Birds, planes, and DNA Barcoding



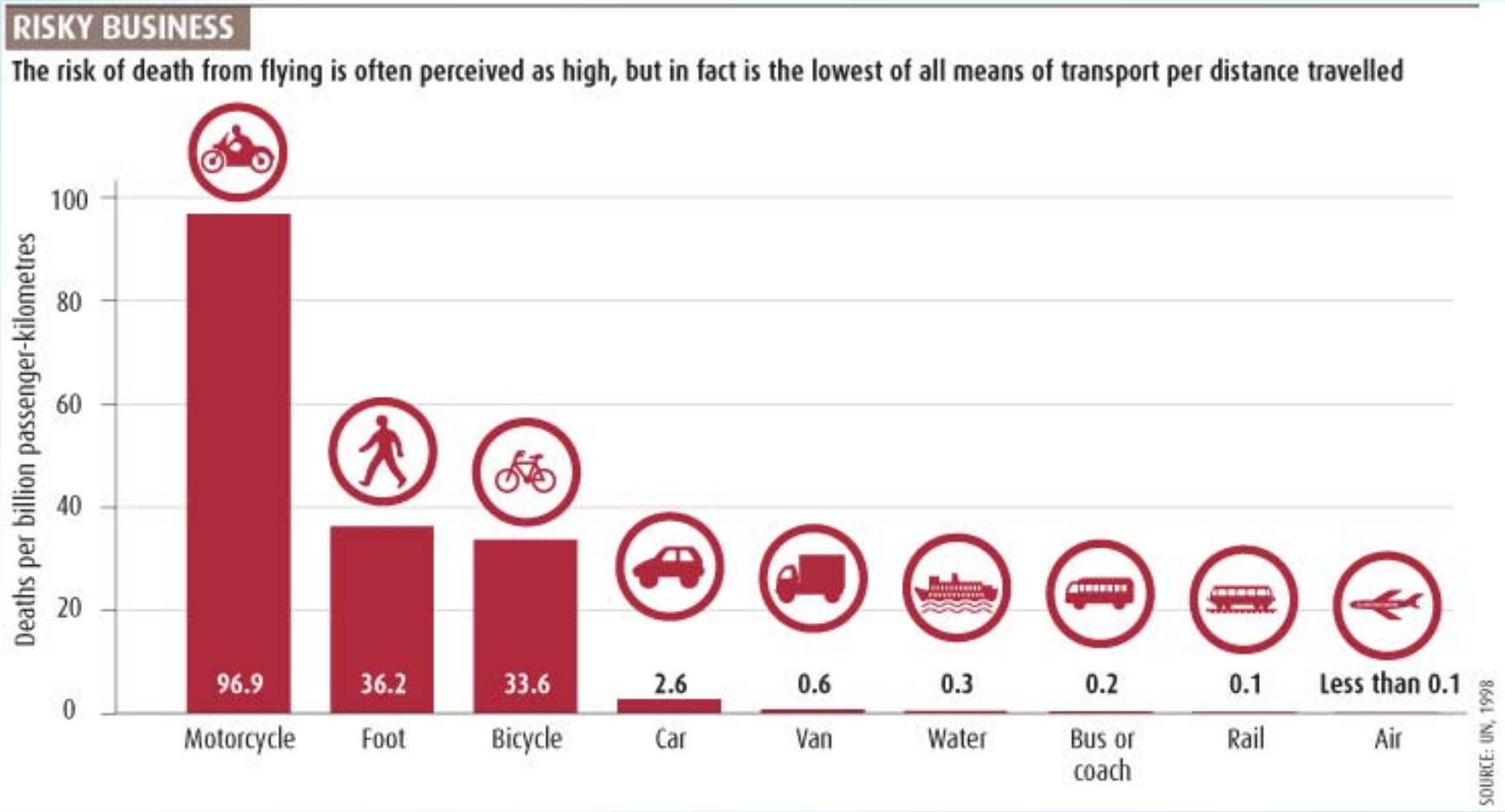


Evgeny V. Zakharov

Canadian Centre for DNA Barcoding
University of Guelph

October 08th, 2014
Ingersol, ON

The safest mode to travel



…and yet

and again…

News

Passenger: Flight From Westchester Co. Airport ‘Started Swerving’ After Bird Strike

April 25, 2012 7:30 PM

Share 40 Tweet 4 Share 14

JetBlue Flight 571 had to make an emergency landing at Westchester County Airport after encountering a bird strike Tuesday evening. (credit: CBS 2).

CBS New York Buy Tickets Order Food More FOLLOW US LOGIN

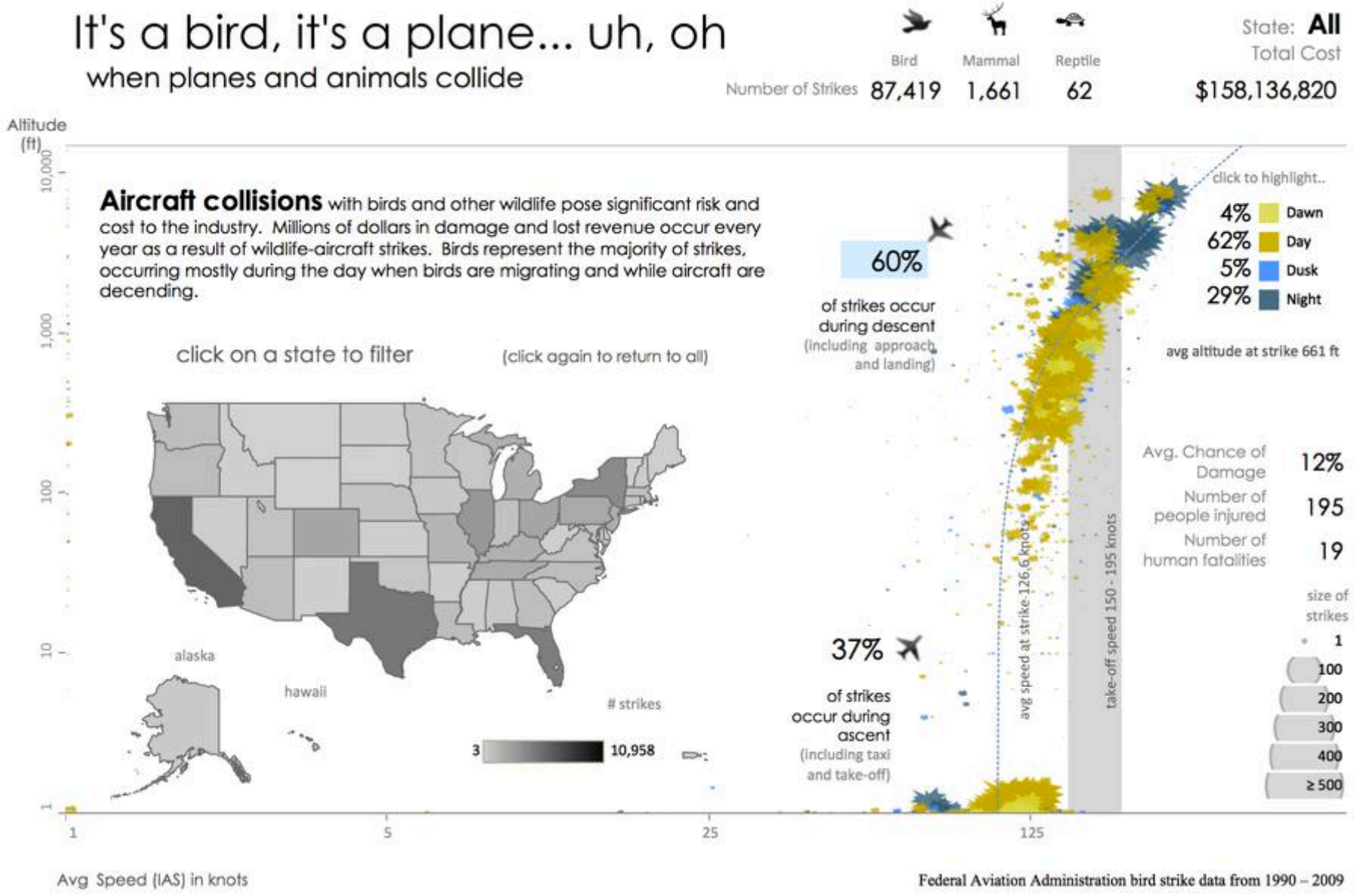
ExpressJet Airlines 5400, which took off from Greensboro, North Carolina, landed at the airport safely after reporting a bird strike around 7:18 a.m. There’s evidence a bird struck the aircraft’s nose gear, according to the FAA.

The incidents all occurred on or near Runway 31, WCBS 880’s Marla Diamond reported.

No damage or injuries were reported, said The Port Authority of New York and New Jersey, which runs the airport.

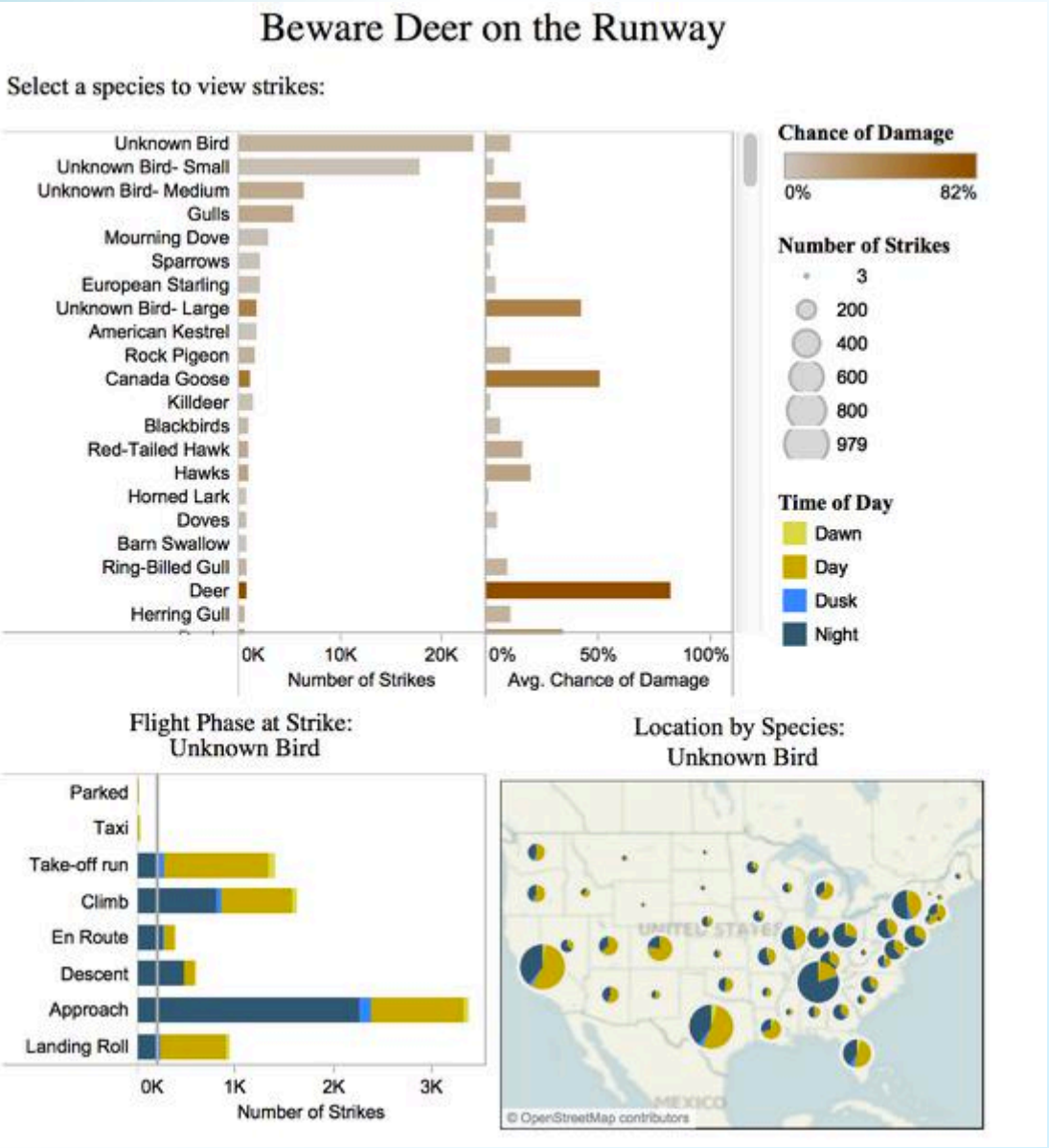
The FAA is investigating and the Smithsonian will examine the remains, Papa reported.

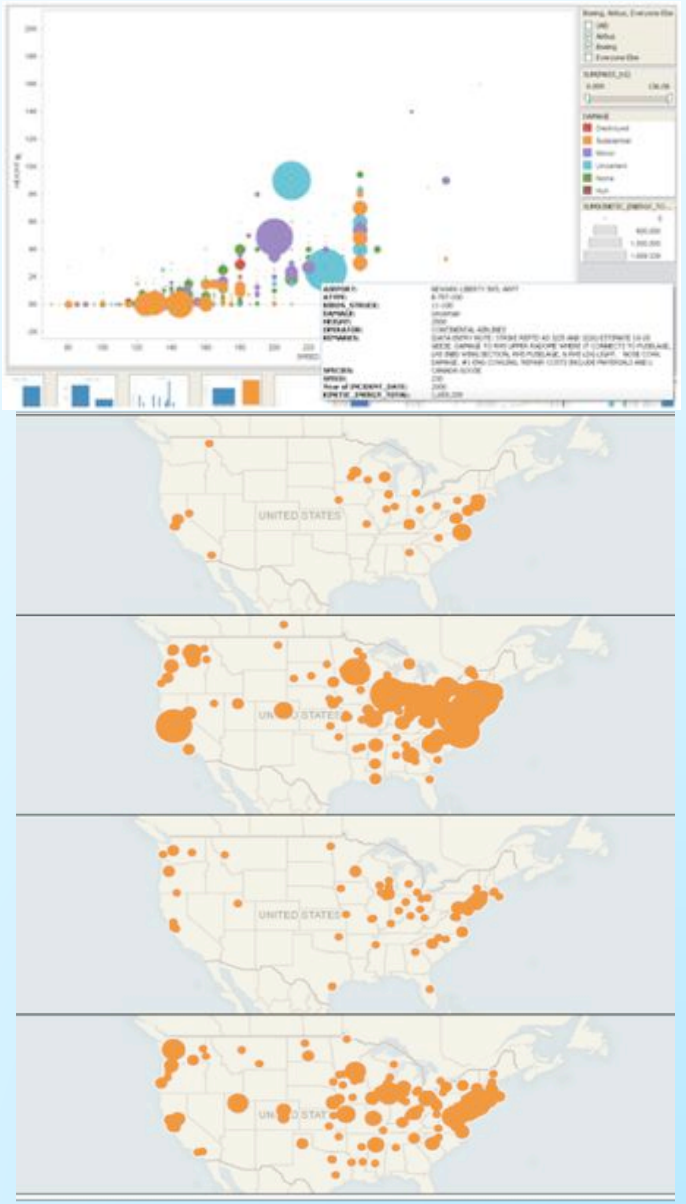
Understanding Birdstrike Occurrence Data



Author: Kelly Martin (http://www.tableausoftware.com/public/gallery/when-birds-and-planes-collide)

Understanding Birdstrike Occurrence Data





Author: Ellie Fields (http://www.tableausoftware.com/public/gallery/bird-strikes)

Dealing with the unknowns

Dealing with the unknowns

Tools and Technology Article

Using DNA Barcodes to Identify Bird Species Involved in Birdstrikes

CARLA J. DOVE,[1] *Smithsonian Institution, Department of Vertebrate Zoology, Mail Routing Code 116, P.O. Box 37012, Washington, D.C. 20013-7012, USA*

NANCY C. ROTZEL,[2] *Smithsonian Institution, Department of Vertebrate Zoology, Mail Routing Code 116, P.O. Box 37012, Washington, D.C. 20013-7012, USA*

MARCY HEACKER, *Smithsonian Institution, Department of Vertebrate Zoology, Mail Routing Code 116, P.O. Box 37012, Washington, D.C. 20013-7012, USA*

LEE A. WEIGT, *Smithsonian Institution, Laboratories of Analytical Biology, Mail Routing Code 543, Suitland, MD 20746, USA*

ABSTRACT We determined effectiveness of using mitochondrial DNA barcodes (*cytochrome c oxidase subunit 1* [CO1]) to identify bird–aircraft collision (birdstrike) cases that lacked sufficient feather evidence for morphological diagnosis. From September through December 2006, 821 samples from birdstrike events occurring in the United States were submitted for DNA analysis. We successfully amplified a CO1 DNA barcode product from 554 (67.5%) of the samples; 267 (32.5%) did not contain viable DNA and depended on morphological methods (microscopy) for Order or Family level identification. We deemed 19 cases inconclusive either because the DNA barcode recovered from the sample did not meet our 98% match criteria when compared to the Barcode of Life Database (BoLD) or because the DNA barcode matched to a set of ≥2 closely related species with overlapping barcodes, preventing complete species identification. Age of the sample (≤6 months) did not affect DNA viability, but initial condition of the sample and the collection method was critical to DNA identification success. The DNA barcoding approach has great potential in aiding in identification of birds (and wildlife) for airfield management practices, particularly in regions of the world that lack the vast research collections and individual expertise for morphologic identifications. (JOURNAL OF WILDLIFE MANAGEMENT 72(5):1231–1236; 2008)

DOI: 10.2193/2007-272

KEY WORDS barcodes, BoLD, birdstrikes, CO1, DNA, identification.

Outline

DNA Barcoding 101

- **Rationale**
- **What is DNA barcoding ?**
- **History of DNA barcoding**
- **How does it work ?**
- **Why do we care ?**
- **What are the applications of DNA barcoding?**

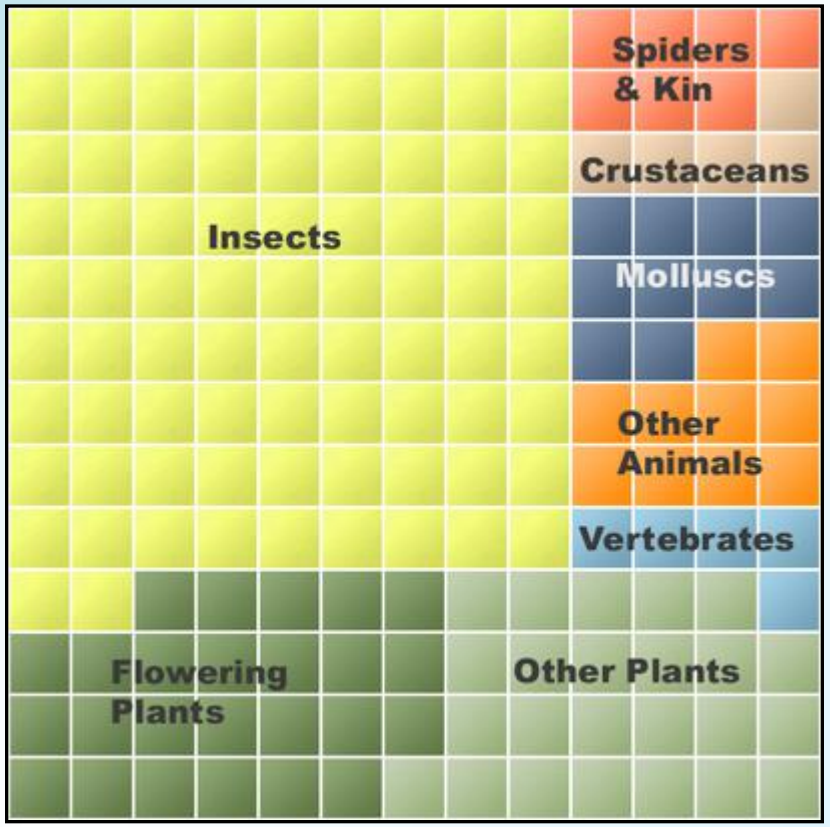
Canadian Center for DNA Barcoding - building reference library

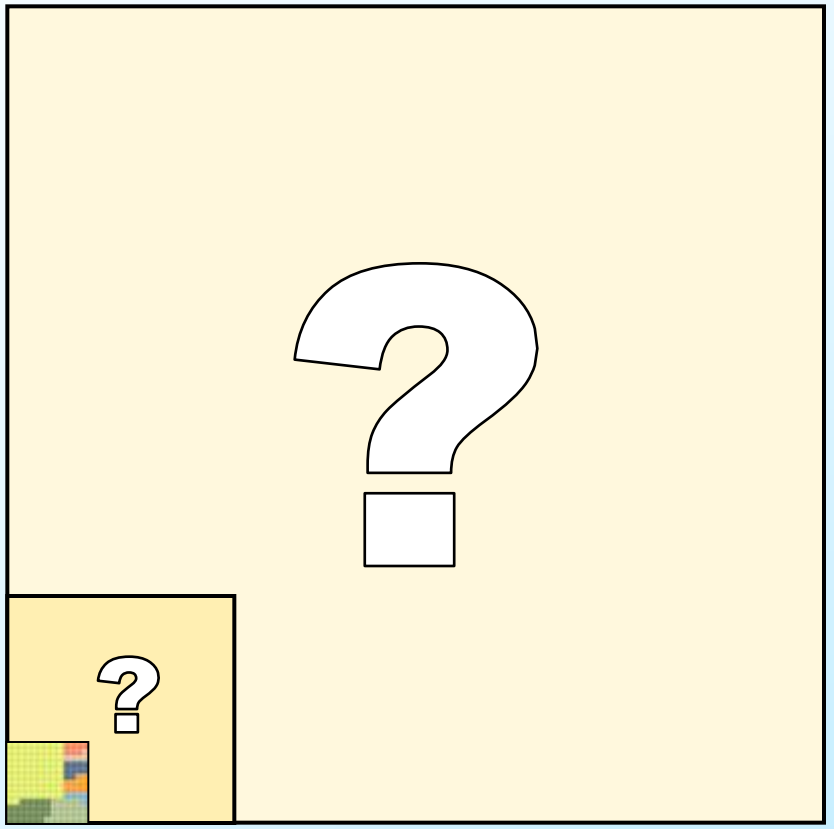
- **DNA Barcode Factory**
- **International Barcode of Life**
- **New frontiers**

Birdstrike species identification through DNA Barcoding

- **Challenges and examples**
- **10 years and counting**

DNA Barcoding 101. How many species on Earth.





Known biodiversity:
1.7 million species of plants and animals

Estimated biodiversity:
10 million to 100 million species

Our Planet:
Home to 10-100 million species

Our Minds:
Able to recognize and recall perhaps 1,000 species


What Species?

What Species?

Imagine...

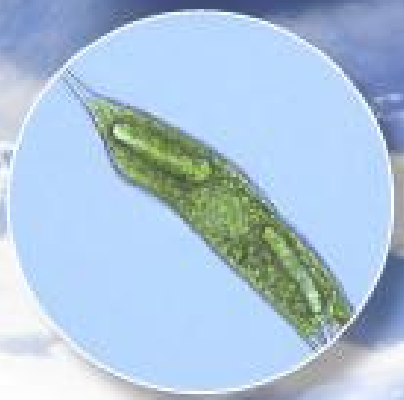
a world in which you can know the name of

any animal

any plant

any fungus

any organism



Matter, Cosmos, Biodiversity

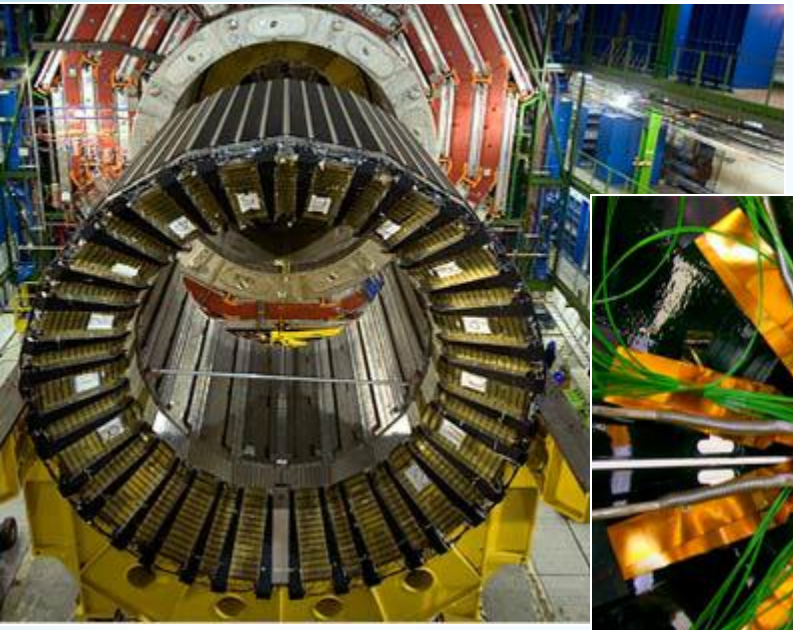
Michael Faraday

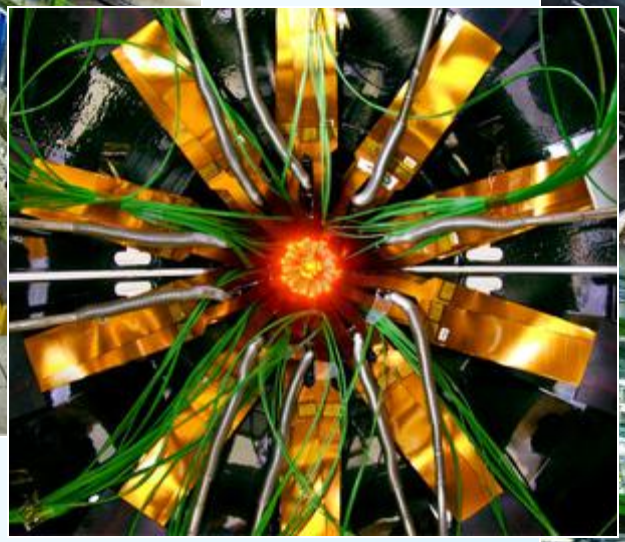
John Herschel

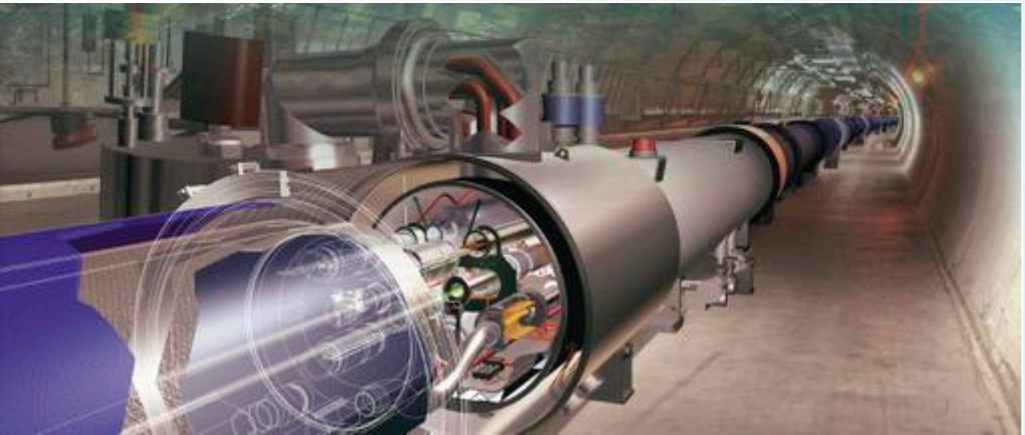
Charles Darwin

Census of Matter 1850

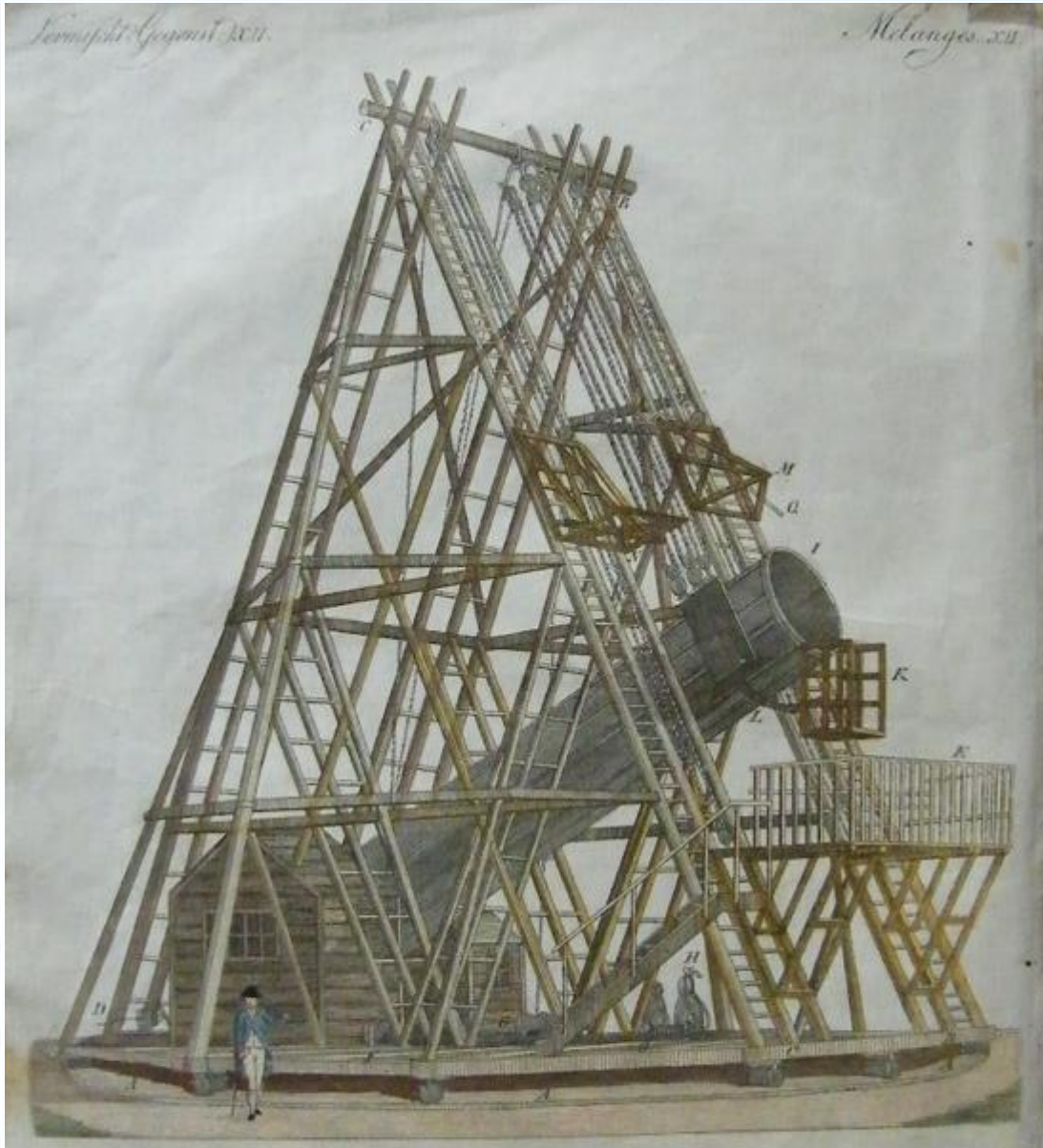
Census of Matter 2010



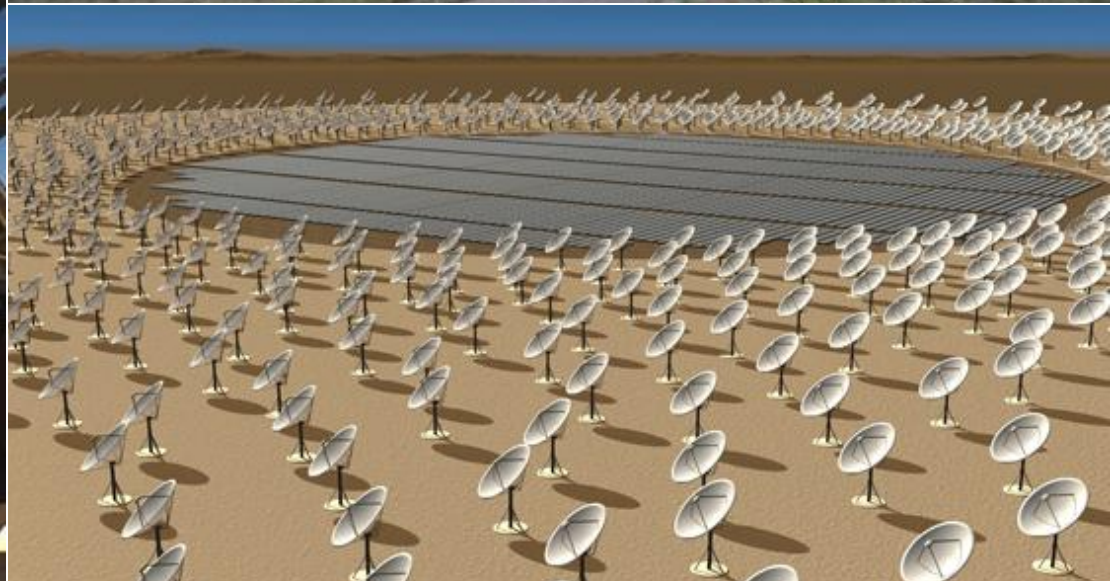




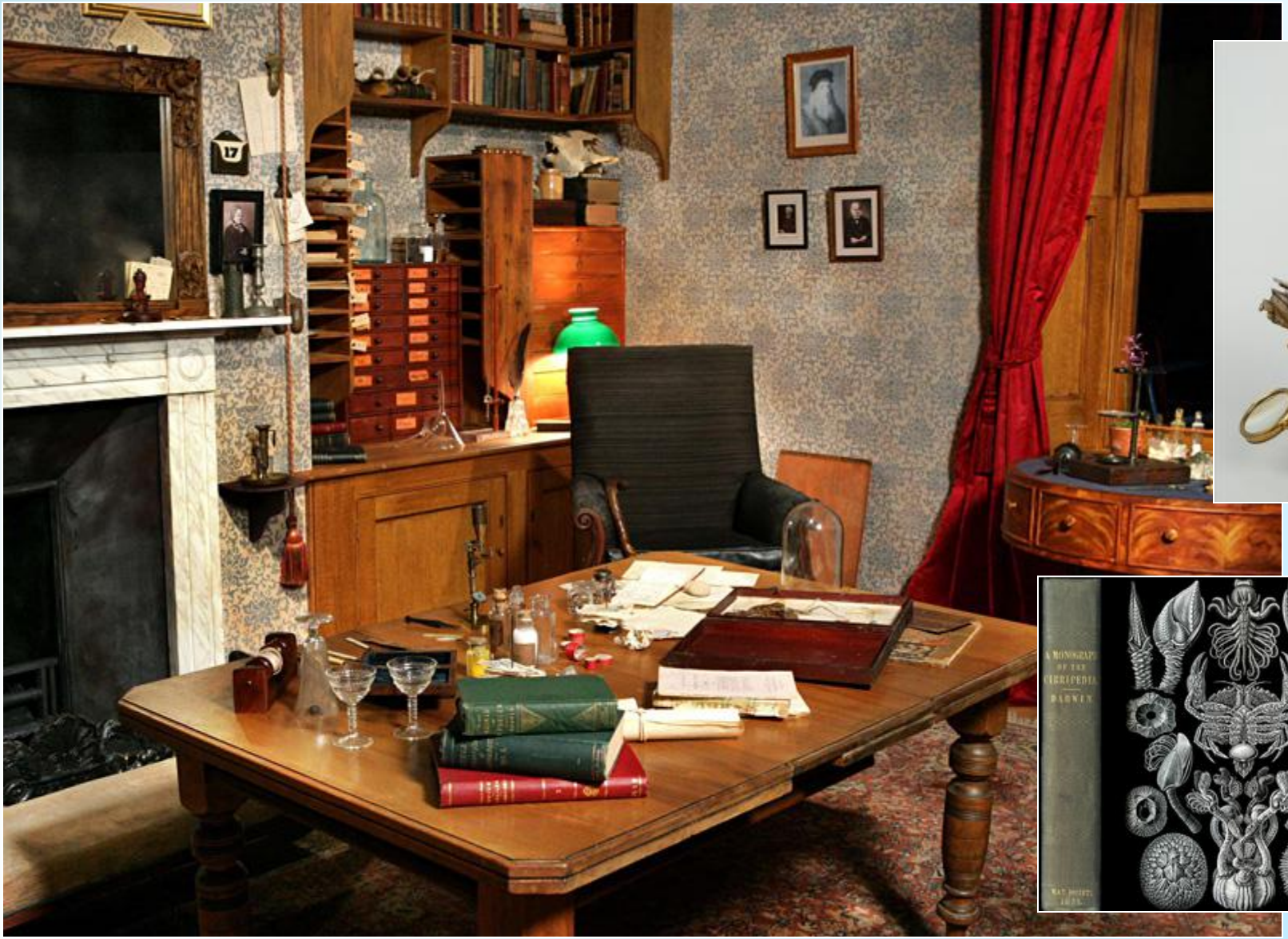
Census of the Cosmos 1850

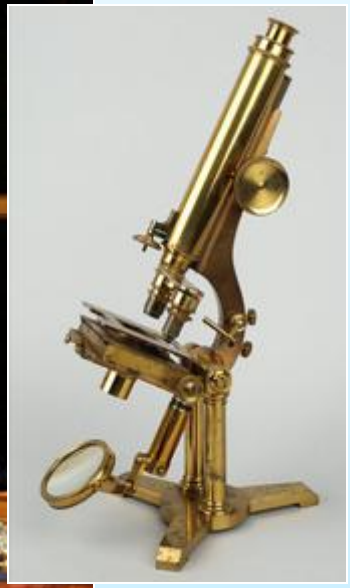


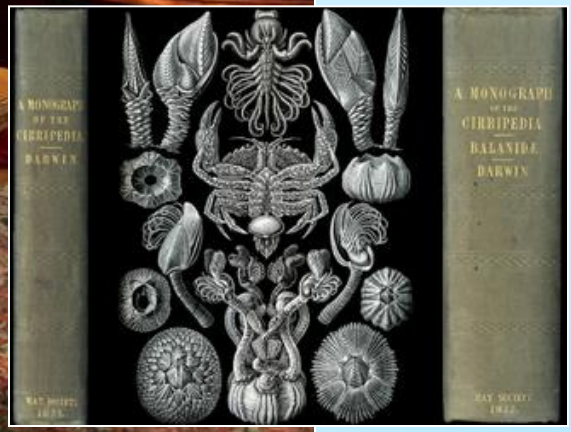
Census of the Cosmos 2010



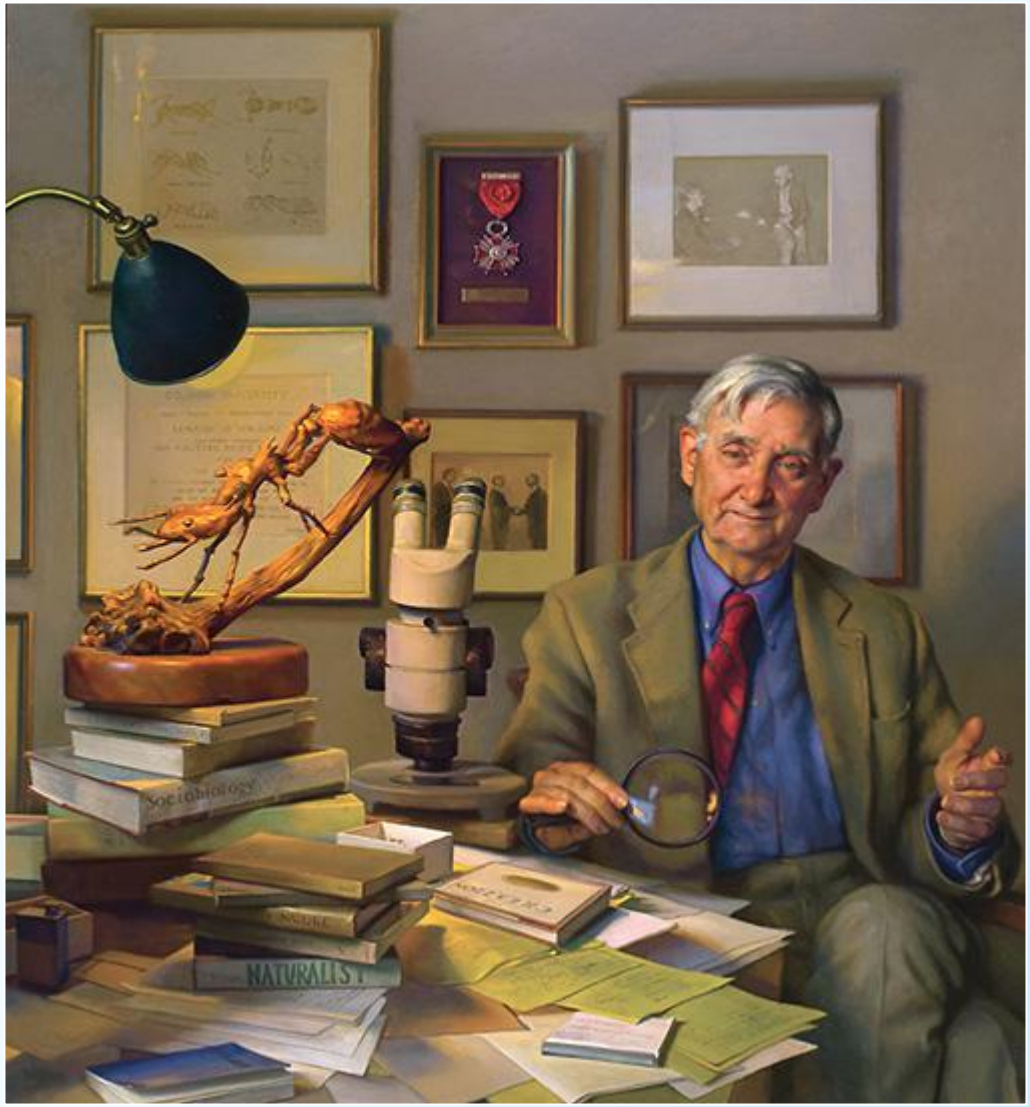
Census of Biodiversity 1840



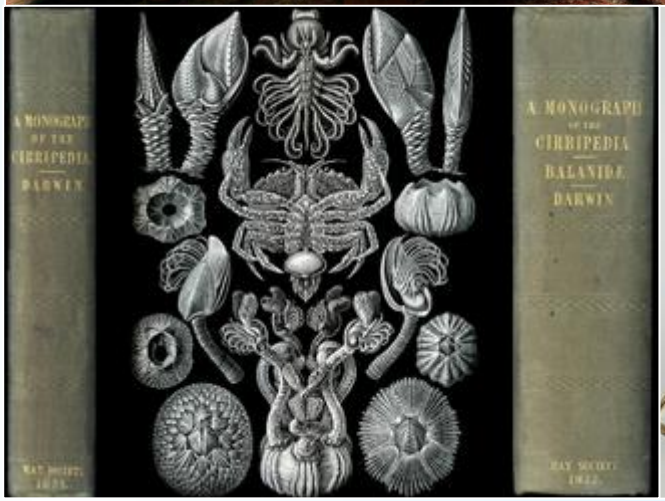


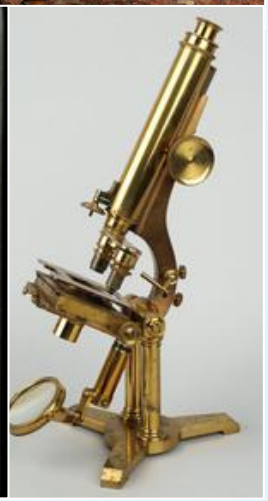


Census of Biodiversity 2003

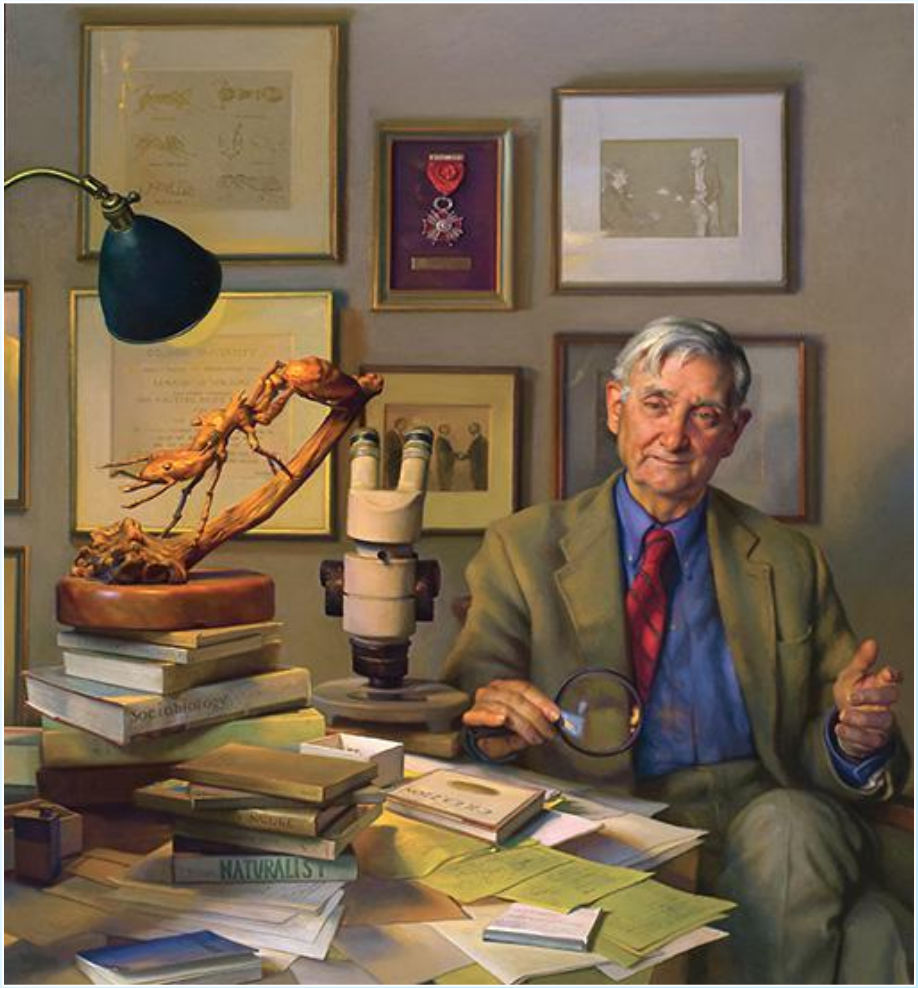


Census of Biodiversity 1840





Census of Biodiversity 2003



1.7M species described in 250 years =

Many centuries to complete catalogue of life

Building A DNA-Based ID System For All Eukaryotes

DNA-based Identification System

Universal Product Code

- Ten unique states
- Twelve distinct positions

DNA Barcode

- Four unique states
- Over 600 positions

It happened once, it will happen again!

DNA fingerprinting revolutionized criminal forensics, DNA testing

Sperling FA, Anderson GS, Hickey DA. (1994)

A DNA-based approach to the identification of insect species used for postmortem interval estimation.

J Forensic Sci. 39: 418-427.

DNA fragments were amplified using the polymerase chain reaction (PCR), followed by direct DNA sequencing of the amplification products. Based on these abundant DNA sequence differences, we can unambiguously identify the immature larval stages of these insects.

These DNA sequence differences were also used to predict species-specific, diagnostic restriction sites in the amplified DNA, and these predictions were verified by digestion with nine restriction enzymes.

DNA Barcoding is only a decade old

THE ROYAL SOCIETY FirstCite® e-publishing

Received 29 July 2002
Accepted 30 September 2002
Published online

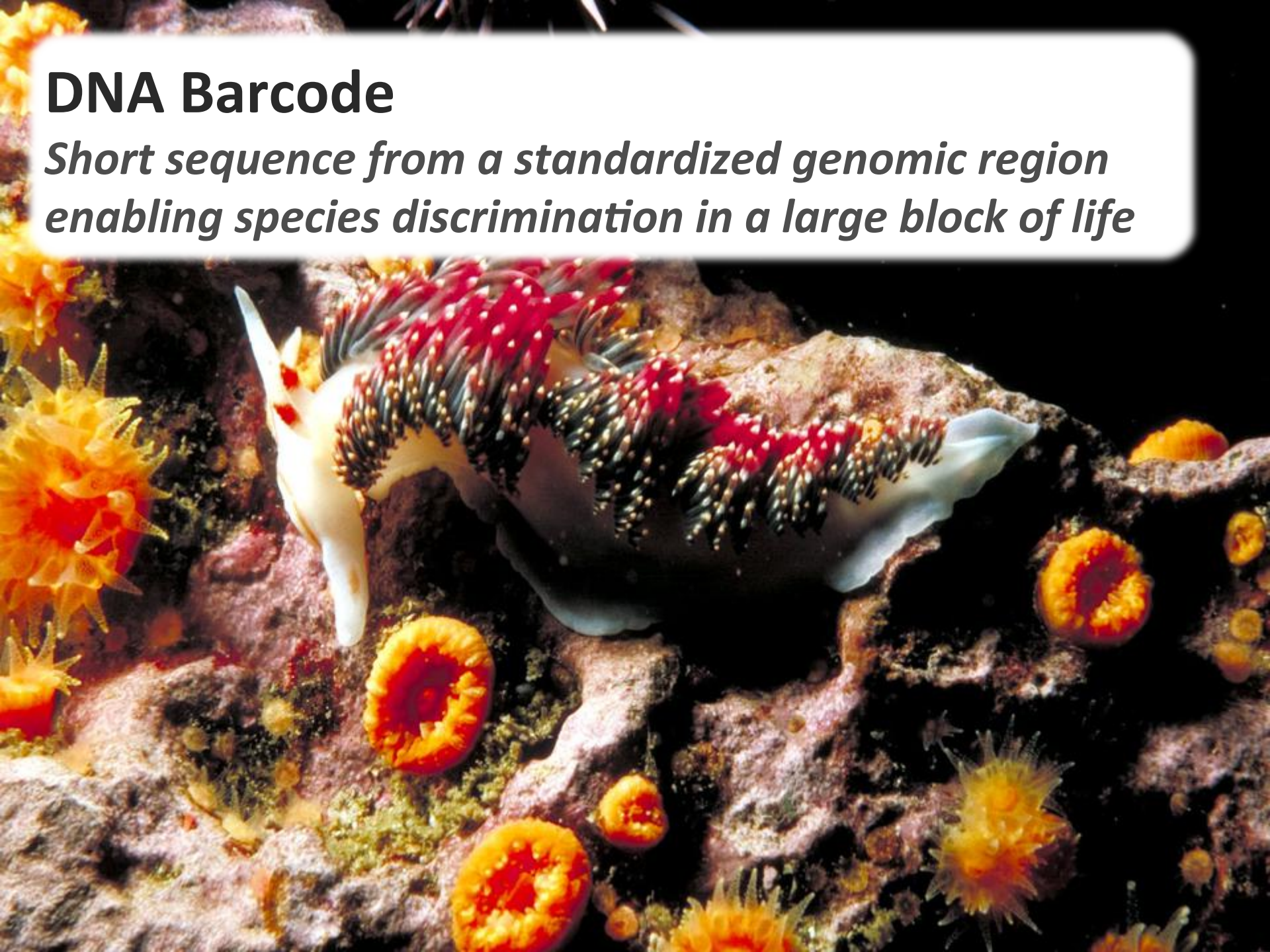
Biological identifications through DNA barcodes

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

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

Although much biological research depends upon species diagnoses, taxonomic expertise is collapsing. We are convinced that the sole prospect for a sustainable identification capability lies in the construction of systems that employ DNA sequences as taxon 'barcodes'. We establish that the mitochondrial gene cytochrome *c* oxidase I (COI) can serve as the core of a global bioidentification system for animals. First, we demonstrate that COI profiles, derived from the low-density sampling of higher taxonomic categories, ordinarily assign newly analysed taxa to the appropriate phylum or order. Second, we demonstrate that species-level assignments can be obtained by creating comprehensive COI profiles. A model COI profile, based upon the analysis of a single individual from each of 200 closely allied species of lepidopterans, was 100% successful in correctly identifying subsequent specimens. When fully developed, a COI identification system will provide a reliable, cost-effective and accessible solution to the current problem of species identification. Its assembly will also generate important new insights into the diversification of life and the rules of molecular evolution.

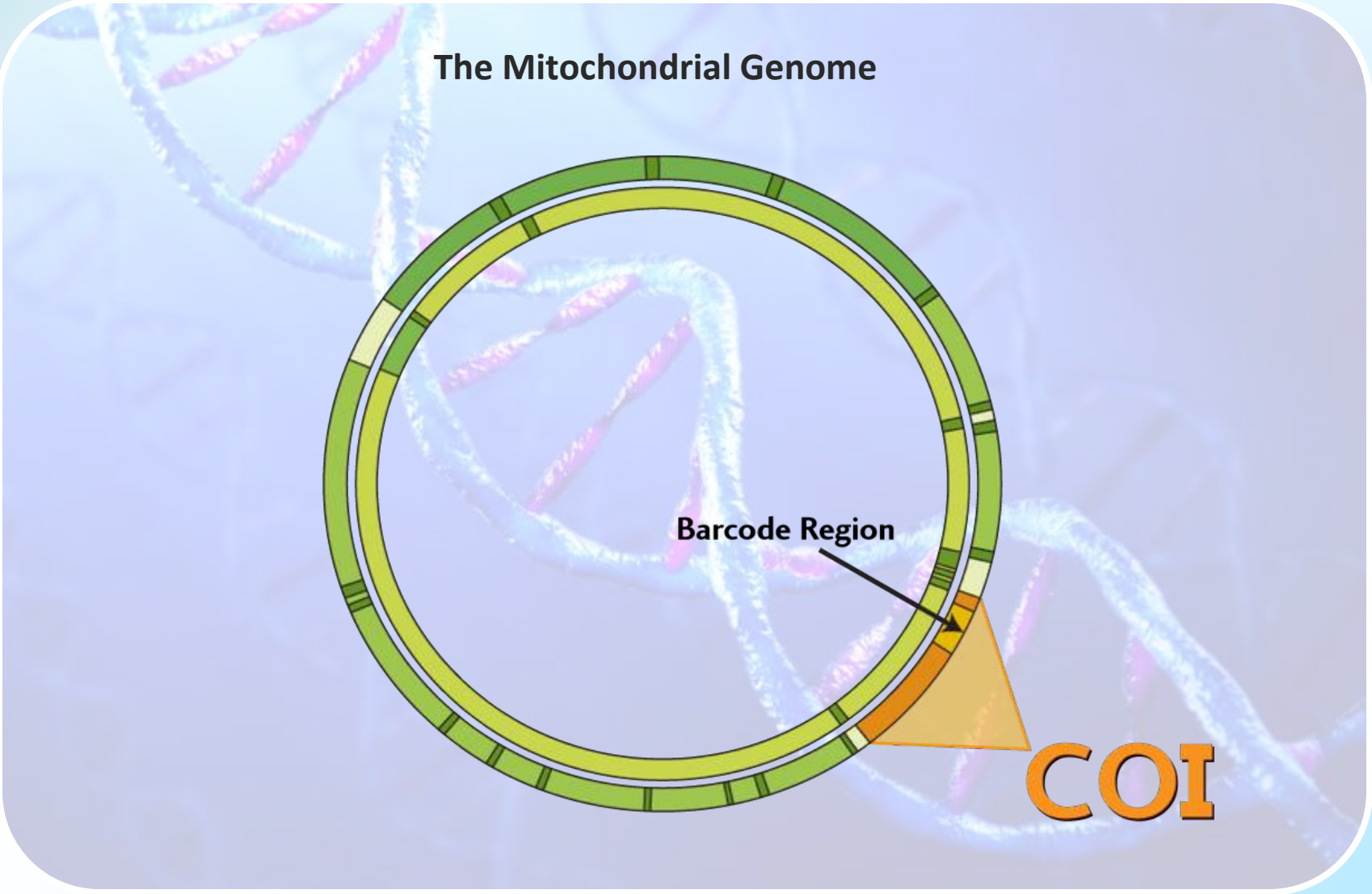
Keywords: molecular taxonomy; mitochondrial DNA; animals; insects; sequence diversity; evolution



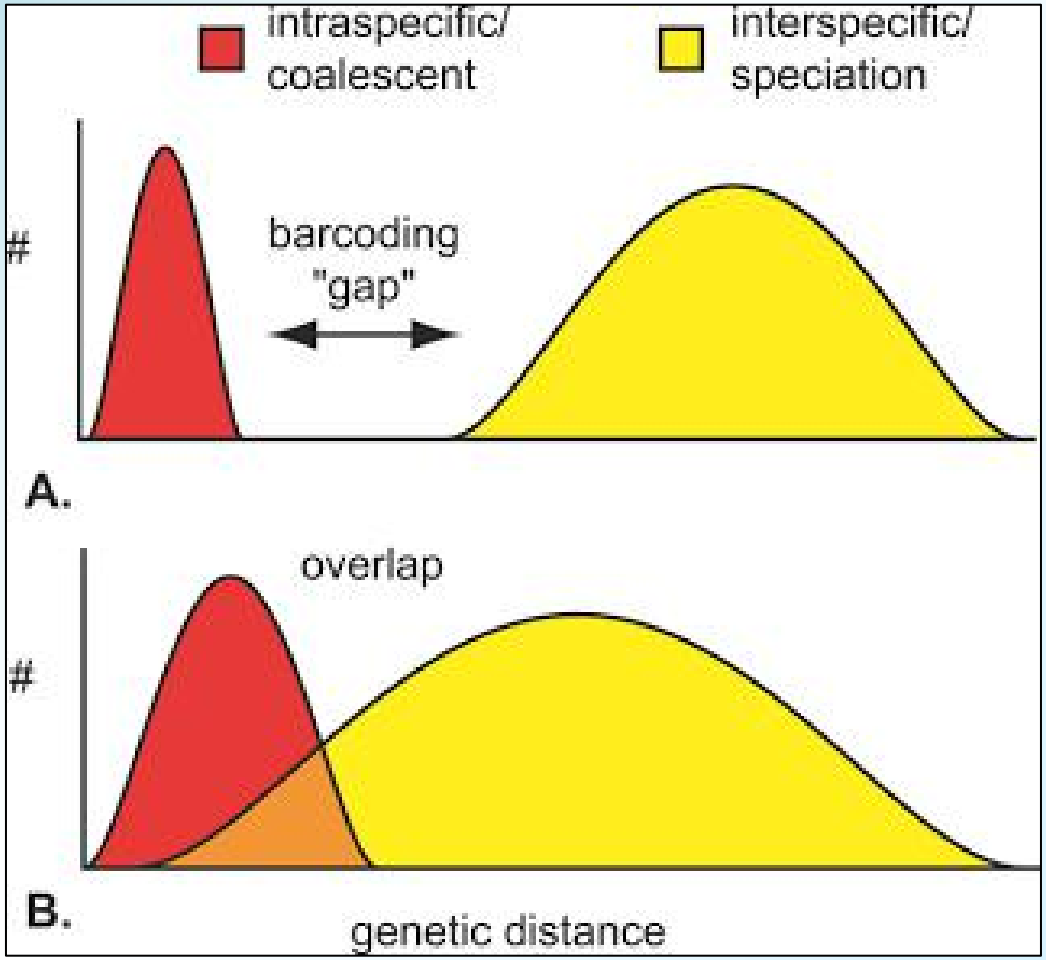
DNA Barcode

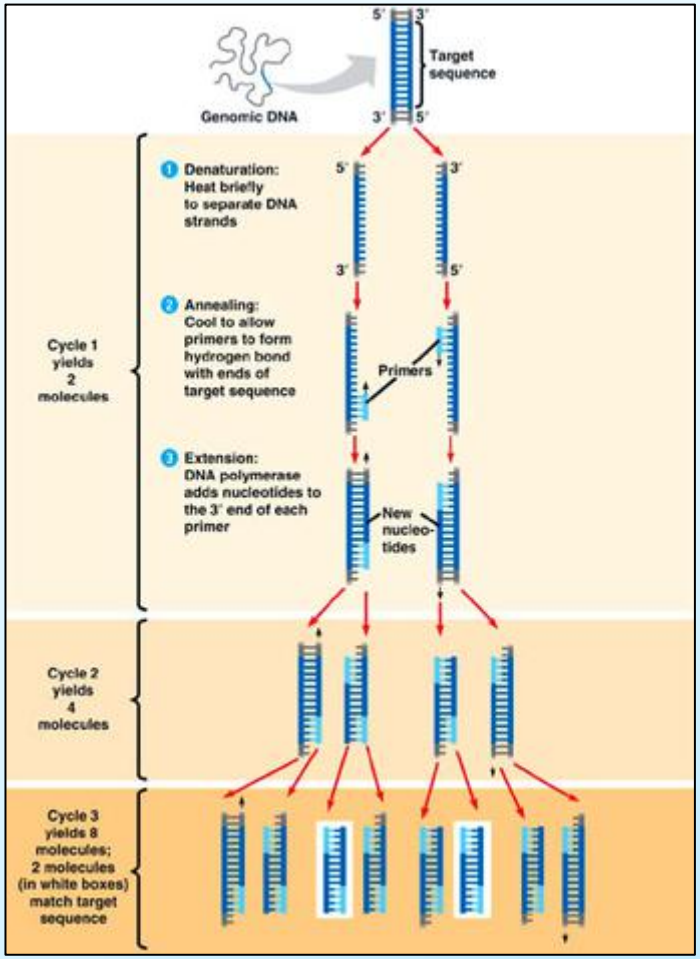
Short sequence from a standardized genomic region enabling species discrimination in a large block of life

Barcode Region for Animals and some Protists



DNA Barcoding 101: Why does it work?





Simple yet powerful premise:

genetic variation among species exceeds variation within species

Primer specificity and targeted gene provide flexibility for diverse applications

The most amazing DNA barcoding facility yet…

Things could not get any better? They did!


November 2010


February 2011


March 2011

Centre for Biodiversity Genomics (CBG), a $14M 40,000 sq. ft. expansion to the research infrastructure


April 2012

This changes everything. Again.

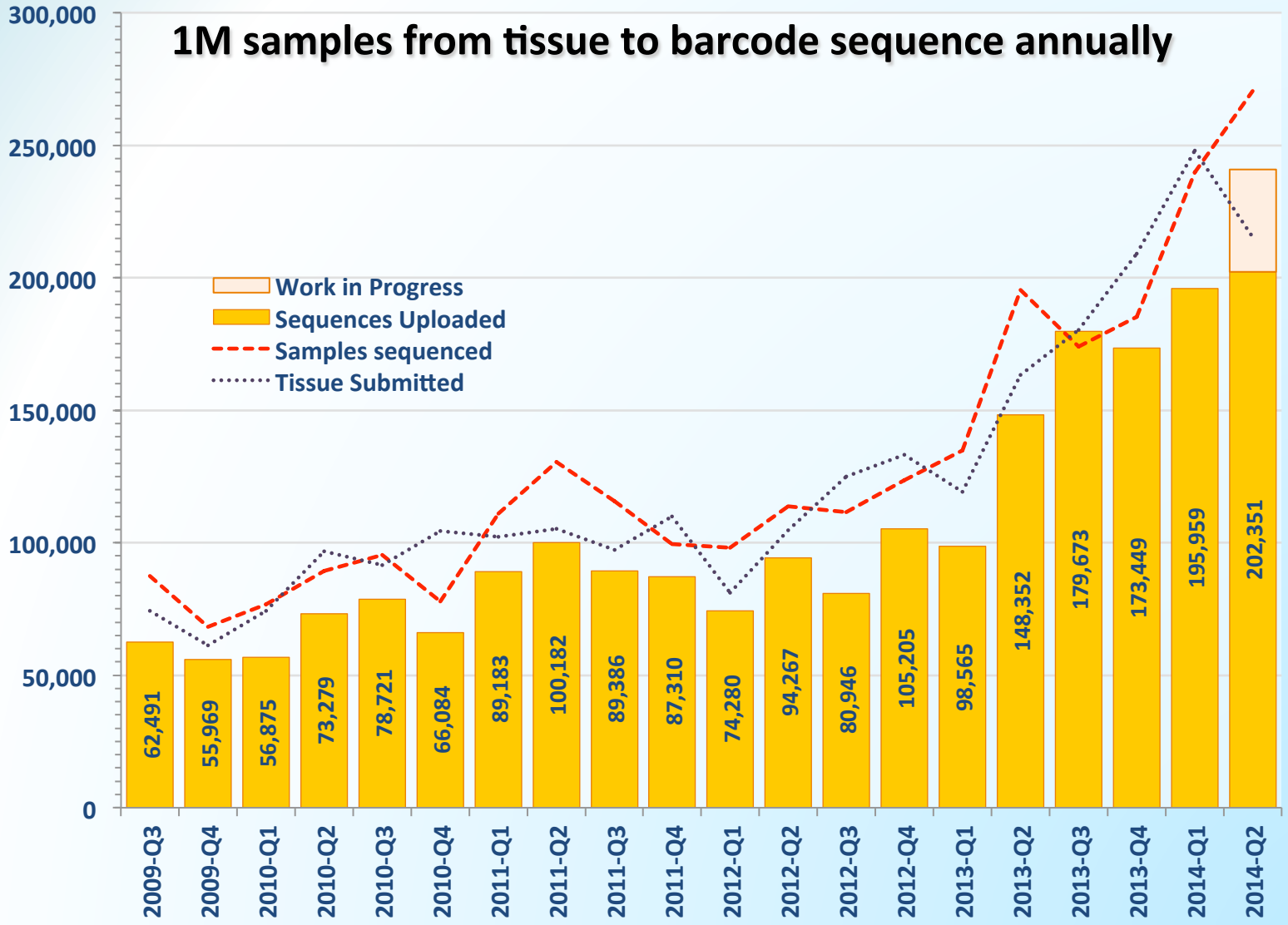
- Five Biomek Liquid Handling systems
  - Two NX
  - Two FX robots with single Cytomat hotels
  - One FX robot with two 96 channel heads, two cytomat hotels and integrated robotic centrifuge bring total extraction capacity up to 700K samples per year
- Five 3730XL DNA sequencers (up to 2M sequencing reads per year)
- Twenty eight 96-well and eight 384-well Eppendorf thermocyclers (5,700 samples can be PCRed concurrently)
- Two 96-well real-time PCR machines
- Fourteen -80oC freezers storing over 1M DNA extract.
- Ion Torrent PGM next generation sequencer

- Over 2.5M barcodes generated by the CCDB (over 70% of the total barcodes gathered globally)
- Surpassed 800K finished barcodes per year

DNA Barcoding facility in Guelph

CCDB
Canadian Centre for DNA Barcoding

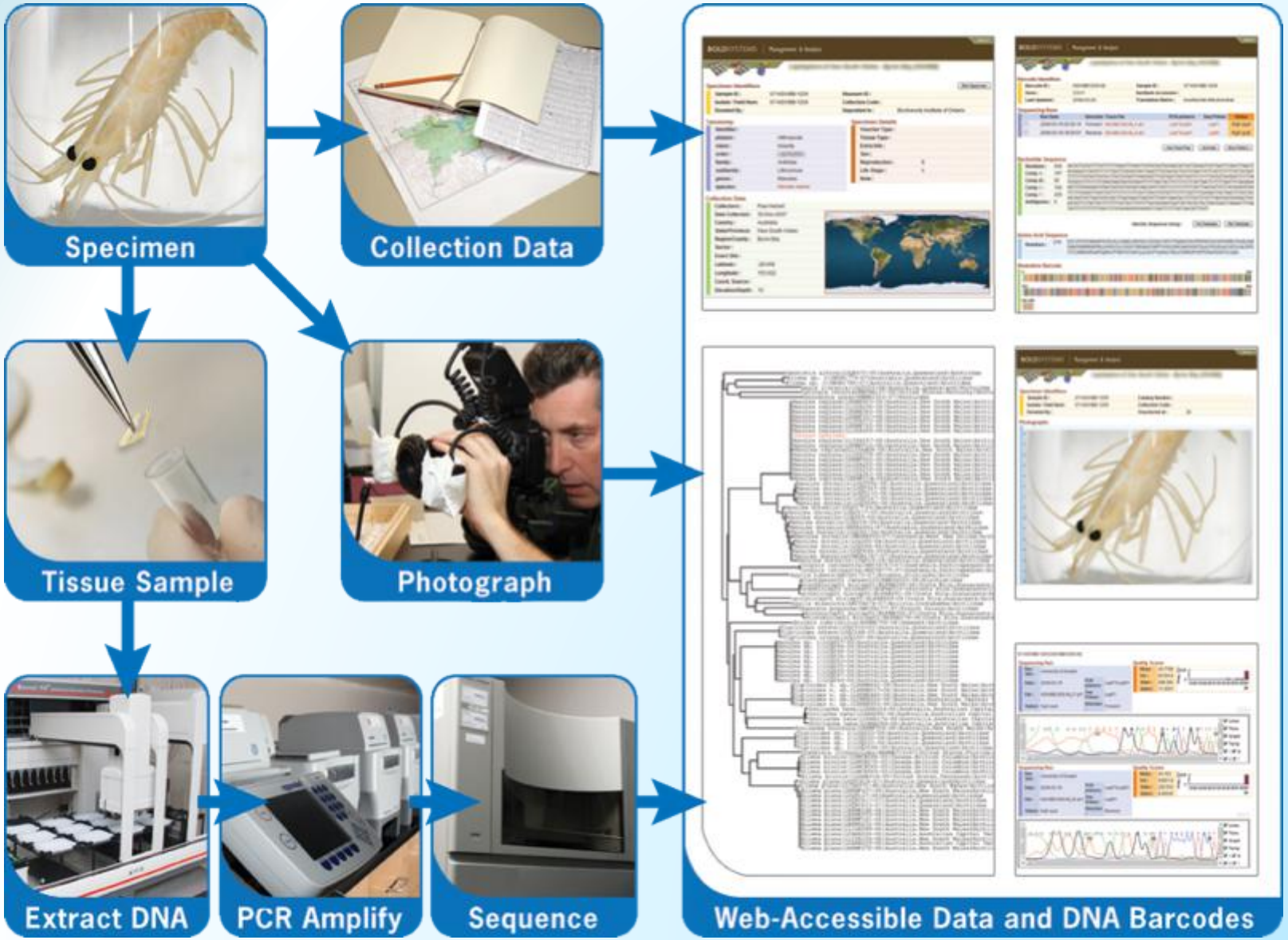
Analytical throughput at the CCDB



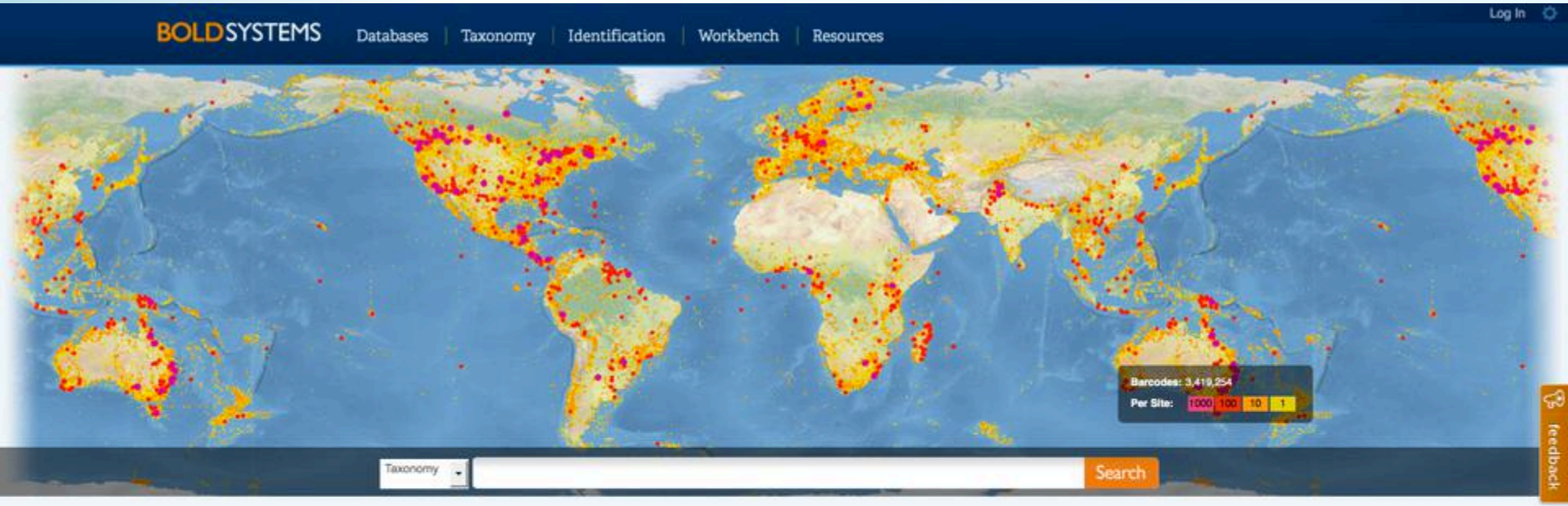
CCDB Today

- Current projects supported by the CCDB
- Analytical capacity and sample supply Chain
- Research infrastructure upgrades
- Traceability (user alerts)
- Error proofing
- MTP-384 and new workflows
- Reducing laboratory waste
- Sequence autoediting

Obtaining A Barcode Record



Partnership with BOLD, online library for DNA barcodes



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 Numbers (BINs), sequence clusters that closely approximate species.

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 environment that securely supports the assembly and validation of DNA barcodes and ancillary sequences.

DNA Barcode Library

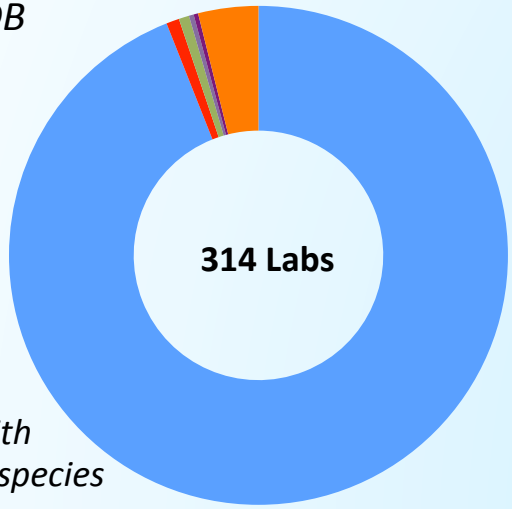
BOLD

Records: **3.4M** barcodes (+0.2M),
4.5M specimens (+0.3M)
3.1M specimens public (+XM)

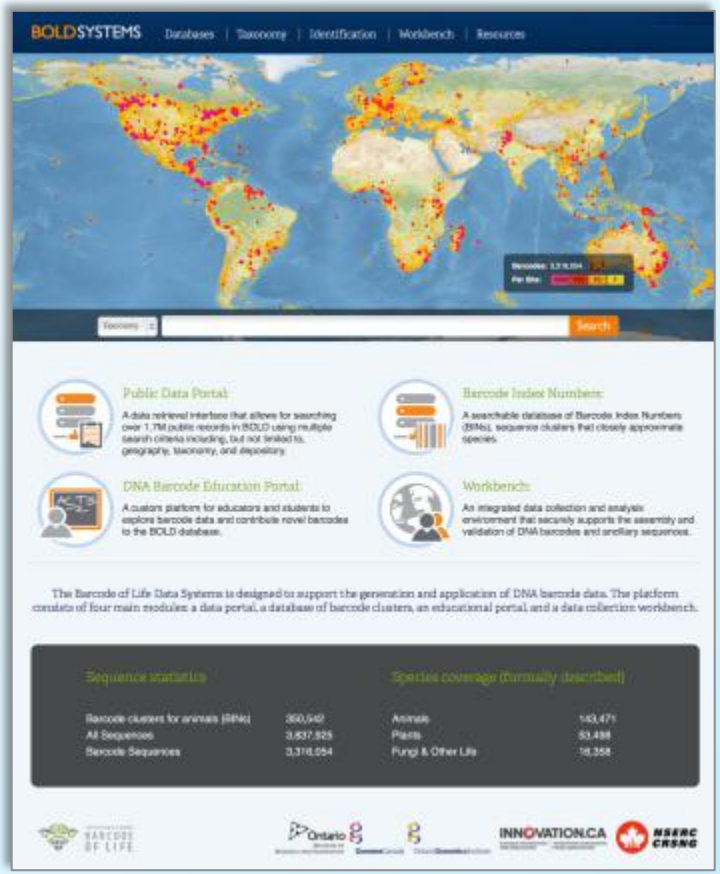
Species: **350K** animals (+24K),
54K plants (+2K),
16K fungi (+3K)

Citations: **1051 BOLD** citations (Google Scholar)

Sequencing: ***73%*** *CCDB*



GBOL Labs activating with
13K Sequences from 3K species



Imagine...

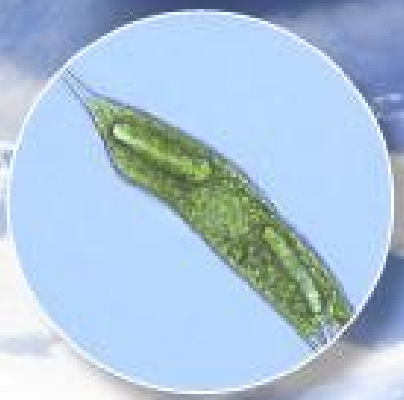
a world in which you can know the name of

any animal

any plant

any fungus

any organism



Imagine...

international BARCODE OF LIFE

Making Every Species Count

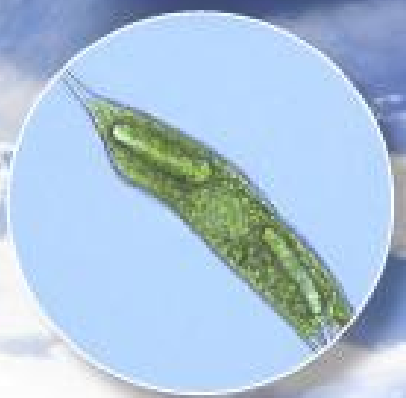
a world in which you can know the name of

any animal

any plant

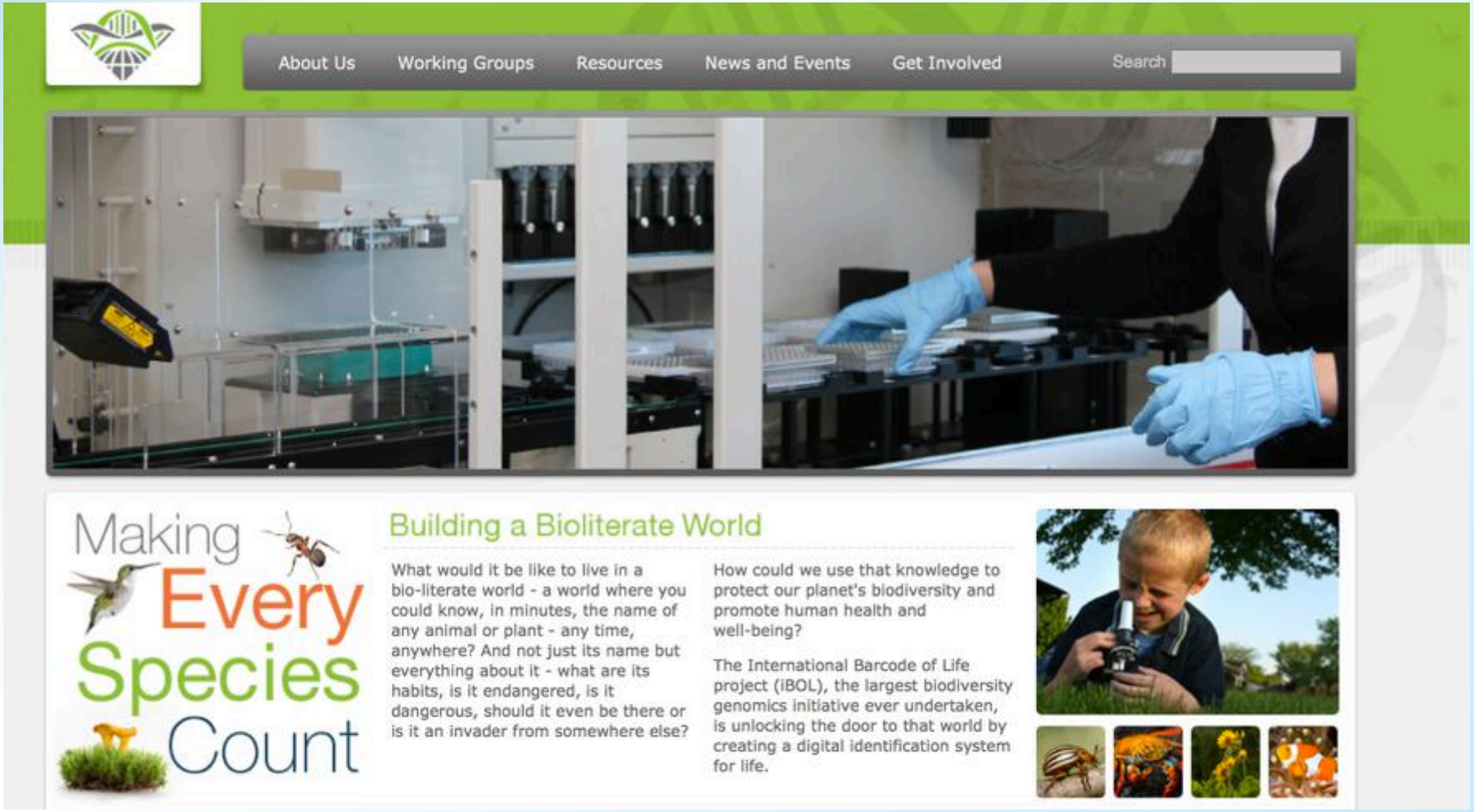
any fungus

any organism



This is the world that iBOL and CCDB are helping to build

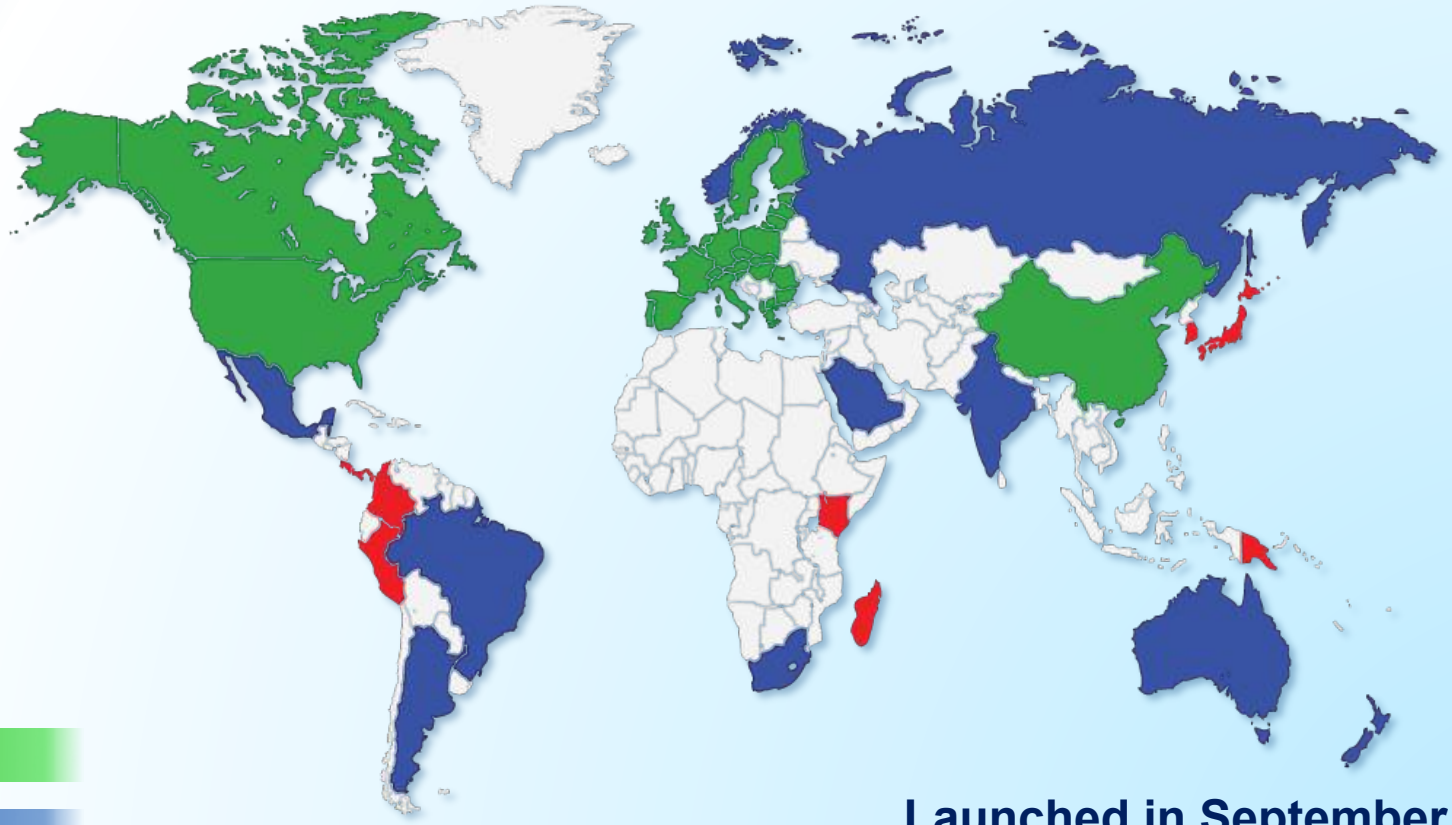
International Barcode of Life, ibol.org



International Barcode of Life

international BARCODE OF LIFE

5 Years
5M Specimens
500K Species



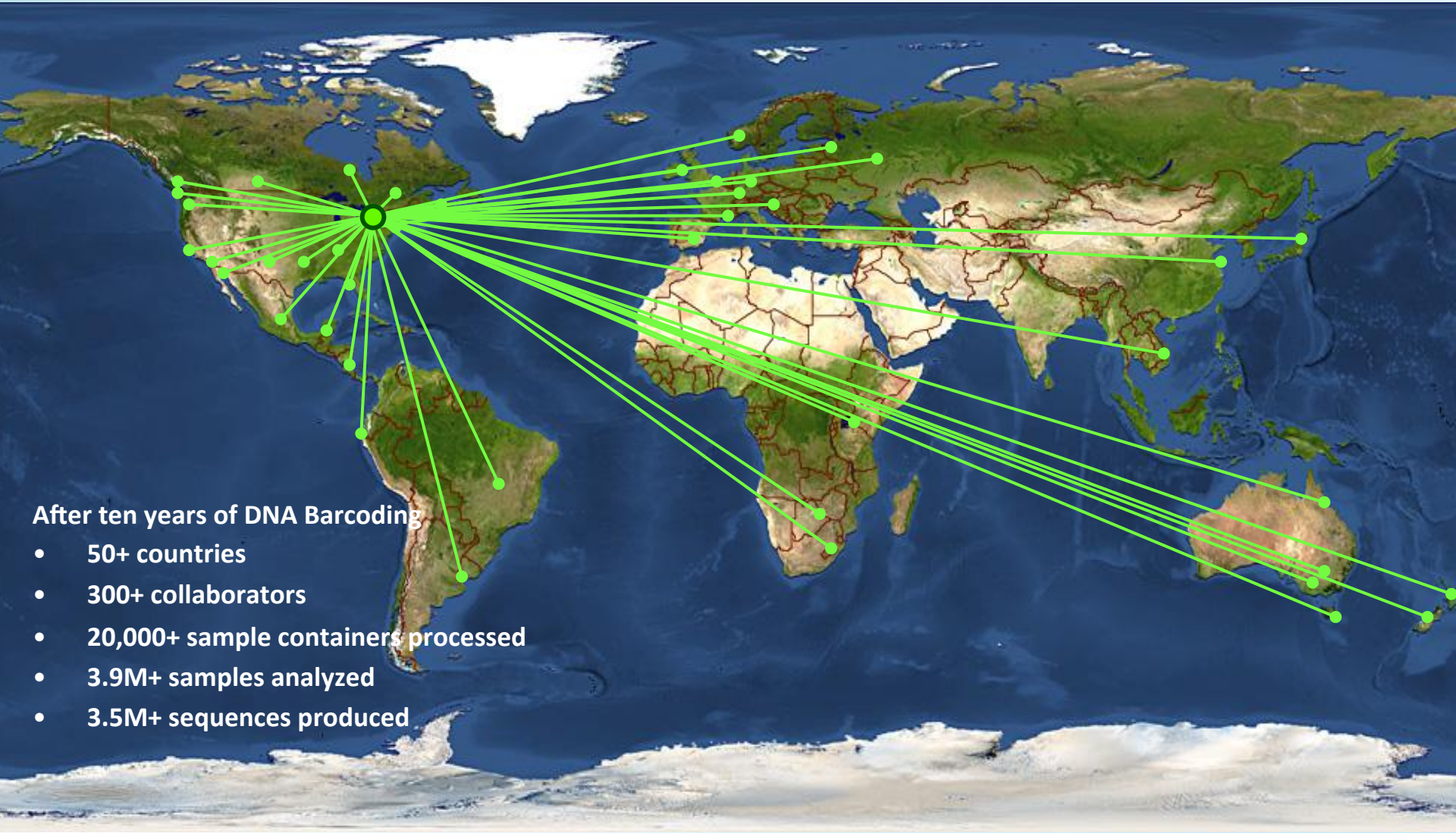
Central Node

Regional Node

National Node

Launched in September 2009
Nearly 4M specimens acquired and
500K Species

CCDB – main analytical hub for iBOL



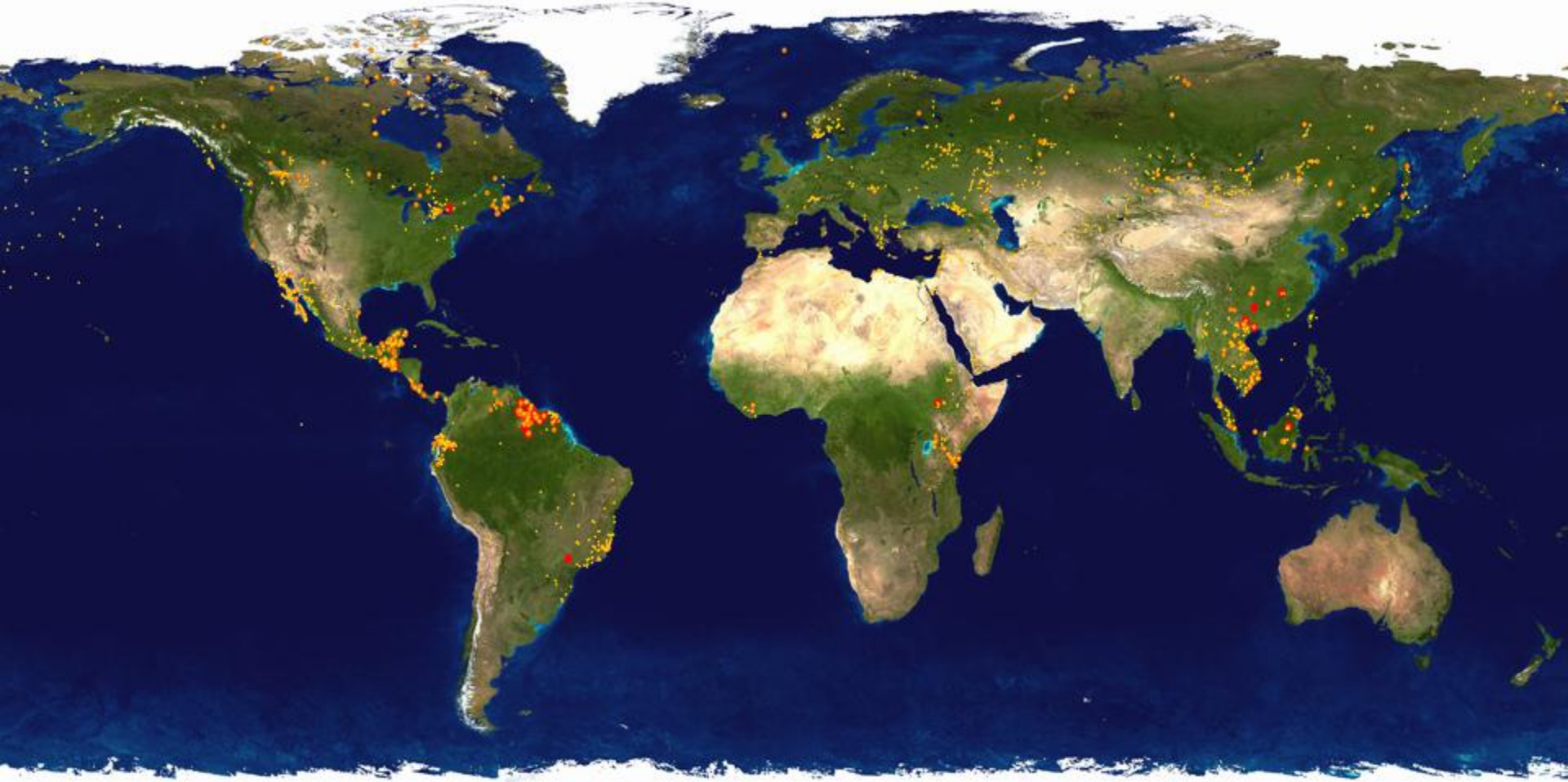
BOLD, online library for DNA barcodes

640 CPU Cores, 200 Terabytes of Storage

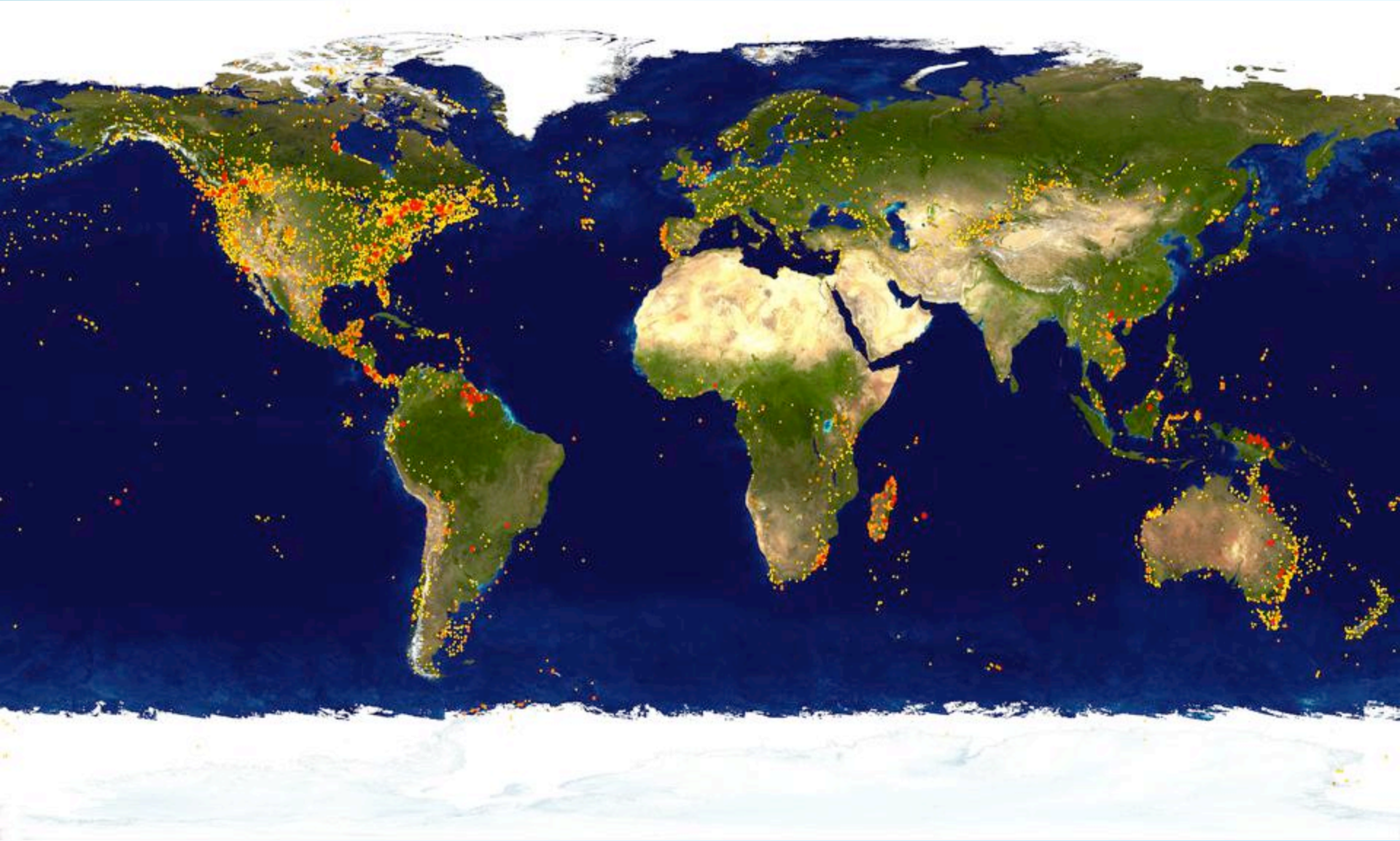
1 Architect, 6 Programmers, 2 Project Managers, 2 Bioinformaticians

7 Data Managers

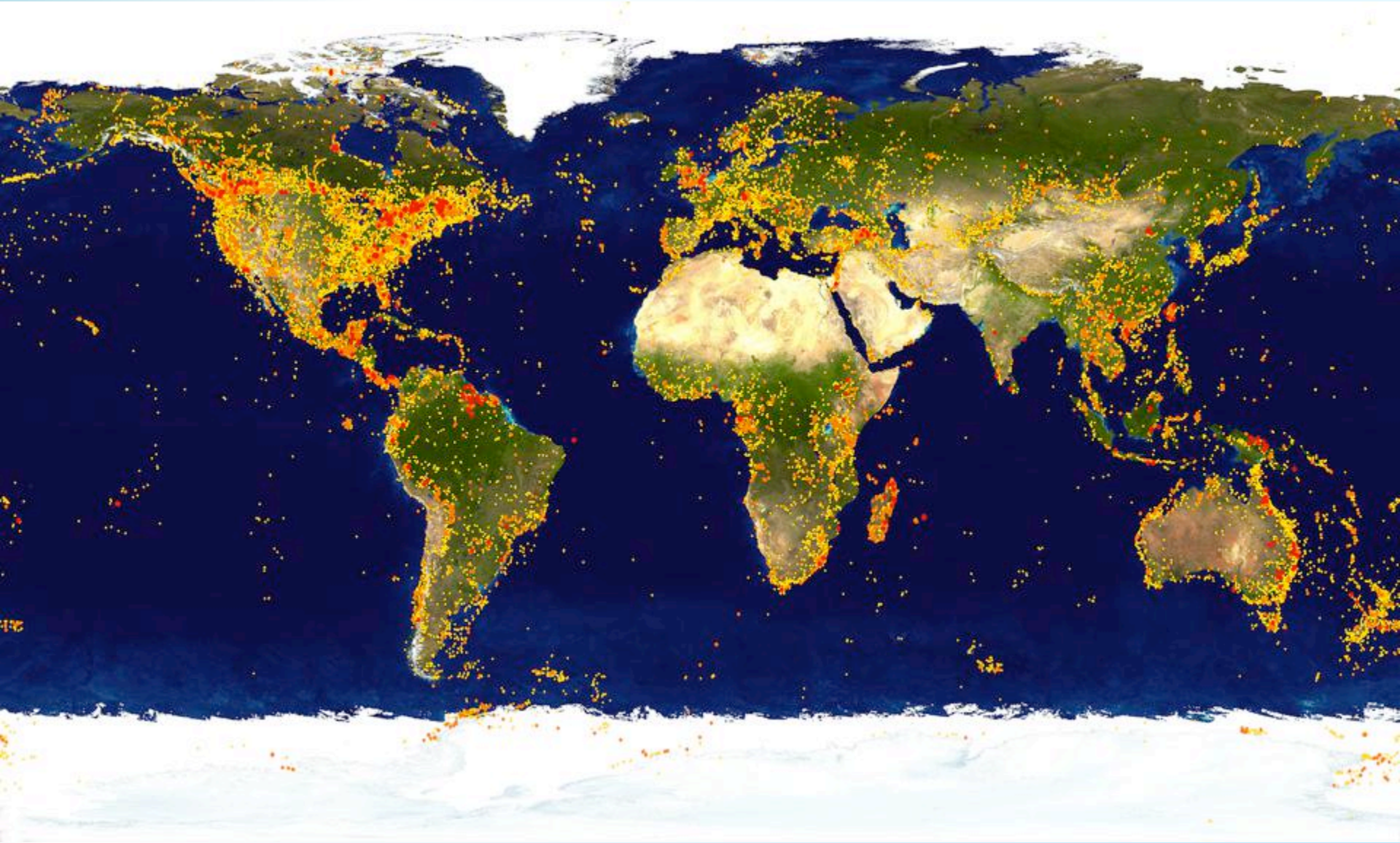
Barcode Data on BOLD - 2007



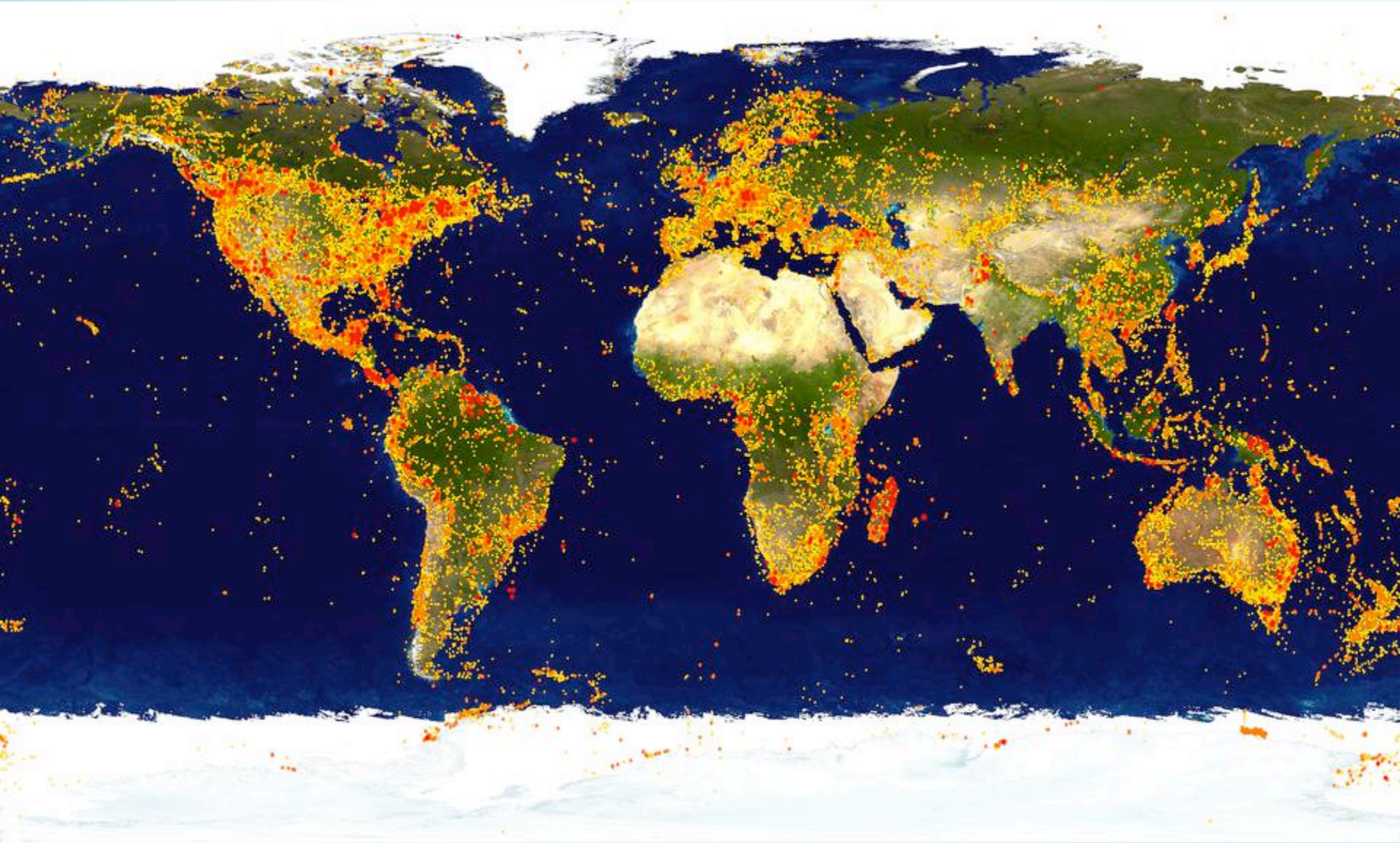
Barcode Data on BOLD - 2009



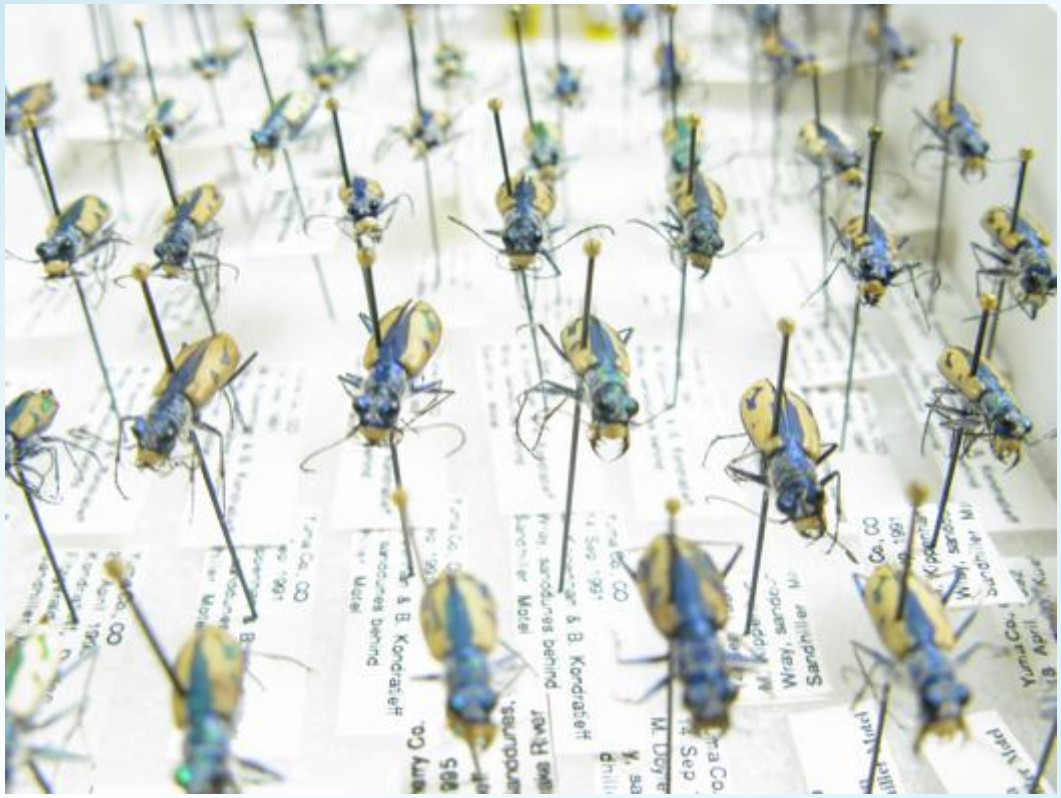
Barcode Data on BOLD - 2012



Barcode Data on BOLD - 2014



DNA Barcoding and Specimen Vouchers

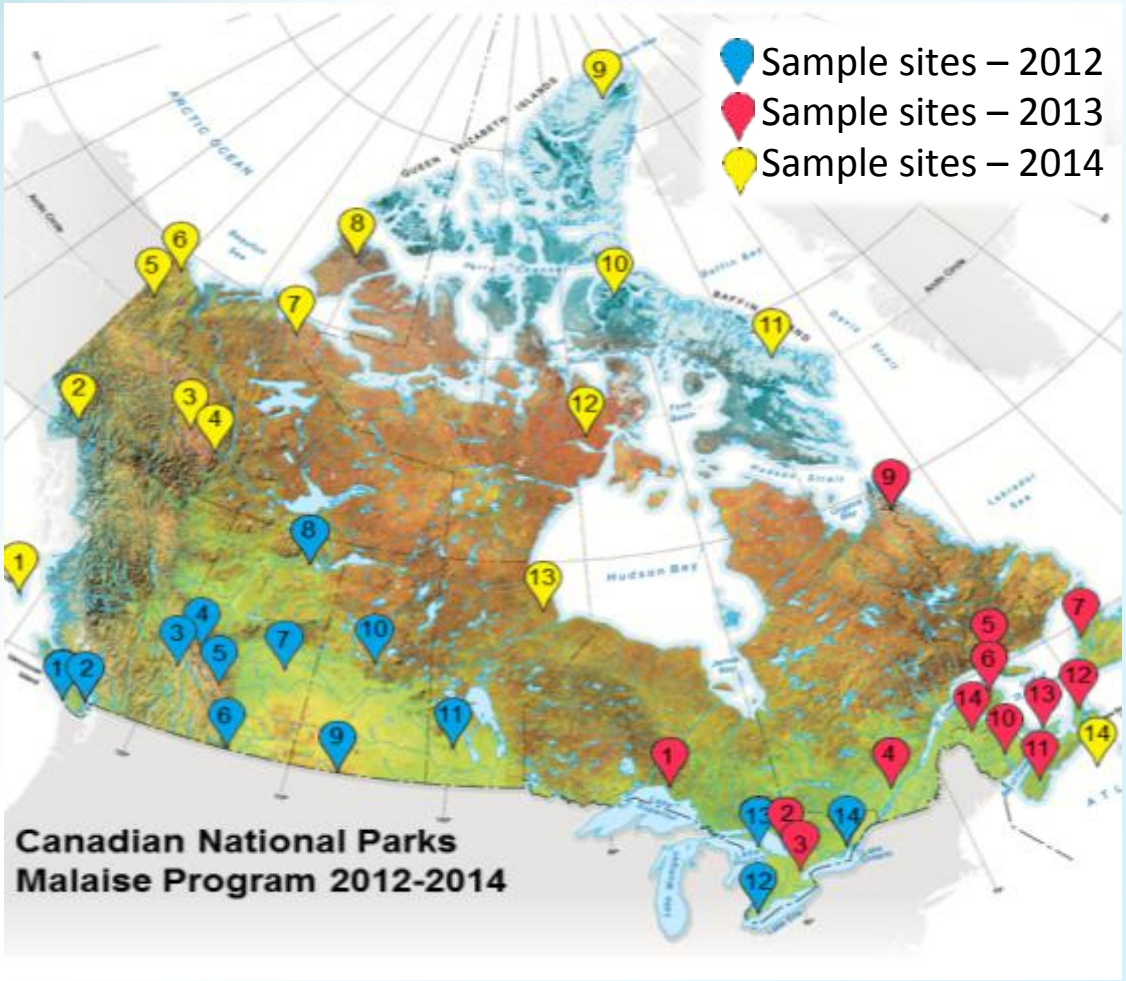


Museum sampling

Permanent connection with the vouchers

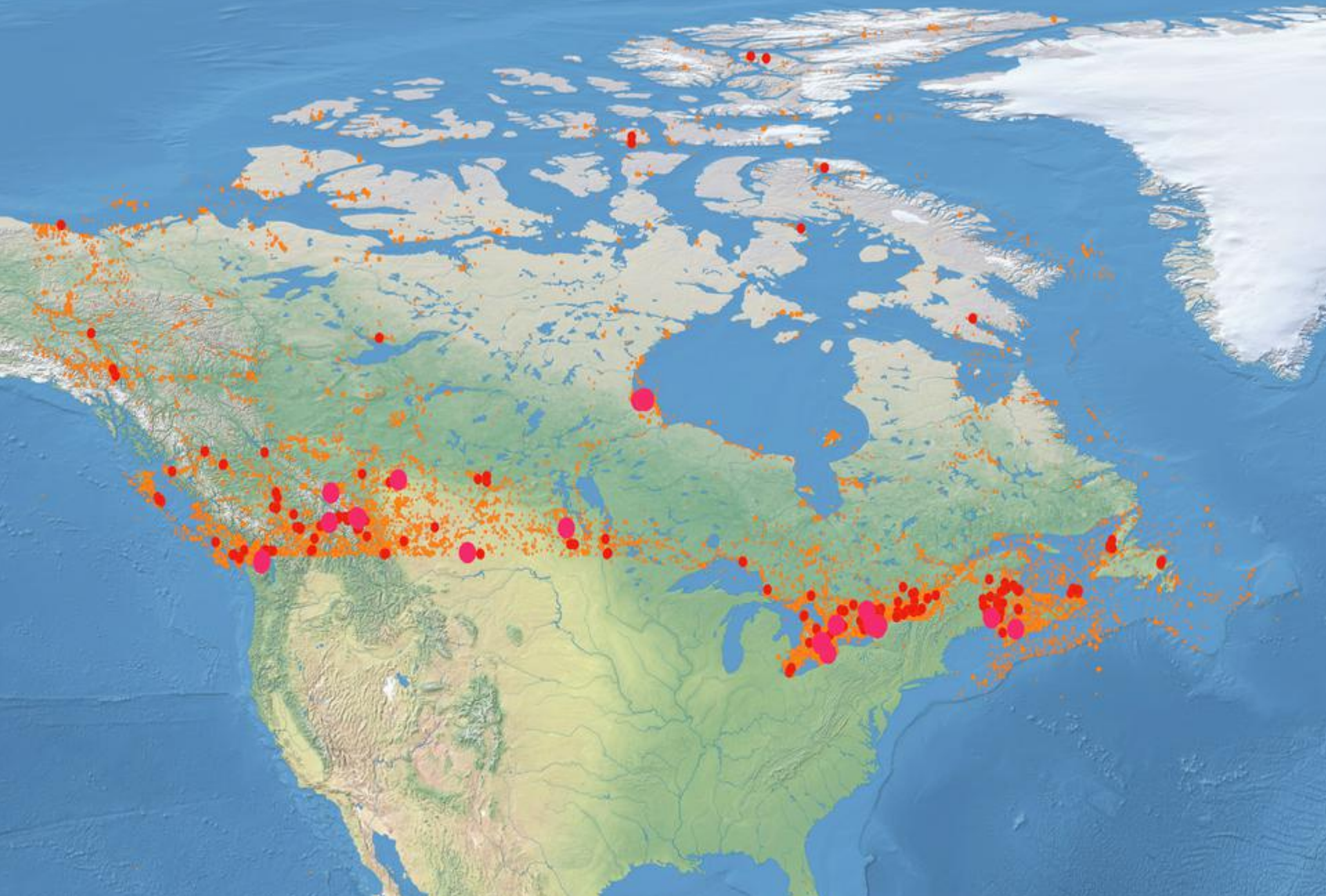
Specimen Collection – Mobile Lab

Specimen Collection – Canadian National Parks



- 25 national parks were sampled in 2014 including 14 remote parks
- All 45 national parks have now been sampled
- 99% of 2013 specimens have been analyzed
- Data now available for 600K specimens from 32 national parks

Barcode Closure for Canada



Towards a Complete Barcode Library for Canadian Life

Vascular Plants: 6K

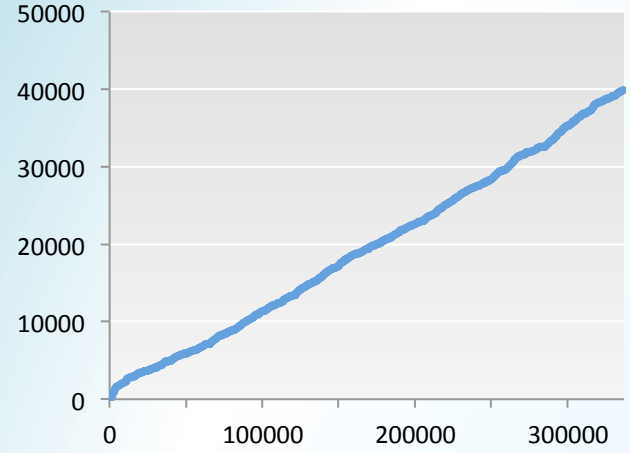
Invertebrates: 70K

Other Plants: 3K

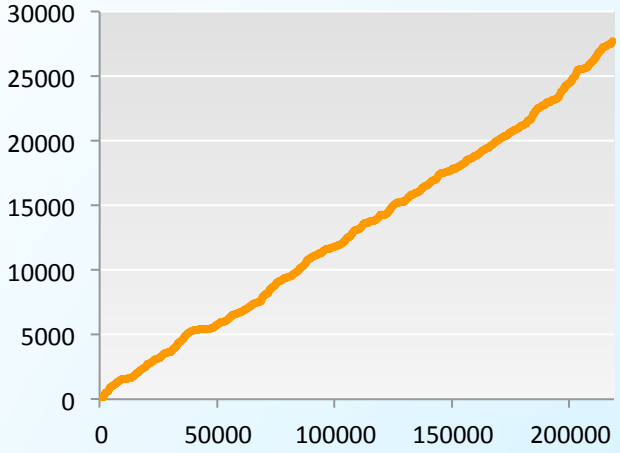
Vertebrates: 2K ✓ **DONE!**

350K Animal Species Analyzed by October 2014

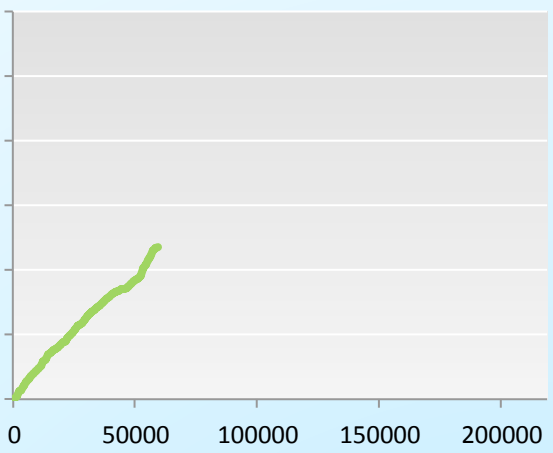
North America



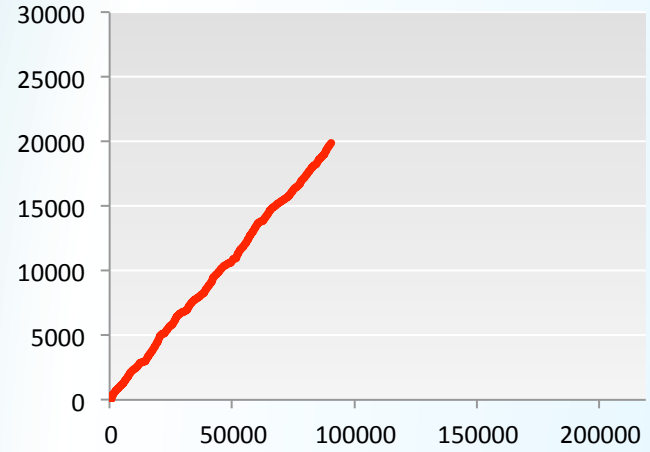
South America



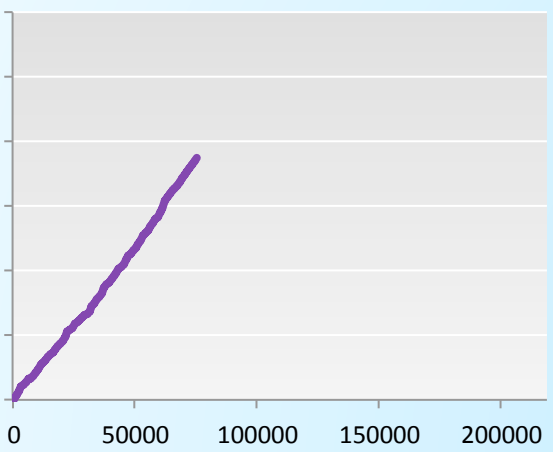
Europe



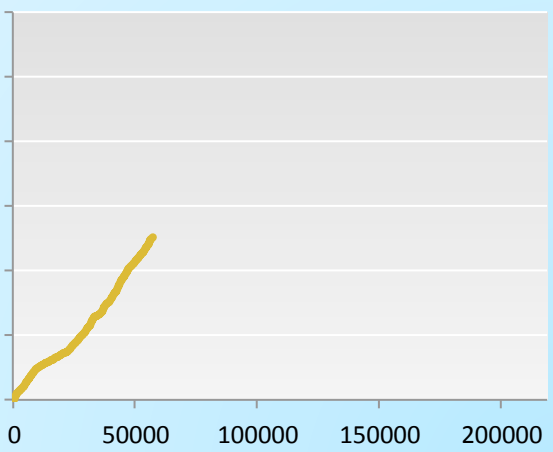
Australia



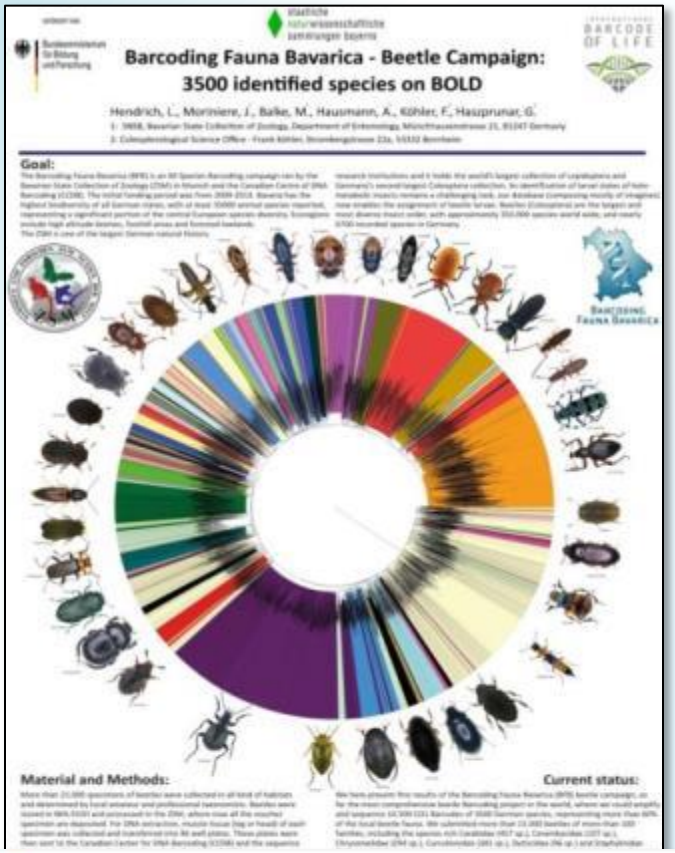
Asia



Africa



Scientific impact of iBOL and CCDB work



Barcoding Fauna Bavarica - Beetle Campaign:
3500 identified species on BOLD

Hendrich, L., Morinière, J., Balke, M., Hausmann, A., Köhler, F., Hausprunar, G.

Goal:

Material and Methods:

Current status:

OPEN ACCESS Freely available online

PLOS ONE

A Transcontinental Challenge — A Test of DNA Barcode Performance for 1,541 Species of Canadian Noctuoidea (Lepidoptera)

Reza Zahiri[1]*, J. Donald Lafontaine[2], B. Christian Schmidt[3], Jeremy R. deWaard[1], Evgeny V. Zakharov[1], Paul D. N. Hebert[1]

1 Biodiversity Institute of Ontario, University of Guelph, Guelph, Ontario, Canada, 2 Canadian National Collection of Insects, Arachnids, and Nematodes, Agriculture and Agri-Food Canada, Ottawa, Ontario, Canada, 3 Canadian National Collection of Insects, Arachnids, and Nematodes, Canadian Food Inspection Agency, Ottawa, Ontario, Canada

- Beetles of Germany – **3.5K** (Hendrich et al. in press)
- Noctuoid Moths of Canada – **1.5K** (Zahiri et al. 2014)
- Beetles of Finland – **1.6K** (Pentinsaari et al. in press)
- Moths of Austria/Finland – **1K** (Huemer et al. in review)
- Bees of Germany – **600** (Schmidt et al. in prep)
- Spiders of Canada – **1K** (Blagoev et al. in prep)

Barcode coverage for birds makes identification easy

Molecular Ecology Notes (2007) 7, 535–543 doi: 10.1111/j.1471-8286.2006.01670.x

BARCODING

Comprehensive DNA barcode coverage of North American birds

KEVIN C. R. KERR,* MARK Y. STOECKLE,† CARLA J. DOVE,‡ LEE A. WEIGT,‡ CHARLES M. FRANCIS§ and PAUL D. N. HEBERT*

*Department of Integrative Biology, Biodiversity Institute of Ontario, University of Guelph, Guelph, Ontario, Canada N1G 2W1, †Program for the Human Environment, The Rockefeller University, New York, NY 10021, USA ‡Smithsonian Institution, National Museum of Natural History, Washington, DC 20013-7012, USA, §National Wildlife Research Centre, Canadian Wildlife Service, Environment Canada, Ottawa, Ontario, Canada K1A 0H3

Abstract

DNA barcoding seeks to assemble a standardized reference library for DNA-based identification of eukaryotic species. The utility and limitations of this approach need to be tested on well-characterized taxonomic assemblages. Here we provide a comprehensive DNA barcode analysis for North American birds including 643 species representing 93% of the breeding and pelagic avifauna of the USA and Canada. Most (94%) species possess distinct barcode clusters, with average neighbour-joining bootstrap support of 98%. In the remaining 6%, barcode clusters correspond to small sets of closely related species, most of which hybridize regularly. Fifteen (2%) currently recognized species are comprised of two distinct barcode clusters, many of which may represent cryptic species. Intraspecific variation is weakly related to census population size and species age. This study confirms that DNA barcoding can be effectively applied across the geographical and taxonomic expanse of North American birds. The consistent finding of constrained intraspecific mitochondrial variation in this large assemblage of species supports the emerging view that selective sweeps limit mitochondrial diversity.

Keywords: Aves, cryptic species, cytochrome *c* oxidase, DNA barcoding, intraspecific mitochondrial variation, selective sweeps

Received 25 October 2006; revision accepted 20 November 2006

Degraded sample? No problem!

Molecular Ecology Notes (2006) 6, 959–964 doi: 10.1111/j.1471-8286.2006.01470.x

BARCODING

A minimalist barcode can identify a specimen whose DNA is degraded

MEHRDAD HAJIBABAEI,* M. ALEX SMITH,* DANIEL H. JANZEN,† JOSEPHINE J. RODRIGUEZ,‡ JAMES B. WHITFIELD‡ and PAUL D. N. HEBERT*

*Biodiversity Institute of Ontario, Department of Integrative Biology, University of Guelph, Guelph, Ontario, Canada, N1G 2W1, †Department of Biology, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA, ‡Department of Entomology, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA

Abstract

A DNA barcode based on 650 bp of mitochondrial gene cytochrome *c* oxidase I is proving to be highly functional in species identification for various animal groups. However, DNA degradation complicates the recovery of a full-length barcode from many museum specimens. Here we explore the use of shorter barcode sequences for identification of such specimens. We recovered short sequences — i.e. ~100 bp — with a single PCR pass from more than 90% of the specimens in assemblages of moth and wasp museum specimens from which full barcode recovery was only 50%, and the latter were usually less than 8 years old. Short barcodes were effective in identifying specimens, confirming their utility in circumstances where full barcodes are too expensive to obtain and the identification comparisons are within a confined taxonomic group.

Keywords: COI, DNA barcoding, DNA degradation, fish, Lepidoptera, museum specimens, parasitic wasps, taxonomy

Received 3 March 2006; revision received 21 April 2006; accepted 24 May 2006

Degraded sample? No problem!

Extinct *Urania* moth

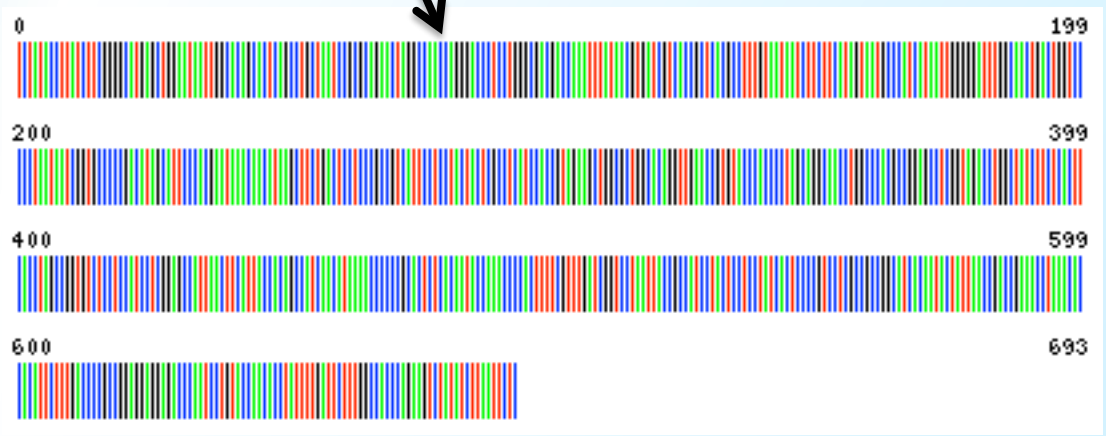
Cultural artifacts

Barcoding type specimens

Dry shark fins

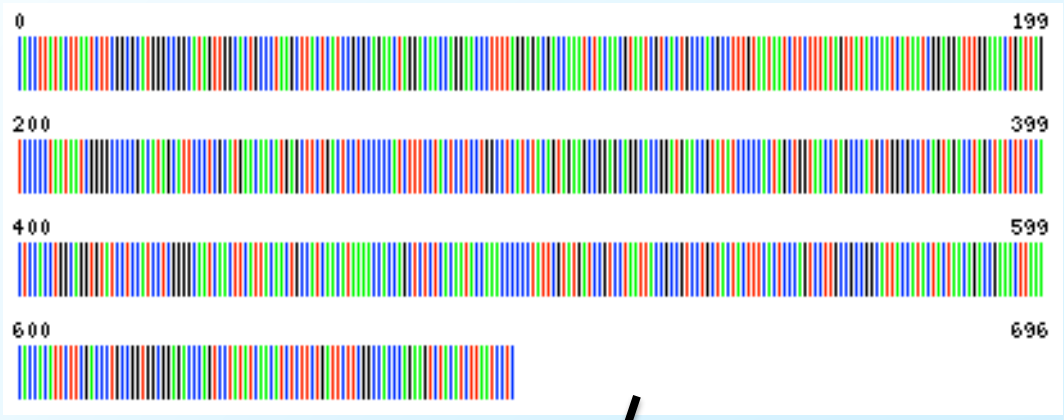
700yo fish bone

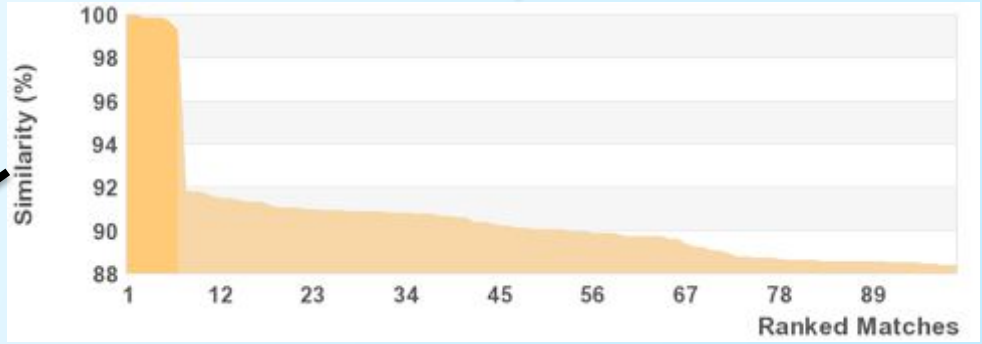
No morphology? No problem!



American wigeon
(*Anas americana*)

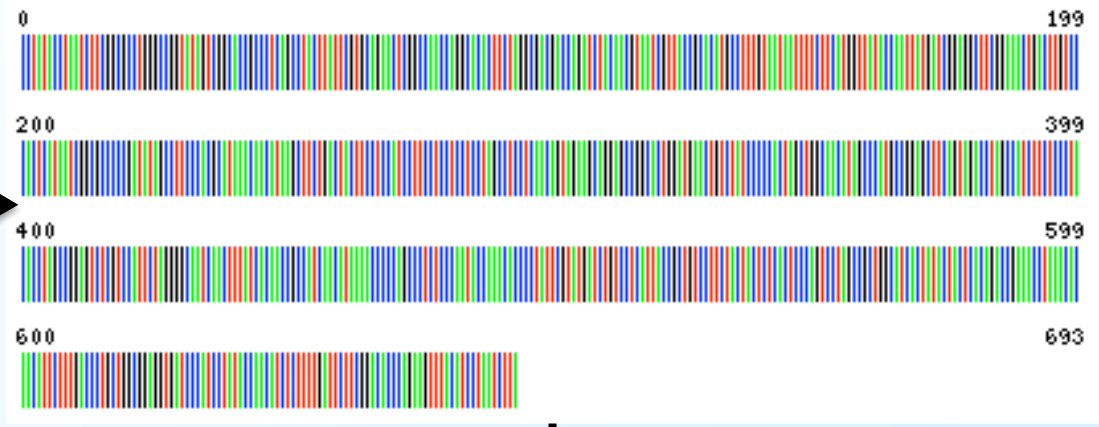
No morphology? No problem!

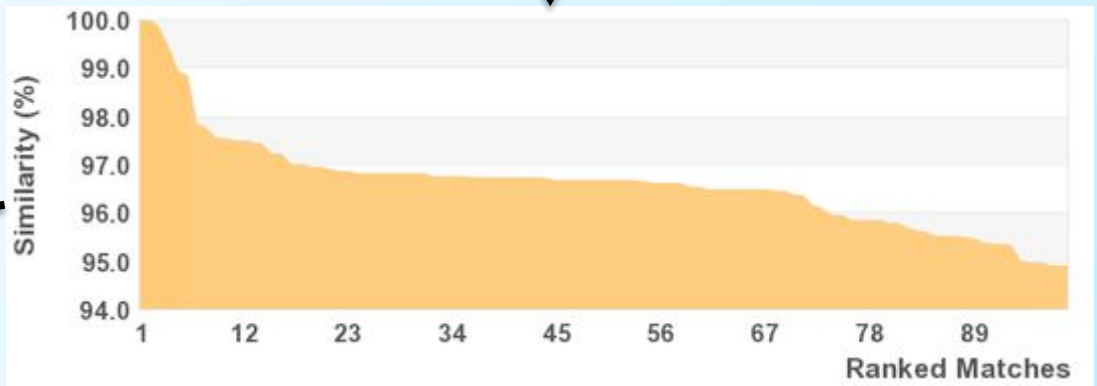




Pied-billed grebe (*Podilymbus podiceps*)

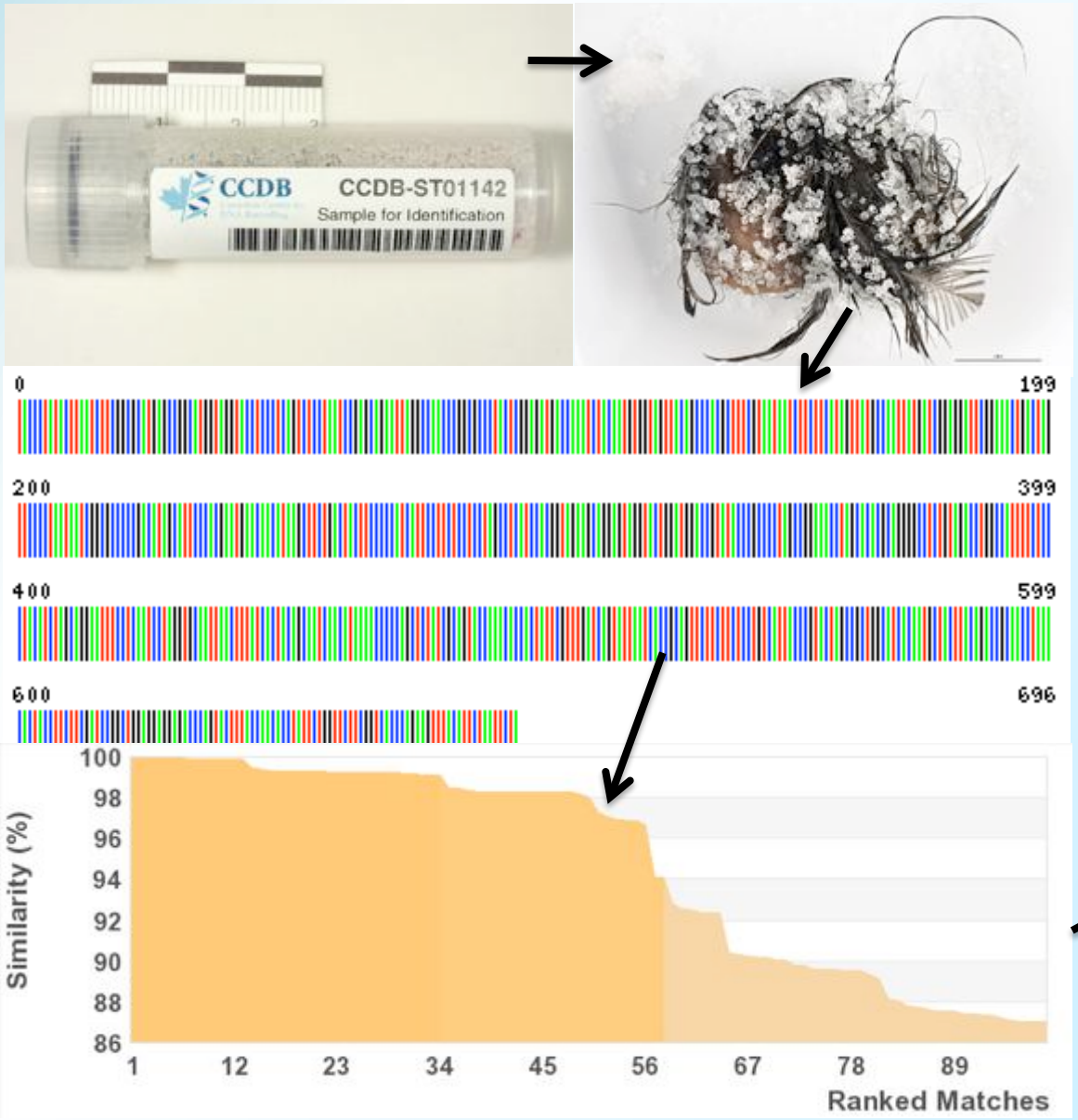
No morphology? No problem!





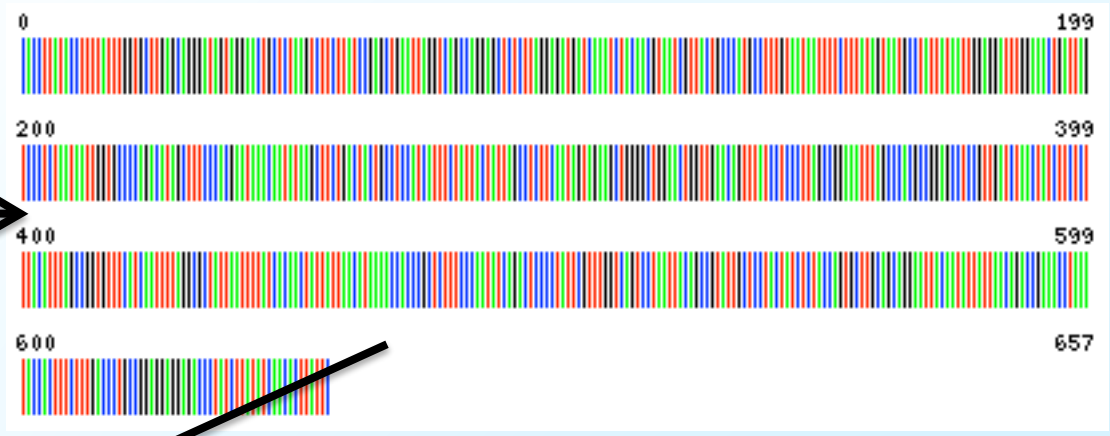
Redtailed hawk (*Buteo jamaicensis*)

No morphology? No problem!



Barn swallow
(*Hirundo rustica*)

It is not a bird! And it is not from BC!



Query: ABCBF208-14|Tadarida brasiliensis

Top Hit: Chordata - Chiroptera - Tadarida brasiliensis (100%)

Search Request:

Type : COI FULL DATABASE (includes records without species designation)

Search Result:

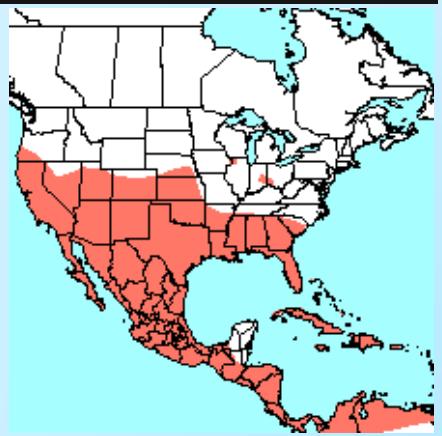
Tree Based Identification

Similarity scores of the top 99 matches:

TOP 20 Matches :

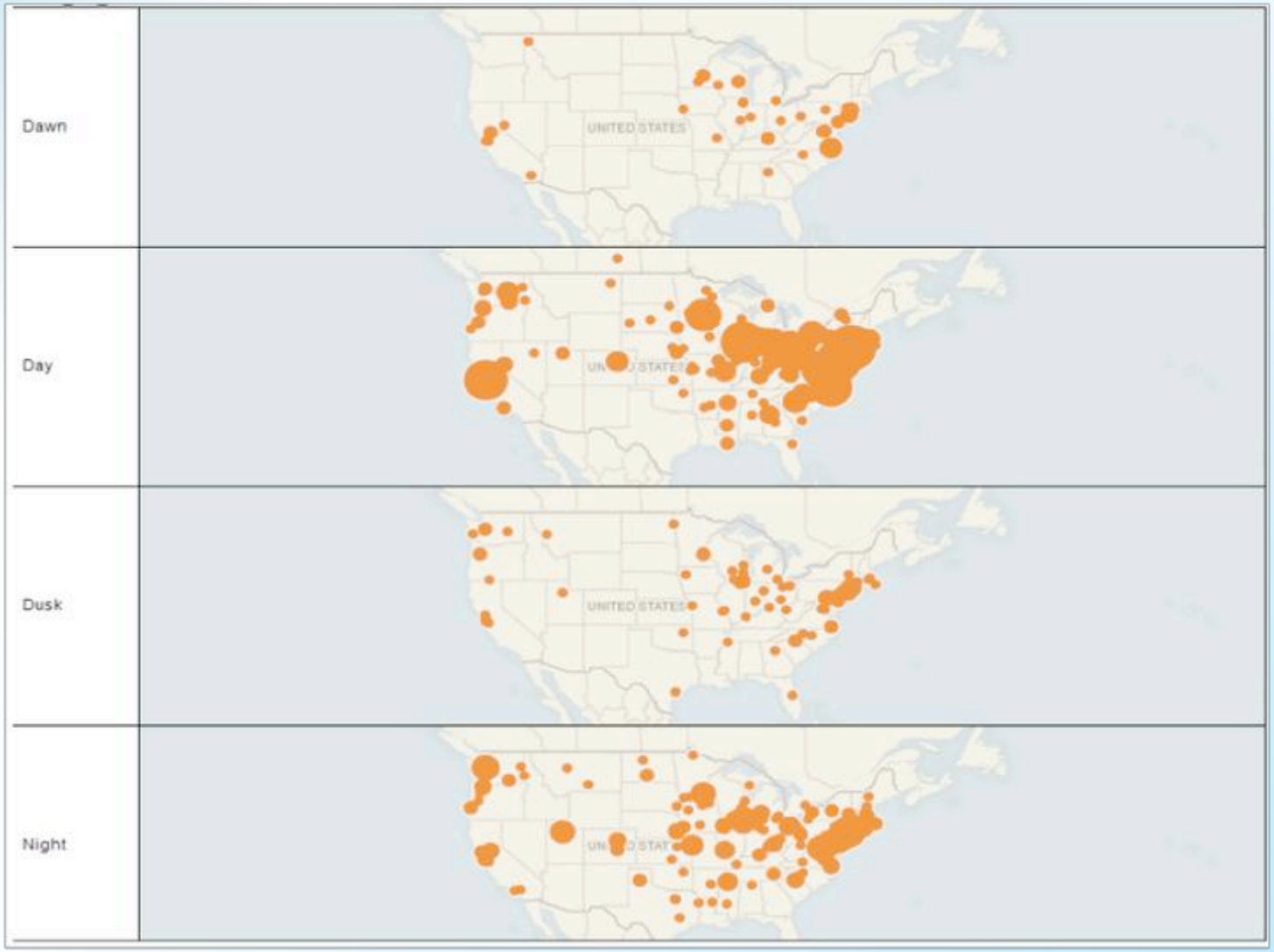
Display option: Top 20

Phylum	Class	Order	Family	Genus	Species	Similarity (%)	Status
Chordata	Mammalia	Chiroptera	Molossidae	Tadarida	*brasiliensis*	100	Published
Chordata	Mammalia	Chiroptera	Molossidae	Tadarida	*brasiliensis*	100	Published
Chordata	Mammalia	Chiroptera	Molossidae	Tadarida	*brasiliensis*	100	Published
Chordata	Mammalia	Chiroptera	Molossidae	Tadarida	*brasiliensis*	100	Published
Chordata	Mammalia	Chiroptera	Molossidae	Tadarida	*brasiliensis*	100	Published
Chordata	Mammalia	Chiroptera	Molossidae	Tadarida	*brasiliensis*	100	Published



Mexican free tail bat (*Tadaria brasiliensis*)

Increased species-level identification…



…better insights from bird strike “bid data”

The path forward

Molecular Ecology Resources (2010) doi: 10.1111/j.1755-0998.2010.02884.x

DNA BARCODING

Birdstrikes and barcoding: can DNA methods help make the airways safer?

JOHN WAUGH,* MAX W. EVANS,† CRAIG D. MILLAR‡ and DAVID M. LAMBERT§

- Continue educating key stakeholders about benefits of birdstrike monitoring and reporting
- Capture as much data as possible
- Take advantage of all available approaches to obtain species level identification
- Employ standard sampling kit and protocols

Information Division – LifeScanner

App development

Kit design

Marketing Campaign

Kits Available in Canada & USA

Availability to Educational Community

Development of sequencing device

2014

2015

2016

Outreach

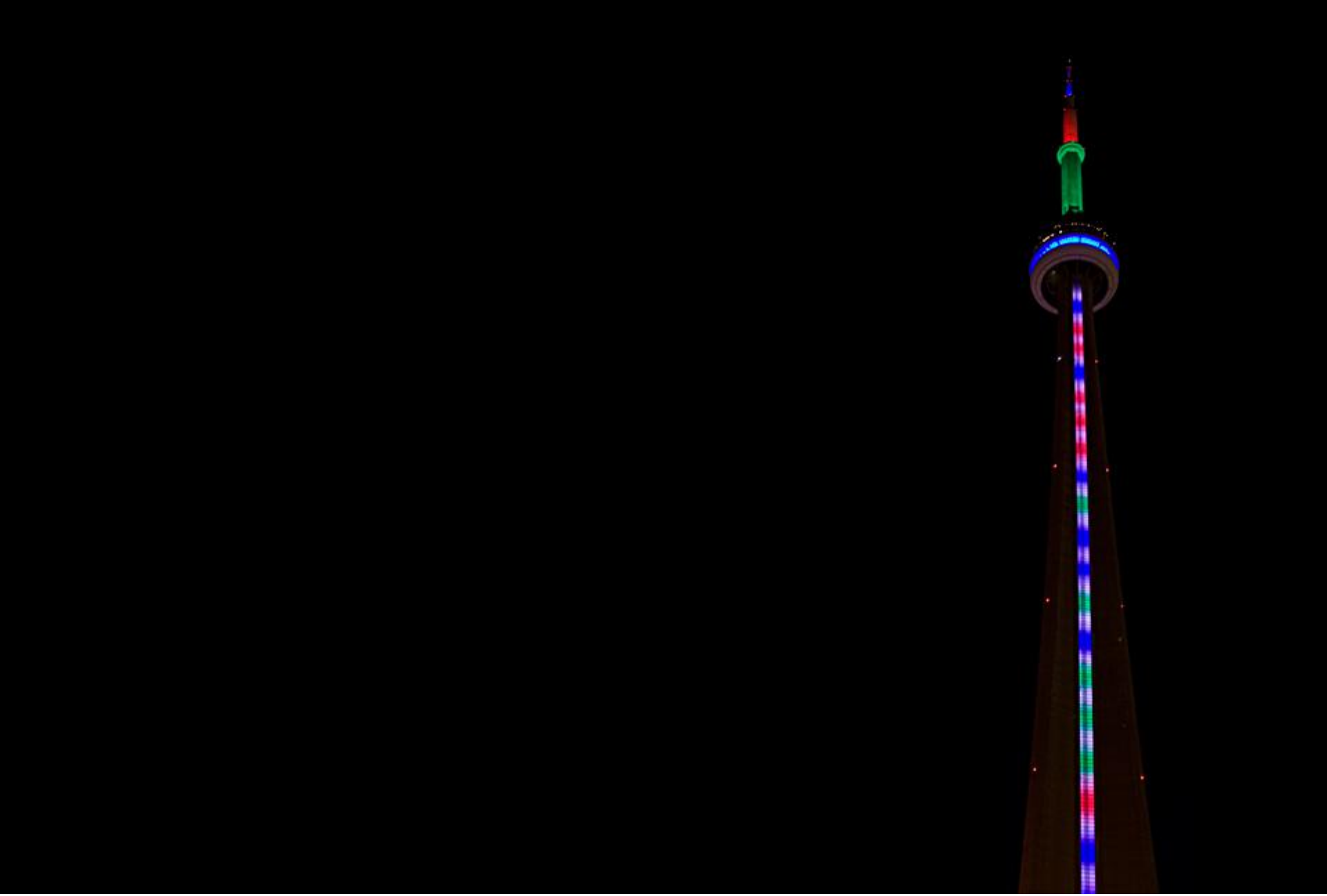
international BARCODE OF LIFE

http://ibol.org/news-and-events/newsletter/
http://dna-barcoding.blogspot.com

Meet the CCDB Team…

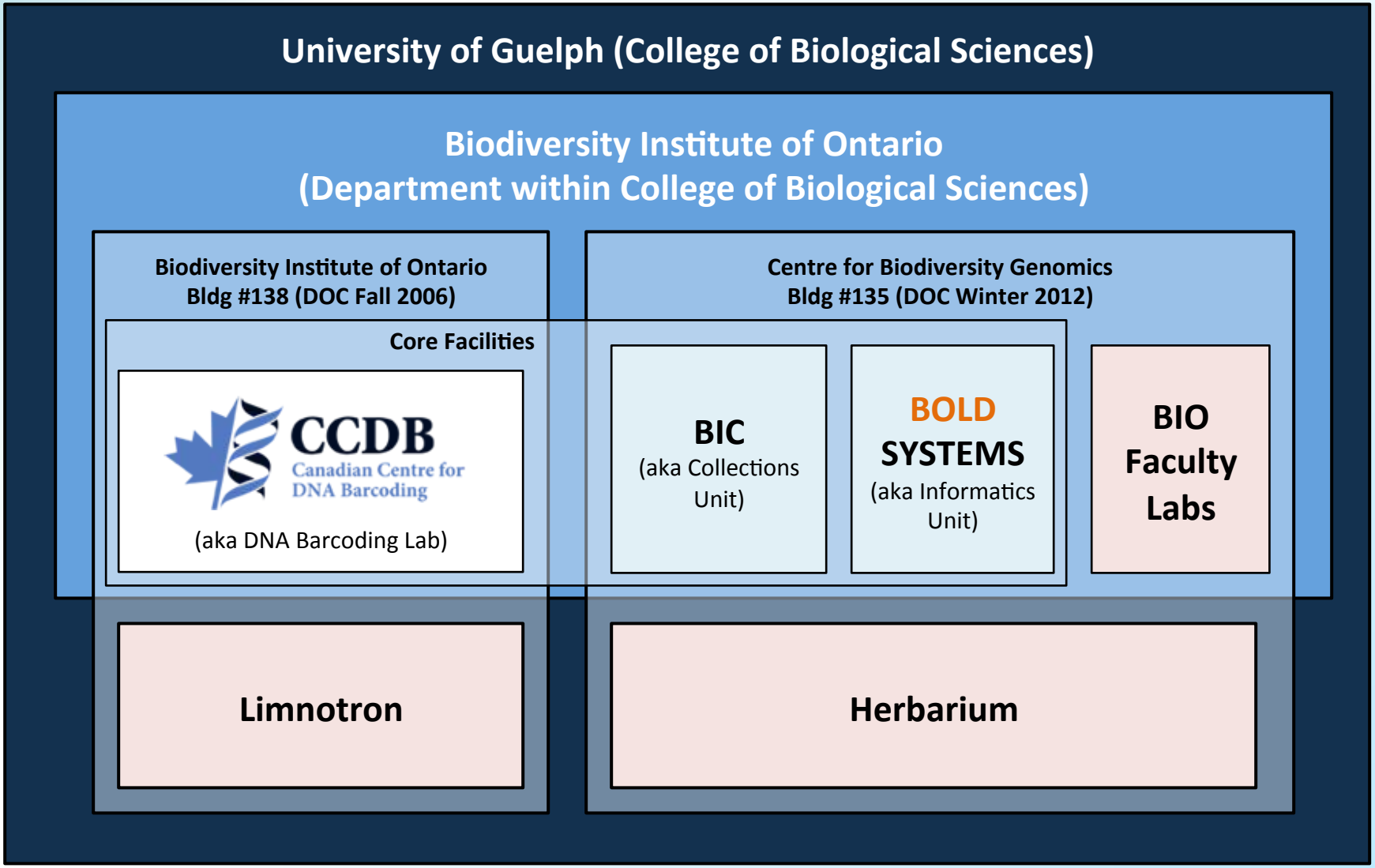
UNIVERSITY of GUELPH BIO CCDB international BARCODE OF LIFE

Funding Provided By:
GenomeCanada
Ontario Ministry of Research and Innovation
OntarioGenomicsInstitute
NSERC CRSNG



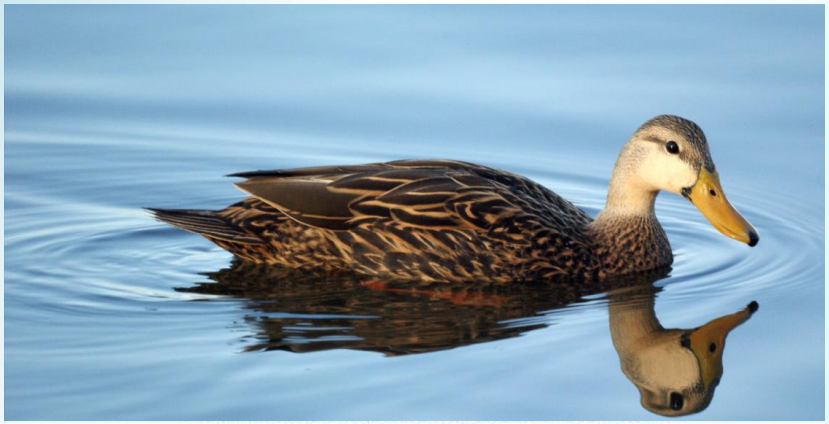
Funding Provided By:

Introductions – Administrative Divisions



The Ugly Duckling

The Ugly Duckling



MUSCOVY DUCK

Cairina moschata

- Has a body like a duck
- Behaves like a goose
- Roosts like a chicken
- Has a breast like a turkey

BIO
Biodiversity Institute of Ontario

[redacted] 2010

Rabbi [redacted]
[redacted]
Brooklyn, NY 11219
Phone: 718 [redacted]
E-mail: [redacted]

Dear Rabbi [redacted]

The DNA barcoding analysis of the duck meat from KIRYAS JOEL POULTRY 845-783-8085 (as per Certificate of Chain of Custody), showed that this meat is derived from either a pure-breeding line of the Muscovy duck (*Carina moschata*) or from a hybrid bird that had a Muscovy duck as its maternal parent.

Please refer to the full report for details.

Sincerely,

Digitally signed by Natalia Ivanova
DN: cn=Natalia Ivanova, c=CA, o=Biodiversity Institute of Ontario, email=nivanova@uoguelph.ca
Date: 2010.0[redacted] 11:06:02 -05'00'

Natalia Ivanova, PhD
Lead DNA Scientist
Canadian Center for DNA Barcoding

Biodiversity Institute of Ontario
University of Guelph
579 Gordon Street
Guelph, Ontario, Canada
N1G 2W1

Phone: +1 (519) 824-4120 x56395
Fax: +1 (519) 824-5703
Website: www.biodiversity.ca

The Ugly Duckling

February 02, 2010

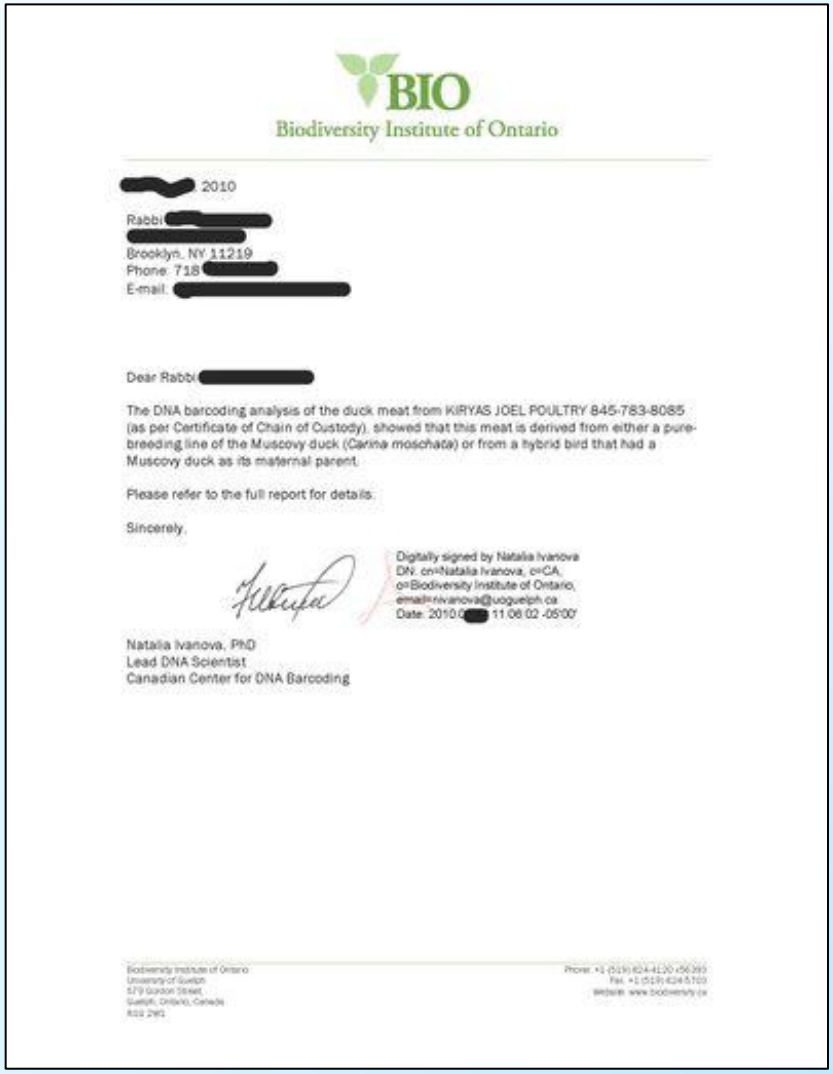
Satmar Poultry Slaughterhouse Duck Recall

A non-kosher duck breed sold as kosher.

After initially denying the problem, last week under pressure Kiryas Joel Poultry issued a duck recall:

A proclamation titled "Notice" in "Der Blatt" of January 15, 2010, KJ Poultry of Monroe, N.Y., signed by management, requests the return of ducks purchased "due to the investigation underway of the ducks delivered by a farmer that may have mixed breeds." The announcement advises that: "Until such time that a clear determination can be made, the ducks are not to be eaten and should be returned."

The determination KJ Poultry wanted has now been made, and the problem, such that it is, is real.



BIO
Biodiversity Institute of Ontario

[redacted] 2010

Rabbi [redacted]
[redacted]
Brooklyn, NY 11219
Phone: 718 [redacted]
E-mail: [redacted]

Dear Rabbi [redacted]

The DNA barcoding analysis of the duck meat from KIRYAS JOEL POULTRY 845-783-8085 (as per Certificate of Chain of Custody), showed that this meat is derived from either a pure-breeding line of the Muscovy duck (Carina moschata) or from a hybrid bird that had a Muscovy duck as its maternal parent.

Please refer to the full report for details.

Sincerely,

Digitally signed by Natalia Ivanova
DN: cn=Natalia Ivanova, c=CA, o=Biodiversity Institute of Ontario, email=nivanova@uoguelph.ca
Date: 2010.[redacted] 11:06:02 -05'00'

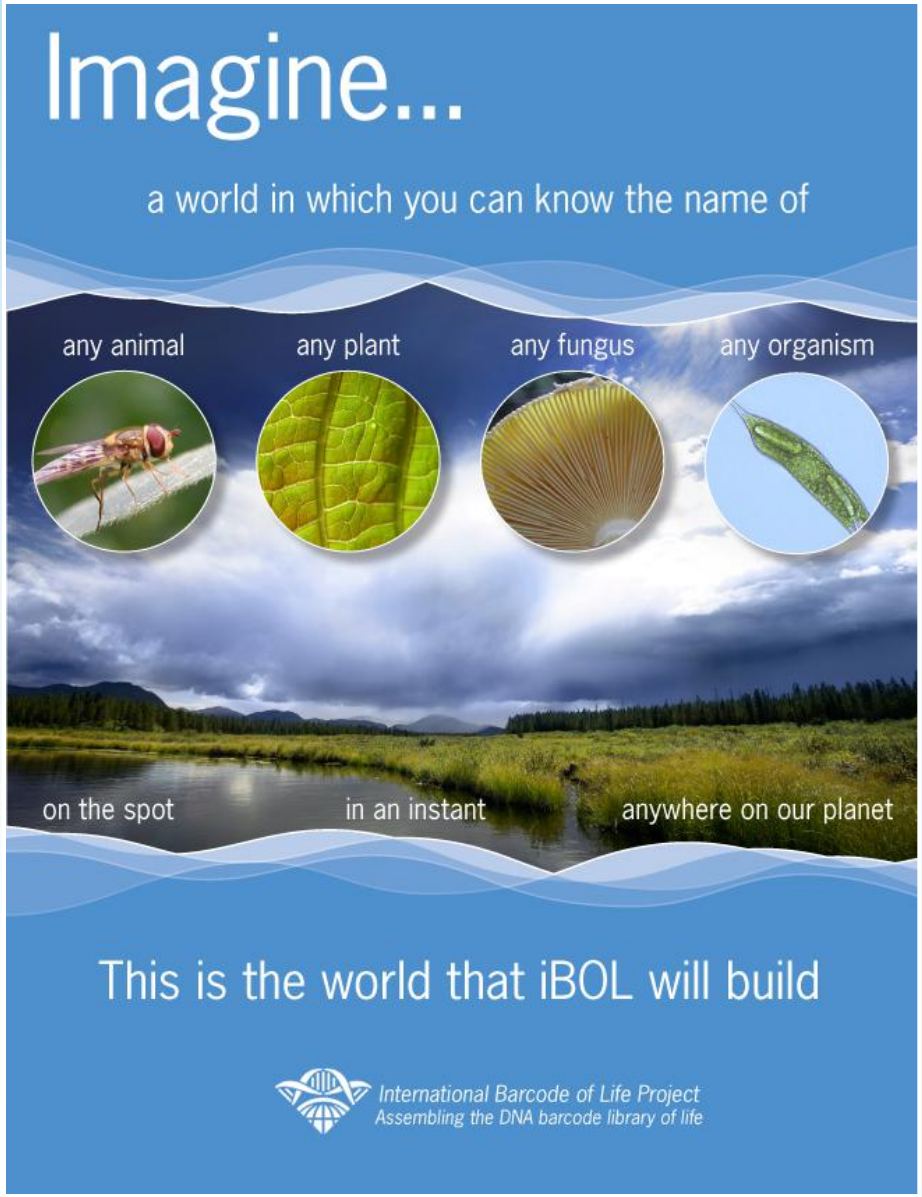
Natalia Ivanova, PhD
Lead DNA Scientist
Canadian Center for DNA Barcoding

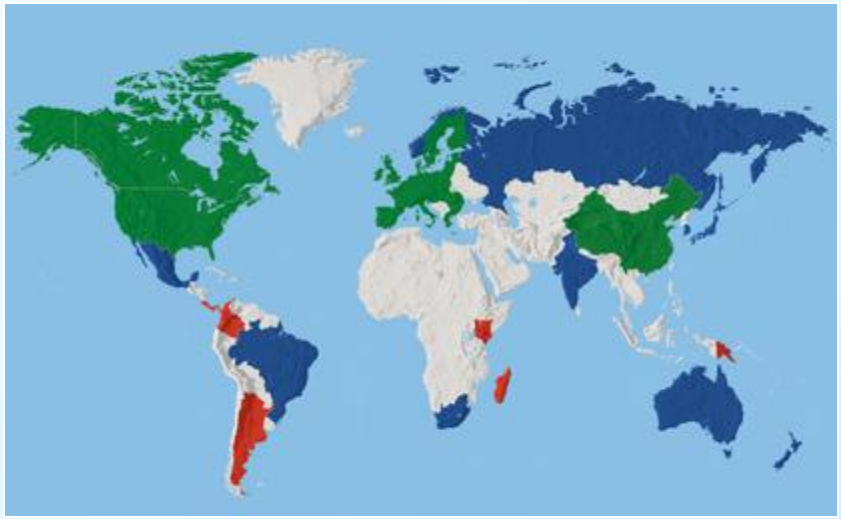
Biodiversity Institute of Ontario
University of Guelph
579 Gordon Street
Guelph, Ontario, Canada
N1G 2W1

Phone: +1 (519) 824-4120 x56395
Fax: +1 (519) 824-5703
Website: www.biodiversity.ca

http://failedmessiah.typepad.com/.a/6a00d83451b71f69e20128774b52fc970c-400wi

International Barcode of Life Project

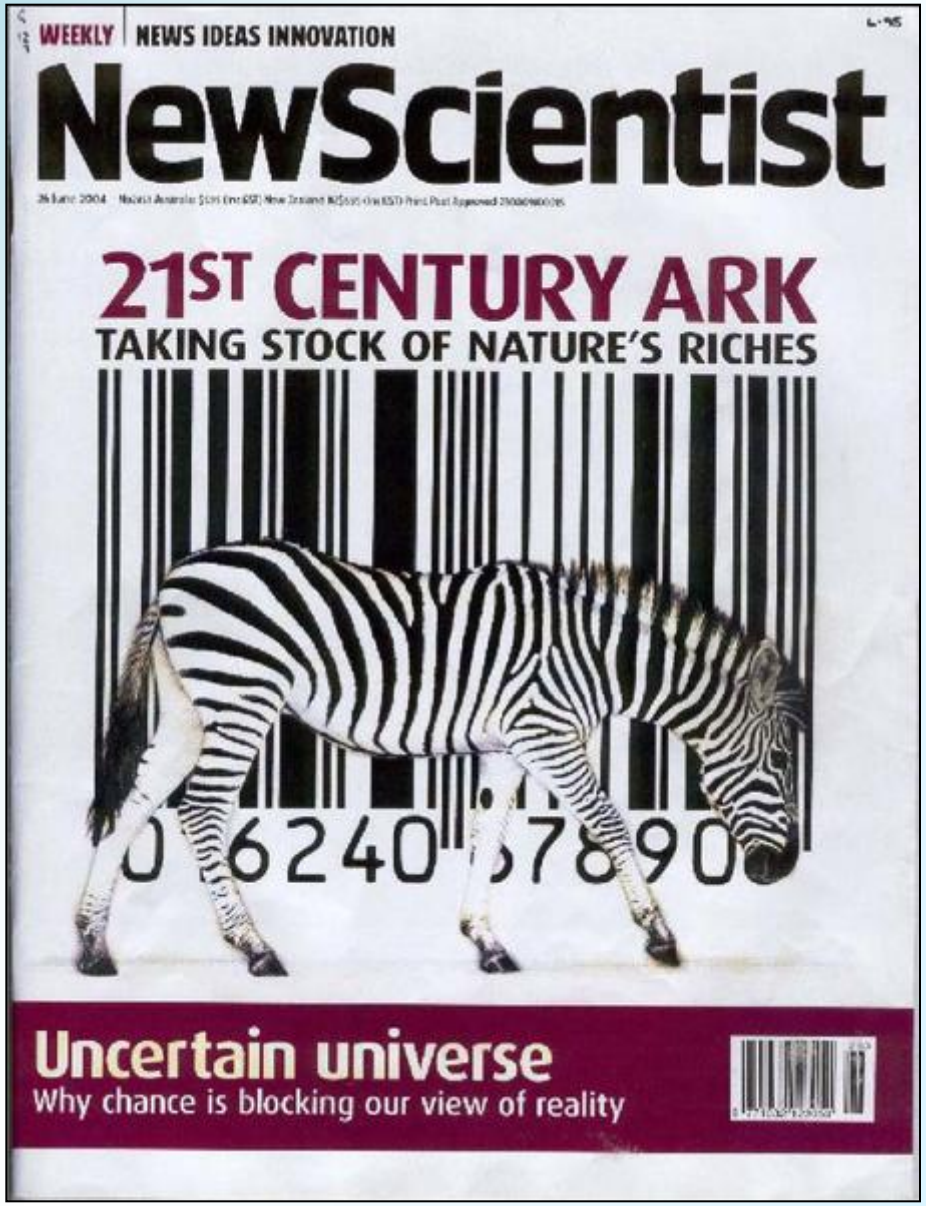




5 Years

5M Specimens

500K Species


WEEKLY | NEWS IDEAS INNOVATION
NewScientist
21ST CENTURY ARK
TAKING STOCK OF NATURE'S RICHES
Uncertain universe
Why chance is blocking our view of reality


SCIENCE NEWS
pheromone keeps bees young
defense against emphysema
random numbers at work
stress is a killer
bar code kingdom