

# Whole Organism Research Data



David Lazarus  
Museum für Naturkunde Berlin



# Outline

1. Types of Data

2. Data Storage and Retrieval

3. Data Management at the MfN

1. General activities

2. Own work as detailed example

4. Problems, Prospects and Future Developments



# I. Types of Data

- Whole Organism = Data linked to taxa names (species or other units): 'taxonomic data'
- Primary data in many biologic fields
- Taxonomic names are tags for scientific hypotheses
- Complex, heirarchical structure
- Taxa names, and much associated data, from human observation, not instruments



# 1.2 Where Created/Used

- Taxonomy and systematics
- Evolutionary biology
- Ecology
- Biodiversity/Conservation biology
- Paleontology
- Pathology/Disease control



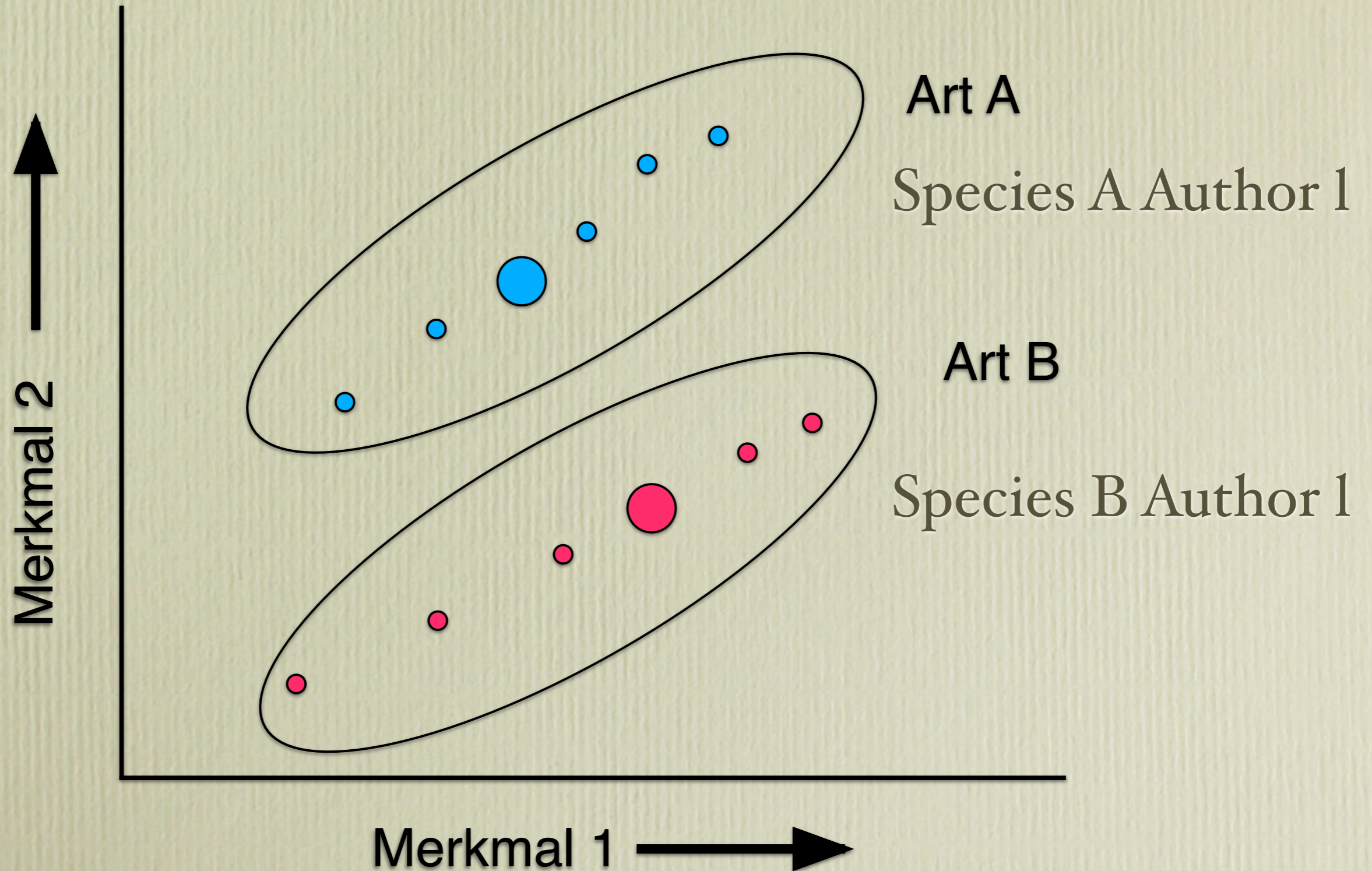
# I. 3 Names as Tags for Hypotheses

- Species have complex structure, indirectly inferred from very incomplete observational data
- Taxa name is tag for hypothesised structure and evolutionary history of natural populations
- Practice founded in theory (gene flow, development, reproductive barriers, evolution etc.) -> taxonomy is research, not 'stamp collecting'
- Hypotheses change with time, name tag follows hypothesis encompassing reference 'type' specimen



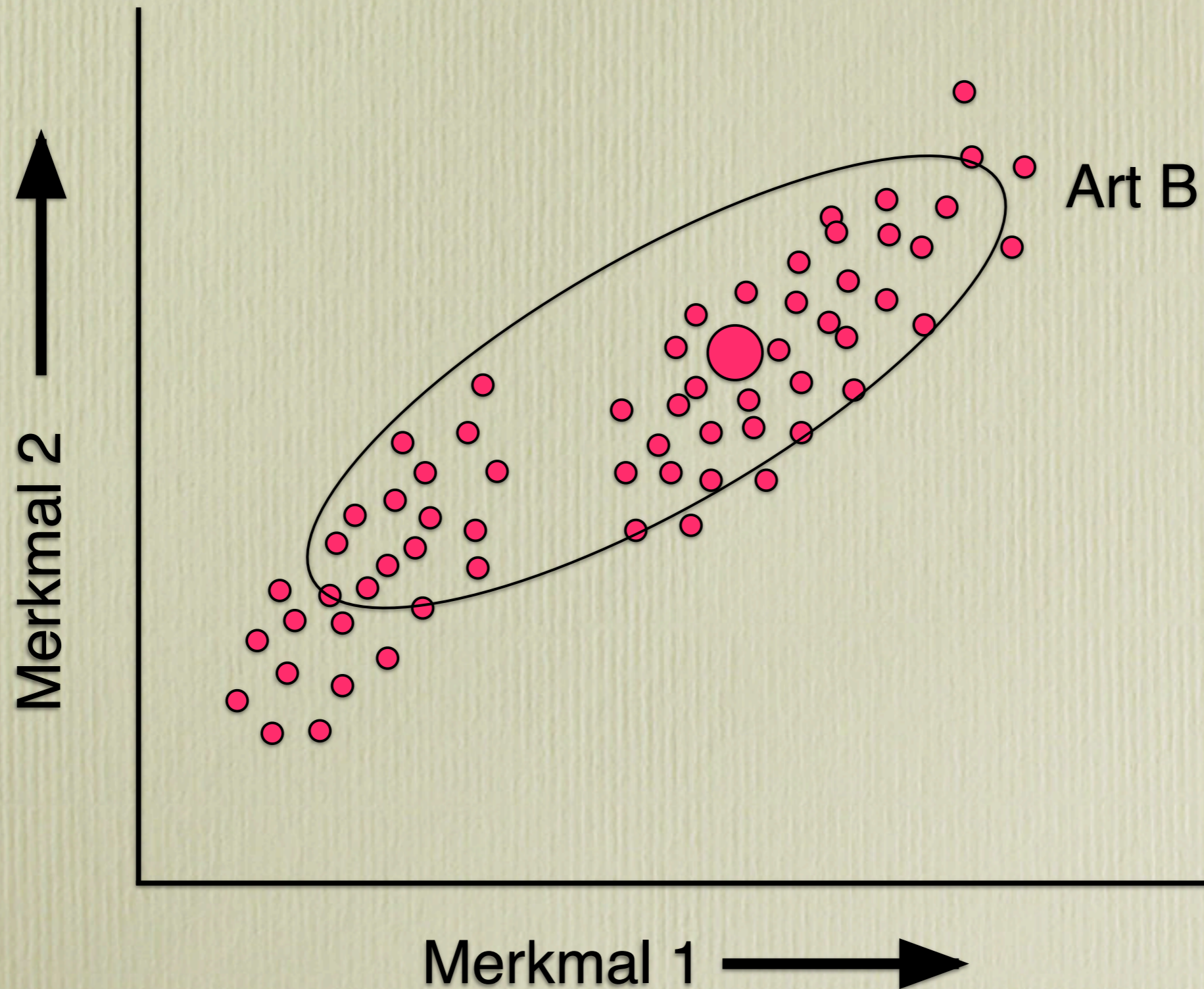


# Changing Species Concepts and Name Use Rule - 1 of 3



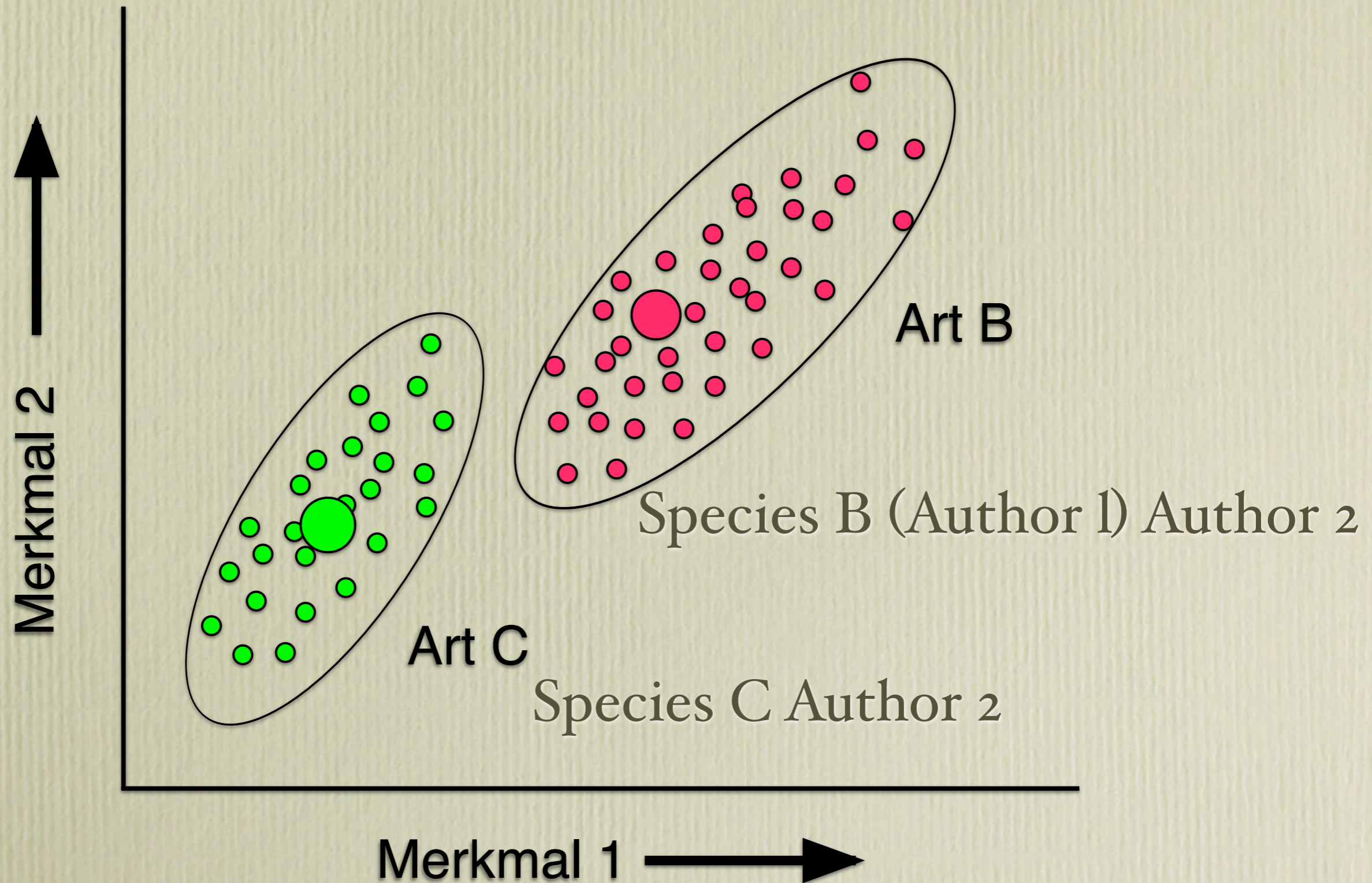


# Changing Species Concepts and Name Use Rule - 2 of 3





# Changing Species Concepts and Name Use Rule - 3 of 3





# Complex Name Relationships

- As evolutionary hypotheses, placement of species in genus may often change
- Genus concepts can also change
- Subspecies may be changed to species (and vice versa)
- etc



***Cycladophora davisiana***

*Pterocodon davisianus*

*Pterocanium davisianus*

*Eucyrtidium davisianum*

*Stichopilium davisianum*

*Theocalyptra davisiana*

*Diplocyclas davisiana*

*Cycladophora davisiana davisiana*

*Theocalyptra davisiana davisiana*

but not

*Cycladophora davisiana cornutoides*

= *Cycladophora cosma cosma*





# 1.4 Data Structure

- Taxa and data associated to them
- Taxa names are complex concepts 'embedded' in data records
- Associated data can include:
  - individual specimen characteristics (morphologic, genetic)
  - geographic and (paleontology) geologic occurrences
  - population to species level information (abundance, variation, ecology)
  - (museum) storage, use, curation



# Individual Specimen Data

*Genbank sequence for  
species of Radiolaria*

## Pterocanium trilobum gene for 18S rRNA, partial sequence, isolate: 1051

GenBank: AB246682.1

[FASTA](#) [Graphics](#)

[Go to:](#)

LOCUS AB246682 1718 bp DNA linear INV 04-AUG-2006  
DEFINITION Pterocanium trilobum gene for 18S rRNA, partial sequence, isolate:  
1051.

ACCESSION AB246682  
VERSION AB246682.1 GI:110278367

KEYWORDS .

SOURCE Pterocanium trilobum  
ORGANISM [Pterocanium trilobum](#)  
Eukaryota; Rhizaria; Polycystinea; Nassellaria; Theoperidae;  
Pterocanium.

REFERENCE 1

AUTHORS Kunitomo,Y., Sarashina,I., Iijima,M., Endo,K. and Sashida,K.  
TITLE Molecular phylogeny of acantharian and polycystine radiolarians  
based on ribosomal DNA sequences, and some comparisons with data  
from the fossil record

JOURNAL Eur. J. Protistol. 42 (2), 143-153 (2006)

PUBMED [17070759](#)

REFERENCE 2 (bases 1 to 1718)

AUTHORS Kunitomo,Y., Sarashina,I., Iijima,M., Endo,K. and Sashida,K.  
TITLE Direct Submission

JOURNAL Submitted (10-JAN-2006) Yoshiki Kunitomo, University of Tsukuba,  
Department of Earth Evolution Sciences, Graduate School of Life and  
Environmental Sciences; Department of Earth Evolution Sciences,  
Graduate School of Life and Environmental Sciences, University of  
Tsukuba, tsukuba, Ibaraki 305-8572, Japan  
(E-mail:yoshiki@arsia.geo.tsukuba.ac.jp, Tel:81-29-853-4427,  
Fax:81-29-851-9764)

FEATURES

Location/Qualifiers  
source 1..1718  
/organism="Pterocanium trilobum"  
/mol\_type="genomic DNA"  
/isolate="1051"  
/specimen\_voucher="PT-05-1-051 (Department of Earth  
Evolution Sciences, Univ. of Tsukuba)"  
/db\_xref="taxon:[366611](#)"  
[rRNA](#)  
<1..>1718  
/product="18S ribosomal RNA"

ORIGIN

```
1  tgaaactgcg aatggctcat tataacagtt atagtttatt tgatggttcc aattacgtgg
61  ataaccgtag taattctaga gctaatacat gctaaatacc ctggctttgc tgtcgggggtg
121 catttattag atgtcgaacc caatgctctt cggagcttat tgaggactca taataactct
181 cgcagatcgc aactgtgctg gcgtatcatt caagtttctg acctatcagt tagtcggtag
241 tgtattggac tactgaggct ctgacgggta acggagaatt agggttcgat tccggagagg
301 gagcctgaga aacggctacc acatctaagg aaggcagcag gcgcgtaact tacccaatcc
361 tgactcaggg aggtagtgac aagaaatcac aatatttcgc ccaattgggc aaaataattg
421 tattgagaac aggctaaaca tgtaacgat aaaccattgg agggcaagtc tggtgccagc
481 agccgcggta attccagctc caatagcgtg tactaacggt gttgcagtta aaaagctcgt
541 agttggattt ttagaagcgt gtaagcgtag tgaccttcg tgtctcttaa cgtgagcacc
601 gcttttattt tgtcccggcg tttggttggc attagtttgc cgatctaatt gttggggcgt
661 gtttactttg agaaaattag agtgttcaaa acagtcatac gattgtatat catagcatgg
721 aataataact gatgacctcg gtctctttt gttggattca gatggctgat gtaatgattg
781 atagggttag tcggaggtat tagtattctg tcgtagagg tgaaattcat ggatcggcag
841 aagactgaca actgcgaaag catttatcta ggatttgctc tttgatcaag aacgaaagtt
901 gggggatcga agacgatcag ataccgctgt agtctcaact ataaactata ccagctaggg
961 attgqcgagc gtttgacttg accttgctag aaccttatga gaaatcaaag catttgggtt
```



# Occurrence Data

valid_taxon_name	taxon_name	longitude	latitude	sample_age_ma	taxon_abundance
Anthocyrtidium angulare	Anthocyrtidium angulare	59.8645	16.6778	1.07109	P
Cycladophora cosma cosma	Corythospyris cosma cosma	-148.446	54.3654	0.117784	F
Cycladophora cosma cosma	Cycladophora davisiana cornutoides	151.629	33.926	0.130287	R
Cycladophora cosma cosma	Corythospyris cosma cosma	168.337	51.4504	0.159848	F
Cycladophora cosma cosma	Cycladophora davisiana cornutoides	143.236	40.6298	0.247479	R
Cycladophora cosma cosma	Cycladophora davisiana cornutoides	153.838	38.6268	0.357647	C
Cycladophora cosma cosma	Cycladophora davisiana cornutoides	159.796	43.927	0.40937	F
Cycladophora cosma cosma	Cycladophora davisiana cornutoides	-12.164	48.9158	0.525654	C
Cycladophora cosma cosma	Cycladophora davisiana cornutoides	-12.164	48.9158	0.536611	F
Cycladophora cosma cosma	Cycladophora davisiana cornutoides	-12.164	48.9158	0.572769	R
Cycladophora cosma cosma	Cycladophora davisiana cornutoides	143.236	40.6298	0.743379	R
Cycladophora cosma cosma	Corythospyris cosma cosma	-148.446	54.3654	0.763864	C
Cycladophora cosma cosma	Cycladophora davisiana cornutoides	153.838	38.6268	0.855503	R
Cycladophora cosma cosma	Cycladophora davisiana cornutoides	143.929	39.7355	0.862313	R
Cycladophora cosma cosma	Corythospyris cosma cosma	168.337	51.4504	14.4762	C
Cycladophora cosma cosma	Cycladophora cosma cosma	80.595	-61.579	16.8873	R
Lampromitra coronata	Lampromitra coronata	78.9981	-58.441	20.5411	F
Lychnocanoma elongata	Lychnocanoma elongata	158.512	-0.4987	17.4018	T
Lychnocanoma elongata	Lychnocanium bipes	158.61	-1.7422	24.9266	F
Lychnocanoma elongata	Lychnocanium bipes	158.61	-1.7422	25.1064	F



# Species Level Data

## *Ecologic classification of diatom species for lake quality surveys*

Table 3. Trophic and pH classifications for diatom taxa compiled from literature and supplemented by distributions observed in synoptic survey.

Taxon	Classification*	
	Trophic	pH
<i>Anomeoneis sphaerophora</i> (Ehr.) Pfitz. var. <i>sphaerophora</i>	●	alb
<i>Anomeoneis sphaerophora</i> var. <i>sculpta</i> O. Müll.	●	alb
<i>Anomeoneis vitrea</i> (Grun.) Ross var. <i>vitrea</i>	o-e	acp
<i>Asterionella formosa</i> Hass. var. <i>formosa</i>	m-e	acp
<i>Asterionella ralfsii</i> W. Sm. var. <i>ralfsii</i>	m	acb
<i>Caloneis ventricosa</i> (Ehr.) Meist. var. <i>ventricosa</i>	●	alp
<i>Caloneis ventricosa</i> var. <i>truncatula</i> (Grun.) Meist.	●	alp
<i>Capartogramma crucicula</i> (Grun. ex Cl.) Ross var. <i>crucicula</i>	u	●
<i>Chaetoceros</i> sp.	e	alp
<i>Cocconeis placentula</i> Ehr. var. <i>placentula</i>	e	cir
<i>Cocconeis placentula</i> var. <i>lineata</i> (Ehr.) V. H.	o	alp
<i>Coscinodiscus lacustris</i> Grun. var. <i>lacustris</i>	●	alp-cir
<i>Cyclotella comta</i> (Ehr.) Kütz. var. <i>comta</i>	●	alp
<i>Cyclotella meneghiniana</i> Kütz. var. <i>meneghiniana</i>	e	alp
<i>Cyclotella pseudostelligera</i> Hust. var. <i>pseudostelligera</i>	m-e	cir
<i>Cyclotella stelligera</i> Cl. u. Grun. var. <i>stelligera</i>	o-e	cir
<i>Cyclotella striata</i> (Kütz.) Grun. var. <i>striata</i>	●	alp-cir
<i>Cyclotella aspera</i> (Ehr.) H. Perag. var. <i>aspera</i>	●	alp



# 1.5 Data Sources

- Taxa names as conclusions from taxonomic research (published professional papers)
- Individual scientific observations from research papers, environmental agency surveys, etc
- Museum collections provide raw material used in research, and collection information can, with additional evaluation, become research data



## 2. Data Storage & Retrieval

- Taxonomic names and metadata
- Data associated with names



## 2. Taxonomic Name Data

- Catalogs with text description, illustrations, link to publication - still mostly printed but major effort to put online
- 'Bar-code' molecular definitions - typically in molecular databases linked to taxa name
- Taxa names, as scientific hypotheses, inherently global, and database efforts equally so
- Hundreds of specialised catalogs plus portals that integrate them
- 'Name Servers' that provide standardised, cross-referenced names as common key fields



# 2. On-line Taxonomic Catalog 1 of 2

Popular, well known  
groups with extensive  
associated data

*- tho not what bait best used to catch  
trophy specimens....*

*www.fishbase.org*

*Esox lucius* Linnaeus, 1758

Northern pike

Uploa  
| All pic

Picture by Jean-Fran

#### Classification / Names

Actinopterygii (Ray-finned fishes) > [Esoci](#)  
[Common names](#) | [Synonyms](#) | [Catalog of F](#)

#### Main reference

[Crossman, E.J. 1996. \(Ref. 26373\)](#)  
[Other references](#) | [Biblio](#) | [Coordinator](#) | [Co](#)

#### Size / Weight / Age

Max length : 137 cm FL male/unsexed; (R  
TL male/unsexed; (Ref. [556](#)); common len  
[40637](#)); max. published weight: 35 kg; max

#### Length at first maturity

L<sub>m</sub> [37.6](#), range 25 - 63 cm

#### Environment

Freshwater; brackish; demersal; potamodro  
[1998](#))

#### Climate / Range

Subtropical; 10°C - 28°C (Ref. [12741](#)); 74°N - 36°N, 167°W - 180°E

#### Distribution

Circumpolar in fresh water. North America: Atlantic, Arctic, Pacific, Great Lakes and Mississippi River basins from Labrador to Alaska and south to Pennsylvania, Missouri and Nebraska, USA (Ref. [5723](#)). Eurasia: Caspian, Black, Baltic, White, Barents, Arctic, North and Aral Seas and Atlantic basins, southwest to Adour drainage; Mediterranean basin in Rhône drainage and northern Italy. Widely distributed in central Asia and Siberia eastward to Anadyr drainage (Bering Sea basin). Historically absent from Iberian Peninsula, Mediterranean France, central Italy, southern and western Greece, eastern Adriatic basin, Iceland, western Norway and northern Scotland. Widely introduced and translocated throughout Europe (Ref. [59043](#)). Several countries report adverse ecological impact after introduction (Ref. [1739](#)).

[Countries](#) | [FAO areas](#) | [Ecosystems](#) | [Occurrences](#) | [Introductions](#)

#### Short description

[Dorsal spines](#) (total): 6 - 8; [Dorsal soft rays](#) (total): 17-25; [Anal spines](#): 4 - 7; [Anal soft rays](#): 10 - 22; [Vertebrae](#): 57 - 65. Diagnosed from all other freshwater fishes in Europe by the combination of the





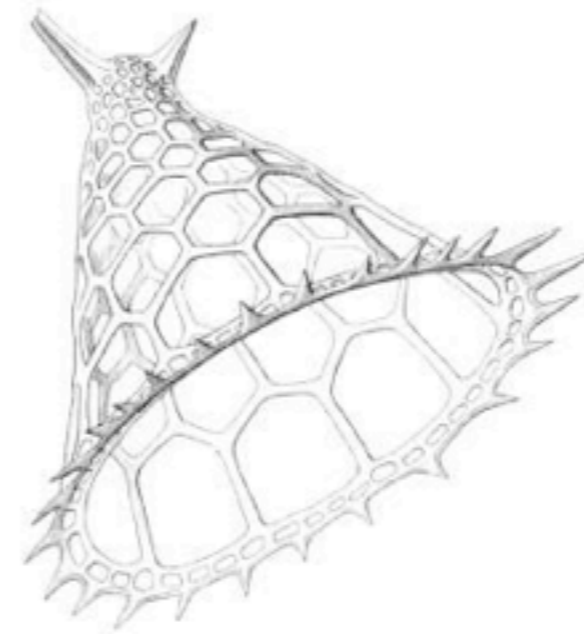
# 2. On-line Taxonomic Catalog 2 of 2

Less well known  
groups with little  
associated data

*(may in fact be synonym to  
another species)*

[Radiolaria.org](http://Radiolaria.org) | [Home](#) > [Cenozoic](#) > [Miocene](#) >

[List species](#)



***Clathrocyclas alcmenae* Haeckel, 1887**

**Description** - [Add description](#)

Shell conical, with two indistinct strictures. Length of the three joints = 2: 10: 1, breadth = 2: 12: 14. Cephalis hemispherical, with two divergent, pyramidal horns; frontal horn as long as the cephalis, occipital horn of twice the length. Thorax conical, with large hexagonal meshes, increasing gradually in size towards the short abdomen, which is represented only by a single circular girdle of small, square, abdominal pores. Coronal of the peristome with twenty to thirty triangular, divergent, nearly horizontal feet, as long as the cephalis.

**Dimensions**

Length of the three joints, a 0.02, b 0.1, c 0.01; breadth, a 0.02, b 0.12, c 0.14.

**Habitat** Central Pacific, Station 272, depth 2600 fathoms.

*Haeckel 1887*

[Description](#)

[Images](#)

[Synonyms](#)

[References](#)

[Distribution](#)

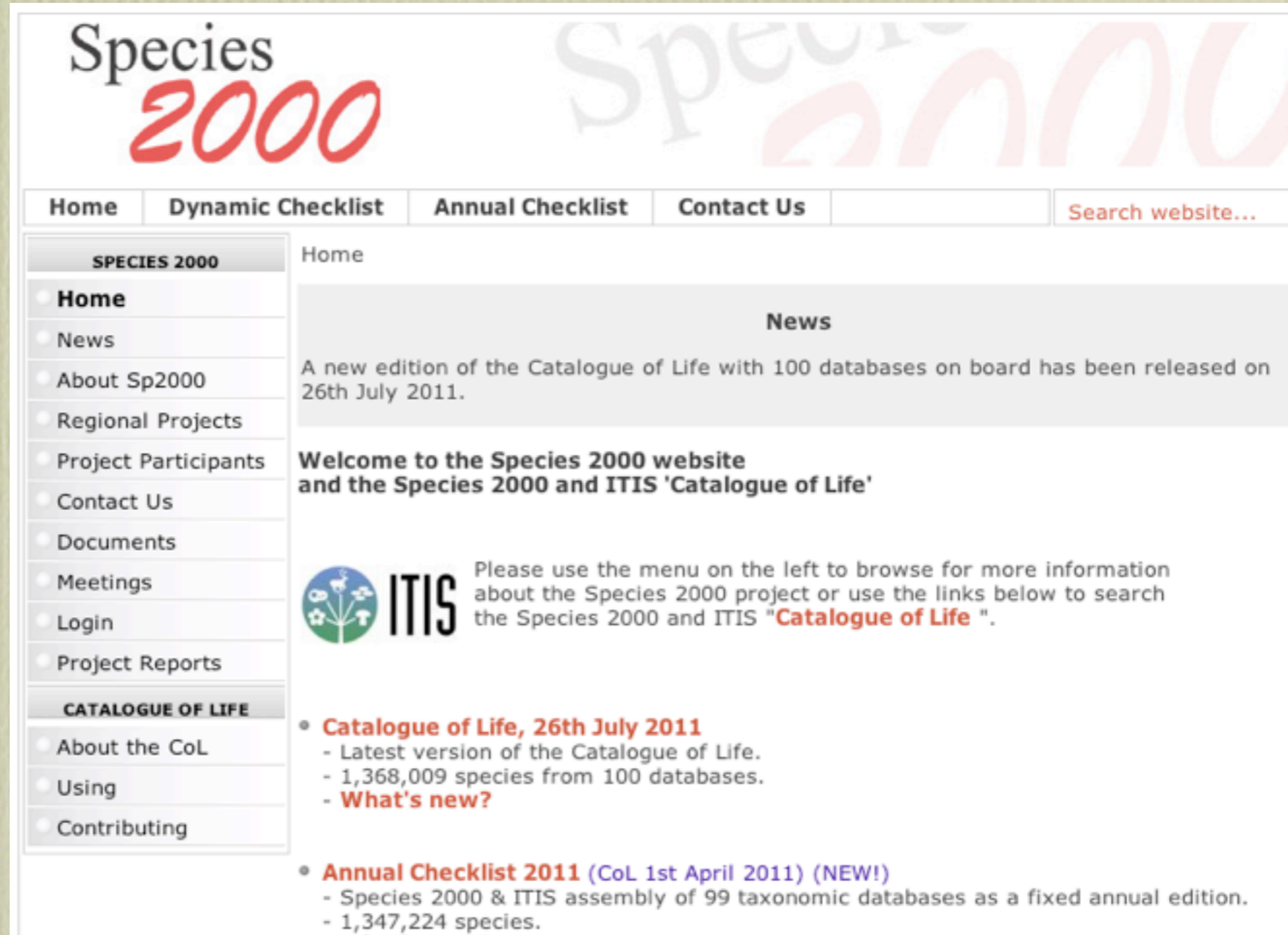
[Discussion / Comments](#)

[Web links](#)



# Name Servers

- Harvest names from many source catalog databases
- Provide common resolution to link names between databases
- Serve names to Associated data databases, e.g. OBIS, EOL



The screenshot shows the Species 2000 website. At the top left is the logo "Species 2000" with "Species" in black and "2000" in red. To the right is a search bar with the text "Search website...". Below the logo is a navigation menu with links: Home, Dynamic Checklist, Annual Checklist, and Contact Us. A sidebar on the left contains a menu for "SPECIES 2000" with links: Home, News, About Sp2000, Regional Projects, Project Participants, Contact Us, Documents, Meetings, Login, and Project Reports. Below this is a menu for "CATALOGUE OF LIFE" with links: About the CoL, Using, and Contributing. The main content area has a "Home" heading, a "News" section with a paragraph about a new edition of the Catalogue of Life, and a "Welcome to the Species 2000 website and the Species 2000 and ITIS 'Catalogue of Life'" section. Below this is the ITIS logo and a paragraph encouraging users to use the menu on the left. At the bottom, there are two bullet points: "Catalogue of Life, 26th July 2011" and "Annual Checklist 2011 (CoL 1st April 2011) (NEW!)", each with sub-bullets.

Species 2000

Home Dynamic Checklist Annual Checklist Contact Us Search website...

SPECIES 2000

- Home
- News
- About Sp2000
- Regional Projects
- Project Participants
- Contact Us
- Documents
- Meetings
- Login
- Project Reports

CATALOGUE OF LIFE


- About the CoL
- Using
- Contributing

Home

News

A new edition of the Catalogue of Life with 100 databases on board has been released on 26th July 2011.

Welcome to the Species 2000 website and the Species 2000 and ITIS 'Catalogue of Life'

 Please use the menu on the left to browse for more information about the Species 2000 project or use the links below to search the Species 2000 and ITIS "Catalogue of Life".

- **Catalogue of Life, 26th July 2011**
  - Latest version of the Catalogue of Life.
  - 1,368,009 species from 100 databases.
  - **What's new?**
- **Annual Checklist 2011 (CoL 1st April 2011) (NEW!)**
  - Species 2000 & ITIS assembly of 99 taxonomic databases as a fixed annual edition.
  - 1,347,224 species.



## 2. Data Associated to Taxa

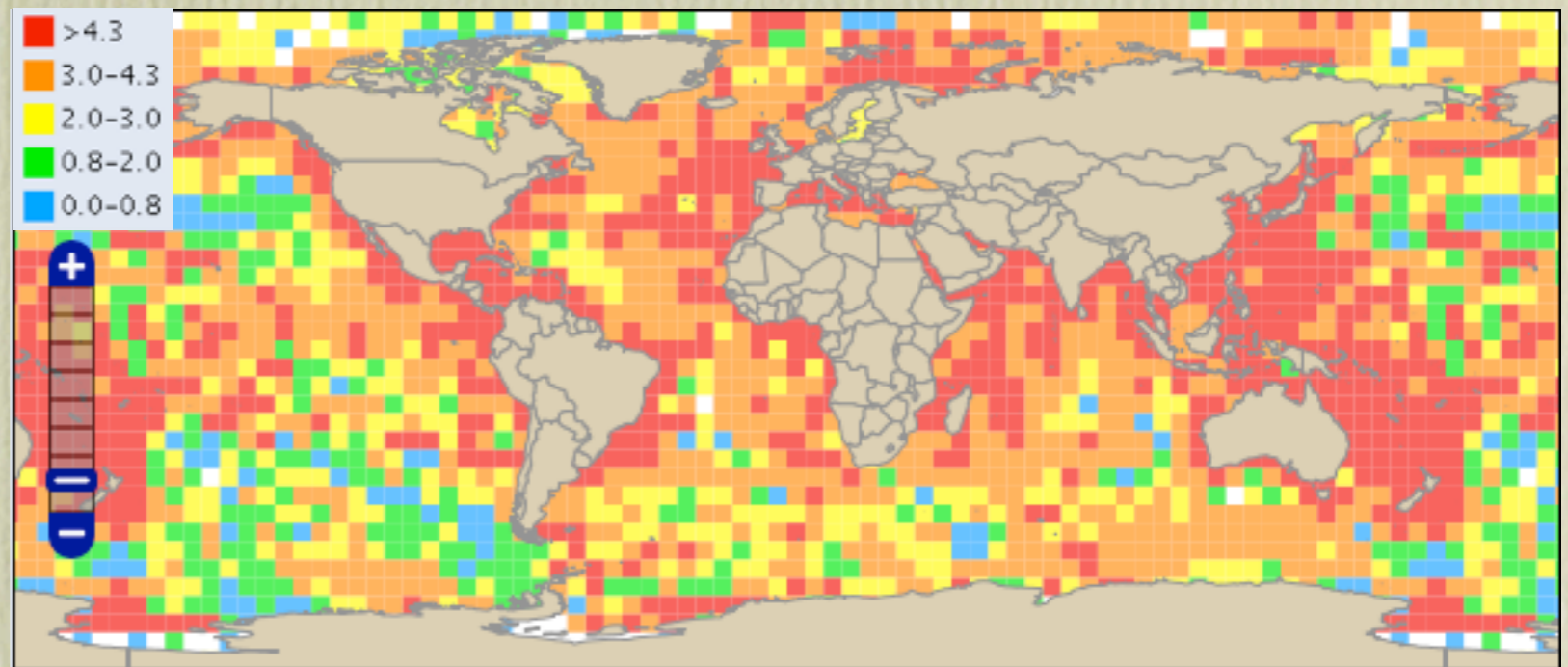
- Most data stored only locally [spreadsheets] by researcher, no management other than central backups
- ‘Supplemental’ data files at journal websites, World Data Center archives
- Global molecular databases
- Regional and higher taxon specific databases (government agencies, society databases)
- Global syntheses of published occurrence data (PBDB, Neptune, OBIS)
- Museum collection databases
- Portals to above



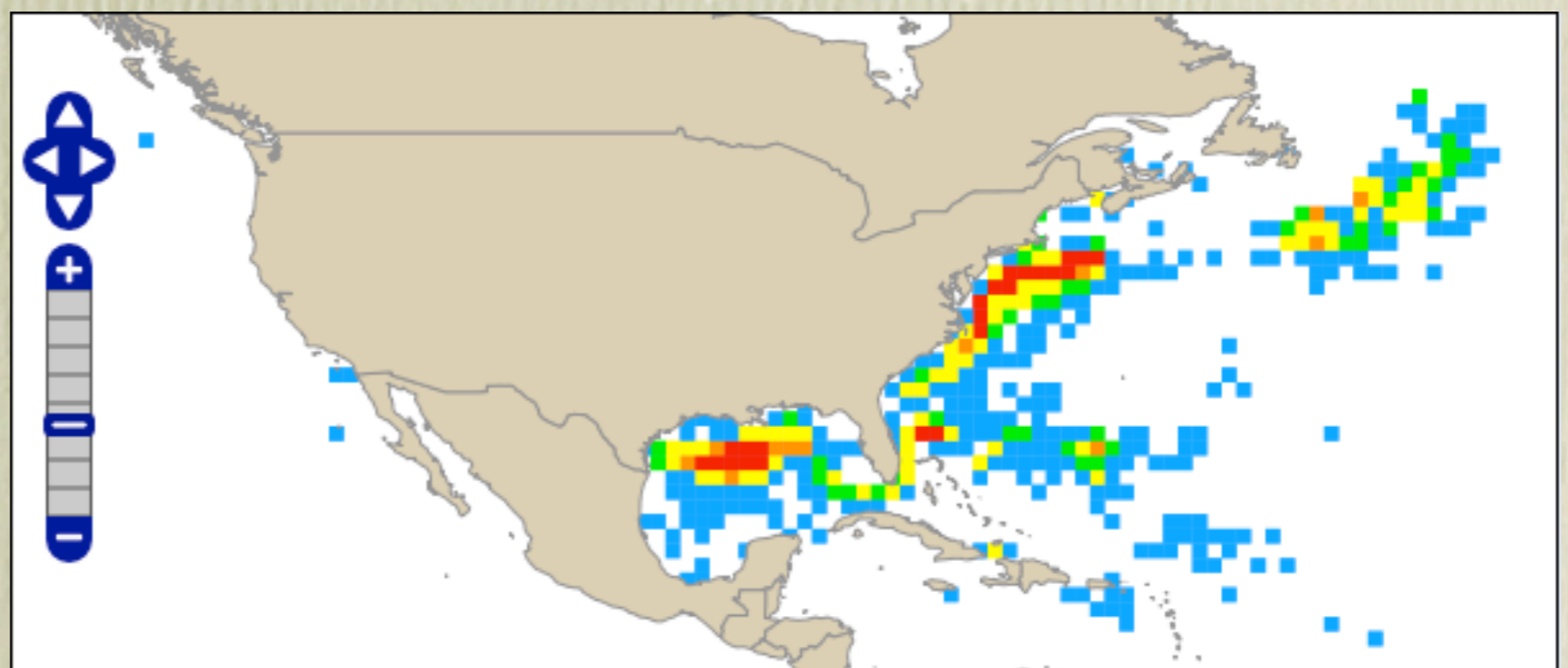
# Biogeographic Syntheses

- > 30 million records,  
>100,000 species,  
ca. 1,000 datasets
- National  
environmental &  
fishery agencies etc.,  
Museum collections,  
Research  
organisations

*Total Diversity (Shannon-Wiener)*



*Thunnus thynnus (bluefin Tuna)*





## 2. Portals - GBIF

- Most general portal to taxonomic information
- Similar sources as for OBIS but also terrestrial, fresh water data
- >300 million records
- 'Nodes' for specific higher taxa, IT development, other activities

... free and open access to biodiversity data  
GLOBAL BIODIVERSITY INFORMATION FACILITY

Search

HOME SPECIES COUNTRIES DATAS

Species: *Crocidura monticola* Peters, 1870  
Sunda Shrew

» Kingdom: Animalia » Phylum: Chordata » Class: Mammalia » Order: Soricomorpha » Family: Soricidae » Genus: *Crocidura* » S

**Actions for *Crocidura monticola***

**Explore:** [Occurrences \(42 records\)](#) [Names and classification](#)

**List:** [Countries with occurrences](#) [Datasets with occurrences](#)

**Download:** [Darwin Core records](#) [One-degree cell density overlay for Google Earth](#) [Placemarks for Go](#)

**Names and classification**

According to [Catalogue of Life: 2007 Annual Checklist: The Integrated Taxonomic Information System](#)

Name [Crocidura monticola Peters, 1870](#)

Classification » Kingdom: [Animalia](#) » Phylum: [Chordata](#) » Class: [Mammalia](#) » Order: [Soricomorpha](#) » Family: [Soricidae](#) » Genus: [Crocidura](#) » Species: [Crocidura monticola](#)

Status [Accepted name](#)

Synonyms [Crocidura neglecta](#), [Crocidura minuta](#)

Common names [English : Sunda Shrew](#)

Record identifier [ITS-633637](#)

Record URL <http://www.itis.gov/servlet>


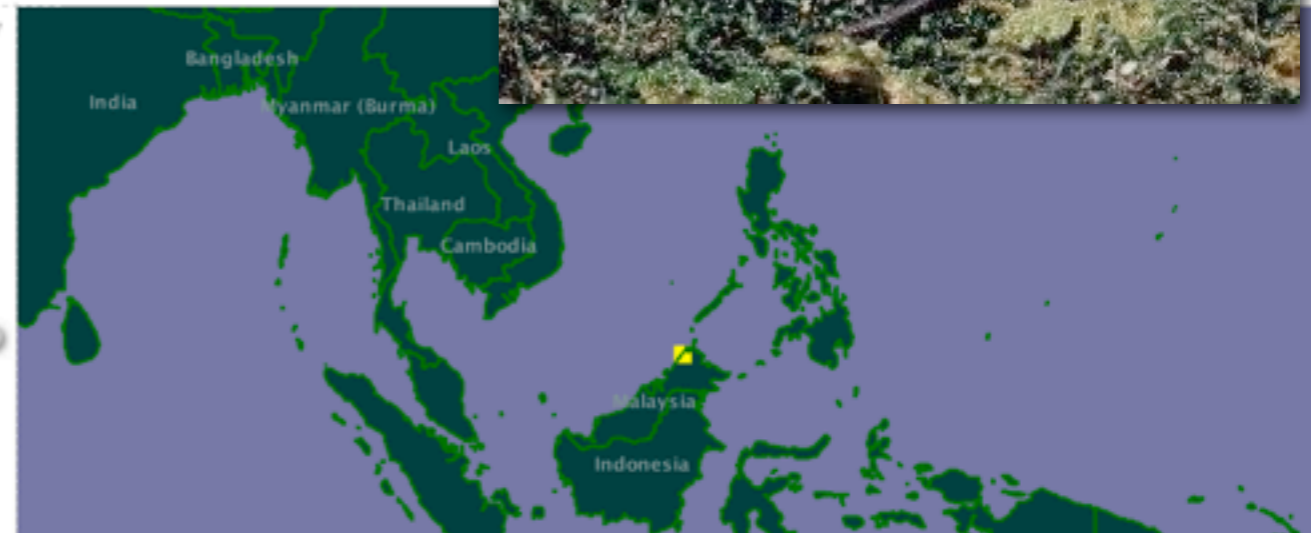
Review date [28-Nov-2006](#)

Feedback [Feedback to Catalogue Peters, 1870](#)  
Please note that this feedback concerns species page instead.

**Occurrence overview**

77

27





# 3. Data Management at the MfN

- Collection Databases
  - 1<sup>rst</sup> general system (Specify 6, USA); now in beta
  - GBIF: Collection records partially linked
  - GBIF Node: add paleontology to GBIF (ABCDEFGFG)
- Primary, individually held research results (unmanaged)
- Research Data : Contributions to global efforts
  - Taxa name databases
  - Neptune, PBDB, EOL, WoRMS
  - Document access - BHL



# History of Neptune and NSB

- Initiated in early 1990s as local microfossil occurrence database by Lazarus & ETH (Zürich) team, as counterpart to Sepkoski database & to fill gap in ODP capabilities (Janus shipboard data only, mostly raw data archive)



*ETH*



*Ames*

*MfN Berlin*



- Recast (ca 2005) as standard sql internet database by Chronos Iowa (Cervato, Diver, Fils and others) - funding ended ca. 2008
- ‘Researcher’ version NSB (for ‘Neptune Sandbox Berlin’) initiated 2010 by Lazarus and Diver w. funding from CEES Oslo (Stenseth, Liow)



# Papers Citing Neptune/Results

- Ezard, T. H. G., Aze, T., Pearson, P. N. & Purvis, A. 2011. Interplay between changing climate and species' ecology drives macroevolutionary dynamics. *Science*, **332**, 349-351.
- Powell, M. G. & MacGregor, J. 2011. A geographic test of species selection using planktonic foraminifera during the Cretaceous/Paleogene mass extinction. *Paleobiology*, **37**, 426-437.
- Renaudie, J., Danelian, T., Saint, M. S., Le, C. L. & Tribovillard, N. 2010. Siliceous phytoplankton response to a middle Eocene warming event recorded in the tropical Atlantic (Demerara Rise, ODP Site 1260A). *Palaeogeography Palaeoclimatology Palaeoecology*, **286**, 121-134.
- Marx, F. G. & Uhen, M. D. 2010. Climate, critters, and cetaceans: Cenozoic drivers of the evolution of modern whales. *Science*, **327**, 993-996.
- Crux, J. A., Gary, A., Gard, G. & Ellington, W. E. 2010. Recent advances in the application of biostratigraphy to hydrocarbon exploration and production. In: *Application of Modern Stratigraphic Techniques: Theory and Case Histories* (ed. by K. T. Ratcliffe & B. A. Zaitlin), pp. 57-80. SEPM.
- Hayward, B. W., Sabaa, A. T., Thomas, E., Kawagata, S., Nomura, R., Schröder-Adams, C., Gupta, A. K. & Johnson, K. 2010. Cenozoic record of elongate, cylindrical, deep-sea benthic foraminifera in the Indian Ocean (ODP Sites 722, 738, 744, 758 and 763). *Journal of Foraminiferal Research*, **40**, 113-133.
- Hayward, B. W., Johnson, K., Sabaa, A. T., Kawagata, S. & Thomas, E. 2010. Cenozoic record of elongate, cylindrical, deep-sea benthic foraminifera in the North Atlantic and equatorial Pacific Oceans. *Marine Micropalaeontology*, **74**, 75-95.
- Liow, L. H., Skaug, H. J., Ergon, T. & Schweder, T. 2010. Global occurrence trajectories of microfossils: environmental volatility and the rise and fall of individual species. *Paleobiology*, **36**, 224-252.
- Frada, M., Percopo, I., Young, J., Zingone, A., de Vargas, C. & Probert, I. 2009. First observations of the heterococcolithophore-holococcolithophore life cycle combinations in the family Pontosphaeraceae (Calcihaptophycidae, Haptophyta). *Marine Micropalaeontology*, **71**, 20-27.
- Rabosky, D. L. & Sorhannus, U. 2009. Diversity dynamics of marine planktonic diatoms across the Cenozoic. *Nature*, **247**, 183-187.
- Fils, D., Cervato, C., Reed, J., Diver, P., Tang, X., Bohling, G. & Greer, D. 2009. CHRONOS architecture: Experiences with an open-source services-oriented architecture for geoinformatics. *Comp. Geosci.* **35**(4), 774-782.
- Cermeño, P. & Falkowski, P. G. 2009. Controls on Diatom Biogeography in the Ocean. *Science*, **325**, 1539-1541.
- Henderiks, J. & Pagani, M. 2008. Coccolithophore cell size and the Paleogene decline in atmospheric CO<sub>2</sub>. *Earth and Planetary Science Letters*, **269**, 575-583.
- Foote, M., Crampton, J. S., Beu, A. G. & Cooper, R. A. 2008. On the bidirectional relationship between geographic range and taxonomic duration. *Paleobiology*, **34**, 421-433.
- Cody, R. D., Levy, R. H., Harwood, D. M. & Sadler, P. M. 2008. Thinking outside the zone: High-resolution quantitative diatom biochronology for the Antarctic Neogene. *Palaeogeog., Palaeoclim., Palaeoecol.*, **260**, 92-121.
- Allen, A. P. & Savage, V. M. 2007. Setting the absolute tempo of biodiversity dynamics. *Ecology Letters*, **10**, 637-646.
- McPeck, M. 2007. The macroevolutionary consequences of ecological differences among species. *Palaeontology*, **50**, 111s129.
- Liow, L. H. & Stenseth, N. C. 2007. The rise and fall of species: implications for macroevolutionary and macroecological studies. *Proc. Royal Soc. B*, **274**, 2745-2752.
- Kucera, M. & Schönfeld, J. 2007. The origin of modern oceanic foraminiferal faunas and Neogene climate change. In: *Deep-Time Perspectives on Climate Change: Marrying the Signal from Computer Models and Biological Proxies* (ed. by M. Williams, A. M. Haywood, F. J. Gregory & D. N. Schmidt), pp. 409-425. The Geological Society, London.
- Finkel, Z. V., Sebbo, J., Feist-Burkhardt, S., Irwin, A. J., Katz, M. E., Schofield, O., Young, J. & Falkowski, P. G. 2007. A universal driver of macroevolutionary change in the size of marine phytoplankton over the Cenozoic. *Proc. Natl. Acad. Sci. USA*, **104**, 20416-20420.
- Muttoni, G. & Kent, D. V. 2007. Widespread formation of cherts during the early Eocene climatic optimum. *Palaeogeog., Palaeoclim., Palaeoecol.*, **253**, 348-362.
- Lazarus, D., Bittniok, B., Diester-Haass, L., Meyers, P. & Billups, K. 2006. Comparison of radiolarian and sedimentologic paleoproductivity proxies in the latest Miocene-Recent Benguela Upwelling System. *Marine Micropaleontology*, **60**, 269-294.
- Allen, A. P., Gillooly, J. F., Savage, V. M. & Brown, J. H. 2006. Kinetic effects of temperature on rates of genetic divergence and speciation. *Proc. Nat. Acad. Sci. USA*, **103**(24), 9130-9135.
- Allen, A. P. & Gillooly, J. F. 2006. Assessing latitudinal gradients in speciation rates and biodiversity at the global scale. *Ecology Letters*, **9**, 947-954.
- Spezzaferri, S. & Spiegler, D. 2005. Fossil planktic foraminifera (an overview). *Paläontologische Zeitschrift*, **79**, 149-166.
- Finkel, Z. V., Katz, M. E., Wright, J. D., Schofield, O. & Falkowski, P. G. 2005. Climatically driven macroevolutionary patterns in the size of marine diatoms over the Cenozoic. *Proc. Natl. Acad. Sci. USA*, **102**, 8927-8932.
- Tozzi, S., Schofield, O. & Falkowski, P. 2004. Historical climate change and ocean turbulence as selective agents for two key phytoplankton functional groups. *Marine Ecology Progress Series*, **274**, 123-132.
- Thierstein, H. R., Cortés, M. & Haidar, A. 2004. Plankton community behavior on ecological and evolutionary time-scales: when models confront evidence. In: *Coccolithophores: From Molecular Processes to Global Impact* (ed. by H. Thierstein & J. R. Young), pp. 455-480. Springer, Berlin.
- Falkowski, P. G., Katz, M. E., Knoll, A. H., Quigg, A., Raven, J. A., Schofield, O. & Taylor, F. J. R. 2004. The evolution of modern eukaryotic phytoplankton. *Science*, **305**, 354-360.
- Katz, M. E., Finkel, Z. V., Grzebyk, D., Knoll, A. & Falkowski, P. G. 2004. Evolutionary trajectories and biogeochemical impacts of marine eukaryotic phytoplankton. *Ann Rev. Ecol. Syst.*, **35**, 523-556.
- Schmidt, D. N., Thierstein, H. R. & Bollmann, J. 2004. The evolutionary history of size variation of planktic foraminiferal assemblages in the Cenozoic. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **212**, 159-180.
- Cervato, C. & Burckle, L. 2003. Pattern of first and last appearance in diatoms: Oceanic circulation and the position of polar fronts during the Cenozoic. *Paleoceanography*, **18**, web.
- Lazarus, D. B. 2002. Environmental control of diversity, evolutionary rates and taxa longevities in Antarctic Neogene radiolaria. *Paleontologica Electronica*, **5**, 1-30.
- Spencer-Cervato, C. 1999. The Cenozoic deep sea microfossil record: explorations of the DSDP/ODP sample set using the Neptune database. *Palaeontologica Electronica*, **2**, web.
- Lazarus, D. B., Spencer-Cervato, C., Pianka-Biolzi, M., Beckmann, J. P., von Salis, K., Hilbrecht, H. & Thierstein, H. R. 1995. *Revised chronology of Neogene DSDP holes from the world ocean. Ocean Drilling Program*: College Station. ca. 250 p.
- Lazarus, D. B. 1994. The Neptune Project - a marine micropaleontology database. *Math. Geol.*, **26**, 817-832.
- Spencer-Cervato, C., Thierstein, H. R., Lazarus, D. B. & Beckmann, J. P. 1994. How synchronous are Neogene marine plankton events? *Paleoceanography*, **9**, 739-763.
- Spencer-Cervato, C., Lazarus, D. B., Beckmann, J. P., Perch-Nielsen, K. v. S. & Biolzi, M. 1993. New calibration of Neogene radiolarian events in the North Pacific. *Marine Micropaleontology*, **21**, 261-294.



# Papers Citing Neptune/Results

- Ezard, T. H. G., Aze, T., Pearson, P. N. & Purvis, A. 2011. Interplay between changing climate and species' ecology drives macroevolutionary dynamics. *Science*, **332**, 349-351.
- Powell, M. G. & MacGregor, J. 2011. A geographic test of species selection using planktonic foraminifera during the Cretaceous/Paleogene mass extinction. *Paleobiology*, **37**, 426-437.
- Renaudie, J., Danelian, T., Saint, M. S., Le, C. L. & Tribovillard, N. 2010. Siliceous phytoplankton response to a middle Eocene warming event recorded in the tropical Atlantic (Demerara Rise, ODP Site 1260A). *Palaeogeography Palaeoclimatology Palaeoecology*, **286**, 121-134.
- Marx, F. G. & Uhen, M. D. 2010. Climate, critters, and cetaceans: Cenozoic drivers of the evolution of modern whales. *Science*, **327**, 993-996.
- Crux, J. A., Gary, A., Gard, G. & Ellington, W. E. 2010. Recent advances in the application of biostratigraphy to hydrocarbon exploration and production. In: *Application of Modern Stratigraphic Techniques: Theory and Case Histories* (ed. by K. T. Ratcliffe & B. A. Zaitlin), pp. 57-80. SEPM.
- Hayward, B. W., Sabaa, A. T., Thomas, E., Kawagata, S., Nomura, R., Schröder-Adams, C., Gupta, A. K. & Johnson, K. 2010. Cenozoic record of elongate, cylindrical, deep-sea benthic foraminifera in the Indian Ocean (ODP Sites 722, 738, 744, 758 and 763). *Journal of Foraminiferal Research*, **40**, 113-133.
- Hayward, B. W., Johnson, K., Sabaa, A. T., Kawagata, S. & Thomas, E. 2010. Cenozoic record of elongate, cylindrical, deep-sea benthic foraminifera in the North Atlantic and equatorial Pacific Oceans. *Marine Micropalaeontology*, **74**, 75-95.
- Liow, L. H., Skaug, H. J., Ergon, T. & Schweder, T. 2010. Global occurrence trajectories of microfossils: environmental volatility and the rise and fall of individual species. *Paleobiology*, **36**, 224-252.
- Frada, M., Percopo, I., Young, J., Zingone, A., de Vargas, C. & Probert, I. 2009. First observations of the heterococcolithophore-holococcolithophore life cycle combinations in the family Pontosphaeraceae (Calcihaptophycidae, Haptophyta). *Marine Micropalaeontology*, **71**, 20-27.
- Rabosky, D. L. & Sorhannus, U. 2009. Diversity dynamics of marine planktonic diatoms across the Cenozoic. *Nature*, **247**, 183-187.
- Fils, D., Cervato, C., Reed, J., Diver, P., Tang, X., Bohling, G. & Greer, D. 2009. CHRONOS architecture: Experiences with an open-source services-oriented architecture for geoinformatics. *Comp. Geosci.* **35**(4), 774-782.
- Cermeño, P. & Falkowski, P. G. 2009. Controls on Diatom Biogeography in the Ocean. *Science*, **325**, 1539-1541.
- Henderiks, J. & Pagani, M. 2008. Coccolithophore cell size and the Paleogene decline in atmospheric CO<sub>2</sub>. *Earth and Planetary Science Letters*, **269**, 575-583.
- Foote, M., Crampton, J. S., Beu, A. G. & Cooper, R. A. 2008. On the bidirectional relationship between geographic range and taxonomic duration. *Paleobiology*, **34**, 421-433.
- Cody, R. D., Levy, R. H., Harwood, D. M. & Sadler, P. M. 2008. Thinking outside the zone: High-resolution quantitative diatom biochronology for the Antarctic Neogene. *Palaeogeog., Palaeoclim., Palaeoecol.*, **260**, 92-121.
- Allen, A. P. & Savage, V. M. 2007. Setting the absolute tempo of biodiversity dynamics. *Ecology Letters*, **10**, 637-646.
- McPeck, M. 2007. The macroevolutionary consequences of ecological differences among species. *Palaeontology*, **50**, 111s129.
- Liow, L. H. & Stenseth, N. C. 2007. The rise and fall of species: implications for macroevolutionary and macroecological studies. *Proc. Royal Soc. B*, **274**, 2745-2752.
- Kucera, M. & Schönfeld, J. 2007. The origin of modern oceanic foraminiferal faunas and Neogene climate change. In: *Deep-Time Perspectives on Climate Change: Marrying the Signal from Computer Models and Biological Proxies* (ed. by M. Williams, A. M. Haywood, F. J. Gregory & D. N. Schmidt), pp. 409-425. The Geological Society, London.
- Finkel, Z. V., Sebbo, J., Feist-Burkhardt, S., Irwin, A. J., Katz, M. E., Schofield, O., Young, J. & Falkowski, P. G. 2007. A universal driver of macroevolutionary change in the size of marine phytoplankton over the Cenozoic. *Proc. Natl. Acad. Sci. USA*, **104**, 20416-20420.
- Muttoni, G. & Kent, D. V. 2007. Widespread formation of cherts during the early Eocene climatic optimum. *Palaeogeog., Palaeoclim., Palaeoecol.*, **253**, 348-362.
- Lazarus, D., Bittniok, B., Diester-Haass, L., Meyers, P. & Billups, K. 2006. Comparison of radiolarian and sedimentologic paleoproductivity proxies in the latest Miocene-Recent Benguela Upwelling System. *Marine Micropaleontology*, **60**, 269-294.
- Allen, A. P., Gillooly, J. F., Savage, V. M. & Brown, J. H. 2006. Kinetic effects of temperature on rates of genetic divergence and speciation. *Proc. Nat. Acad. Sci. USA*, **103**(24), 9130-9135.
- Allen, A. P. & Gillooly, J. F. 2006. Assessing latitudinal gradients in speciation rates and biodiversity at the global scale. *Ecology Letters*, **9**, 947-954.
- Spezzaferri, S. & Spiegler, D. 2005. Fossil planktic foraminifera (an overview). *Paläontologische Zeitschrift*, **79**, 149-166.
- Finkel, Z. V., Katz, M. E., Wright, J. D., Schofield, O. & Falkowski, P. G. 2005. Climatically driven macroevolutionary patterns in the size of marine diatoms over the Cenozoic. *Proc. Natl. Acad. Sci. USA*, **102**, 8927-8932.
- Tozzi, S., Schofield, O. & Falkowski, P. 2004. Historical climate change and ocean turbulence as selective agents for two key phytoplankton functional groups. *Marine Ecology Progress Series*, **274**, 123-132.
- Thierstein, H. R., Cortés, M. & Haidar, A. 2004. Plankton community behavior on ecological and evolutionary time-scales: when models confront evidence. In: *Coccolithophores: From Molecular Processes to Global Impact* (ed. by H. Thierstein & J. R. Young), pp. 455-480. Springer, Berlin.
- Falkowski, P. G., Katz, M. E., Knoll, A. H., Quigg, A., Raven, J. A., Schofield, O. & Taylor, F. J. R. 2004. The evolution of modern eukaryotic phytoplankton. *Science*, **305**, 354-360.
- Katz, M. E., Finkel, Z. V., Grzebyk, D., Knoll, A. & Falkowski, P. G. 2004. Evolutionary trajectories and biogeochemical impacts of marine eukaryotic phytoplankton. *Ann Rev. Ecol. Syst.*, **35**, 523-556.
- Schmidt, D. N., Thierstein, H. R. & Bollmann, J. 2004. The evolutionary history of size variation of planktic foraminiferal assemblages in the Cenozoic. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **212**, 159-180.
- Cervato, C. & Burckle, L. 2003. Pattern of first and last appearance in diatoms: Oceanic circulation and the position of polar fronts during the Cenozoic. *Paleoceanography*, **18**, web.
- Lazarus, D. B. 2002. Environmental control of diversity, evolutionary rates and taxa longevities in Antarctic Neogene radiolaria. *Paleontologica Electronica*, **5**, 1-30.
- Spencer-Cervato, C. 1999. The Cenozoic deep sea microfossil record: explorations of the DSDP/ODP sample set using the Neptune database. *Palaeontologica Electronica*, **2**, web.
- Lazarus, D. B., Spencer-Cervato, C., Pianka-Biolzi, M., Beckmann, J. P., von Salis, K., Hilbrecht, H. & Thierstein, H. R. 1995. *Revised chronology of Neogene DSDP holes from the world ocean. Ocean Drilling Program*: College Station. ca. 250 p.
- Lazarus, D. B. 1994. The Neptune Project - a marine micropaleontology database. *Math. Geol.*, **26**, 817-832.
- Spencer-Cervato, C., Thierstein, H. R., Lazarus, D. B. & Beckmann, J. P. 1994. How synchronous are Neogene marine plankton events? *Paleoceanography*, **9**, 739-763.
- Spencer-Cervato, C., Lazarus, D. B., Beckmann, J. P., Perch-Nielsen, K. v. S. & Biolzi, M. 1993. New calibration of Neogene radiolarian events in the North Pacific. *Marine Micropaleontology*, **21**, 261-294.



# NSB Features & Plans

- Long-term support via in-house staff (me)
- Gradual improvement in data quality; coverage; retrieval/analytic functions
- New external funding proposals planned:
  - Speed-up improvement of database
  - Create new website for managed access



# Integrated Ocean Drilling Program (IODP) PCG & TNLs

- Multiple institutions, platforms in IODP; continued problems with paleontology data management led to 2006 Houston meeting on IODP paleontology data
- Paleontology Coordination Group (PCG) formed, first meeting 2007 Berlin.
  - *Advisory function, external scientist (Lazarus) and IODP staff as co-chairs*
- Goals - improve paleontologic data entry, data management, data quality within all aspects of IODP



# Taxonomic Name Lists (TNLs)

- First priority for PCG : creation of scientifically useful taxonomic name lists for IODP databases, data-entry applications, etc.
  - *Name, status (valid, status/synon. to, author, year)*
- Current IODP taxonomy data undocumented mixture of valid names, synonyms, open-nomenclature, misspellings, etc without any quality control
- Community editors/teams found for Planktonic foraminifera, calcareous nannofossils, diatoms and radiolarians, dinoflagellates, contracts signed fall 2008; final lists now -ready





# TNL Statistics

*(initial distribution files)*

	<i>N names*</i>	<i>Main Worker(s)</i>
<i>Diatom</i>	2835	Iwai, Suto, Akiba, Harwood
<i>Radiolarian</i>	3691	Lazarus, Suzuki
<i>Plank. Foram.</i>	3933	Huber, Hooks
<i>Calc. Nanno.</i>	4235	Wise, Young, Lees
<i>Dinoflagellates</i>	1975	Feist-Burkhardt
<i>Totals</i>	16,669	10 taxonomists + Diver

*\*ODP+Neptune combined, raw*



# Radiolarian TNL, 'final'

	<i>Original</i>	<i>Final</i>	<i>N change</i>	<i>%change</i>
<i>Total Records</i>	17155	17348	193	<b>1.13</b>
<i>Status records</i>	3416	4260	844	<b>24.71</b>
<i>Valid species</i>	1164	1639	475	<b>40.81</b>
Recent		421		
Neogene		393		
Paleogene		469		
Mesozoic		356		
<i>Synonyms</i>	643	975	332	<b>51.63</b>
<i>Generic</i>	466	821	355	<b>76.18</b>
<i>Unknown</i>	740	345	-395	<b>-53.38</b>
<i>Questionable</i>	388	417	29	<b>7.47</b>
<i>Subspecies*</i>	0	24	24	<b>(-)</b>
<i>Incorrect group</i>	2	39	37	<b>1850</b>



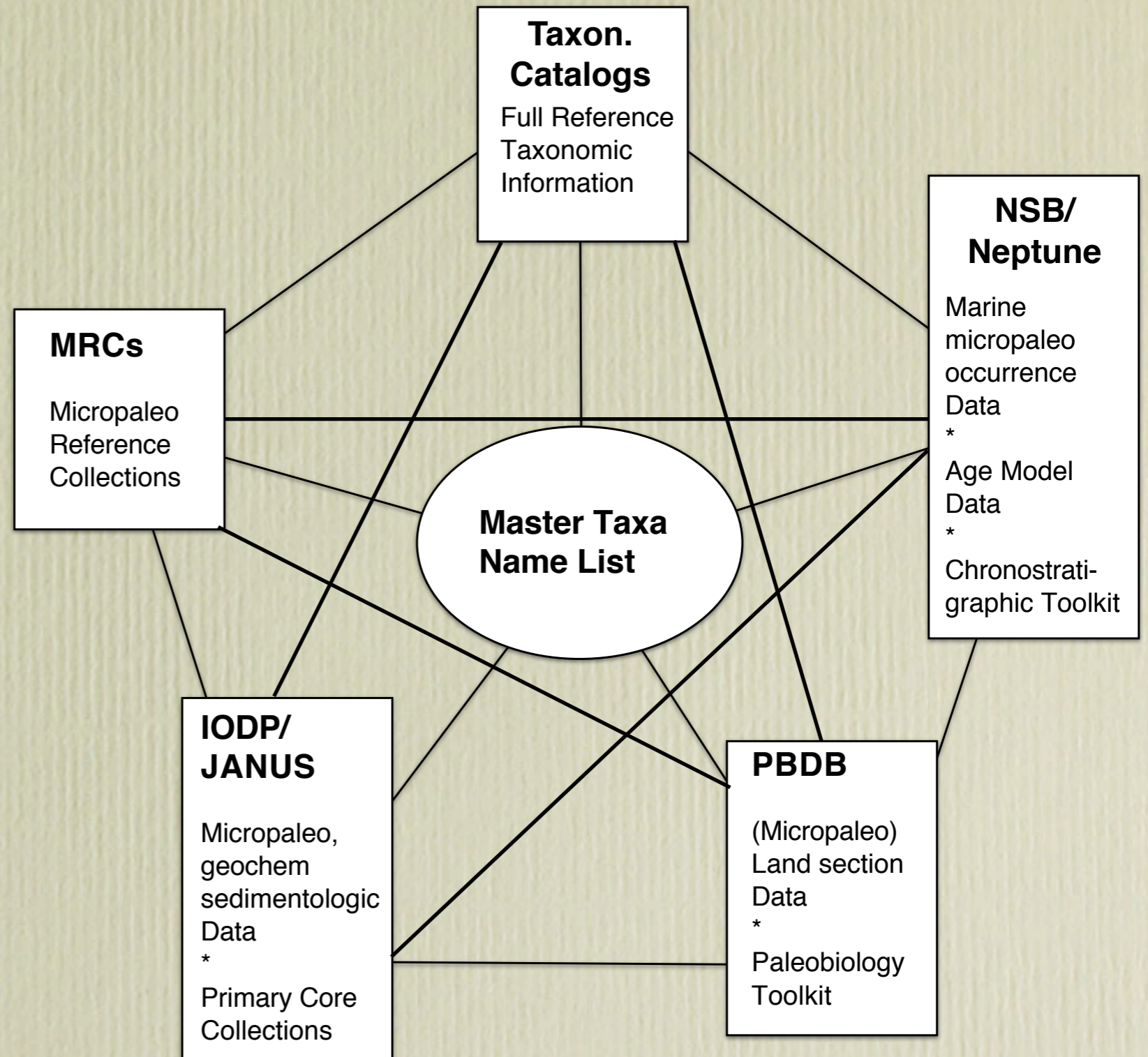
# Radiolarian TNL, 'final'

	<i>Original</i>	<i>Final</i>	<i>N change</i>	<i>%change</i>
<i>Total Records</i>	17155	17348	193	<b>1.13</b>
<i>Status records</i>	3416	4260	844	<b>24.71</b>
<i>Valid species</i>	1164	1639	475	<b>40.81</b>
Recent		421		
Neogene		393		
Paleogene		469		
Mesozoic		356		
<i>Synonyms</i>	643	975	332	<b>51.63</b>
<i>Generic</i>	466	821	355	<b>76.18</b>
<i>Unknown</i>	740	345	-395	<b>-53.38</b>
<i>Questionable</i>	388	417	29	<b>7.47</b>
<i>Subspecies*</i>	0	24	24	<b>(-)</b>
<i>Incorrect group</i>	2	39	37	<b>1850</b>



# TNL as Micropaleontology Name Server

- Federated system for managing deep-sea micropaleontology data
- All individual dbs exist, a few are provisionally linked
- Central TNL db manages 'key field' for meaningful data exchange
- Central TNL db development recently funded by IODP, now in development





# 4. Problems and Prospects

- Global taxonomic research data management as a core institutional responsibility
  - Budget to support work (IT, data entry, etc)
  - Appropriate recognition in evaluations
- Training (new) taxonomic biologists with needed skills
- Data Archive Centers need enhancements in taxonomic data management
- Better coordination and oversight of current piece-meal efforts



Thanks for Listening!