Non-Research Impacts of Major Research Infrastructures: Lessons from the DNA Barcoding Initiative

David E. Schindel, Executive Secretary
Consortium for the Barcode of Life
National Museum of Natural History
Smithsonian Institution
SchindelD@si.edu;
http://www.barcoding.si.edu
202/633-0812; fax 202/633-2938

Overview

- NSF Review Criteria: Congressional pressure
  - Intrinsic scientific merit
  - Wider impact
- Science & Technology Centers
  - Research and Infrastructure
  - Education
  - Technology transfer
- Two distributed research infrastructures
  - DNA Barcodes
  - Scientific Collections

Non-research Impact

- Non-academic sectors
  - Government regulation
  - Other secondary uses for data
- Technology development
- S&T Workforce Training
  - Undergraduate research experiences
  - Doctoral/Post-doctoral training & mobility
- K-12 curriculum and instruction
- General public: museums, media

Why Bother?

- Widens base of stakeholders, funders
- Widens pool of participants
- Generates unanticipated research connections
- Positive media coverage
- Increased visibility, credibility of field
- Attracts more and better students
- Increases motivation of host organization
- Drives cultural change across a field

NSF S&T Centers

- Technology transfer:
  - CCD sensors
  - Image processing software
- Public awareness:
  - Museum exhibits, Radio reports
  - Sunday newspaper supplements
- Workforce training
  - Undergraduate courses, summer projects
  - Interdisciplinary/Mobility experiences for doctoral students

Pre-University Education

- Curriculum and instructional materials
  - Not driven by science, not for outsiders
  - Hire a specialist, rely on partners
- Scientists in the classroom
  - Potential for local impact, visibility
  - Not easily scaleable; proceed with caution
  - Requires staff assigned to outreach
- Teachers in the laboratory

Species Identification Matters

- Basic research on evolution, ecology
- Understanding ecosystem services
- Endangered/protected species
- Agricultural pests/beneficial species
- Disease vectors/pathogens
- Invasive species (e.g., in ballast water)
- Environmental quality indicators
- Managing for sustainable harvesting
- Consumer protection, ensuring food quality
- Fidelity of seedbanks, culture collections

Encyclopedia of Life

Taxonomic revisions, biotic surveys

Published species descriptions

Public database records

Not yet collected

Encyclopedia of Life

Taxonomic revisions, biotic surveys

Published species descriptions

Public database records

Description/revision not yet published

Data not yet released

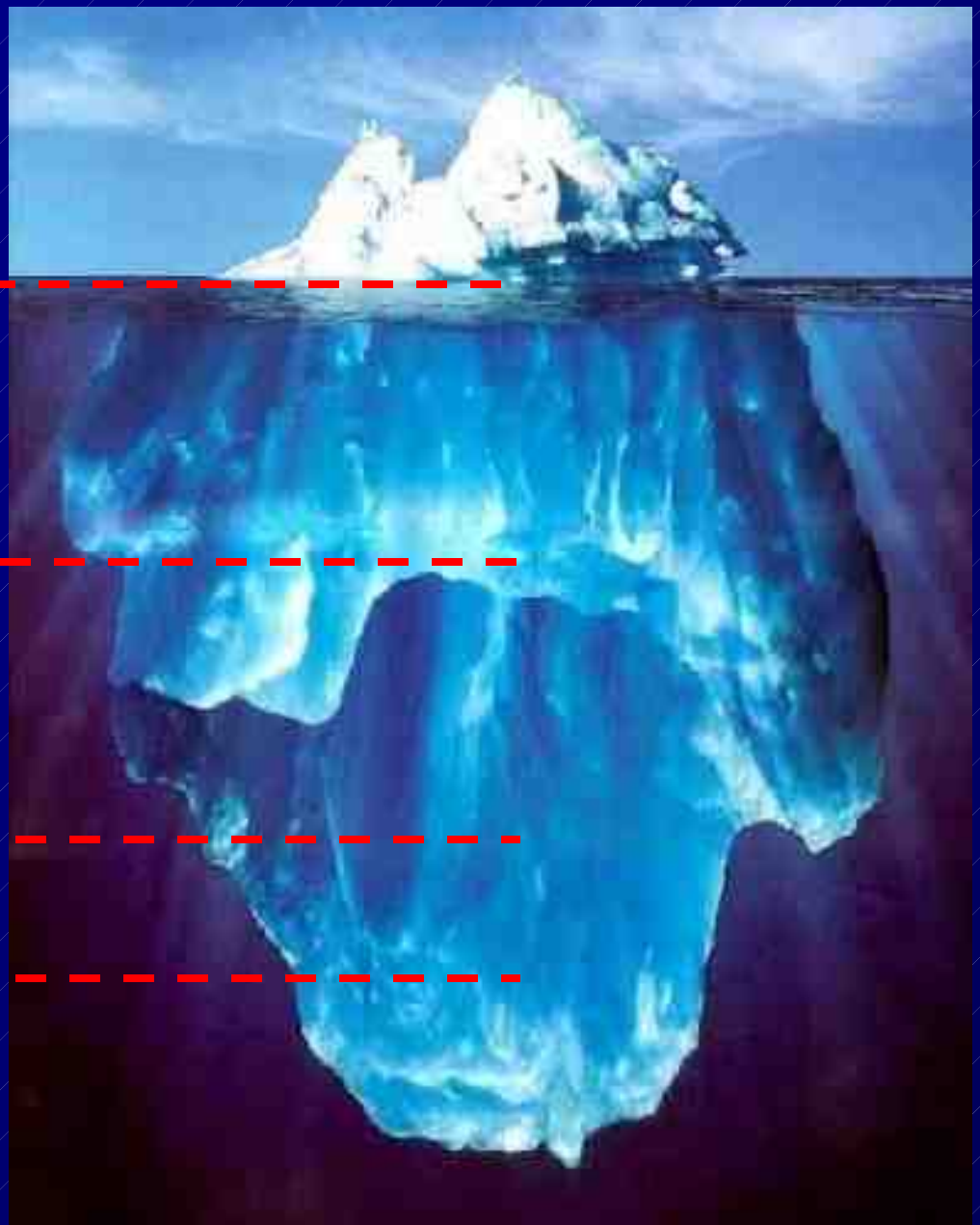
Not yet described

Not yet in specimen catalog

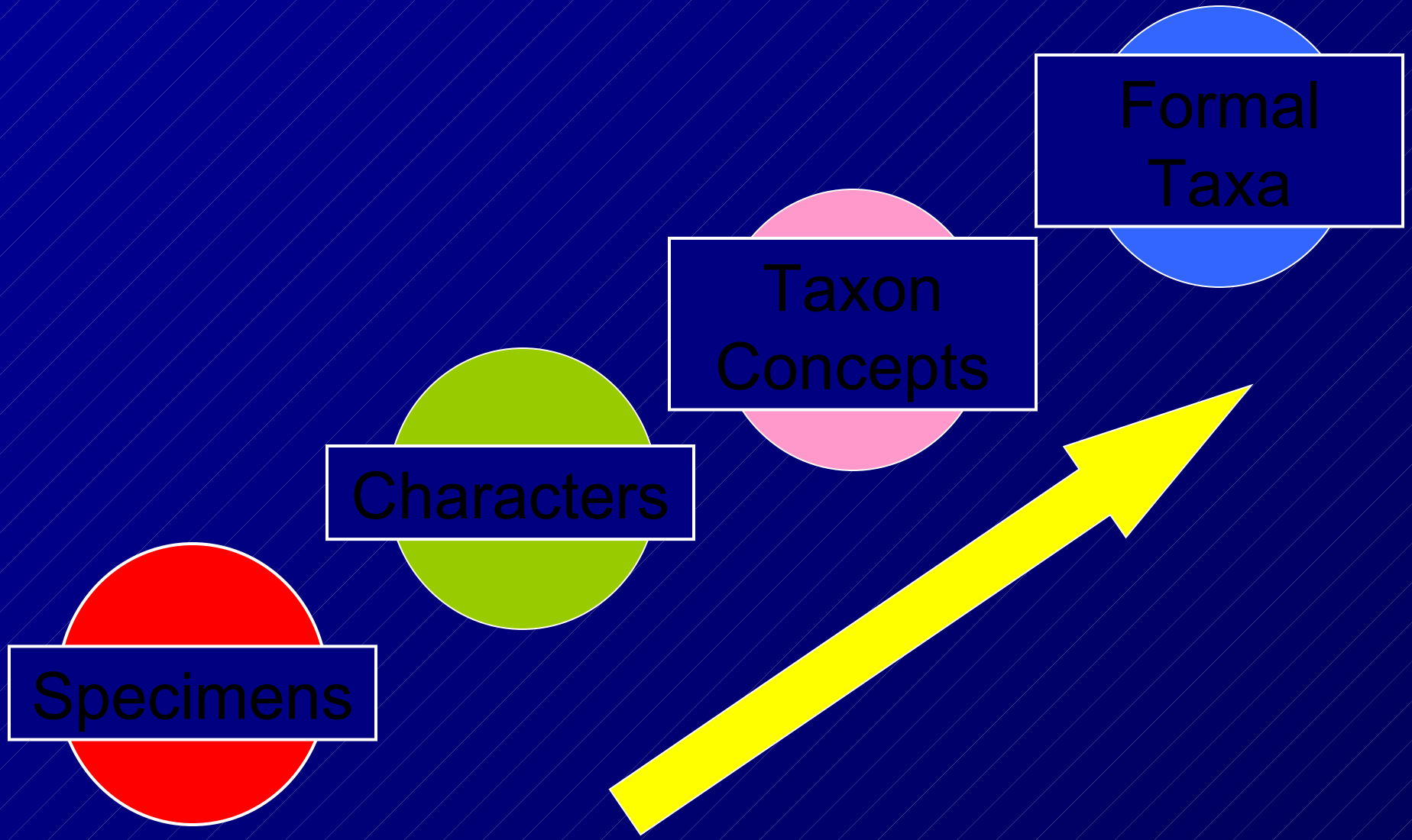
Not yet examined

Not yet curated

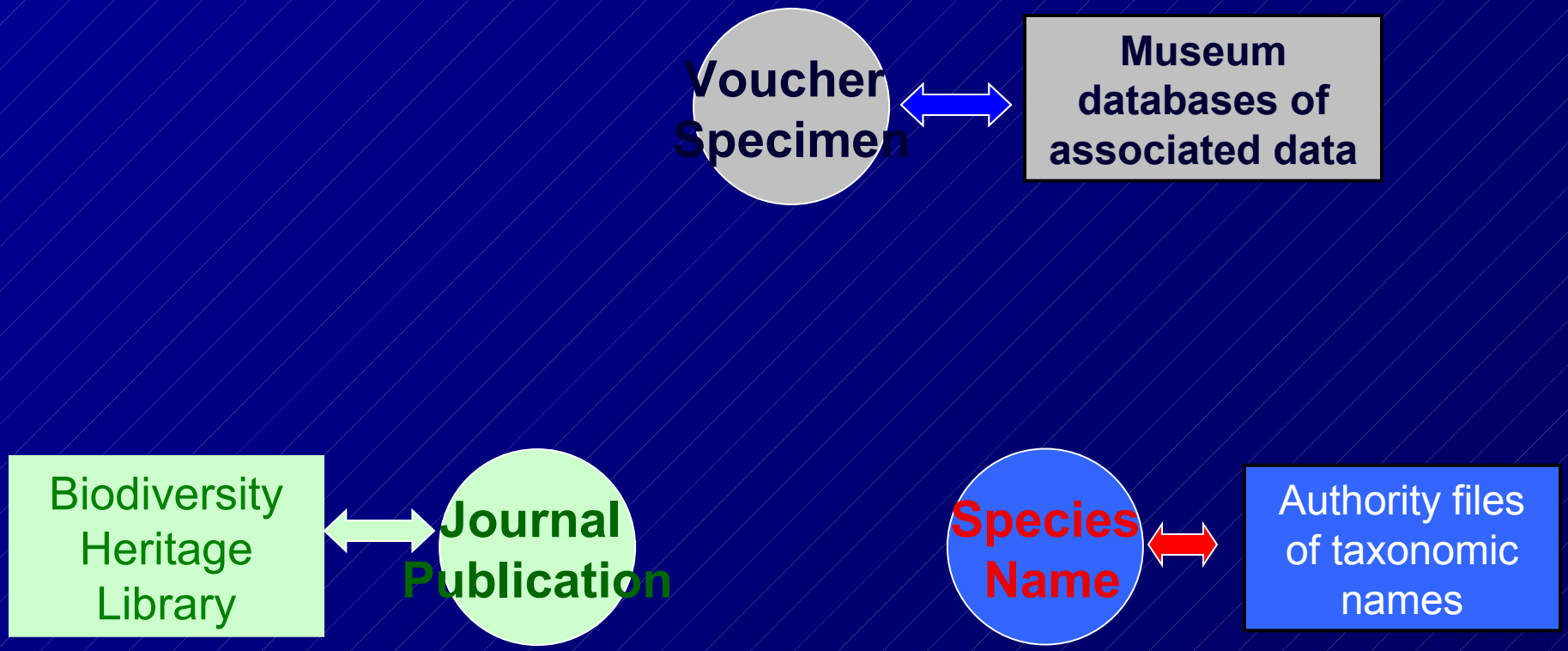
Not yet collected



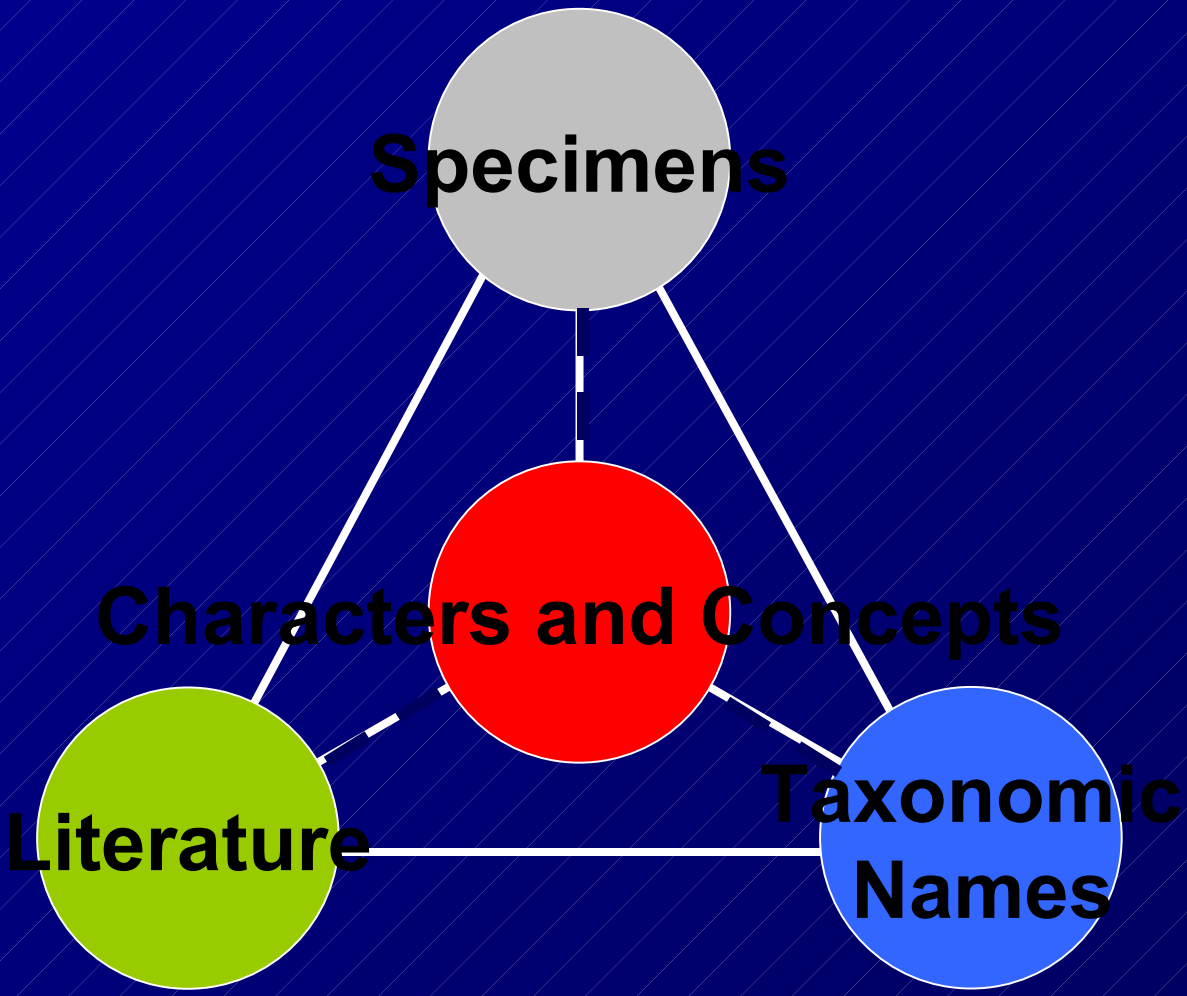
Taxonomic Processes



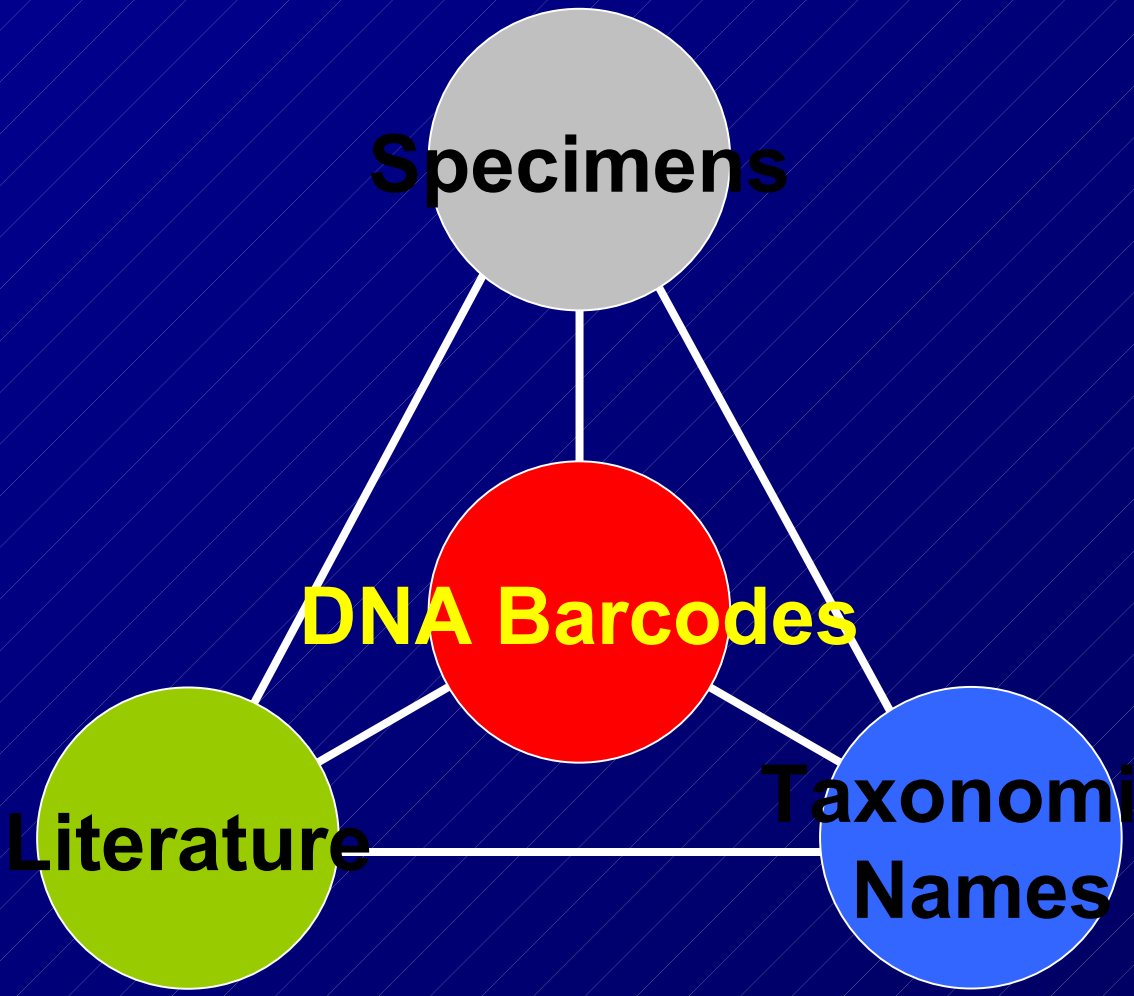
Growth of Biodiversity Databases



Taxonomic Inference



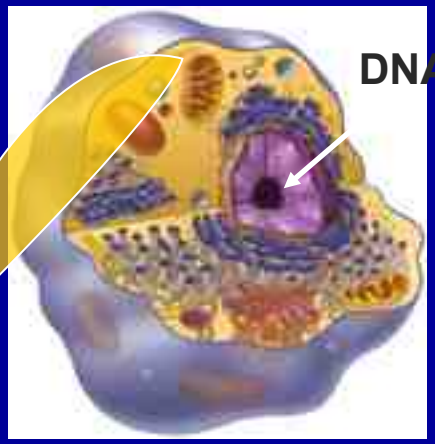
DNA Barcodes as Key Variable



A DNA barcode is a short gene sequence taken from standardized portions of the genome, used to identify species

Associating Life Stages, Processed Parts, Dimorphic Genders

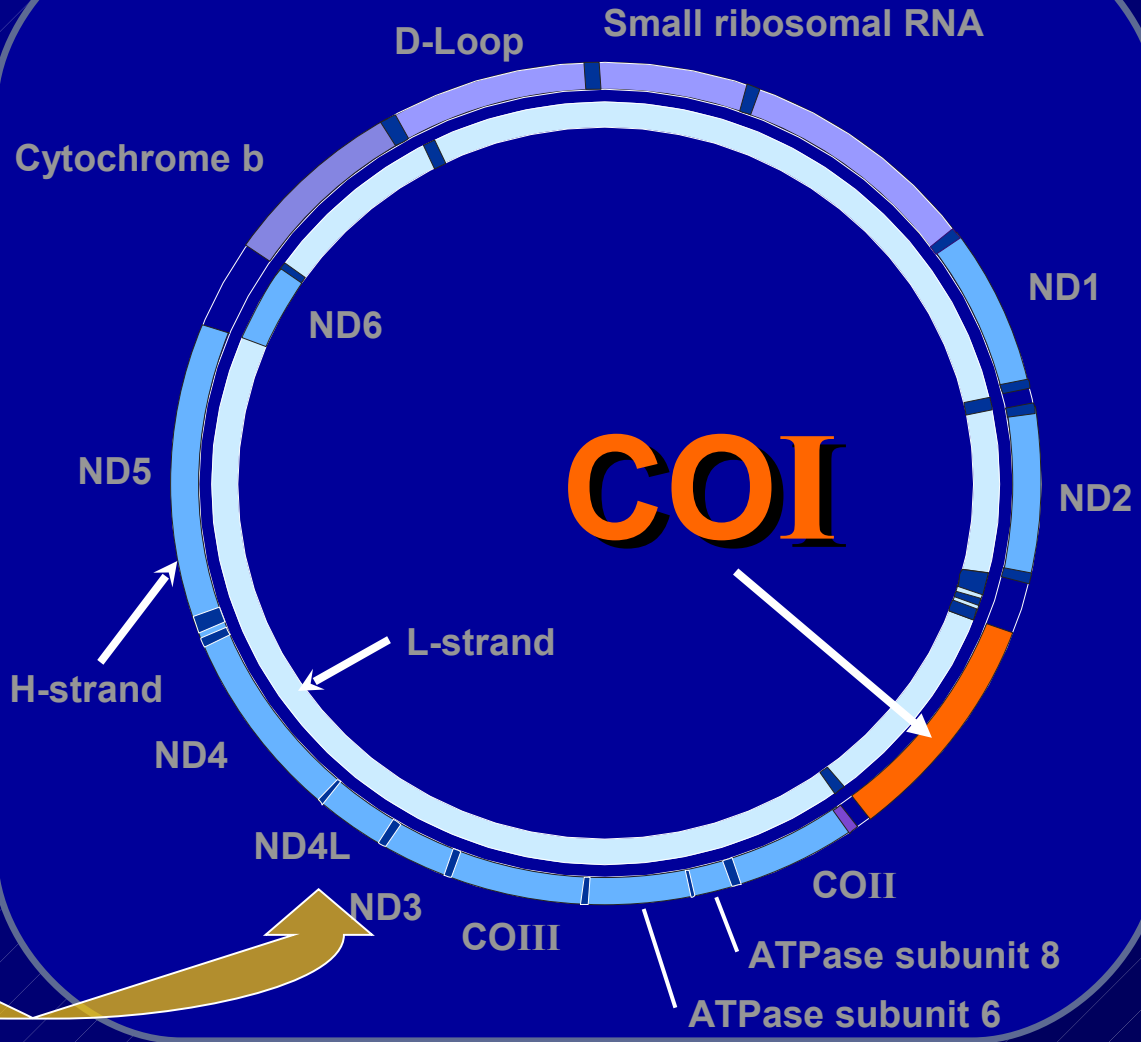
An Internal ID System for All Animals



Typical Animal Cell

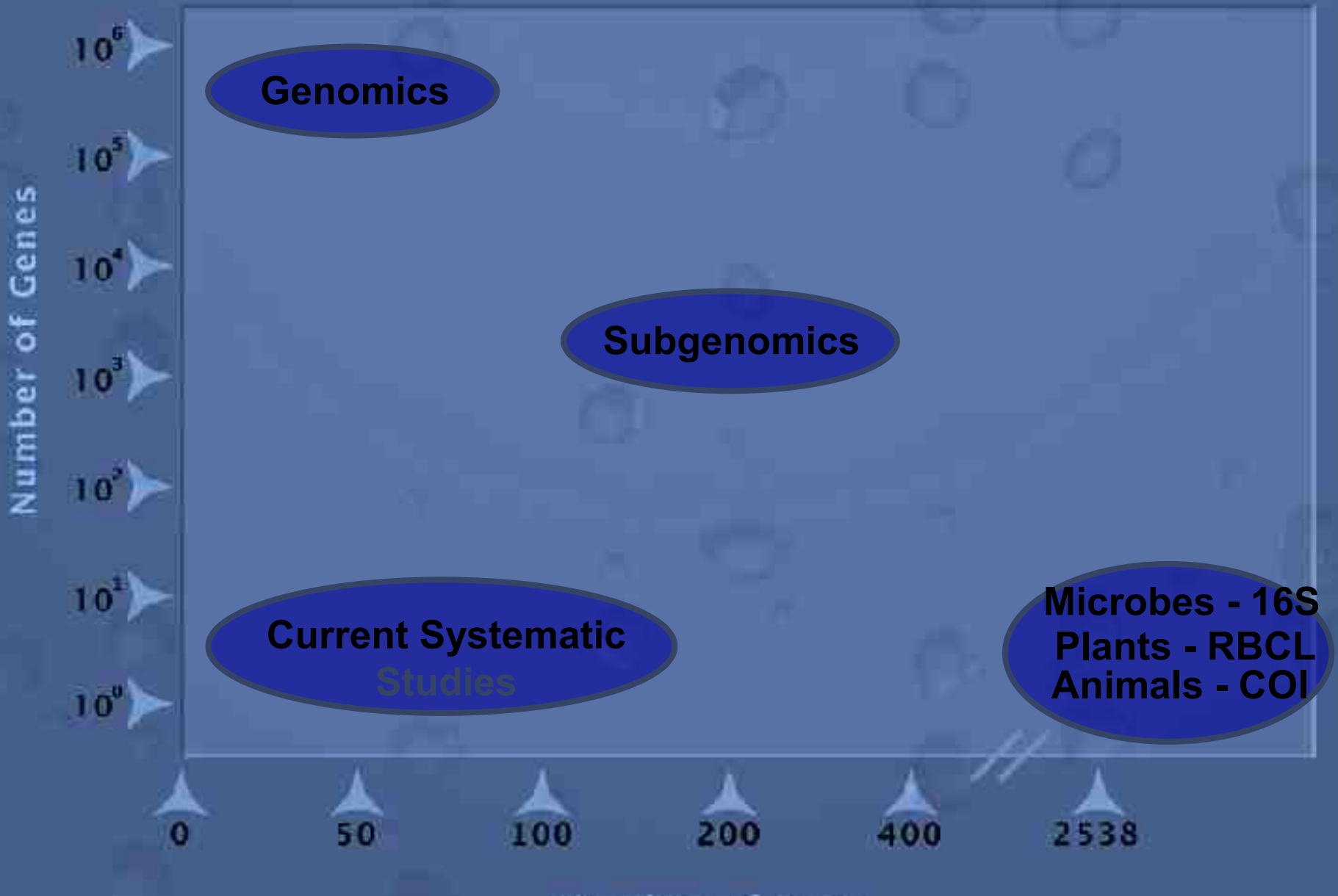
Mitochondrion

The Mitochondrial Genome



Non-COI regions for other taxa

- Land plants:
  - Chloroplast *matK* and *rbcL* approved Nov 09
  - Non-coding plastid and nuclear regions being explored
- Fungi and protists:
  - CBOL Working Groups convened
  - Recommendations expected in 2010

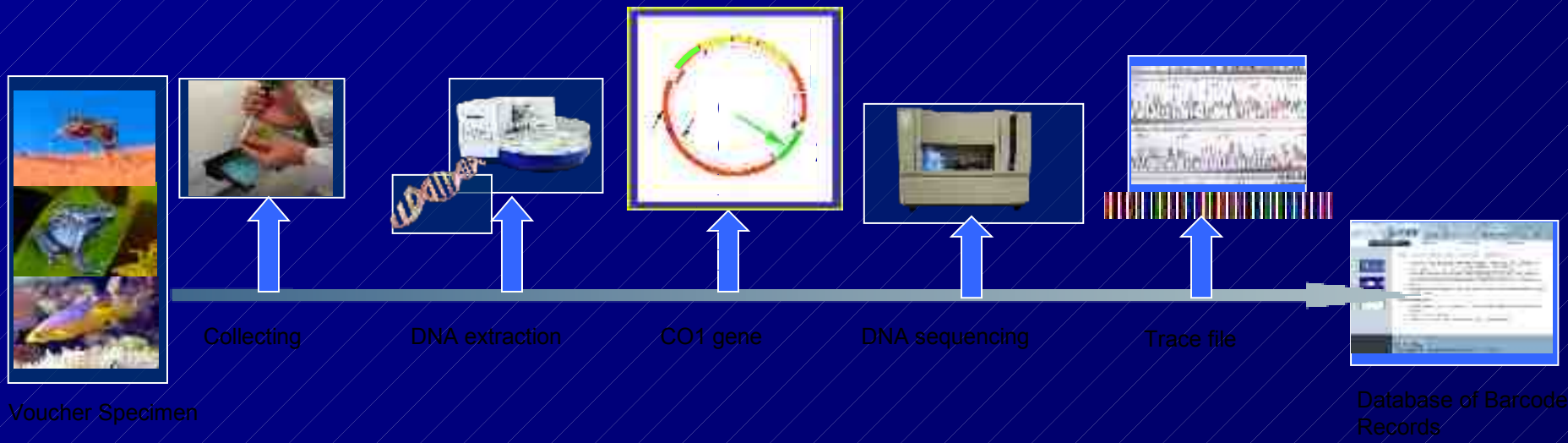

Genomics
Subgenomics
Current Systematic Studies
Microbes - 16S
Plants - RBCL
Animals - COI
Number of Genes
Number of Taxa
0
50
100
200
400
2538

How Barcoding Works

- PHASE 1: Build a barcode reference library:
  - Well-identified specimen
  - Tissue subsample
  - DNA extraction, PCR amplification
  - DNA sequencing
  - Data submission to GenBank
- PHASE 2: Use it to identify unknowns:
  - Any unidentified juvenile, adult, fragment, product
  - Tissue sample, DNA, sequencing
  - Comparison with sequences in reference library

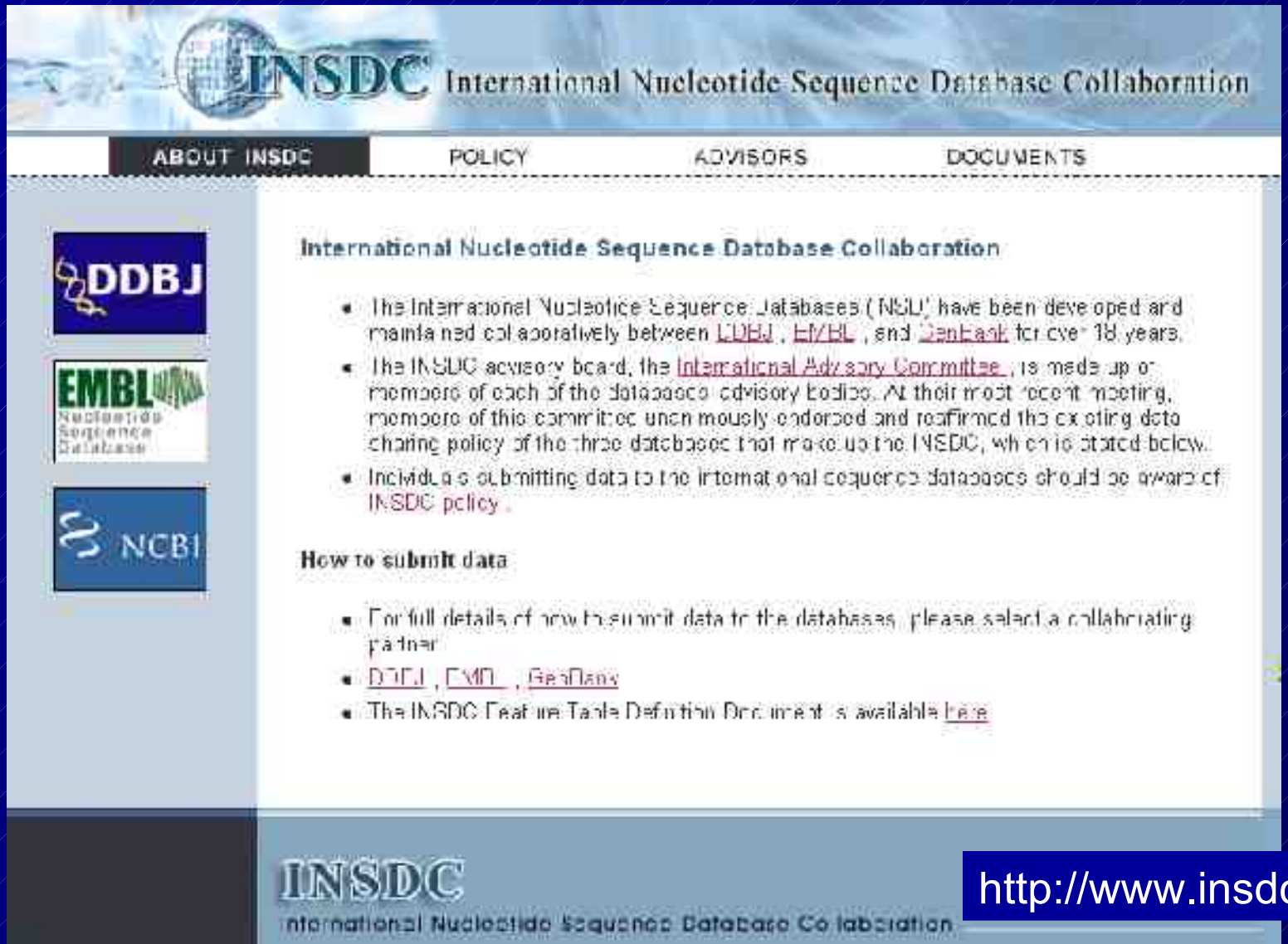
How Barcoding is Done

From specimen to sequence to species



GenBank, EMBL, and DDBJ

Global, Open Access to Barcode Data



http://www.insdc.org/

Barcode of Life Community

800,000 DNA barcode records from 75,000 species

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

ABB

BOLDSYSTEMS

Canadian Barcode of Life Network

MBI

TBI

CCDB

BIO

Consortium for the Barcode of Life (CBOL)

- Established May 2004 with Sloan Foundation grant
- Secretariat opens at Smithsonian, September 2004
- Now in its third two-year funding period
- Workshops, Working Groups, networking, representation/marketing
- Now an international affiliation of 200+ members in 50+ countries:
  - Natural history museums, biodiversity organizations
  - Users: e.g., government agencies
  - Private sector biotech companies, database providers

CBOL Member Organizations: 2009

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

Outreach Activities

- Cape Town, South Africa, April 2006, SANBI
  - Scale insects in African agriculture
- Nairobi, Kenya, October 2006
  - Commercial fisheries in Rift Valley lakes
- Brazil, March 2007
  - Hardwood tree species
  - Endangered mammals, reptiles, amphibians
- Taiwan, September 2007
- Nigeria, October 2008
- Beijing, May 2009
- India, March 2010

Leading Labs Network

CBOL's Global Projects

- Fish Barcode of Life (FISH-BOL)
  - 30,000 marine/freshwater species by 2010
- All Birds Barcoding Initiative (ABBI)
  - 10,000 species by 2010
- Tephritid fruit flies
  - 2,000 pest/beneficial species and relatives by 2008
- Mosquitoes
  - 3,300 species by 2008
- Endangered species
- Trees of the world

Adoption by Regulators

- Food and Drug Administration
  - Reference barcodes for commercial fish
- NOAA/NMFS
  - $100K for Gulf of Maine pilot project
  - FISH-BOL workshop with agencies, Taipei, Sept 2007
- Federal Aviation Administration – $500K for birds
- Environmental Protection Agency
  - $250K pilot test, water quality bioassessment
- FAO International Plant Protection Commission
  - Proposal for Diagnostic Protocols for fruit flies
- CITES, National Agencies, Conservation NGOs
  - International Steering Committee, identifying pilot projects

Current Norm: High throughput

Large labs, hundreds of samples per day

ABI 3100 capillary

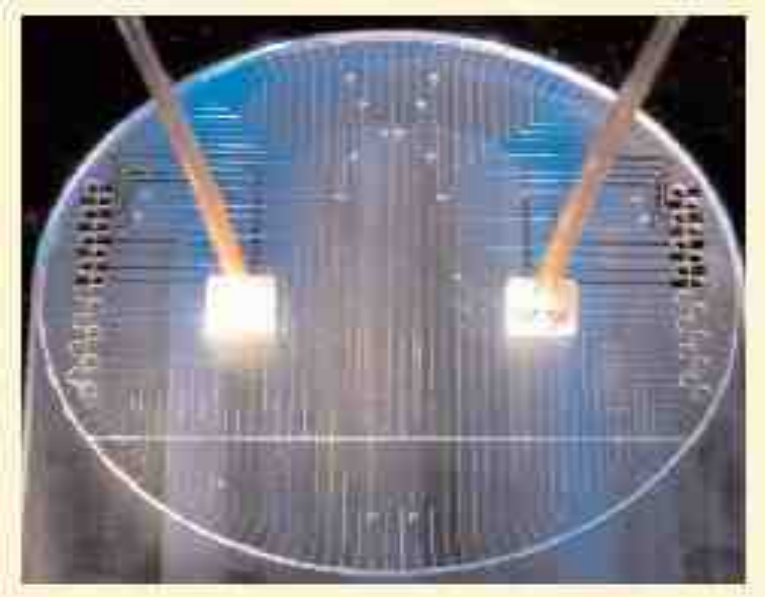
automated sequencer

Large capacity PCR and

sequencing reactions

Emerging Norm: Table-top Labs

Faster, more portable: Hundreds of samples per hour



Integrated DNA microchips

Table-top microfluidic systems

Producing Barcode Data: 201?

Barcode data anywhere, instantly

- Data in seconds to minutes
- Pennies per sample
- Link to reference database
- A taxonomic GPS
- Usable by non-specialists

Applied Biosystems

Technology Development Partnership Goal

The DNA Sequencing Lab of 2013?

engaging schools in UK plant barcoding

Karen James
Natural History Museum, London

Project elements

- DNA barcoding (data pipeline)
- workshop planning and implementation (NHM staff)
- preliminary surveys/feasibility studies
- child-friendly keys
- management of expectations
- new facilities, on-site curation
- quality control
- working with educators
- online support (wikis, tutorials etc.)

Lessons learned

- successful shift in funding approach (private educational charity, ~£400K over 3 yrs)
- balancing scientific and educational aims (feedback from science teachers)
- need for detailed project planning to ensure educational and scientific continuity (e.g. mapping barcoding workflow onto the workshop narrative)

Canadian National Fish Survey

- Secondary and university students
- Hands-on learning:
  - Market and restaurant sampling program
  - Classroom sample preparation
  - Submit specimen data online
  - Sample processing at university
- Online learning
  - Data analysis on public workbench
  - Communication with professional taxonomists
- Consumer impact through reporters, government agencies

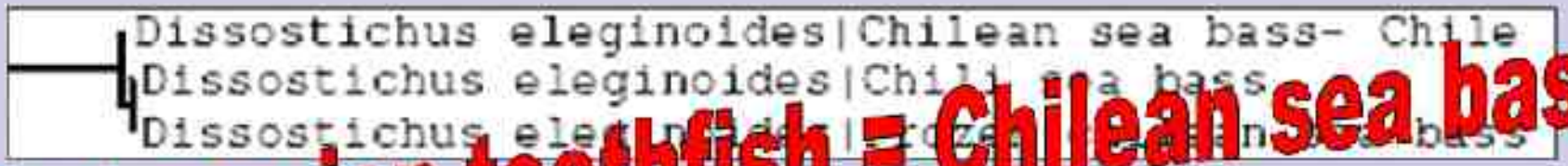
.: Economic Impact?

Acadian Redfish

Red Snapper

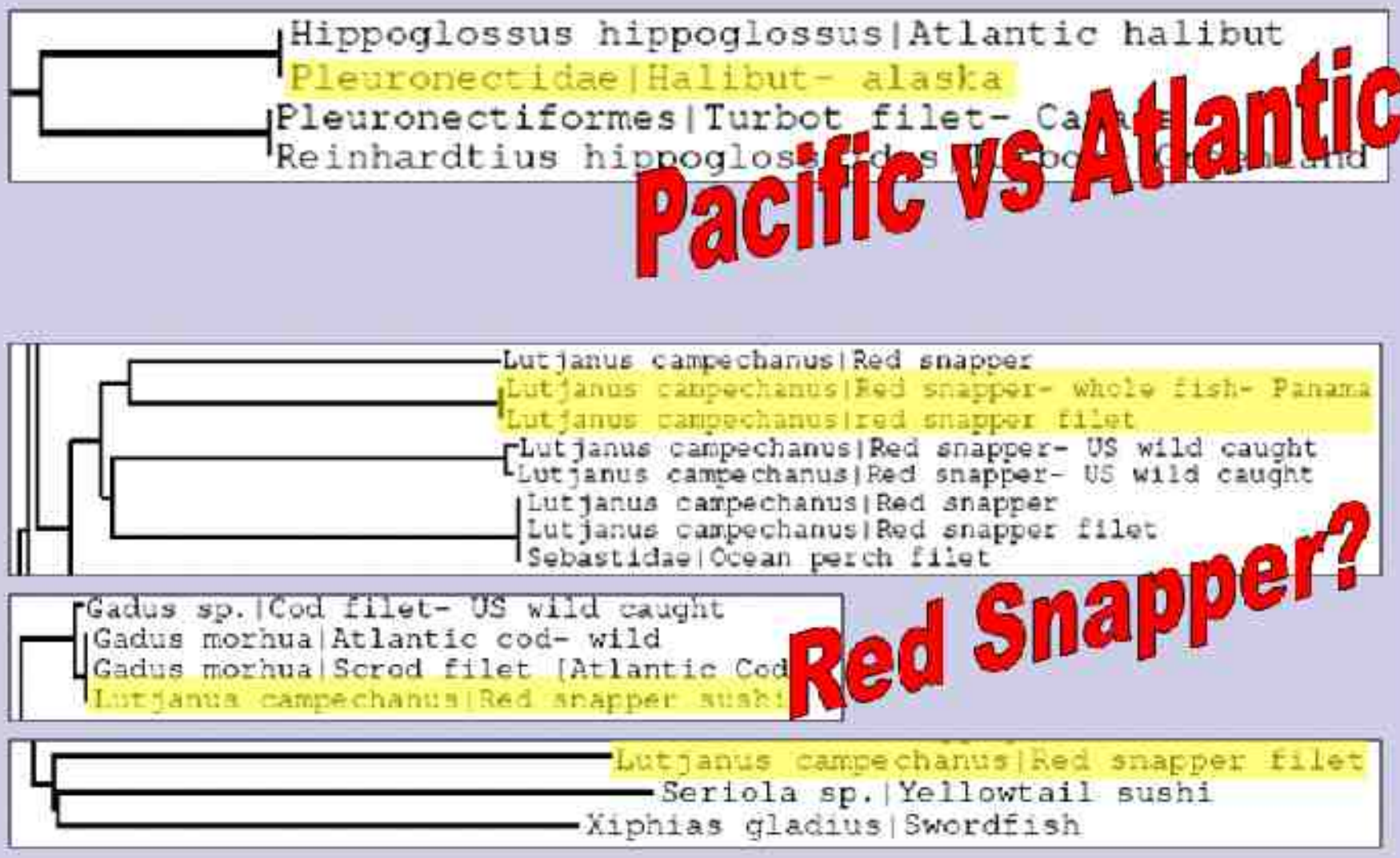
.: What was right?

Basa = Dory = Sutchi catfish



Patagonian toothfish = Chilean sea bass

.: What was not?


Hippoglossus hippoglossus|Atlantic halibut
Pleuronectidae|Halibut- alaska
Pleuronectiformes|Turbot filet- Ca
Reinhardtius hippoglos
Pacific vs Atlantic
Lutjanus campechanus|Red snapper
Lutjanus campechanus|Red snapper- whole fish- Panama
Lutjanus campechanus|red snapper filet
Lutjanus campechanus|Red snapper- US wild caught
Lutjanus campechanus|Red snapper- US wild caught
Lutjanus campechanus|Red snapper
Lutjanus campechanus|Red snapper filet
Sebastidae|Ocean perch filet
Gadus sp.|Cod filet- US wild caught
Gadus morhua|Atlantic cod- wild
Gadus morhua|Scrod filet [Atlantic Cod
Lutjanus campechanus|Red snapper sushi
Red Snapper?
Lutjanus campechanus|Red snapper filet
Seriola sp.|Yellowtail sushi
Xiphias gladius|Swordfish

canada.com
WHERE PERSPECTIVES CONNECT

Tuesday » June 9 » 2009

There's something fishy about this

Investigation finds more than one-quarter of 21 retail fish samples in Metro Vancouver were mislabelled

Larry Pynn
Vancouver Sun

Thursday, June 04, 2009

More than one-quarter of 21 retail fish samples analysed for DNA in Metro Vancouver were mislabelled, an investigation by The Vancouver Sun and the University of Guelph has revealed.

Two cases involved the use of cheaper products as cod at fish-and-chip outlets: Coney Island Seafood on Marine Drive in White Rock sold Southeast Asian catfish, a species known to be raised in aquaculture; and Speeds Pub in Ladner sold pollock.

Tests at sushi restaurants produced similar results: the high-end Takumi Japanese Restaurant in West Vancouver sold bastard halibut (a flounder) as halibut sushi, while Bon Sushi in Surrey near White Rock sold hake as crabmeat.

CREDIT: Ian Smith, Vancouver Sun

Neil Parker of Surrey puts salt on his grandson's fish and chips in White Rock. The cod in the advertised cod and chips at Coney Island Seafood turned out to be catfish, a recent DNA test found.

Taxonomy

Name, rank and serial number

Biologists want to barcode half a million species in the next five years

THE tale of the unknown goby began in 1982 when Benjamin Victor, of the Ocean Science Foundation in Irvine, California, discovered an unusual fish in a reef in Panama. With only a single specimen he was hard pressed to prove it was a new species, so the fish remained, unnamed, on his desk for 25 years. Then, last year, he was sent an unusual fish larva. Using a new kind of DNA identification called barcoding he showed that it was a younger version of his mystery goby and that both specimens were, indeed, a new species.

DNA barcoding was invented by Paul Hebert of the University of Guelph, in Ontario, Canada, in 2003. His idea was to generate a unique identification tag for each species based on a short stretch of DNA. Separating species would then be a simple

which there are at least 3,500 species, many of them hard to tell apart.

So far Dr Linton's team has used the COI gene to distinguish 390 species of mosquito, of which 7% have turned out to be new species. *Anopheles oswaldoi*, for example, was known to be a carrier of malaria in northern, but not southern, Brazil. That was puzzling. DNA barcoding, however, has shown that *A. oswaldoi* is actually four species, of which only one carries malaria. That explains the geographical discrepancy and should also assist efforts to curb the disease in Brazil by allowing the real culprit to be studied in detail.

Fly titles

The mosquito initiative has also had a piece of luck. Using some chemical wiz-

as medicines. In doing so, they have had to identify a new kind of barcode, as the COI gene is not found in plants.

Another group that could benefit from barcoding are customs officers, says Mark Blaxter, an evolutionary biologist at the University of Edinburgh. For those struggling to prevent the importation of pests or endangered wildlife, rapid and accurate identification tools are essential—particularly when perishable goods are being held up. America's Department of Agriculture is creating barcodes for the world's fruit flies. These are important agricultural pests and often arrive in the country as hard-to-identify larvae, or eggs, on fruit. Another group at the National Chung Hsing University in Taiwan (where hundreds of newly minted experts in the field have just met for the Second International Barcode of Life Conference) have created a prototype barcoding biochip. This is a collection of miniature DNA test sites on a sliver of glass that will rapidly discriminate between four species of fruit flies.

Barcoding's ease of use is also attracting interest from other government agencies. America's Federal Aviation Administration and its air force are working on bird

SCIENCE JOURNAL DECEMBER 4, 2009

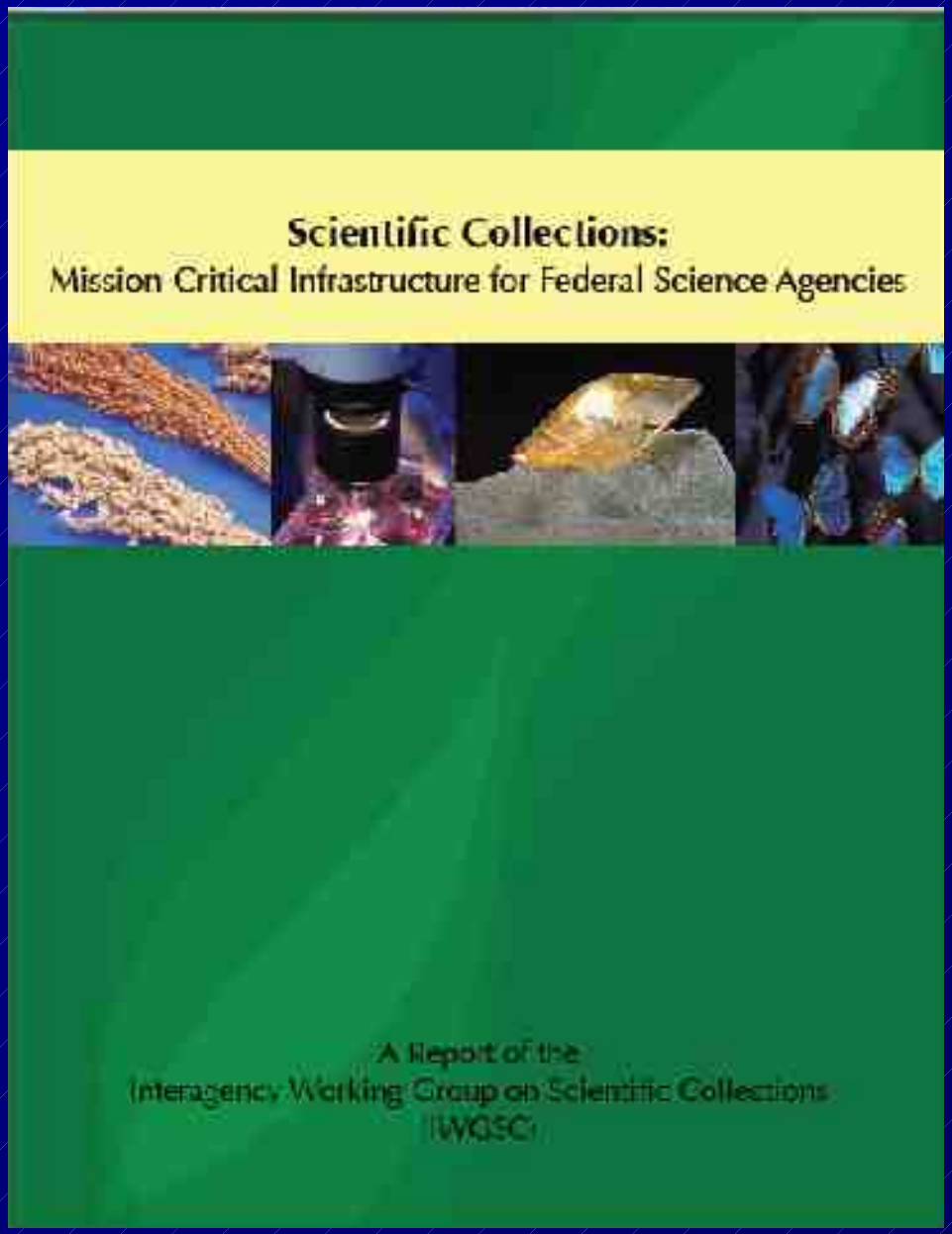
DNA 'Barcodes' Surface Fishy Imposters on Menus

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

By ROBERT LEE HOTZ

Bloomberg News

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.



http://www.ostp.gov/galleries/NSTC%20Reports/Revision_1-22_09_CL.pdf

OECD Global Science Forum, April 2009, Paris

Science Collections International

An international coordinating mechanism

Richard Lane,
Natural History Museum, London

Background

- **OSTP has defined scientific collections as critical scientific infrastructure for the Nation**
- IWGSC Federal survey of object-based collections
- NSF survey of non-Federal collections
- OECD initiative to create Scientific Collections International (SciColl) for coordination activities

Research Impacts of Collections

- Validation/re-testing of original findings
- Data extraction using new analytical instrumentation
  - DNA extraction from fossil bones, sediment
- ‘Off-label’ use by other research disciplines
  - Pathogen DNA from field biologists
  - Isotopic analysis of shells and bones

Non-Research Impacts of Scientific Collections

- Reference material for regulation and enforcement
- Calibration of scientific instruments
- Museum exhibits and education programming
- Headline news
- Science reporting
- “Night at the Museum” feature films

The Planetary Genome Project

Building the Synoptic Collection of Earth's Genetic Resources

One of three "Big Ideas" selected for long-term development

Smithsonian
National Museum of Natural History

Five Year Goals

- Global Consortium
  - Collaborations with ISBER and FIBO?
- “Life on Ice”
- Bioinformatics Program
- Exhibits, Web, Fellowships and Traineeships

Outreach and Education

Training the next generation
Educating the Public

Fellowships, traineeships

Temporary/Traveling exhibition
"Biodiversity Rescue Mission"
for tweens and teens

Web – Discoveries, blogs, field, fellowships, public participation in fieldwork