

# Learning to Count (Organismal Abundance and eDNA)

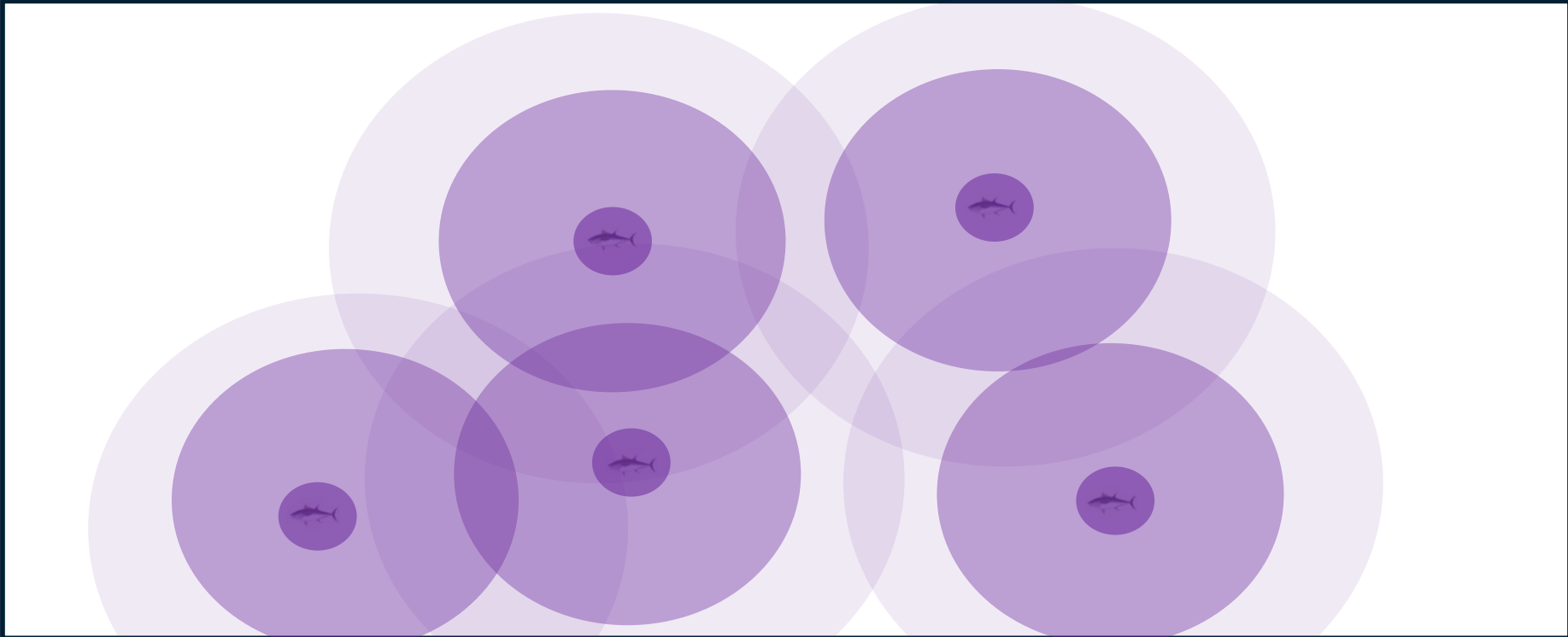
Ryan P. Kelly  
University of Washington  
Andrew (Ole) Shelton  
NOAA Fisheries

Wellington, NZ  
Feb 2025



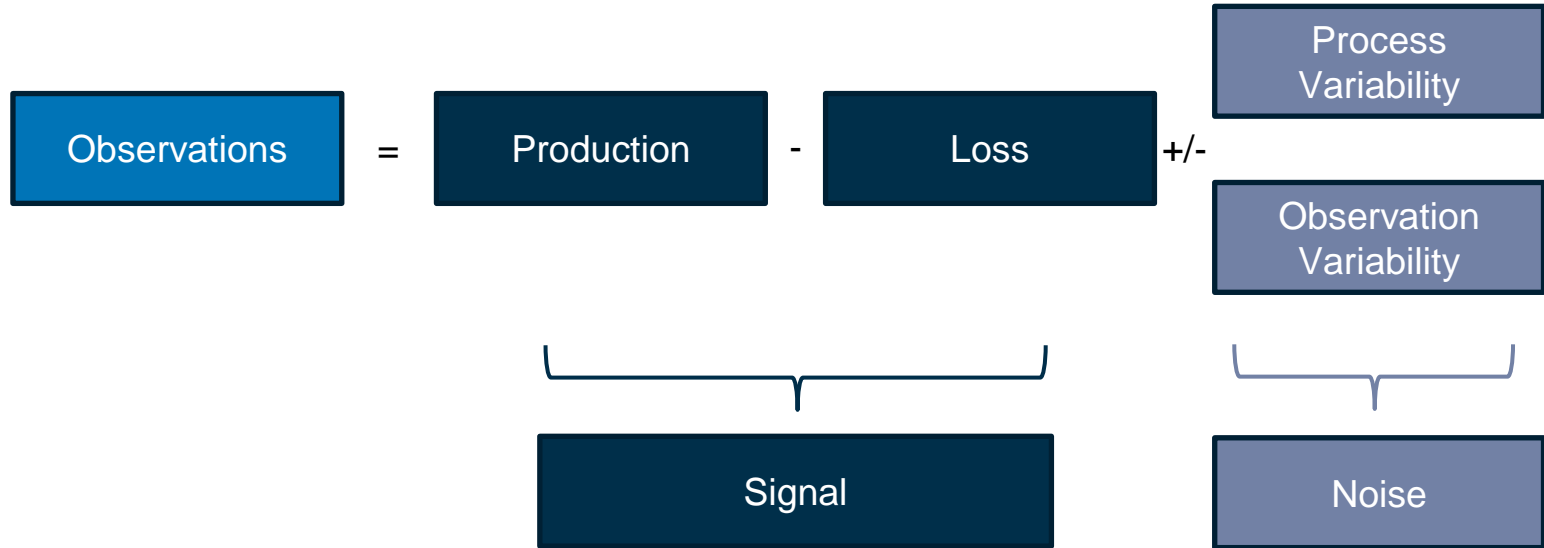
More of a species = more of its DNA (in the world)

---



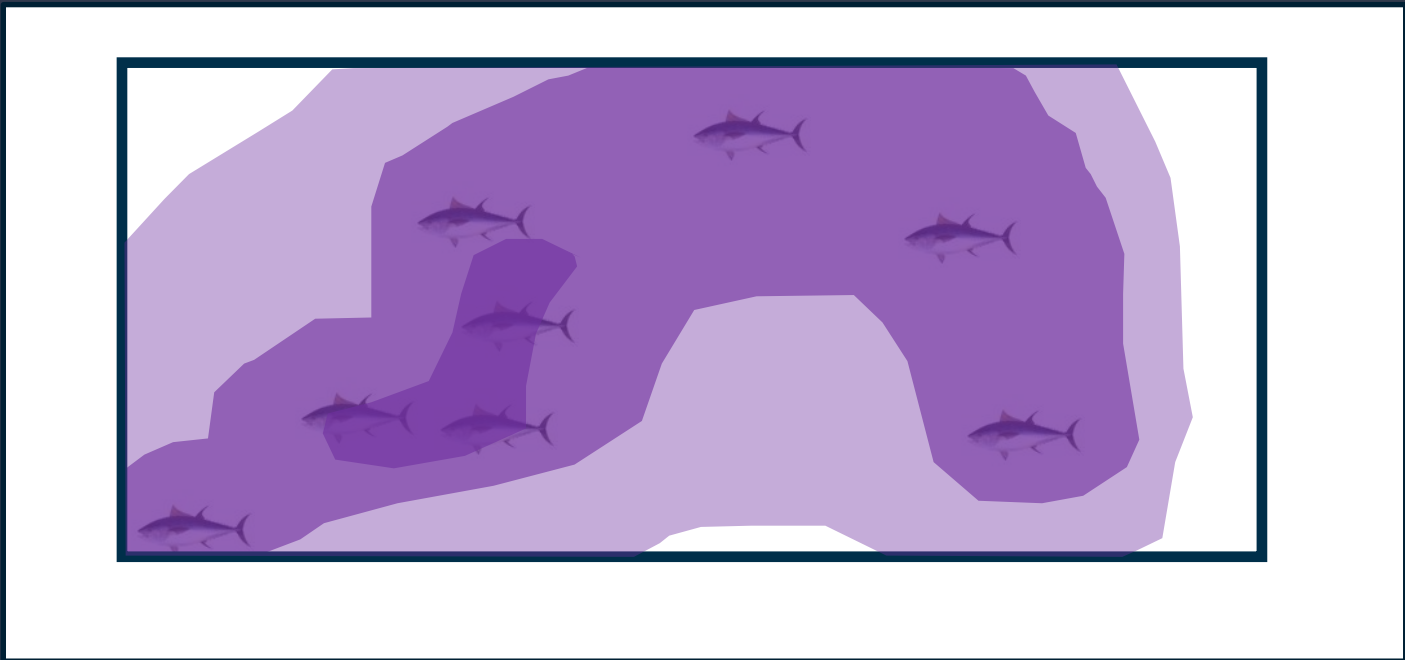
# What we observe $\neq$ what is there

---



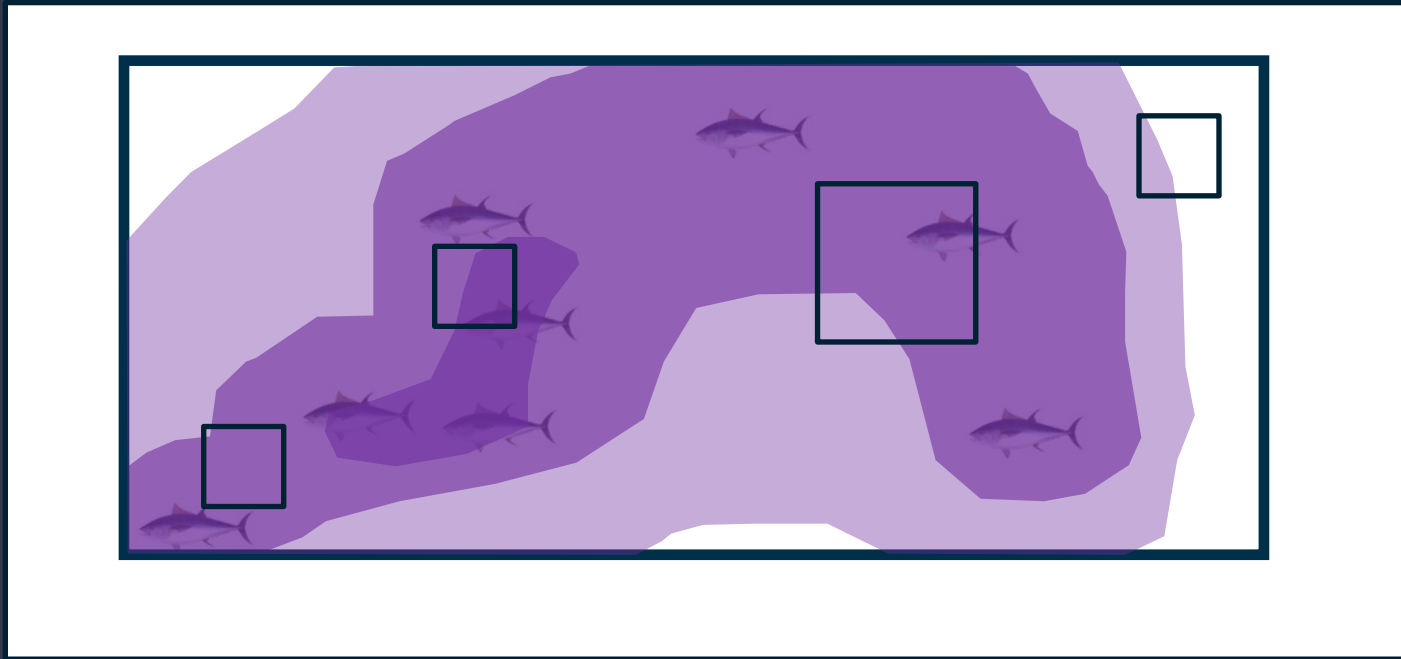
Signal: organisms are discrete; DNA is more continuous

---



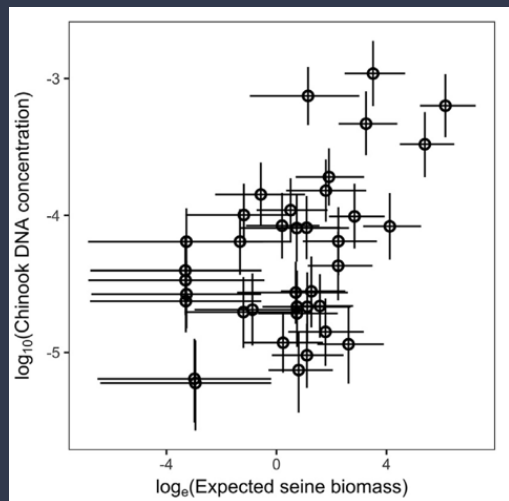
# Different sampling frames = different relationships

---



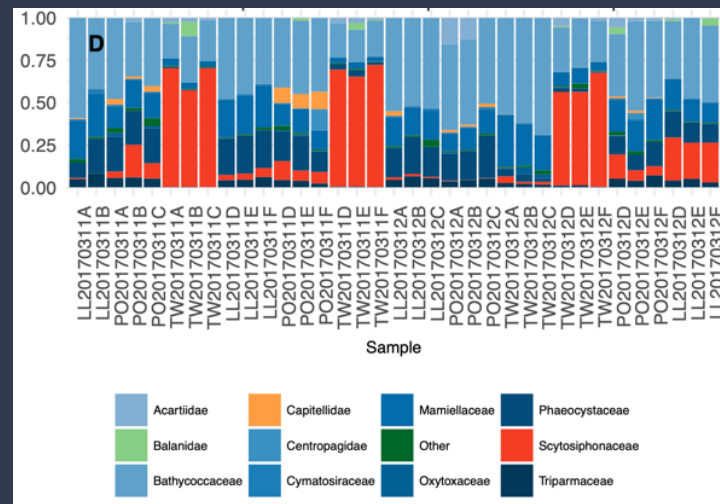
# Different Observation Models

qPCR / ddPCR vs. Metabarcoding



Shelton et al. 2019 Biol. Conserv.

Raw signal should reflect abundance +/- noise



Kelly et al. 2019 PeerJ

Raw signal should NOT reflect abundance +/- noise

- Amp bias
- Compositionality



- Calibrate (e.g., with mocks)
- Pull out of proportion space with additional data

# Real-World Applications

---

## Single-Species

- qPCR and biomass at different spatial scales
- qPCR + traps to estimate animal density

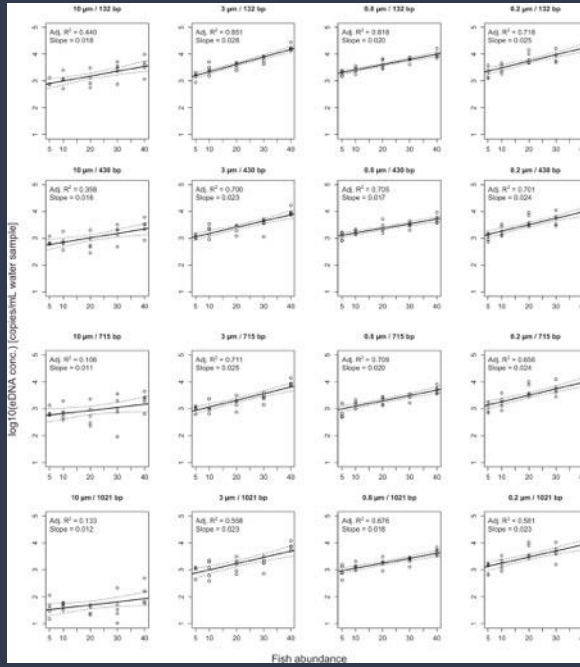
## Multi-Species

- 12S metabarcoding + trawl to estimate fish abundance
- 12S metabarcoding + qPCR to estimate eDNA concentrations for many species
- 12S metabarcoding + visual survey to estimate biomass of many species

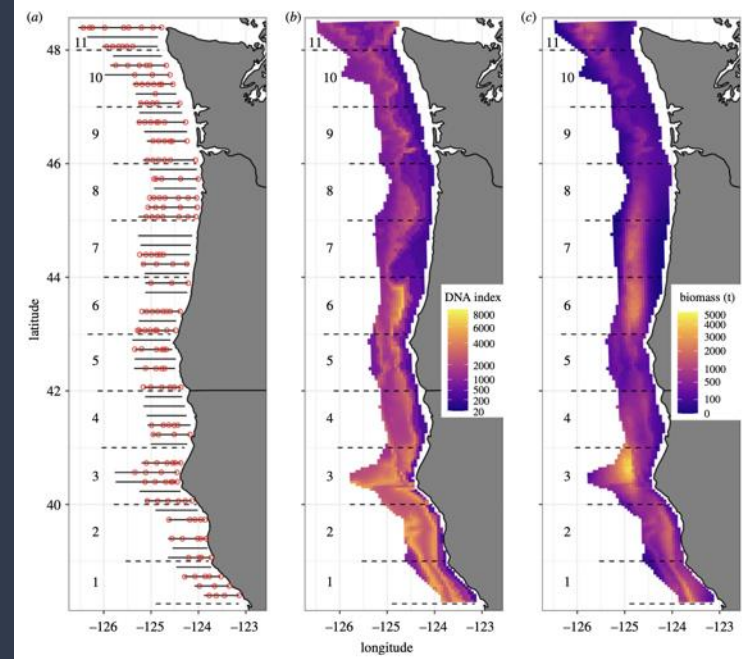
# Single Species – qPCR / ddPCR

Quantitative estimates of eDNA concentration

Tiny Scale



Large Scale



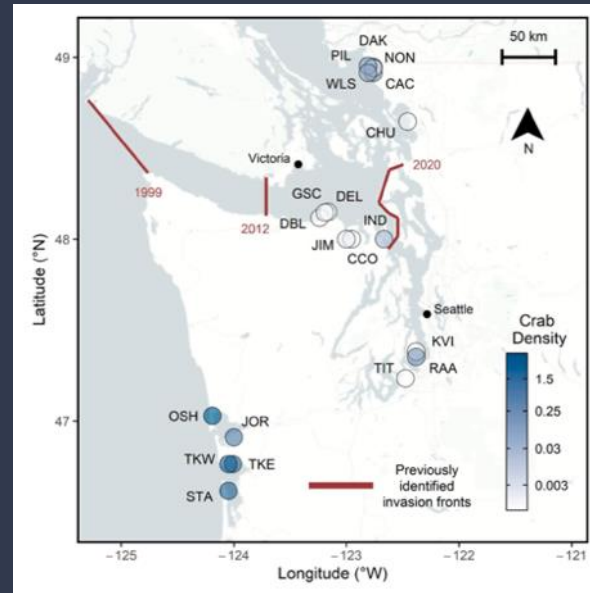
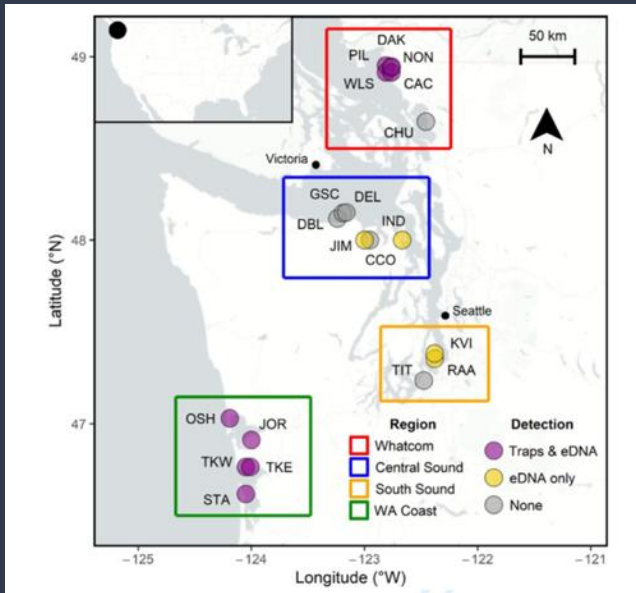
Jo and Yamanaka (2022) – Zebrafish aquarium study across pore sizes + fragment sizes.

Shelton et al. (2022): smoothed qPCR signal reflects Hake (*Merluccius productus*) eDNA distribution

# Single Species – multiple data sources

Quantitative estimates of species abundance

Green Crab (*Carcinus maenas*): qPCR + Traps



eDNA + one other data source:

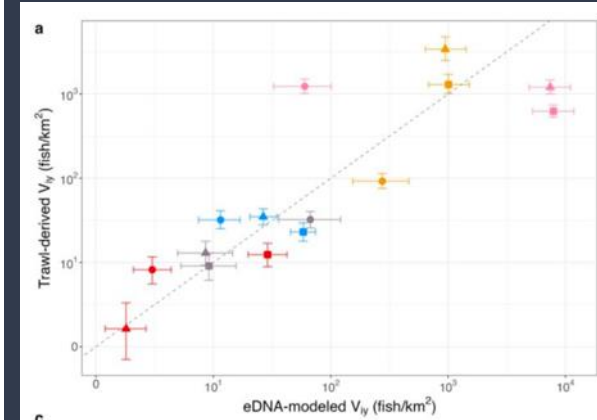
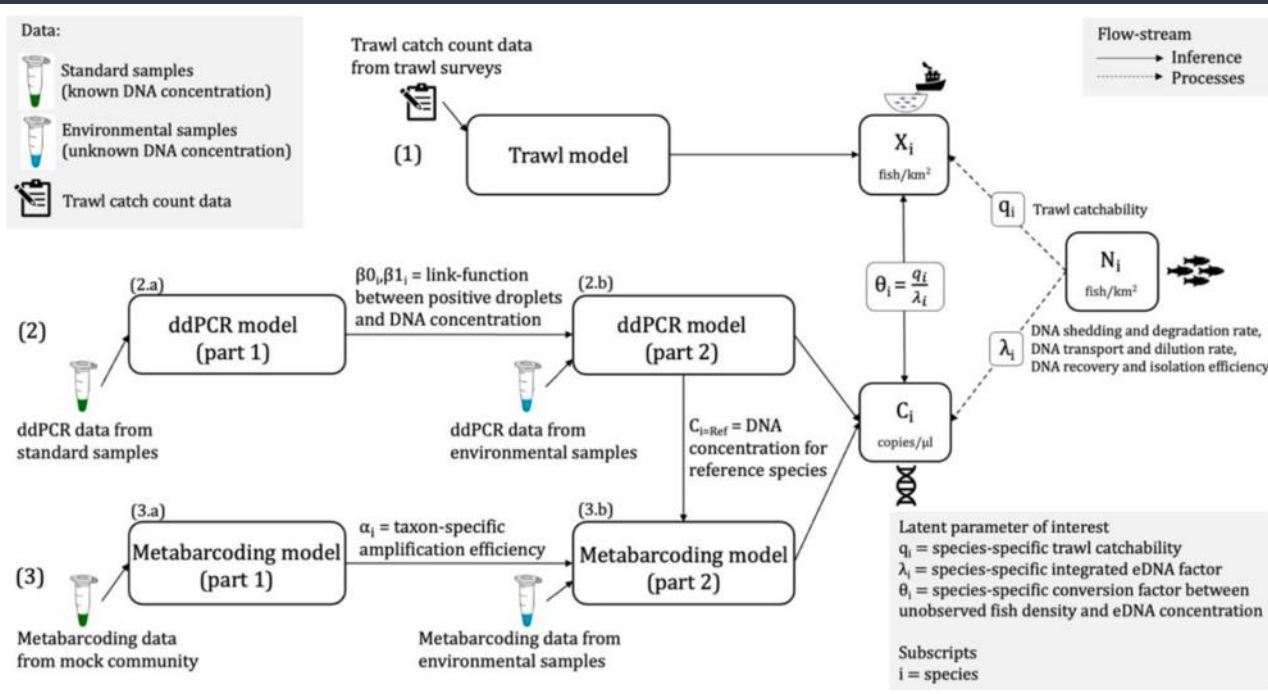


agkeller@berkeley.edu

# Multispecies (Metabarcoding)

Metabarcoding (proportions) + dPCR (ref species quantity) + trawls

Quantitative estimates of species abundance



Reciprocal out-of-sample cross-validation





# Thanks!

---

Eily Allan

Megan Shaffer

Gled Guri

Kate Bertko

Pedro Brendão

Shana Hirsch

Aden Ip

Owen Liu

Kim Parsons

Krista Nichols

Olivia Scott

Stephanie Matthews

Len Thomas

Jon Whitney

Jamie Gove

Jameal Samhour

The David and Lucile Packard Foundation

Oceankind

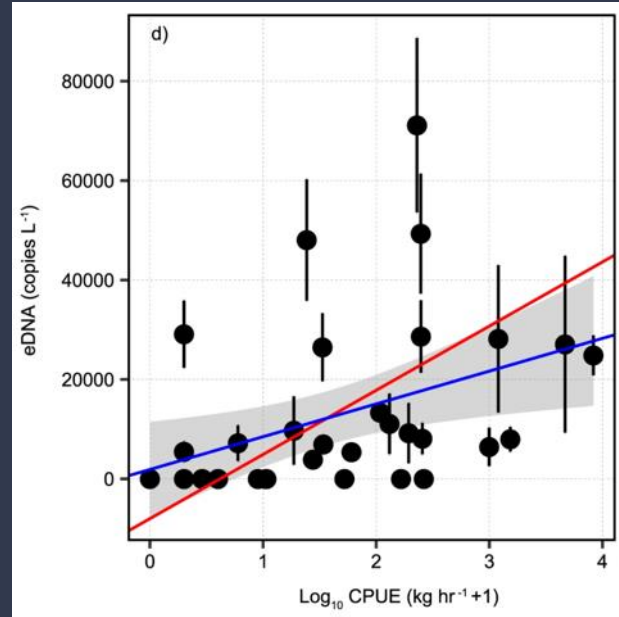
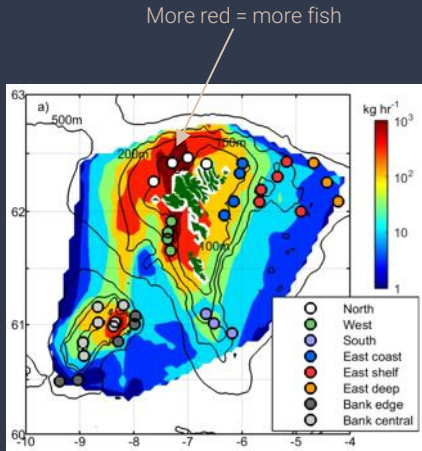
Office of Naval Research

NOAA Fisheries

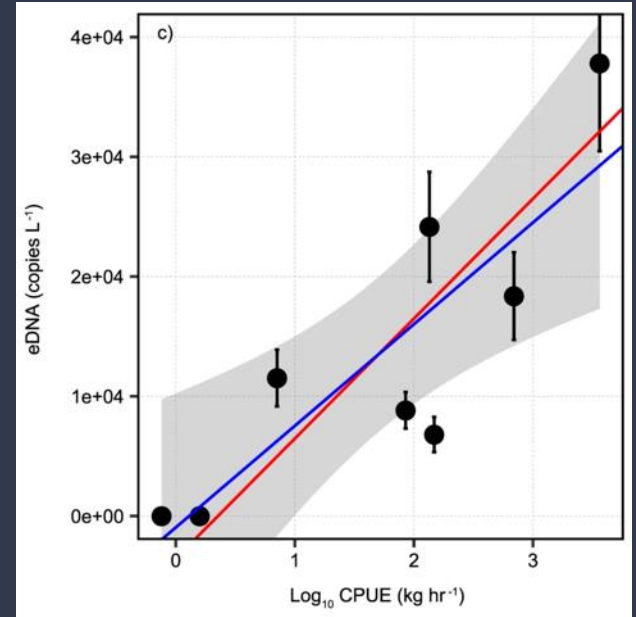
# Extra Slides

---

# Different sampling frames = different relationships



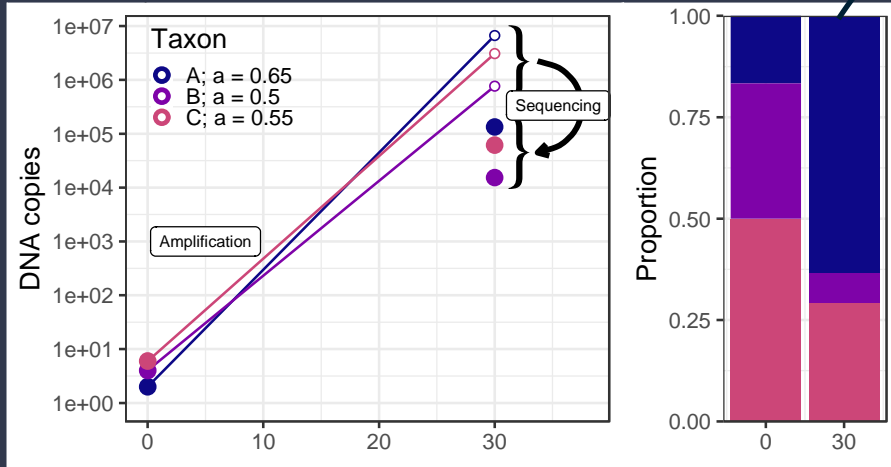
Site Level



Regional Level

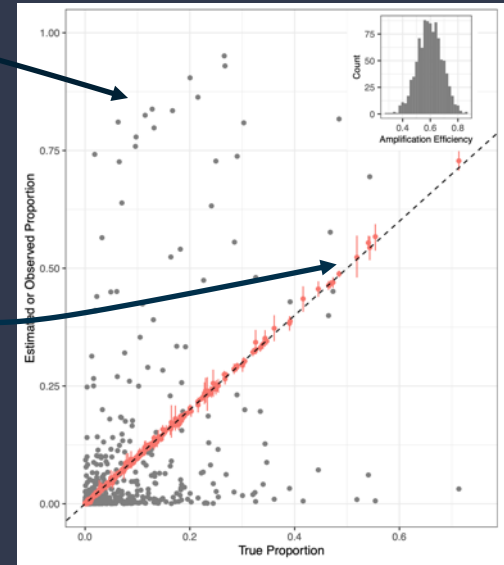
# Important Note on Metabarcoding

Amplification Bias: Different species may/do amplify at different rates



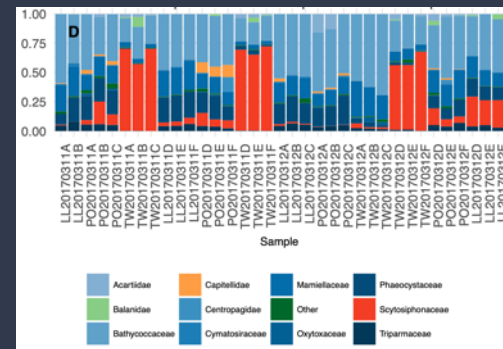
Uncalibrated

Calibrated



# Other Important Note on Metabarcoding

The resulting dataset is compositional and contains little/no info about absolute concentration



$\mathbf{Y} \sim \text{Multinomial}(\boldsymbol{\mu}, N),$

$$\mu_i = \frac{e^{\nu_i}}{\sum_{i=1}^I e^{\nu_i}},$$

$$\nu_i = \beta_i + N_{\text{PCR}}\alpha_i + \epsilon_i,$$

$$\epsilon_i \sim N(0, \tau_i),$$

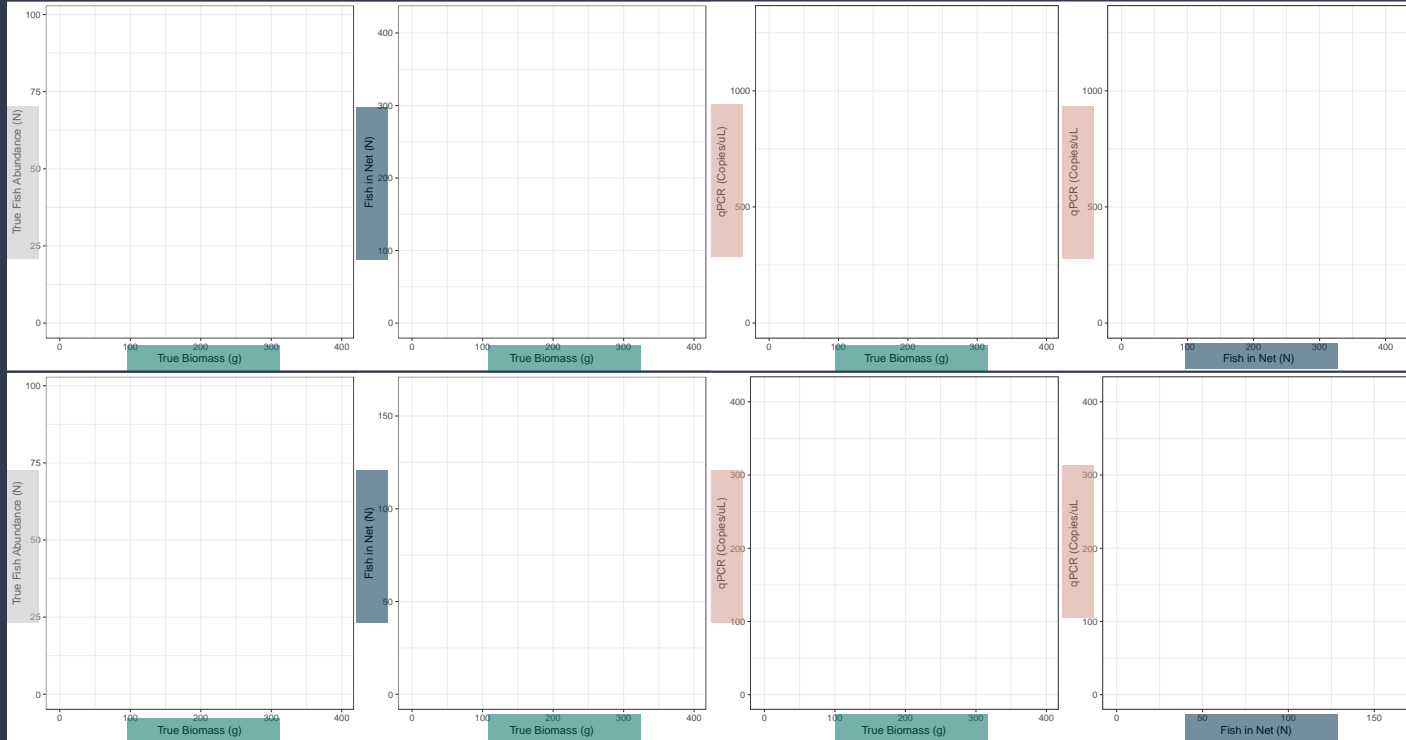
- We model the PCR process for each species in a composition
- Getting to absolute abundance requires additional information
  - e.g., visual counts, trawl, trap, whatever

# Compounding Noise

Variable mass-per-fish

1:1 eDNA copies per unit mass

Moderate  
Observation  
Variability



Low  
Observation  
Variability

The spatial/temporal scale of process + observation variability determine what correlations might exist