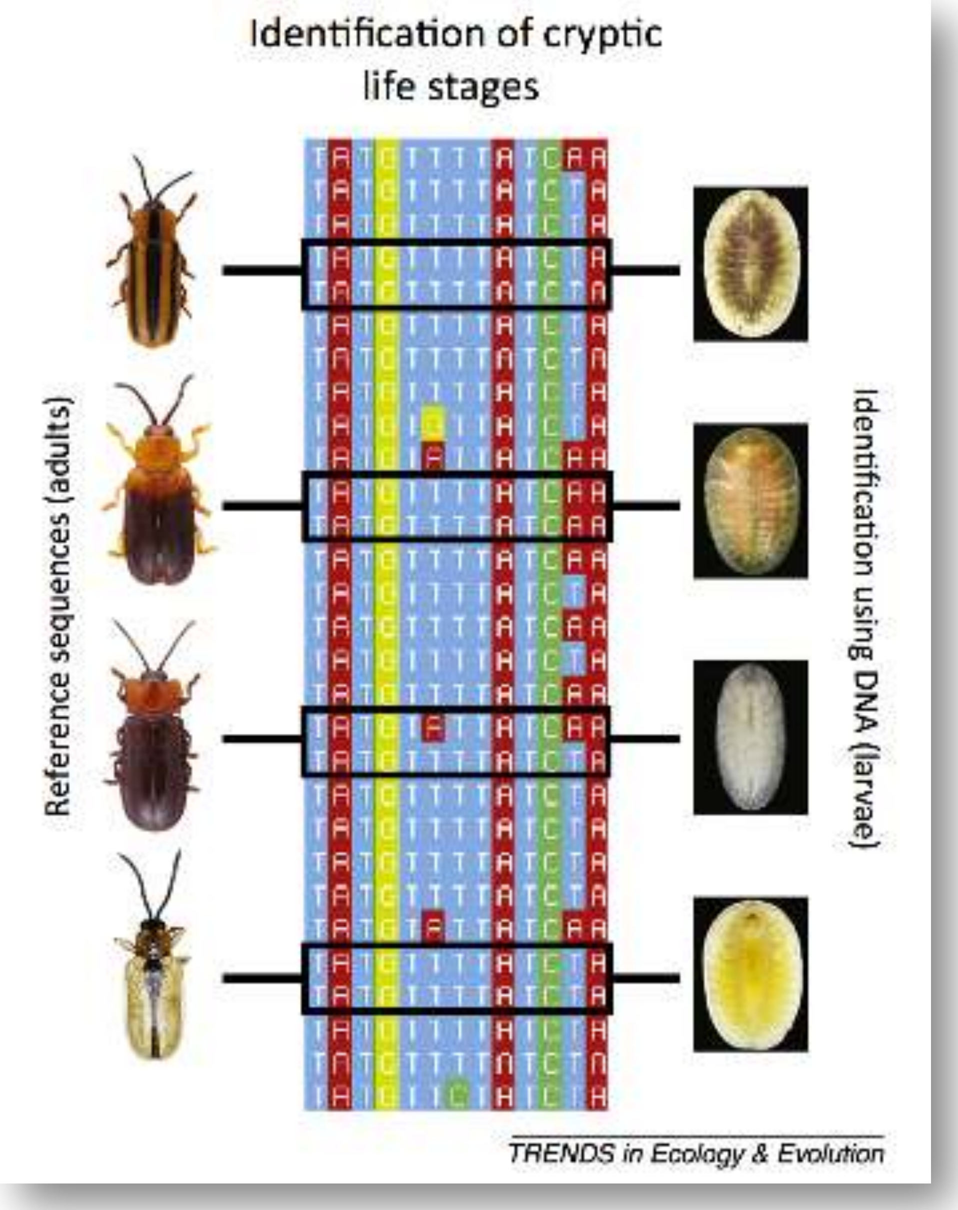
- ✓ Possibilità di discriminare *sibling or cryptic species*
- ✓ Chiarire le relazioni tra variabili morfologiche entrospecie (es. individuare diversi stadi di sviluppo)
- ✓ Può aiutare nell'identificazione di interazioni tra specie (es. specificità ospite parassita)
- ✓ Applicato all'entomologia e all'entomologia forense
- ✓ Primo passo verso l'ideazione di un sistema palmare
- ✓ Valorizzazione delle collezioni museali
- ✓ Frodi commerciali e genetica della conservazione
- ✓ Identificazione delle specie/individuo sulla base di parti di individui

DNA Barcoding (COI)– Punti di forza

Grazie all'analisi del DNA barcode:

- Associazione di stadi larvali immaturi con la loro forma adulta
- Identificazioni delle relazioni tra pianta e erbivoro

Nuovi orizzonti sulla determinazione dei confini di specie



TRENDS in Ecology & Evolution

DNA Barcoding (COI)– Punti di forza

- ✓ Possibilità di discriminare *sibling or cryptic species*
- ✓ Chiarire le relazioni tra variabili morfologiche entrospecie (es. individuare diversi stadi di sviluppo)
- ✓ Può aiutare nell'identificazione di interazioni tra specie (es. specificità ospite parassita)
- ✓ Applicato all'entomologia e all'entomologia forense
- ✓ Primo passo verso l'ideazione di un sistema palmare
- ✓ Valorizzazione delle collezioni museali
- ✓ Frodi commerciali e genetica della conservazione
- ✓ Identificazione delle specie/individuo sulla base di parti di individui

DNA Barcoding (COI)– Punti di forza

- ✓ Possibilità di discriminare *sibling or cryptic species*
- ✓ Chiarire le relazioni tra variabili morfologiche entrospecie (es. individuare diversi stadi di sviluppo)
- ✓ Può aiutare nell'identificazione di interazioni tra specie (es. specificità ospite parassita)
- ✓ Applicato all'entomologia e all'entomologia forense
- ✓ Primo passo verso l'ideazione di un sistema palmare
- ✓ Valorizzazione delle collezioni museali
- ✓ Frodi commerciali e genetica della conservazione
- ✓ Identificazione delle specie/individuo sulla base di parti di individui

DNA Barcoding (COI)– Punti di forza

- Monitoraggio di vettori di malattie
- Tipizzazione di organismi invasivi
- Ricerca di organismi candidati come agenti per il controllo biologico
- Completamento delle filogenesi soprattutto per gli insetti che tipicamente si trovano in cadaveri in decomposizione

DNA Barcoding (COI)– Punti di forza

- ✓ Possibilità di discriminare *sibling or cryptic species*
- ✓ Chiarire le relazioni tra variabili morfologiche entrospecie (es. individuare diversi stadi di sviluppo)
- ✓ Può aiutare nell'identificazione di interazioni tra specie
- ✓ Applicato all'entomologia e all'entomologia forense
- ✓ Primo passo verso l'ideazione di un sistema palmare
- ✓ Valorizzazione delle collezioni museali
- ✓ Frodi commerciali e genetica della conservazione
- ✓ Identificazione delle specie/individuo sulla base di parti di individui

DNA Barcoding (COI)– Punti di forza

DNA barcode come base per la creazione di un sistema portatile di identificazione di specie direttamente sul campo

DNA Barcoding (COI)– Punti di forza

- ✓ Possibilità di discriminare *sibling or cryptic species*
- ✓ Chiarire le relazioni tra variabili morfologiche entrospecie (es. individuare diversi stadi di sviluppo)
- ✓ Può aiutare nell'identificazione di interazioni tra specie
- ✓ Applicato all'entomologia e all'entomologia forense
- ✓ Primo passo verso l'ideazione di un sistema palmare
- ✓ Valorizzazione delle collezioni museali
- ✓ Frodi commerciali e genetica della conservazione
- ✓ Identificazione delle specie/individuo sulla base di parti di individui

DNA Barcoding (COI)– Punti di forza

- ✓ Possibilità di discriminare *sibling or cryptic species*
- ✓ Chiarire le relazioni tra variabili morfologiche entrospecie (es. individuare diversi stadi di sviluppo)
- ✓ Può aiutare nell'identificazione di interazioni tra specie
- ✓ Applicato all'entomologia e all'entomologia forense
- ✓ Primo passo verso l'ideazione di un sistema palmare
- ✓ Valorizzazione delle collezioni museali
- ✓ Frodi commerciali e genetica della conservazione
- ✓ Identificazione delle specie/individuo sulla base di parti di individui

DNA Barcoding (COI)– Punti di forza

- Detect mislabeling and illegal trade of endangered and/or protected species
- Monitoring of conservation actions

SCIENCE JOURNAL | DECEMBER 4, 2009

DNA 'Barcodes' Surface Fishy Imposters on Menus

By ROBERT LEE HOTZ

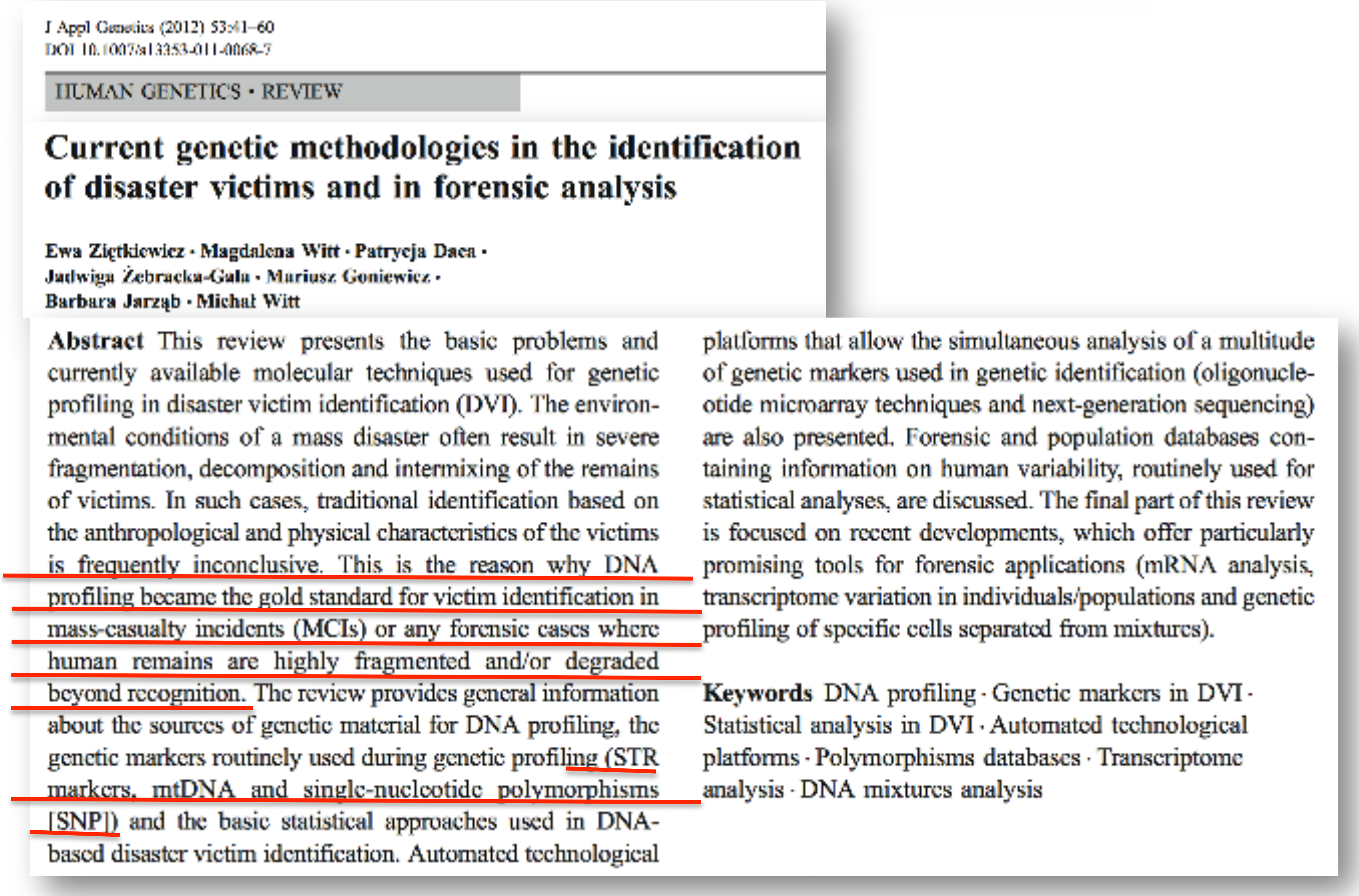
Researchers Use Gene Segments to Settle Restaurant Mysteries, Check Stream Quality and Take Inventory of All Living Things

Researchers using a new DNA test recently discovered that fish ordered from menus in New York and Denver might not always be the species served. Sampling the fare at 31 sushi bars, scientists at the American Museum of Natural History found that customers who ordered tuna were sometimes served a cheaper substitute, an endangered species or a fish banned in several countries as a health hazard.

DNA Barcoding (COI)– Punti di forza

- ✓ Possibilità di discriminare *sibling or cryptic species*
- ✓ Chiarire le relazioni tra variabili morfologiche entrospecie (es. individuare diversi stadi di sviluppo)
- ✓ Può aiutare nell'identificazione di interazioni tra specie
- ✓ Applicato all'entomologia e all'entomologia forense
- ✓ Primo passo verso l'ideazione di un sistema palmare
- ✓ Valorizzazione delle collezioni museali
- ✓ Frodi commerciali e genetica della conservazione
- ✓ Identificazione delle specie/individui sulla base di parti di individui

DNA Barcoding – Punti di forza e applicazioni



J Appl Genetics (2012) 53:41–60
DOI 10.1007/s13353-011-0068-7

HUMAN GENETICS • REVIEW

Current genetic methodologies in the identification of disaster victims and in forensic analysis

Ewa Ziętkiewicz · Magdalena Witt · Patrycja Daca ·
Jadwiga Żebracka-Gala · Mariusz Goniewicz ·
Barbara Jarząb · Michał Witt

Abstract This review presents the basic problems and currently available molecular techniques used for genetic profiling in disaster victim identification (DVI). The environmental conditions of a mass disaster often result in severe fragmentation, decomposition and intermixing of the remains of victims. In such cases, traditional identification based on the anthropological and physical characteristics of the victims is frequently inconclusive. This is the reason why DNA profiling became the gold standard for victim identification in mass-casualty incidents (MCIs) or any forensic cases where human remains are highly fragmented and/or degraded beyond recognition. The review provides general information about the sources of genetic material for DNA profiling, the genetic markers routinely used during genetic profiling (STR markers, mtDNA and single-nucleotide polymorphisms [SNP]) and the basic statistical approaches used in DNA-based disaster victim identification. Automated technological platforms that allow the simultaneous analysis of a multitude of genetic markers used in genetic identification (oligonucleotide microarray techniques and next-generation sequencing) are also presented. Forensic and population databases containing information on human variability, routinely used for statistical analyses, are discussed. The final part of this review is focused on recent developments, which offer particularly promising tools for forensic applications (mRNA analysis, transcriptome variation in individuals/populations and genetic profiling of specific cells separated from mixtures).

Keywords DNA profiling · Genetic markers in DVI · Statistical analysis in DVI · Automated technological platforms · Polymorphisms databases · Transcriptome analysis · DNA mixtures analysis

DNA Barcoding – Altre Applicazioni

Epidemiologia e salute pubblica

✓ Identificazione di organismi vettori di malattie (e.g. vettori della malaria - zanzare, rabbia - mammiferi, borreliosi – zecche)

✓ Pianificazione di campagne di vaccinazione e mappe di distribuzione di rischio

✓ Diagnosi delle parassitosi

Difesa dei consumatori

✓ Identificazione di specie presenti nei cibi freschi o processati

✓ Qualità, origine e presenza di patogeni

Implicazioni SANITARIE

- Specie tossiche
- Provenienti da aree da cui è vietata l'importazione o che richiedono controlli sanitari
- Specie esotiche (trasmissione di zoonosi)
- Specie allergeniche o a basso valore nutritivo

Implicazioni ETICHE

- Religione
- Costo
- Legislazione

DNA Barcoding – Altre Applicazioni

Norme doganali

- ✓ Identificazione di specie Aliene invasive
- ✓ Controllo del rispetto delle normative di import/export
- ✓ Identificazione del contrabbando di specie protette

Implicazioni nella GENETICA DELLA CONSERVAZIONE

- Controllo di **FRODI COMMERCIALI**

Monitoraggi ambientali e forensi

- ✓ Lista delle specie e relativa abbondanza
- ✓ Valutazione dello stato di salute del territorio
- ✓ Investigazioni sul commercio illegale

Implicazioni nella GENETICA DELLA CONSERVAZIONE

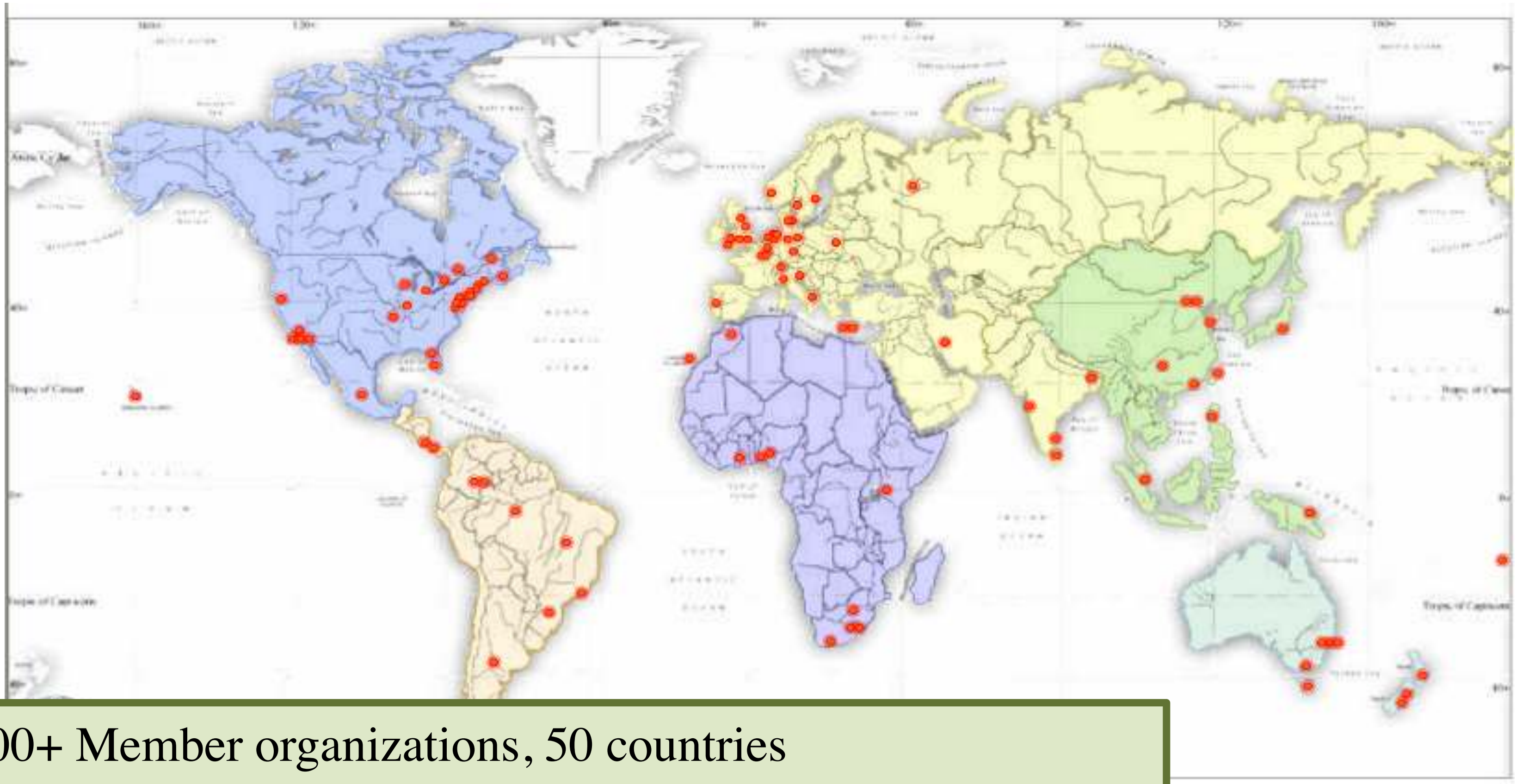
- **TRACCIABILITA' (origine, wild us aquaculture)**
- **Monitoraggio specie a rischio di estinzione**

Le ricerche del progetto **DNA Barcode** sono gestite e coordinate dal Consorzio **CBOL (*Consortium of the Barcode Of Life*)**, costituito nel **2003** presso la **Smithsonian Institution di Washington**

CBOL acts to organize, catalyse, and integrate barcoding activities around the world.

CBOL CORPORATION

CBOL Member Organizations: 2009



- 200+ Member organizations, 50 countries
- 35+ Member organizations from 20+ developing countries

Promuove l'iniziativa internazionale per lo sviluppo del DNA barcode come metodo standard per l'identificazione di species

Barcode of Life Community

CONSORTIUM FOR THE BARCODE OF LIFE

- Promote barcoding as a global standard
- Build participation
- Working Groups
- BARCODE standard
- International Conferences
- Increase production of public BARCODE records

Networks, Projects, Organizations

ECBOL

iBOL
INTERNATIONAL BARCODE OF LIFE

FISH-BOL

ABBI

BOLDSYSTEMS

Canadian
Barcode of Life
Network

MBI
Mosquito Barcoding Initiative

TBI

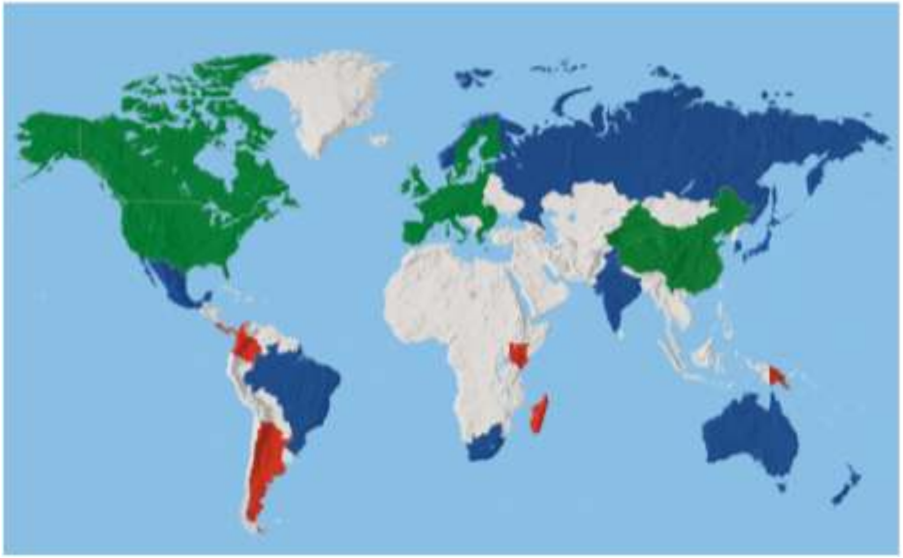
CCDB

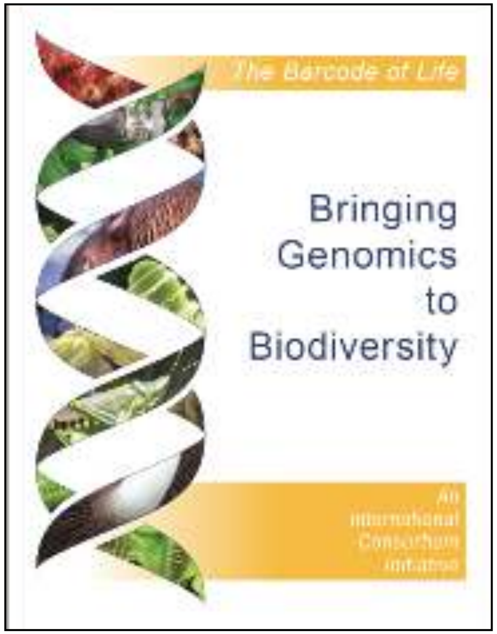
BIO

A global science project

Making Every Species Count

international
BARCODE
OF LIFE



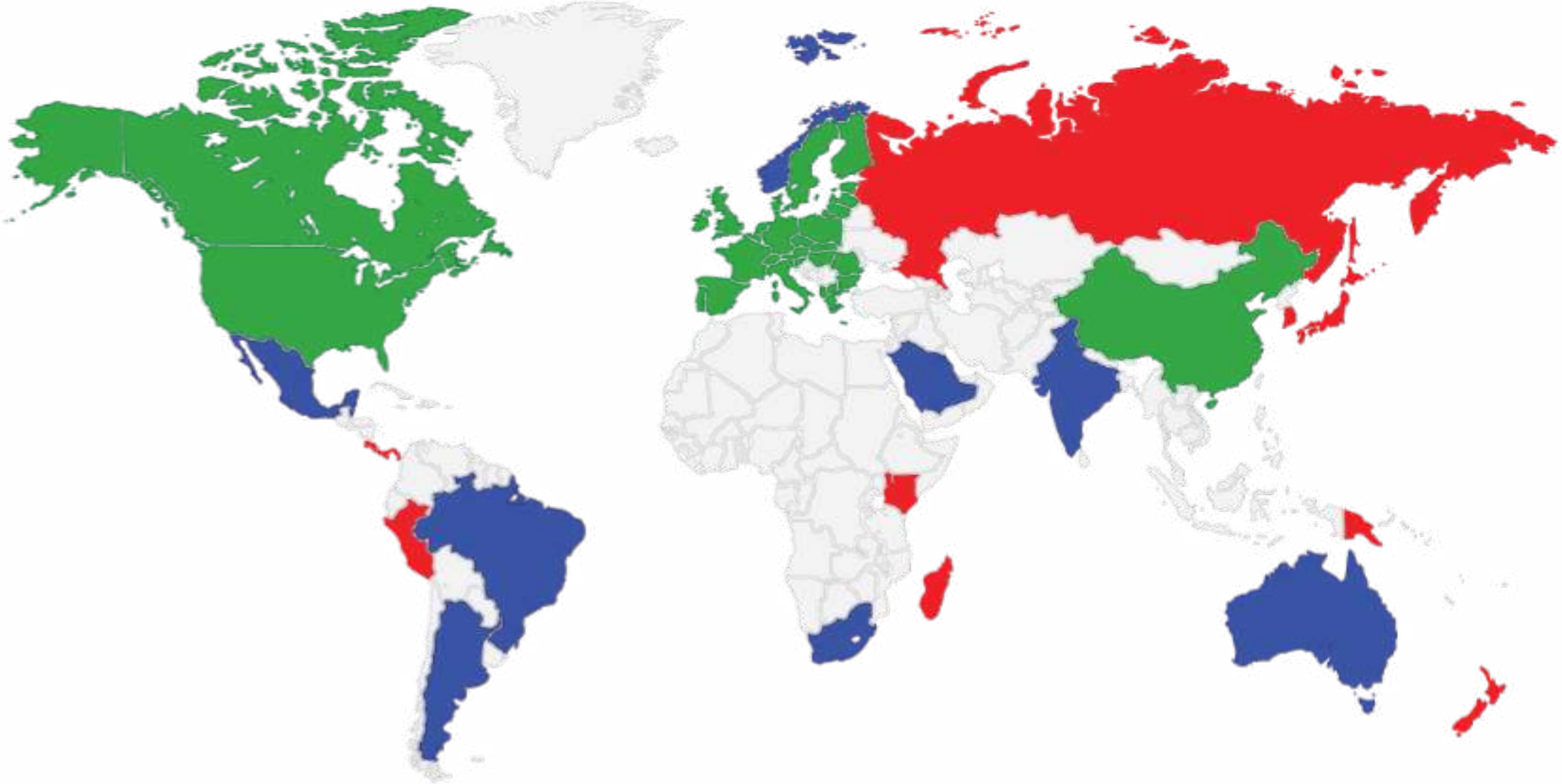


The Barcode of Life

Bringing
Genomics
to
Biodiversity

An
International
Consortium
Initiative

iBOL structure: participating nations

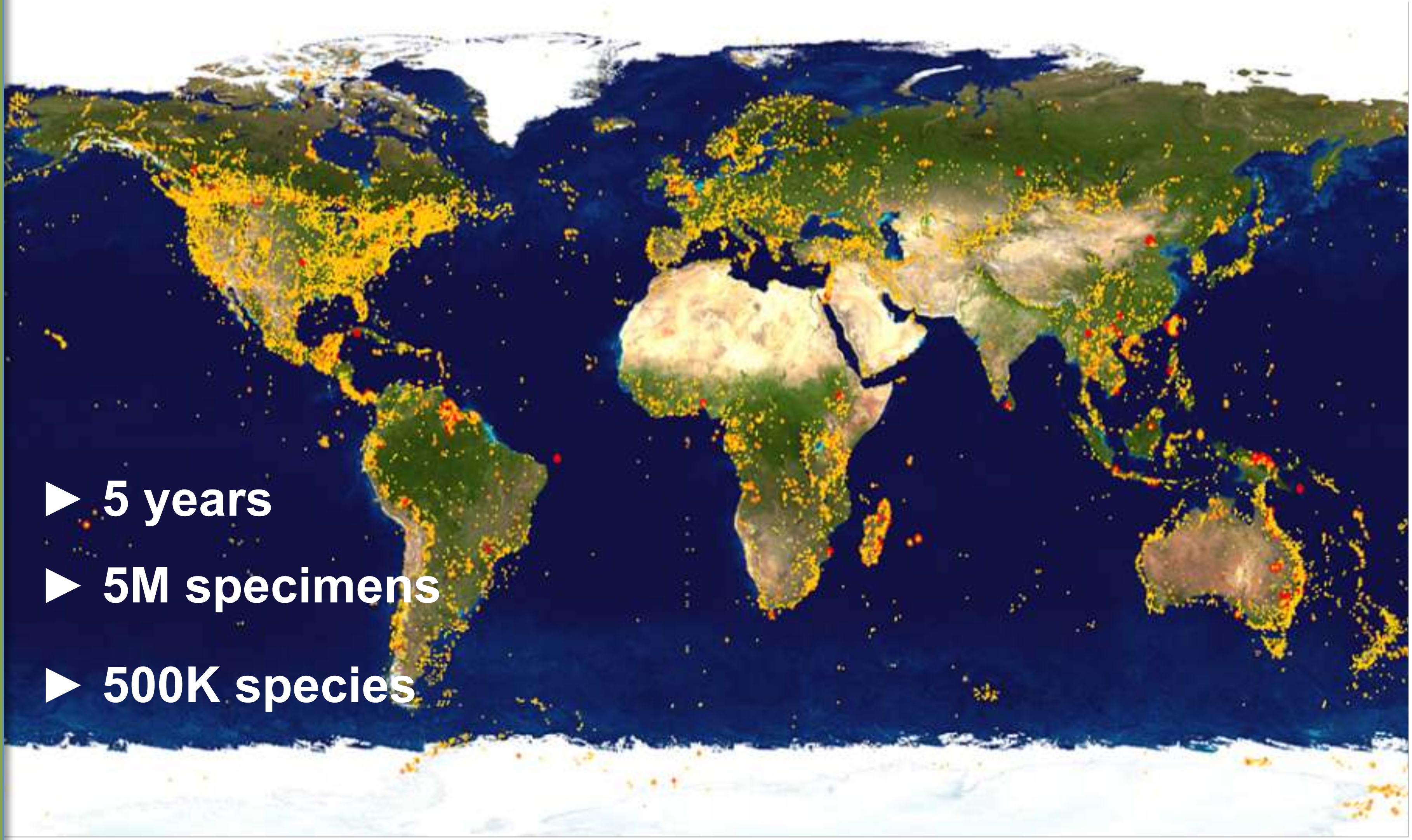


Central Nodes ($25M) **Regional Nodes ($5M)** **National Nodes ($1M)**

Differenti ruoli di ogni nazione

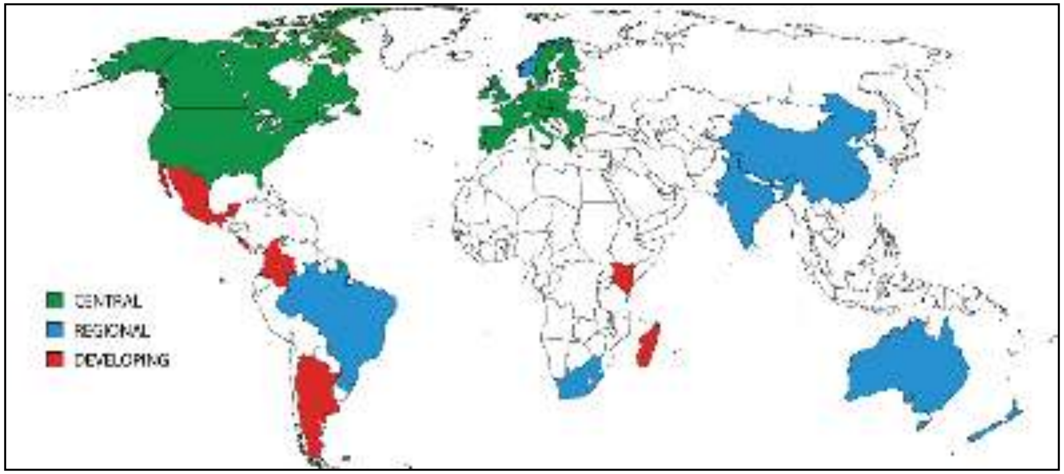
Official launch of iBOL – CN Tower, Toronto, September 25, 2010

iBOL launches with 1M records, 100K species



- ► 5 years
- ► 5M specimens
- ► 500K species

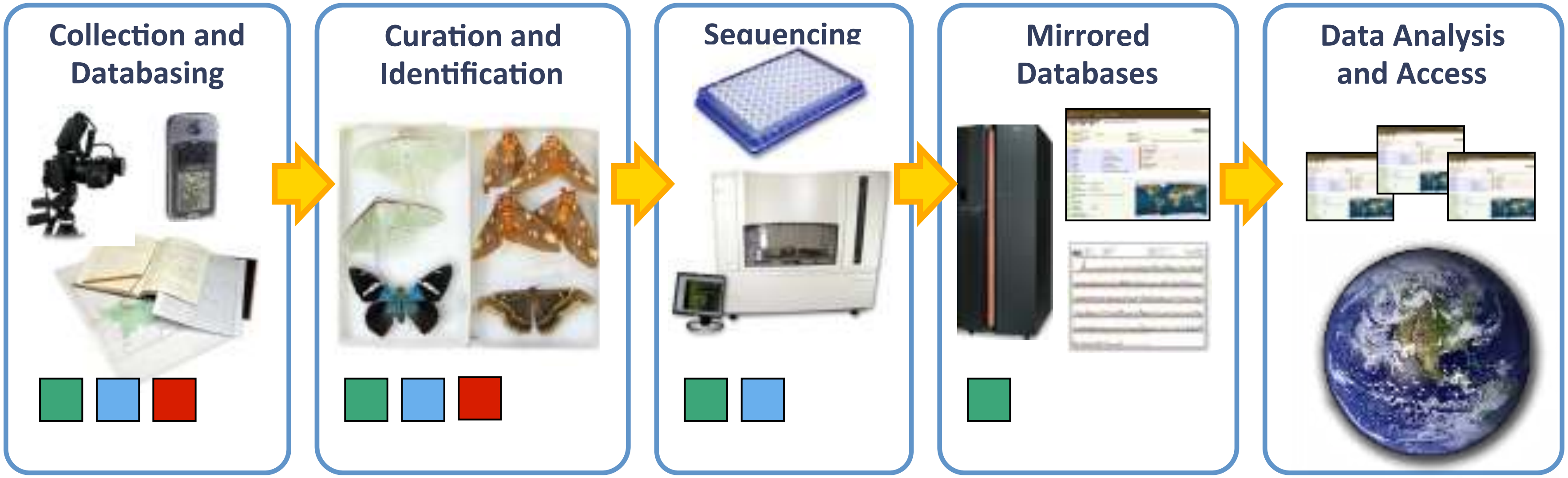
ICI is an alliance of researchers and biodiversity organisations in 2 nations.
All nations active in specimen assembly, curation and data analysis.
Sequencing and informatics support by regional and central nodes.



🟩 **Central Nodes**

🟦 **Regional Nodes**

🟥 **Developing Nodes**



Collection and Databasing
🟩 🟦 🟥

Curation and Identification
🟩 🟦 🟥

Sequencing
🟩 🟦

Mirrored Databases
🟩

Data Analysis and Access

1. **CBOL** (the Consortium for the Barcode Of Life).
2. **iBOL** (International Barcode Of Life).

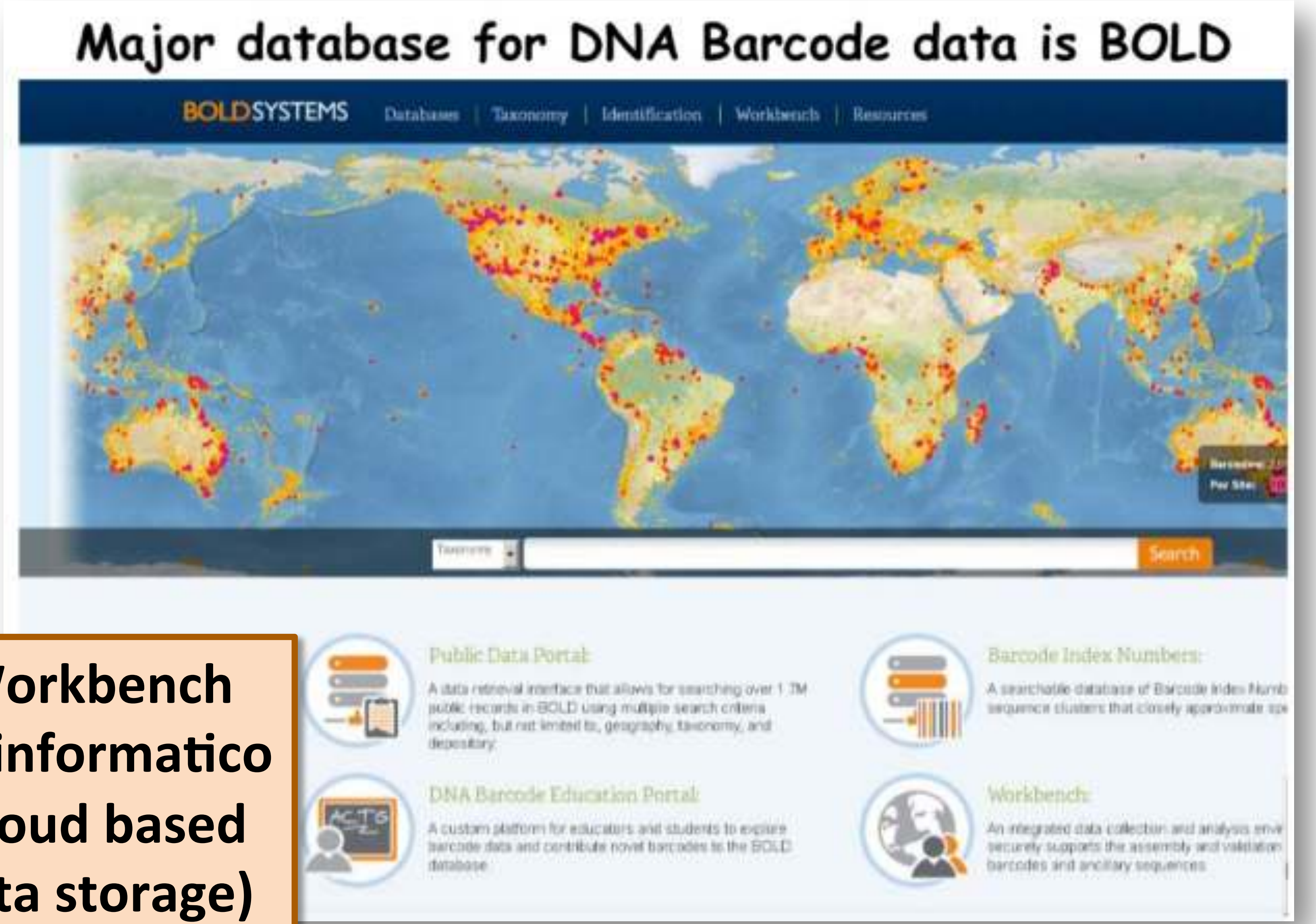
   La più grande iniziativa genomica sulla biodiversità (25 Nazioni).
3. **CCDB** (Canadian Centre for DNA Barcoding).

   Stabilito nel 2004, iniziativa internazionale per lo sviluppo del DNA barcode
4. **GenBank** (Online sequence database)
5. **EOL** (Enciclopedia)
6. **GBIF** (Global Biodiversity Information Facility)
7. **CBG** (Center for Biodiversity Genomics)

1. **CBOL** (the Consortium for the Barcode Of Life).
2. **iBOL** (International Barcode Of Life).
   La più grande iniziativa genomica sulla biodiversità (25 Nazioni).
3. **CCDB** (Canadian Centre for DNA Barcoding).
   Stabilito nel 2004, iniziativa internazionale per lo sviluppo del DNA barcode
4. **GenBank** (Online sequence database)
5. **EOL** (Enciclopedia)
6. **GBIF** (Global Biodiversity Information Facility)
7. **CBG** (Center for Biodiversity Genomics)

BOLD's partners

Barcode Of Life Data (BOLD) system



Major database for DNA Barcode data is BOLD

Public Data Portal:

A data retrieval interface that allows for searching over 1.7M public records in BOLD using multiple search criteria including, but not limited to, geography, taxonomy, and depository.

Barcode Index Numbers:

A searchable database of Barcode Index Numb sequence clusters that closely approximate spe

DNA Barcode Education Portal:

A custom platform for educators and students to explore barcode data and contribute novel barcodes to the BOLD database.

Workbench:

An integrated data collection and analysis envi securely supports the assembly and validation barcodes and ancillary sequences.

Workbench bioinformatico (cloud based data storage)

Barcode Of Life Data (BOLD) system

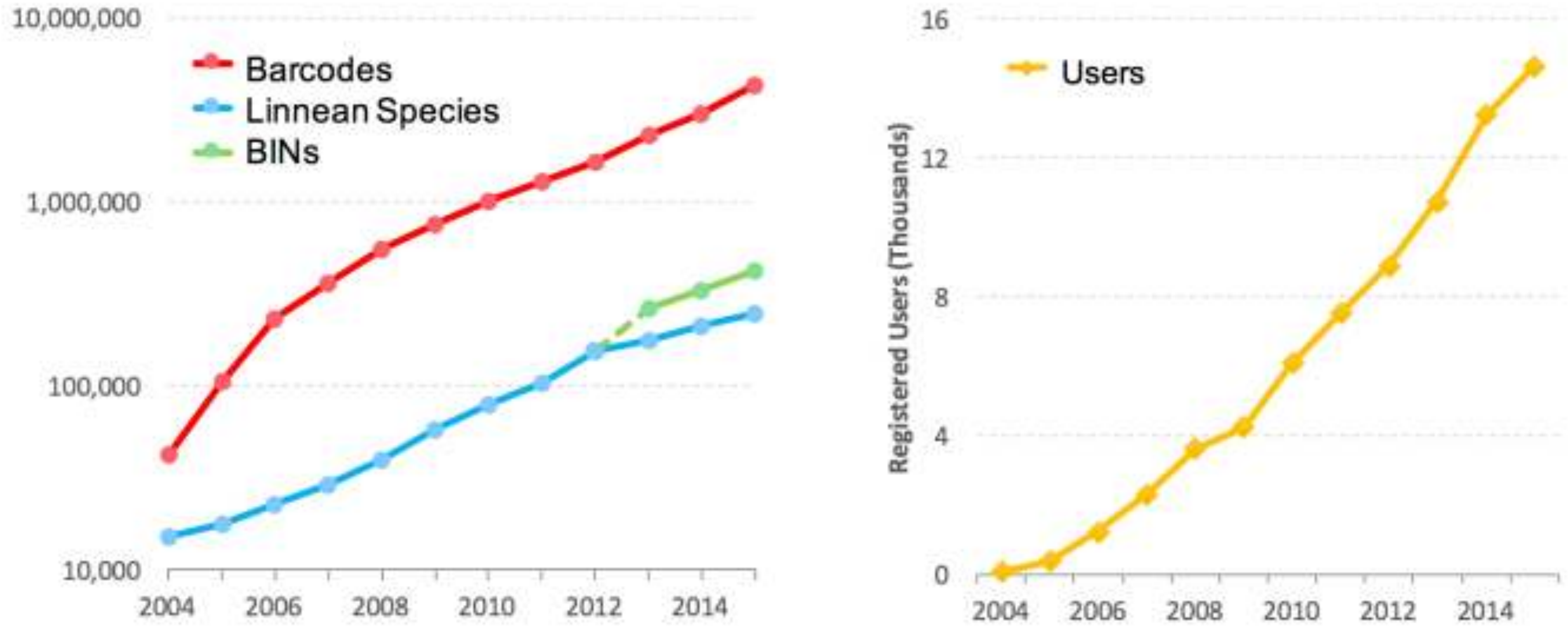
- Sviluppato nel Centre for Biodiversity Genomics in Canada
- Ha lo scopo di acquisire, stoccare, analizzare e pubblicare tutti i DNA barcodes.
- Si occupa inoltre di assemblare tutte le informazioni morfologiche e genetiche disponibili associate ad un determinato barcode.

4 moduli:

1. **Data Portal** = portale con interfaccia grafica che consente la ricerca multicriterio di dati pubblici
2. **Educational portal** = piattaforma educativa per insegnanti e studenti.
3. **Registry of BINs** (putative species) = un database in cui è possibile cercare direttamente i numeri di barcode.
4. **Data collection and analysis workbench** = supporto per l’assemblaggio dei dati e la loro validazione.

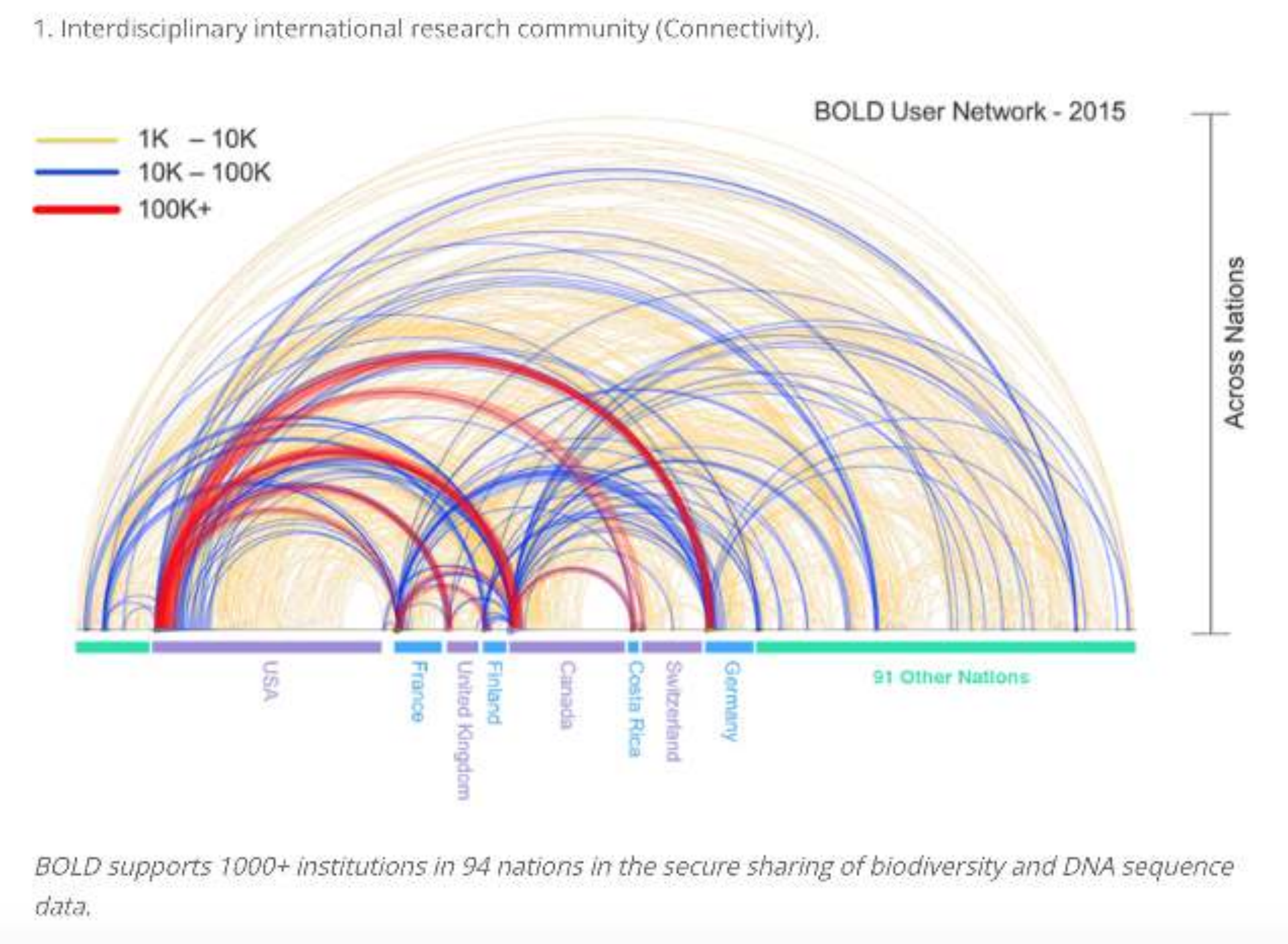
Barcode Of Life Data (BOLD) system

2. Growth of Database (Data Volume & Users)



BOLD has grown to become a key resource in the biodiversity genomics research field as can be seen from the rapid rise in users and volume of data hosted.

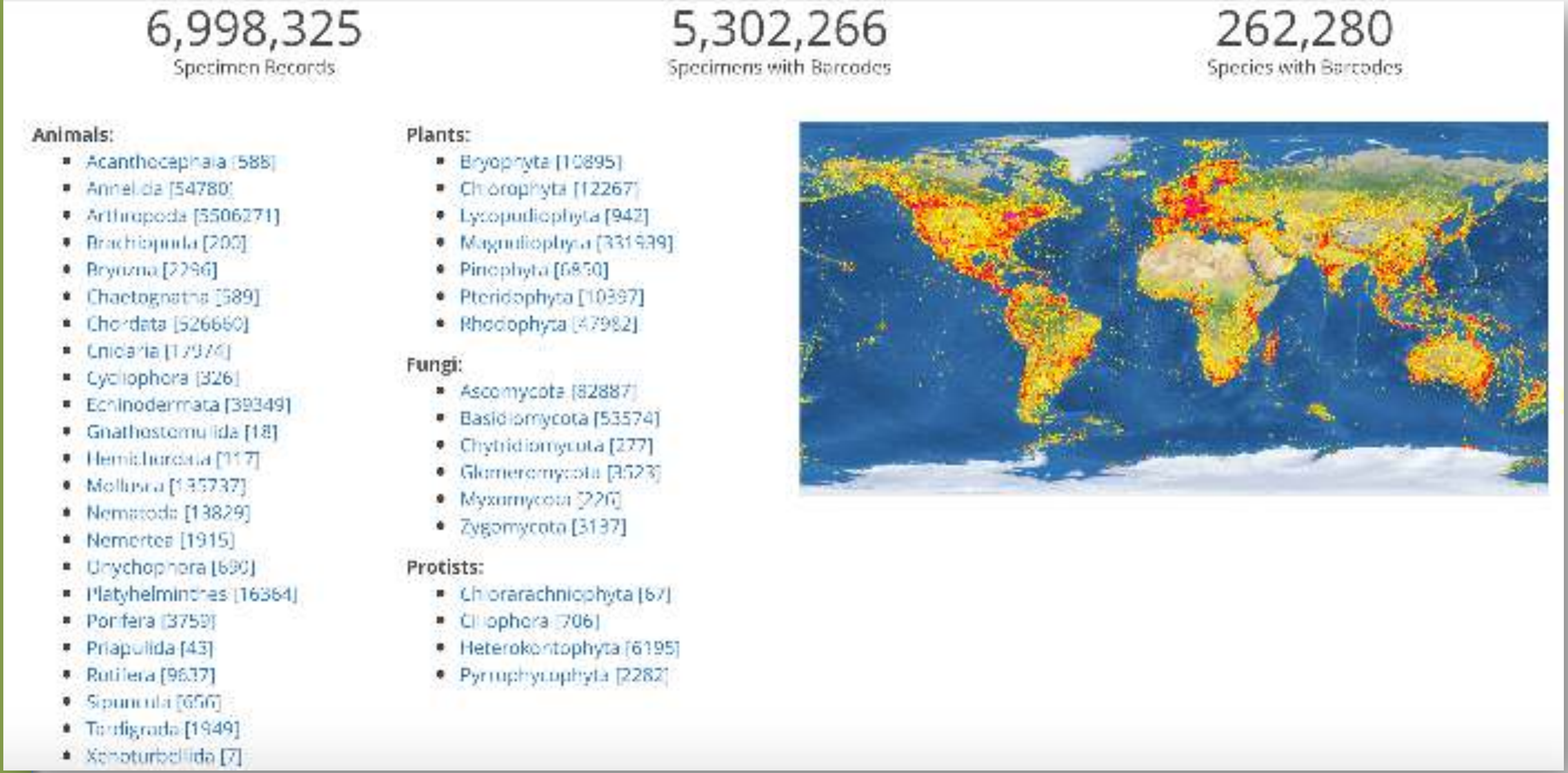
Barcode Of Life Data (BOLD) system



1. Interdisciplinary international research community (Connectivity).

BOLD supports 1000+ institutions in 94 nations in the secure sharing of biodiversity and DNA sequence data.

Barcode Of Life Data (BOLD) system



6,998,325
Specimen Records

5,302,266
Specimens with Barcodes

262,280
Species with Barcodes

Animals:

- Acanthocephala [588]
- Annelida [54780]
- Arthropoda [5506271]
- Brachiopoda [200]
- Bryozoa [2296]
- Chaetognatha [589]
- Chordata [526660]
- Cnidaria [17974]
- Cycliophora [326]
- Echinodermata [39349]
- Gnathostomulida [18]
- Hemichordata [117]
- Mollusca [135737]
- Nematoda [13829]
- Nemertea [1915]
- Onychophora [690]
- Platyhelminthes [16364]
- Porifera [3759]
- Priapulida [43]
- Rotifera [9637]
- Sipuncula [656]
- Tardigrada [1949]
- Xenoturbellida [7]

Plants:

- Bryophyta [10895]
- Chlorophyta [12267]
- Lycopodiophyta [942]
- Magnoliophyta [331939]
- Pinophyta [6850]
- Pteridophyta [10397]
- Rhodophyta [47982]

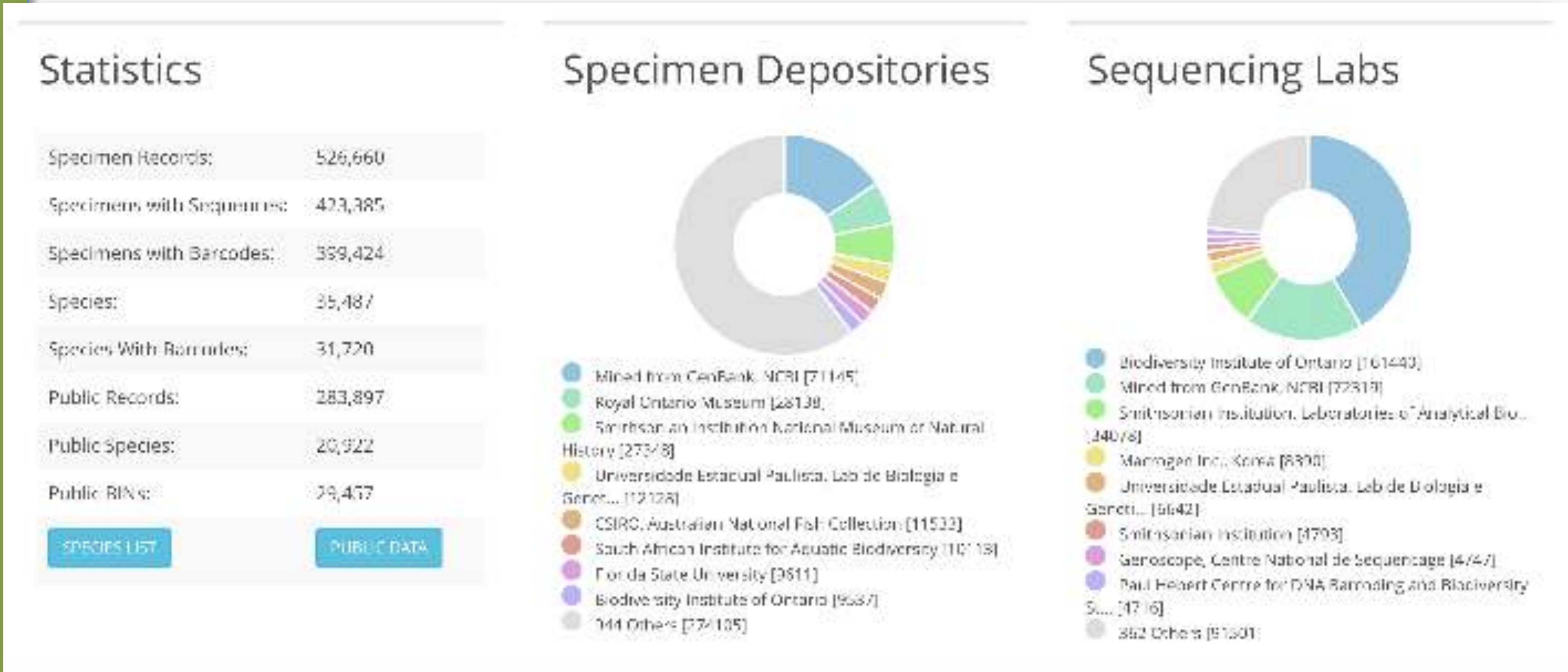
Fungi:

- Ascomycota [82887]
- Basidiomycota [53574]
- Chytridiomycota [277]
- Glomeromycota [3523]
- Myxomycota [226]
- Zygomycota [3137]

Protists:

- Chlorarachniophyta [67]
- Ciliophora [706]
- Heterokontophyta [6195]
- Pyrrophycophyta [2282]

Barcode Of Life Data (BOLD) system



Statistics

Specimen Records:	526,660
Specimens with Sequences:	423,385
Specimens with Barcodes:	399,424
Species:	35,487
Species With Barcodes:	31,720
Public Records:	283,897
Public Species:	20,922
Public BINs:	29,437

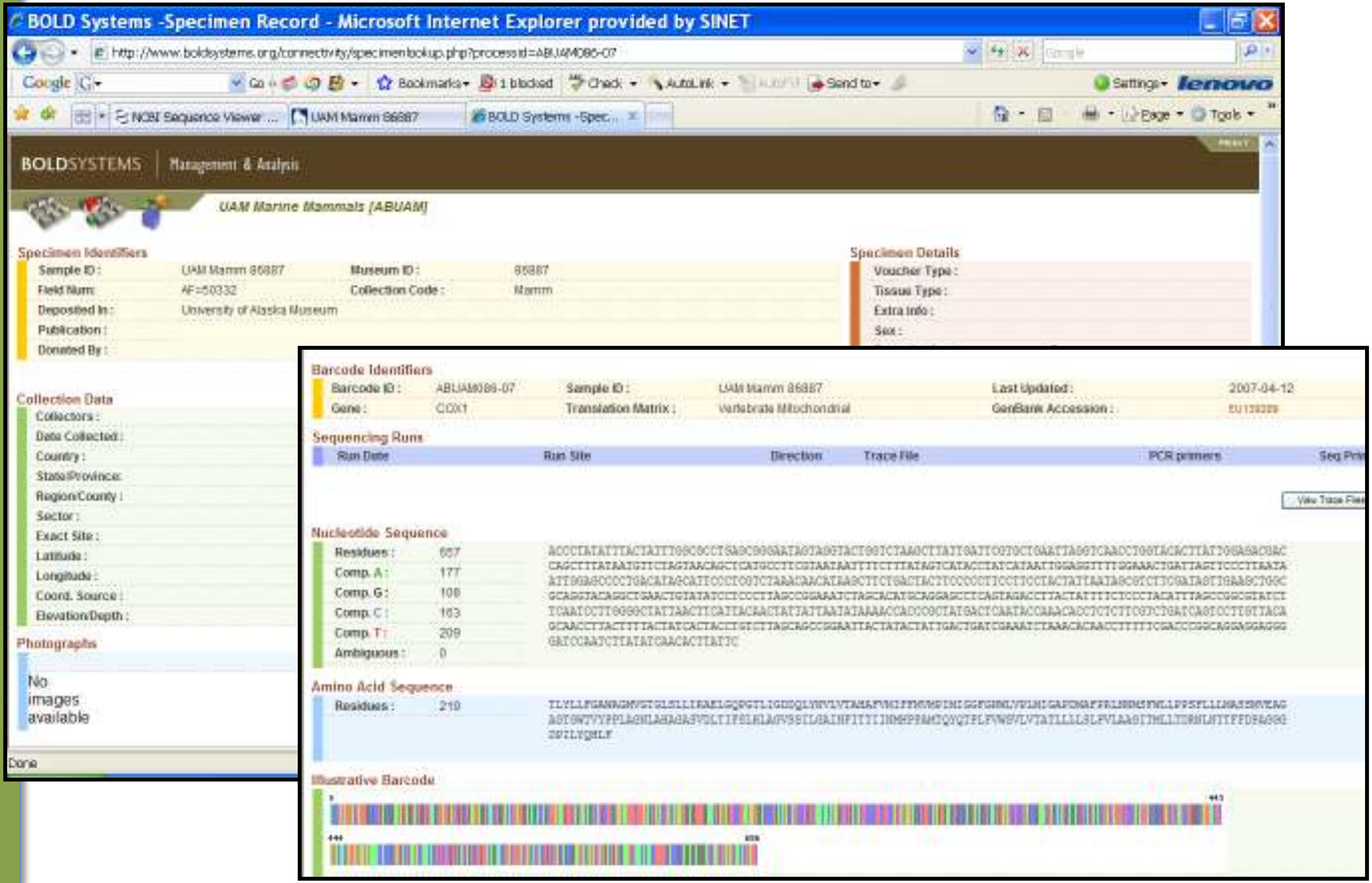
Specimen Depositories

- Mined from GenBank, NCBI [71145]
- Royal Ontario Museum [28139]
- Smithsonian Institution National Museum of Natural History [27348]
- Universidade Estadual Paulista, Lab de Biologia e Genet... [12128]
- CSIRO, Australian National Fish Collection [11533]
- South African Institute for Aquatic Biodiversity [10113]
- Florida State University [9611]
- Biodiversity Institute of Ontario [9537]
- 344 Others [274105]

Sequencing Labs

- Biodiversity Institute of Ontario [161443]
- Mined from GenBank, NCBI [72319]
- Smithsonian Institution, Laboratories of Analytical Bio... [34076]
- Macrogen Inc., Korea [8300]
- Universidade Estadual Paulista, Lab de Biologia e Geneti... [5642]
- Smithsonian Institution [4793]
- Genoscope, Centre National de Sequencage [4747]
- Paul Hebert Centre for DNA Barcoding and Biodiversity S... [4716]
- 352 Others [91501]

Linkout from GenBank to BOLD



UAM Marine Mammals [ABUAM]

Specimen Identifiers

Sample ID :	UAM Mamm 86887	Museum ID :	86887
Field Num:	AF-50332	Collection Code :	Mamm
Deposited In :	University of Alaska Museum		
Publication :			
Donated By :			

Specimen Details

Voucher Type :
Tissue Type :
Extra Info :
Sex :

Collection Data

Collectors :
Date Collected :
Country :
State/Province:
Region/County :
Sector :
Exact Site :
Latitude :
Longitude :
Coord. Source :
Elevation/Depth :

Photographs

No images available

Barcode Identifiers

Barcode ID :	ABUAM086-07	Sample ID :	UAM Mamm 86887	Last Updated :	2007-04-12
Gene :	COX1	Translation Matrix :	Vertebrate Mitochondrial	GenBank Accession :	

Sequencing Runs

Run Date	Run Site	Direction	Trace File	PCR primers	Seq Prim

Nucleotide Sequence

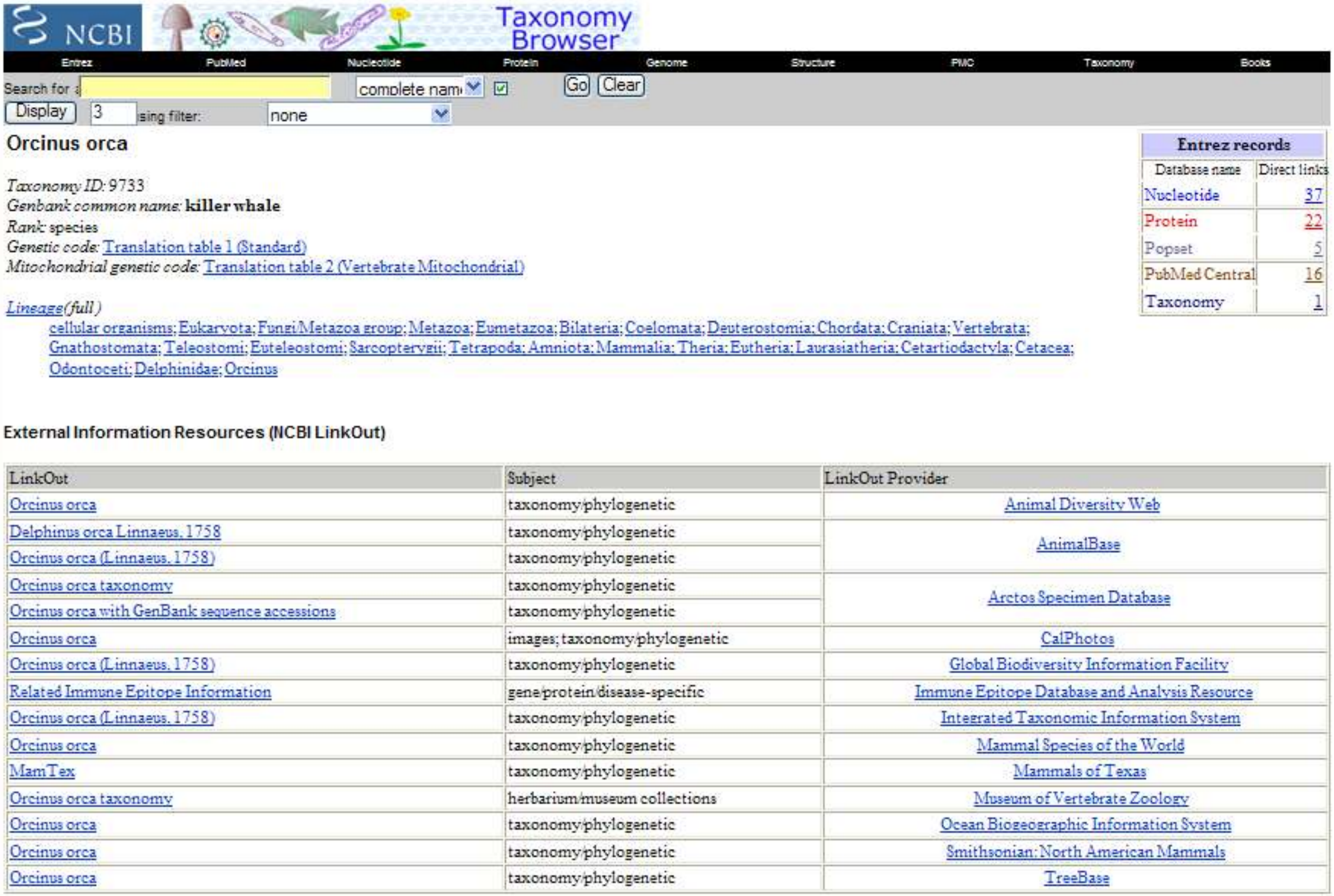
Residues :	657
Comp. A :	177
Comp. G :	108
Comp. C :	163
Comp. T :	209
Ambiguous :	0

Amino Acid Sequence

Residues :	219

Illustrative Barcode

Linkout from GenBank to Taxonomy



Orcinus orca

Taxonomy ID: 9733
Genbank common name: **killer whale**
Rank: species
Genetic code: Translation table 1 (Standard)
Mitochondrial genetic code: Translation table 2 (Vertebrate Mitochondrial)

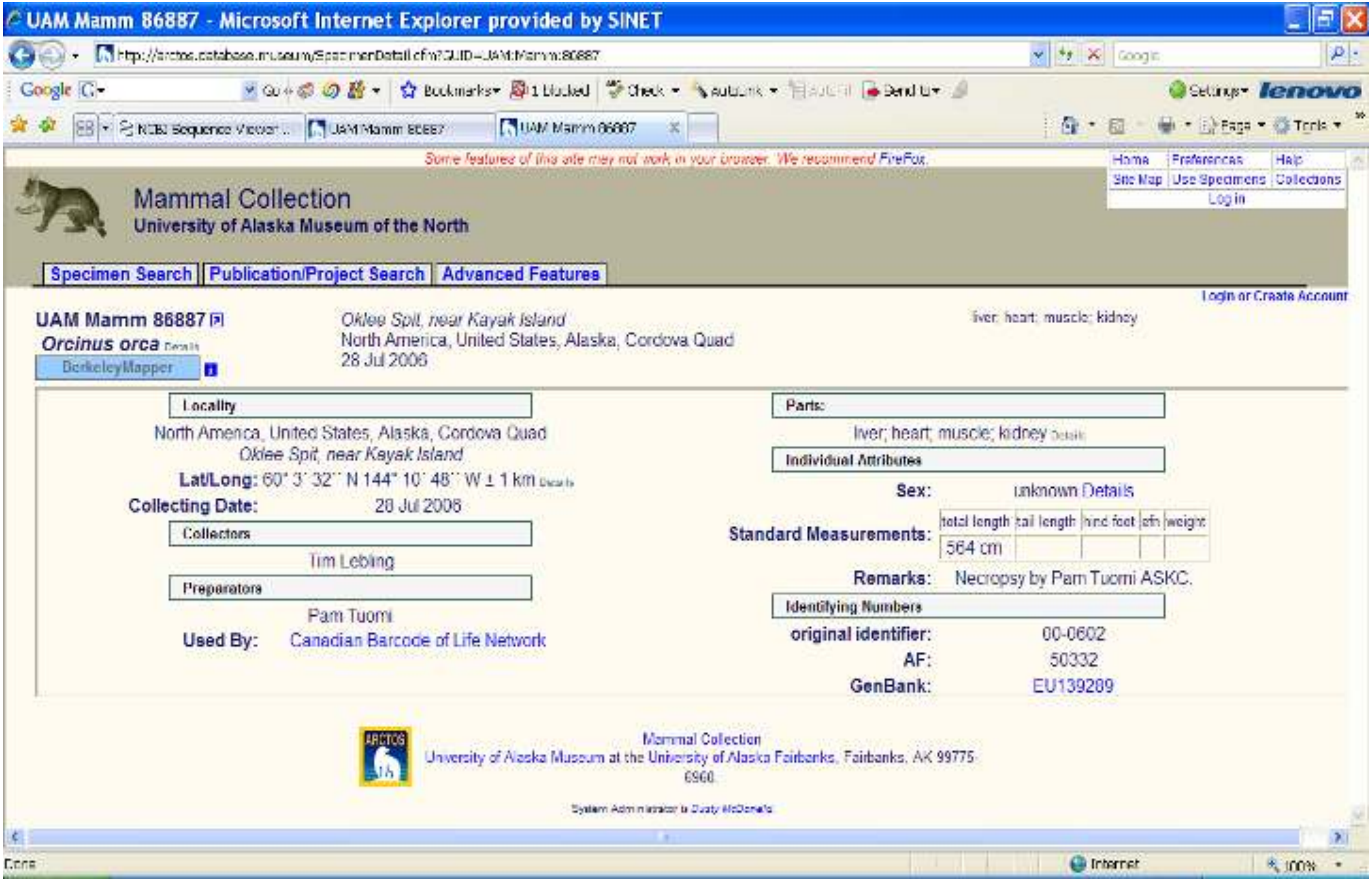
Lineage (full)
cellular organisms; Eukaryota; Fungi/Metazoa group; Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata; Craniata; Vertebrata; Gnathostomata; Teleostomi; Euteleostomi; Sarcopterygii; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Laurasiatheria; Cetartiodactyla; Cetacea; Odontoceti; Delphinidae; Orcinus

Entrez records	
Database name	Direct links
Nucleotide	37
Protein	22
Popset	5
PubMed Central	16
Taxonomy	1

External Information Resources (NCBI LinkOut)

LinkOut	Subject	LinkOut Provider
Orcinus orca	taxonomy/phylogenetic	Animal Diversity Web
Delphinus orca Linnaeus, 1758	taxonomy/phylogenetic	AnimalBase
Orcinus orca (Linnaeus, 1758)	taxonomy/phylogenetic	
Orcinus orca taxonomy	taxonomy/phylogenetic	Arctos Specimen Database
Orcinus orca with GenBank sequence accessions	taxonomy/phylogenetic	
Orcinus orca	images; taxonomy/phylogenetic	CalPhotos
Orcinus orca (Linnaeus, 1758)	taxonomy/phylogenetic	Global Biodiversity Information Facility
Related Immune Epitope Information	gene/protein/disease-specific	Immune Epitope Database and Analysis Resource
Orcinus orca (Linnaeus, 1758)	taxonomy/phylogenetic	Integrated Taxonomic Information System
Orcinus orca	taxonomy/phylogenetic	Mammal Species of the World
MamTex	taxonomy/phylogenetic	Mammals of Texas
Orcinus orca taxonomy	herbarium/museum collections	Museum of Vertebrate Zoology
Orcinus orca	taxonomy/phylogenetic	Ocean Biogeographic Information System
Orcinus orca	taxonomy/phylogenetic	Smithsonian: North American Mammals
Orcinus orca	taxonomy/phylogenetic	TreeBase

Link from GenBank to Museums



Some features of this site may not work in your browser. We recommend FireFox.

Mammal Collection
University of Alaska Museum of the North

UAM Mamm 86887
Orcinus orca Details

Oklee Spit, near Kayak Island
North America, United States, Alaska, Cordova Quad
28 Jul 2006

liver; heart; muscle; kidney

Locality

North America, United States, Alaska, Cordova Quad
Oklee Spit, near Kayak Island

Lat/Long: 60° 3' 32'' N 144° 10' 48'' W ± 1 km Details

Collecting Date: 28 Jul 2006

Collectors

Tim Lebling

Preparators

Pam Tuomi

Used By: Canadian Barcode of Life Network

Parts:

liver; heart; muscle; kidney Details

Individual Attributes

Sex: unknown Details

Standard Measurements:

total length	tail length	hind foot	efn	weight
564 cm				

Remarks: Necropsy by Pam Tuomi ASKC.

Identifying Numbers

original identifier: 00-0602

AF: 50332

GenBank: EU139289

Mammal Collection
University of Alaska Museum at the University of Alaska Fairbanks, Fairbanks, AK 99775-6960.

System Administrator is Dusty McDonald

Es. di ricerca per specie o famiglia

Family name

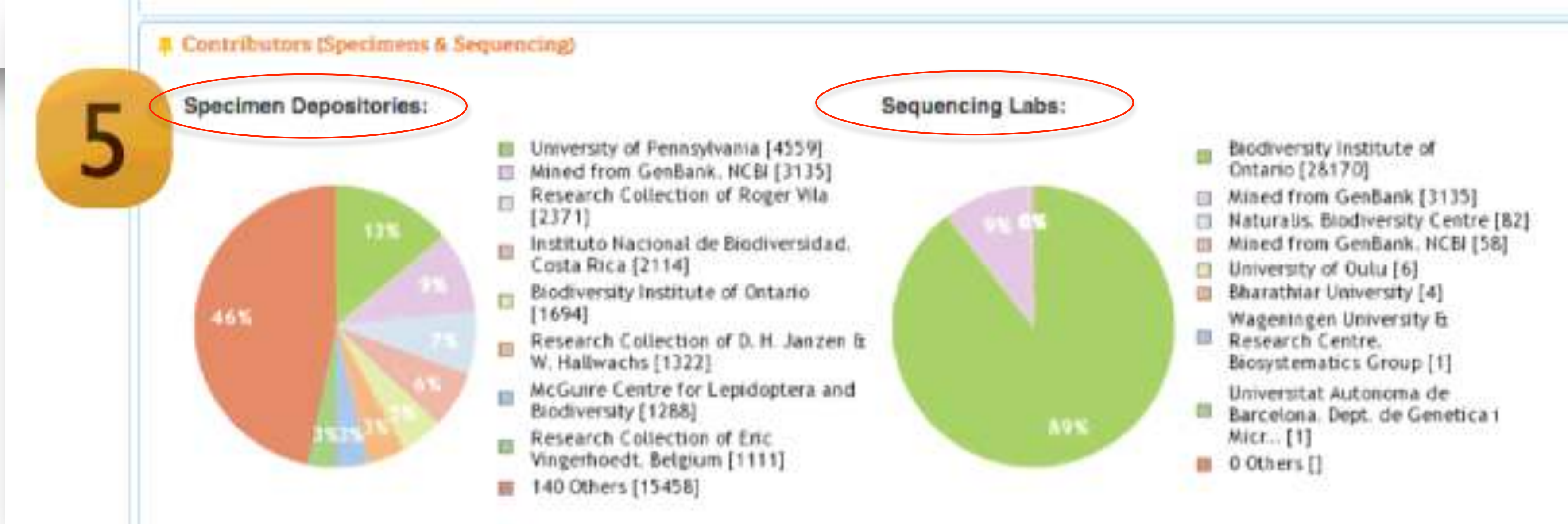
Taxon description

Bold stats

Nymphalidae {family} - Arthropoda; Insecta; Lepidoptera;

Nymphalidae is the largest family of butterflies with about 6,000 species distributed throughout most of the world. These are usually medium sized to large butterflies. Most species have a reduced pair of forelegs and many hold their colourful wings flat when resting. They are also called **brush-footed butterflies** or **four-footed butterflies**. Many species are brightly colored and include popular species such as the emperors, admirals, tortoiseshells, and fritillaries. However, the underwings are in contrast often dull and in some species look remarkably like dead leaves, or are much paler, producing a cryptic effect that helps the surroundings. full article at Wikipedia

Specimen Records:	35,153	Public Records:	11,963
Specimens with Sequences:	31,644	Public Species:	1,916
Specimens with Barcodes:	29,893	Public BINs:	1,474
Species:	4,450		
Species With Barcodes:	4,079		



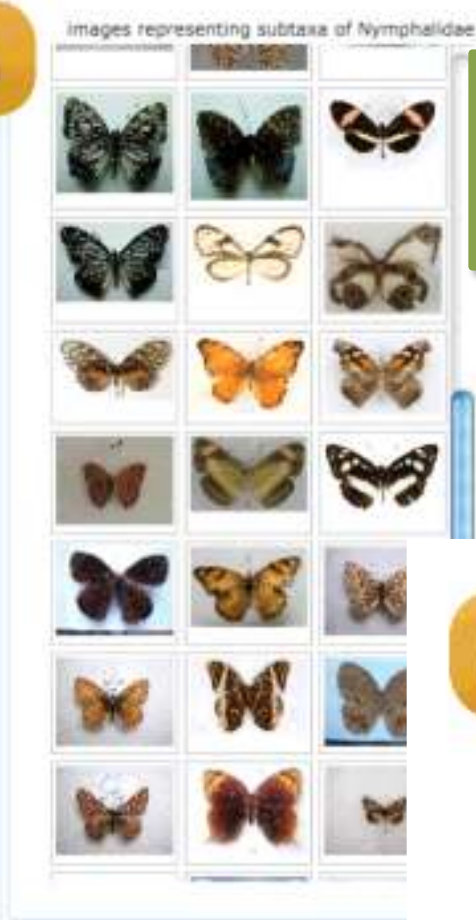
Specimen Depositories:

- 🟩 University of Pennsylvania [4559]
- 🟪 Mined from GenBank, NCBI [3135]
- ☐ Research Collection of Roger Vila [2371]
- 🟥 Instituto Nacional de Biodiversidad, Costa Rica [2114]
- 🟨 Biodiversity Institute of Ontario [1694]
- 🟧 Research Collection of D. H. Janzen & W. Hallwachs [1322]
- 🟦 McGuire Centre for Lepidoptera and Biodiversity [1288]
- 🟩 Research Collection of Eric Vingerhoedt, Belgium [1111]
- 🟥 140 Others [15458]

Sequencing Labs:

- 🟩 Biodiversity Institute of Ontario [28170]
- ☐ Mined from GenBank [3135]
- ☐ Naturalis Biodiversity Centre [82]
- 🟥 Mined from GenBank, NCBI [58]
- 🟨 University of Oulu [6]
- 🟧 Bharathiar University [4]
- 🟦 Wageningen University & Research Centre, Biosystematics Group [1]
- 🟩 Universitat Autonoma de Barcelona, Dept. de Genetica i Micr... [1]
- 🟥 0 Others []

6



Imaggini di campioni di riferimento di altri subtaxa

7

Sample ID: INB0004226627

Vaucher

8

Collected from 90 countries. Top 20: Show All Countries

Costa Rica	6581	Democratic Republic of the Congo	756	Papua New Guinea
Canada	2482	Brazil	746	Germany
United States	1558	Romania	644	Italy

Siti di campionamento per la specie vaucher

Es. di ricerca BIN

Barcode Index Number Registry For BOLD:AAB5726

Back to Last Page

N URI:	BOLD:AAB5726	Average Distance:	0.58% (p-dist)
:	Pending	Maximum Distance:	2.19% (p-dist)
mber Count:	83 [74 Public]	Distance to Nearest Neighbor:	1.76% (p-dist)
rcode Compliant Members:	56		
Founding Record:			

NEAREST NEIGHBOR (NN) DETAILS:

Nearest BIN URI:	BOLD:ABZ4800	Average Distance:	0.17% (p-dist)
Member Count:	2	Maximum Distance:	0.17% (p-dist)
Nearest Member:	GBCH1066-06	Distance Variance:	0% (p-dist)
Nearest Member Taxonomy:	Arthropoda, Arachnida, Araneae, Agelenidae, Agelenopsis, Agelenopsis kastoni		

	da [83]
ss:	Arachnida [83]
er:	Araneae [80]
	Opiliones [3]
Family:	Agelenidae [77]
Subfamily:	
Genus:	Agelenopsis [75]
Species:	Agelenopsis potteri [64]

Agelenopsis potteri Details

Tags: Validated

Comments:
Renee Labbee: Specimen ID verified. Sequence matching.

You have access to 17 of these records.

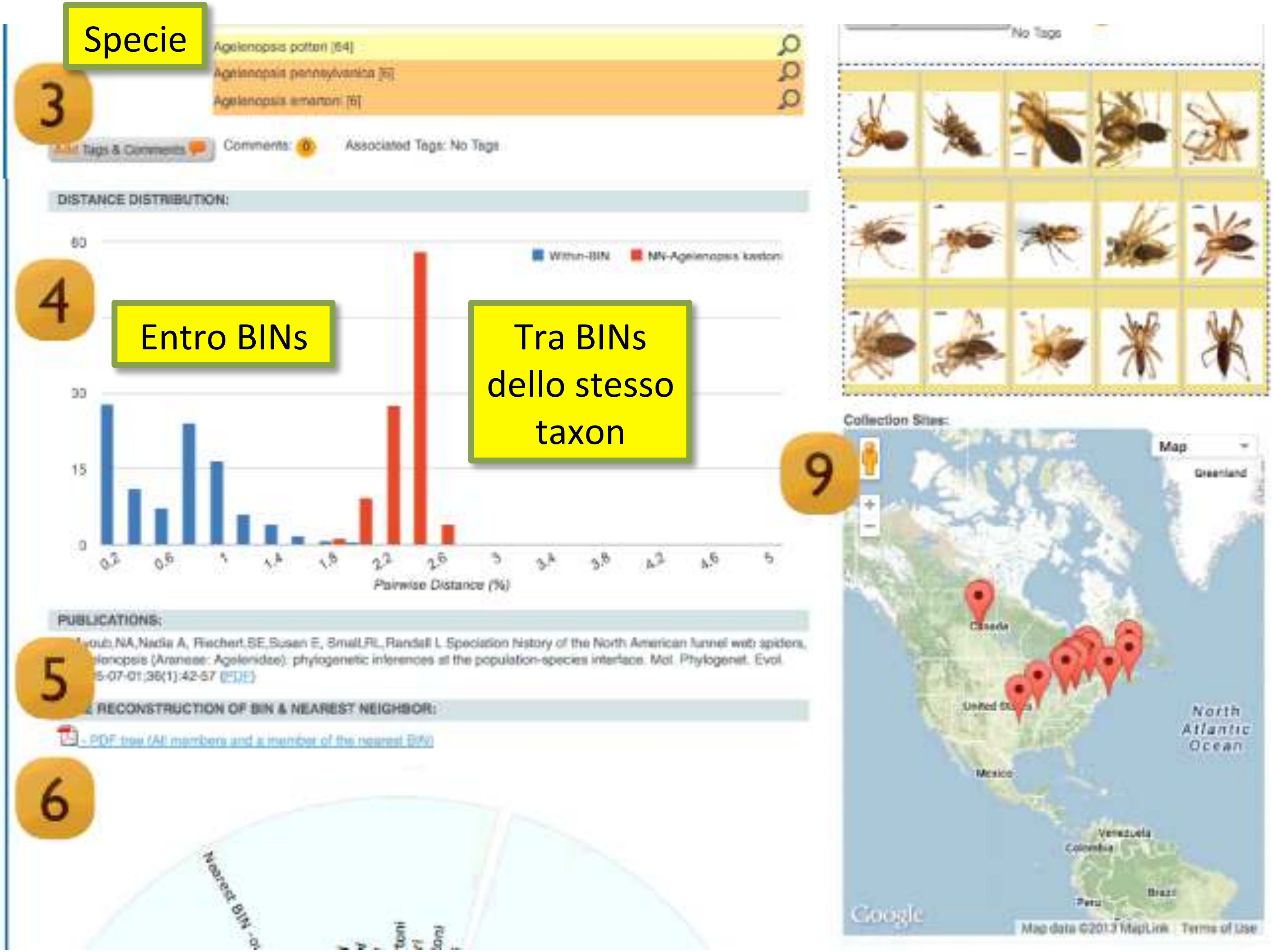
View Records

BIN COMPLIANT WITH METADATA REQUIREMENTS

Specimen Images:

enopsis potteri)
Commons - Attribution Non-Commercial
Like BY-NC-SA
otography Group, Biodiversity Institute of

Comments: 0
Associated Tags:
No Tags



Specie

Agelenopsis potteri [64]

Agelenopsis pennsylvanica [6]

Agelenopsis emertoni [6]

Comments: 0 Associated Tags: No Tags

DISTANCE DISTRIBUTION:

Entro BINs

Tra BINs dello stesso taxon

PUBLICATIONS:

youb,NA,Nadia A, Riechert,SE,Susan E, Small,RL,Randall L Speciation history of the North American funnel web spiders, elenopsis (Araneae: Agelenidae): phylogenetic inferences at the population-species interface. Mol. Phylogenet. Evol. 5-07-01;36(1):42-57 (PDF)

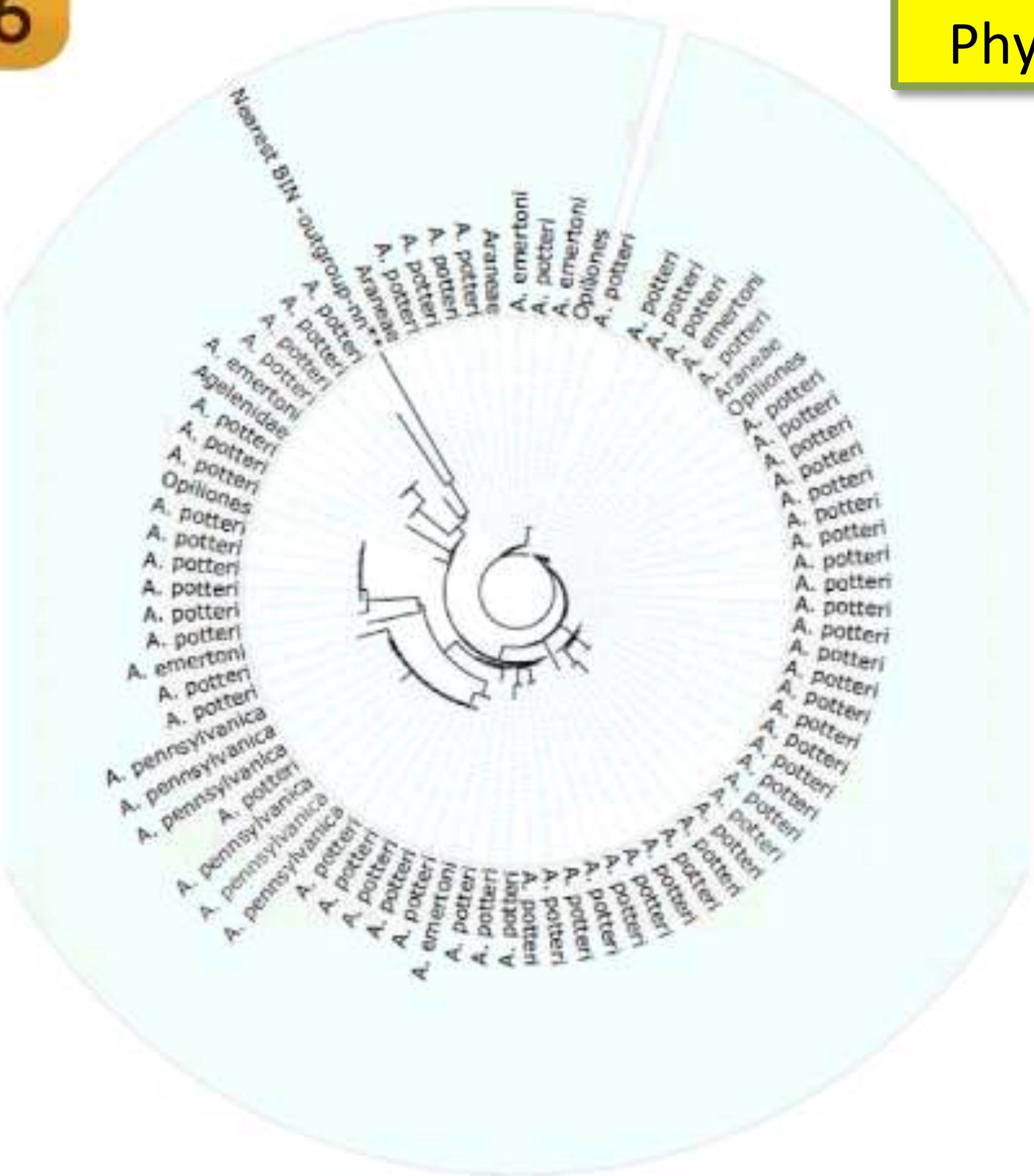
RECONSTRUCTION OF BIN & NEAREST NEIGHBOR:

- PDF tree (All members and a member of the nearest BIN)

Collection Sites:

6

Phylogeny


Nearest BIN -outgroup-nn**
Araneae
A. potteri
A. potteri
A. potteri
A. potteri
A. potteri
Araneae
A. emertoni
A. potteri
A. emertoni
Opiliones
A. potteri
A. potteri
A. potteri
A. potteri
A. emertoni
A. potteri
Araneae
Opiliones
A. potteri
A. potteri
A. potteri
A. potteri
A. potteri
A. potteri
A. potteri
A. potteri
A. potteri
A. potteri
A. potteri
A. potteri
A. potteri
A. potteri
A. potteri
A. potteri
A. potteri
A. potteri
A. potteri
A. potteri
A. potteri
A. potteri
A. potteri
A. potteri
A. potteri
A. potteri
A. emertoni
A. potteri
A. potteri
A. potteri
A. potteri
A. pennsylvanica
A. pennsylvanica
A. pennsylvanica
A. potteri
A. pennsylvanica
A. pennsylvanica
A. pennsylvanica
A. potteri
A. potteri
A. emertoni
A. potteri
A. potteri
A. potteri
A. potteri
A. potteri
A. potteri
Opiliones
A. potteri
A. potteri
A. potteri
Agelenidae
A. emertoni
A. potteri
A. potteri
A. potteri
A. potteri

Collection sites

8

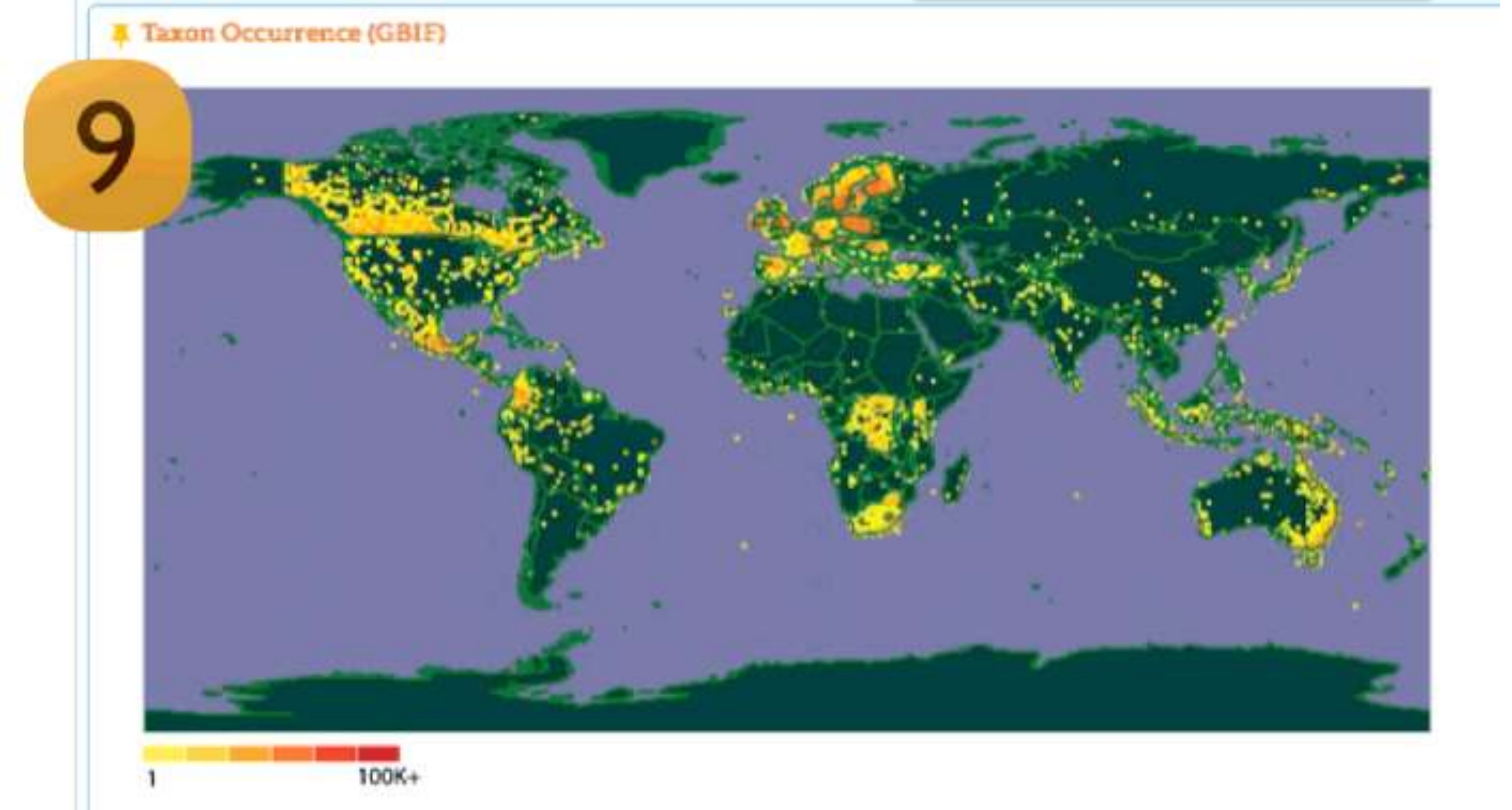
Collection Sites

Collected from 90 countries. Top 20: Show All Countries

Costa Rica	6581	Democratic Republic of the Congo	756	Papua New Guinea	320
Canada	2482	Brazil	746	Germany	317
United States	1558	Romania	644	Italy	294
Mexico	1436	China	506	Cameroon	289
Spain	1417	Australia	438	Malaysia	272
Russia	1003	Panama	350	Kenya	251
Argentina	879	Madagascar	334		

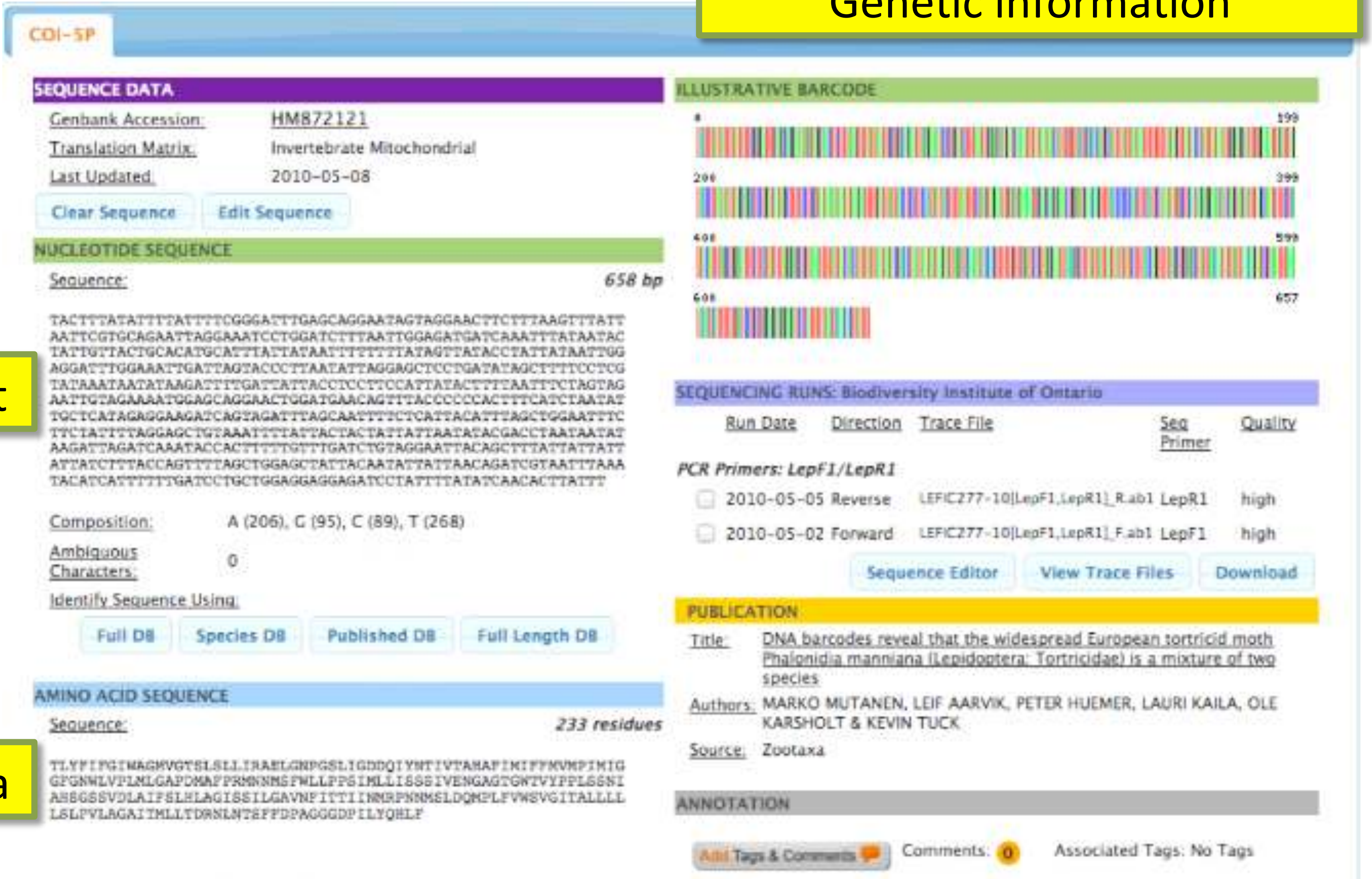
Taxon occurence

9



Taxon Occurrence (GBIF)

Genetic information



SEQUENCE DATA

Genbank Accession: HM872121
Translation Matrix: Invertebrate Mitochondrial
Last Updated: 2010-05-08

NUCLEOTIDE SEQUENCE

nt

Sequence: *658 bp*

```
TACTTTATATTTTATTTTCGGGATTTGAGCAGGAATAGTAGGAACTTCTTTAAGTTTATT
AATTCGTGCAGAATTAGGAAATCCTGGATCTTTAATTGGAGATGATCAAATTTATAATAC
TATTGTTACTGCACATGCATTTATTATAATTTTTTTTATAGTTATACCTATTATAATTGG
AGGATTTGGAAATTGATTAGTACCCTTAATATTAGGAGCTCCTGATATAGCTTTTCCTCG
TATAAATAATATAAGATTTTGATTATTACCTCCTTCCATTATACTTTTAATTTCTAGTAG
AATTGTAGAAAATGGAGCAGGAACTGGATGAACAGTTTACCCCCCACTTTCATCTAATAT
TGCTCATAGAGGAAGATCAGTAGATTTAGCAATTTTCTCATTACATTTAGCTGGAATTTC
TTCTATTTTAGGAGCTGTAAATTTTATTACTACTATTATTAATATACGACCTAATAATAT
AAGATTAGATCAAATACCACTTTTTGTTTGATCTGTAGGAATTACAGCTTTATTATTATT
ATTATCTTTACCAGTTTTAGCTGGAGCTATTACAATATTATTAACAGATCGTAATTTAAA
TACATCATTTTTTGATCCTGCTGGAGGAGGAGATCCTATTTTATATCAACACTTATTT
```

Composition: A (206), G (95), C (89), T (268)
Ambiguous Characters: 0
Identify Sequence Using:

AMINO ACID SEQUENCE

aa

Sequence: *233 residues*

```
TLYFIFGIWAGMVGTSLSLLIRAELGNPGSLIGDDQIYNTIVTAHAFIMIFFMVMPIMIG
GFGNWLVPLMLGAPDMAFPRMNNMSFWLLPPSIMLLISSSIVENGAGTGWTVYPPLSSNI
AHSGSSVDLAIFSLHLAGISSILGAVNFITTIINMRPNNMSLDQMPLFVWSVGITALLLL
LSLPVLAGAITMLLTDRNLNTSFFDPAGGGDPILYQHLF
```

ILLUSTRATIVE BARCODE

SEQUENCING RUNS: Biodiversity Institute of Ontario

	Run Date	Direction	Trace File	Seq Primer	Quality
PCR Primers: LepF1/LepR1					
☐	2010-05-05	Reverse	LEFIC277-10[LepF1,LepR1]_R.ab1	LepR1	high
☐	2010-05-02	Forward	LEFIC277-10[LepF1,LepR1]_F.ab1	LepF1	high

PUBLICATION

Title: DNA barcodes reveal that the widespread European tortricid moth Phalonidia manniana (Lepidoptera: Tortricidae) is a mixture of two species
Authors: MARKO MUTANEN, LEIF AARVIK, PETER HUEMER, LAURI KAILA, OLE KARSHOLT & KEVIN TUCK
Source: Zootaxa

ANNOTATION

Comments: 0 Associated Tags: No Tags