Connectivity using DNA – the basics

Greg Rouse
Scripps Institution of Oceanography

WORKSHOP ON
THE DESIGN OF "IMPACT REFERENCE ZONES" AND "PRESERVATION REFERENCE ZONES" IN DEEP-SEA MINING CONTRACT AREAS

Sept 27-29 2017

What are species? Not easy to define across the diversity of life.

Evolutionary trees (phylogenies) vs populations; matter of scale.

DNA barcoding 'gaps' & delimiting 'species'.

Rise of DNA means 'species' discovered are unnamed= **Dark taxa**

Biogeography versus phylogeography; also a matter of scale.

Visualizing phylogenies versus populations = trees vs networks.

Some connectivity examples from deep sea organisms.

The word **Species** matters

Species affect:

- Estimates of species diversity/ biodiversity;
- Status of diagnosable populations;
- Understanding patterns of gene flow within these units= connectivity;
- Delineation of areas of endemism;
- Units to receive protection under local, national, or international legal instruments

Cracraft, J. "Species Concepts in Theoretical and Applied Biology: A Systematic Debate with Consequences." In Species Concepts: A Debate, ed. Q. D. Wheeler and R. Meier, 3-14. New York: Columbia University Press, 2000.

The plurality of species concepts= Dozens!

- Agamospecies concept
- **Biological species concept**
- Cohesion species concept
- Cladistic species concept
- Composite species concept
- Ecological species concept
- Evolutionary significant unit concept
- Evolutionary species concept
- Genealogical concordance concept
- Genetic species concept
- Genotypic cluster definition concept
- Hennigian species concept
- Internodal species concept
- Morphological species concept
- Non-dimensional species concept
- Phenetic species concept
- **Phylogenetic species concept**
  - Diagnosable version concept
  - Monophyly version concept
  - Diagnosable & monophyly version concept
- Polythetic species concept
- Recognition species concept
- Reproductive competition species concept
- Successional species concept
- Taxonomic species concept

Pleijel F. 1999. Phylogenetic Taxonomy, a Farewell to Species, and a revision of *Heteropodarke* (Hesionidae, Polychaeta, Annelida). Syst Biol 48(4):755-789.

There is no clear universal answer except to say that there are **different** kinds of species in nature

The current system of nomenclature; we have to name species

Linnean rank			
kingdom	Animalia	Animalia	Animalia
phylum	Mollusca	Chordata	Chordata
class	Gastropoda	Reptilia	Mammalia
order	Mesogastropoda	Ornithischia (Predentata)	Primates
family	Cypraeidae	Ceratopsidae	Hominidae
genus	*Cypraea*	*Triceratops*	*Homo*
species	*tigris*	*horridus*	*sapiens*

Rules on names, ranks etc. laid out in

International Code of Zoological Nomenclature,
International Code of Botanical Nomenclature,
International Code of Nomenclature of Bacteria etc.

These set guidelines and publish reports containing the rules of nomenclature.

Comparison of Biodiversity on Land and in the Ocean

http://www.marinespecies.org/aphia.php?p=stats

- **Accepted named marine (animal) species = 198,355 (8-3-2017)**
- Current described land (animal) species ~1.8 million
- How many more?

How many species are there?

Species	Earth			Ocean		
	Catalogued	Predicted	±SE	Catalogued	Predicted	±SE
Eukaryotes						
Animalia	953,434	7,770,000	958,000	171,082	2,150,000	145,000
Chromista	13,033	27,500	30,500	4,859	7,400	9,640
Fungi	43,271	611,000	297,000	1,097	5,320	11,100
Plantae	215,644	298,000	8,200	8,600	16,600	9,130
Protozoa	8,118	36,400	6,690	8,118	36,400	6,690
Total	1,233,500	8,740,000	1,300,000	193,756	2,210,000	182,000
Prokaryotes						
Archaea	502	455	160	1	1	0
Bacteria	10,358	9,680	3,470	652	1,320	436
Total	10,860	10,100	3,630	653	1,320	436
Grand Total	**1,244,360**	**8,750,000**	**1,300,000**	**194,409**	**2,210,000**	**182,000**

Predictions for prokaryotes represent a lower bound because they do not consider undescribed higher taxa. For protozoa, the ocean database was substantially more complete than the database for the entire Earth so we only used the former to estimate the total number of species in this taxon. All predictions were rounded to three significant digits.
doi:10.1371/journal.pbio.1001127.t002

Mora C, Tittensor DP, Adl S, Simpson AGB, Worm B (2011) How Many Species Are There on Earth and in the Ocean?. PLoS Biol 9(8): e1001127. doi:10.1371/journal.pbio.1001127

PLOS | BIOLOGY

Marine animal 'species' diversity

- **Mora C, Tittensor DP, Adl S, Simpson AGB, Worm B (2011) How Many Species Are There on Earth and in the Ocean? PLoS Biol**

- "our results suggest that some 86% of existing species on Earth and 91% of species in the ocean still await description."

- if 198,355 known then **~2.2 million still to name**

- **Given how little we have sampled (especially deep sea) we have only guesses as how many species there are.**

- **Only a small fraction has actually been named!**

What are species? Not easy to define across the diversity of life.

Evolutionary trees (phylogenies) vs populations; matter of scale.

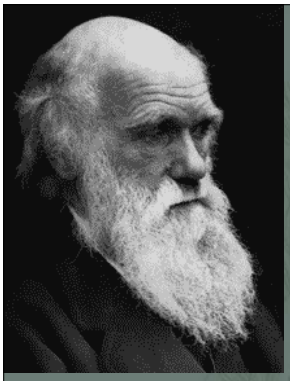
DNA barcoding 'gaps' & delimiting 'species'.

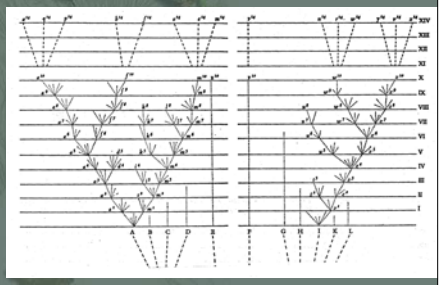
Rise of DNA means 'species' discovered are unnamed= **Dark taxa**

Biogeography versus phylogeography; also a matter of scale.

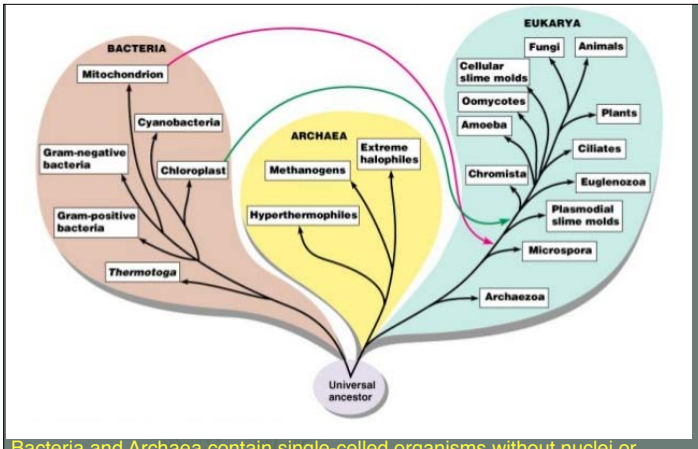
Visualizing phylogenies versus populations = trees vs networks.

Some connectivity examples from deep sea organisms.





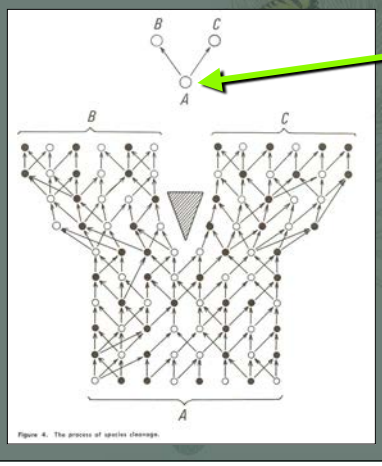
From ‘Origin of Species’ 1859



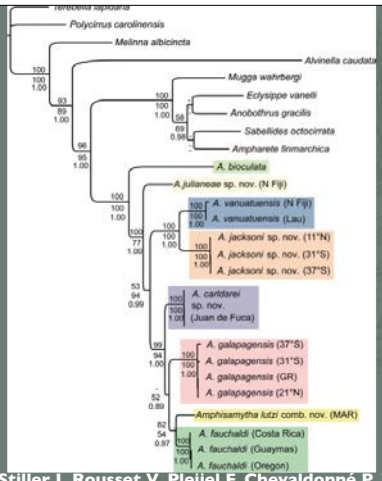
Bacteria and Archaea contain single-celled organisms without nuclei or organelles; collectively, they are called **prokaryotes**.

Fungi, protists, animals, and plants contain organisms with cells having nuclei and organelles; collectively, they are called **eukaryotes**.

Phylogeography



Put a magnifying glass here



Amphisamytha vents and seeps

Maximum likelihood tree from sequences

Stiller J, Rousset V, Pleijel F, Chevaldonné P, Vrijenhoek RC, Rouse GW (2013) Phylogeny, biogeography and systematics of hydrothermal vent and methane seep *Amphisamytha* (Ampharetidae, Annelida), with descriptions of three new species. Systematics and Biodiversity 11: 35–65

What are species? Not easy to define across the diversity of life.

Evolutionary trees (phylogenies) vs populations; matter of scale.

DNA barcoding ‘gaps’ & delimiting ‘species’.

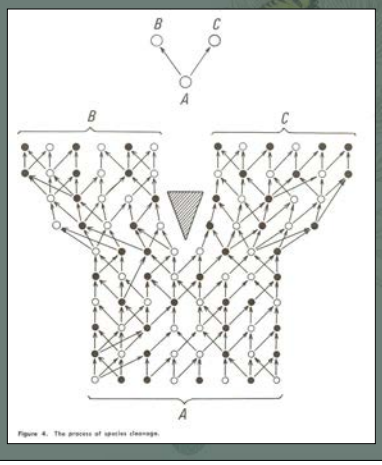
Rise of DNA means ‘species’ discovered are unnamed= **Dark taxa**

Biogeography versus phylogeography; also a matter of scale.

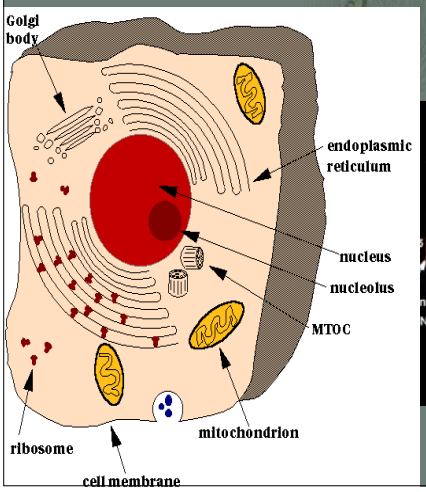
Visualizing phylogenies versus populations = trees vs networks.

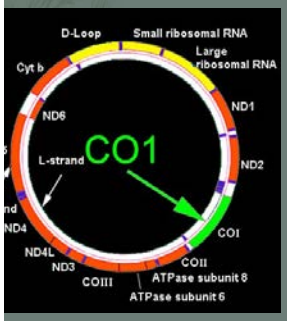
Some connectivity examples from deep sea organisms.

Telling ‘species’ apart

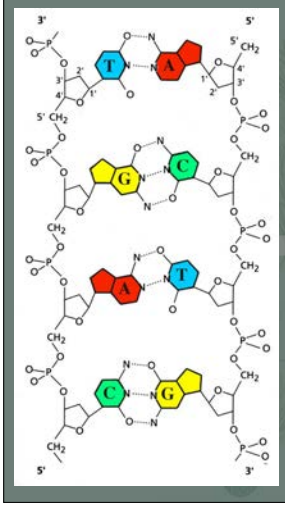


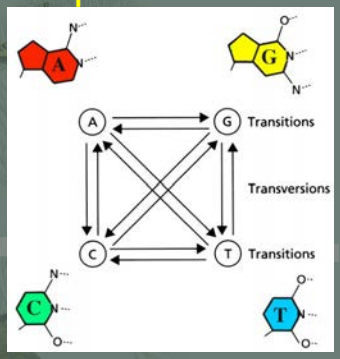
The chosen gene for animals





Some basic principles for DNA





- four bases of two kinds

AT and CG

genetic distance = n_d/n

- Estimate the mean number of changes per site since 2 taxa (sequences) split = number of differences / total length

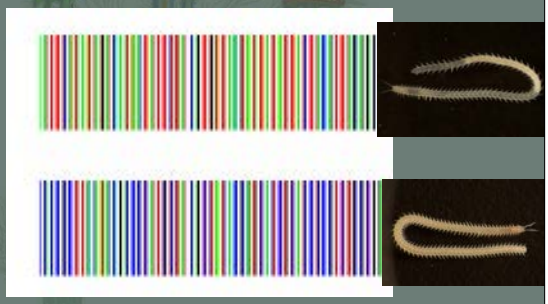
2 changes/40 sites;
D=0.05

= 5% distance

DNA Barcoding

Identify samples to 'species' based on a short standard DNA sequence

- Identifies
- Short, quick
- Standardised
- **Cheap!**



DNA Barcoding

THE ROYAL SOCIETY

Received 29 July 2002
Accepted 30 September 2002
Published online 8 January 2003

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

THE ROYAL SOCIETY biology letters

Barcoding animal life: cytochrome *c* oxidase subunit 1 divergences among closely related species

Paul D. N. Hebert*, Sujeevan Ratnasingham and Jeremy R. deWaard

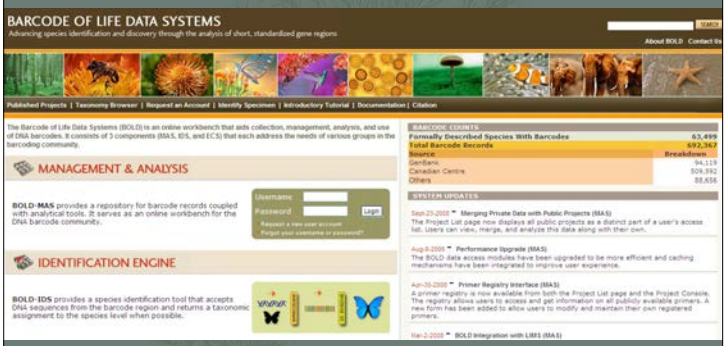
Department of Zoology, University of Guelph, Guelph, Ontario N1G 2W1, Canada
**Author for correspondence (phebert@uoguelph.ca).*

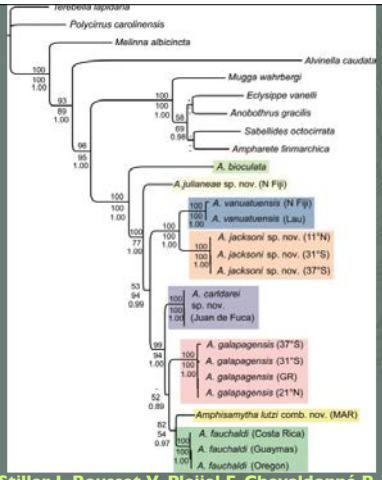
Rec'd 09.03.03; *Accept* 28.03.03; *Online* 15.05.03

Proposed a CO1-based (~650bp of the 5' end) global identification system of animals

98% of congeneric species in 11 animal phyla showed >2% sequence divergence in CO1

Barcode of Life Data Systems
BOLD - project management



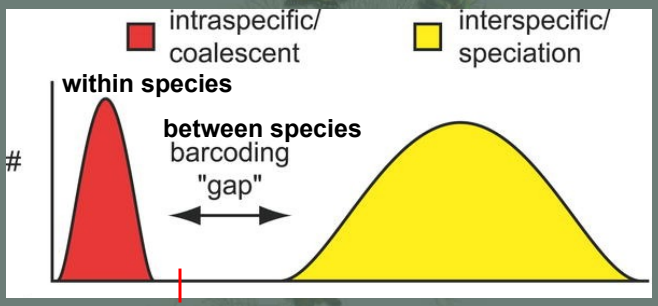


Amphisamytha vents & seeps

DNA reveals cryptic species

Stiller J, Rousset V, Pleijel F, Chevaldonné P, Vrijenhoek RC, Rouse GW (2013) Phylogeny, biogeography and systematics of hydrothermal vent and methane seep *Amphisamytha* (Ampharetidae, Annelida), with descriptions of three new species. Systematics and Biodiversity 11: 35–65

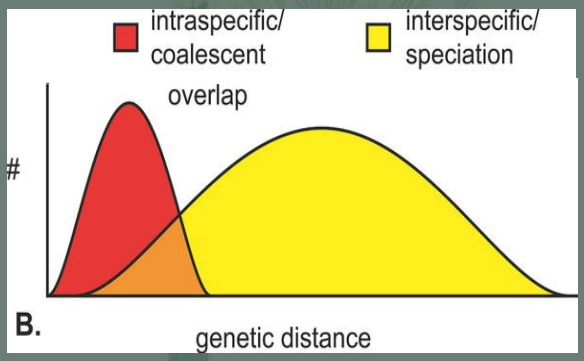
Does DNA Barcoding work?



Barcoding rests on the idea that **between** species genetic distance is larger, than **within** species variation.

Meyer and Paulay, PLoS Biology (2006)

Overlap the reality (often)



What are species? Not easy to define across the diversity of life.

Evolutionary trees (phylogenies) vs populations; matter of scale.

DNA barcoding ‘gaps’ & delimiting ‘species’.

Rise of DNA means ‘species’ discovered are unnamed= Dark taxa

Biogeography versus phylogeography; also a matter of scale.

Visualizing phylogenies versus populations = trees vs networks.

Some connectivity examples from deep sea organisms.

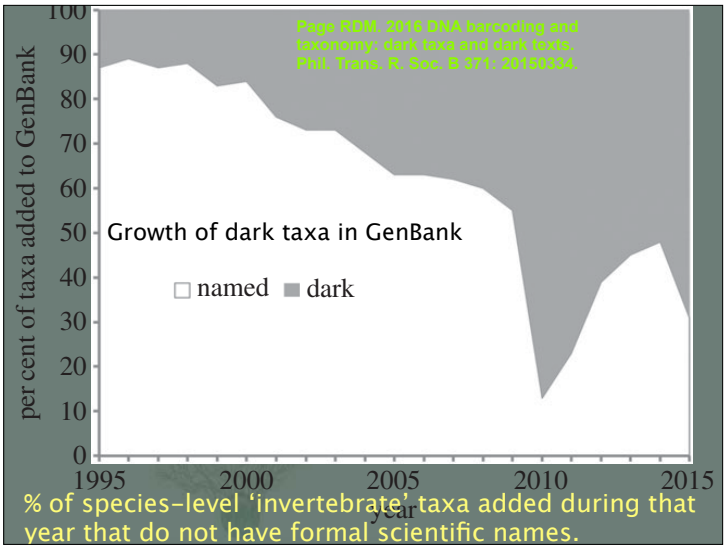
Lamellibrachia sp. MNHN IU-2013-79 cytochrome oxidase subunit I (COI) gene, partial cds; mitochondrial

GenBank: KP203875.1

FASTA Graphics PopSet

Go to:

```
LOCUS       KP203875                618 bp    DNA     linear   INV 30-NOV-2015
DEFINITION  Lamellibrachia sp. MNHN IU-2013-79 cytochrome oxidase subunit I
            (COI) gene, partial cds; mitochondrial.
ACCESSION   KP203875
VERSION     KP203875.1
KEYWORDS    .
SOURCE      mitochondrion Lamellibrachia sp. MNHN IU-2013-79
  ORGANISM  Lamellibrachia sp. MNHN IU-2013-79
            Eukaryota; Metazoa; Lophotrochozoa; Annelida; Polychaeta; Palpata;
            Canalipalpata; Sabellida; Siboglinidae; Lamellibrachia.
REFERENCE   1  (bases 1 to 618)
  AUTHORS   Samadi,S., Puillandre,N., Boisselier,M.-C., Chen,W.-J., Corbari,L.,
            Maestrati,P., Hana,R., Pante,E., Thubaut,J., Zuccon,D. and
            Hourdez,S.
  TITLE     Everything is not everywhere: the heterogeneity of the deep-sea
            makes it potentially vulnerable to mining activities in Papua
            New-Guinea
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 618)
  AUTHORS   Samadi,S., Puillandre,N., Boisselier,M.-C., Chen,W.-J., Corbari,L.,
            Maestrati,P., Hana,R., Pante,E., Thubaut,J., Zuccon,D. and
            Hourdez,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-NOV-2014) Systematique and Evolution, Museum National
            d'Histoire Naturelle, 43 rue Cuvier, Paris 75005, France
COMMENT     ##Assembly-Data-START##
            Sequencing Technology :: Sanger dideoxy sequencing
            ##Assembly-Data-END##
FEATURES             Location/Qualifiers
     source          1..618
                     /organism="Lamellibrachia sp. MNHN IU-2013-79"
                     /organelle="mitochondrion"
                     /mol_type="genomic DNA"
                     /specimen_voucher="MNHN:IU:2013-79"
                     /db_xref="taxon:1784442"
     gene            <1..>618
                     /gene="COI"
     CDS             <1..>618
                     /gene="COI"
                     /codon_start=3
                     /transl_table=5
                     /product="cytochrome oxidase subunit I"
                     /protein_id="ALA45055.1"
                     /translation="TULVATSMSLLIRAELGQPGTLLGDDQIYNCLITAHGLLMWFFV
                     VLPILMGGFGNWLVPLMLGAPDMAFPRINNLSFWLIPPAVILVNSAFIEKGAGTGWT
                     VYPPLASNIAHAGPCIDLAIFALHLSGISSILASINFITTVMNMRWPGMLEMVPLFV
                     WSVKLTAVLLLLSIPVLAGGLTMLLTDRNLNTSFFDPAGGGDPVL'
ORIGIN
        1 gaacaggtct agtagcoact agaataagac tactaattcg agctgaactt ggtcagccog
       61 gaacactctt agggagtgac caaattaaca attgcattat tacagctcac gggttctaa
      121 taatattctt tgtagttctc cctattttaa tagggggttt tggaaattga ctagttccct
      181 taatacttgg agctcctgat atagctttcc cacgaatcaa taatctagga ttttgactaa
      241 ttcccccggc agtaatctta ctagtgatat cogctttaat tgaaaaagct gccggaacag
```



% of species-level 'invertebrate' taxa added during that year that do not have formal scientific names.

What are species? Not easy to define across the diversity of life.

Evolutionary trees (phylogenies) vs populations; matter of scale.

DNA barcoding 'gaps' & delimiting 'species'.

Rise of DNA means 'species' discovered are unnamed= Dark taxa

Biogeography versus phylogeography; also a matter of scale.

Visualizing phylogenies versus populations = trees vs networks.

Some connectivity examples from deep sea organisms.

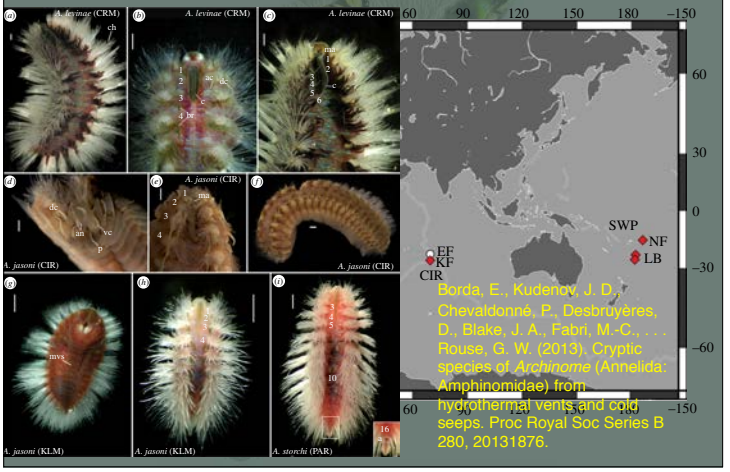
Historical Biogeography

- Long time scales (many millions of years)
- Large spatial scales (e.g. continents, oceans)
- Historical explanations (e.g. geological events)

Phylogeography

- DNA population studies
- Species ranges, connectivity of populations
- Intermediate time scales
- migrations, glaciations, bottlenecks

Archinome & vents & seeps

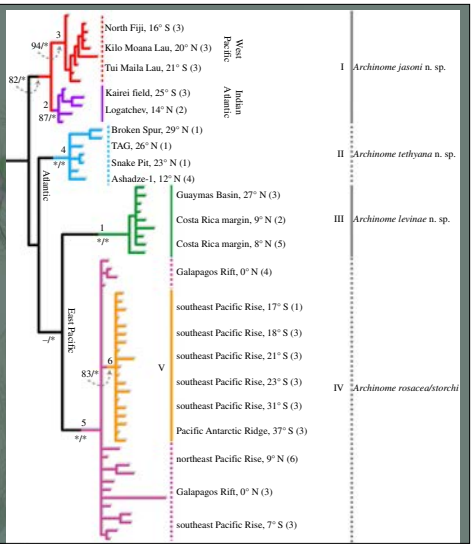


Borda, E., Kudenov, J. D., Chevaldonné, P., Desbruyères, D., Blake, J. A., Fabri, M.-C., . . . Rouse, G. W. (2013). Cryptic species of *Archinome* (Annelida: Amphinomidae) from hydrothermal vents and cold seeps. Proc Royal Soc Series B 280, 20131876.

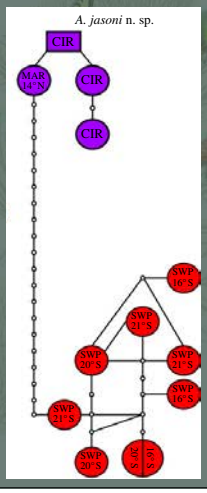
Archinome & vents & seeps

Borda, E., Kudenov, J. D., Chevaldonné, P., Desbruyères, D., Blake, J. A., Fabri, M.-C., . . . Rouse, G. W. (2013). Cryptic species of *Archinome* (Annelida: Amphinomidae) from hydrothermal vents and cold seeps. Proc Royal Soc Series B 280, 20131876.

DNA reveals cryptic species



Archinome jasoni
West Pacific
to Indian
to MAR?

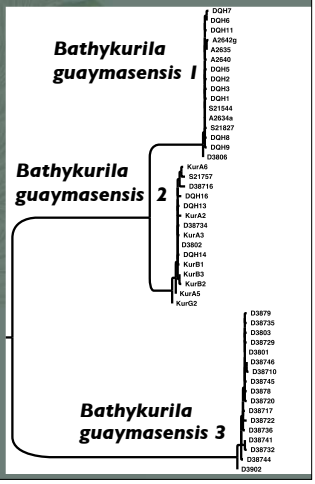


Borda, E., Kudenov, J. D., Chevaldonné, P., Desbruyères, D., Blake, J. A., Fabri, M.-C., . . . Rouse, G. W. (2013). Cryptic species of *Archinome* (Annelida: Amphinomidae) from hydrothermal vents and cold seeps. Proc Royal Soc Series B 280, 20131876.

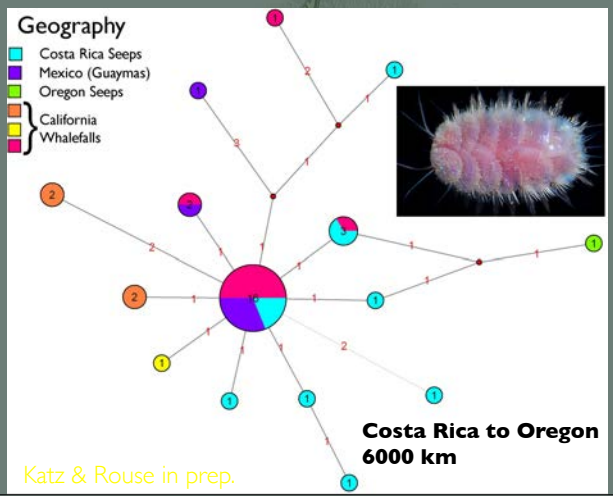
Bathykurila guaymasensis
whalefalls, seeps, vents

DNA reveals
cryptic species

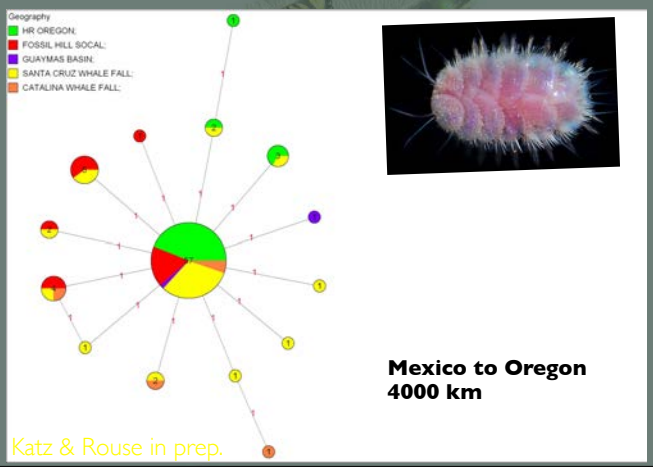
Katz & Rouse in prep.



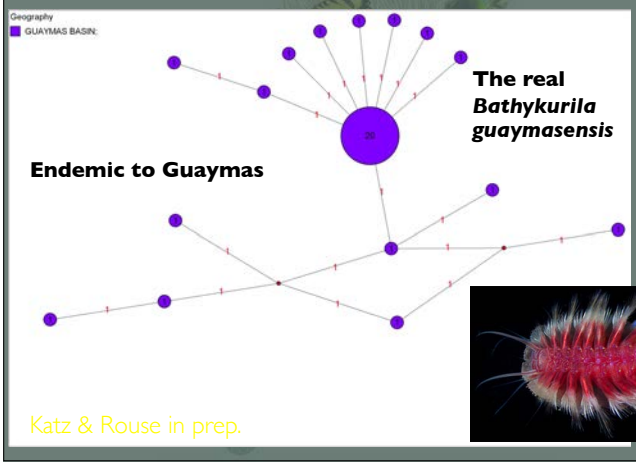
Bathykurila guaymasensis whalefalls, seeps, vents



Bathykurila guaymasensis whalefalls, seeps, vents



Bathykurila guaymasensis whalefalls, seeps, vents



The real *Bathykurila guaymasensis*

Endemic to Guaymas

Katz & Rouse in prep.

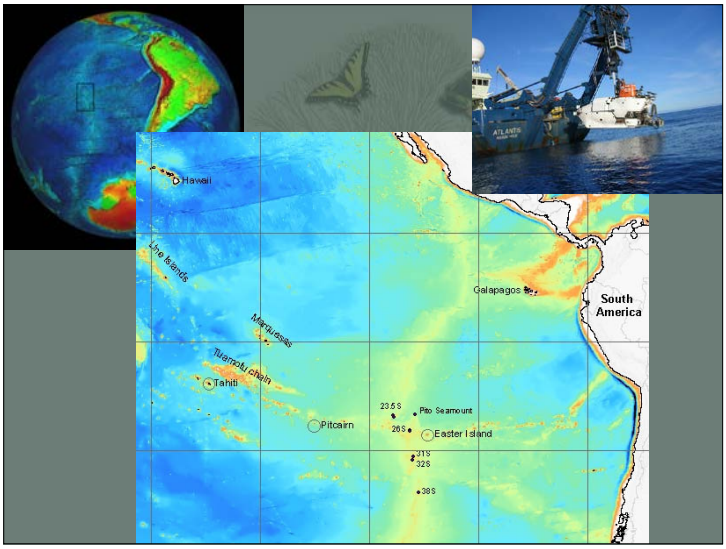
Comparative phylogeography

Hydrothermal vent fields located along the East Pacific Rise (EPR) and Galapagos Rift (GAR), where hundreds to thousands of kilometers can separate island-like populations.

Long-distance dispersal occurs via larval stages, but larval life histories differ among these taxa.

Hydrothermal vent annelids provide insights into dispersal modes and barriers to gene flow.

L. A. Hurtado, R. A. Lutz & R. C. Vrijenhoek. 2004. Distinct patterns of genetic differentiation among annelids of eastern Pacific hydrothermal vents. Molecular Ecology 13: 2603-2615.



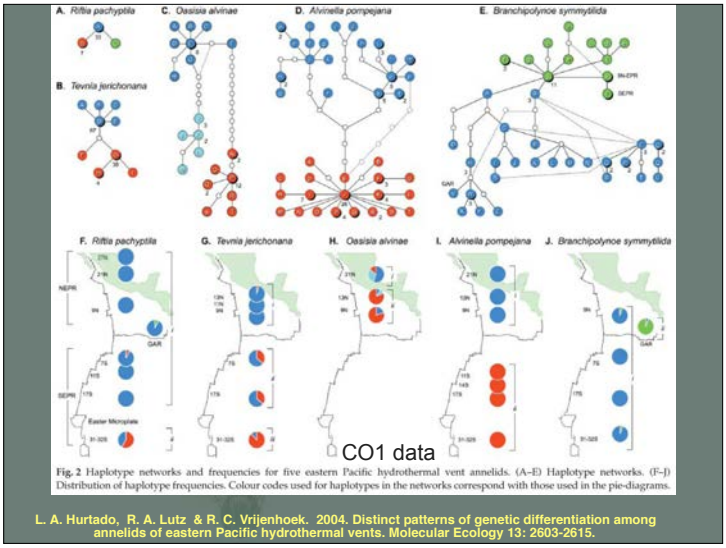


Fig. 2 Haplotype networks and frequencies for five eastern Pacific hydrothermal vent annelids. (A–E) Haplotype networks. (F–J) Distribution of haplotype frequencies. Colour codes used for haplotypes in the networks correspond with those used in the pie-diagrams.

L. A. Hurtado, R. A. Lutz & R. C. Vrijenhoek. 2004. Distinct patterns of genetic differentiation among annelids of eastern Pacific hydrothermal vents. Molecular Ecology 13: 2603-2615.

Conclusions

DNA data provides a cost-effective tool for monitoring biodiversity

Soon more 'species' will be known from DNA data than have been named

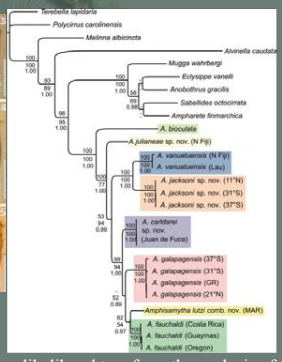
Nevertheless, DNA data can identify 'species' and also allow insights into connectivity of populations within a given 'species'

A baseline reference dataset, across proposed mining areas, linking DNA to morphology where possible, is fundamental to monitoring

Scaling up the data acquisition beyond DNA barcodes? New technology will also make this cost effective.

Amphisamytha vents and seeps

Stiller J, Rousset V, Pleijel F, Chevaldonné P, Vrijenhoek RC, Rouse GW (2013) Phylogeny, biogeography and systematics of hydrothermal vent and methane seep *Amphisamytha* (Ampharetidae, Annelida), with descriptions of three new species. Systematics and Biodiversity 11: 35–65



Maximum likelihood tree from the analysis of the cor sequences