Census of Marine Zooplankton

CENSUS OF MARINE LIFE

DNA barcoding of North Atlantic zooplankton

Ann Bucklin
University of Connecticut, Groton, CT USA

Brian D. Ortman (Univ. British Columbia)
Robert M. Jennings (Univ. Massachusetts – Boston) and
Lisa M. Nigro (Univ. North Carolina)
All formerly at University of Connecticut, Groton, CT USA

and Nancy J. Copley
Woods Hole Oceanographic Institution, Woods Hole, MA USA

Zooplankton images by Russell R. Hopcroft *(Univ. of Alaska)*
and Laurence P. Madin *(Woods Hole Oceanographic Inst.)*

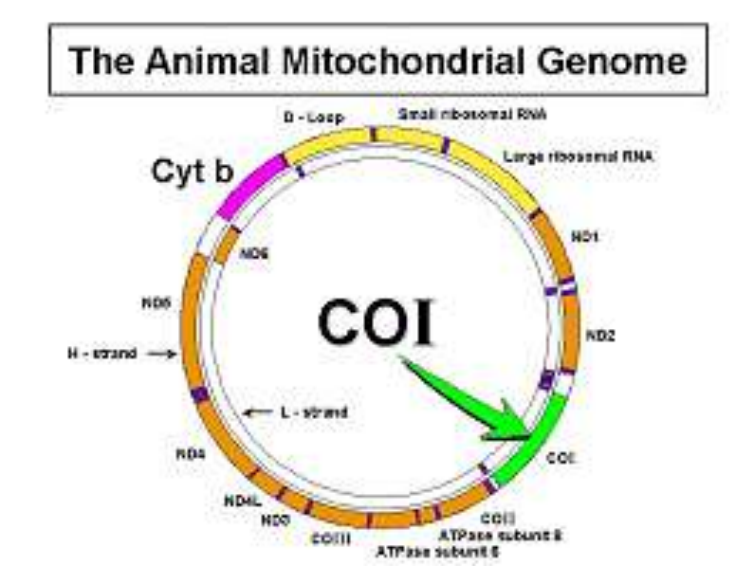
CMarZ Symposium and Steering Group Meeting
IOCAS, Qingdao, China – May 11-13, 2010

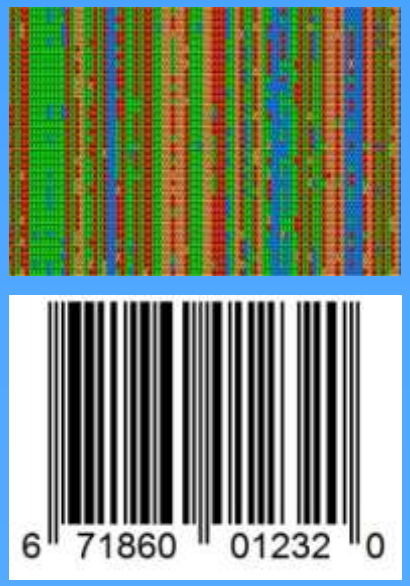
What is DNA Barcoding?

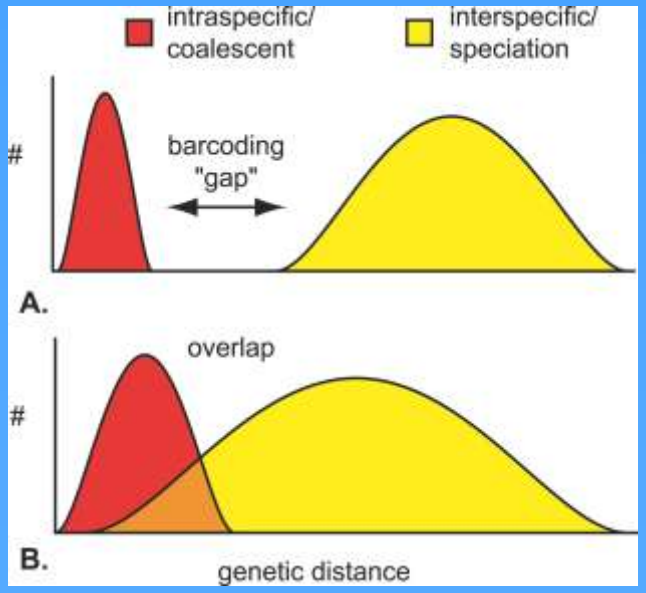
- **Definition: Derivation of short DNA sequence(s) that enables species identification or recognition in a particular domain of life (e.g., eucaryotes).**

- **Focus to date: For animals, a 658 base-pair fragment of the mitochondrial gene, cytochrome oxidase c subunit I (COI). Also called a "gold standard"barcode.**

- **Species identification using barcodes relies on "barcode gap" or non-overlapping distribution of distances within and between species).**

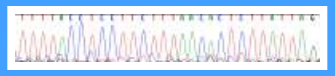






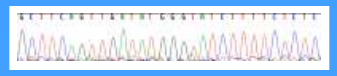
The Rosetta Stone

Hippopodius hippopus



Sapphirina metallina

Limacina helicina



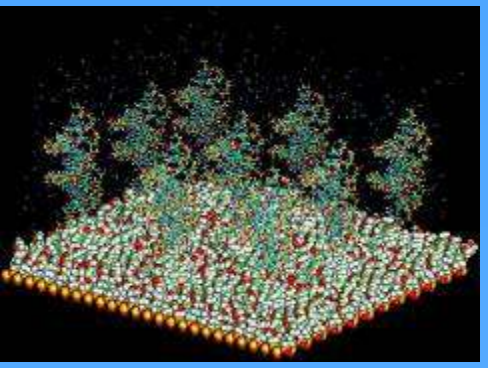
Salpa cylindrica

- **DNA barcode library will serve as Rosetta Stone for decoding zooplankton species diversity by linking species names, morphology, and DNA sequence variation.**
- **Taxonomically comprehensive and geograpically extensive (global) barcode database will allow identification of known species with barcodes using only the DNA sequence.**
- **Rapid analysis of known species diversity, distribution, and abundance may be done based only on DNA sequences.**

CENSUS OF MARINE LIFE

Why Barcode Zooplankton?

CMarZ

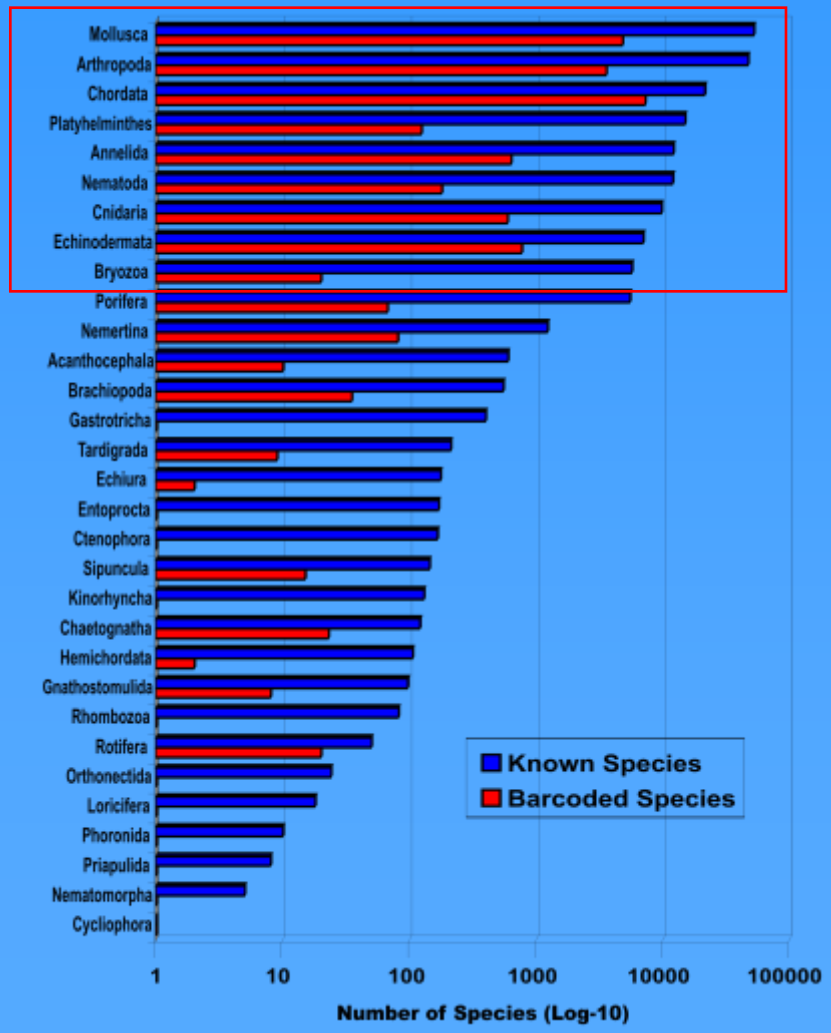


- **DNA barcodes aid in species identification, because organisms are frequently rare, fragile, and/or small.**
- **Morphological identification is difficult and mistakes are likely due to simple or evolutionarily-conserved body plans.**
- **Within-species variation: barcodes can describe population genetic and phylogeographic patterns, reveal taxonomically-significant geographic variation and cryptic species within taxa with circumglobal or disjunct geographic distributions.**
- **Between-species variation: DNA barcodes can reconstruct relationships among closely-related species and reveal processes associated with speciation.**

Barcoding Marine Life

MarBOL (www.marinebarcoding.org) is working to determine COI barcodes for all ~230,000 known marine metazoan species; currently barcodes have been determined for ~10% of species.

95%



Phylum	Known Species	Barcoded Species (#)	Barcoded Species (%)
Acanthocephala	600	10	1.7%
Annelida	12,148	635	5.2%
Arthropoda	47,217	3,580	7.6%
Brachiopoda	550	35	6.4%
Bryozoa	5,700	20	0.4%
Chaetognatha	121	23	19.0%
Chordata	21,517	7,279	33.8%
Cnidaria	9,795	594	6.1%
Ctenophora	166	0	0.0%
Cyclophora	1	1	100.0%
Echinodermata	7,000	771	11.0%
Echiura	176	2	1.1%
Entoprocta	170	0	0.0%
Gastrotricha	400	0	0.0%
Gnathostomulida	97	8	8.2%
Hemichordata	106	2	1.9%
Kinorhyncha	130	0	0.0%
Loricifera	18	0	0.0%
Mollusca	52,525	4,813	9.2%
Nematoda	12,000	180	1.5%
Nematomorpha	5	0	0.0%
Nemertina	1,230	81	6.6%
Orthonectida	24	0	0.0%
Phoronida	10	0	0.0%
Platyhelminthes	15,000	124	0.8%
Porifera	5,500	67	1.2%
Priapulida	8	1	12.5%
Rhombozoa	82	0	0.0%
Rotifera	50	20	40.0%
Sipuncula	144	15	10.4%
Tardigrada	212	9	4.2%
TOTALS	192,702	18,270	9.5%

Bucklin, Steinke, Blanco-Bercial (In review) Ann.Rev.Mar.Sci.

CENSUS OF MARINE LIFE

Species Diversity of Holozooplankton

	Phylum		Taxon	Species
1	Foraminifera	1	Foraminifera	49
2	Actinopoda	2	Acantharea	150
		3	Polycystinea (Radiolaria)	350
3	Cercozoa	4	Phaeodarea (Radiolaria)	350
4	Ciliophora	5	Aloricate Ciliata	150
		6	Tintinnida	300
5	Cnidaria	7	Hydromedusae	842
		8	Siphonophora	160
		9	Cubomedusae	18
		10	Scyphomedusae	161
6	Ctenophora	11	Ctenophora	90
7	Rotifera	12	Rotifera	50
8	Platyhelminthes	13	Platyhelminthes	3
9	Nematomorpha	14	Nectonema	5
10	Nemertea	15	Nemertinea	99
11	Annelida	16	Polychaeta	110
12	Mollusca	17	Gastropoda	144
		18	Cephalopoda	370
13	Arthropoda	19	Cladocera	8
		20	Ostracoda	169
		21	Isopoda	20
		22	Copepoda	2000
		23	Mysidacea	700
		24	Amphipoda	400
		25	Euphausiacea	86
		26	Decapoda	50
		27	Insecta	5
14	Chaetognatha	28	Chaetognatha	93
15	Chordata	29	Appendicularia	64
		30	Pyrosoma	8
		31	Doliolida	17
		32	Salpidae	45
	TOTALS			***7,066***

Barcodes for Cnidaria - Medusozoa

B.D. Ortman* et al. (In press) DSR-II

Francesc Pages, Brian Ortman, Dhugal Lyndsay

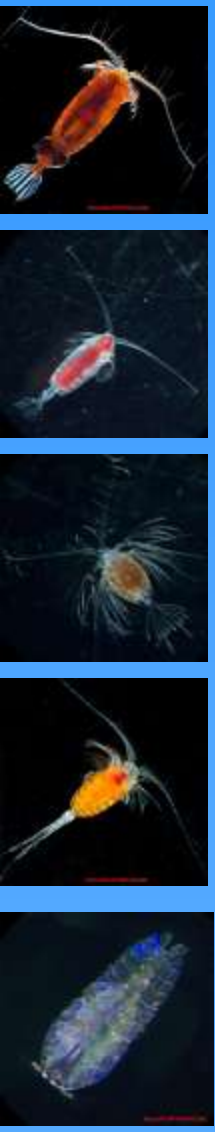
228 barcodes for 95 species, including:

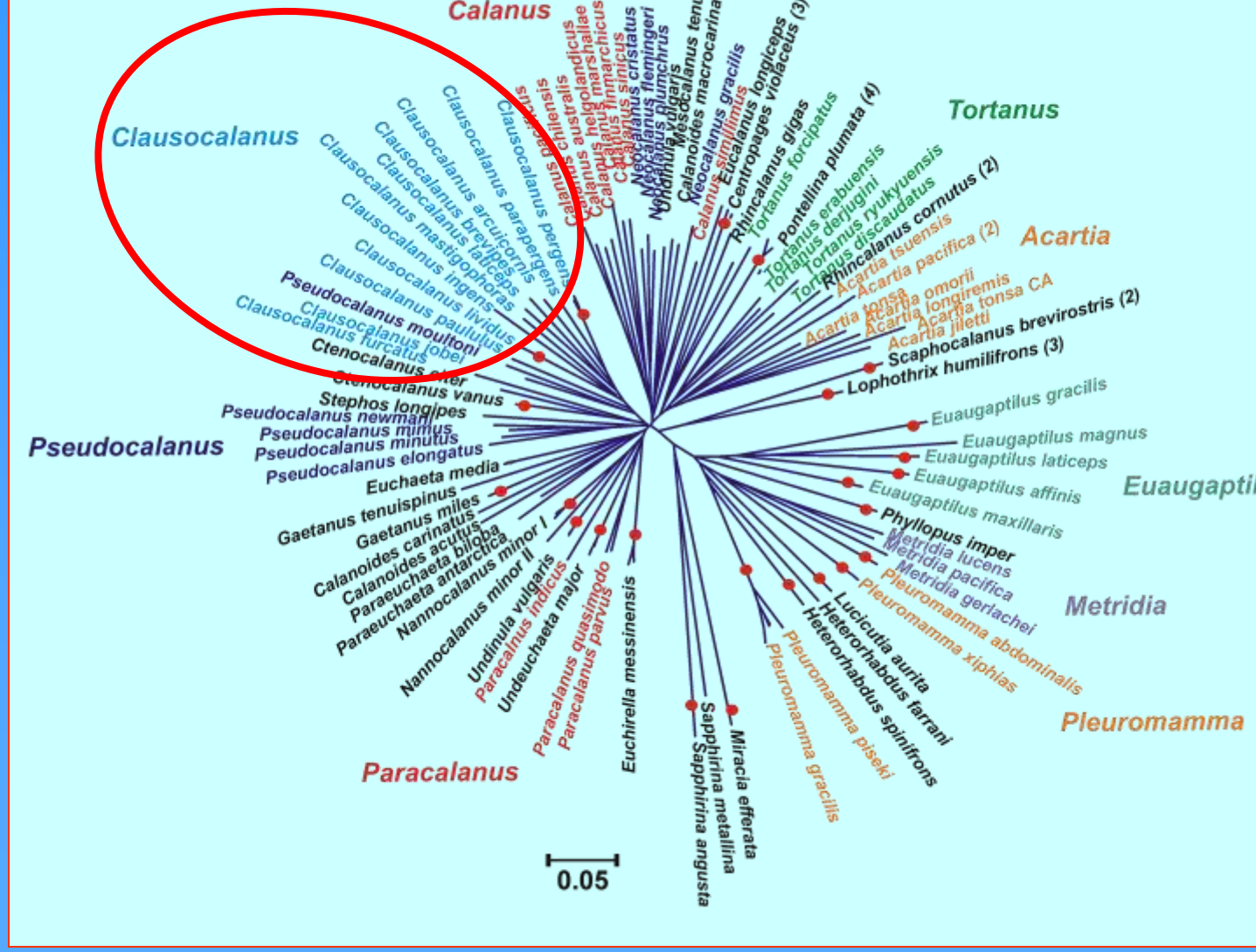
- **Siphonophores**
- **Hydromedusae**
- **Scyphomedusae**
- **Cubomedusa**

*Current address: Univ. British Colombia

Barcodes for Calanoid Copepods

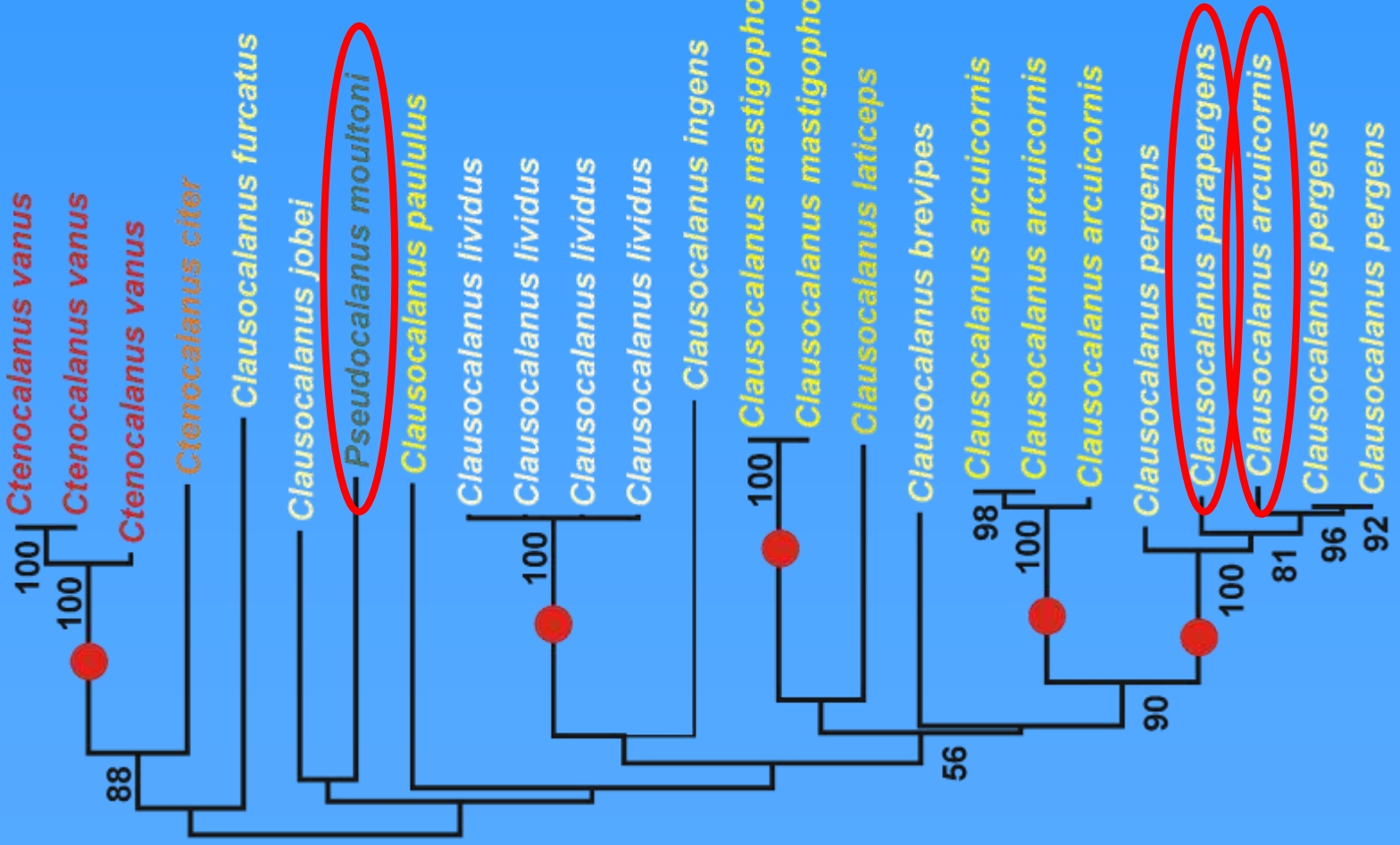
150 species in Neighbor Joining tree with Kimura-2-Parameter distances, bootstrapped 1000X





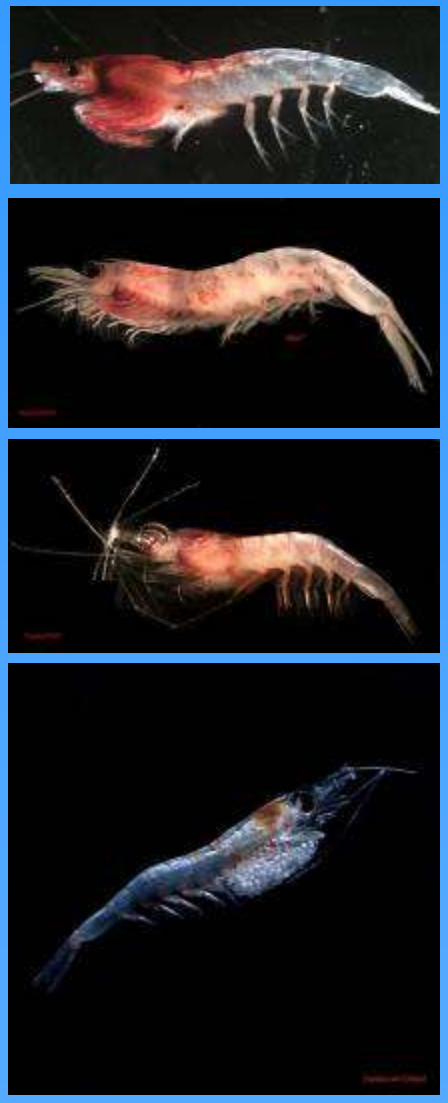
Barcodes for Copepods: Clausocalanus

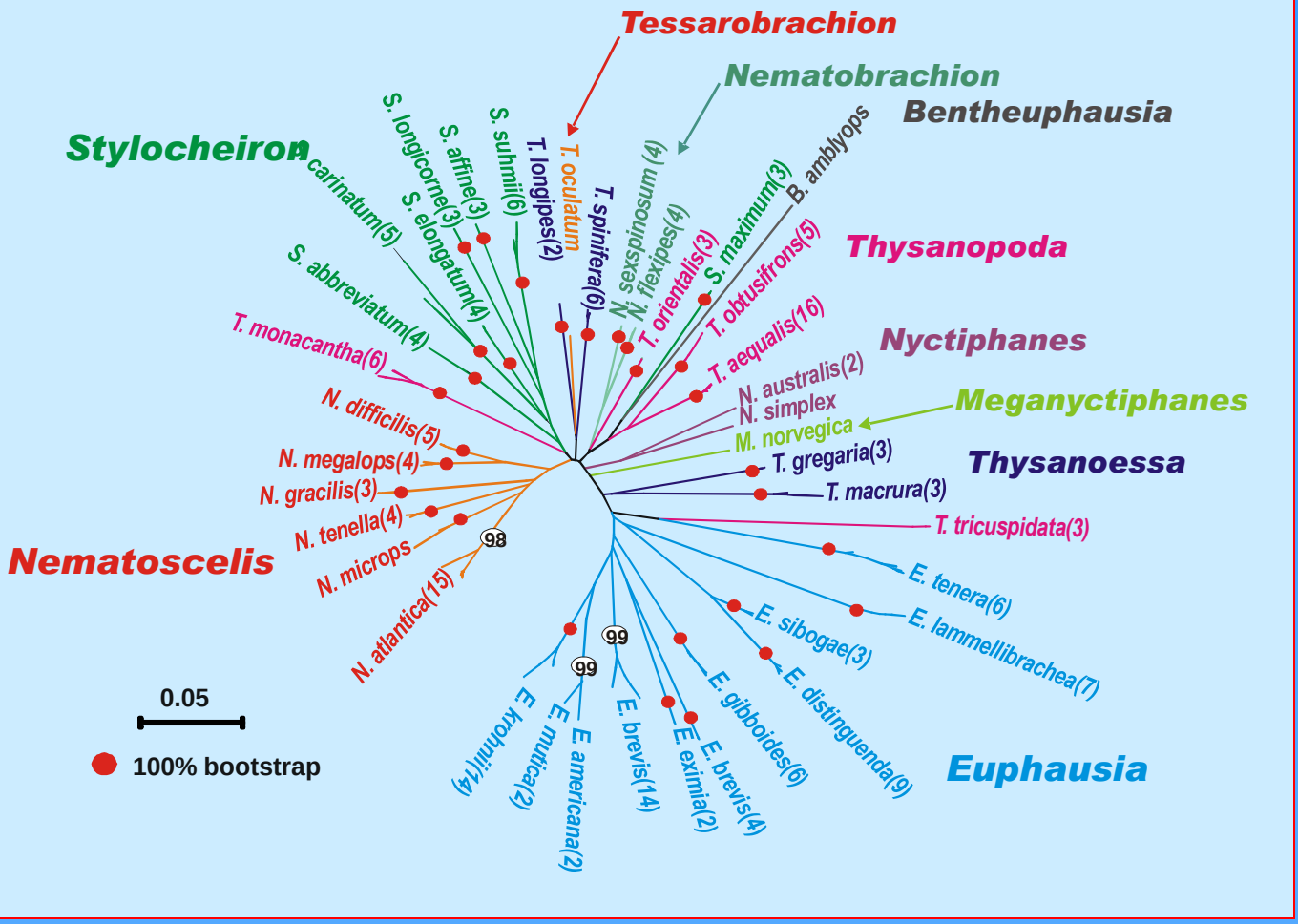
COI resolves species accurately and reliably, but does not reliably resolve evolutionary relationships among species.



Barcodes for Euphausiids

Bucklin, Wiebe, et al., 2007, J. Plankton Res.



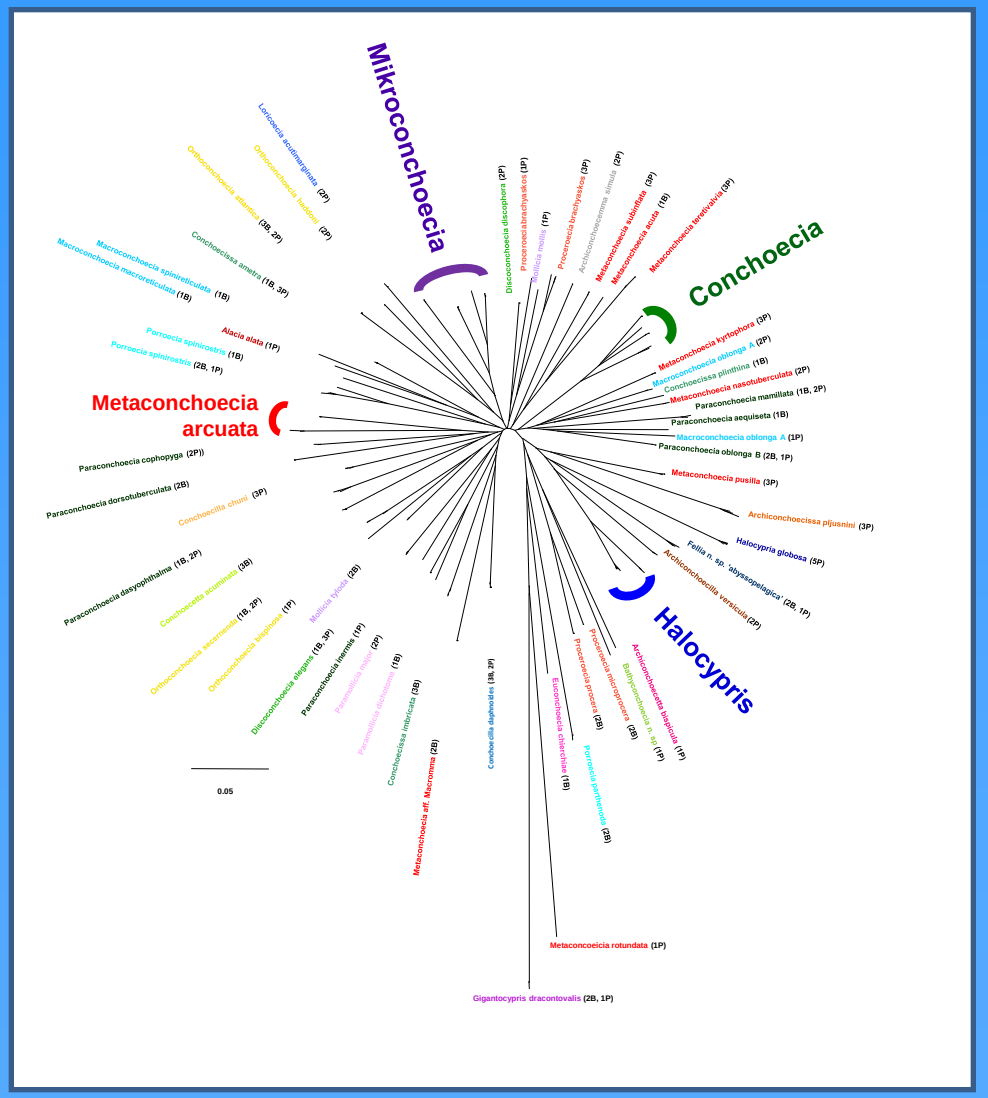


Barcodes for 193 individuals of 40 species in Neighbor Joining tree using Kimura-2-Parameter distances, bootstrapped 1000X

Barcodes for Ostracods

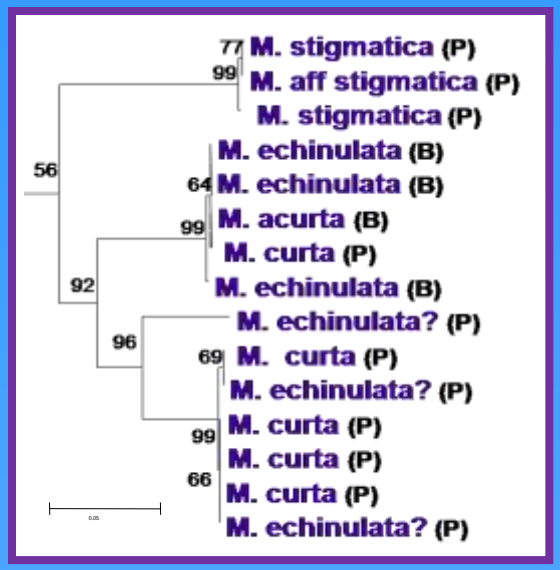
Lisa M. Nigro[1], Martin V. Angel[2] and Ann Bucklin[1]
[1]Department of Marine Sciences, University of Connecticut, USA
[2]National Oceanography Centre, Southampton, UK

Collections of >80 species from Sargasso Sea and Eastern Atlantic regions.
Neighbor Joining Tree with Kimura-2-Parameter (K2P) distances.

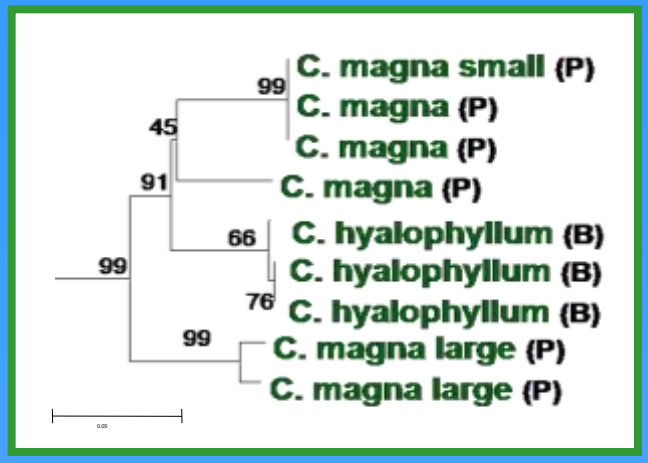


Barcodes for Ostracods

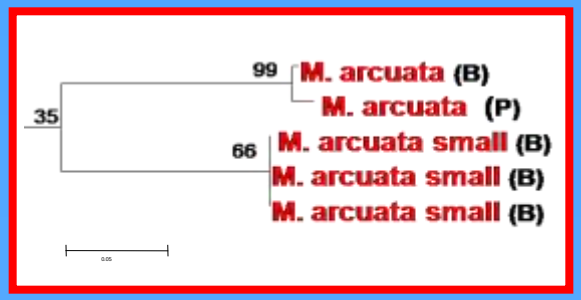
Lisa M. Nigro[1], Martin V. Angel[2] and Ann Bucklin[1]
[1]Department of Marine Sciences, University of Connecticut, USA
[2]National Oceanography Centre, Southampton, UK



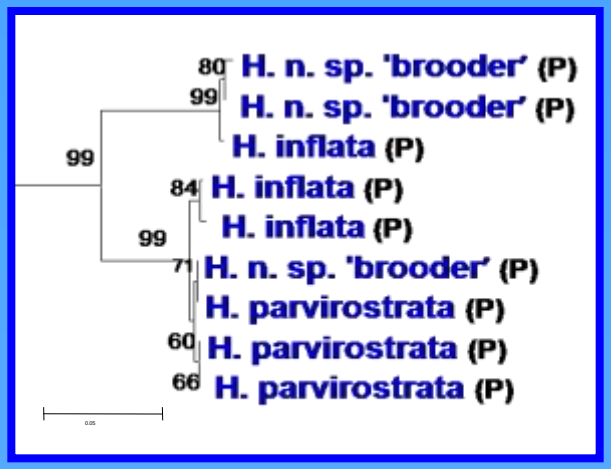
Mikroconchoecia



Conchoecia



Metaconcoecia arcuata

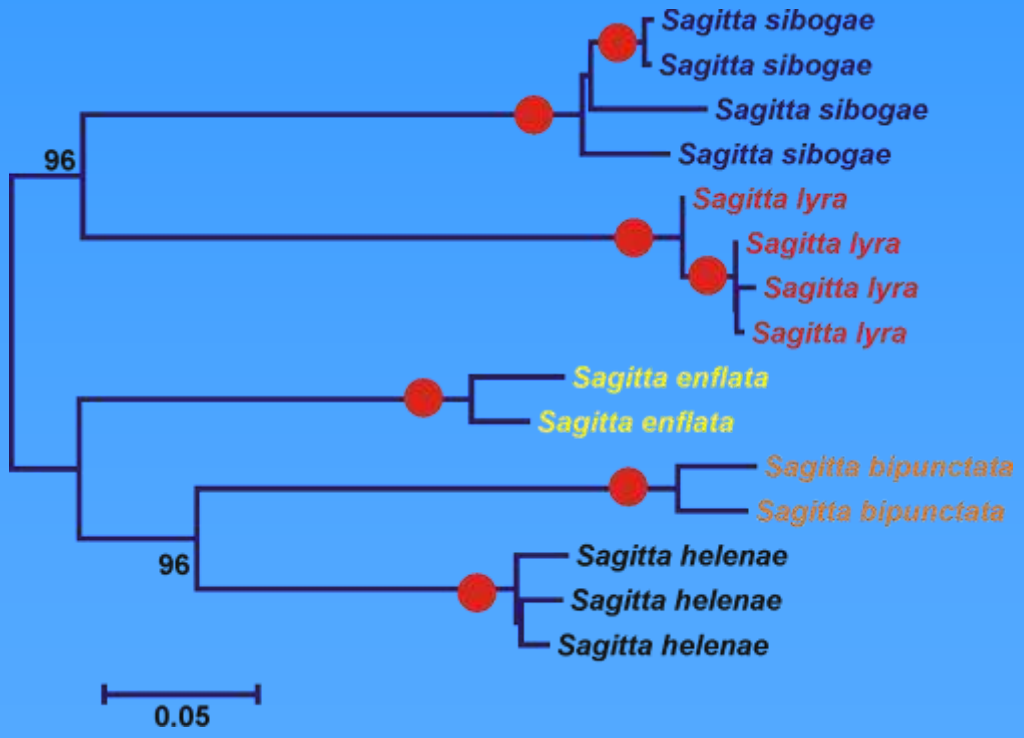


Halocypris

Barcodes for Chaetognaths:Sagitta

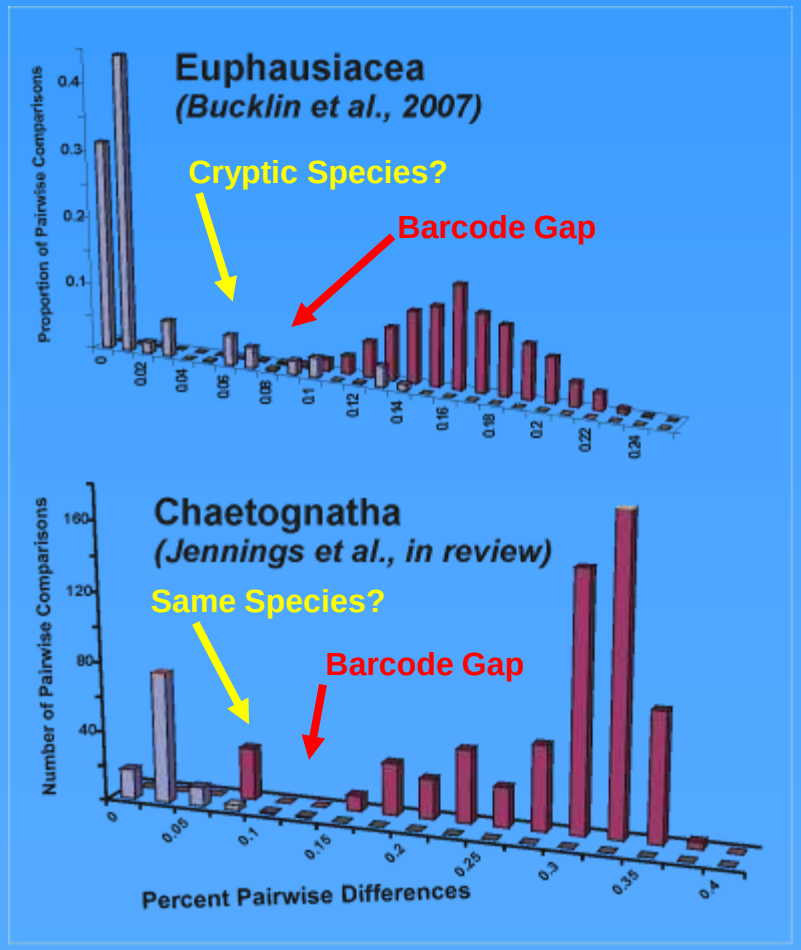
R.M. Jennings et al. (in press) DSR-II

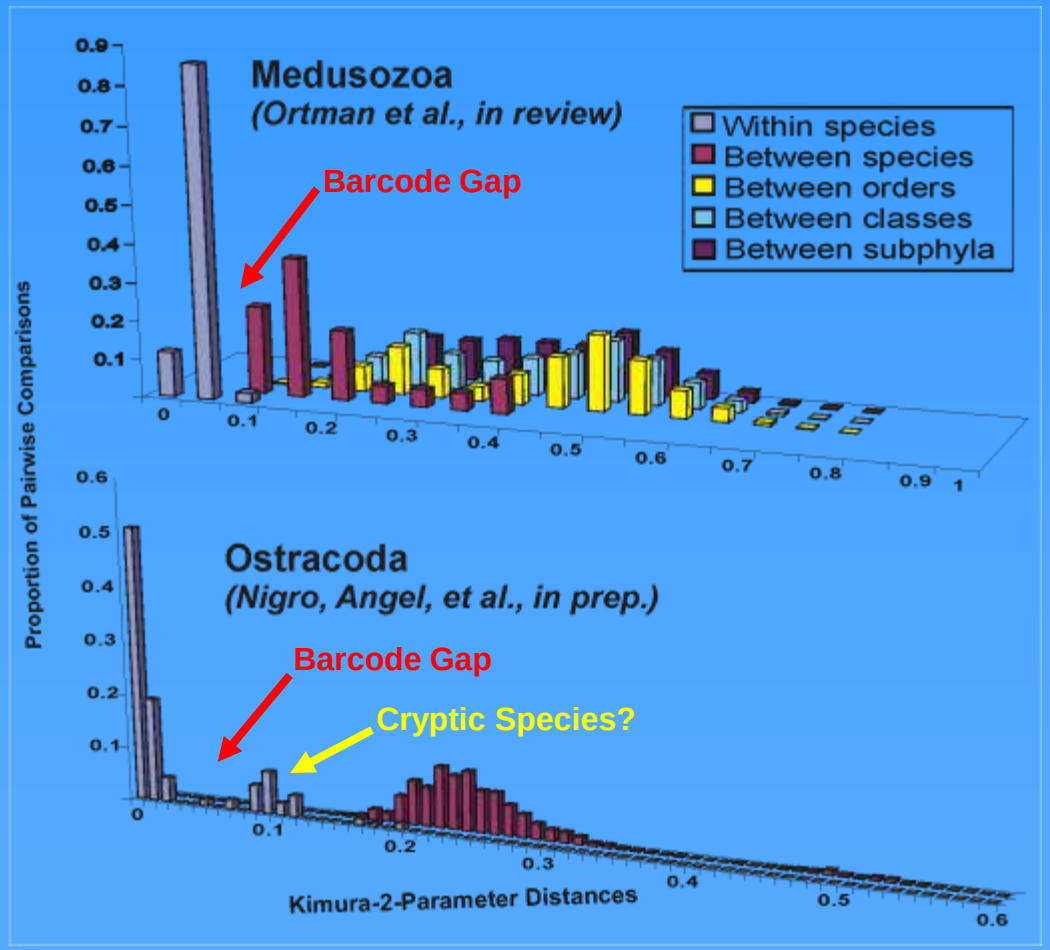
DNA barcodes reliably and accurately discriminated species of *Sagitta* based on Neighbor Joining K2P tree analysis.



DNA Barcode Gap for Zooplankton

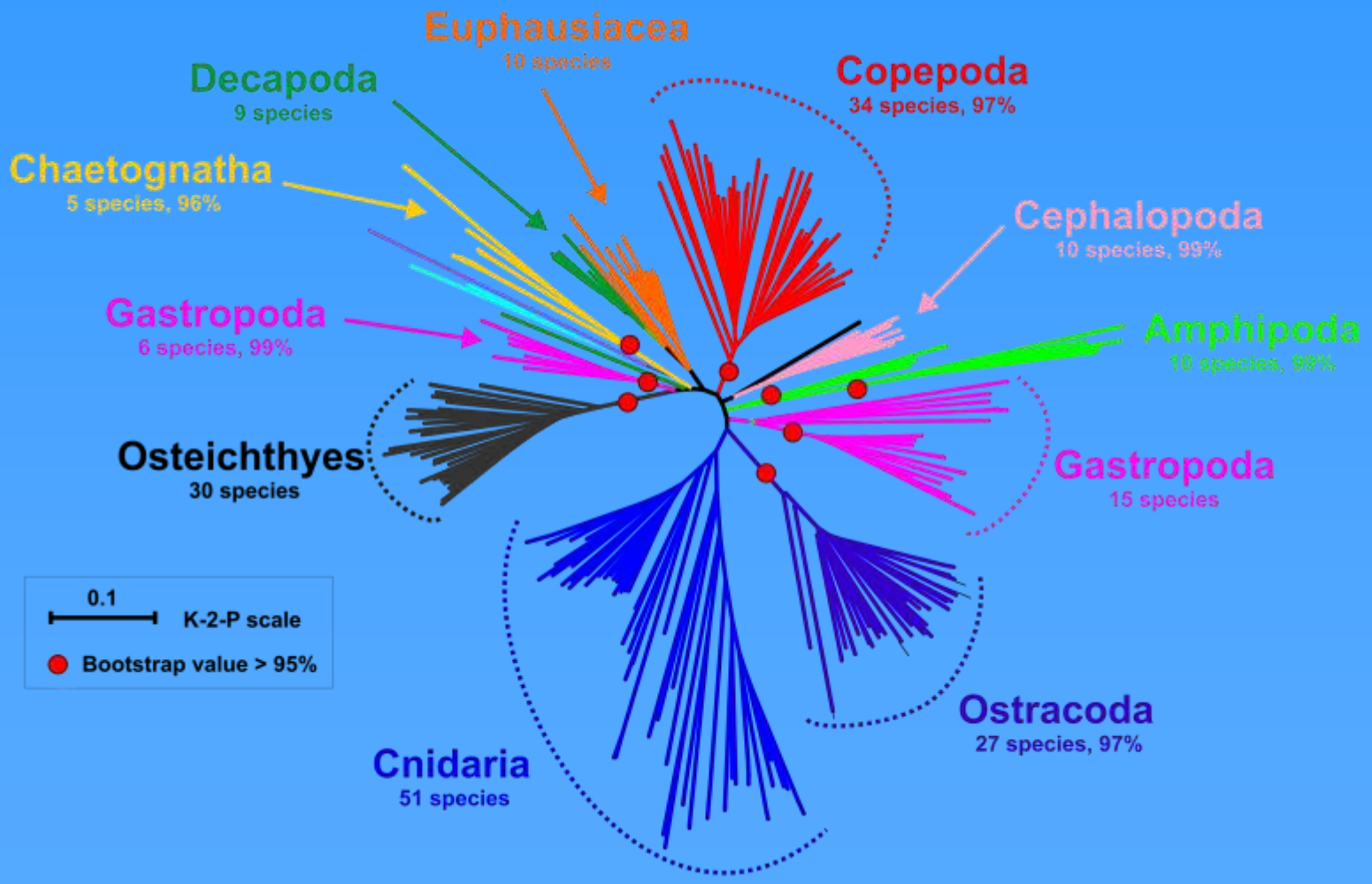
COI barcode sequence variation within and between species shows barcode gaps for all zooplankton groups analyzed.





Sargasso Sea Barcodes: NJ Tree

Distance-based analysis resolves branches between major zooplankton groups in Neighbor Joining (NJ) tree; 348 barcodes for 198 species; Kimura-2-Parameter distances, 1000X bootstrapping



Barcode Vector Analysis

Sirovich, Stoeckle, and Zhang (2009) PLoS-One


OPEN ACCESS Freely available online
PLoS one

A Scalable Method for Analysis and Display of DNA Sequences

Lawrence Sirovich[1]*, Mark Y. Stoeckle[2], Yu Zhang[1]

Abstract

Background: Comparative DNA sequence analysis provides insight into evolution and helps construct a natural classification reflecting the Tree of Life. The growing numbers of organisms represented in DNA databases challenge tree-building techniques and the vertical hierarchical classification may obscure relationships among some groups. Approaches that can incorporate sequence data from large numbers of taxa and enable visualization of affinities across groups are desirable.

Methodology/Principal Findings: Toward this end, we developed a procedure for extracting diagnostic patterns in the form of indicator vectors from DNA sequences of taxonomic groups. In the present instance the indicator vectors were derived from mitochondrial cytochrome c oxidase I (COI) sequences of those groups and further analyzed on this basis. In the first example, indicator vectors for birds, fish, and butterflies were constructed from a training set of COI sequences, then correlations with test sequences not used to construct the indicator vector were determined. In all cases, correlation with the indicator vector correctly assigned test sequences to their proper group. In the second example, this approach was explored at the species level within the bird grouping; this also gave correct assignment, suggesting the possibility of automated procedures for classification at various taxonomic levels. A false-color matrix of vector correlations displayed affinities among species consistent with higher-order taxonomy.

Conclusions/Significance: The indicator vectors preserved DNA character information and provided quantitative measures of correlations among taxonomic groups. This method is scalable to the largest datasets envisioned in this field, provides a visually-intuitive display that captures relational affinities derived from sequence data across a diversity of life forms, and is potentially a useful complement to current tree-building techniques for studying evolutionary processes based on DNA sequence data.

Introduction

1) Group and order FASTAs by taxonomic similarity.

2) Align sequences; select 500 bp COI domain starting 100 bp downstream of LCOI-1490.

3) Transform data

A → [1,0,0,0]
C → [0,1,0,0]
G → [0,0,1,0]
T → [0,0,0,1]

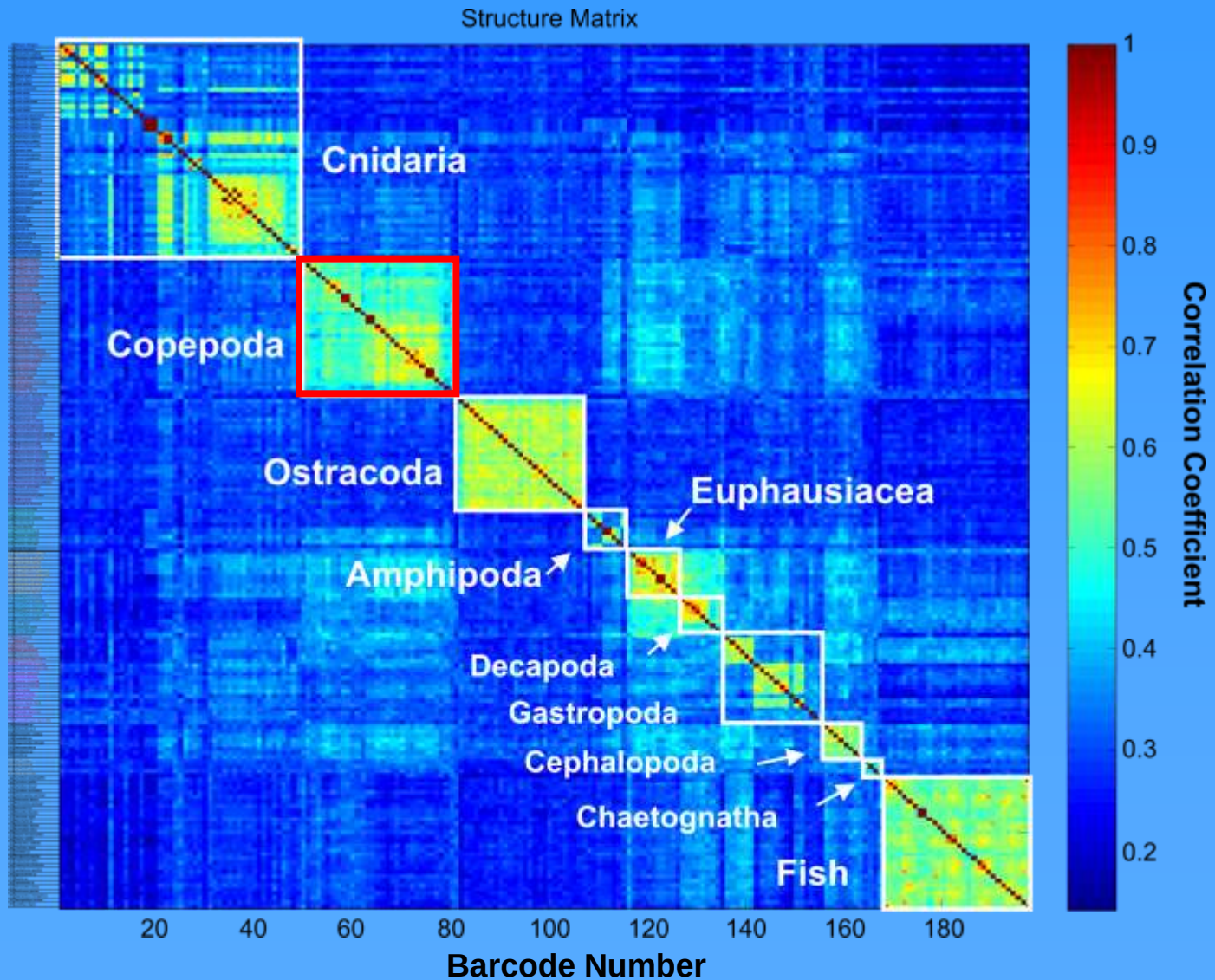
ATTC→ [1,0,0,0, 0,0,0,1, 0,0,0,1, 0,1,0,0]

4) Compute Hamming distance, dH = number of substitutions between 2 sequences; normalized due to standard domain for analysis.

5) Compute correlation coefficient for all sequence pairs.

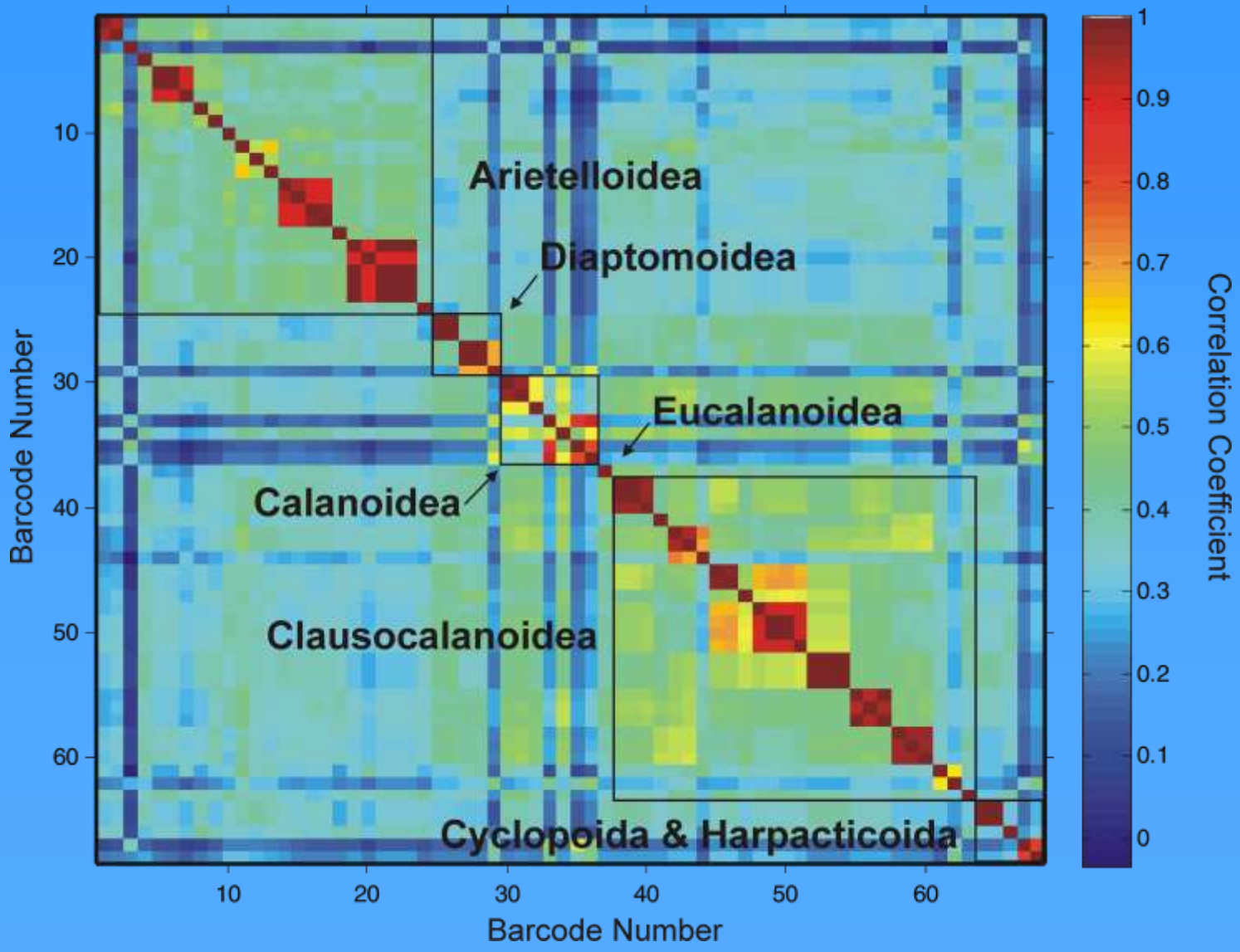
Sargasso Sea Barcodes: Vector Analysis

Vector analysis and heat map display (Klee diagram) clearly resolves major groups; 348 barcodes for 198 species; analysis method from Sirovich et al. (2009, 2010)

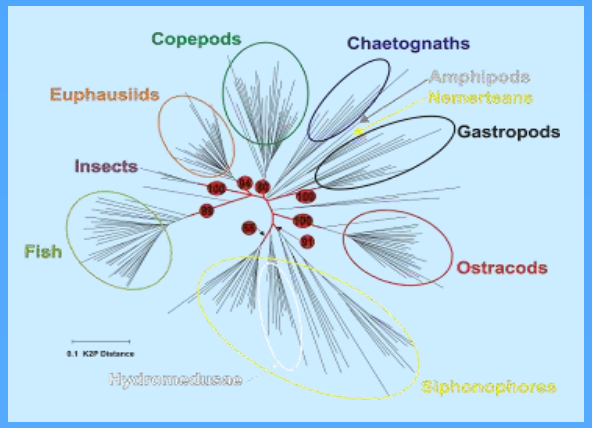


Sargasso Sea Copepods: Vector Analysis

Subset of Sargasso Sea data: 69 barcodes for 34 copepod species
Vector analysis is scalable and zoomable. (Sirovich et al., 2009)



Environmental Barcoding of Zooplankton

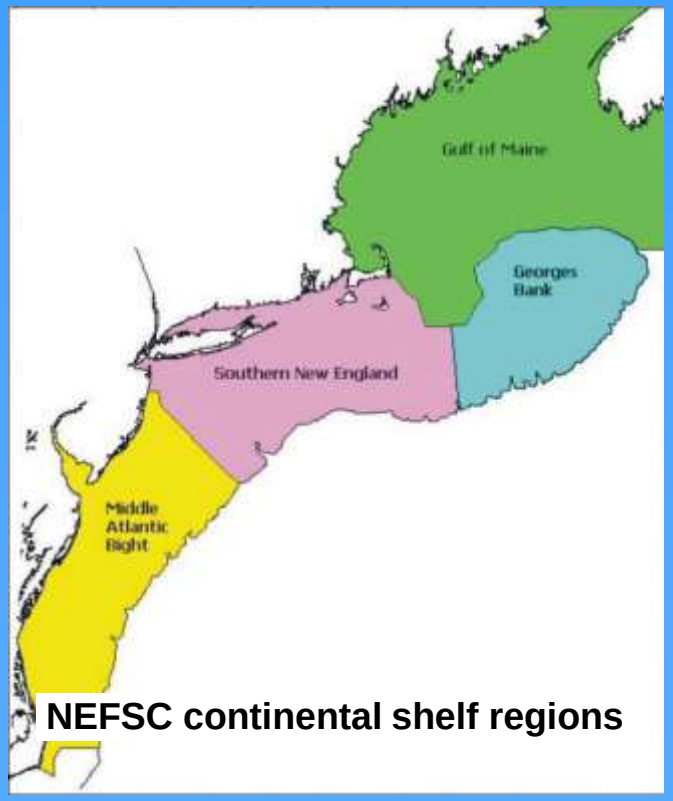


- Environmental barcoding is sequencing the barcode gene from bulk environmental samples and identifying species from a library of barcodes for known species.
- DNA or rRNA extracted from bulk samples; used for amplification of COI or construction of COI cDNA libraries.
- High throughput DNA sequencing used for exhaustive analysis of DNA or cDNA libraries.
- Database of DNA barcodes allows accurate identification of known species, estimation of new or undescribed species.

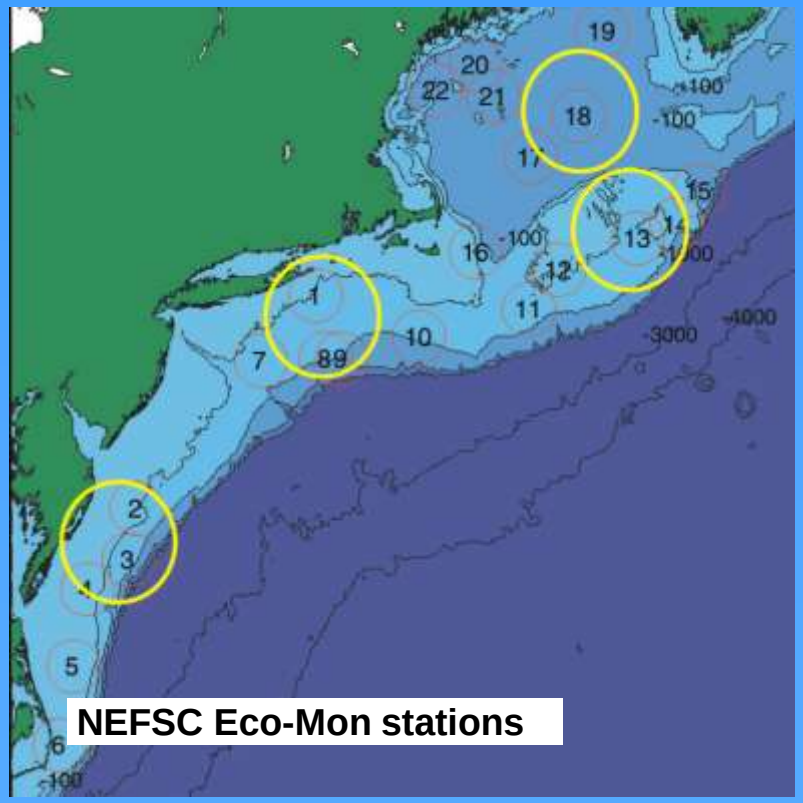
Ecosystem Monitoring with Barcodes

Northeast Fisheries Science Center Ecosystem Monitoring Program

Zooplankton samples from a fisheries Ecosystem Monitoring Program (EcoMon) are being used to create a DNA barcode database for 200 species collected from four regions of the Northwest Atlantic continental shelf ecosystem based on stratified random sampling.

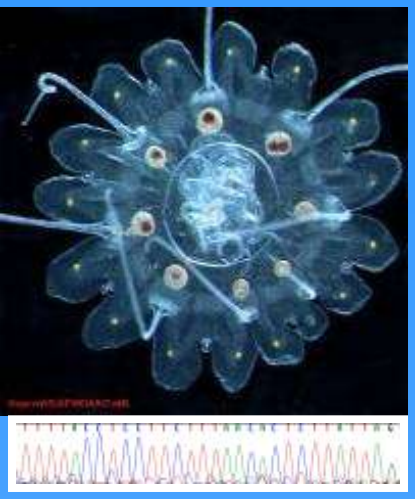


NEFSC continental shelf regions



NEFSC Eco-Mon stations

Conclusions



- DNA barcode library for described species of North Atlantic marine holozooplankton is being done by taxon-by-taxon barcoding, with CMarZ expert taxonomists working closely with DNA barcoders.
- CMarZ has pioneered at-sea taxonomic analysis and regional approaches to DNA barcoding.
- Environmental barcoding (sequencing DNA or rRNA from unsorted bulk zooplankton samples) will allow rapid determination of DNA sequences.
- Barcodes can provide accurate, routine identification of species only if unknown barcodes are present in database.
- Ecosystem monitoring using DNA barcodes is possible now and will be practical and cost-effective very soon.

Acknowledgements

R/V *RH Brown* - Apr 2006

UConn “Team DNA”

CMarZ Steering Group members

NMFS Northeast Fisheries Science Center

- Jerry Prezioso and Eco-Mon

Barcode data providers:

- UConn Team DNA
- Lisa Nigro (UNC, USA)
- Brian Ortman (UBC, Canada)
- CJ Sweetman (VIMS, USA)
- Rob Jennings (UMass-Boston, USA)

With appreciation for support from:

- Alfred P. Sloan Foundation
- NOAA Office of Ocean Exploration
- US NSF Biological Oceanography