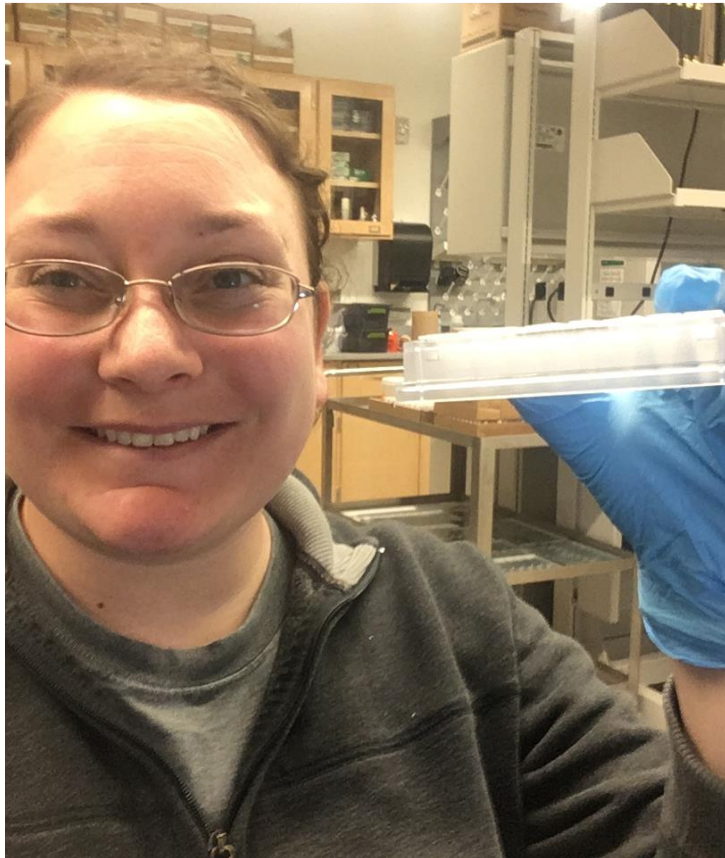


# Challenges to DNA barcoding: *an ecologist's perspective*



**Susan Daniel, Lyubov Burlakova,  
and Alexander Karatayev:** *Great  
Lakes Center, SUNY Buffalo State*

**Isabel Porto-Hannes:** *University of  
Buffalo*

**Paul Hebert:** *University of Guelph*

**Michael E. Pfrender:** *University of  
Notre Dame*

**David Lodge:** *Cornell University*

**Lars Rudstam:** *Cornell University*

**Anett Trebitz:** *U.S. EPA Office of  
Research & Development*

**Sara Westergaard:** *U.S.  
Environmental Protection Agency*

# The Potential of DNA Barcoding



**DNA Barcoding is the use of a short DNA fragment as a genetic tag for species identification**



Early detection of species

Including invasive species, nuisance species, rare taxa, etc.



Increase the scope of biological surveillance

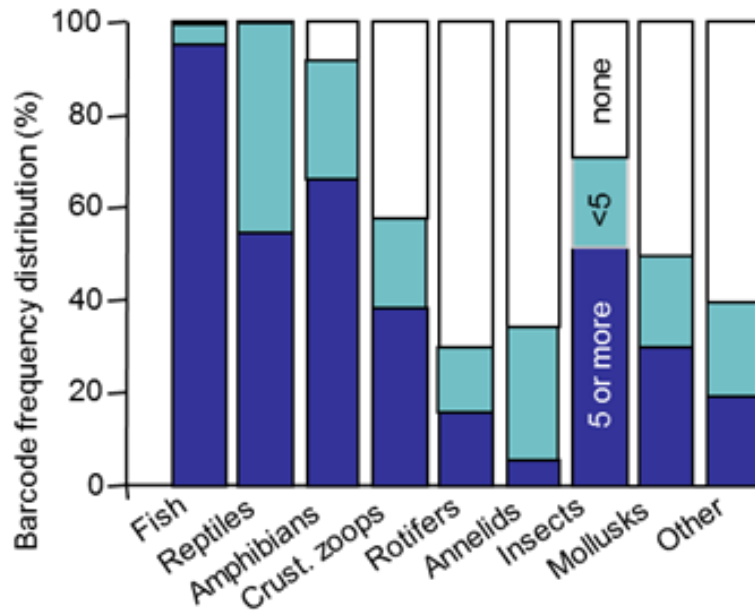


Allows taxonomists to focus reference specimens for establishing reliable databases versus routine identification (Valentini, Pompanon, and Taberlet 2009)



Helps distinguish species that are: (Decaëns et al. 2013)

Morphologically similar  
High in phenotypic variability  
Ambiguous juvenile stages  
Damaged



Trebitz et al. 2015 Sci. Rep. 5:12162.

# Project Goals



Generate voucher collection with high quality photos



Clear up common taxonomic questions



Identify locations with cryptic species to describe new species (Metabarcoding)



Fill in the gaps of missing taxa in genetic barcode reference libraries (Genetic Barcoding)

*Main gene: Cytochrome oxidase I, COI*  
*Others: 16S, ND1*

# Collaboration



Dr. Valerie Brady (University of Minnesota) shared ~400 specimens of Mollusca and Annelida (Great Lakes Coastal Wetland). In return, we shared over 1000 specimens



Graduate student Adam Frankiewicz (U Minnesota Duluth) collected and sent for genetic analysis 16 of the missing 17 species from the family Pisiidae.

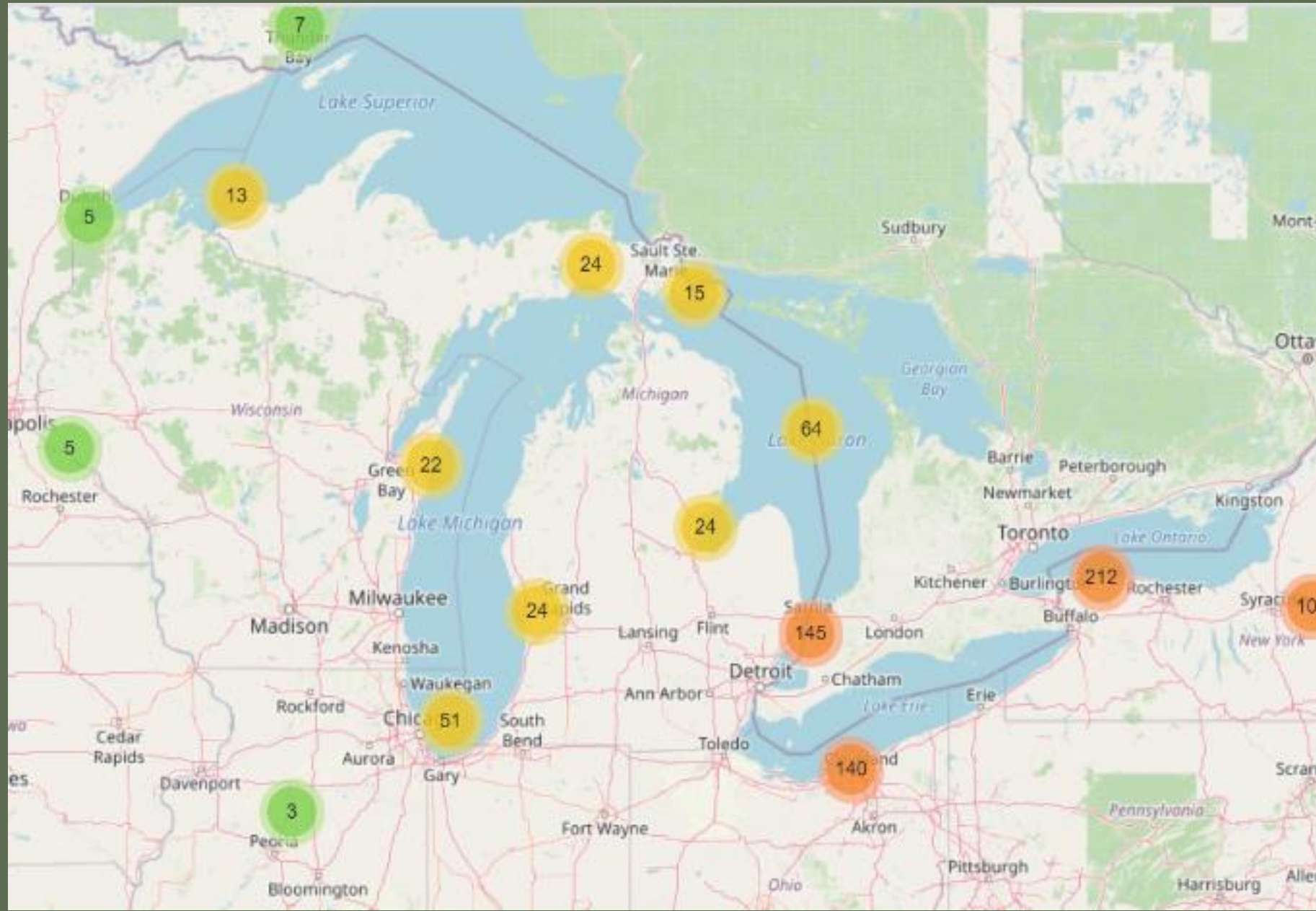


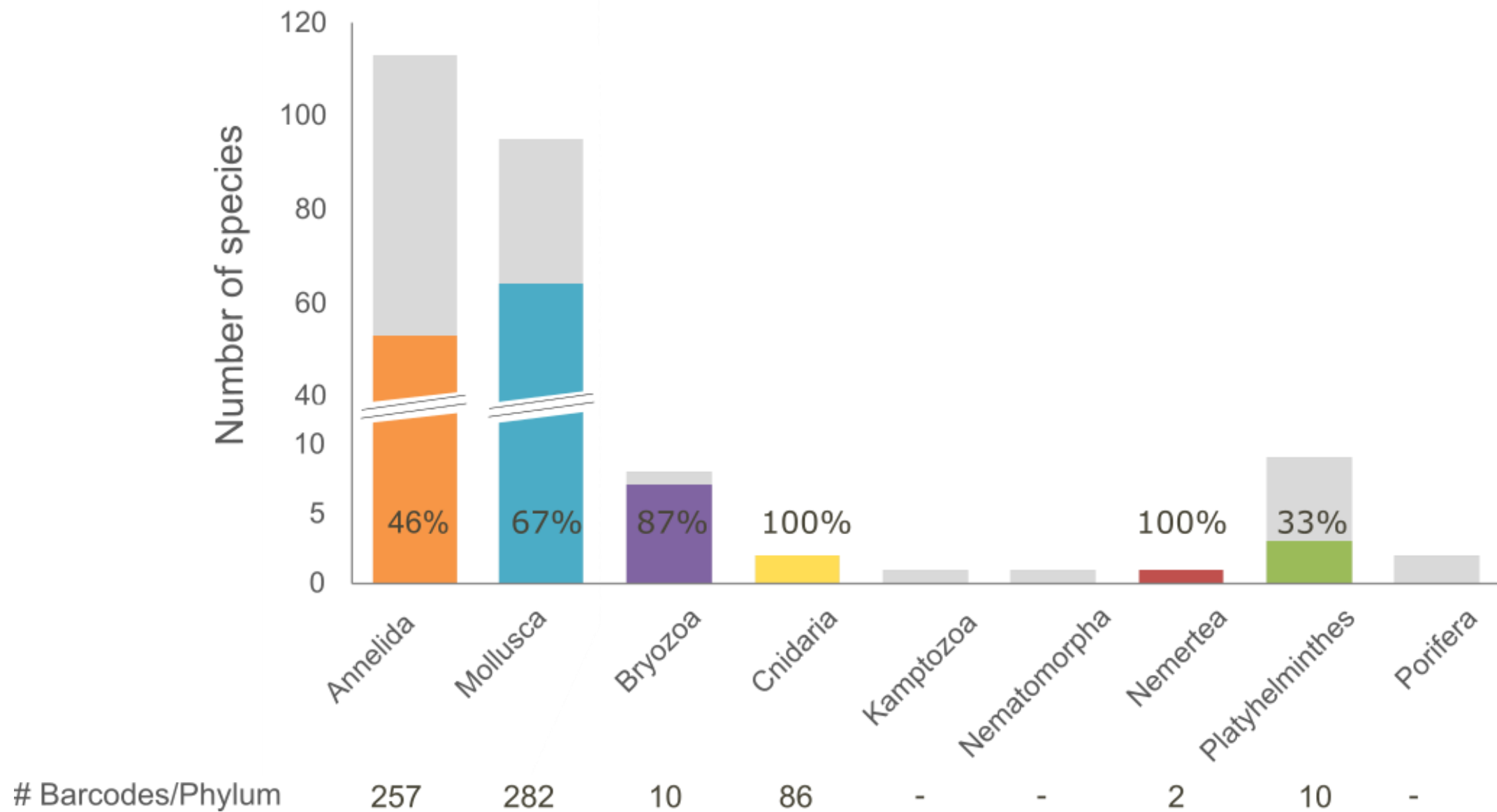
Cornell teams (Lindsay Schaffner, James Watkins, Chris Marshal, Beth Whitmore, Joe Connolly and Lars Rudstam) shared specimens



Mandi Ohar from U.S. FWS shared Annelida from over two years of river/embayment collections







Bold Systems v4

boldsystems.org

ABP

Paused

SUNY Buffalo StateGreat Lakes Center...E. H. Butler Library...Google ScholarMacro Photograph...BingSusie Databases En...Google HangoutsCSMI Erie 2019 Pro...

BOLDSYSTEMS

DATABASESIDENTIFICATIONTAXONOMYWORKBENCHRESOURCESLOGIN

BARCODE OF LIFE DATA SYSTEM v4

Advancing biodiversity science through DNA-based species identification.

EXPLORE THE DATA

DESIGNED TO SUPPORT THE GENERATION & APPLICATION OF DNA BARCODE DATA

BOLD is a cloud-based data storage and analysis platform developed at the Centre for Biodiversity Genomics in Canada. It consists of four main modules, a data portal, an educational portal, a registry of BINs (putative species), and a data collection and analysis workbench.

**Level 1  
Most Stringent**

- Collected within the Great Lakes basin.
- COI seq 500 bp long, species name, voucher data, collection record, identifier, primers, trace files
- Barcodes are from taxa on the EPA Water Life list
- Identified to species or genus
- Identified by morphological examination
- Vouchered or include photographs showing diagnostic features
- Barcodes are publicly available

**Level 2  
Relaxed Stringency**

- Same as Level I, except:
- Can be from outside the Great Lakes watershed

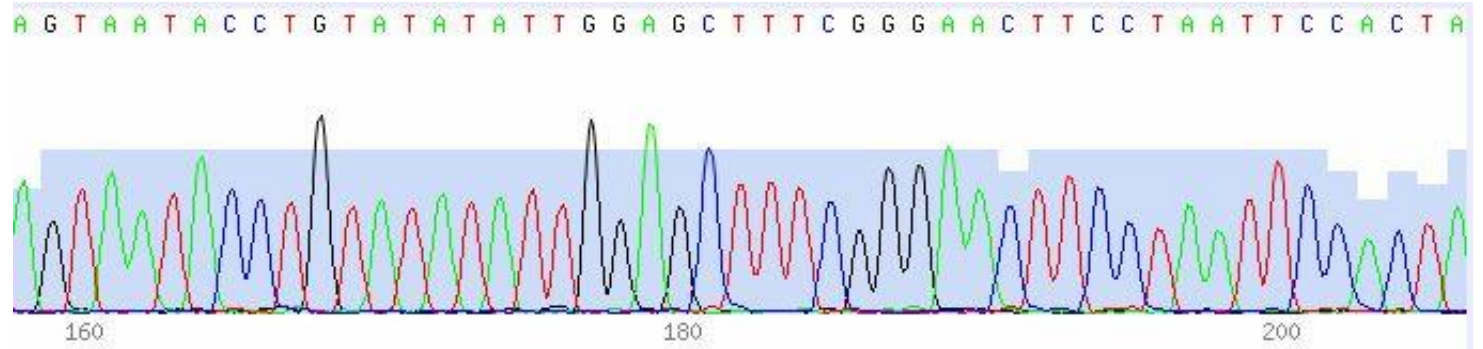
**Level 3  
Lowest Stringency**

- Voucher data/picture optional
- Collection record (at least collection date and location with GPS coordinates)
- Identifier of the specimen optional
- Primer information optional
- Trace file(s) optional

From Great Lakes DNA Barcode Database  
*Daniel et al. in preparation*

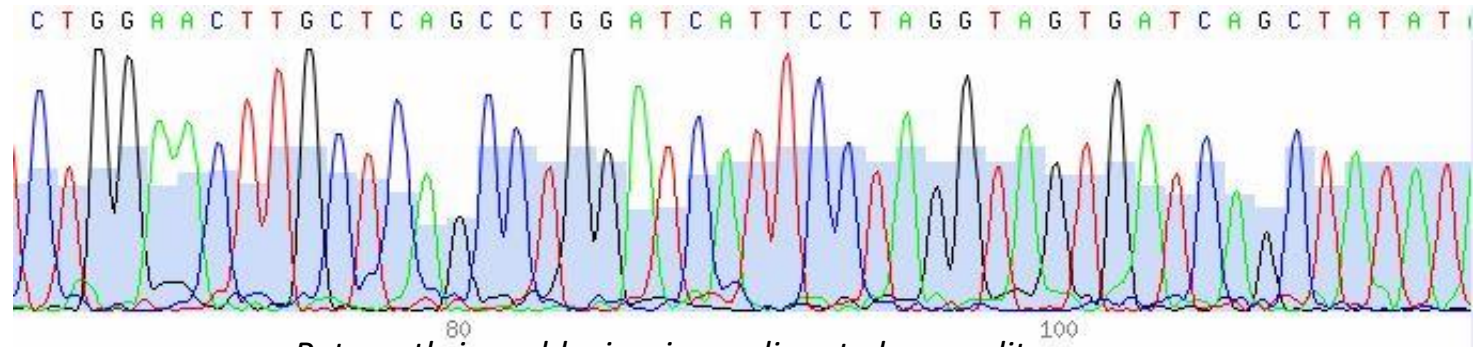


# The Good



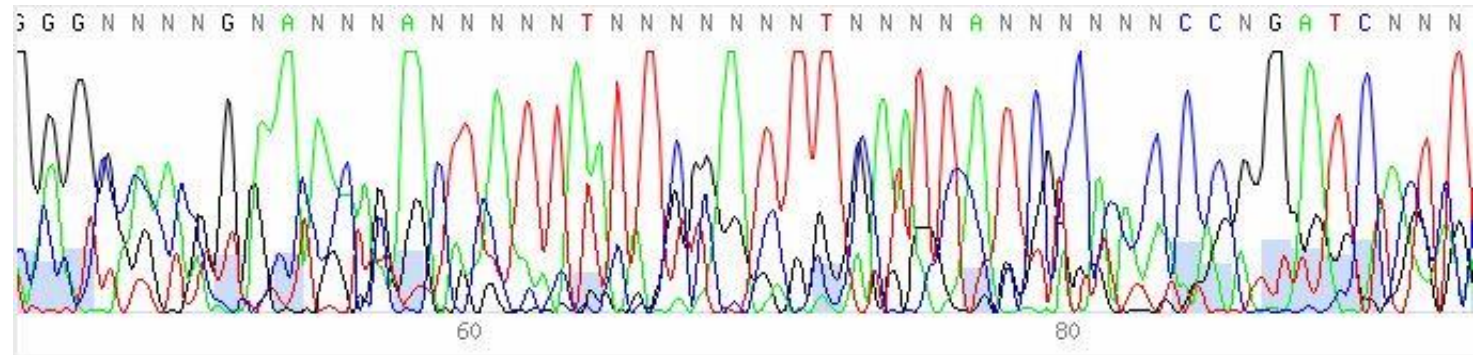
*Potamothenrix moldaviensis*, high quality sequence

# The Bad



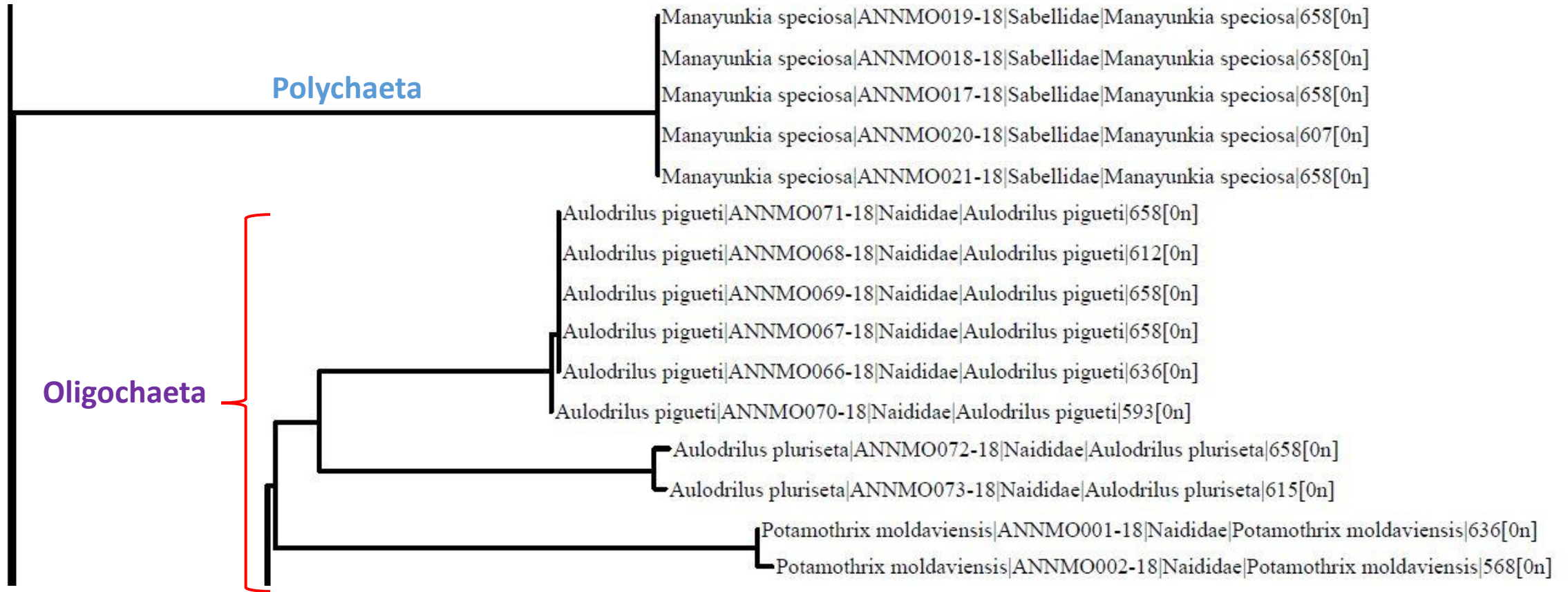
*Potamothenrix moldaviensis*, medium to low quality sequence

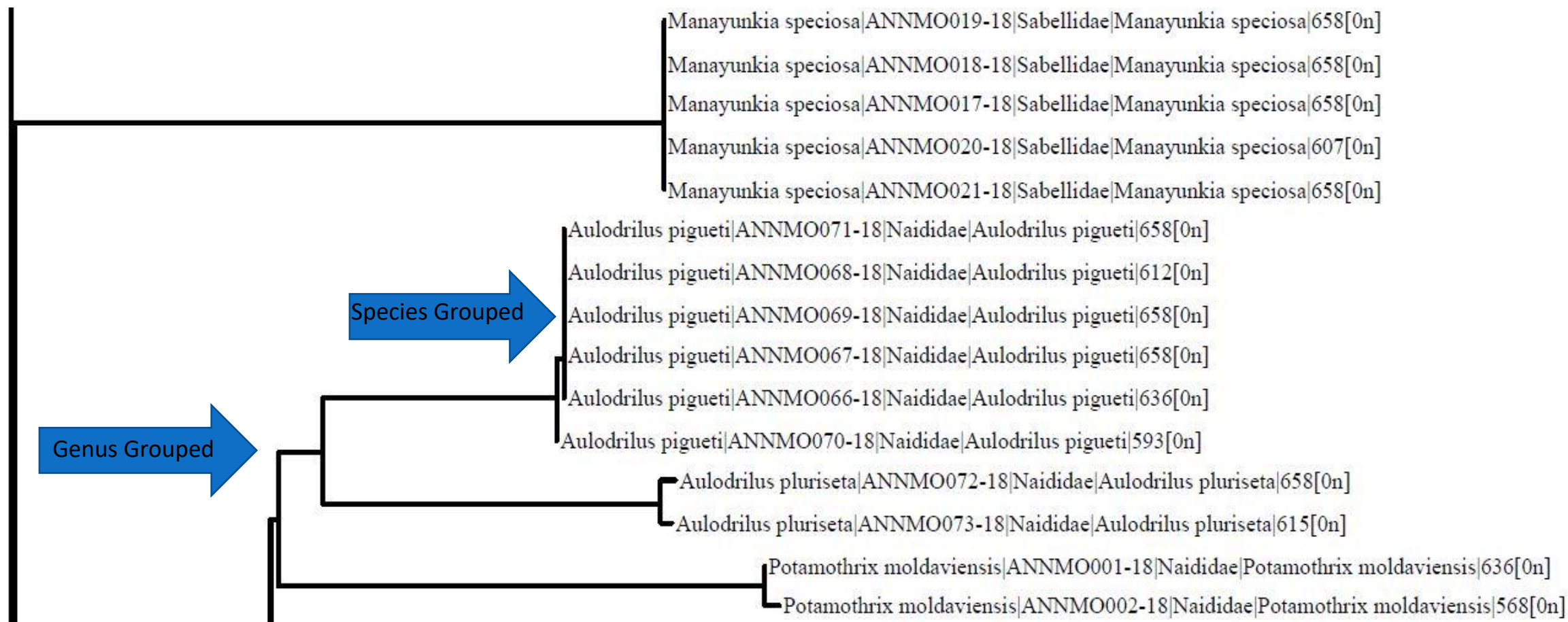
# The Ugly



*Limnodrilus claparedianus*, co-amplification, failed quality

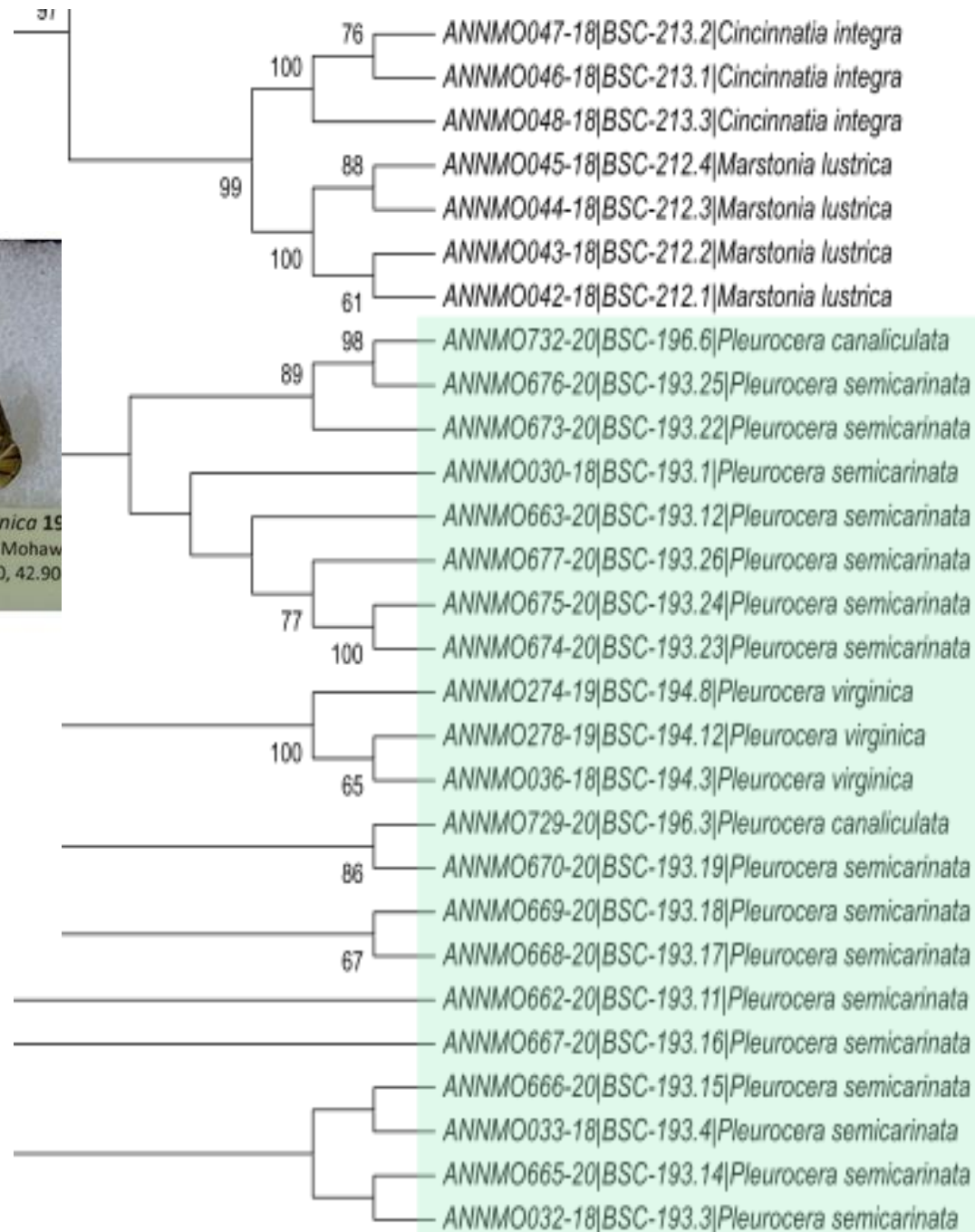
# “The Good”

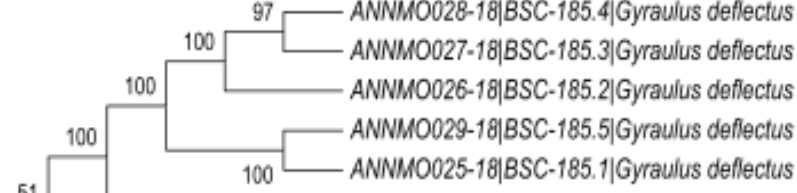






“The Bad”





KUNGL.  
VETENSKAPS-  
AKADEMIEN  
THE ROYAL SWEDISH ACADEMY OF SCIENCES



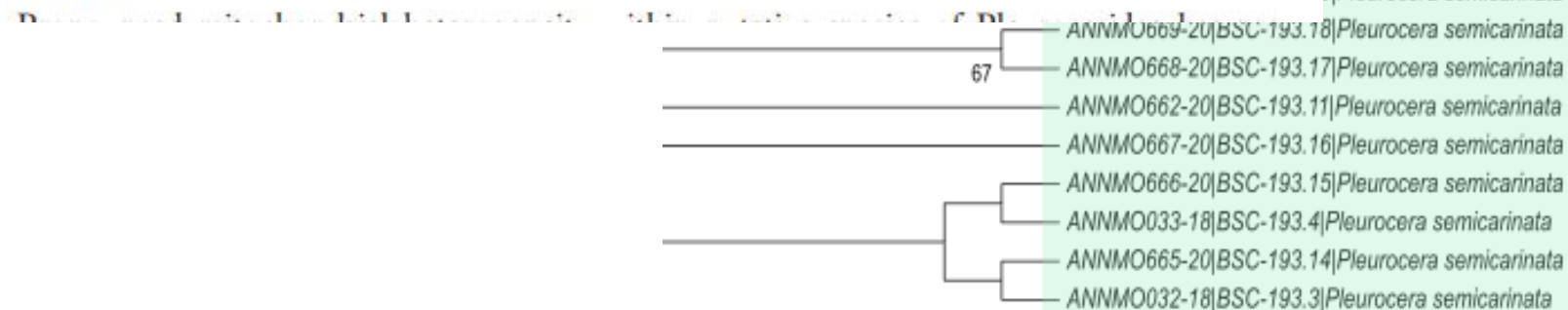
*Laevapex fuscus*  
*Probythinella emarginata*  
*Probythinella emarginata*  
*Cincinnatia integra*  
*Cincinnatia integra*  
*Cincinnatia integra*  
*Marstonia lustrica*  
*Marstonia lustrica*  
*Marstonia lustrica*  
*Marstonia lustrica*  
*Pleurocera canaliculata*  
5|*Pleurocera semicarinata*  
2|*Pleurocera semicarinata*  
*Pleurocera semicarinata*  
2|*Pleurocera semicarinata*  
5|*Pleurocera semicarinata*  
4|*Pleurocera semicarinata*  
3|*Pleurocera semicarinata*  
*Pleurocera virginica*  
2|*Pleurocera virginica*  
*Pleurocera virginica*  
*Pleurocera canaliculata*  
9|*Pleurocera semicarinata*

## Morphology, molecules and taxonomy: extreme incongruence in pleurocerids (Gastropoda, Cerithioidea, Pleuroceridae)

NATHAN V. WHELAN & ELLEN E. STRONG

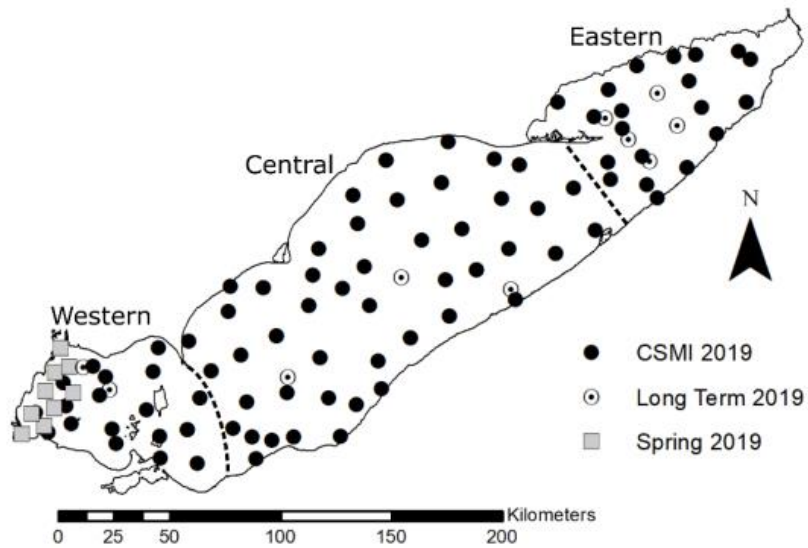
Submitted: 2 April 2015  
Accepted: 18 July 2015  
doi:10.1111/zsc.12139

Whelan, N.V., Strong, E.E. (2015). Morphology, molecules and taxonomy: extreme incongruence in pleurocerids (Gastropoda, Cerithioidea, Pleuroceridae). —*Zoologica Scripta*, 00, 000–000.

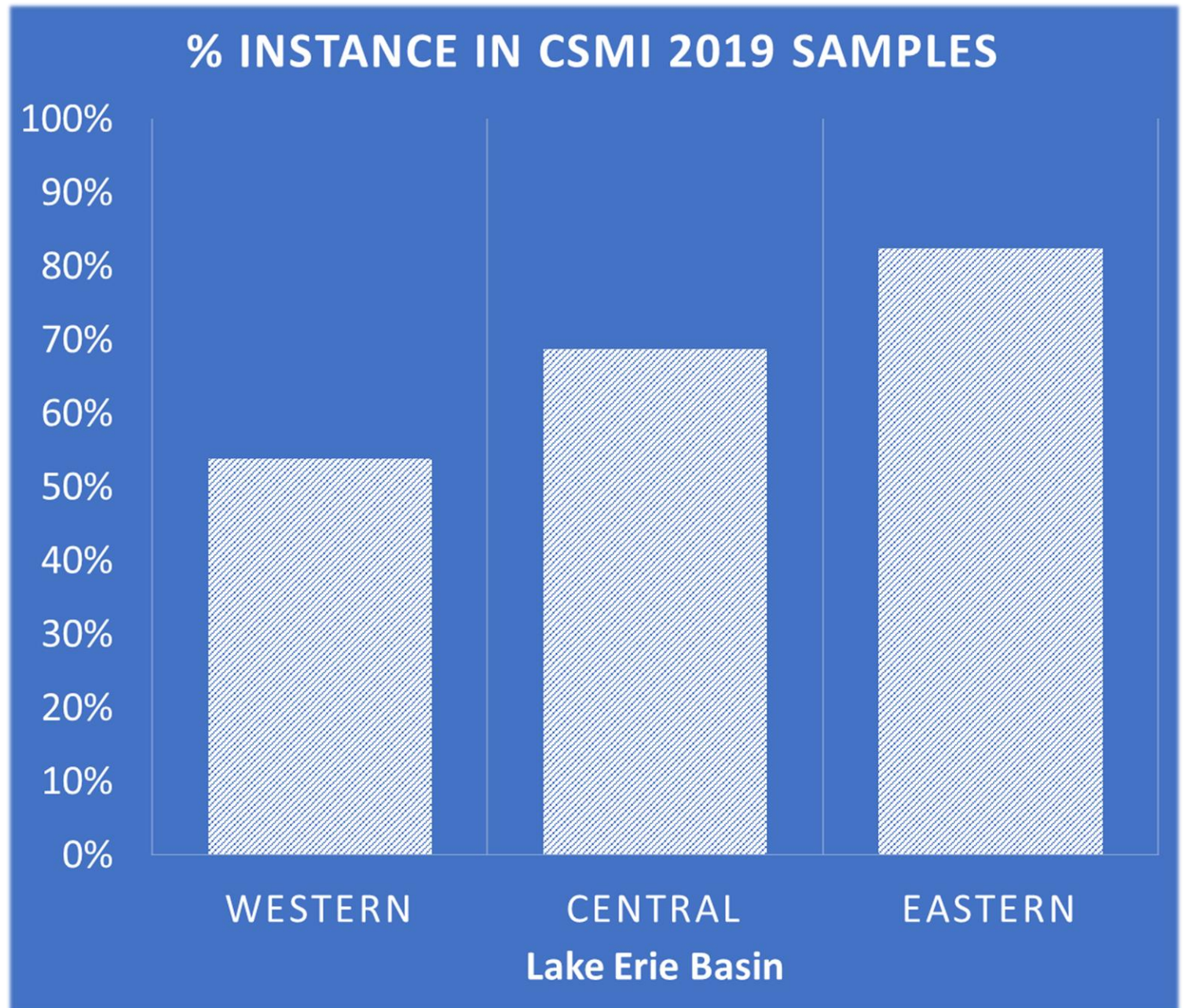




## The case of *Limnodrilus hoffmeisteri*

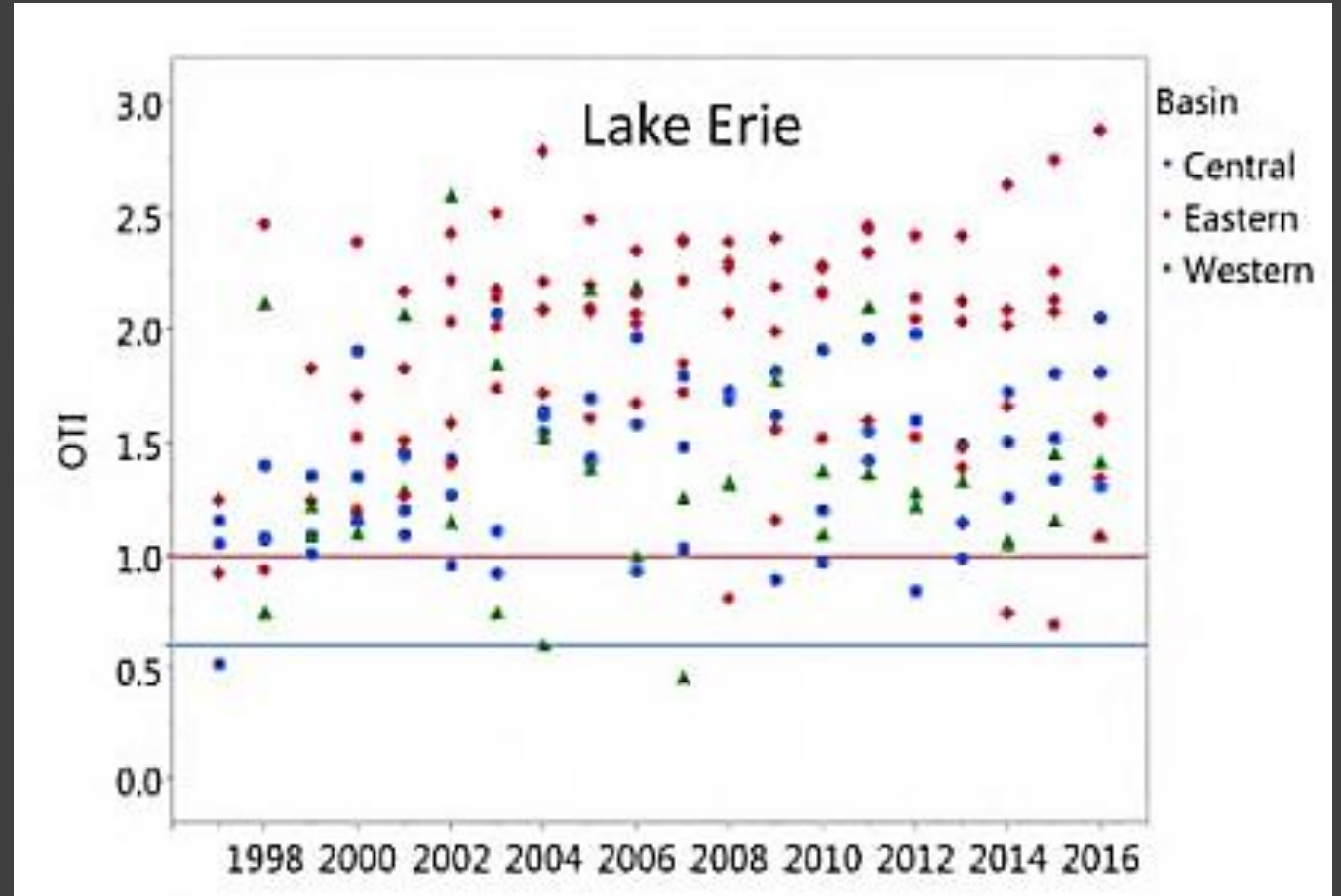


77 total stations,  
Western 26  
Central 32  
Eastern 19



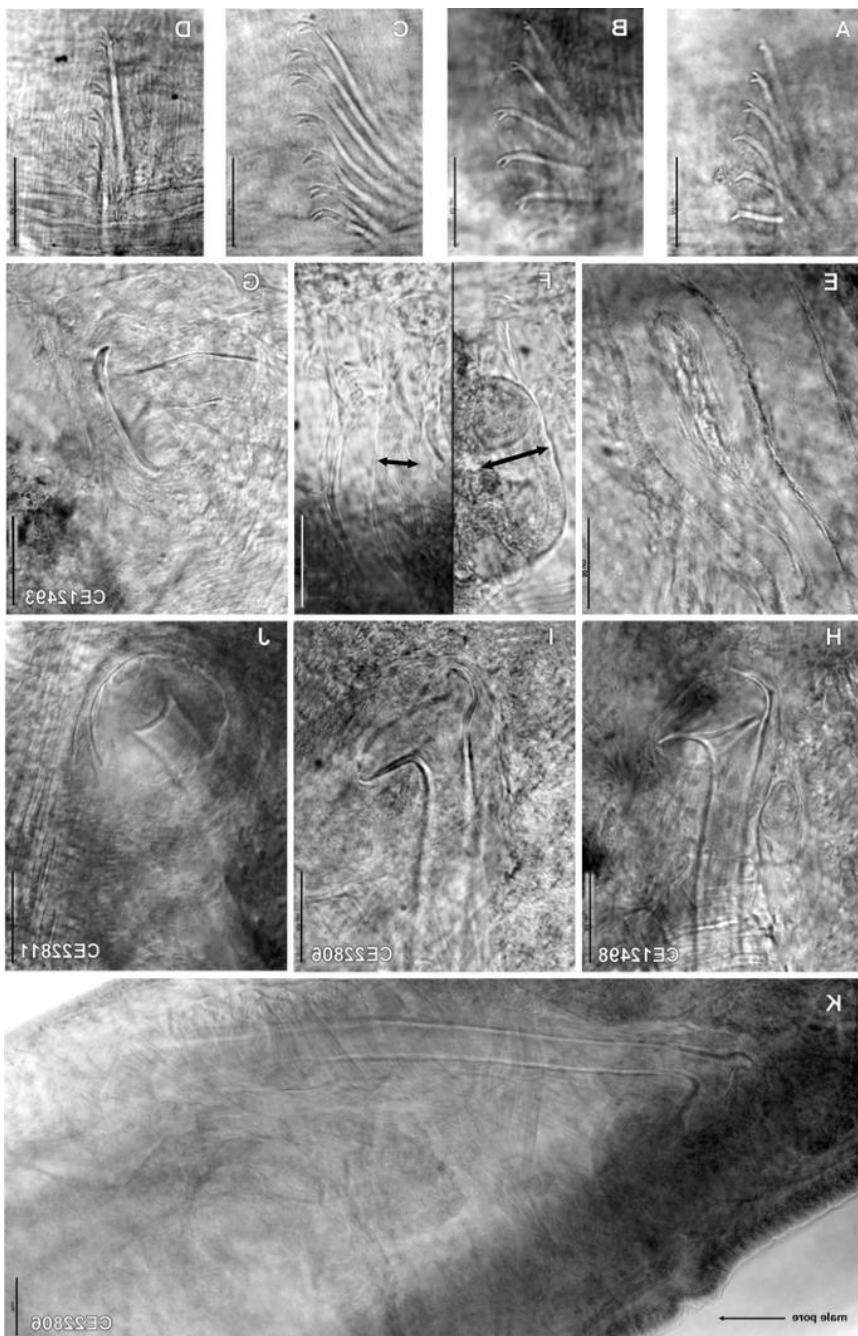
# Oligochaeta Tropic Index (Milbrink, 1983)

*Limnodrilus hoffmeisteri* is the highest tropic rank aside from *Tubifex tubifex* whose rank is determined by density of *L. hoffmeisteri*

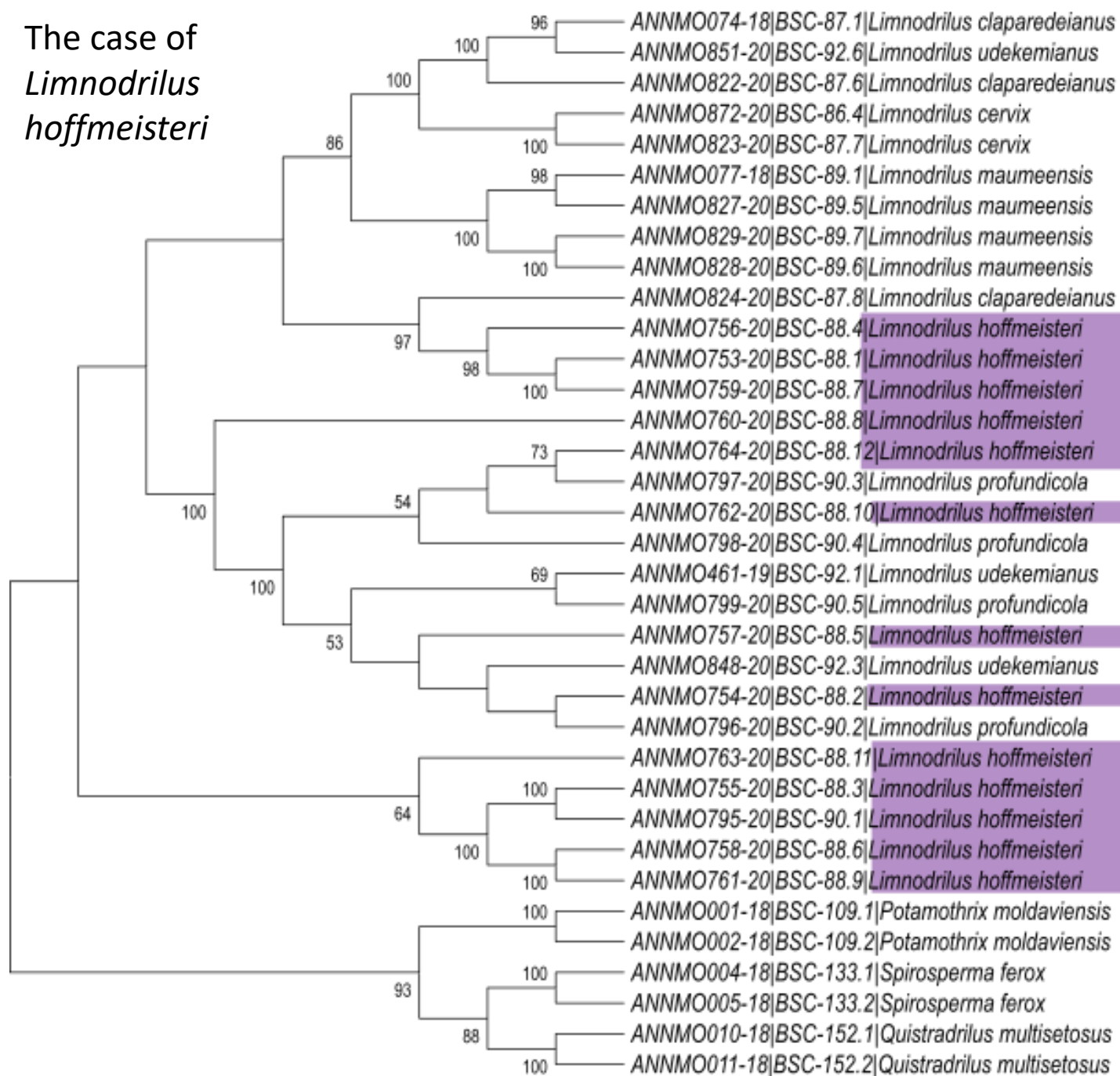


State of the Great Lakes 2019 Technical Report





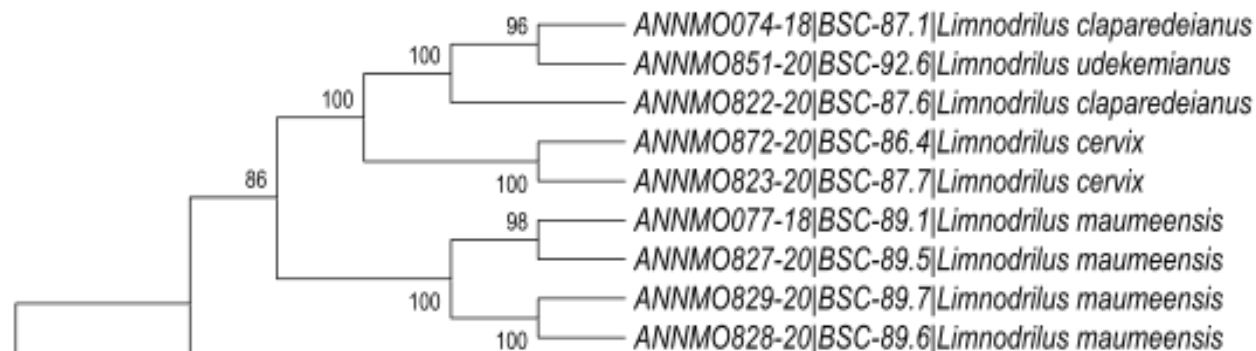
## The case of *Limnodrilus* *hoffmeisteri*



*Limnodrilus hoffmeisteri*

Org Divers Evol (2017) 17:477–495

DOI 10.1007/s13127-016-0317-z




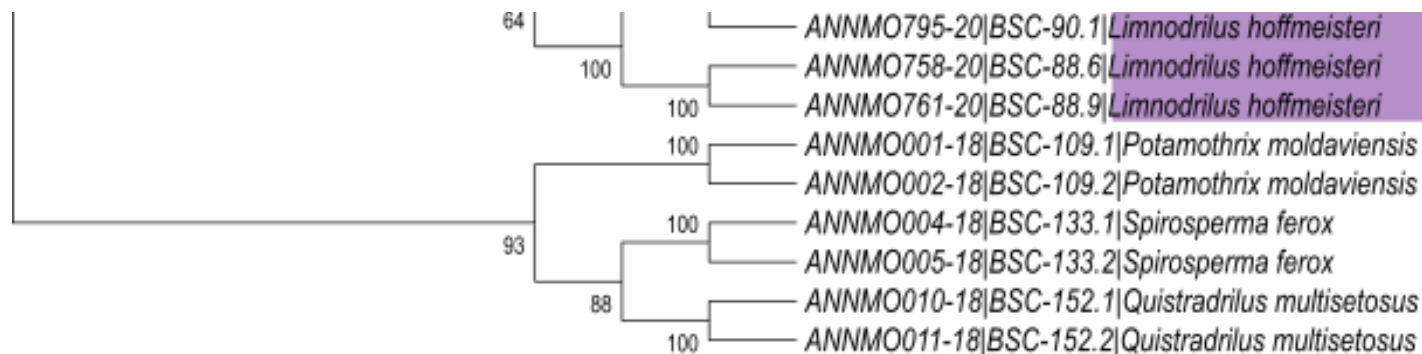
CrossMark

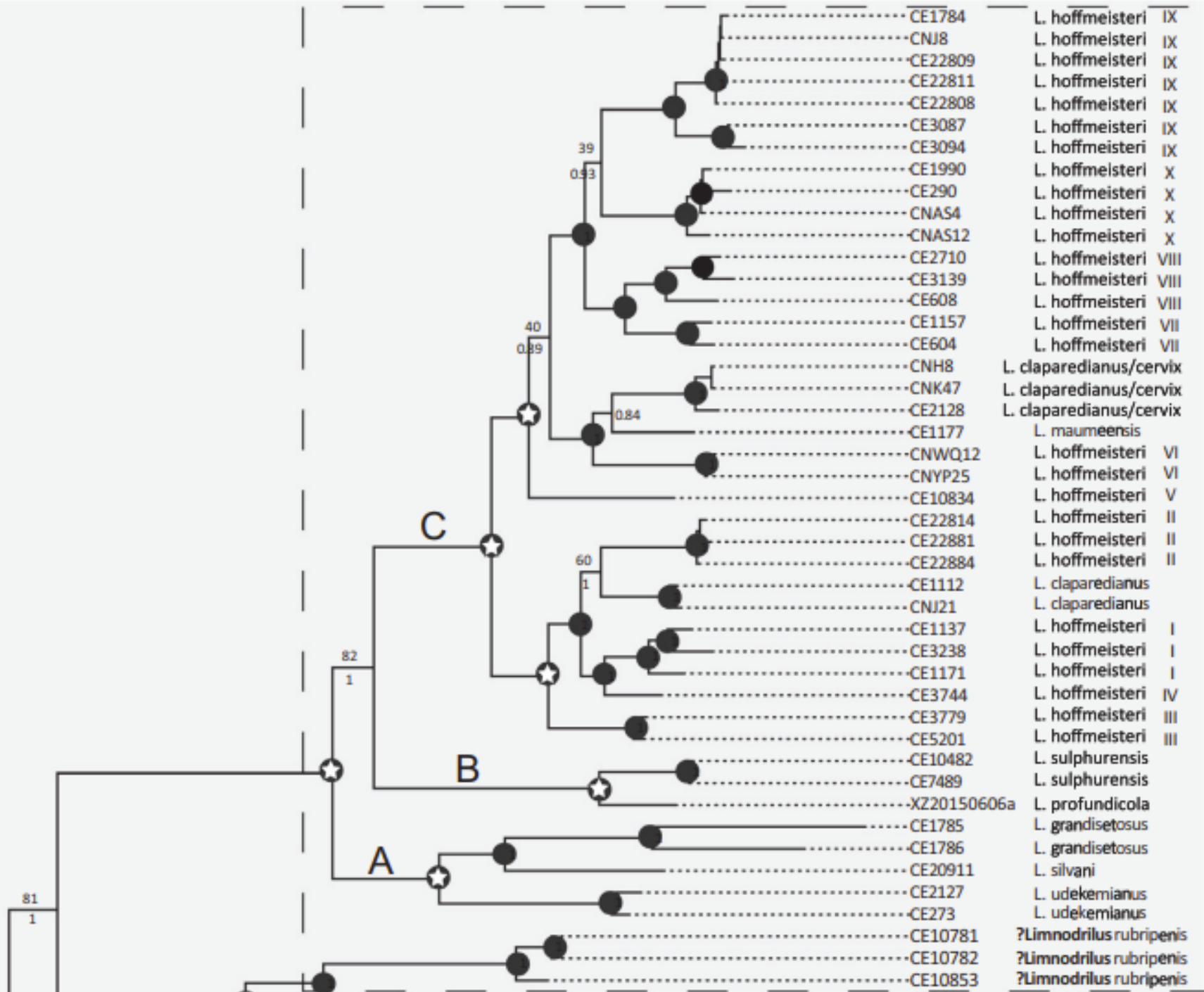
ORIGINAL ARTICLE

**ORGANISMS  
DIVERSITY &  
EVOLUTION**

# Extensive cryptic diversity in the cosmopolitan sludge worm *Limnodrilus hoffmeisteri* (Clitellata, Naididae)

Yingkui Liu<sup>1</sup> · Steven V. Fend<sup>2</sup> · Svante Martinsson<sup>1</sup> · Christer Erséus<sup>1</sup> 








Tubificinae

*Liu et al., 2017*




### Seven loci

Cytochrome oxidase I,  
COI, 12S, 16S rDNA,  
Histone-3 (H3), Internal  
Transcribed Spacer  
region (ITS), 18S, and  
partial 28S rDNA






Order	Family	Species	Mean Intra-Sp 	Max Intra-Sp 	Nearest Neighbour	Nearest Species	Distance to NN 
Haplotaxida	Naididae	Limnodrilus cervix	2.89	5.82	ANNMO851-20	Limnodrilus udekemianus	0
Haplotaxida	Naididae	Limnodrilus claparedianus	N/A	0	ANNMO756-20	Limnodrilus hoffmeisteri	17.45
Haplotaxida	Naididae	Limnodrilus hoffmeisteri	19.82	51.18	ANNMO848-20	Limnodrilus udekemianus	0
Haplotaxida	Naididae	Limnodrilus maumeensis	1.21	2.59	ANNMO870-20	Limnodrilus cervix	14.8
Haplotaxida	Naididae	Limnodrilus profundicola	0.28	0.48	ANNMO762-20	Limnodrilus hoffmeisteri	0
Haplotaxida	Naididae	Limnodrilus udekemianus	18.73	23.29	ANNMO754-20	Limnodrilus hoffmeisteri	0

- n=49 specimens
- Only 4 additional records from outside GLC available
- Method: CO1 gene, BOLD Aligner Amino Acid based alignment, Kimura 2 Parameter

Order	Family	Species	Mean Intra-Sp 	Max Intra-Sp 	Nearest Neighbour	Nearest Species	Distance to NN 
Haplotaxida	Naididae	Limnodrilus cervix	2.89	5.82	ANNMO851-20	Limnodrilus udekemianus	0
Haplotaxida	Naididae	Limnodrilus claparedianus	N/A	0	ANNMO756-20	Limnodrilus hoffmeisteri	17.45
Haplotaxida	Naididae	Limnodrilus hoffmeisteri	19.82	51.18	ANNMO848-20	Limnodrilus udekemianus	0
Haplotaxida	Naididae	Limnodrilus maumeensis	1.21	2.59	ANNMO870-20	Limnodrilus cervix	14.8
Haplotaxida	Naididae	Limnodrilus profundicola	0.28	0.48	ANNMO762-20	Limnodrilus hoffmeisteri	0
Haplotaxida	Naididae	Limnodrilus udekemianus	18.73	23.29	ANNMO754-20	Limnodrilus hoffmeisteri	0

- n=49 specimens
- Only 4 additional records from outside GLC available
- Method: CO1 gene, BOLD Aligner Amino Acid based alignment, Kimura 2 Parameter

Order	Family	Species	Mean Intra-Sp 	Max Intra-Sp 	Nearest Neighbour	Nearest Species	Distance to NN 
Haplotaxida	Naididae	Aulodrilus pluriseta	0.27	1.4	ANNMO070-18	Aulodrilus pigueti	17.77
Haplotaxida	Naididae	Branchiura sowerbyi	0	0	ANNMO919-20	Aulodrilus pigueti	20.91
Haplotaxida	Naididae	Potamothrix moldaviensis	1.65	3.26	ANNMO893-20	Potamothrix vej dovskyi	20.11
Haplotaxida	Naididae	Potamothrix vej dovskyi	0	0	ANNMO001-18	Potamothrix moldaviensis	20.11
Haplotaxida	Naididae	Quistradrilus multisetosus	0	0	ANNMO909-20	Aulodrilus limnobius	20.2
Haplotaxida	Naididae	Spirosperma ferox	0.07	0.16	ANNMO909-20	Aulodrilus limnobius	20.94
Haplotaxida	Naididae	Spirosperma nikolskyi	0.12	0.5	ANNMO006-18	Spirosperma ferox	22.97
Sabellida	Sabellidae	Manayunkia speciosa	0	0	ANNMO003-18	Potamothrix moldaviensis	37.2
Unionida	Unionidae	Simpsonaias ambigua	0	0	ANNMO020-18	Manayunkia speciosa	43.31

N=73 specimens

Some  
uncomfortable  
questions  
arise

---

Is the CO1 gene good enough for some of these groups (family or genus level)?

---

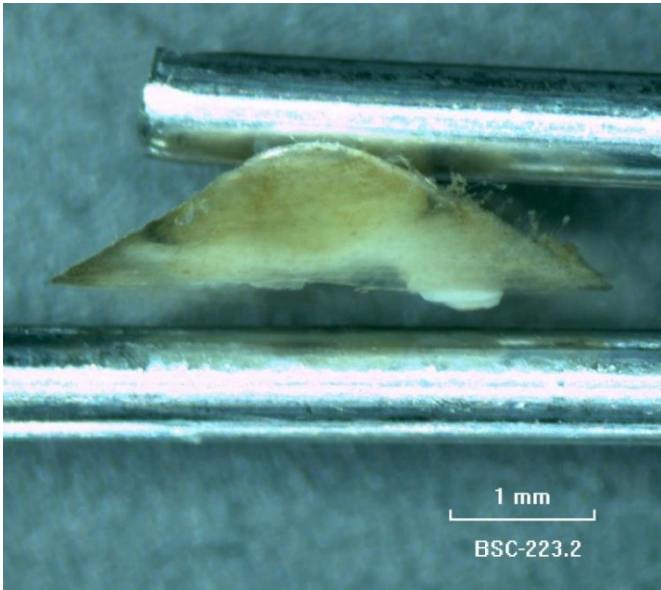
Is it an issue of sequence quality, quantity, or identification?

---

Are we just not ready to ask these questions yet?

---

How do we move past this and resolve some of these questions without a massive sampling effort?



# Specimen $\neq$ Sequence

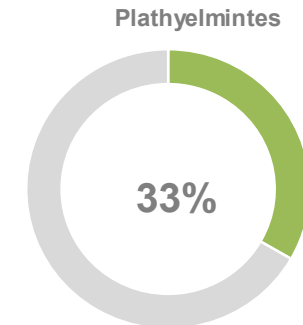
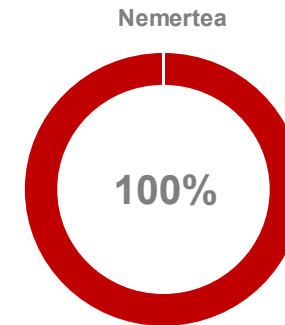
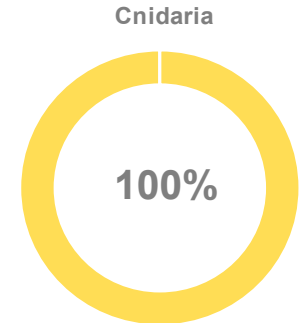
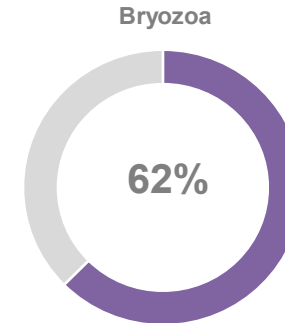
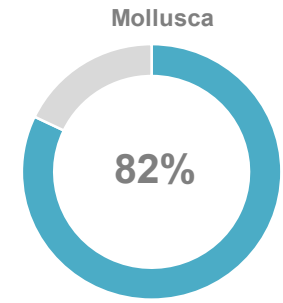
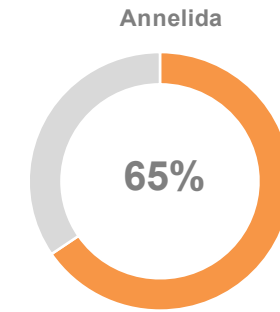
## *and other challenges*

- Low sequencing success
- Insufficient primers for many less studied groups
- Working around the operculum
- Extracting enough usable tissue
- Immature oligochaete
- Licensing information for photos
- Props to orient specimens properly... a paperclip can do wonders!



# What does this mean for *Lake Erie and Monitoring?*

- eDNA is not the answer for some biological surveys
  - Building of reference libraries
  - One gene region seems to be inadequate for important taxa
- Our current methods seem to be missing some of the diversity due to cryptic species or immature specimens.
- Because of this, we need to continue our efforts!
- Though, we have come far, Annelida only had 3% species at the start of this grant.



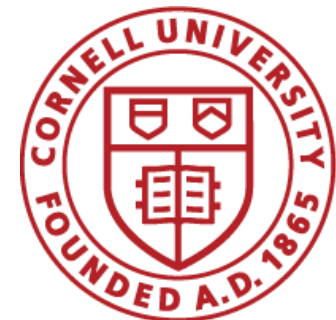
Updated August  
2021 from all  
available records  
on BOLD

# Acknowledgements

- Funding: **U.S. EPA, project "DNA Barcode Reference Library: Mollusca, Annelida, and minor phyla"** Great Lakes National Program Office (GLNPO)
- US EPA Research Vessel Lake Guardian-Captain and crew
- All my authors and collaborators on this project
  - Jakob Boehler, Ron Griffiths, Jim Watkins, Kristy Deiner, Paul Simonin, Lindsay Schaffner, etc. there are so many!
- GLC staff (Kit Hastings, Erik Hartnett, Susan Dickinson, and Brianne Tulumello) and students
- Valerie Brady, Adam Frankiewicz and their collaborators
- Mandi Ohar, U.S. FWS



MVP



# Works Cited

- Decaëns, T., Porco, D., Rougerie, R., Brown, G.G. and James, S.W., 2013. Potential of DNA barcoding for earthworm research in taxonomy and ecology. *Applied Soil Ecology*, 65, pp.35-42.
- Vivien, R., Wyler, S., Lafont, M. and Pawlowski, J., 2015. Molecular barcoding of aquatic oligochaetes: implications for biomonitoring. *PloS one*, 10(4), p.e0125485.
- Jo, H., Gim, J.A., Jeong, K.S., Kim, H.S. and Joo, G.J., 2014. Application of DNA barcoding for identification of freshwater carnivorous fish diets: Is number of prey items dependent on size class for *Micropterus salmoides*?. *Ecology and evolution*, 4(2), pp.219-229.
- Willerslev, E., Hansen, A.J., Binladen, J., Brand, T.B., Gilbert, M.T.P., Shapiro, B., Bunce, M., Wiuf, C., Gilichinsky, D.A. and Cooper, A., 2003. Diverse plant and animal genetic records from Holocene and Pleistocene sediments. *Science*, 300(5620), pp.791-795.
- Valentini, A., Pompanon, F. and Taberlet, P., 2009. DNA barcoding for ecologists. *Trends in Ecology & Evolution*, 24(2), pp.110-117.
- Handelsman, J., 2004. Metagenomics: application of genomics to uncultured microorganisms. *Microbiology and molecular biology reviews*, 68(4), pp.669-685.
- Trebitz, A.S., Hoffman, J.C., Grant, G.W., Billehus, T.M. and Pilgrim, E.M., 2015. Potential for DNA-based identification of Great Lakes fauna: match and mismatch between taxa inventories and DNA barcode libraries. *Scientific reports*, 5, p.srep12162.
- Riesenfeld, C.S., Schloss, P.D. and Handelsman, J., 2004. Metagenomics: genomic analysis of microbial communities. *Annu. Rev. Genet.*, 38, pp.525-552.
- Kathman, R.D. and Brinkhurst, R.O., 1998. Guide to the freshwater oligochaetes of North America. Aquatic Resources Center.
- Clarke, A.H., 1981. The freshwater molluscs of Canada. Canada Communication Group Pub.
- Jokinen, E.H., 1992. The freshwater snails (mollusca: gastropoda) of New York
- Smith, D.G. and Pennak, R.W., 2001. Pennak's freshwater invertebrates of the United States: Porifera to Crustacea. John Wiley & Sons.
- Klemm, D.J. ed., 1985. A guide to the freshwater Annelida (Polychaeta, Naidid and tubificid Oligochaeta, and Hirudinea) of North America. Kendall Hunt Publishing Company.