

# A VALIDATED GENETIC REFERENCE DATABASE FOR MARINE ANIMALS OF THE NORTHWEST ATLANTIC OCEAN: ACTUAL AND FUTURE IMPROVEMENTS

Claude Nozères

 @cnozeres

# A REFERENCE DATABASE — WHY BUILD IT?

## **Double validation: Taxonomy + Genomics**

Some species are problematic, making it difficult to use names, but using genus or other general level may be insufficient for ecological analyses

## **Regional patterns**

Effort varies across fisheries surveys, making it a challenge to show spatio-temporal patterns if not sure of the specimens and names involved

## **Biodiversity Informatics**

Online tools make it possible to standardize, publish, analyze, and compare authoritative information on taxa

# OUTLINE



**Past** – the work that lead to development of current databases



**Present** – current application: corrections and discoveries



**Next** – larger scale, automated analyses, publish and share

# TALE TAKES PLACE IN A FARAWAY SEA

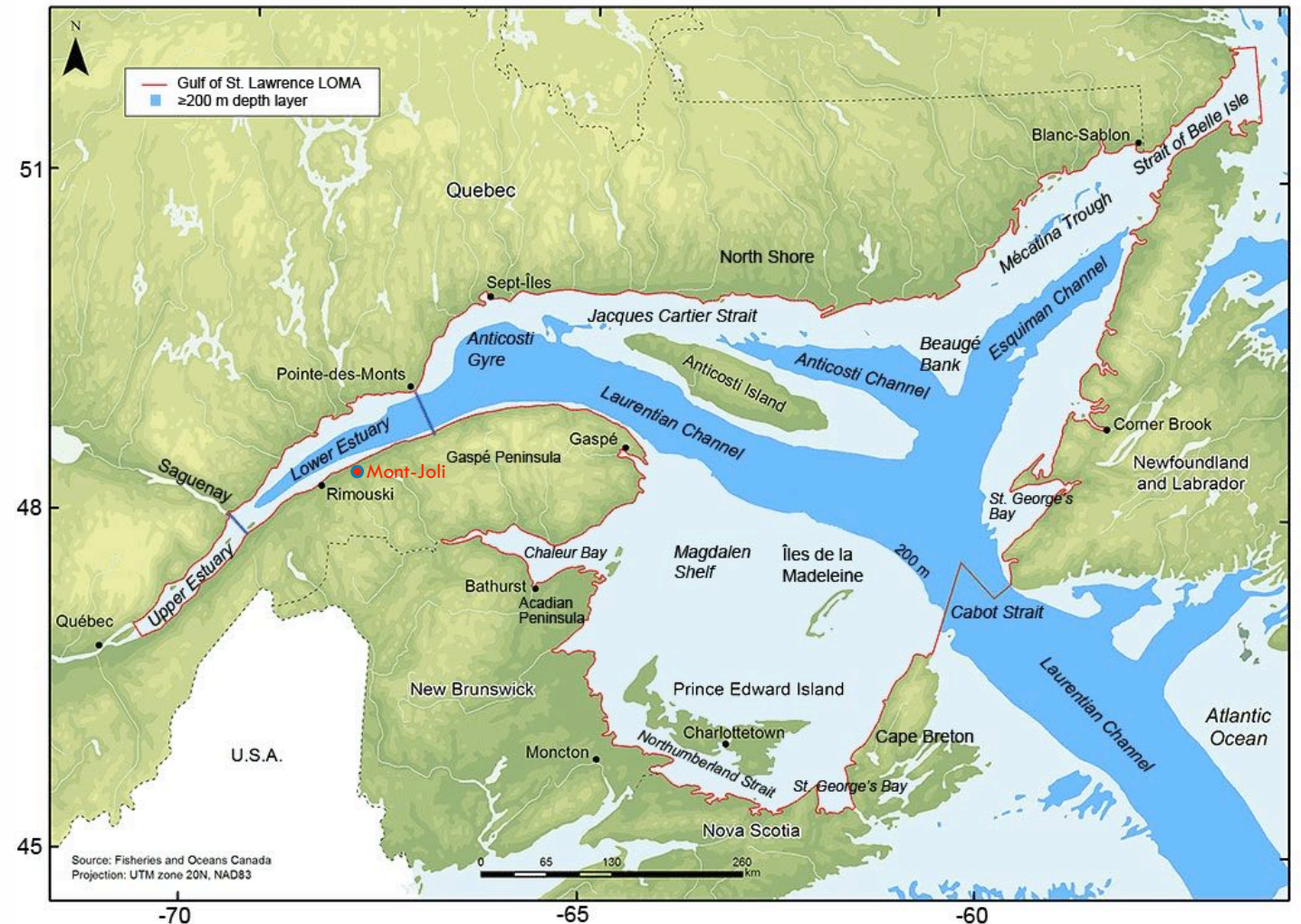
Originally from Vancouver

Studied grey seals at Guelph

Came to Mont-Joli to study  
marine mammal trophic  
ecology & foodwebs

Realized we needed validated  
information on prey species  
present in the region

Have been reviewing marine  
species here for 20 years



Source: Fisheries and Oceans Canada

<https://www.dfo-mpo.gc.ca/oceans/management-gestion/gulf-golfe-eng.html>



# PAST — BUILD REFERENCE BASES FOR PREDATORS

Student project (U. Laval, 2001-2006):

- 1) Analyze lipid signatures in beluga blubber — a **non-destructive** approach to reveal diet
- 2) Analyze signatures of all potential prey species to build a **prey reference database**
- 3) Using a combination of multivariate tools, model the **contributions of prey** in the beluga

Innovative statistics

World-class chemical analyses

Extensive sample database

Weak link — how to know if species were certain?

## Fatty acid signature classification patterns amongst beluga whales (*Delphinapterus leucas*), harbour seals (*Phoca vitulina*), and potential prey species in the St. Lawrence estuary, Canada

Claude Nozères<sup>1,2</sup>, Véronique Lesage<sup>2</sup>, Sara J. Iverson<sup>3</sup>, Mike O. Hammill<sup>2</sup>, and Helga Guderley<sup>1</sup>  
(1) Université Laval, Ste-Foy, Québec, (2) Maurice-Lamontagne Institute, Fisheries and Oceans Canada, Mont-Joli, Québec, (3) Dalhousie University, Halifax, Nova Scotia



### ABSTRACT

Fatty acid (FA) signature analysis is a tool that can reveal foraging patterns among marine mammals and their prey. Gas chromatography was used to investigate the FA composition of 660 animals sampled in the St. Lawrence estuary including harbour seals (n=43), beluga whales (n=81), and 12 prey species (N=533). Of 66 detected FAs, those >0.4% by mass of the total FAs were retained to produce 16 logratio variables of potential dietary importance. Exploratory principal components analysis (PCA) produced 3 factors accounting for 80% of the total variance. Factor loadings and scores grouped belugas and seals together on factor 1, distinct from all marine fish and invertebrates in the major saturated and monounsaturated FAs. Factors 2 and 3 were comprised of diel-linked monounsaturated and polyunsaturated, respectively. Predictive classification was evaluated with discriminant function (DF) analysis, using the 13 highest-contributing communalities >0.7. FA logratio variables to correctly classify 90% of all individuals by species. Belugas and seals were separated from prey on the first DF, while being separated from each other and linked to prey on the remaining 12 DFs. These results are consistent with previous findings that the FA signatures of marine mammal blubber reflect but do not match that of their prey. An additional DF analysis using only the 3 factors retained from PCA, classified 57% of individuals by species. Examination of the misclassifications revealed the variability within species groups by age-class or season, and similarities in diet among certain species, as suggested by field observations. Future prey modeling should incorporate the variability in seasonal, regional and length-class differences when possible, in order to estimate group contributions to diet using blubber analyses.

### INTRODUCTION

The potential for using FA signatures in foraging ecology studies has been apparent for a number of years. To use the method quantitatively, it is important to compile the signatures for all species directly involved in the food web of seals and whales. However, a comparative analysis of only the key species for a given region under study may also be effective and serves as a valuable beginning in the assessment of both key indicator FAs as well as potential trophic interactions among populations of marine mammals and their prey.

### MATERIALS AND METHODS

Blubber samples from stranded belugas (n=81) and biopsied harbour seals (n=43) were collected from 1980 to 2001. A total of 533 whole specimens of 12 prey species of fish and shrimp (TABLE 1) were collected from scientific missions and tidal fish weirs throughout three subregions, from the upper estuary to the northern Gulf of St. Lawrence (FIGURES 3, 4) in 2000-2001. FA composition of samples was analyzed using gas chromatography (Budge et al. 2002). Of the 66 detected FAs, the 16 that were each >0.4% by mass of total FAs were retained for statistical analysis. These FA data were transformed to produce 16 logratio variables that were subsequently used in principal components analysis (PCA). Further analysis via discriminant functions (DFA) was performed using the 13 highest-contributing logratio variables (communalities > 0.7) to produce predictive classifications using the SPSS 11 software package. Additional insight regarding diet and classification of prey species was obtained from the examination of prey stomach contents and unpublished data.



FIGURE 1. STRANDED BELUGA WHALES WITH BLEBBS COLLECTED, AUGUST 2000.

FIGURE 2. STRANDED BELUGA WHALES WITH BLEBBS COLLECTED, AUGUST 2000.

### CLASSIFICATIONS (DFA)

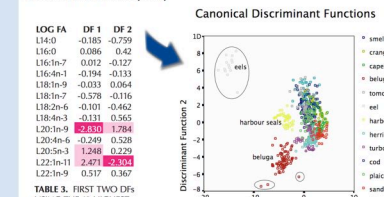


FIGURE 3. PLOT OF ALL INDIVIDUALS FOR DF1 AND DF2. OUTLIERS CIRCLED: 3 BELUGA SAMPLES, AND ALL EELS.

### EXPLORATORY (PCA)

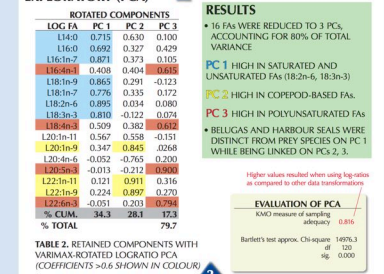


FIGURE 4. EXPLORATORY PCA PLOT OF LOGRATIO FAS (COMMUNALITIES > 0.7) SHOWN IN COLOUR.

COMMON NAME	SCIENTIFIC NAME	N	COMMENT
Beluga whale	<i>Delphinapterus leucas</i>	81	inner blubber
Harbour seal	<i>Phoca vitulina</i>	43	inner blubber
American eel	<i>Anguilla rostrata</i>	14	at sea only to species
American sand lance	<i>Ammodytes americanus</i>	42	marine pelagic
Ribbon barnaculus	<i>Acetopterus (Notolepis) rissio</i>	18	marine pelagic
Atlantic herring	<i>Clupea harengus</i>	72	marine pelagic
Redfish	<i>Sebastes mento</i>	127	marine pelagic
Capelin	<i>Mallotus villosus</i>	71	marine pelagic
Winter flounder	<i>Paralichthys americanus</i>	43	marine groundfish
American plaice	<i>Hippoglossoides platessoides</i>	40	marine groundfish
Greenland halibut (turbot)	<i>Paralichthys platessoides</i>	26	marine groundfish
Atlantic cod	<i>Gadus morhua</i>	14	marine groundfish
Atlantic tomcod	<i>Micropogonias tomcod</i>	29	estuarine groundfish
Sand shrimp	<i>Crangon septemspinatus</i>	28	estuarine invertebrate

TABLE 1. MARINE MAMMAL AND MAJOR PREY SPECIES SAMPLED FROM THE ESTUARY AND GULF OF ST. LAWRENCE.

### EVALUATION OF DFA

SPECIES	% CORRECT	% MISCLASSIFIED AS OTHER SPECIES
Beluga whale	90.5	9.5
Harbour seal	90.5	9.5
Capelin	90.5	9.5
Atlantic herring	90.5	9.5
Redfish	90.5	9.5
Winter flounder	90.5	9.5
American plaice	90.5	9.5
Greenland halibut (turbot)	90.5	9.5
Atlantic cod	90.5	9.5
Atlantic tomcod	90.5	9.5
Sand shrimp	90.5	9.5

TABLE 2. SUMMARY OF SPECIES PREDICTIONS USING ALL 13 DFs.

### RESULTS

- THE 13 HIGHEST-CONTRIBUTING FAS CORRECTLY CLASSIFIED NEARLY 90% OF ALL SPECIMENS BY SPECIES CLASS.
- 16 FAS WERE REDUCED TO 3 PCs, ACCOUNTING FOR 80% OF TOTAL VARIANCE.
- PC 1 HIGH IN SATURATED AND UNSATURATED FAS (18:2n-6, 18:3n-3).
- PC 2 HIGH IN COPPOD-BASED FAS.
- PC 3 HIGH IN POLYUNSATURATED FAS.
- BELUGAS AND HARBOUR SEALS WERE DISTINCT FROM PREY SPECIES ON PC 1 WHILE BEING LINKED ON PC 2, 3.

### DISCUSSION AND CONCLUSIONS

- Results are consistent with previous findings that the FA signatures of marine mammal blubber reflect, but do not match, that of their prey. This is also consistent with the need for establishing 'calibration factors' (Iverson et al. 2004) before making direct comparisons between predators and their prey.
- In general, the FA signatures observed here complemented stomach contents analysis in suggesting dietary differences between prey species. However, sampling origin and data treatment affected results.
- FA signatures of several prey species resembled each other, consistent with shared dietary habits. FA signatures also differed within some species, again consistent with differences in diet by size and region. These results indicate the use of FA in revealing trophic similarities and differences among individuals.
- Variability within and among prey species suggests that a FA signature database should be established on a regional basis, and is complemented by tools such as stomach contents and isotopic tracers.

### ACKNOWLEDGEMENTS

We thank Shelley Lang and all the lab technicians at Dalhousie University for conducting the fatty acid sample analyses. Thanks also to everyone who assisted in field sampling. This research was funded by the Department of Fisheries and Oceans Canada, Species at Risk Program.

### REFERENCES

- Budge, S. M., S. J. Iverson, W. D. Bowen, and R. G. Ackman. 2002. Among- and within-species variability in fatty acid signatures of marine fish and invertebrates on the Scotian Shelf, Georges Bank, and southern Gulf of St. Lawrence. *Canadian Journal of Fisheries and Aquatic Sciences* 59:886-898.
- Dalgaard, J., M. S. John, G. Katner, D. Muller-Naveira, and W. Hagen. 2003. Fatty acid trophic markers in the pelagic marine environment. *Advances in Marine Biology* 46:223-340.
- Field, A. 2001. *Discovering Statistics using SPSS for Windows*. SAGE Publications Ltd, London, UK.
- Iverson, S. J., Field, C., Bowen, W. D., and Blanchard, W. 2004. Quantitative fatty acid signature analysis: a new method of estimating predator diets. *Ecological Monographs* in press.

# PAST — DOCUMENT SPECIES

To ensure reference samples were accurately identified, began to **photograph the species** encountered on N. Gulf of St. Lawrence surveys

**Produced posters, guides and photo catalogues** to help folks identify samples in captures on surveys

**Reviewing capture data** across surveys and years, would sometimes detect patterns related to sampling effort, not necessarily ecology

[https://figshare.com/authors/Claude\\_Nozeres/101073](https://figshare.com/authors/Claude_Nozeres/101073)

[https://www.researchgate.net/profile/Claude\\_Nozeres](https://www.researchgate.net/profile/Claude_Nozeres)



# PAST — DFO AND BIODIVERSITY REFERENCE WORK

During the *Census of Marine Life* (2000-2010), DFO was a pioneer in marine biodiversity informatics – one of the largest contributors to occurrences on OBIS, photos and notes on WoRMS, and sequences to BOLD

## National working groups and datasets\*

### Image data standards

Best practices

### Physical sample collections

Collections management

### Taxonomic standards

from ITIS to WoRMS (World Register of Marine Species)

**CaRMS** (Canadian Register of Marine Species): names, lists, photos

**OBIS Canada** (Ocean Biogeographic Information System): occurrences

**Centre for Marine Biodiversity:** website, resources, meetings

\* inactive since 2015 due to retirements, especially of **Mary Kennedy**, BIO



# PAST — EXPLORATORY GENOMICS

## GenBank (2000s)

Researchers contact us for specimens, post sequences, publish articles

## BOLDsystems (barcoding) (mid-2000s)

Sample on survey ships, post sequences online with **photos** and **maps**, email **corrections** to website, publish articles (2010s):

**Carr et al. (2011)** *A Tri-Oceanic Perspective: DNA Barcoding Reveals Geographic Structure and Cryptic Diversity in Canadian Polychaetes.*

<https://doi.org/10.1371/journal.pone.0022232>

**McCusker et al. (2012)** *Barcoding Atlantic Canada's commonly encountered marine fishes.* <https://doi.org/10.1111/1755-0998.12043>

**Hotke (2015)** *DNA Barcode Variability in Canadian Cnidaria.*  
<http://hdl.handle.net/10214/9236>

**Radulovici et al. (2015)** *The Application of DNA Barcodes for the Identification of Marine Crustaceans from the North Sea and Adjacent Regions.*  
<https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0139421>

**BOLD SYSTEMS** DATABASES IDENTIFICATION TAXONOMY WORKBENCH RESOURCES LOGIN

### PUBLIC DATA PORTAL - Specimen Record

Enter search terms PUBLIC DATA SEARCH

Specimen: [DWC](#) [XML](#) [TSV](#)  
Sequence: [FASTA](#) [TRACE](#)  
Combined: [XML](#) [TSV](#)

Record Details For SCFAC445-06 [Back to Search: Record](#)

#### IDENTIFIERS

Sample ID:	06-255	Museum ID:	ARC 25998
Field ID:	06-255	Collection Code:	
Deposited In:	Dalhousie University		

#### TAXONOMY

Phylum:	Chordata	Subfamily:	Gaidropsarinae
Class:	Actinopterygii	Genus:	Enchelyopus
Order:	Gadiformes	Species:	Enchelyopus cimbrius
Family:	Gadidae	Subspecies:	
BIN ID:	BOLD:AAB4061		

\* Barcode Index Numbers (BIN): Clustered barcode sequences that create OTUs (operational taxonomic units) closely reflective of species groupings

#### SPECIMEN DETAILS

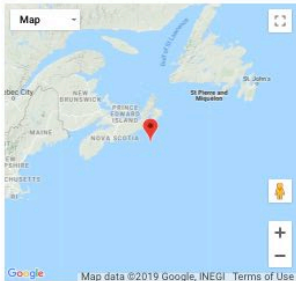
Voucher Status:	Reproduction:		
Tissue Descriptor:	Sex:		
Brief Note:	Life Stage:		
Detailed Notes:	Latitudes and Longitudes given here are averages. Range of lats and longs is 42.810N and 58.378W to 45.558N 63.204W		

#### COLLECTION DETAILS

Country:	Canada	Date Collected:	
Province/State:	Nova Scotia	Collectors:	
Region/County:			
Sector:			
Exact Site:	Eastern Scotian Shelf		
Latitude:	44.472	Elevation:	
Longitude:	-60.766	Elv. Accuracy:	
Coord. Source:		Depth:	
Coord. Accuracy:		Depth Accuracy:	

#### COLLECTION SITE

Map



Map data ©2019 Google, INEGI Terms of Use

#### ATtribution

Specimen Depository:	Dalhousie University
Sequencing Center:	Biodiversity Institute of Ontario
Photography:	CBG Photography Group, Centre for Biodiversity Genomics
Collectors:	
Specimen Identification:	Lou Van Guelpen
Project Manager:	Paul Bentzen
Sequencing Support:	

[http://www.boldsystems.org/index.php/Public\\_RecordView?processid=SCFAC445-06](http://www.boldsystems.org/index.php/Public_RecordView?processid=SCFAC445-06)



# PRESENT — LINKED TOOLS

**Data standards:** DarwinCore by TDWG

**Genetics:** BOLDsystems workbench

**Taxonomy:** FishBase to WoRMS to Colppensci

**Occurrences:** OBIS to GBIF

**Citizen Science:** iNaturalist to GBIF

**Literature:** WoRMs to BHL

**R packages:** rOpenSci to query all above



Many thanks for the leadership by the Chair, James Macklin, Agriculture and Agri-Food Canada



<http://www.boldsystems.org/index.php/>

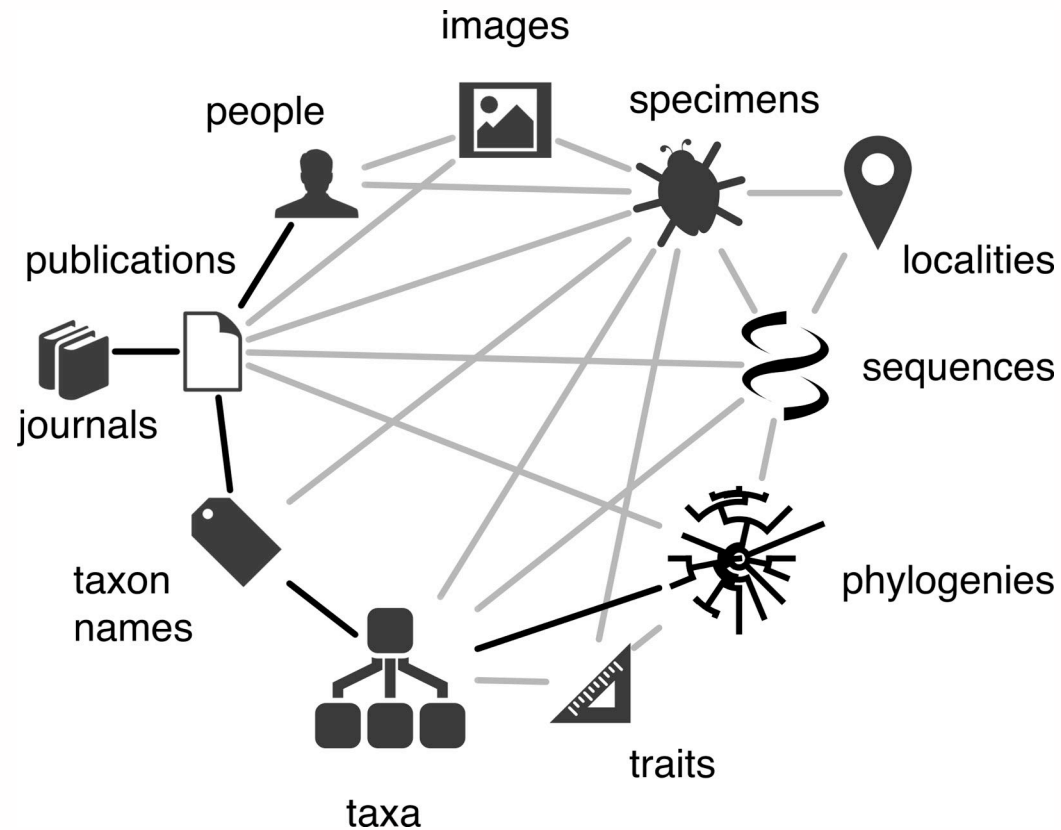


**Biodiversity  
Heritage  
Library**

<https://www.biodiversitylibrary.org/>



# PRESENT — LINKED DATA



Page R (2016) Towards a biodiversity knowledge graph. Research Ideas and Outcomes 2: e8767. <https://doi.org/10.3897/rio.2.e8767>

## The FAIR data principles

### FINDABLE

- F1. (meta)data are assigned a globally unique and eternally persistent identifier.
- F2. data are described with rich metadata.
- F3. (meta)data are registered or indexed in a searchable resource.
- F4. metadata specify the data identifier.

### ACCESSIBLE

- A1 (meta)data are retrievable by their identifier using a standardized communications protocol.
  - A1.1 the protocol is open, free, and universally implementable.
  - A1.2 the protocol allows for an authentication and authorization procedure, where necessary.
- A2 metadata are accessible, even when the data are no longer available.

### INTEROPERABLE

- I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
- I2. (meta)data use vocabularies that follow FAIR principles.
- I3. (meta)data include qualified references to other (meta)data.

### RE-USABLE

- R1. meta(data) have a plurality of accurate and relevant attributes.
  - R1.1. (meta)data are released with a clear and accessible data usage license.
  - R1.2. (meta)data are associated with their provenance.
  - R1.3. (meta)data meet domain-relevant community standards.

<https://www.force11.org/group/fairgroup/fairprinciples>

# PRESENT — REGIONAL REFERENCE NAME LIST

Using online linked tools and datasets, make a merged name lists

- removed synonyms, general taxon levels and errors

Resulted in about 2300 species:

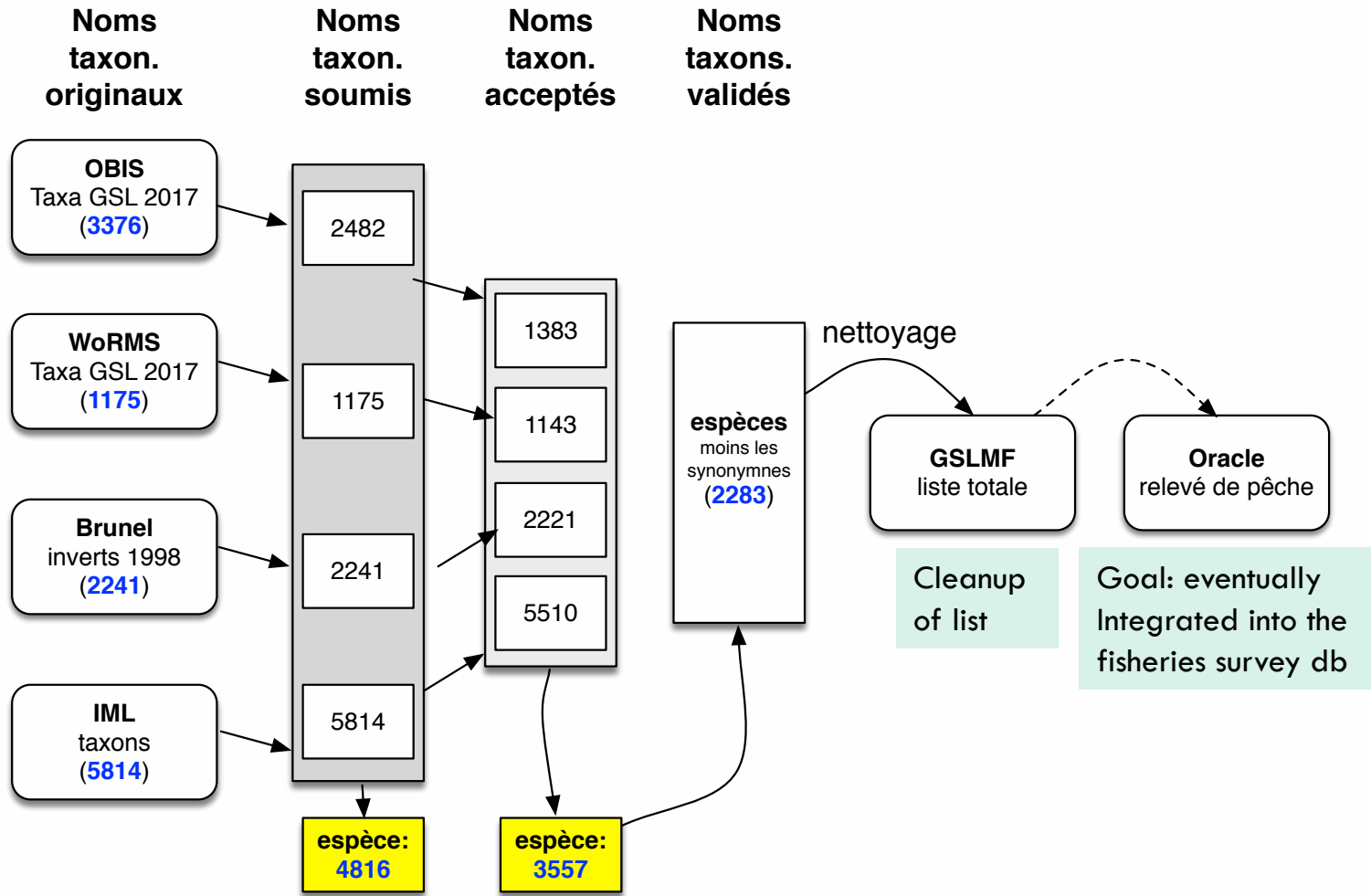
*Preliminary checklist of marine animal species of the Gulf of St. Lawrence, Canada, based on 4 sources (2017)*

See worksheet on ResearchGate:

<https://www.doi.org/10.13140/RG.2.2.10056.62727>

Can now compare this total species list with the BOLD species list (over 800) , to find taxa still in need of sequencing – globally, or regionally for the Gulf

# PRESENT – TAXON CHECKLIST WORKFLOW 2017



Phylum	Species
Arthropoda	758
Annelida	355
Mollusca	319
Chordata	226
Bryozoa	166
Cnidaria	157
Platyhelminthes	93
Echinodermata	68
Porifera	54
Nematoda	40
Acanthocephala	14
Nemertea	12
Chaetognatha	5
Sipuncula	5
Ctenophora	4
Brachiopoda	3
Entoprocta	2
Hemichordata	2
Cephalorhyncha	1
Phoronida	1
<b>Total</b>	<b>2285</b>

Note: 2 new taxa in 2017



# PRESENT — BARCODE LIBRAIRES

Barcoded taxa sometimes revealed surprises

**1 name = 1 genetic grouping (BIN)**

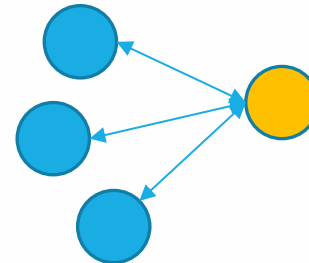
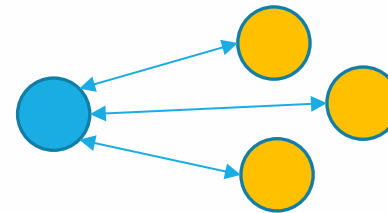
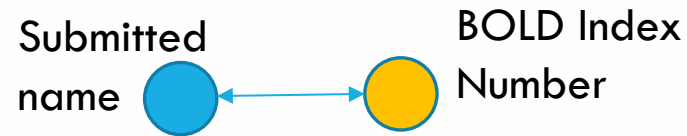
- Normal (expected)

**1 name = several groupings (BINs)**

- Suggestive of undiscovered species, possibly too difficult to identify traditionally

**Several names = 1 grouping (BIN)**

- Suggestive of mistaken identifications (estimates of up to 20% of submitted)



# PRESENT — DOUBLE VALIDATION

**Because of BOLD**, began to review several species and their distributions

**Different names, same group**

- Photos revealed misidentifications →
- Fisheries workers may have inadequate guides

**Same name, different groups and areas**

- Hidden species, not yet noticed
- Needed worldwide data to make it obvious

SPECIMEN IMAGES:



Not a polar sculpin, *Cottunculus microps*, but a snowflake sculpin, *Arctediellus uncinatus*

[http://www.boldsystems.org/index.php/Public\\_RecordView?processid=SCFAD142-09](http://www.boldsystems.org/index.php/Public_RecordView?processid=SCFAD142-09)

# PRESENT — COMPILE AND CONFIRM NAMES

Examine checklists for unusual patterns in distributions:

**Gulf of St. Lawrence marine fauna under review** (October 2019)

<https://www.doi.org/10.13140/RG.2.2.28150.75846>

**= 94 names** that merit further scrutiny

Then evaluate list with online tools for:

- 1) taxonomic status (WoRMS)
- 2) genetic sequences (BOLD)
- 3) historical references (BHL)

Examples of groups under review as a result of NW Atlantic checklists:

sponges, soft corals, tunicates,  
shrimps, crabs, amphipods,  
echinoderms, whelks, polychaetes,  
fishes (commercial & non-commercial,  
pelagic & demersal)

# PRESENT — DISCOVERY WITH GENETICS & MAPS



*Liparis bathyarcticus* Parr, 1931

**not** *Liparis gibbus* (Arctic), **nor** *Liparis liparis* (European)

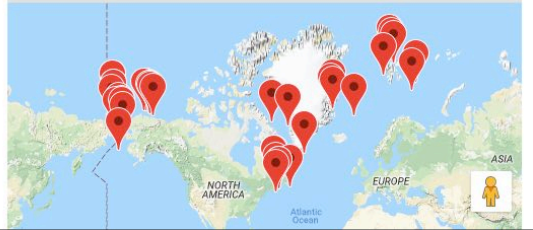
- Determined in 2013, distinguishing it from *L. gibbus*, as it was called up until then, and earlier from *L. liparis*

<http://www.marinespecies.org/aphia.php?p=taxdetails&id=867958>

Phylum:	Chordata [55]	🔍
Class:	Actinopterygii [55]	🔍
Order:	Scorpaeniformes [55]	🔍
Family:	Liparidae [55]	🔍
Subfamily:		
Genus:	Liparis [55]	🔍
Species:	Liparis bathyarcticus [53]	🔍
	Liparis gibbus [1]	🔍
	Liparis tunicatus [1]	🔍
Subspecies:		

COLLECTION SITE:

Map



[http://www.boldsystems.org/index.php/  
Public\\_BarcodeCluster?clusteruri=BOLD:  
AAB4896](http://www.boldsystems.org/index.php/Public_BarcodeCluster?clusteruri=BOLD:AAB4896)



# PRESENT — REDISCOVERY WITH GENETICS & MAPS

Two common NW Atlantic species, forgotten in the literature for decades,  
confirmed with local specimens and sequences



***Eumicrotremus terraenovae* Myers & Böhlke, 1950**

**not** *Eumicrotremus spinosus* (European)

<http://www.marinespecies.org/aphia.php?p=taxdetails&id=159521>



***Myxine limosa* Girard, 1859**

**not** *Myxine glutinosa* (European)

<http://www.marinespecies.org/aphia.php?p=taxdetails&id=271309>

# PRESENT — REDISCOVERY WITH PHOTOS & MAPS

**Reviewed checklist** of shrimps following a recent Gulf of St. Lawrence study (Tamdrari et al. 2018)  
<https://doi.org/10.3354/meps12566>

Were **no records** in recent decades of a common **warmwater** species, *Dichelopandalus leptoceros*

**Online photos** revealed the **similarity** in colour and striping with **coldwater** *Pandalus montagui*

Review of captures in 2017 **confirmed** presence in historical range of Canada (e.g., Murillo et al. 2018): *D. leptoceros* was being misidentified as *P. montagui*

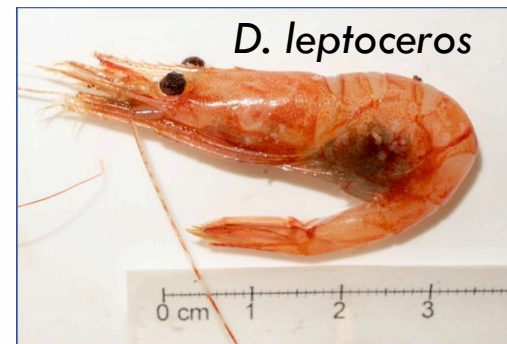
[https://www.researchgate.net/publication/325079607\\_Cruise\\_Report\\_for\\_the\\_CCGS\\_Alfred\\_Needler\\_Maritimes\\_Region\\_Research\\_Vessel\\_Summer\\_Multi-species\\_Survey\\_June\\_28\\_to\\_August\\_14\\_2017\\_Benthic\\_Invertebrates/citations](https://www.researchgate.net/publication/325079607_Cruise_Report_for_the_CCGS_Alfred_Needler_Maritimes_Region_Research_Vessel_Summer_Multi-species_Survey_June_28_to_August_14_2017_Benthic_Invertebrates/citations)

**Absence** is now confirmed for the **Northern Gulf**, and **presence** renewed for **Southern Gulf**

- If *P. montagui* were to decline with warming water, we can now **distinguish** this trend from the presence or range expansion of *D. leptoceros*
- No public sequence yet — needs to be barcoded



<https://www.gbif.org/species/2222507>



<http://www.marinespecies.org/carms/aphia.php?p=image&tid=158356&pic=135462>



<http://www.marinespecies.org/carms/photogallery.php?album=2016&pic=39062>

# NAME CHANGES, RECORD ERRORS...

So what? At least a codfish is still  
a codfish, right?

***Gadus morhua*** – Atlantic Cod

***Gadus macrocephalus*** – Pacific Cod

***Gadus ogac*** – Greenland Cod (Arctic & Atlantic)

Also:

***Boreogadus saida*** - Arctic Cod

***Arctogadus glacialis*** – Polar Cod

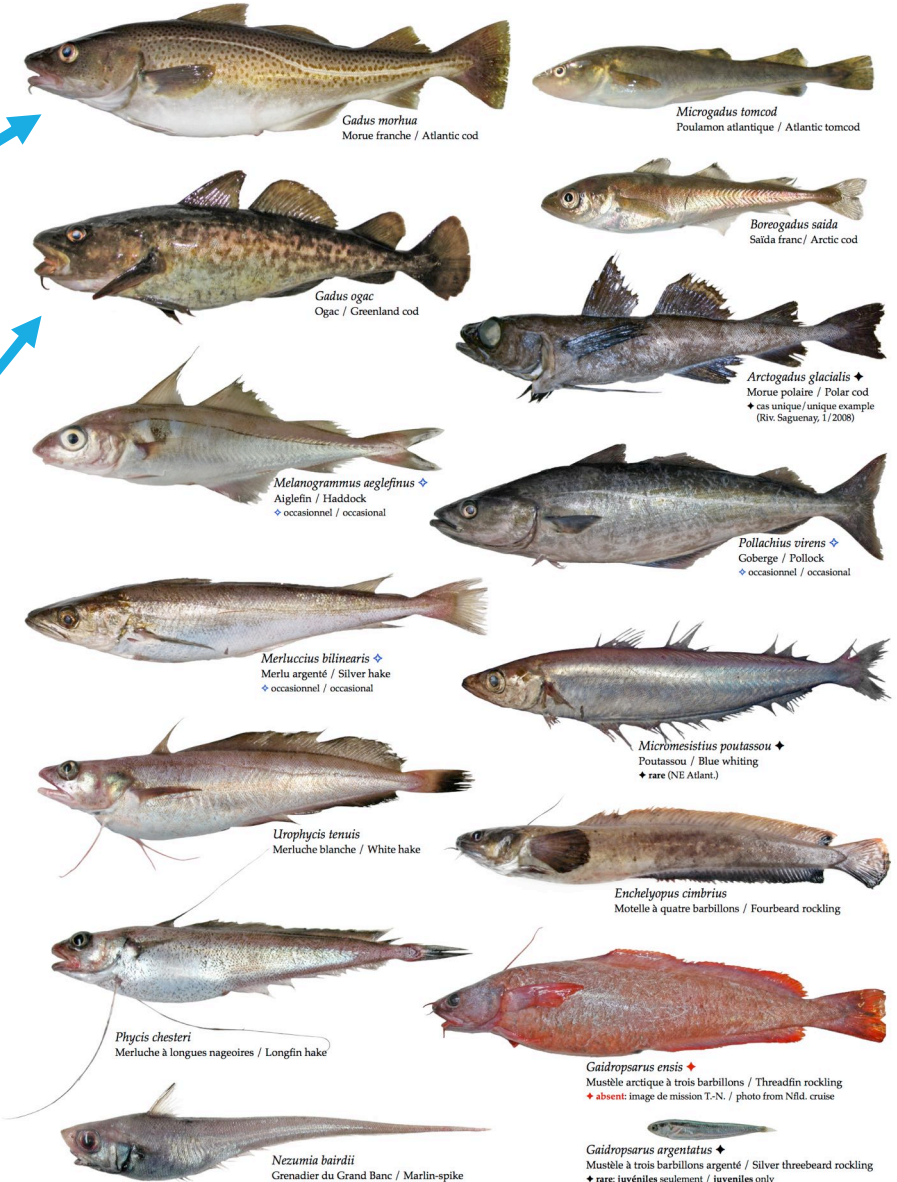
Straightforward, yet easy to mix up...

GADIFORMES

Morues du Saint-Laurent  
provenant des missions (2006-2007)



Codfishes of the St. Lawrence  
as seen on missions (2006-2007)





# WHAT NAME AM I?

To a **local user**, a cod remains the **same**  
To an analyst with **global** data, 'It's a trap!'

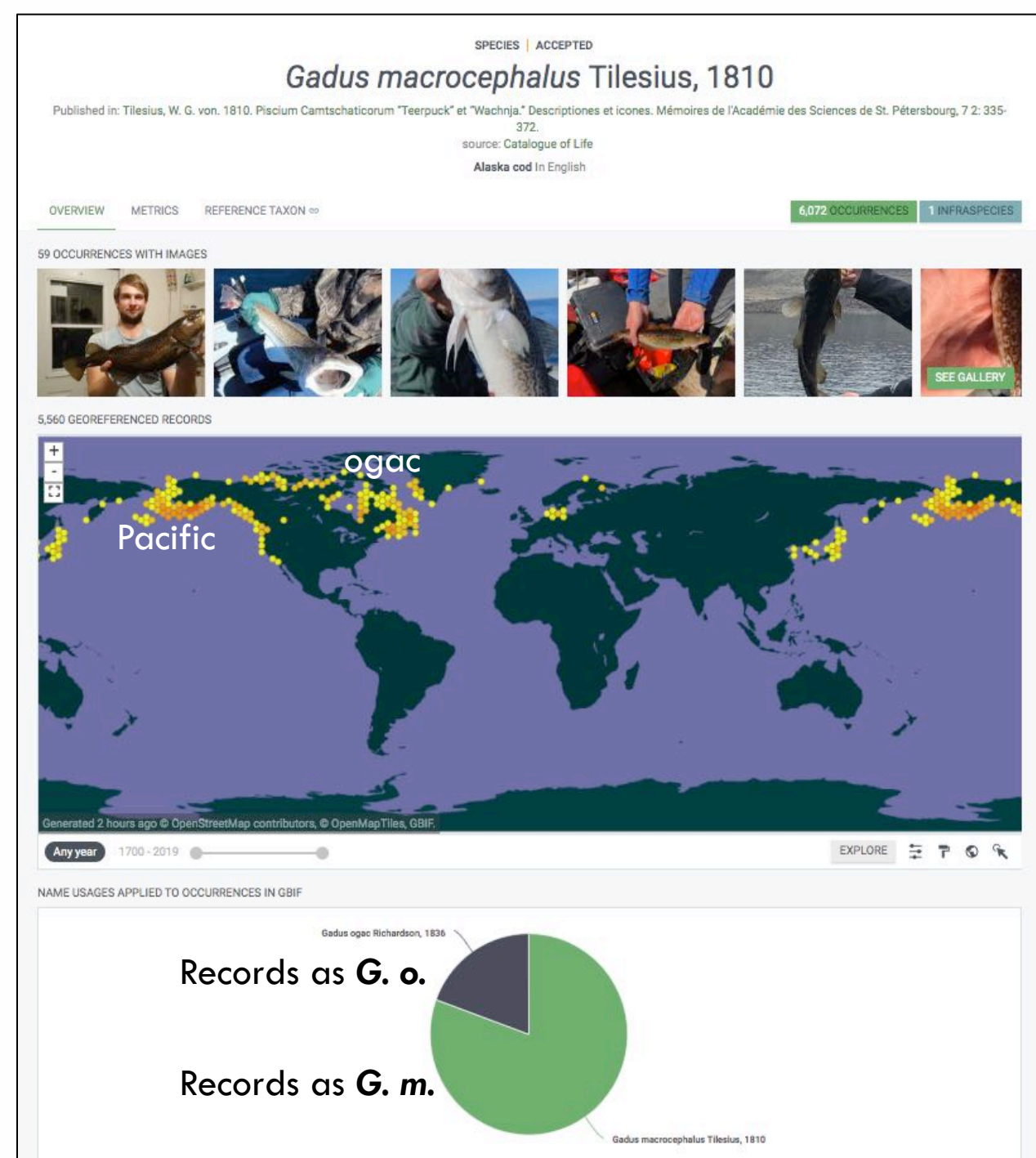
***Gadus ogac*** (Greenland Cod) is now accepted  
as same species as ***Gadus macrocephalus***  
(Pacific Cod) = 2 names for 1 species

Yet, very different distribution and ecology

Scenario: recent records show that Pacific Cod  
sequences are detected in the Arctic, which  
was not observed before


Conclusion: Pacific Cod has recently migrated  
to the Arctic, possibly due to climate change

<https://www.gbif.org/species/2415853>






# PRESENT — ISSUES IN TAXONOMY ACROSS REGIONS



**Polar Biology**  
November 2017, Volume 40, Issue 11, pp 2291–2296 | [Cite as](#)

## First records of Pacific sand lance (*Ammodytes hexapterus*) in the Canadian Arctic Archipelago

Authors [Authors and affiliations](#)

Marianne Falardeau , Caroline Bouchard, Dominique Robert, Louis Fortier

Short Note  
First Online: 17 June 2017

3 Shares 313 Downloads 2 Citations

### Abstract

An increasing number of boreal marine species are expected to invade the warming Arctic Ocean with the potential to displace endemic species. We provide first evidence that Pacific sand lance (*Ammodytes hexapterus*) is expanding its range in the Canadian Arctic Archipelago, a region far outside the species temperate-boreal traditional range south of the Bering Strait. To the best of our knowledge, supported by local Inuit knowledge, the species was not present in the area until the present decade. We observed an increasing density of larval Pacific sand lance with time over the 2011–2016 period, suggesting that environmental conditions are becoming increasingly favorable for the species to reproduce in the Central Canadian Arctic. The northward distribution change of Pacific sand lance is occurring earlier than predicted by current models and could trigger abrupt shifts in Arctic marine food webs if the boreal invader displaces polar cod, a key prey species for top predators in Arctic marine ecosystems.

### Keywords

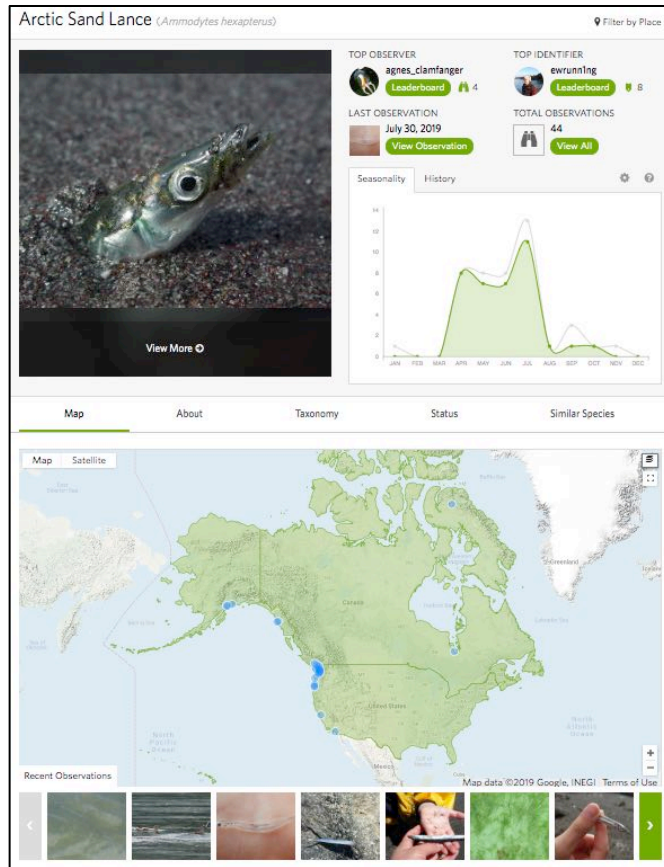
Sandeel Northwest Passage Climate-related invasion Climate change Fish communities Range shift

Through genetics, a Pacific fish species was detected as expanding its distribution into the Arctic...

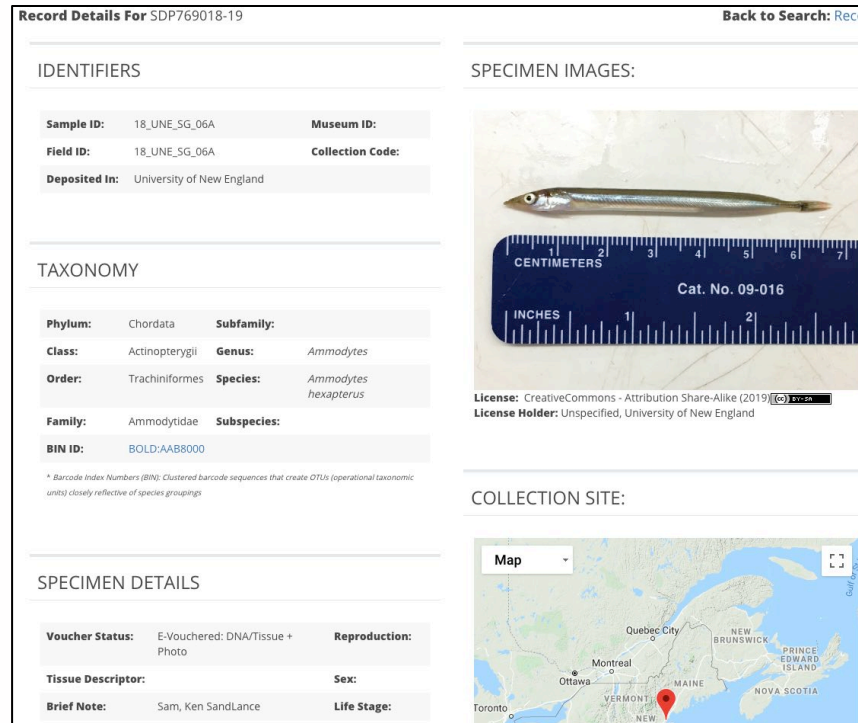
...or is this a real-life case of unnoticed changes in names and taxa?

Falardeau et al. (2017) Polar Biol. 40: 2291–2296  
<https://doi.org/10.1007/s00300-017-2141-0>

# PRESENT — A 'PACIFIC FISH', IN ARCTIC & ATLANTIC



<https://www.inaturalist.org/taxa/130238-Ammodytes-hexapterus>



<http://www.boldsystems.org/index.php/PublicRecordView?processid=SDP769018-19>

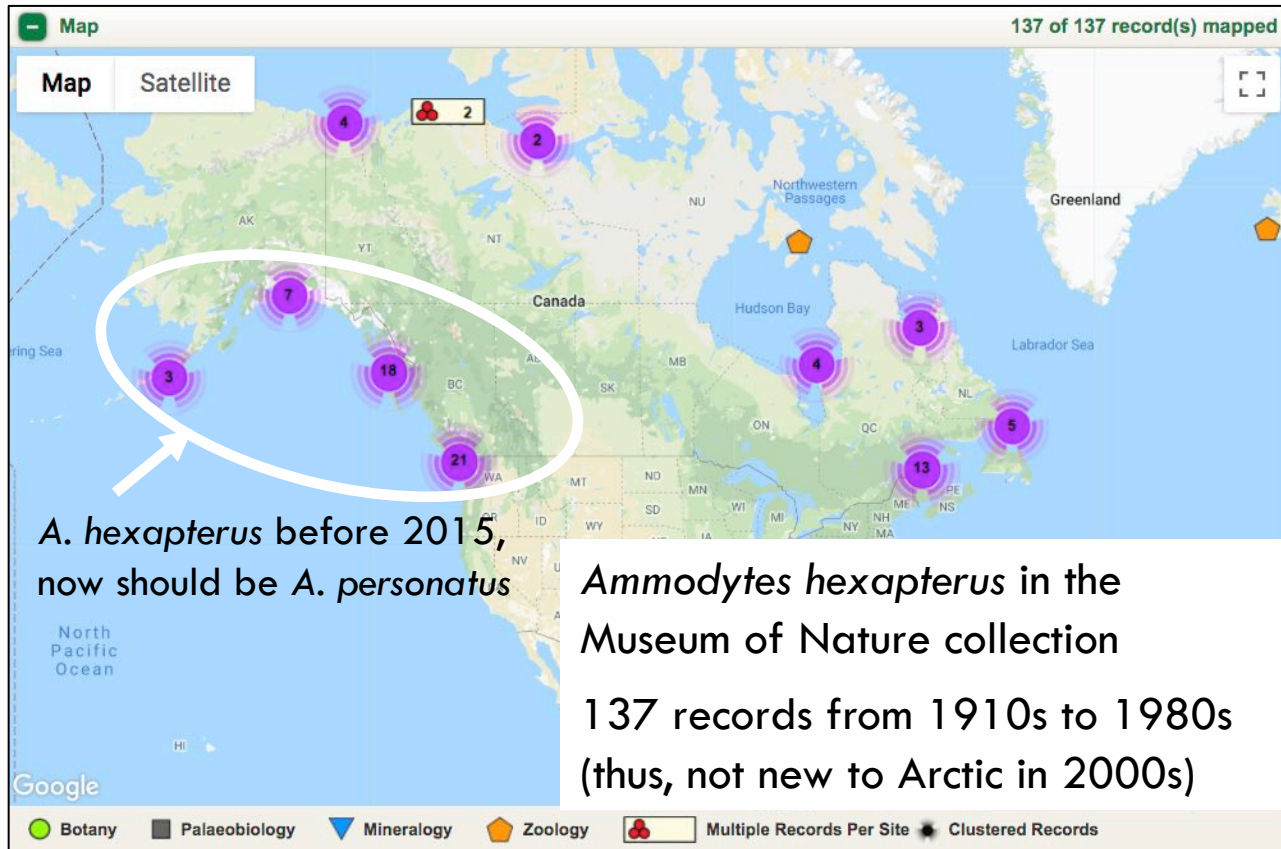
In a revision by Orr et al. (2015), *Ammodytes hexapterus* is now the **Arctic Sand Lance**, and the **Pacific Sand Lance** becomes *A. personatus* <https://spo.nmfs.noaa.gov/sites/default/files/orr.pdf>

Depending on resource consulted, may **still** see *A. hexapterus* listed as a Pacific Sand Lance

While *A. hexapterus* is not a new species to the Arctic, the analyses by Falardeau et al. (2017) do now give us better distribution and genetics data for the region (Mecklenburg et al. 2018)

<https://oaarchive.arctic-council.org/handle/11374/2116>

# PRESENT — SUMMARY: BE WARY WITH BIG DATA



The choice of record, name, sequence, or taxon concept used in analyses would change conclusions for ecosystem monitoring

<http://collections.nature.ca/en/Search/Search>



# NEXT — DATA, FROM ARTISANAL TO INDUSTRIAL

**Large-scale &  
Low-cost  
Sequencing**

**Sampling &  
Analyses**

**Genetics &  
Ecology**



Biosciences have traditionally not been part of the megascience club. But it very much should be, says Paul Hebert at [#biodiversitynext](#). Especially when facing the current [#biodiversity](#) decline. We need industrial scale analysis of life on earth.



<https://twitter.com/eurotaxonomy/status/1186572924100272128>  
5:20 AM · Oct 22, 2019 · Twitter for Android

## Guelph professor preserving life for future generations through \$180M DNA barcoding project

BIOSCAN, a seven-year barcoding project involves 1,000 researchers from more than 30 countries is another step towards a 20-year project in 2026 to complete the inventory of life on planet Earth

Jun 25, 2019 5:00 PM By: Anam Khan



<https://www.guelphtoday.com/local-news/guelph-professor-preserving-life-for-future-generations-through-180m-dna-barcoding-project-1539497>



# BUILDING ON BOLD

**BOLD SYSTEMS** DATABASES IDENTIFICATION TAXONOMY WORKBENCH RESOURCES LOGIN Q

## BARCODE OF LIFE DATA SYSTEMS v4 beta

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.

Please note that this version of BOLD is in beta and will contain bugs. Users can help address these bugs by testing the system and reporting issues to [support@boldsystems.org](mailto:support@boldsystems.org). This version is very different from the prior one but has access to all the same data.

#### DATA PORTAL

A data retrieval interface that allows for searching over 1.7M public records in BOLD using multiple search criteria including, but not limited to, geography, taxonomy, and depository.

#### EDUCATION PORTAL

A custom platform for educators and students to explore barcode data and contribute novel barcodes to the BOLD database.

#### BIN DATABASE

A searchable database of Barcode Index Numbers (BINs), sequence clusters that closely approximate species.

#### WORKBENCH

A data collection and analysis environment that supports the assembly and validation of DNA barcodes and other sequences.

## Reference Library Storage

Adapted from: Sujeewan Ratnasingham et al. (2019) The mBRAVE Platform: Automated NGS/HTS Data Analysis. <https://doi.org/10.3897/biss.3.37986>. Presented at Biodiversity\_Next, October 2019, Leiden

## mBRAVE

### Multiplex Barcode Research And Visualization Environment

mBRAVE is a multi-user platform supporting the storage, validation, analysis, and publication of highly multiplexed projects based on high-throughput sequencing (HTS) instruments. This system builds on the [BOLD Platform](#) to support species identification and discovery for HTS data.

**Login** **Register**

mBRAVE is now accepting all new registrants.

## Reference Library Use/Generation

# NEXT — INDUSTRIAL + ARTISANAL



**David Shorthouse**

@dpsSpiders

Manager, Biodiversity Data / Agriculture and Agri-Food Canada. Opinions mine and not those of my employer.

📍 Ottawa, Ontario 🔗 [orcid.org/0000-0001-7618...](https://orcid.org/0000-0001-7618...) 📅 Joined June 2009

317 Following 683 Followers



**David Shorthouse**

@dpsSpiders

At [#BiodiversityNext](#) a word was used to describe traditional, alpha taxonomy that's been bothering me: artisanal. While that may be true, it's precisely because of the quality of the craft, its potential for connectivity, that massively scalable, collaborative products are made.

7:43 PM · Oct 25, 2019 · [Twitter Web App](#)

<https://twitter.com/dpsSpiders/status/1187877258809237505>

# NEXT — COMPLEMENTARY TOOLS & TEAMS



“Once a taxon has been firmly identified by taxonomic experts and its barcode sequence has been deposited in GenBank or BOLD, eDNA might eventually reduce the need for large teams of expert taxonomists to carry out routine biodiversity monitoring. Yet, the routine application of metabarcoding for Arctic monitoring requires overcoming various limitations. For example, here the eDNA metabarcoding identified *Acartia tonsa*, a potential invader that has been previously recorded in the ecoregions of ports connected to Churchill (Chan et al., 2012). However, the current available COI sequences for *Acartia tonsa* form several distinct clades, some of which cluster with *Acartia hudsonica*, raising the possibility that the eDNA sequences assigned to *A. tonsa* actually belong to the native *A. hudsonica*. **Thus, taxonomic expertise remains crucial for reducing biases of species distributions related to increasing use of large-scale eDNA metabarcoding**” p. 7772, Lacoursière-Roussel et al. (2018) <https://doi.org/10.1002/ece3.4213>



# NEXT — COLLECTIONS MANAGEMENT

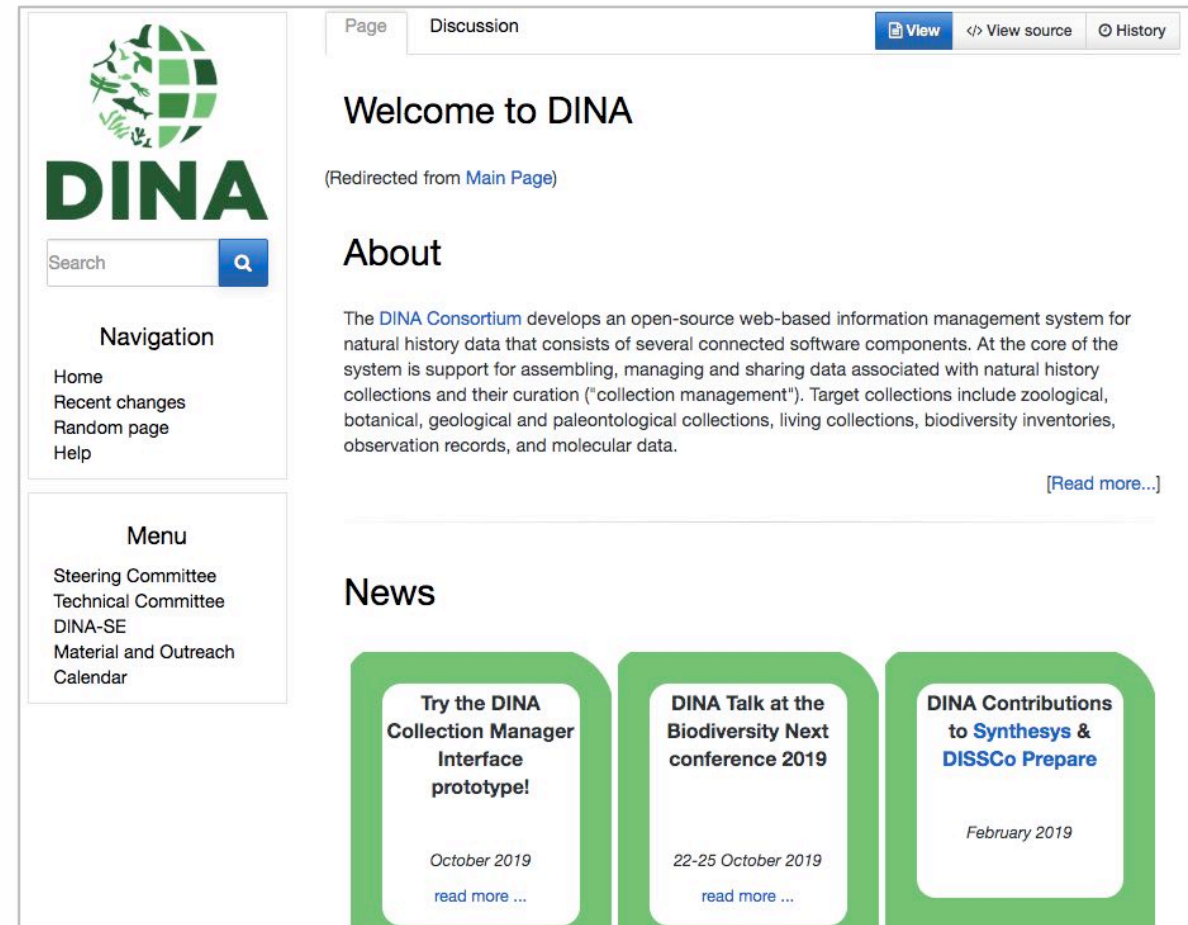
## DINA: Open Source, Open Services

A project to **modernize** the management of biological specimens, collections, media, and sequences

**Modular**, web-based software

Integrated with **biological standards**

**International consortium**, led by Falko Glöckler (Nature Museum Berlin) and James Macklin (Agriculture and Agri-Food Canada)



[https://www.dina-project.net/wiki/Welcome to DINA](https://www.dina-project.net/wiki/Welcome_to_DINA)

# A VALIDATED GENETIC REFERENCE DATABASE FOR MARINE ANIMALS OF THE NORTHWEST ATLANTIC OCEAN – SUMMARY

**Review names** to ensure we work with accurate baselines

**Prepare to examine** patterns that may be uncovered during the review

**Take small steps** while reviewing groups and building a reference database of names and sequences

**Be ready to accommodate** change when faced with new information

**Publish** data and **cite** data (reveal your findings and learn from others)



[https://imgs.xkcd.com/comics/desert\\_island.png](https://imgs.xkcd.com/comics/desert_island.png)



# A VALIDATED GENETIC REFERENCE DATABASE FOR MARINE ANIMALS OF THE NORTHWEST ATLANTIC OCEAN – **SUMMARY**

We are better equipped to study regional species by validating data on: **names, specimens, photos** – and now with **sequences**

- Developing and maintaining this work requires effort
- Online resources make it easier (work smarter, not harder)
- Future: perform large-scale analyses (find more patterns faster)
- Will still need experts to review results

# NEXT — DFO BIODIVERSITY REFERENCE BASES

## Working group

- Inform personnel to available online resources

## Workshops (online or physical)

- Training with linked data – using web, R, and GitHub

## Publish more of what we know so far

- Regional photo catalogues
- Regional checklists
- Regional sequences

## Receive credit

- **Cite** publications and datasets using DOIs (Digital Object Identifiers)
- Post to tracking services: **ORCID** (author ID), **Bloudhound** (specimens)  
<https://orcid.org/>    <https://bloudhound-tracker.net/>



Biodiversity Next - Building a global infrastructure for biodiversity data.  
**Together.** (Leiden, Oct. 2019)

<https://biodiversitynext.org/>