EPA United States Environmental Protection Agency

Lessons learned from broad - spectrum early-detection monitoring in the Laurentian Great Lakes

Anett Trebitz, Joel Hoffman, Erik Pilgrim & Greg Peterson

U.S. EPA Great Lakes Toxicology & Ecology Division (Duluth MN),
U.S. EPA Watershed & Ecosystem Characterization Division (Cincinnati OH)

ICAIS conference, 27-31 October 2019, Montreal Quebec

Context: Great Lakes AIS early detection monitoring

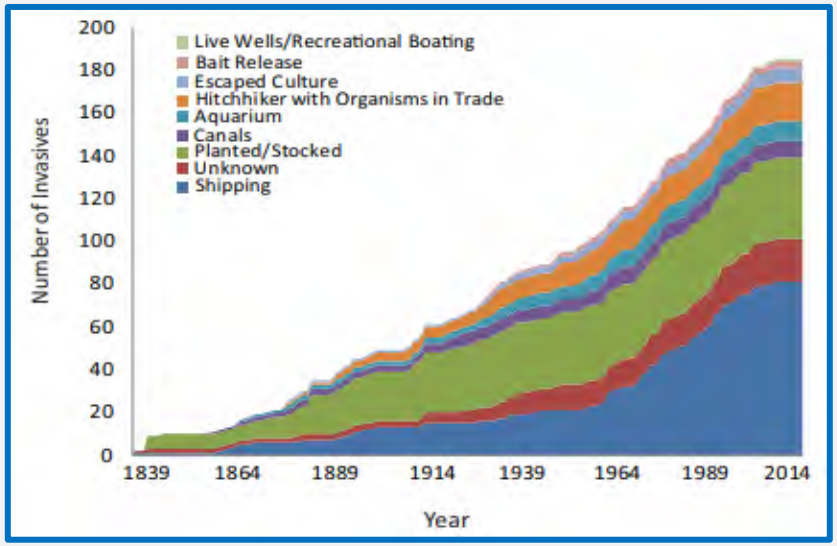
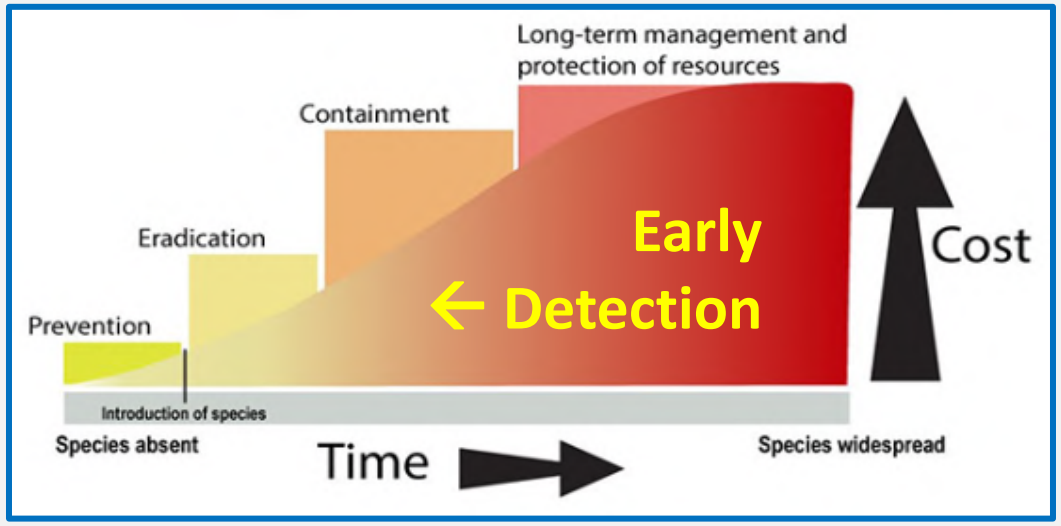


Figure source: State of Great Lakes 2017 tech rept.

AIS continue to arrive with extensive ecological and economic impact. *GL Water Quality Agreement* & *GL Restoration Initiative* call for development of multi-species early detection monitoring (EDM)

Early detection **means finding AIS while still rare**

Broad spectrum **enables discovery of unexpected AIS and informs ecological assessment & understanding**



EPA/ORD role: Tech basis for broad-spectrum EDM

Goals:

- Survey design recommendations
- Survey outcome evaluation tools
- Taxonomic tools, esp. DNA-based

Key elements:

- Complex GL settings with various AIS concerns
- Statistically robust design
- Comprehensive data collection across taxa and habitats
- Deliberate attention to analysis & refinement of survey performance

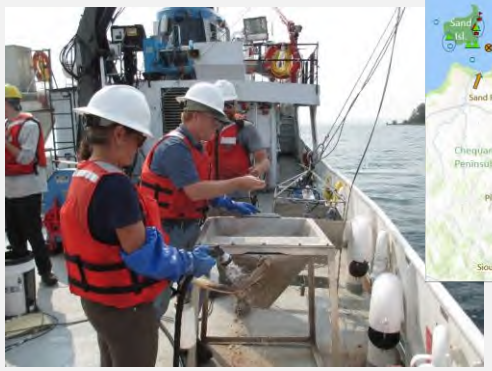
Learning from suite of sampling campaigns

Fish monitoring in GL ports (2006-present)

Partnership with U.S. FWS

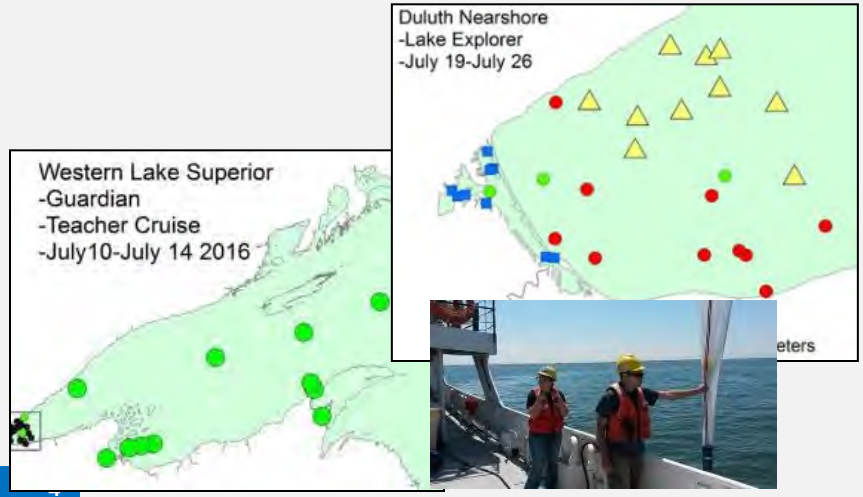
Dreissena in Lake Superior 2017, 2019

Motivated by U.S. NPS concerns



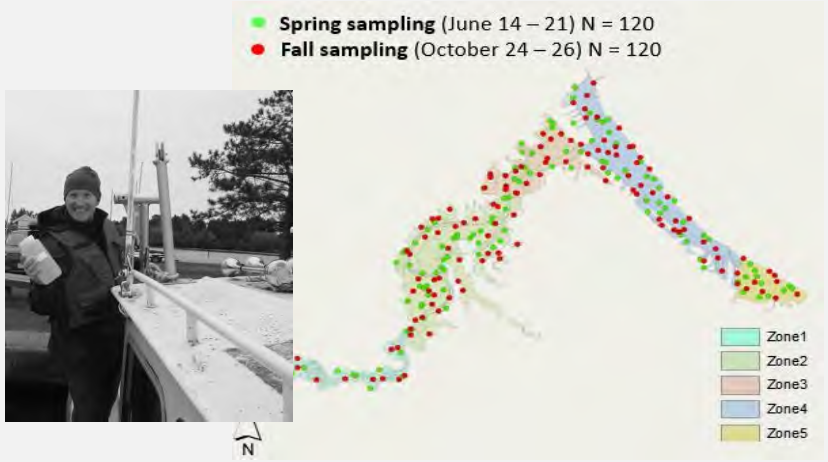
Invertebrate community surveys (2016)

Coord with EPA/GLNPO sampling cruise

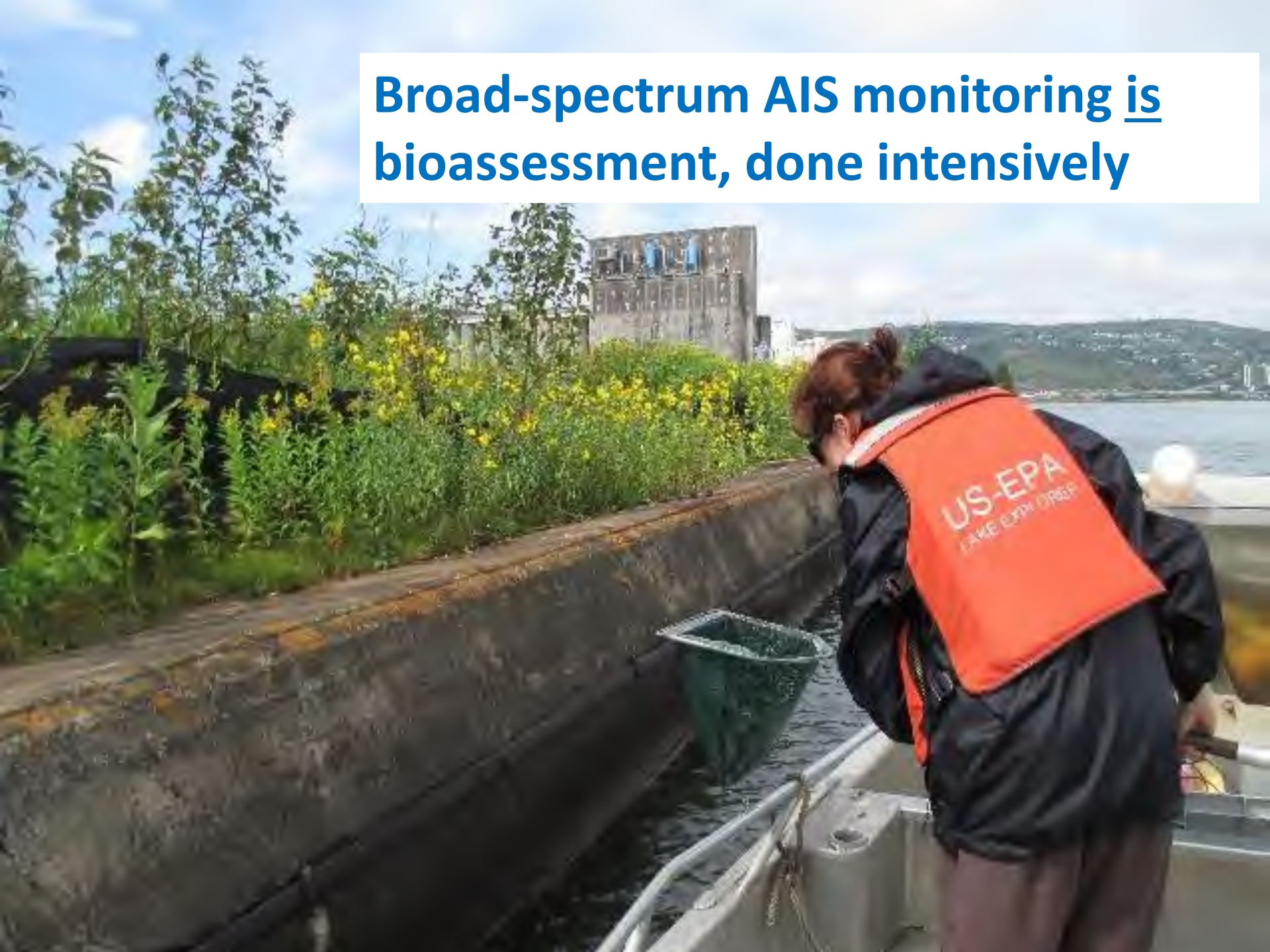


St. Louis River Estuary, 2016 eDNA survey

Partnership with TNC



Broad-spectrum AIS monitoring is bioassessment, done intensively



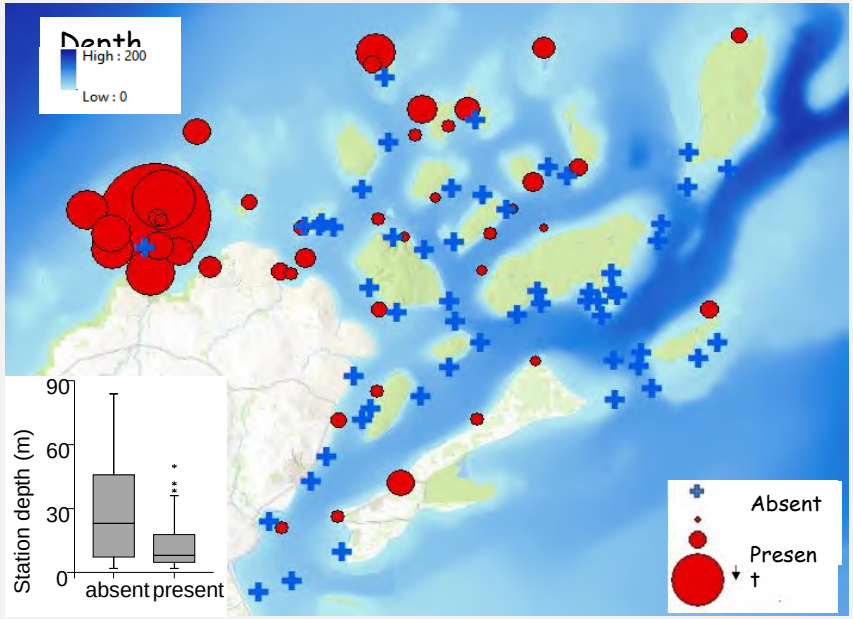
AIS monitoring is bioassessment, done intensively

Going beyond basic biomonitoring:

- Seek rare taxa in addition to common ones (more than bioindicators)
- Cover nooks & crannies with suite of gear (not easily standardized/streamlined)
- Survey effectiveness eval of high importance (was search thorough, even if AIS <u>not</u> found?)

Dreissena EDM example:

- Big field effort: 100 stations, multiple gears, 3 ships x 2 weeks
- Invert community baseline acquired
- Veligers in 44% of zoop tows but <u>very</u> low abundance
- Searched large sample fractions; usual small aliquots would miss
- No hits in eDNA, benthos, or video



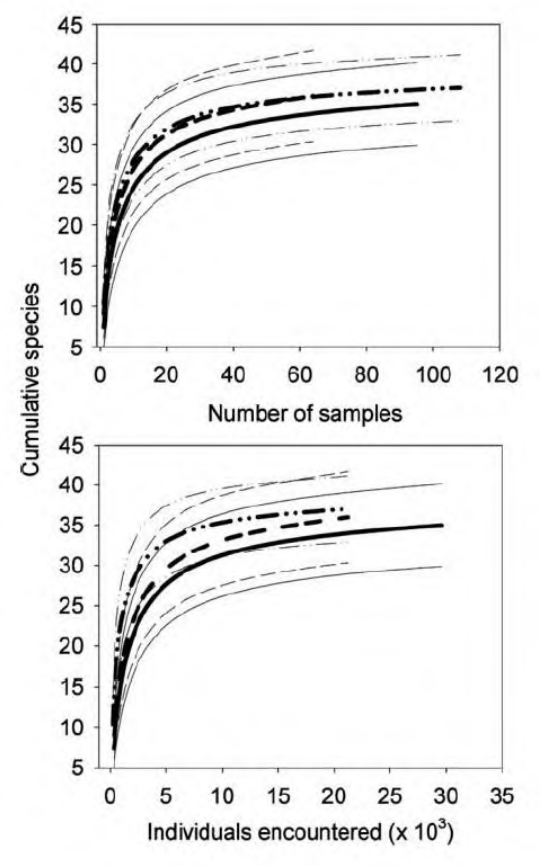
AIS monitoring is resource intensive

Fish (adult/juvenile)

est. 38 species, found 37

+2 samples to get 95%

~75,000 individuals

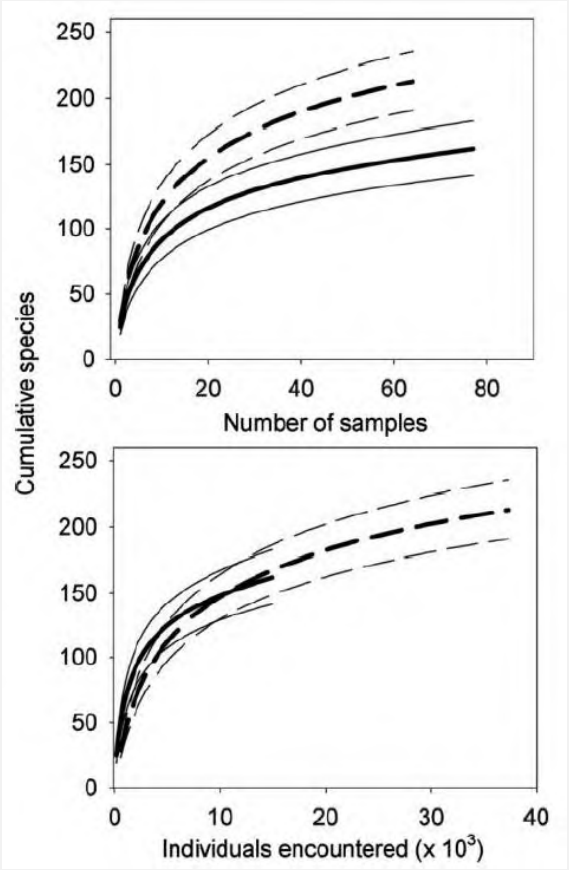


Benthic inverts

est. 205 species, found 162

+131 samples to get 95%

~100,000 individuals

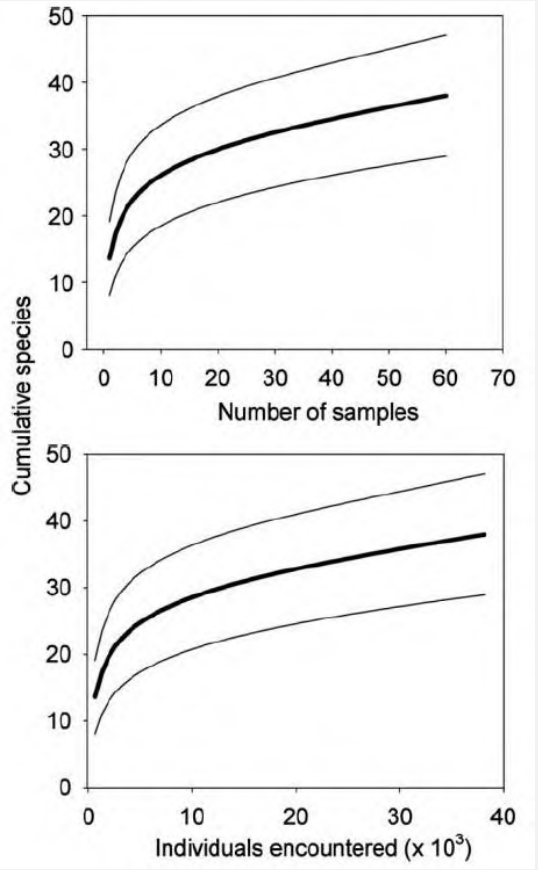


Zooplankton

est. 88 species, found 37

+716 samples to get 95%

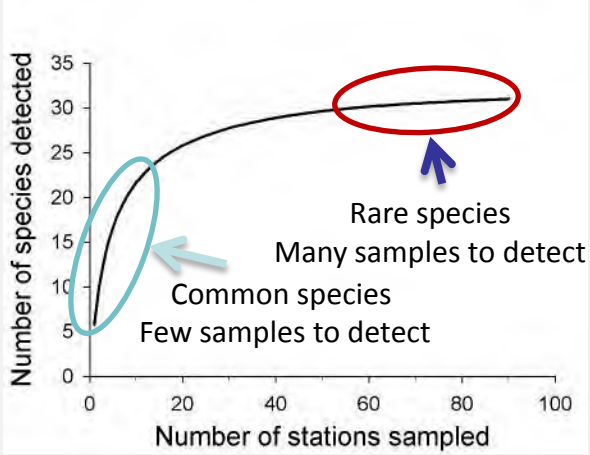
~500,000 individuals

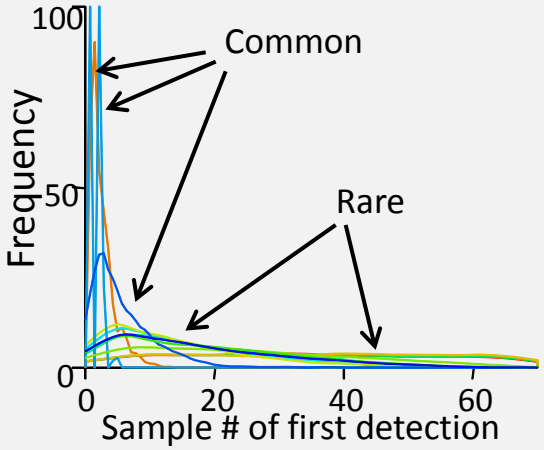


Survey eval tools shouldn't hinge on finding AIS

Species accumulation theory:

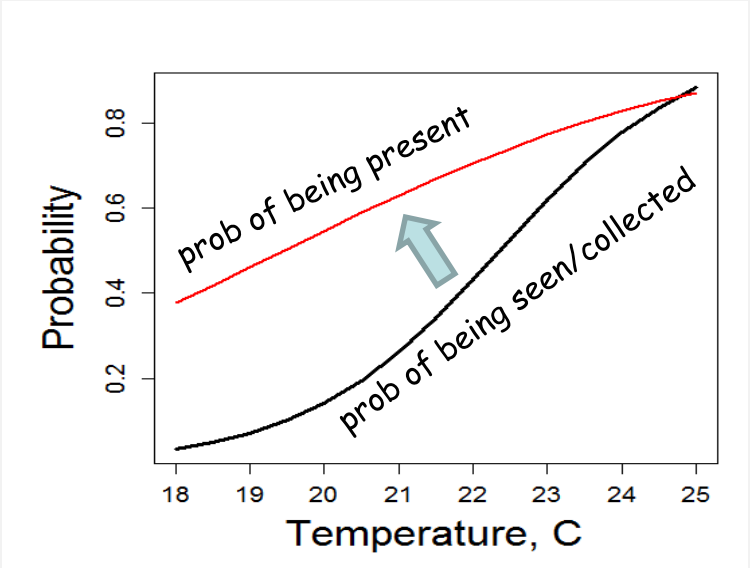
- Builds on ALL taxa in sample
- Distance to asymptote & efficiency reaching it are performance measures





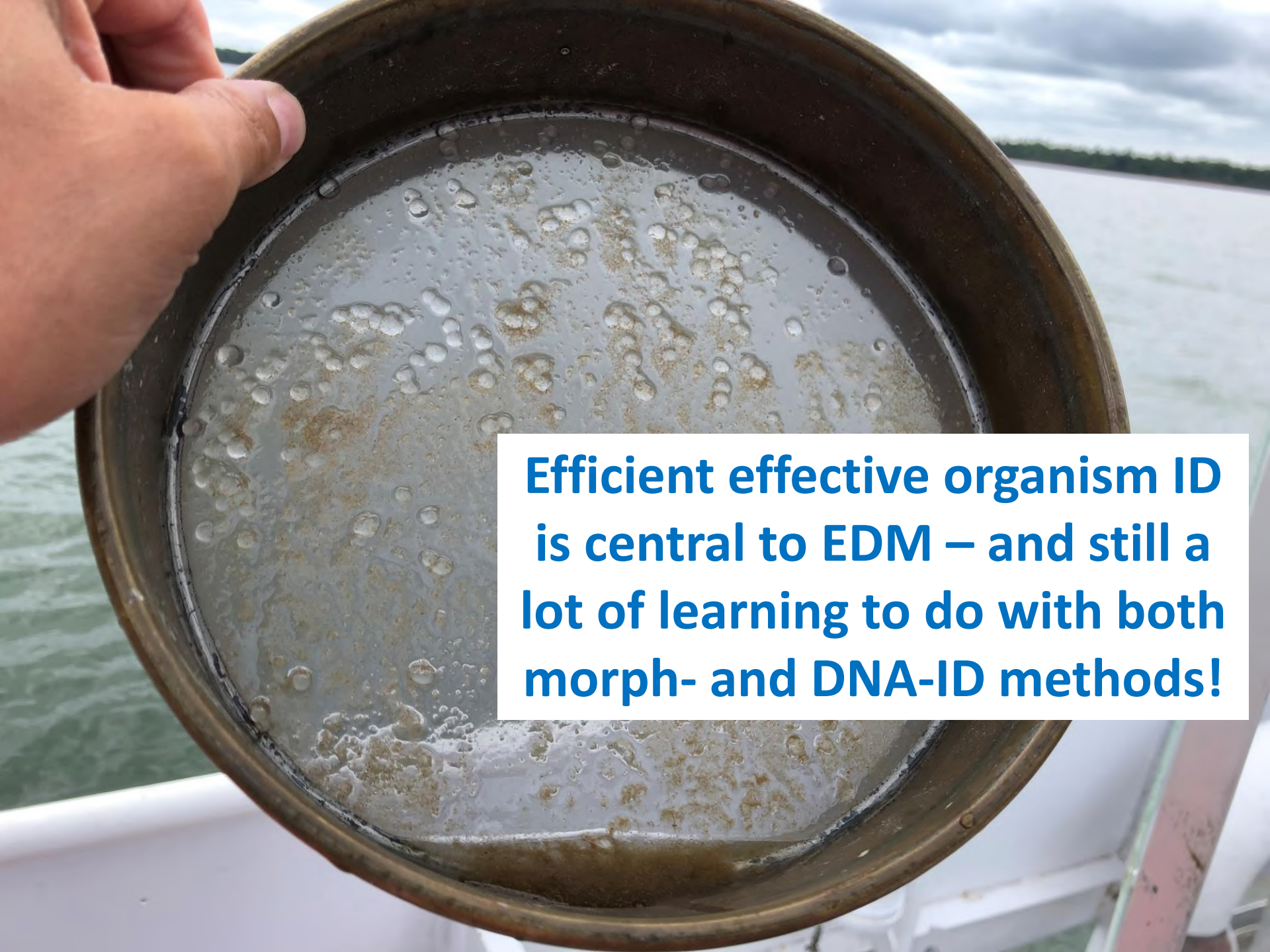
Detection probability: Improved by:

- reducing encounters with common species
- amplifying individual's signal



Occupancy analysis:

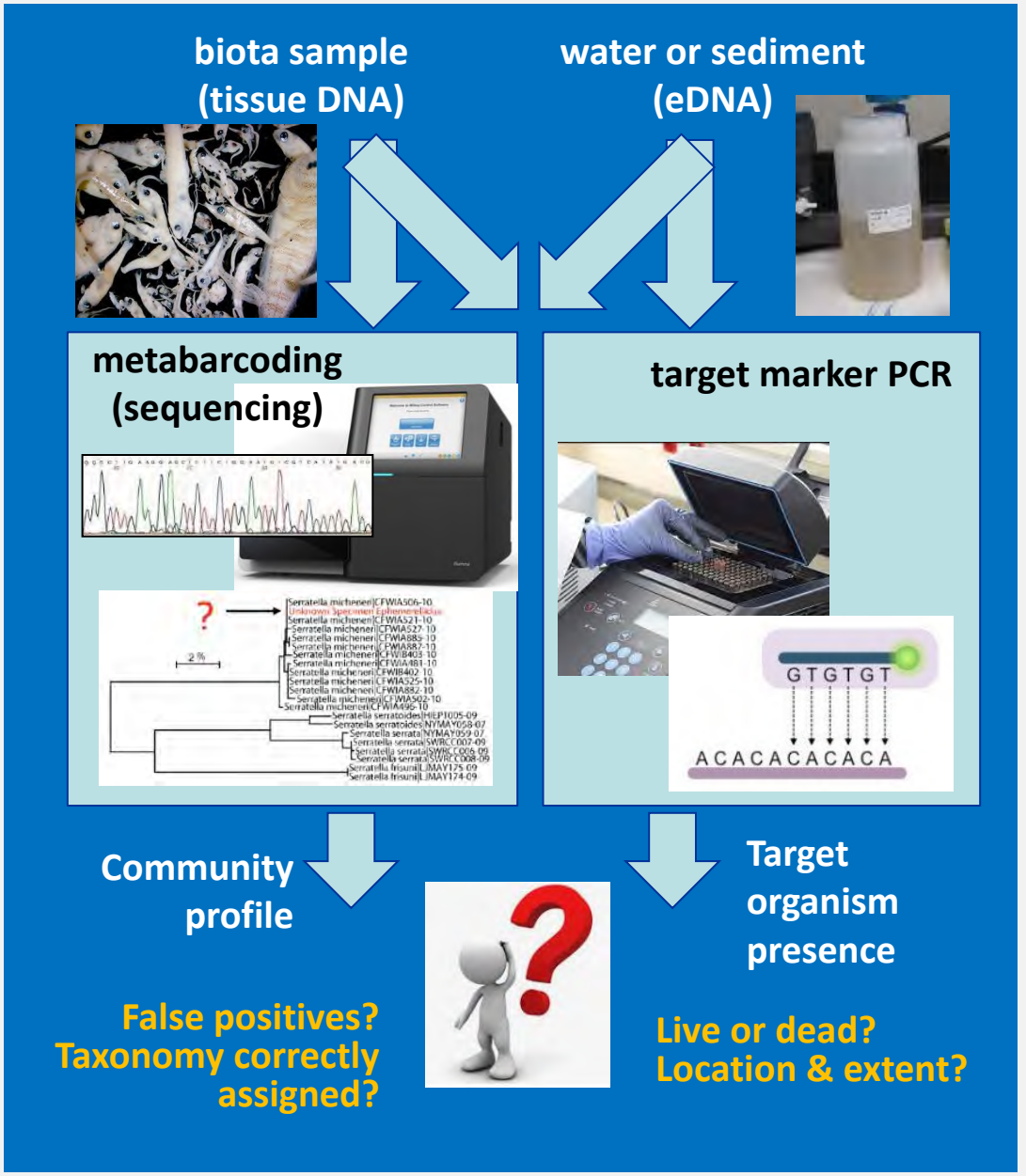
Freq in replicate samples/visits used to index detectability and refine distribution estimates.


Efficient effective organism ID is central to EDM – and still a lot of learning to do with both morph- and DNA-ID methods!

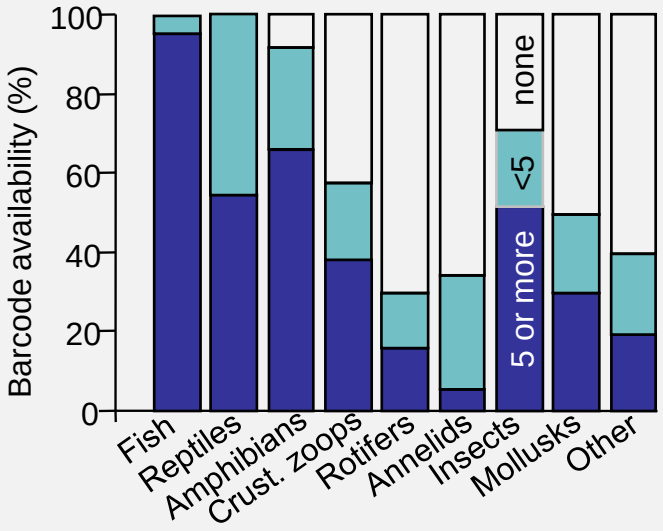
DNA-ID not new, but also still not operationalized

R&D still needed on:

- Field & lab protocols (sample #, size, strategy)
- Info sufficiency (barcodes, target markers)
- Process bias & errors
- Bioinformatics decision points (false presence vs. absence tradeoff?)
- Detection probability (understand, improve, account for)
- Indicators & inference (quantifying, validating)

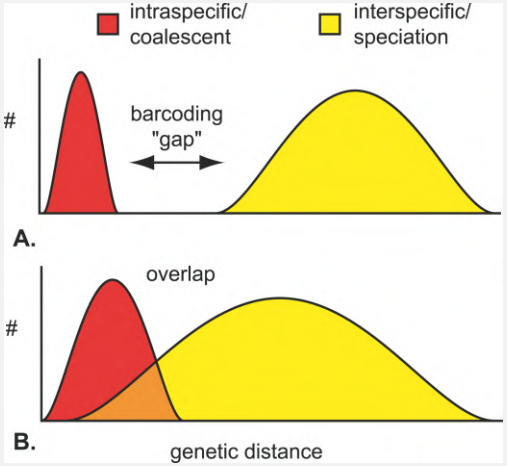


Neither morph- nor DNA-ID *resolve* all taxa



DNA-ID can't resolve when:

- Lack reference barcode – large percentage of GL invertebrates!
- Gene regions overlap – e.g., CO1 confuses *Cottus ricei* (native sculpin) & *C. gobio* (threat-list)



Morph-ID can't resolve when:

- Lack differentiating features (e.g., eggs, early life-stages, damaged specimens)
- Keys not available or not cognizant of non-native look-alikes

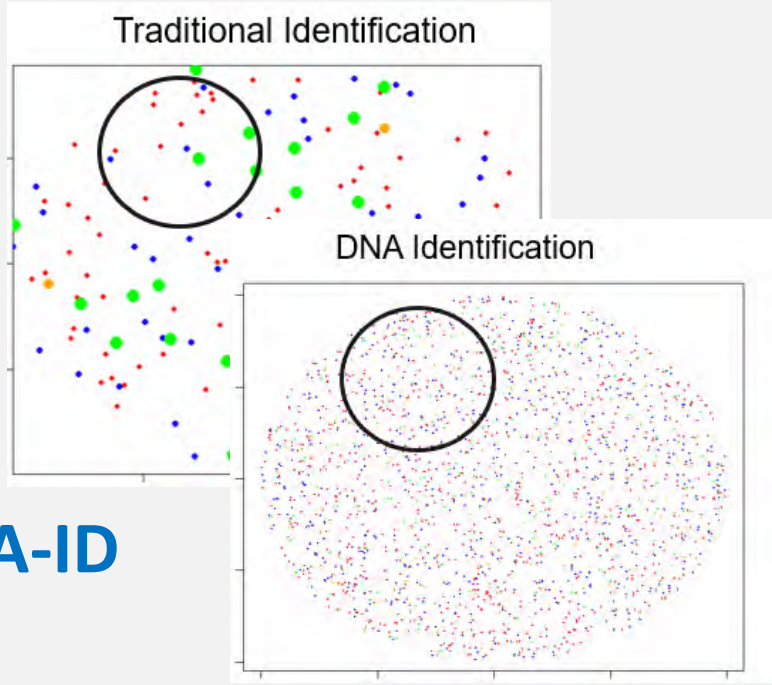
Neither morph- nor DNA-ID *find* all taxa

DNA-ID misses taxa when:

- DNA hard to extract (e.g., mollusks)
- Low biomass

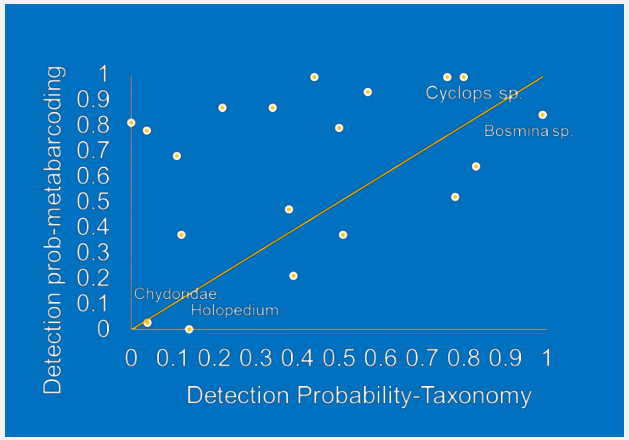
Morph-ID misses taxa when:

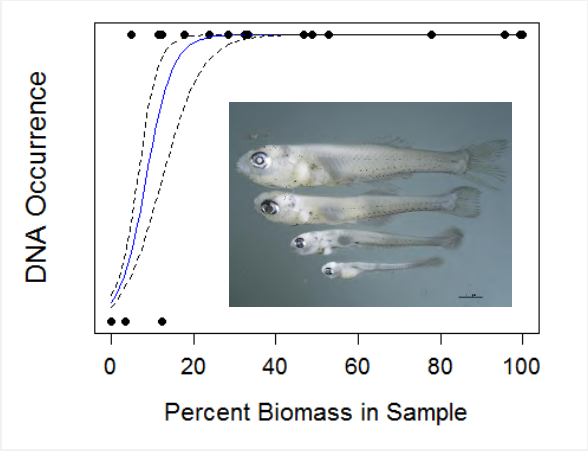
- Deep search is prohibitive (microscopy, expertise, slide-mounting, etc.)



Detectability often higher with DNA-ID than morph-ID but not always!

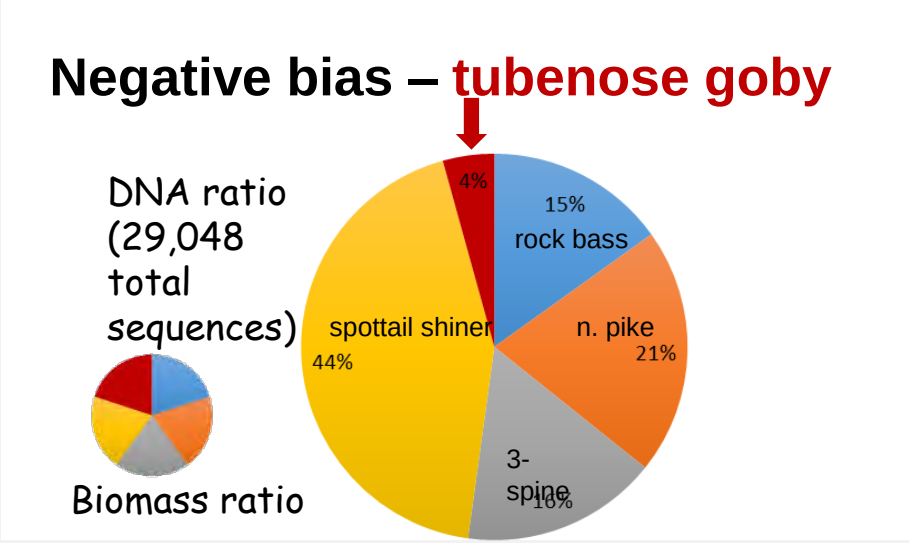
Example L. Superior zoop



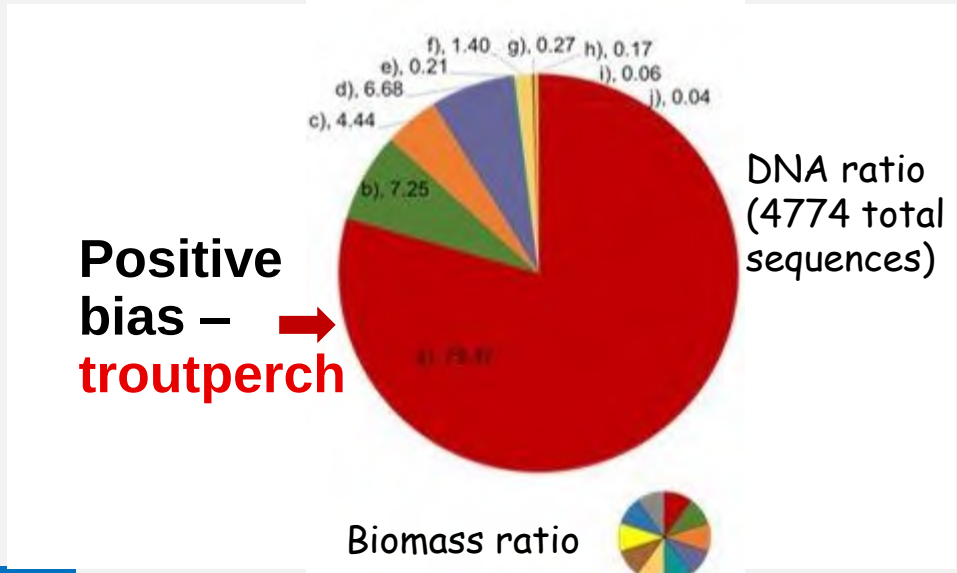


Example SLRE larval fish

Tricky to compare morph vs. DNA-ID beyond P/A



Hatzenbuhler et al. 2017.



Morph- vs. DNA-ID give different answers!

- Inherent diffs – counting individuals vs. DNA copy number
- Organism-dependent diffs – readily shed/extracted vs. recalcitrant DNA
- Sample mix-dependent diffs – PCR amplification & marker binding

Not crucial for AIS find but is for bioindicators & AIS impact


So where are we at right now
for broad-spectrum EDM??

EDM easiest for (adult) fish, harder for inverts

Consider larvae rather than adult fish

- Not amenable to field ID, but…
- Can be more abundant & available than adults
- Larvae often the stage transported & introduced
- DNA methods & barcodes largely complete

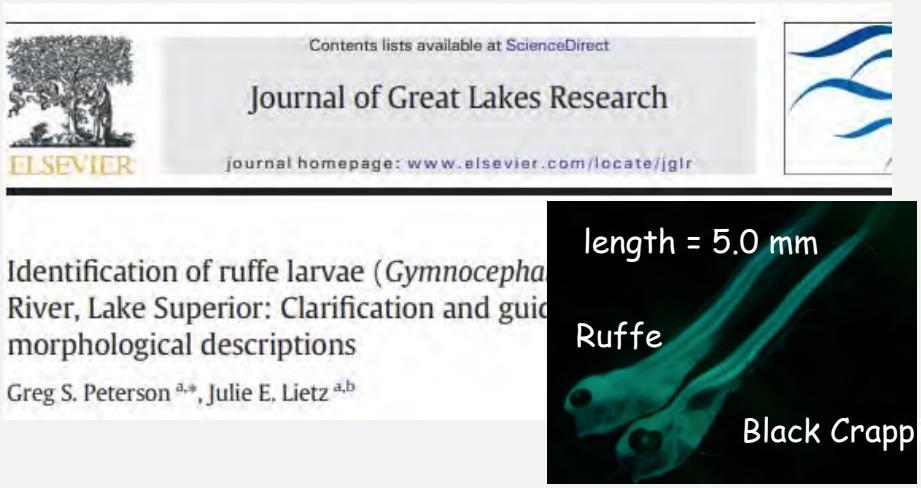
Zoops & benthos harder than fish:

- More habitats to search, more life histories to consider
- Less biogeographic & taxonomic knowledge
- Many more taxa to ID and count
- DNA methods & barcodes need work
- Pairing EDM with other biomonitoring goals has great potential here!

Building fauna knowledge to better recognize AIS

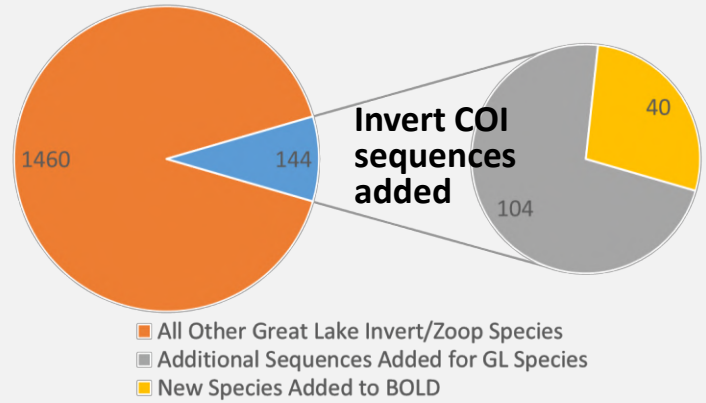
Improving ID-trait knowledge

- Larval ruffe vs. centrarchids or percids
- Tandem morph & DNA work was key!



Building out barcode library, esp. for inverts

- EPA-ORD doing this
- So are EPA-GLNPO funded teams

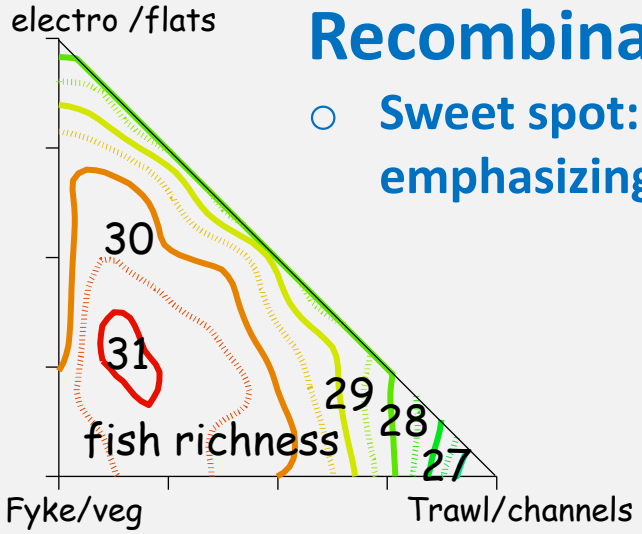


Getting invert info as well compiled as vertebrate info

- Fauna inventory now assembled
- Helps with "what to expect?" & "is this new?"

https://www.glerl.noaa.gov/data/waterlife/
Trebitz et al. 2019, J Great Lakes Res

Efficient effective design – random first, then optimize

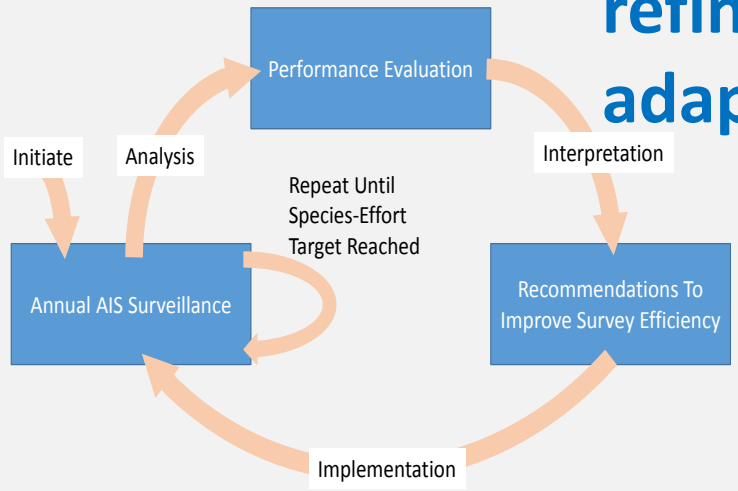


Recombination analyses to assess mix:

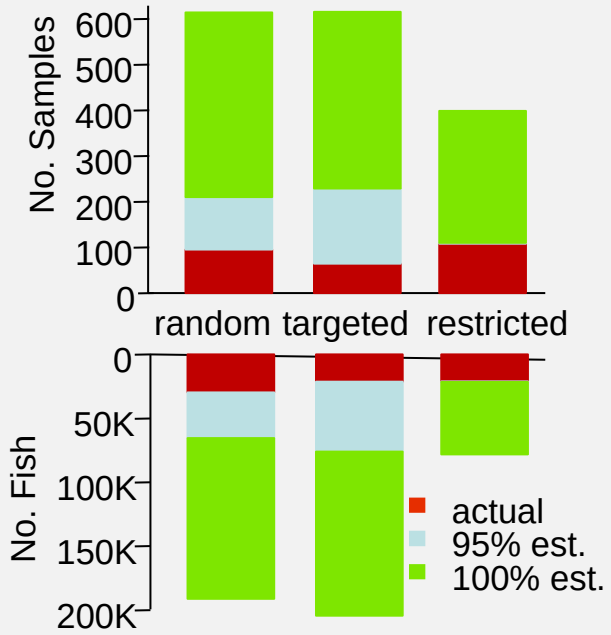
- Sweet spot: Covering all unique habitats while emphasizing most informative subset

Formal design comparisons:

- Random is robust starting point, evolve as info accrues

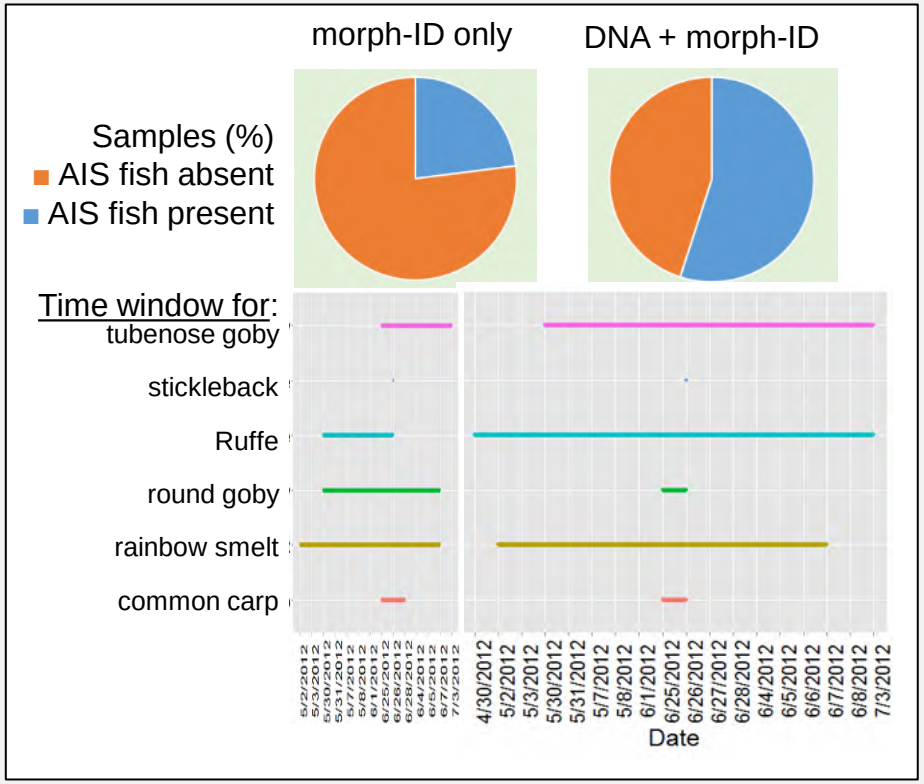


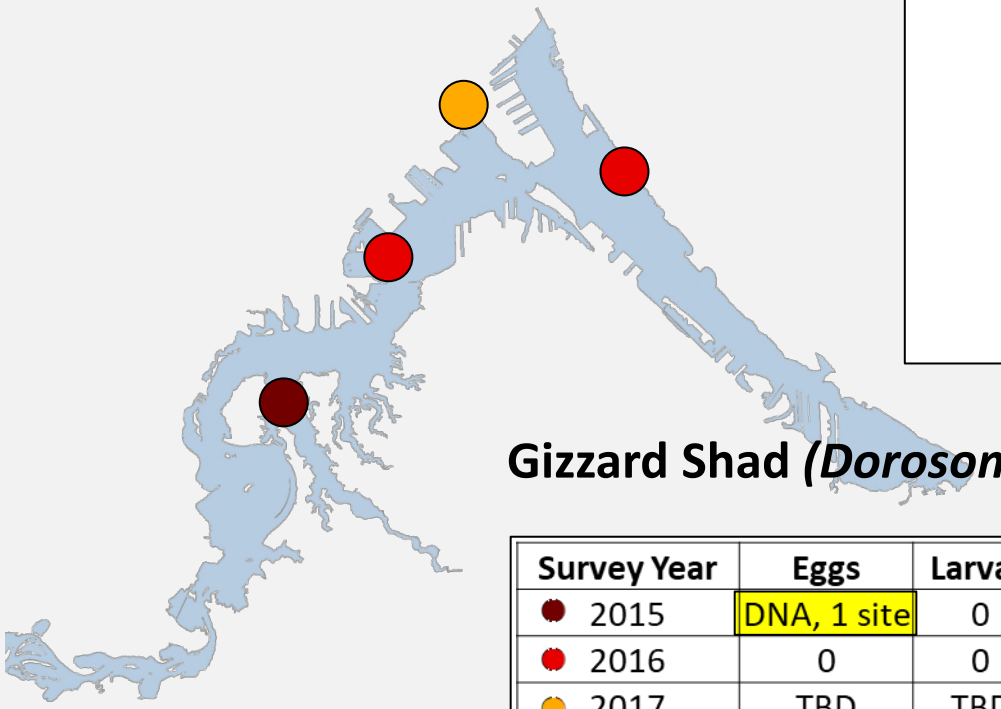
Review & refine adaptively!



Morph & DNA together as effective path forward

DNA-ID and morph-ID together improve AIS detection and learning in time and space



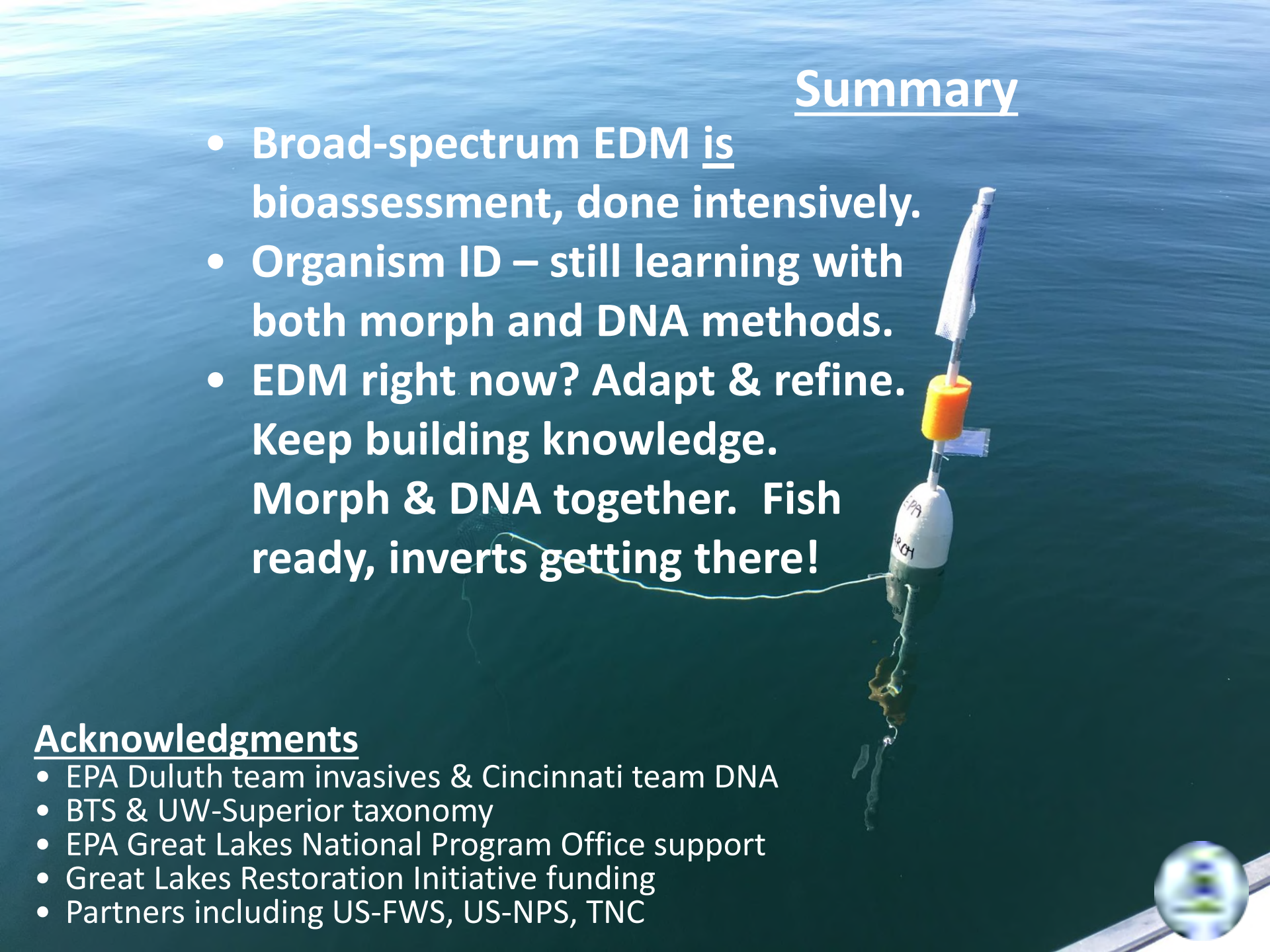


Gizzard Shad *(Dorosoma cepedianum)*

Survey Year	Eggs	Larvae	Adult/Juv	eDNA
2015	DNA, 1 site	0	0	NA
2016	0	0	0	2 sites
2017	TBD	TBD	2 fish, 1 site	NA

Summary

- **Broad-spectrum EDM is bioassessment, done intensively.**
- **Organism ID – still learning with both morph and DNA methods.**
- **EDM right now? Adapt & refine. Keep building knowledge. Morph & DNA together. Fish ready, inverts getting there!**



Acknowledgments

- EPA Duluth team invasives & Cincinnati team DNA
- BTS & UW-Superior taxonomy
- EPA Great Lakes National Program Office support
- Great Lakes Restoration Initiative funding
- Partners including US-FWS, US-NPS, TNC

ABSTRACT: Motivated by decades of ecologic and economic impacts from a growing list of nonindigenous species, the 2012 Great Lakes Water Quality Agreement between Canada and the United States calls for establishment of an aquatic non-indigenous species early detection and rapid response network. This presentation focuses on lessons learned from broad-spectrum (i.e., cross-species) early-detection monitoring conducted by the U.S. Environmental Protection Agency as part of this Great Lakes network. Such monitoring is inherently resource-intensive, with surveys capable of detecting 95% of the species pool taking on the order of 100, 200, and 500 samples for fish, benthic invertebrates, and zooplankton respectively. We have found a random probability design an effective starting point for monitoring; once information concerning species distributions is generated the design can be optimized by emphasizing habitats and collection devices that contribute most strongly to the species pool. Effective tools for generating such information include occupancy modeling and community rarefaction, neither of which hinge on the presence or identity of any particular non-indigenous species. In applying a combination of organism collections identified via morphology and DNA and water samples identified only via contained eDNA, we have learned to temper enthusiasm for DNA metabarcoding with constraints stemming from sequencing difficulties and still limited invertebrate barcode availability. An adaptive monitoring cycle involving repeated assessment, refinement, and outcome communication has proven a helpful framework for broad-spectrum early-detection monitoring in the Great Lakes.