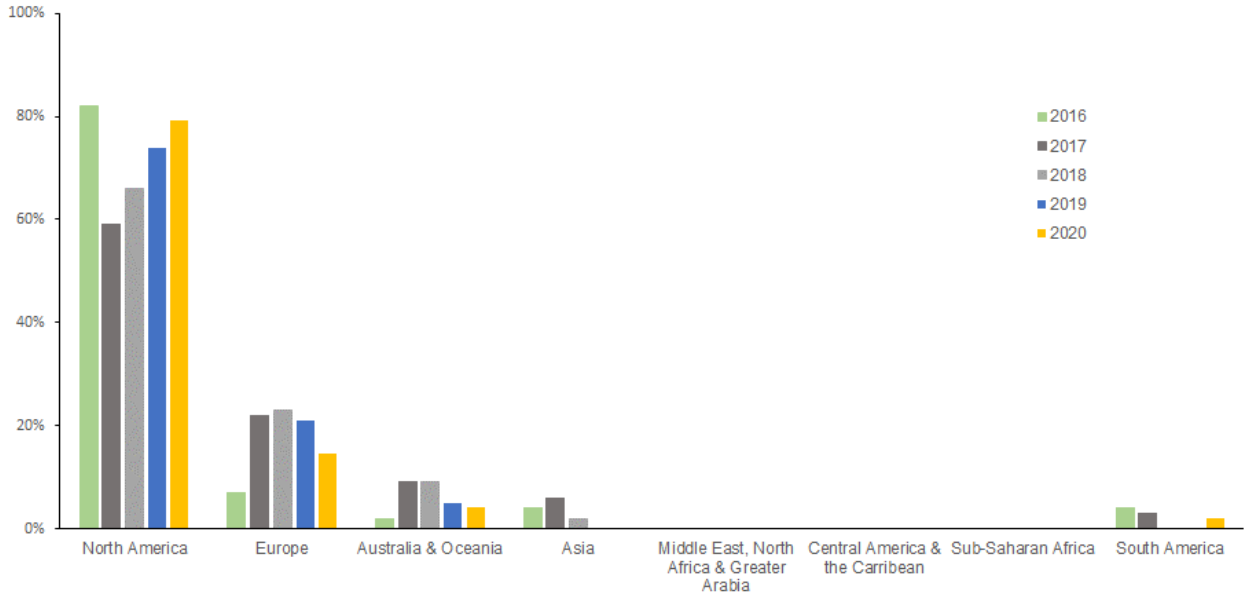
IWEG REPORT 5: APPENDICES

Contents

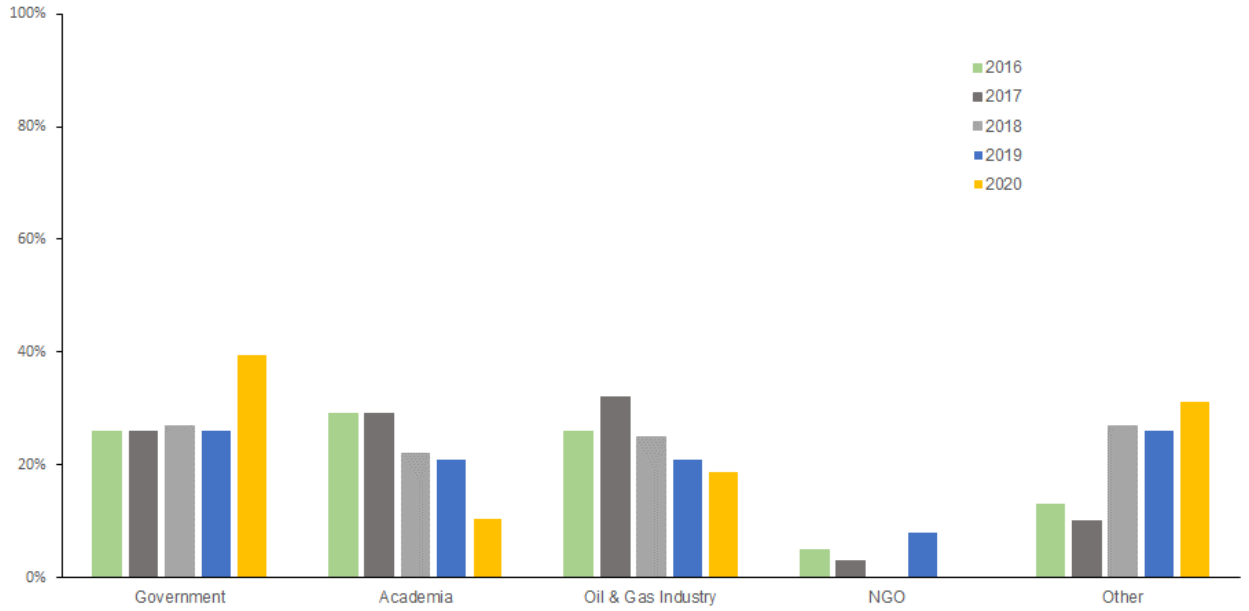
APPENDIX A.

List of questions posed to IWEG participants over the last five years. Responses are summarized in bar graphs after each question.

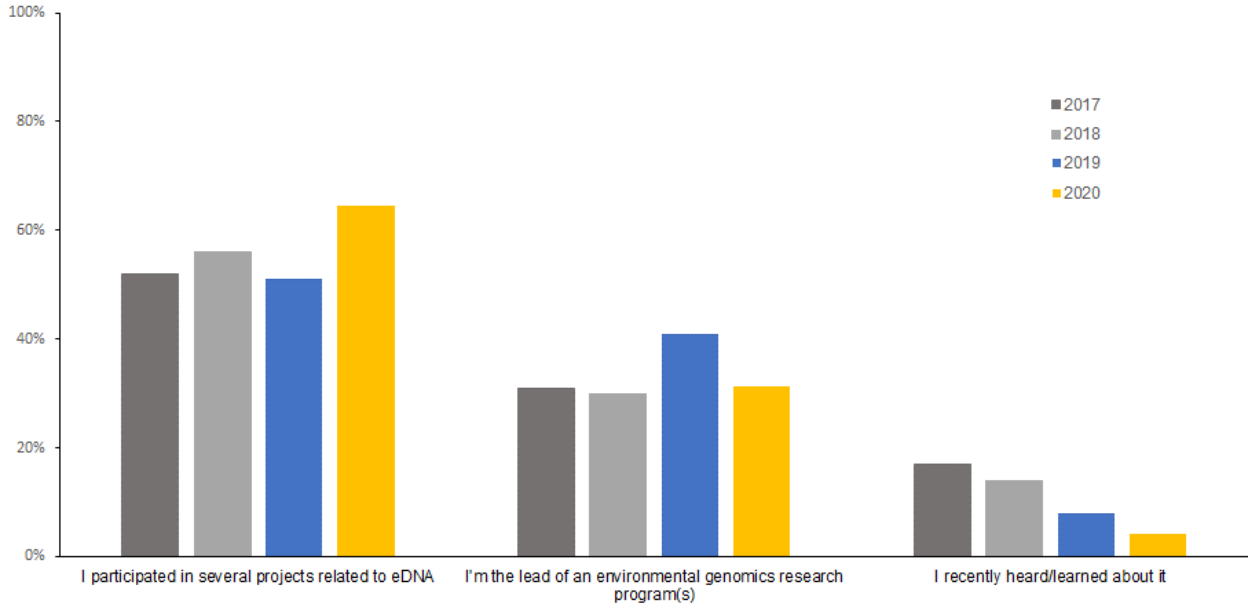
WHICH REGION ARE YOU FROM?



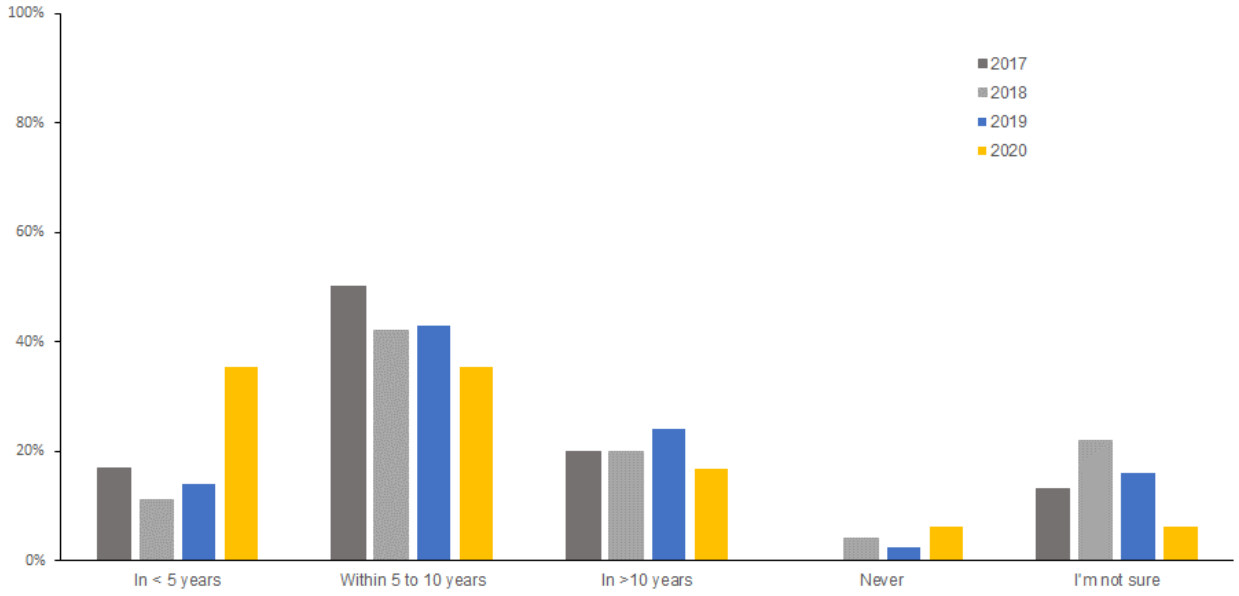
WHAT GROUP DO YOU REPRESENT?



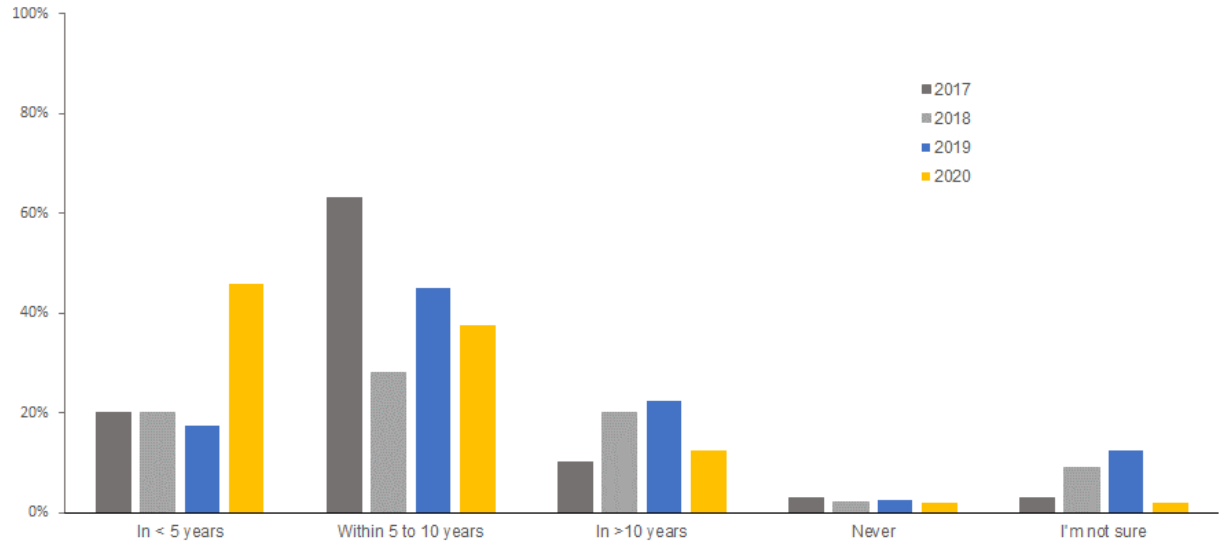
HOW FAMILIAR ARE YOU WITH "ENVIRONMENTAL GENOMICS"?



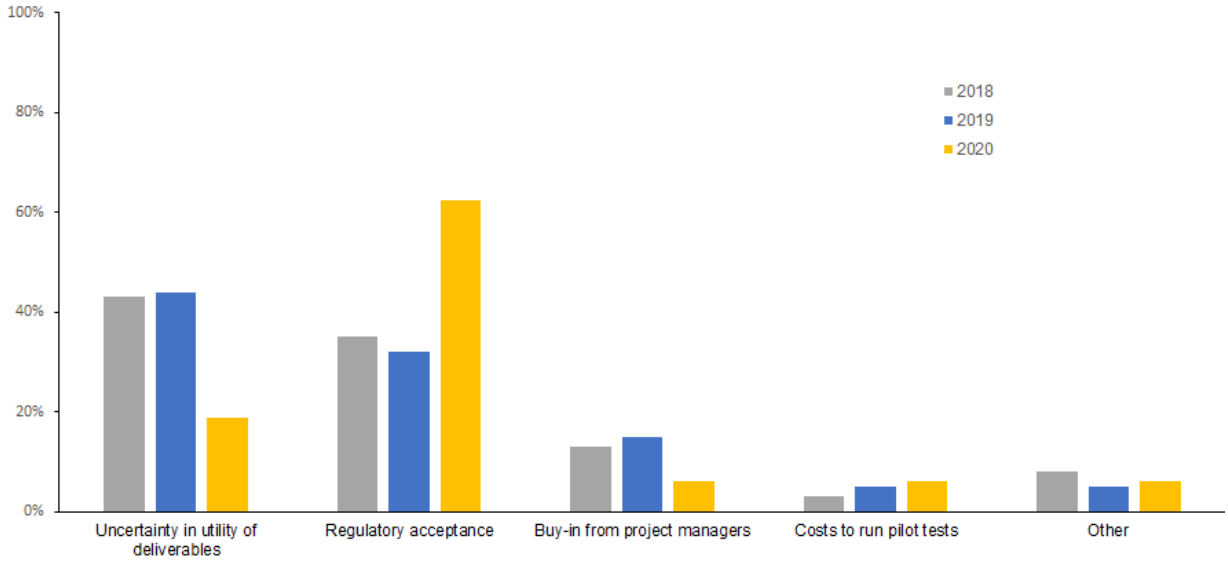
WHEN DO YOU SEE ENVIRONMENTAL GENOMICS BECOMING THE PRIMARY APPROACH TO CHARACTERIZE AND MONITOR BIODIVERSITY IN THE *TERRESTRIAL* ENVIRONMENT?



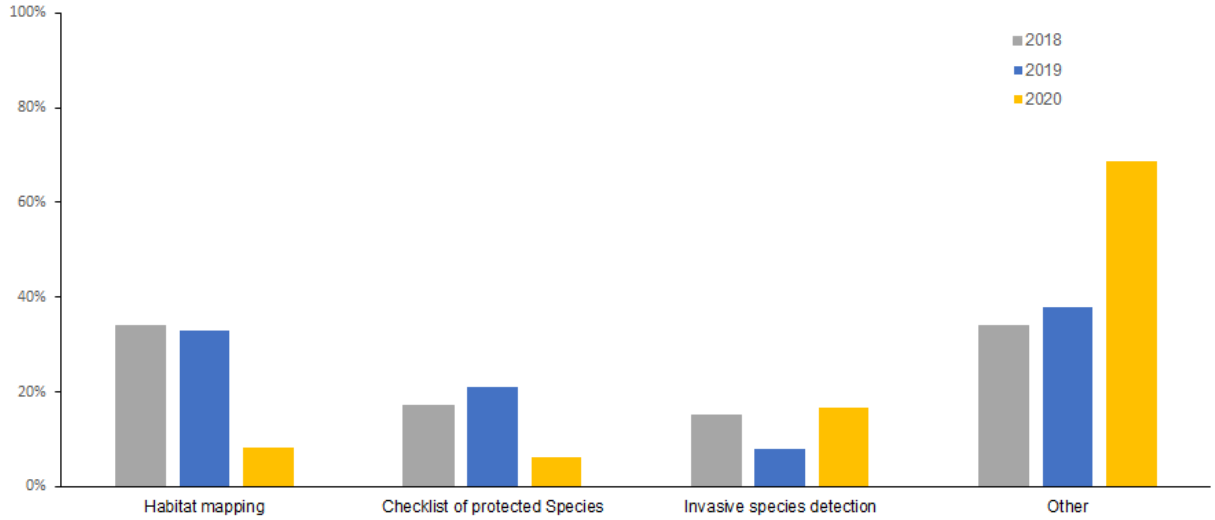
WHEN DO YOU SEE ENVIRONMENTAL GENOMICS BECOMING THE PRIMARY APPROACH TO CHARACTERIZE AND MONITOR BIODIVERSITY IN THE *MARINE/AQUATIC* ENVIRONMENT?



WHAT ARE THE BIGGEST BARRIERS TO INTRODUCING ENVIRONMENTAL GENOMICS IN NEW AND EXISTING ENVIRONMENTAL ASSESSMENT AND MONITORING EFFORTS?

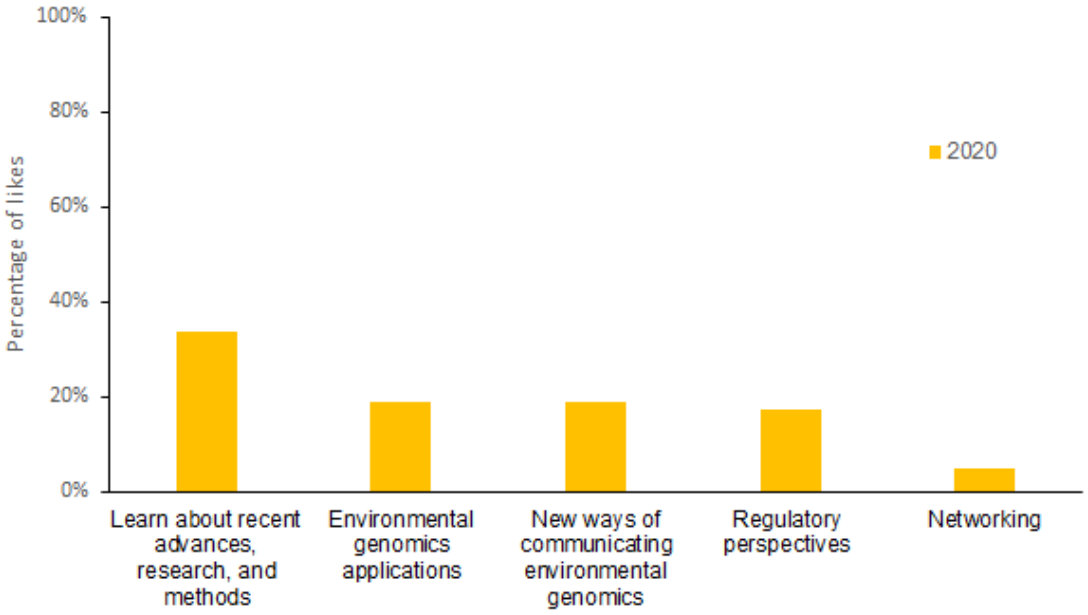


FROM A REGULATORY PERSPECTIVE, WHAT IS THE HIGHEST PRIORITY?



WHAT DO YOU HOPE TO GET OUT OF THIS WORKSHOP?

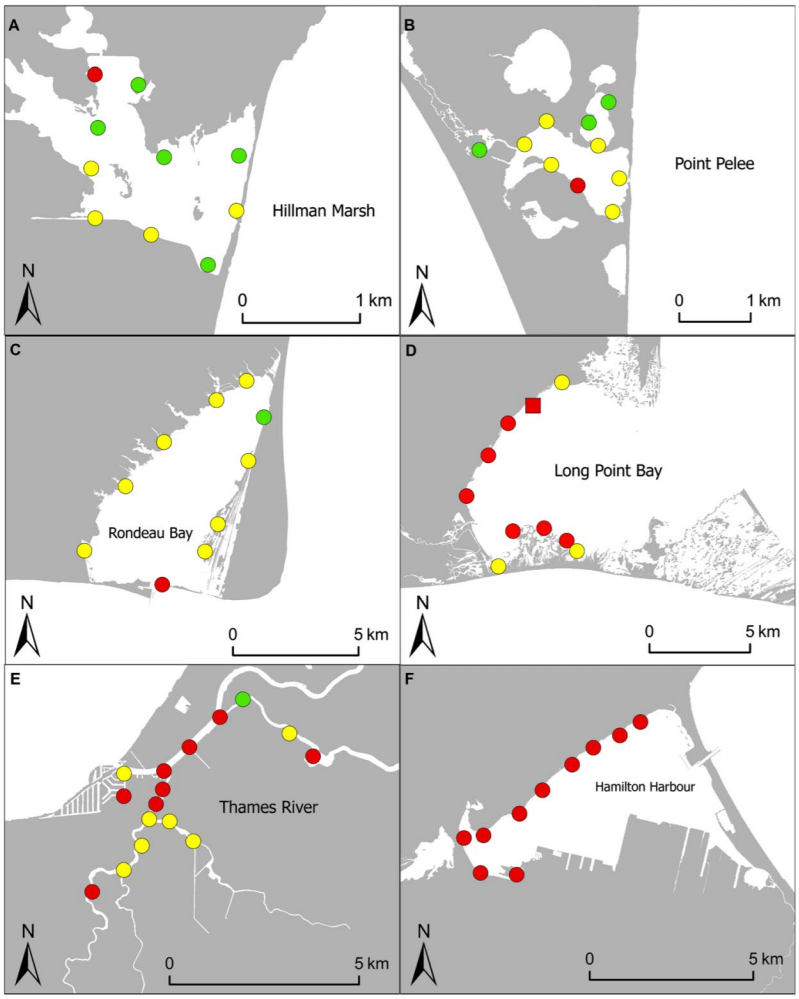
Note that this was an open question, and we report the top five responses based on "likes" from the audience.



APPENDIX B.

EXAMPLE FIGURES USED DURING THE INTERACTIVE SESSION, DETERMINING BEST PRACTICES FOR DATA VISUALIZATION, DURING IWEG 2020.

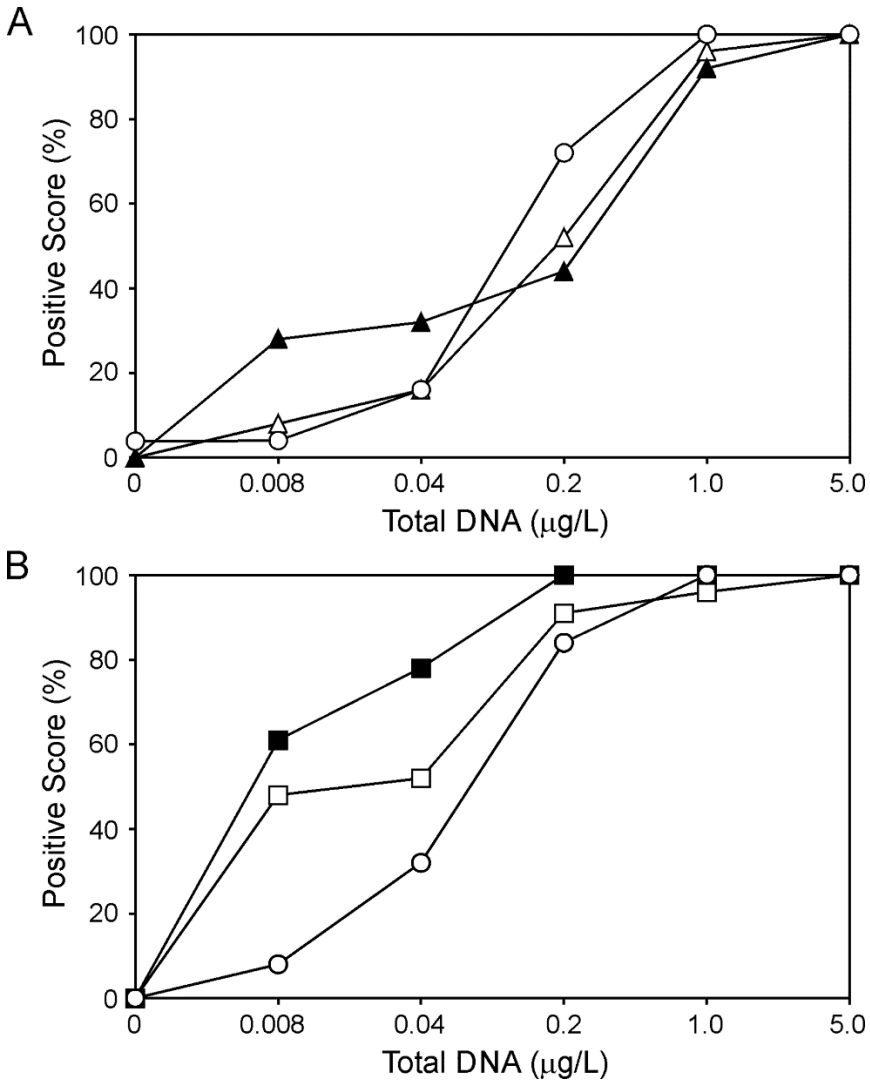
TARGETED DETECTION: MAPS



Maximum qPCR detections of spotted gar eDNA within sampling locations. (A) Hillman Marsh, (B) Point Pelee, (C) Rondeau Bay, (D) Long Point Bay, (E) Thames River, and (F) Hamilton Harbour with detections categorized by samples yielding starting DNA copy numbers greater than 100 copies/5 µL in at least one sample (green), starting copy numbers greater than 10 copies/5 µL and less than 100 copies/5 µL in at least one sample (yellow), and no detection or starting copy numbers less than 10 copies/5 µL in all samples (red). The square symbol in Long Point Bay indicates the location where live spotted gar were captured.

FROM: Boothroyd, M., Mandrak, N. E., Fox, M., & Wilson, C. C. (2016). Environmental DNA (eDNA) detection and habitat occupancy of threatened spotted gar (Lepisosteus oculatus): eDNA Detection of Endangered Species. Aquatic Conservation: Marine and Freshwater Ecosystems, 26(6), 1107–1119. doi: 10.1002/aqc.2617

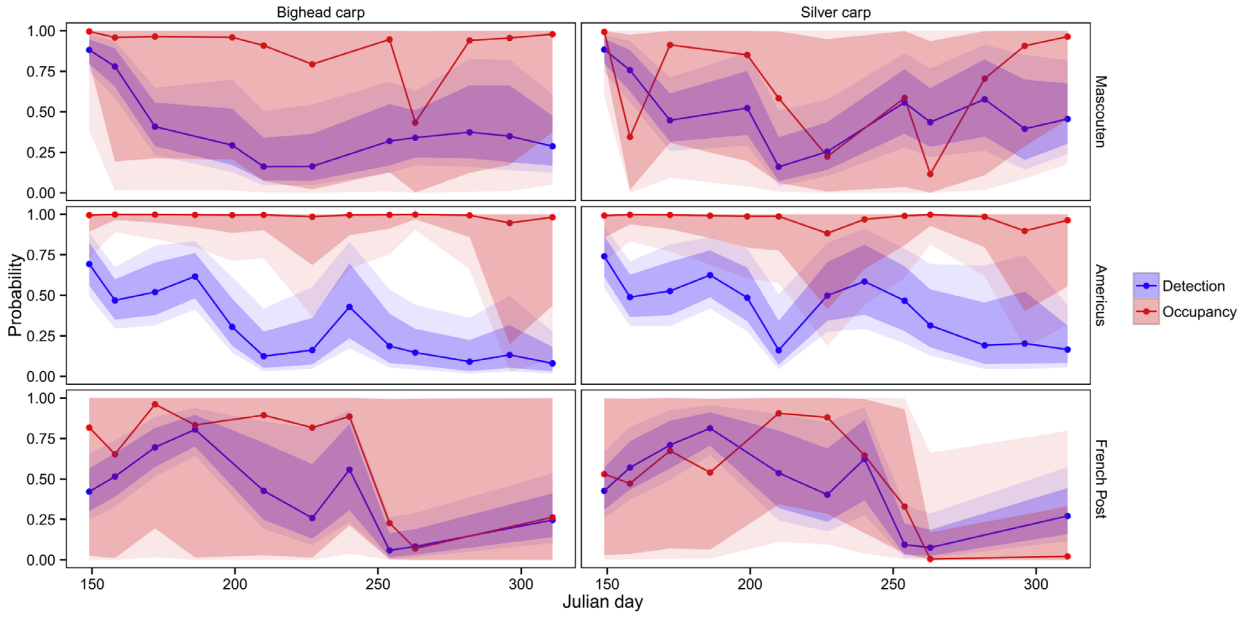
TARGETED DETECTION: ASSAY SENSITIVITY PLOTS



Sensitivity of species-specific and animal group qPCR primer sets for (A) bullfrog or (B) tailed frog determined using a concentration range of total DNA. A 5-fold dilution series (0.008, 0.04, 0.2, 1.0 and 5.0 μg/L) of bullfrog frog total DNA was assessed in the qPCR assay against eLICA1 (white triangle), eLICA2 (black triangle), and eFrog3 (white circle) primer sets in Panel A. Panel B shows the results of qPCR assays against eASMO (white square), eASMO9 (black square), and eFrog3 (white circle) primer sets using tailed frog total DNA using the same 5-fold dilution series (0.008, 0.04, 0.2, 1.0 and 5.0 μg/L). The percentage of reactions demonstrating detection following 50 cycles compared to the total reactions performed (n = 23–26 technical replicates) is shown. Negative control reactions containing no DNA template displayed no positive detection score with any of the assays with the sole exception of eFrog3 at 0.04% for the bullfrog template only.

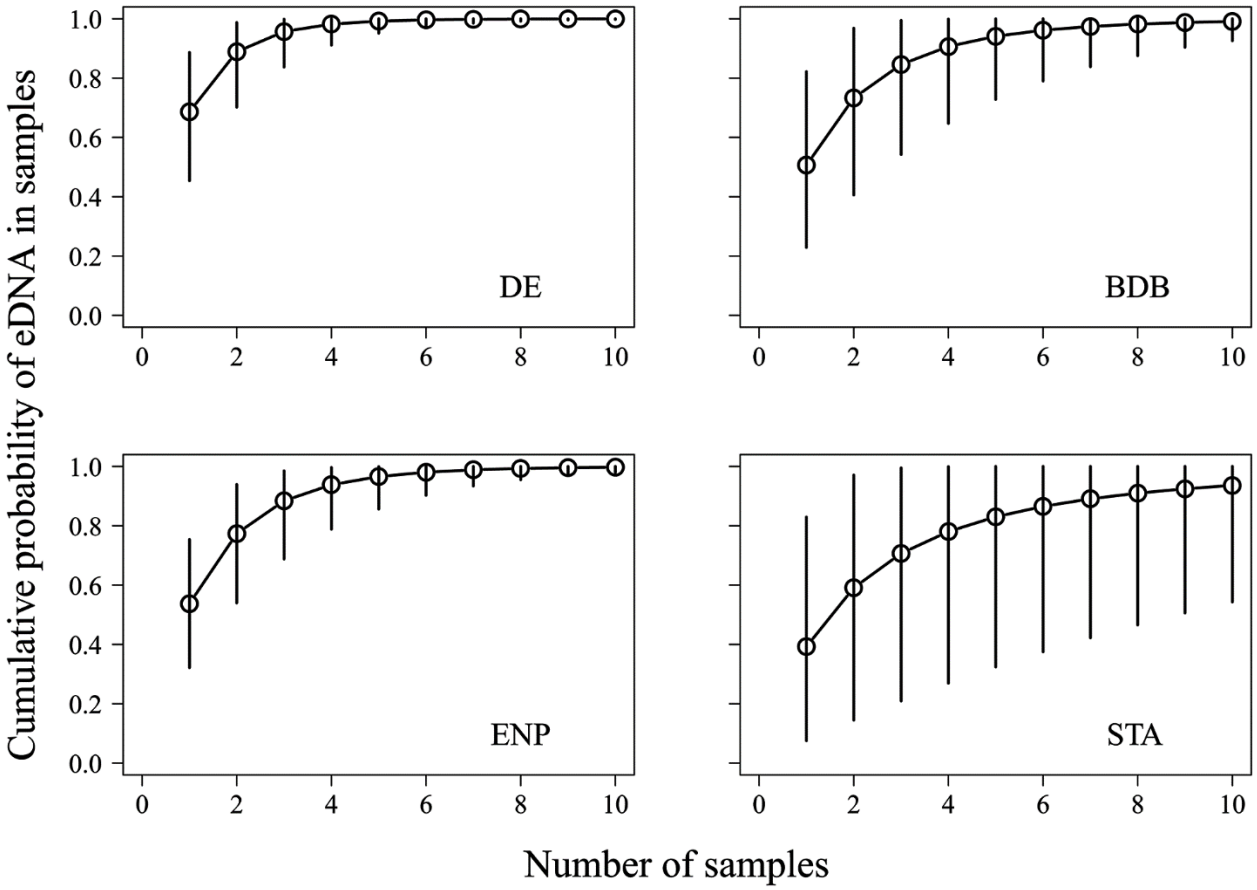
FROM: Veldhoen, N., Hobbs, J., Ikonomou, G., Hii, M., Lesperance, M., & Helbing, C. C. (2016). Implementation of Novel Design Features for qPCR-Based eDNA Assessment. PLOS ONE, 11(11), e0164907. doi: 10.1371/journal.pone.0164907

TARGETED DETECTION: OCCUPANCY AND DETECTION PROBABILITIES



Occupancy ($\psi_{d,t}$) and detection probabilities ($p_{d,t}$) through time for both species. The dots and lines are the medians for the posterior distribution. The lightest shaded regions are the 95% credible intervals and the darker shaded regions are the 80% credible intervals from the posterior distribution.

FROM: Erickson, R. A., Merkes, C. M., Jackson, C. A., Goforth, R. R., & Amberg, J. J. (2017). Seasonal trends in eDNA detection and occupancy of bigheaded carps. Journal of Great Lakes Research, 43(4), 762–770. doi: 10.1016/j.jglr.2017.06.003



The effect of sample size on estimates of the cumulative probability of occurrence of Burmese python environmental DNA (θ*) associated with samples taken from the four field locations analyzed by the three-level occupancy model. BDB, Bird Drive Basin; DE, Deering Estates; ENP, Everglades National Park; STA, Stormwater Treatment Area 5.

FROM: Hunter, M. E., Oyler-McCance, S. J., Dorazio, R. M., Fike, J. A., Smith, B. J., Hunter, C. T., … Hart, K. M. (2015). Environmental DNA (eDNA) sampling improves occurrence and detection estimates of invasive Burmese pythons. PLoS ONE, 10(4), e0121655. doi: 10.1371/journal.pone.0121655

TARGETED DETECTION: DETECTION SUMMARY TABLE

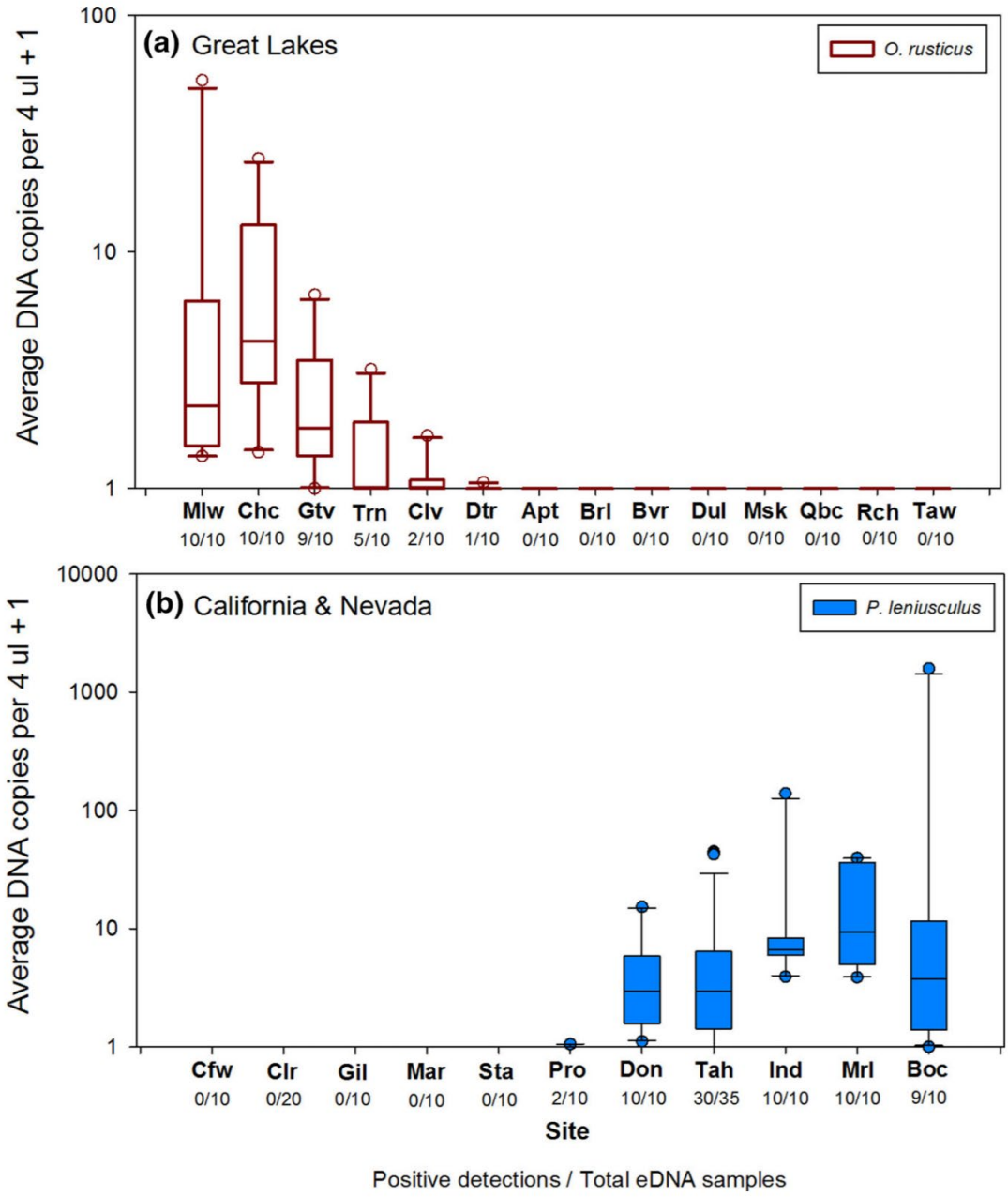
Location	No. of samples	Sampling date(s)	No. of samples with positive detections
Lake Erie basin			
Sandusky Bay	48	29 August 2011	4
	120	30 August 2011	0
	157	31 August 2011	0
North Maumee Bay	48	18 August 2011	2
Lorain	44	1 September 2011	0
Eagle Marsh connection (Ind.)	133	18 October 2010	0
	95	28 June 2011	0
Maumee River (Ind.)	17	18 October 2010	0
	14	28 June 2011	0
Maumee River (Ohio)	50	23 November 2010	0
	25	18 November 2010	0
Total	751		6
Lake St. Clair basin			
Algonac Marsh	20	11 August 2011	0
Belvidere Bay	20	11 August 2011	0
Black Creek	5	11 August 2011	0
Campau Bay	15	11 August 2011	0
Fairhaven Bay	20	11 August 2011	0
Swan Creek	15	11 August 2011	0
Thames River	120	22 September 2011	0
Total	215		0
Lake Michigan basin			
Burns Harbor and Ditch	44	11 August 2010	0
Chicago Area Water Way System (above electric barrier only)	811*	2009–2010	58
Galien River	24	12 July 2011	0
	15	5 October 2010	0
	59	29 September 2010	0
Gary Boat Slip	11	18 August 2010	0
Galiena River (Ind.)	8	5 October 2010	0
Grand River	38	21 June 2011	0
Indiana Harbor	54	18 August 2010	0
Kalamazoo River	120	13 June 2011	0
Kenosha Harbor	24	2 August 2011	0
Kinnickinnic River	6	17 November 2010	0
	16	26 July 2011	0
Menomenee River	12	26 July 2011	0
	10	17 November 2010	0
Milwaukee Harbor	72	26 July 2011	0
	24	17 November 2010	0
Milwaukee River	44	26 July 2011	0
	19	17 November 2010	0
	9	16 November 2010	0
Muskegon River	38	21 June 2011	0
Racine Harbor and Root River	48	2 August 2011	0
Sheboygan River	48	19 July 2011	0
St. Joseph River	120	21 July 2011	0
	60	3 June 2011	0
	65	16 September 2010	0
	57	15 September 2010	0
Total	1856		58
Grand Total	2822		64

*Indicates samples collected and reported fully in Jerde et al. 2011.

Location, sampling effort, date of collections, and the number of samples testing positive for DNA from either bighead or silver carp.

FROM: Jerde, C. L., Chadderton, W. L., Mahon, A. R., Renshaw, M. A., Corush, J., Budny, M. L., … Lodge, D. M. (2013). Detection of Asian carp DNA as part of a Great Lakes basin-wide surveillance program. Canadian Journal of Fisheries and Aquatic Sciences, 70(4), 522–526. Doi: 10.1139/cjfas-2012-0478

TARGETED DETECTION: DNA QUANTIFICATION BOXPLOTS



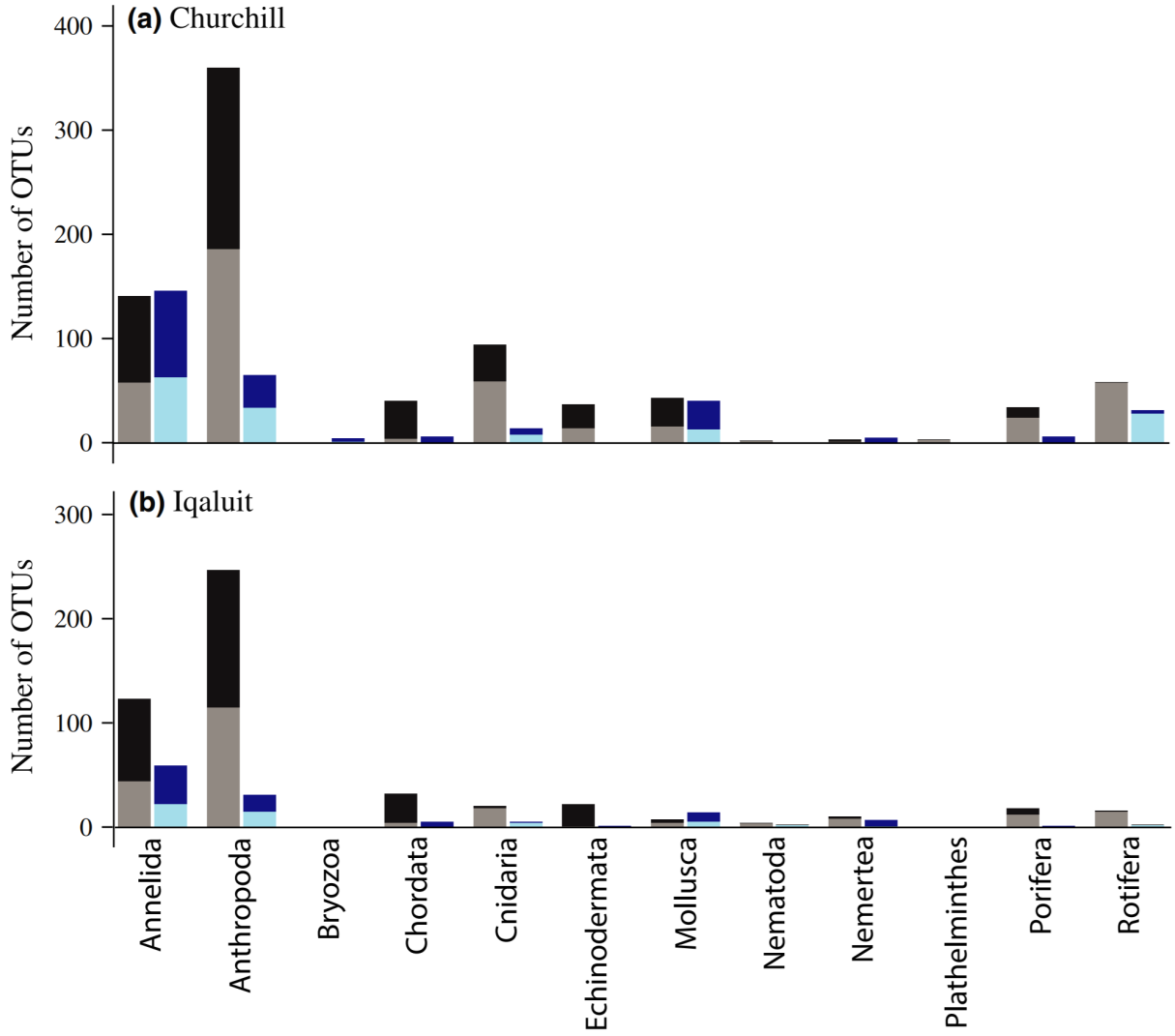
Average eDNA copies per 4 ul + 1 on a log10 axis for both the Great Lakes (a) and California and Nevada (b) regions for the rusty crayfish *Orconectes rusticus* and signal crayfish *Pacifastacus leniusculus*. No *O. rusticus* eDNA was detected from California and Nevada lakes, and reciprocally no *P. leniusculus* eDNA was detected from the Great Lakes region. Sample site abbreviations are given in Electronic Supplementary Table S3 (see also Fig. 1). The number of positive detections per total number of eDNA water samples at each site is given below the x-axis of each plot.

FROM: Larson, E. R., Renshaw, M. A., Gantz, C. A., Umek, J., Chandra, S., Lodge, D. M., & Egan, S. P. (2017). Environmental DNA (eDNA) detects the invasive crayfishes Orconectes rusticus and Pacifastacus leniusculus in large lakes of North America. Hydrobiologia, 800(1), 173–185. doi: 10.1007/s10750-017-3210-7

Kingdom	Phylum	Class	Order	Family	Genus	Species
Animalia	Annelida	2	7	17	16	
	Arthropoda	3	12	35	38	8
	Brachiopoda	1				
	Bryozoa	1	1	3		
	Chaetognatha		1	1	1	
	Chordata	5	17	45	73	36
	Cnidaria	2	5	9	7	
	Ctenophore	1	1	1	1	
	Echinodermata	3	2	3	4	2
	Gastrotricha		1	1		
	Mollusca	4	15	22	19	5
	Nematoda	1	1	1		
	Nemertea	2	1	2	2	
	Platyhelminthes	4	3	2		
	Porifera	3	6	6	6	
	Xenacoelomorpha	1		1		
Chromista	Bigyra	1	2	2	3	
	Cercozoa	4	4	3	4	
	Ciliophora	7	15	24		
	Cryptophyta	3	3	6	6	
	Haptophyta	4	5	5	4	1
	Heliozoa	1	1	1	1	
	Myzozoa	3	10	25	36	2
	Ochrophyta	9	31	27		
	Oomycota	1	3	2		
	Radiozoa	2	2	1	1	
Fungi	Ascomycota	3	4	3		
	Basidiomycota	2	2	1	1	
	Chytridiomycota	1	1	1	1	
	Entomophthoromycota	1	1	1	1	
Plantae	Chlorophyta	7	10	11	17	
	Glaucophyta	1		1	1	
	Rhodophyta	1	5	7	3	
	Streptophyta	1	11	11	2	
Protozoa	Amoebozoa		1	1	1	
	Apusozoa			2	3	
	Choanozoa	2	1	2	3	
	Picozoa	1	1	1		
Total	**38**	**88**	**186**	**287**	**255**	**54**

The number of eukaryotic taxa identified within each taxonomic rank at Coral Bay in west Australia via ToL-metabarcoding.

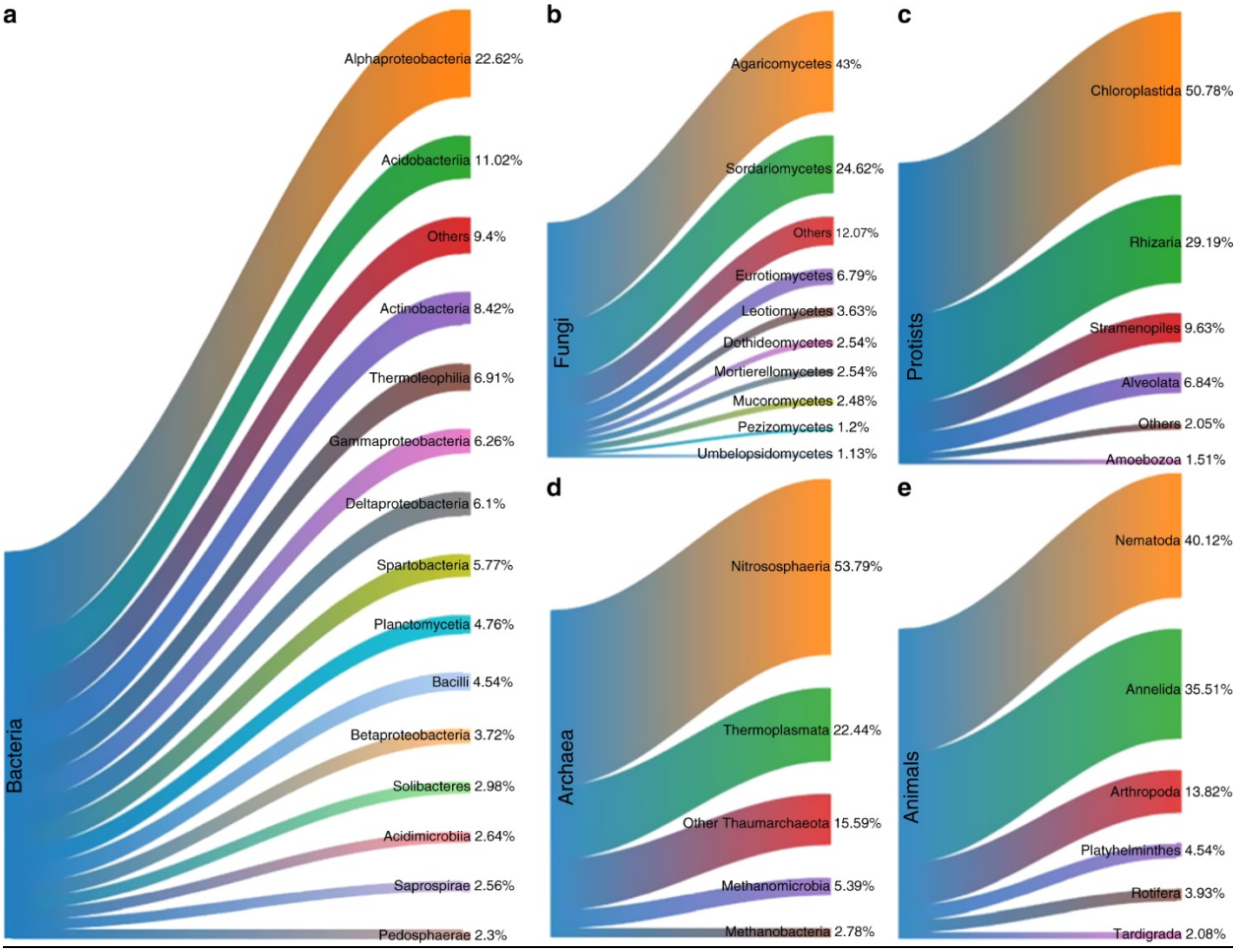
FROM: Stat, M., Huggett, M. J., Bernasconi, R., DiBattista, J. D., Berry, T. E., Newman, S. J., … Bunce, M. (2017). Ecosystem biomonitoring with eDNA: metabarcoding across the tree of life in a tropical marine environment. Scientific Reports, 7(1). doi: 10.1038/s41598-017-12501-5



The number of Operational taxonomic units (OTU) identified at the species level (dark: ≥97% identity) relative to those identified below the species level (lighten: ≥85% and <97% identity) for each phylum and from the COI1 (mlCOIintFjgHCO2198: black and gray) and COI2 (LCO1490-ill_C_R: blue) primer sets separately for both Arctic sampling ports (Churchill and Iqaluit)

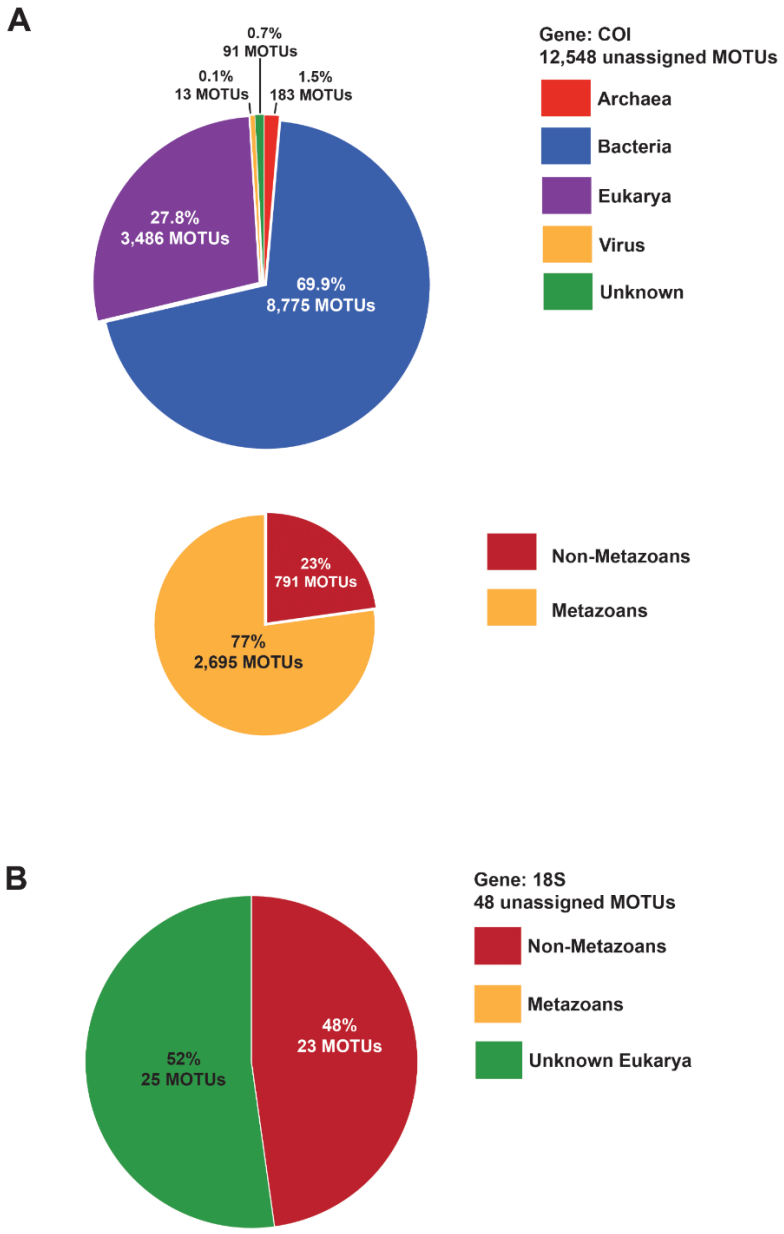
FROM: Lacoursière-Roussel, A., Howland, K., Normandeau, E., Grey, E. K., Archambault, P., Deiner, K., … Bernatchez, L. (2018). EDNA metabarcoding as a new surveillance approach for coastal Arctic biodiversity. Ecology and Evolution, 8(16), 7763–7777. doi: 10.1002/ece3.4213

BIODIVERSITY INVENTORY: SANKEY DIAGRAMS



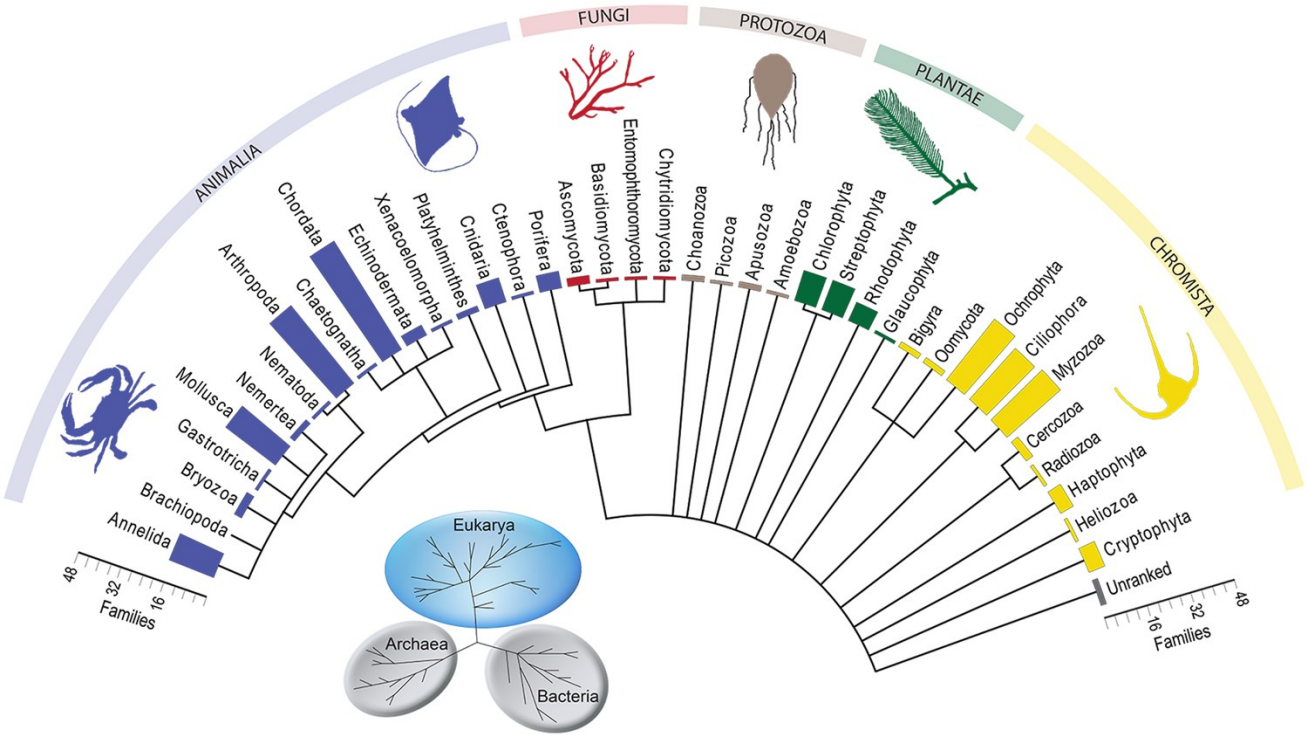
Sankey diagrams of proportional abundances of OTUs from all samples for major soil biota groups. Arms denote proportions of OTUs at the class-level for **a** bacteria; **b** fungi; of major lineages of **c** protists; class-level for **d** archaea; and at the phylum-level for **e** animals.

FROM: George, P. B. L., Lallias, D., Creer, S., Seaton, F. M., Kenny, J. G., Eccles, R. M., … Jones, D. L. (2019). Divergent national-scale trends of microbial and animal biodiversity revealed across diverse temperate soil ecosystems. Nature Communications, 10(1). doi: 10.1038/s41467-019-09031-1



Taxonomic percentages and counts of Molecular Operational Taxonomic units (MOTUs) initially unassigned for COI and 18S, inferred using NCBI BLAST public nucleotide database. (A) Taxonomic percentages for the 12,548 unassigned MOTUs for the COI gene. The top pie chart illustrates the proportions of each domain, while the bottom pie chart estimates the proportions of metazoan versus non-metazoans within the eukaryotes. (B) Taxonomic percentages for the 48 unassigned MOTUs for the 18S gene. The pie chart is divided into non-metazoans versus unknown Eukaryotes. The initially unassigned MOTUs for both genes were identified using the blast-stand alone program with the “blastn” algorithm under default parameters.

FROM: Cowart, D. A., Pinheiro, M., Mouchel, O., Maguer, M., Grall, J., Miné, J., & Arnaud-Haond, S. (2015). Metabarcoding Is Powerful yet Still Blind: A Comparative Analysis of Morphological and Molecular Surveys of Seagrass Communities. PLOS ONE, 10(2), e0117562. doi: 10.1371/journal.pone.0117562



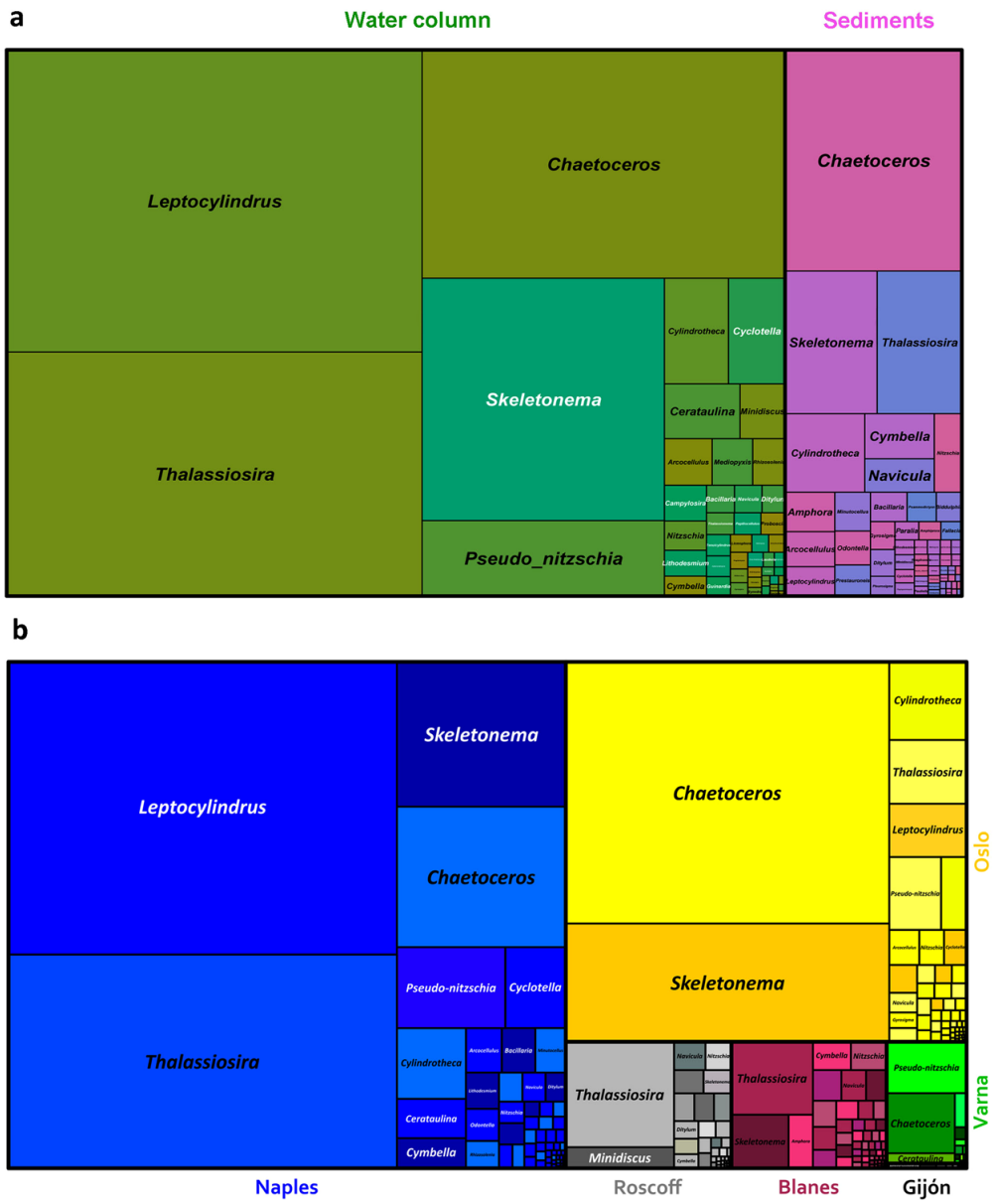
Taxonomic phylogram of eukaryotic diversity at Coral Bay in west Australia derived from ToLmetabarcoding. Bar graphs indicate the number of families in each phyla characterised at Coral Bay, and are coloured according to kingdom.

FROM: Stat, M., Huggett, M. J., Bernasconi, R., DiBattista, J. D., Berry, T. E., Newman, S. J., … Bunce, M. (2017). Ecosystem biomonitoring with eDNA: metabarcoding across the tree of life in a tropical marine environment. Scientific Reports, 7(1). doi: 10.1038/s41598-017-12501-5

	Sampling locations along Eagle Creek (1-6) and Kalamazoo River (7-8)							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Ambloplites rupestris				x	x		x	x
Ameiurus natalis		x	x	x	x	x	x	
Amia calva							x	x
Cyprinella spiloptera							x	
Erimyzon sucetta					x			
Esox americanus	x							
Etheostoma caeruleum					x	x		x
Etheostoma exile		x		x				
Etheostoma nigrum					x	x	x	x
Hypentelium nigricans							x	x
Lepomis cyanellus						x		
Lepomis gibbosus		x	x	x	x	x		
Lepomis gulosus						x		
Lepomis macrochirus		x	x	x	x	x	x	x
Micropterus dolomieu								x
Micropterus salmoides		x		x		x		x
Moxostoma erythrurum						x	x	x
Notemigonus crysoleucas		x	x					
Noturus flavus								x
Perca flavescens		x	x					
Percina maculata								x
Semotilus atromaculatus				x	x	x	x	
Umbra limi	x			x				
Total number of species	2	7	5	8	8	10	9	11

Fish species presence along Eagle Creek and in the Kalamazoo River detected using three genetic markers (Ac16S, Am12S and Cytb). Each column is a sampling location. Each row is the presence (x) or absence of each species.

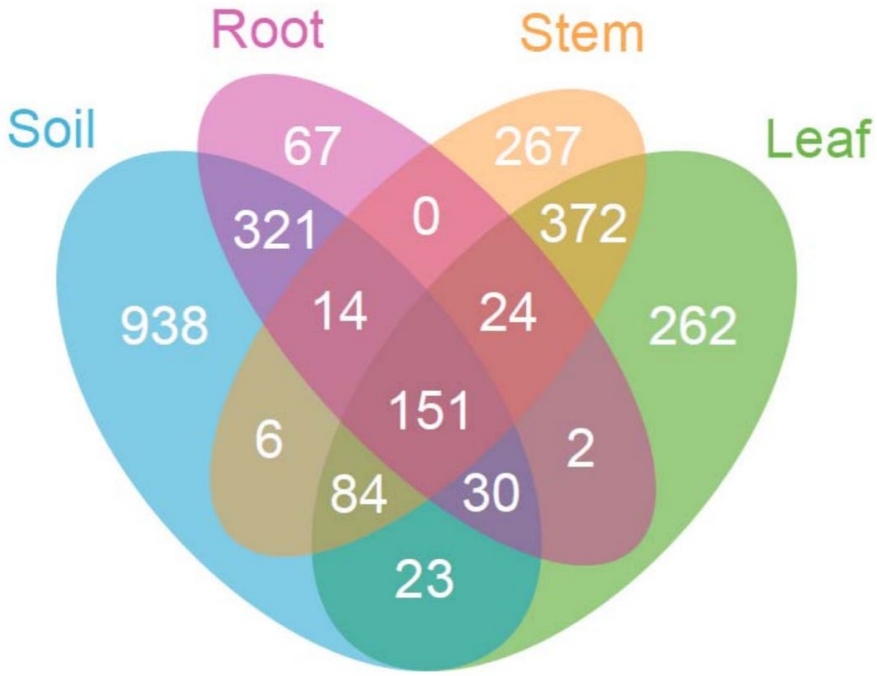
FROM: Li, Y., Evans, N. T., Renshaw, M. A., Jerde, C. L., Olds, B. P., Shogren, A. J., … Pfrender, M. E. (2018). Estimating fish alpha- and beta-diversity along a small stream with environmental DNA metabarcoding. Metabarcoding and Metagenomics, 2, e24262. doi: 10.3897/mbmg.2.24262



Distribution of diatom genera in coastal European seas. (**a**) Distribution of genera in the whole water column (green nuances) and sediment (pink nuances) datasets, both dominated by planktonic genera (*Leptocylindrus*, *Chaetoceros*, *Talassiosira* and *Skeletonema*). (**b**) Different contribution of diatoms and relative abundance of main genera (including water column and sediments) at European sampling sites (Naples=blue, Oslo=yellow, Roscoff=grey, Blanes=purple, Varna=green, Gijon=black). Naples and Oslo samples collected in two different seasons were merged.

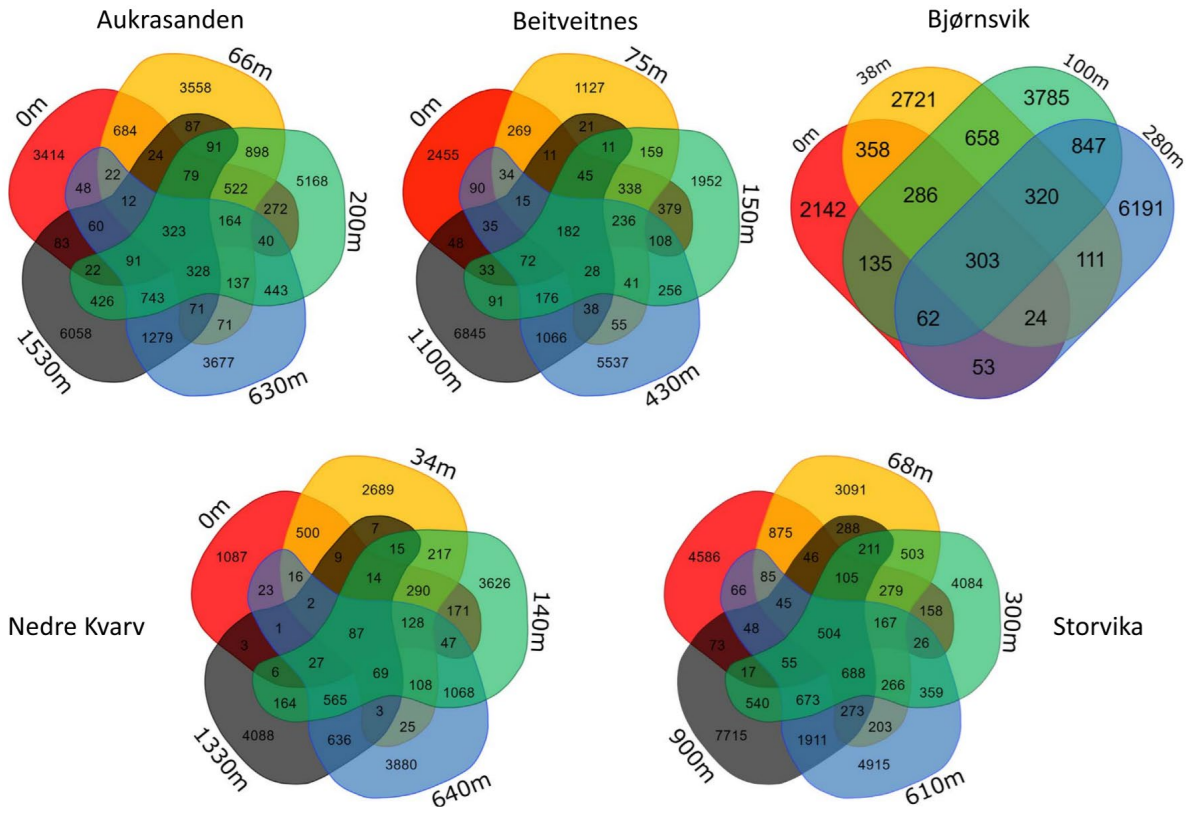
FROM: Piredda, R., Claverie, J.-M., Decelle, J., de Vargas, C., Dunthorn, M., Edvardsen, B., … Zingone, A. (2018). Diatom diversity through HTS-metabarcoding in coastal European seas. Scientific Reports, 8(1). doi: 10.1038/s41598-018-36345-9

COMMUNITY COMPOSITION: VENN DIAGRAMS



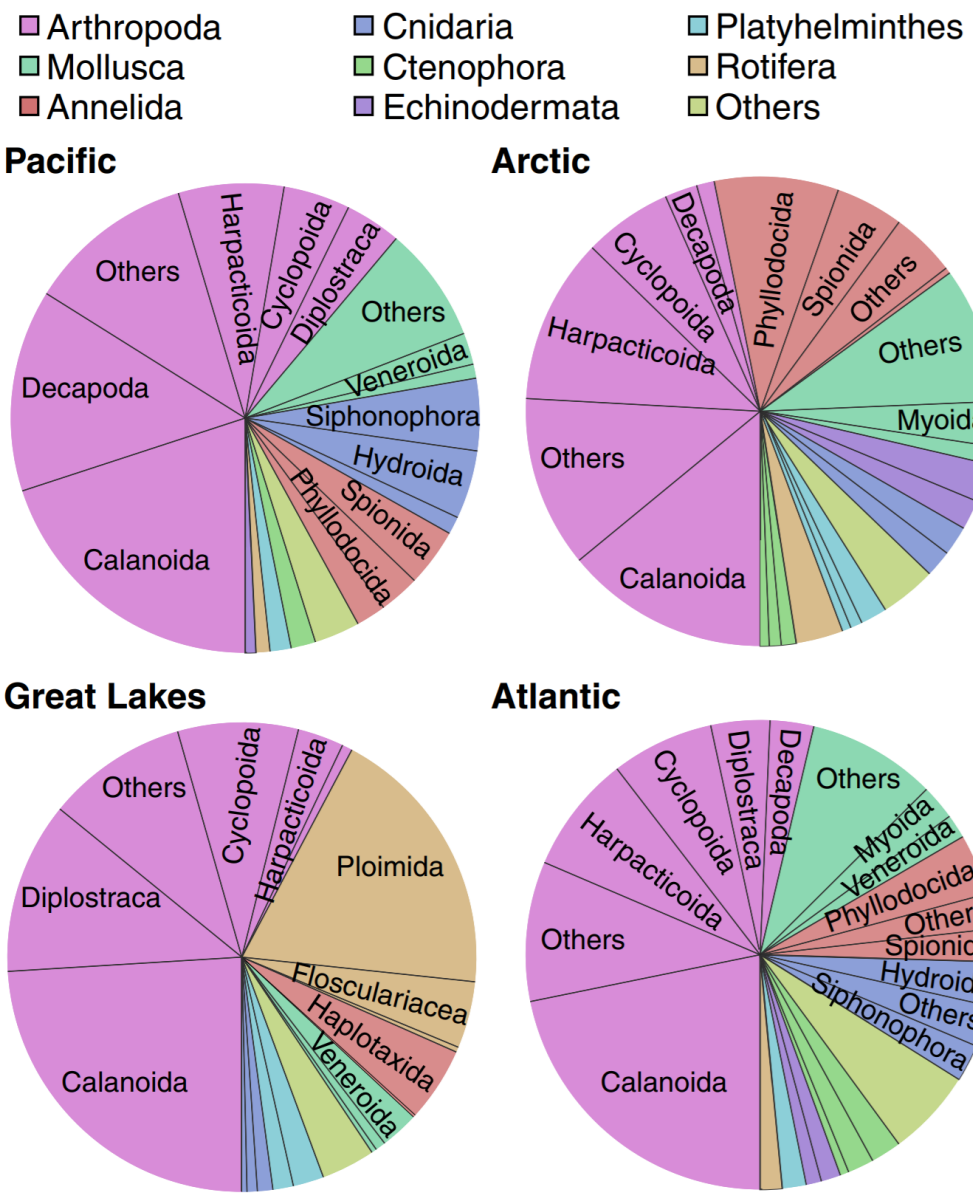
Venn diagram showing the overlap of the fungal communities from the four poplar habitats, based on OTUs. OTU delineation was based on a threshold of <97% sequence similarity.

FROM: Durand, A., Maillard, F., Foulon, J., Gweon, H. S., Valot, B., & Chalot, M. (2017). Environmental Metabarcoding Reveals Contrasting Belowground and Aboveground Fungal Communities from Poplar at a Hg Phytomanagement Site. Microbial Ecology, 74(4), 795–809. doi: 10.1007/s00248-017-0984-0



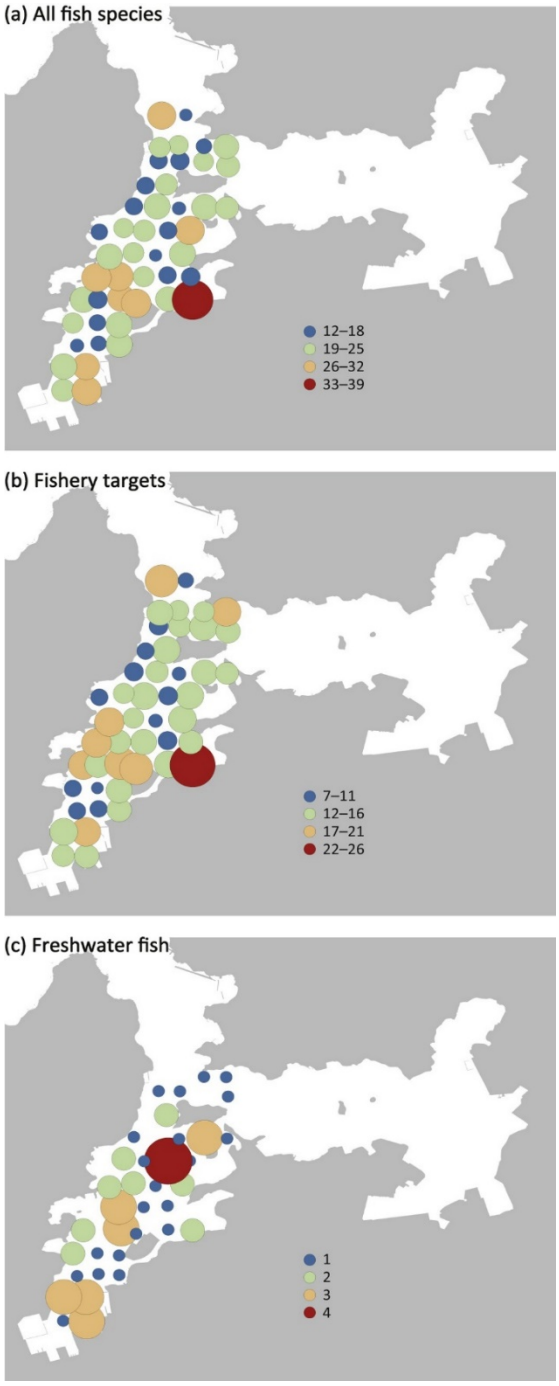
Numbers of bacterial DNA metabarcodes (OTUs – operational taxonomic units) at each farm under study, which were detected exclusively in one single sample and numbers of bacterial DNA metabarcodes which were shared between two or more samples.

FROM: Stoeck, T., Frühe, L., Forster, D., Cordier, T., Martins, C. I., & Pawlowski, J. (2018). Environmental DNA metabarcoding of benthic bacterial communities indicates the benthic footprint of salmon aquaculture. Marine Pollution Bulletin, 127, 139–149.



OTU richness among orders across sampled regions. The proportion of taxa detected from different orders and phyla are shown for each region.

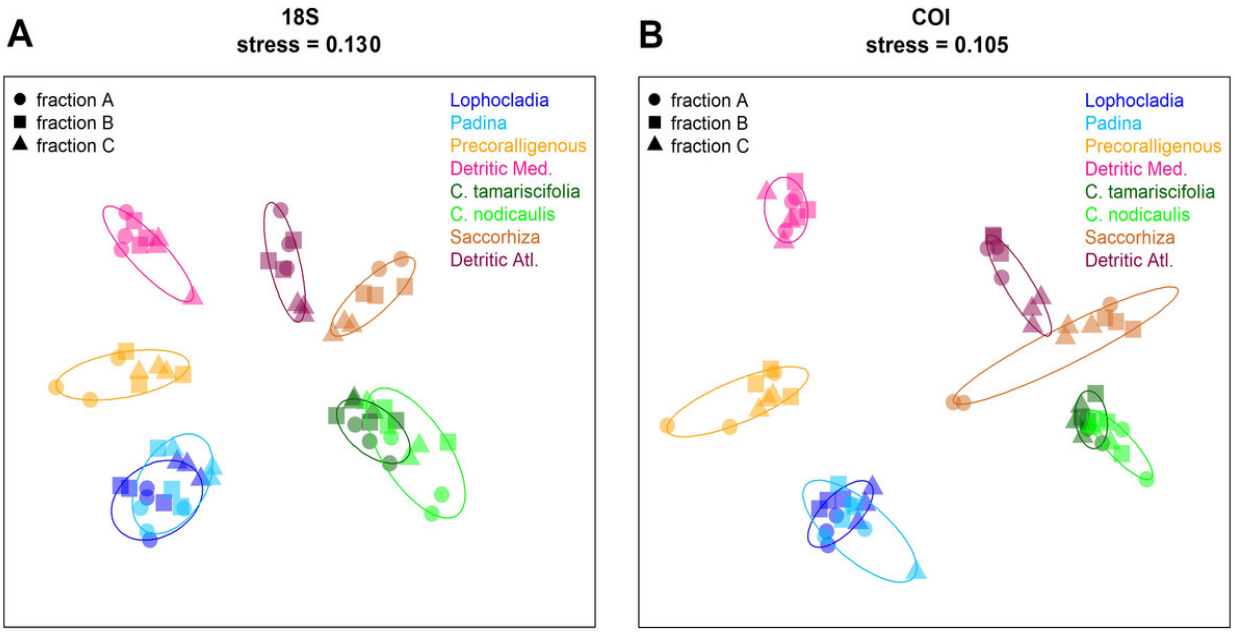
FROM: Chain, F. J. J., Brown, E. A., MacIsaac, H. J., & Cristescu, M. E. (2016). Metabarcoding reveals strong spatial structure and temporal turnover of zooplankton communities among marine and freshwater ports. Diversity and Distributions, 22(5), 493–504. doi: 10.1111/ddi.12427



Circles on each map indicate the number of detected species for (a) all fish species, (b) fishery targets and (c) freshwater fish. Both size and colour reflect the species number. This map was created using QGIS version 2.8 (http://www.qgis.org/en/site/) based on the Administrative Zones Data (2016) [(c) National Spatial Planning and Regional Policy Bureau, Ministry of Land, Infrastructure, Transportation and Tourism.

FROM: Yamamoto, S., Masuda, R., Sato, Y., Sado, T., Araki, H., Kondoh, M., … Miya, M. (2017). Environmental DNA metabarcoding reveals local fish communities in a species-rich coastal sea. Scientific Reports, 7(1). doi: 10.1038/srep40368

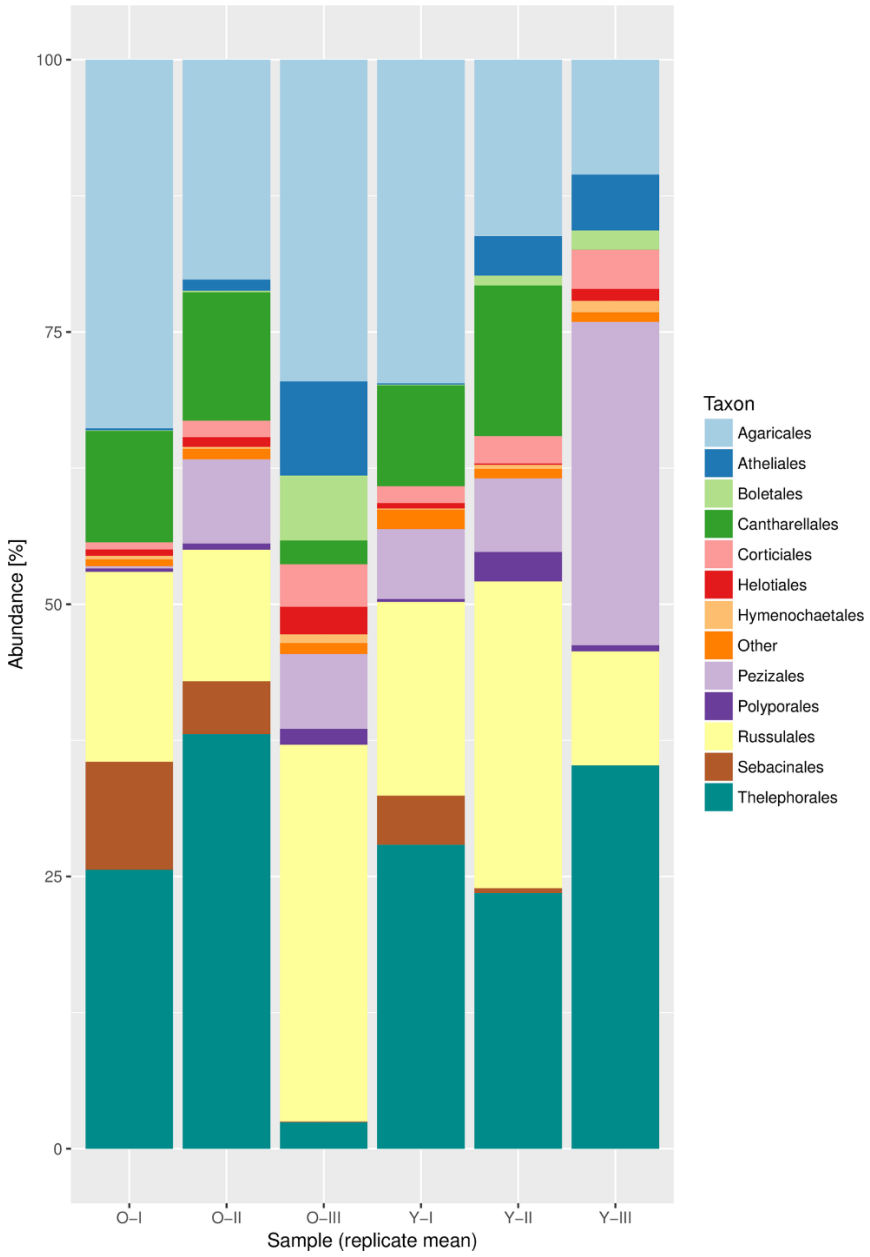
COMMUNITY COMPOSITION: ORDINATIONS



Non-metric multi-dimensional scaling (nmMDS) representation of the samples using the Bray–Curtis coefficient. (A) 18S, (B) COI. Samples are symbol-coded for fraction and colour-coded for community. Stress of the final configurations is indicated.

FROM: Wangensteen, O. S., Palacín, C., Guardiola, M., & Turon, X. (2018). DNA metabarcoding of littoral hard-bottom communities: high diversity and database gaps revealed by two molecular markers. PeerJ, 6, e4705. doi: 10.7717/peerj.4705

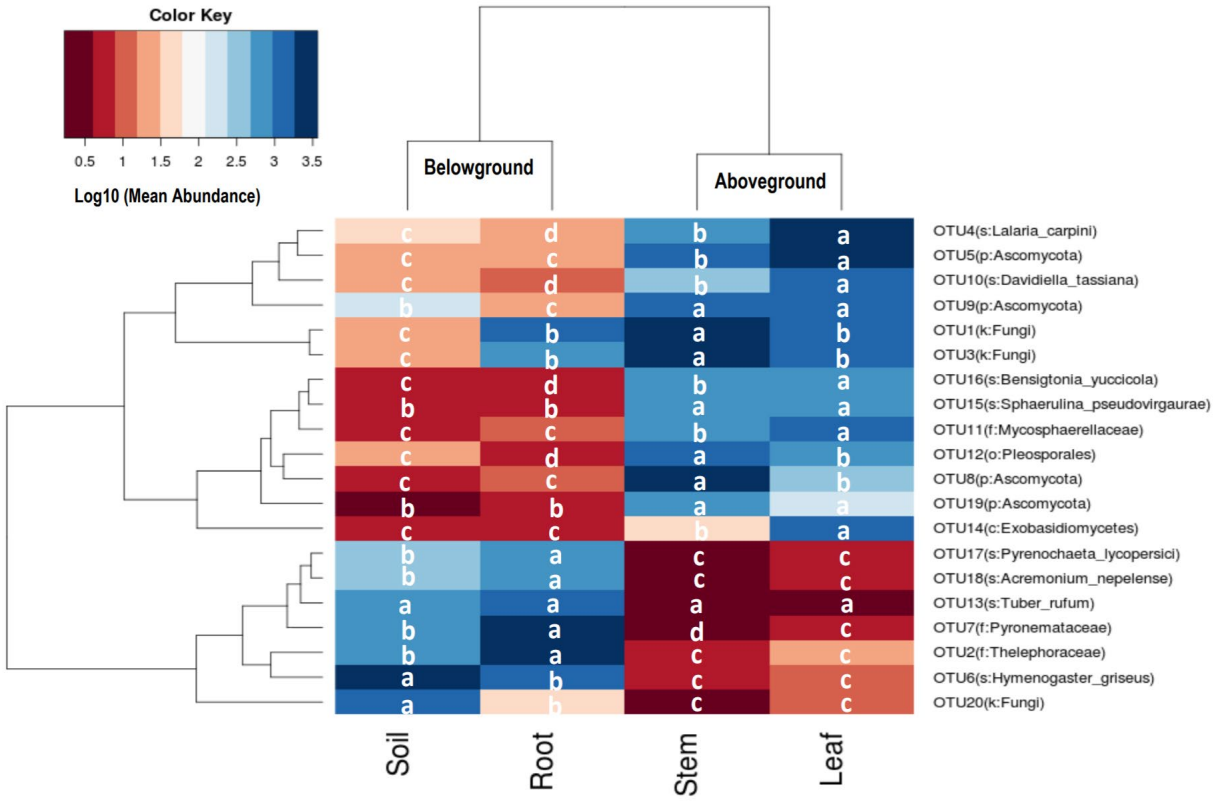
COMMUNITY COMPOSITION: STACKED BAR PLOT



Abundances from normalized read counts for fungal orders detected in roots and surrounding soil from Norway spruce in the Bavarian Alps. Only orders with abundance >1% in at least one sample are shown, other orders are summed under 'Other'. Sample replicate average is shown and labelled according to the age of the host tree ('O' for old, 'Y' for young individuals) and the altitude ('I','II','III', see Fig 1). See S2 Fig for all samples.

FROM: Schön, M. E., Nieselt, K., & Garnica, S. (2018). Belowground fungal community diversity and composition associated with Norway spruce along an altitudinal gradient. PLOS ONE, 13(12), e0208493. doi: 10.1371/journal.pone.0208493

COMMUNITY COMPOSITION: HEAT MAPS



Heat map and hierarchical cluster analysis of the relative abundance of fungal OTUs from the various poplar habitats. Letters indicate significantly different abundances at $p < 0.05$ (Kruskal-Wallis comparison test), n=23 (root and leaf) or n=24 (soil and stem). The dendrogram represents linkage clustering using Euclidean distance measures. OTU delineation was based on a threshold of < 97% sequence similarity. The number associated with the OTU corresponds to the relative abundance rank of that OTU in the total dataset. Assignments between brackets show the lowest taxonomic level associated with the OTU using the UNITE database, k: kingdom, p: phylum, o: order, c: class, f: family, s: genus_species.

FROM: Durand, A., Maillard, F., Foulon, J., Gweon, H. S., Valot, B., & Chalot, M. (2017). Environmental Metabarcoding Reveals Contrasting Belowground and Aboveground Fungal Communities from Poplar at a Hg Phytomanagement Site. Microbial Ecology, 74(4), 795–809. doi: 10.1007/s00248-017-0984-0

APPENDIX C.

METHODS USED TO SUMMARIZE THE DISCUSSION DURING THE INTERACTIVE SESSION, DETERMINING BEST PRACTICES FOR DATA VISUALIZATIONS, AT IWEG 2020 INTO CATEGORIES.

To begin the workshop, two questions were asked of participants: (1) What criteria related to content are most important when evaluating figures and tables presenting environmental genomics results? and (2) What criteria related to STYLE are most important when evaluating figures and tables presenting environmental genomics results? Participants answered via an online survey where comments and positive responses to comments (likes) to questions were recorded. Comments were reviewed and grouped into categories based on the data visualization practice they recommended and the number of responses (comments and likes) in each of the categories was calculated. In total, six categories were identified, and these categories are described in Table 1 below. These categories were used to classify recommendations suggested by participants during subsequent discussions.

TABLE 1 - LIST OF CATEGORIES USED TO CLASSIFY VISUALIZATIONS CRITERIA WITH DESCRIPTIONS AND EXAMPLES INCLUDED FOR EACH CATEGORY.

Category	Categorization Rule
Simplicity	*Description*: minimizing the amount of information in a plot; limiting information to that necessary for the conclusion *Includes*: concise, less is more, not too much information, overcomplicated, not overcrowded or busy, focus on one point at a time
Clarity	*Description*: text and formatting of visualizations are clear and easy to understand *Includes*: descriptive title, clear explanation in figure caption, precise labelling and legends, good use of colour, legible text
Ease of Interpretation	*Description*: how easily a viewer can draw conclusions from the visualization; how intuitive it is to read and interpret the visualization *Includes*: all graphics stand alone with no additional text needed, effective formatting highlights trends and important elements, humans are bad at estimating area (e.g. pie charts)
Context	*Description*: presenting information that is needed to understand what is shown in the visualization; presenting all data required for correct interpretations *Includes*: statistically sound, presenting all information needed for correct interpretation, be specific about what data are being presented
Audience	*Description*: considering the audience/viewers/users in deciding how to present data *Includes*: different or relevant audience, identify audience, use colour-blind friendly colours
Relevance	*Description*: link between data presented, the question being asked, and the conclusions drawn from the data *Includes*: justification of conclusions, express the message, directly address questions, results are framed based on questions