DNA barcoding in fishes: FISH-BOL

GENIMPACT, Tenerife, October 2006

Gary Carvalho
School of Biological Sciences, University of Wales Bangor,
Bangor UK

Molecular Ecology and Fisheries Genetics Lab
University of Wales Bangor
OUR PLANET:  
Home to 10-100 million species

OUR MINDS:  
Able to recognize and recall perhaps 1,000 species
SPECIES IDENTIFICATION IS FAR FROM SIMPLE

NO MASTER KEY

HIGHLY DEPENDANT ON GROUP-SPECIFIC EXPERTS

DIFFICULT ACCESS FOR NON-EXPERTS

“TAXONOMIC IMPEDIMENT”
### Tuna Identification

<table>
<thead>
<tr>
<th>Boat</th>
<th>Bigeye</th>
<th>Southern bluefin</th>
<th>Yellowfin</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>on-board</td>
<td>barcode</td>
<td>on-board</td>
</tr>
<tr>
<td>A</td>
<td>41</td>
<td>30</td>
<td>2</td>
</tr>
<tr>
<td>B</td>
<td>12</td>
<td>12</td>
<td>1</td>
</tr>
<tr>
<td>C</td>
<td>36</td>
<td>36</td>
<td>5</td>
</tr>
<tr>
<td>D</td>
<td>71</td>
<td>57</td>
<td>10</td>
</tr>
<tr>
<td>E</td>
<td>43</td>
<td>43</td>
<td>5</td>
</tr>
<tr>
<td>Total</td>
<td>203</td>
<td>178</td>
<td>23</td>
</tr>
</tbody>
</table>
Conventional Genomics - All Genes, One Species

Molecular Ecology and Fisheries Genetics Lab
University of Wales Bangor
Gaining Barcode Closure for Animals

Horizontal Genomics - One Gene, All Species
Horizontal Genomics for Biodiversity

Identifying Life:
- conservation, management, bioprospecting

Discovering Life:
- new species, species ages, geographic patterns

Evolutionary Rules:
- rate variation, shifts in nucleotide usage, protein diversification
First International Barcode of Life Conference: Feb 5-8, 2005

DNA barcoding of marine fishes: outline

- What is DNA barcoding?
- Introduction to FISH-BOL - European Group
- Marine fish barcoding
  - Case study on Australian fish
  - Applications in fish and fisheries science
DNA barcoding of marine fishes: outline

- What is DNA barcoding?
- Introduction to FISH-BOL- European Group
- Marine fish barcoding
  - Case study on Australian fish
  - Applications in fish and fisheries science
What is DNA barcoding?

“A rigorously standardized sequence of a minimum length and quality from an agreed-upon gene, deposited in a major sequence database, and attached to a voucher specimen whose origins and current status are recorded”

Paul Hebert & Ryan Gregory (2005)
DNA Barcode:
short standardized sequence enabling species
discrimination in a large block of life

DOES NOT AIM TO
DESCRIBE
SPECIES BUT TO
DELIMIT SPECIES
BOUNDARIES
DNA barcoding - diagnostic features

1. **Standardized**-sequence and methodology (but **NOT** only one gene for ALL species)
2. **Short**-sequence (<900 bp)- high throughput
3. **Digital** system vs. analogue gradient
4. **Voucher** specimen
5. **Global** database- high comparability
6. Appropriate **evolutionary rates** for species delimitation
7. Does **NOT** aim to recover species trees or phylogenetic relationships
DNA Barcoding in Animals

An Internal ID System for All Animals

The Mitochondrial Genome

Typical Animal Cell

DNA

Mitochondrion

mtDNA
**DNA BARCODE**

**CYTOCHROME C OXIDASE I (COI) partial sequence (≈650 bp)**

Species-level resolution tested in insects, birds, fish and macroalgae

---

**RECOGNIZED LIMITATIONS**

- Unsufficient resolution for recent species
- Introgresive hybridization
- Not useful for some organisms (e.g. plants, cnidarians)
On the whole good species resolution in all taxa.

<table>
<thead>
<tr>
<th>Taxa</th>
<th>n species / sequences</th>
<th>WITHIN SPECIES</th>
<th>BETWEEN SPECIES</th>
<th>RATIO</th>
<th>MAX WITHIN/ MIN BETWEEN</th>
</tr>
</thead>
<tbody>
<tr>
<td>DECAPODS</td>
<td>54 / 143</td>
<td>0.46</td>
<td>16.9</td>
<td>36</td>
<td>2.4 / 4.9</td>
</tr>
<tr>
<td>FISH</td>
<td>47 / 178</td>
<td>0.21</td>
<td>10.53</td>
<td>50</td>
<td>1.1 / 2.0</td>
</tr>
<tr>
<td>AMPHIPODS</td>
<td>20 / 130</td>
<td>0.7</td>
<td>27.5</td>
<td>39</td>
<td>3.1 / 5.4</td>
</tr>
</tbody>
</table>
DNA Barcoding in Animals

Barcode Target: 648 bp of COI
Barcode of Life Database

Barcode of Life

DNA BARCODE: A short DNA sequence enabling discrimination of species within a given compartment of life

NEW MILESTONES

- Canadian Barcode of Life Network (BOL.ca) Moves to Activate: BOL.ca has gained support and will soon launch efforts to assemble a DNA barcode library for all Canadian eukaryotes. Involving more than 50 researchers, BOL.ca represents the world's first national barcode network (details).

BARCODING ANIMALS

- Quick Links for Active Research on Barcoding Animals: Follow these direct links to some of the research projects and tools accessible through the Barcode of Life Database (BOLD).
  - BOLD Login - Access to active projects for registered BOLD researchers
  - BOLD Guest Pass - Access to completed projects to explore the functionality of the BOLD workflow
  - 1D by Barcode - Obtain an identification from a submitted DNA barcode

MAJOR ANNOUNCEMENTS

- Barcode of Life Database reaches 30,000 sequences
  The Barcode of Life Database (BOLD) has now holds over 30,000 barcode sequences, representing more than 12,000 individual species.

LATEST RESEARCH

- DNA Barcoding of North American Birds
  COT barcodes were published for 250 species of North American birds by Hebert et al. 2004. All currently recognized species were clearly delineated, and four probable new species were identified. These findings suggest that the proposed global barcoding survey of birds will contribute to the discovery of new bird species. A discussion of this work may be found here. The barcode data for this study is
DNA barcoding of marine fishes: outline

- What is DNA barcoding?
- Introduction to FISH-BOL- European Group
- Marine fish barcoding
  - Case study on Australian fish
  - Applications in fish and fisheries science
CONSORTIUM FOR THE BARCODE OF LIFE

- CBOL launched April, 2004.
- Active memberships in 25 countries (and growing).
- Projects to barcode all birds and all fishes on Earth.
- Estimated cost to barcode all animal life: $1 billion.
CBOL-Initiated 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 vertebrates (bushmeat)**
  - Successive phases of African mammals
Projects initiated by others

• CMarZ: Marine habitat, multiple taxa

• All-Leps:
  – 2 families, multiple regions
  – 4 regions/habitats, multiple taxa

• BioCode, Moorea: Single location, multiple habitats, multiple taxa
Consortium for the Barcode of Life (CBOL)
80 organizations from 33 countries

Fish Barcode of Life (FishBoL) recently launched worldwide initiative aiming to implement molecular ID for all fish species
FISH-BOL-European Group

Chair: Gary Carvalho, University of Wales, Bangor

Deputy-Chair Filipe Costa, University of Wales, Bangor
current MEMBERS

- Museums (e.g. Crete, Lisbon, London, Paris)

- Governmental agencies (e.g. Fisheries, Food)

- Academia and research institutes

- SMEs (small-medium enterprises).

Fair geographic coverage for FAO 27 and 37
FUNDING SO FAR:

- In the past, EU-funded projects: FISHTRACE (cyt b) and FISH & CHIPS (COI)

Currently:
- Dispersed, small budget, individual institutional funding in various countries

- Marie Curie Fellowship to FC (only until December 2007)

- Guelph Centre – molecular work on fish from Portugal

- Some funds for DNA barcoding of fish from Antarctic (project with liaison with Peter Smith and Filip Volckaert)

- Attempt for DNA barcoding proposals within ‘Marine Genomics Europe (MGE)’. No progress so far
FUNDING (2)

FUNDING NEEDS:

- Dedicated funding, particularly European-level

- EUROFISHBOL WORKSHOP

POTENTIAL FUNDING SOURCES TO CONSIDER:

Public funding - individual institutional funding, national or regional programs, or European-level programs or projects.

Other sources such as private Foundations, multinational companies, biotechnology companies, etc.?
FUNDING (3)

PROSPECTS:

-in the near future: coordination of individual efforts and available dispersed funds.

-long-term: 7th EU framework project (2007-2013)

-other EU funds, European Commission Agencies

NOTE:

- Dedicated funds hard to obtain
- Targeting various applications of DNA barcoding.
EUROFISHCODE (EU Marie-Curie Fellowship- Dr Filipe Costa)

DNA barcodes for species identification of fish and shellfish in Europe: implementation and application in selected case studies

hosted by: Gary Carvalho (UWB)

Linked with international initiatives:

Consortium for the Barcode of Life (CBOL)
organizations from 33 countries

• Fish Barcode of Life (FishBoL) – recently launched worldwide initiative aiming to implement molecular ID for all fish species

Deposit voucher specimens
Molecular Ecology and Fisheries Genetics Lab
University of Wales Bangor
DNA barcoding of marine fishes: outline

• What is DNA barcoding?
• Introduction to FISH-BOL- European Group
• Marine fish barcoding
  – Case study on Australian fish
  – Applications in fish and fisheries science
Why barcode fish

- Global value at first sale:
  - Capture fisheries in 2000 = c. US$ 81 billion
  - Aquaculture (excl. plants) in 2000 = c. US$ 52 billion

- Ornamental fish:
  - Retail trade in US alone in 2000 = c. US$ 3 billion

- Sports fishing:
  - In Australia in 1984, recreational fishers = c. US$ 2 billion.

Fisheries activities globally generate trade > US$ 200 billion and employ 35 million people (65% marine capture, 15% inland capture, 20% aquaculture).

Phil. Trans. Roy. Soc. Lond. 360, 1847-1857
Twelve species of flathead fish (family Platycephalidae)

Mean genetic distance:
- within species 0.22%
- between species 15.55%

Platycephalus longispinis
725 sequences/individuals,
165 described species (each with two or more individuals)
44 genera (teleosts and elasmobranchs)

Mean % divergence
- Spp. = 0.39%
- Genus = 9.93%
- Family = 15.4%
- Order = 22.18%
- Class = 23.27%

Pairwise distance among members of different taxonomic levels

Kimura 2 parameter
COI neighbour-joining tree of 206 fish species and 754 sequences
Marine fishes- *so far*

- **CSIRO-** Australian fishes- 1,200 specimens = 271 spp.
- **Guelph-** 530 spp. across 3 Oceans (c. 3.5% of global marine species diversity)
- **Portugal-** 47 spp., 35 genera

*Approx. 2% divergence threshold proxy for species delimitation*

**ALL USING ONLY TWO SETS OF PRIMERS**
DNA barcoding of marine fishes: outline

• What is DNA barcoding?
• Introduction to FISH-BOL- European Group
• Marine fish barcoding
  – Case study on Australian fish
  – Applications in fish and fisheries science
Applications of molecular ID (1)

• Identification of fish, fillets, fins, and fragments (marketing, substitutions, quota and bycatch management).

• Identification of processed product e.g. canned fish, dried fish, mixtures (substitutions).
Barcoding Australian Fish

Fish fillet identification
Depleted stocks, strict management measures, incentive for substitution

Marko et al. (2004) Investigating the extent of mislabelling using cyt b sequences
FIGURE 1. Maximum-likelihood tree of cytochrome-\(b\) DNA sequences from retail 'red snapper'.

- 77% mislabelling
- Possible inflation of catch estimates
- Distort the status of fish stocks as perceived by consumers
- Possible rare species, without molecular data yet

Applications of molecular ID (1)

• Identification of threatened, endangered and protected species (*conservation*).
• Identification of fish eggs and fish larvae (*ecosystem research, direct and indirect fisheries management*).
Distribution of eggs in North Sea, based on TaqMan methodology (eggs per m² sea surface) (After Fox et al., Unpublished)

- c. 34% of “cod-like” eggs identified as cod (8% haddock + 58% whiting);
- early stage haddock eggs -spawning stock in Irish Sea

(Fox et al., 2005 Irish Sea)
Applications of molecular ID (2)

- Identification of prey items in stomach contents (*food webs and ecosystem research*).
- Identification of historical, archived and museum material (*taxonomy*).
- Identification of new species and possible fusions, insights into phylogenetic relationships (*fish biology, evolution*).
- Possible production of DNA microarrays from the sequence data
- Documentation of range expansions
Concluding remarks

• Value of standardised and digital approach
• Recognition of limitations- use of multiple target sequences in some situations
• Evolutionary interest- conserved within-species variation- selective sweeps?
• Conservation of biodiversity
• Makes Linnean taxonomic system more accessible
• New funding into taxonomic collections and cataloguing- *resurgence of taxonomy*
• Reconciling alternative technology platforms?
The Future of DNA Barcoding

Barcoding: A Field Guide for the Third Millenium
Thanks to:

• Paul Hebert - *Guelph*
• Bob Ward - *CSIRO, Tasmania*
• Bronwyn Innes, *CSIRO, Tasmania*
• Filipe Costa, *Univ. Nova Lisboa, Portugal*
• Clive Fox, *CEFAS, Lowestoft*
• Martin Taylor, *SBS, University of Wales, Bangor*
• Dr David Schindel, *Executive Secretary, CBOL*
Questions?