Building a Comprehensive and Accessible DNA Barcode Database for Fish and Invertebrates in the San Francisco Estuary

Ravi Nagarajan[1], Ann Holmes[1], Danielle Myers[1], Julia Mouat[2], Alisha Goodbla[1], Mallory Bedwell[3], Melinda Baerwald[3], Brian Schreier[3], Andrew Rypel[4], Larry Brown[5], & Andrea Schreier[1]

Background

- Environmental DNA (eDNA) metabarcoding methods rely on DNA barcode reference sequences for each species.
- Some SFE species already have DNA barcode reference sequences available in public databases, but some do not have have barcode data and many have only partial data.
- We are building a custom DNA barcode reference sequence database for SFE fish and invertebrates.
- This database is the first phase of a multiyear Prop 1 funded project to develop eDNA metabarcoding methods to complement existing SFE monitoring.
- The custom database will be made public as a resource to all scientists and will enable accurate taxonomic assignments from eDNA metabarcoding sequencing data. The data will also be useful for other genetic applications.

Approach

Collect SFE specimens → Extract DNA → PCR amplify ~5 barcode regions → DNA sequencing → Assemble database



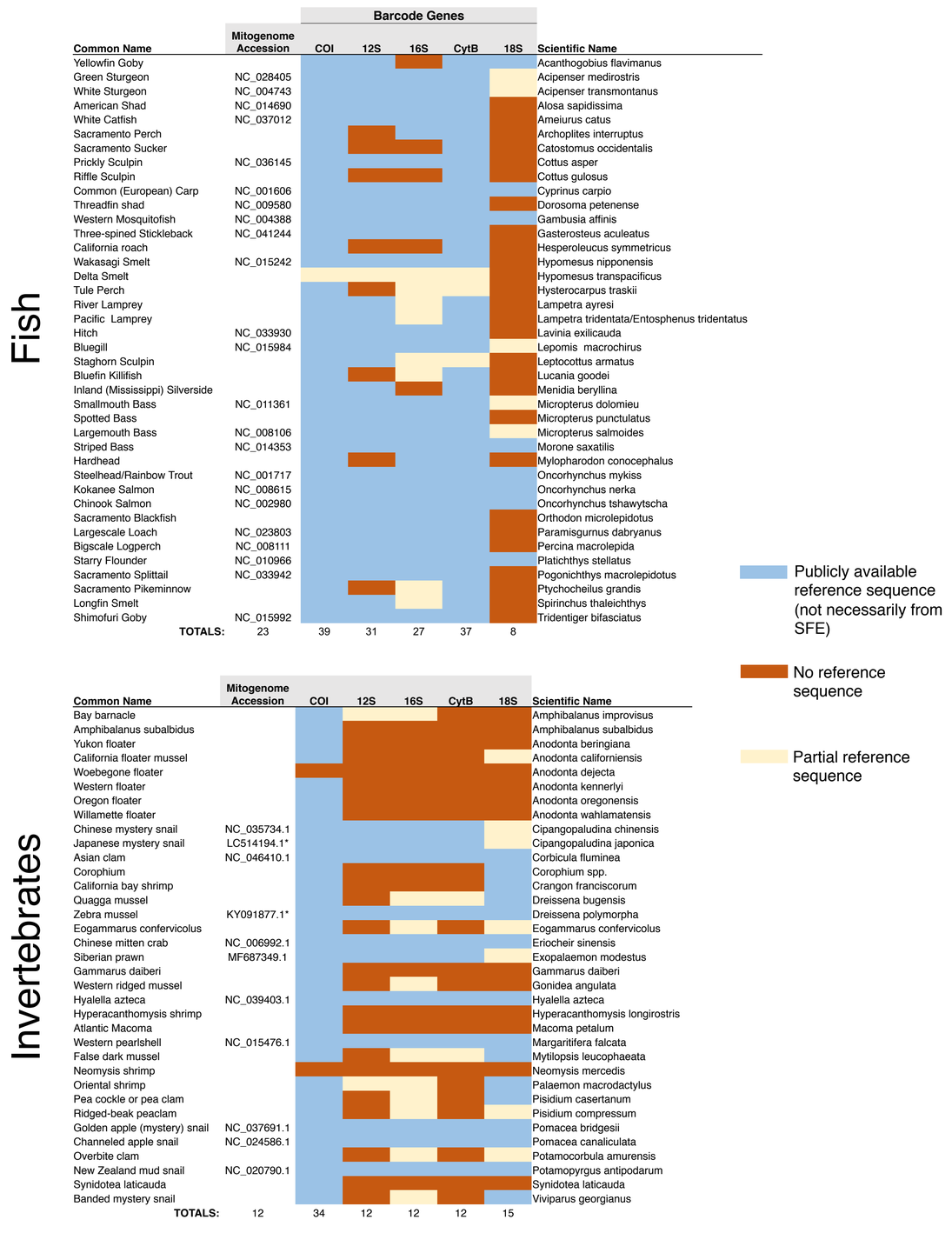
Collection of invasive zebra mussels (*Dreissena polymorpha*) from San Justo Reservoir, CA. Photos credit: Heather Perry, CDFW

- Specimen collection and DNA extractions ongoing
- PCR primer testing complete (see below)
- Goal for database is to sequence ~5 PCR products for each specimen
  - Some barcoding genes contain multiple barcodes
  - eDNA metabarcoding accuracy is improved by sequencing >1 barcode for each sample
  - For reference database sequencing, we are prioritizing PCR primers that will give us sequence information for multiple barcodes from one PCR/sequencing reaction
- DNA sequencing started July 2020 (delayed due to COVID-19 lab shutdown)
- Some extracted DNAs will also be used to create known DNA mixtures to test eDNA metabarcoding protocol prior to using with eDNA samples

PCR primers tested for Sanger DNA sequencing:

Gene	Type	Primer names (Forward/Reverse)	Reference
Cytochrome C Oxidase 1 (COI)	mitochondrial	jgLCO1490/jgHCO2198	Geller et al 2013
12S rRNA	mitochondrial	tRNA-Phe/teleo_R	Doble et al 2019
12S rRNA	mitochondrial	MiFish-U-F/MiFish-U-R	Miya et al. 2015
16s rRNA	mitochondrial	Ac16s-F/Ac16s-R	Evans et al. 2015
16s rRNA	mitochondrial	16s_Metazoa_fw/16s_Metazoa_rev	Shelton et al 2016
16s rRNA	mitochondrial	16S-V5-F/16S-V5-F	Riaz et al. 2011
Cytochrome B (CytB)	mitochondrial	L14912/H15149c	Burgener and Hubner 1998
18S rRNA	nuclear	SSU-F1/SSU-R568	Tanabe et al 2015
18S rRNA	nuclear	SSU-F1289/SSU-R1772	Tanabe et al 2015

Target Species



- Most SFE fish and invertebrate species have existing COI reference sequences
- References for other commonly used barcode genes (12S, 16S, Cytb, and 18S) are still needed for many SFE species
- Our objective is to fill in as many of these gaps as possible

Summary

- Database expected completion by **December 2020.**
- The database will improve the accuracy and utility of eDNA metabarcoding.
- Will enable biodiversity monitoring using eDNA metabarcoding, complementing and augmenting existing IEP monitoring.

Author Affiliations

1. Genomic Variation Laboratory, UC Davis
2. Integrative Genetics and Genomics Graduate Group, UC Davis
3. CA Dept of Water Resources
4. Department of Wildlife, Fish, and Conservation Biology, UC Davis
5. US Geological Survey

References

- Burgener and Hubner (1998) Mitochondrial DNA enrichment for species identification and evolutionary analysis. European Food Research and Technology 207(4) 261-3.
- Doble et al (2019) Testing the performance of environmental DNA metabarcoding for surveying highly diverse tropical fish communities: A case study from Lake Tanganyika. Environmental DNA 2(1) 24-41 DOI: 10.1002/edn3.43.
- Evans et al (2015) Quantification of mesocosm fish and amphibian species diversity via environmental DNA metabarcoding. Mol Ecol Res 16, 29-41.
- Geller et al (2013) Redesign of PCR primers for mitochondrial cytochrome c oxidase subunit I for marine invertebrates and application in all-taxa biotic surveys. Mol Ecol Res 13, 851–861
- Miya et al (2015) MiFish, a set of universal PCR primers for metabarcoding environmental DNA from fishes: detection of more than 230 subtropical marine species. R. Soc. open sci. 2: 150088.
- Riaz et al (2011) ecoPrimers: inference of new DNA barcode markers from whole genome sequence analysis. Nuc Acids Res 39(21). doi:10.1093/nar/gkr732.
- Shelton et al (2016) A framework for inferring biological communities from environmental DNA. Ecological Applications, 26(6), 1645–1659.
- Tanabe et al (2015) Comparative study of the validity of three regions of the 18S-rRNA gene for massively parallel sequencing-based monitoring of the planktonic eukaryote community. Mol Ecol Res 16(2) 402-414 doi: 10.1111/1755-0998.12459.

Acknowledgements

Funding: CDFW/CA Prop 1

Cory Saltzman (CDFW)
Sakura Evans (CDFW)
Martha Volkoff (CDFW)
Heather Perry (CDFW)
Ian Ralston (CDFW)
Chelsea Callahan (CDFW)
Jessica Weidenfeld (UC Davis)
Shayan Kaveh (UC Davis)
Khyana Yearwood (UC Davis)
Serra Perry (UC Davis)
Victor Martinez (UC Davis)
Samantha Barnum (UC Davis)
Ted Grosholz (UC Davis)
Betsy Wells (DWR)
Jesse Adams (DWR)
Suisun Marsh Fish Survey
Teejay O'Rear (UC Davis)
Dylan Stompe (UC Davis)
John Durand (UC Davis)
Lisa Thompson (Sacramento Area Sewer District & Sacramento Regional County Sanitation District)
Cara Wademan (UC Davis)
Grace Auringer (UC Davis)