

1 **SUPPLEMENT 1**

2 **Table of Contents**

3 **1. Methods**

- 4 a. *Study Design*
5 b. *Metabarcoding Collection Isolation, Amplification, and Sequencing*
6 c. *Bioinformatics*
7 d. *Microscopy Identification of Ichthyoplankton*
8 e. *Estimating Abundance (see Supplement 2)*
9 f. *Environmental Covariates*
10 g. *Data analysis*

11 **2. Results**

- 12 a. *Change in Fish Assemblage Structure over Time*
13 b. *Metabarcoding Signal Does Not Degrade Over Time*

14 **3. Figures**

- 15 a. *Figure S1. station Map*
16 b. *Figure S2. Observed Sequence Reads vs. Observed Morphological Counts*
17 c. *Figure S3. Predicted Counts vs. Observed Morphological Counts*
18 d. *Figure S4. Predicted Sequence Reads vs. Observed Sequence Reads*
19 e. *Figure S5. Co-detection of Taxa By Metabarcoding and Microscopy*
20 f. *Figure S6. Temperature Associations in Fish Species*
21 g. *Figure S7. Southern Oceanic Species Drive Fish Community Shifts*
22 h. *Figure S8. Significant Species Occurrence and SST Correlations at each Station*

- 23 i. *Figure S9. Significant Species Occurrence and MWCT Correlations at each*
24 *Station*
- 25 j. *Figure S10. Significant Species Abundance and SST Correlations at each station*
- 26 k. *Figure S11. Significant Species Abundance and MWCT Correlations at each*
27 *station*
- 28 l. *Figure S12. Bar Plot of Significant Species Abundance and SST Correlations*
29 *Across All stations*
- 30 m. *Figure S13. Bar Plot of Significant Species Abundance and MWCT Correlations*
31 *Across All stations*
- 32 n. *Figure S14. Increased Abundance of Southern Mesopelagic Species with Higher*
33 *SST*
- 34 o. *Figure S15. Increased Abundance of Southern and Central Mesopelagic Species*
35 *with Higher SST*
- 36 p. *Figure S16. Heat Map of Abundances Over Time*
- 37 q. *Figure S17. NMDS Ordination of Species and Years*
- 38 r. *Figure S18. NMDS Ordination of Species and Samples*
- 39 s. *Figure S19. Heat Map of San Diego Offshore Abundances Over Time*
- 40 t. *Figure S20. NMDS Ordination of San Diego Offshore Species and Years*
- 41 u. *Figure S21. Heat Map of San Diego Inshore Abundances Over Time*
- 42 v. *Figure S22. NMDS Ordination of San Diego Inshore Species and Years*
- 43 w. *Figure S23. Heat Map of Pt. Conception Abundances Over Time*
- 44 x. *Figure S24. NMDS Ordination of Pt. Conception Species and Years*
- 45 y. *Figure S25. Heat Map of San Nicholas Island Abundances Over Time*

46 z. *Figure S26. NMDS Ordination of San Nicholas Island Species and Years*
47 aa. *Figure S27. Co-occurrence Patterns of Species Controlling for SST*
48 bb. *Figure S28. Co-occurrence Patterns of Species*
49 cc. *Figure S39. Differential Species Abundance Before and After the Marine Heat*
50 *Wave at the San Nicholas Island*
51 dd. *Figure S30. Stable Precision of Amplicon Abundance Over Time*
52 ee. *Figure S31. Stable Precision of Abundance Estimates Over Time*
53

54 **Introduction**

55 This supplemental material provides additional details on the methods, results, and
56 discussion to support the main findings and conclusion of the manuscript.

57 The CalCOFI program (<https://calcofi.com/>) serves to provide fisheries-independent
58 ecosystem assessments of fish assemblages in the Southern California Current and has provided
59 decades of data on ichthyoplankton assemblages (77). The current CalCOFI surveys sample four
60 times per year from the U.S. Mexican Border to Monterey Bay (77). We used this rich sample
61 archive to interrogate ichthyoplankton assemblages from 1996-2019 (See Supplemental
62 Methods). Importantly, the occurrence of a marine heatwave (MHW) within the study region and
63 sampling period provided an additional opportunity to investigate the utility of having a non-
64 destructive means of interrogating the valuable CalCOFI sample archive.

65

66 **Methods**

67 *Study Design*

68 To evaluate the efficacy of metabarcoding methods used to analyze ethanol preserved
69 samples and investigate potential changes in the ichthyoplankton assemblages over decadal
70 scales, we identified ichthyoplankton by metabarcoding and microscopy in ethanol-preserved
71 samples collected over two decades (1996, 1998-2019; Figure S5) of spring CalCOFI cruises (5,
72 27, 35, 38). We note that samples collected in 1997 were stored in <50% ethanol and were
73 discarded due to failed preservation.

74 Samples were collected in late March or early April of each year
75 (calcofi_metadata_analysis_20210907.csv). Here we focus on spring samples because the
76 majority of species in the California Current spawn in spring and historically the annual
77 California Current Ecosystem Report has relied on the spring data (20, 78). This decision is
78 supported by recent work using ichthyoplankton data across the full set of yearly CalCOFI
79 cruises which found little evidence for phenological trends (20, 78), thus aiding the ability to
80 look at impacts from the marine heatwave.

81 Samples were collected from four well-separated stations (up to 370 km apart) from
82 distinct vicinities of the California Current with differing water properties (61) (5, 27, 35, 38)
83 (Figure S1). The northernmost station was located offshore of Point Conception, CA within the
84 California Current (34.14833°N -121.1567°W). The second station was located off San Nicholas
85 Island, CA (33.32333 °N, -119.6667°W) that experiences high variation in annual temperature
86 depending on the respective strengths of the California Current and Southern California Counter
87 Current (61). The third station was a southern coastal inshore station off San Diego, CA
88 (32.84667°N, -117.5383°W) characterized by relatively warmer waters from the California
89 Counter Current with seasonal (spring) upwelling of cool, nutrient-rich water (61). The fourth

90 station was a southern offshore station (31.85000°N , $-119.5683^{\circ}\text{W}$) characterized by sub-tropical
91 oceanic waters (Figure 4).

92 At each station, oblique bongo net tows were conducted from 210 m to the surface using
93 standard CalCOFI methods (15, 64–66). Each side of the bongo net had a 0.71 m-diameter
94 mouth opening and a net size of 0.505 mm mesh. Cod end contents of both bongo nets were
95 preserved at sea. The starboard side was preserved in sodium borate-buffered 2% formaldehyde
96 and the port side was preserved in Tris-buffered 95% ethanol. Ethanol was replaced after 24
97 hours to account for dilution from tissue water loss. Microscopy was conducted to identify
98 species abundance from formaldehyde-preserved samples following standardized CalCOFI
99 techniques (61) while metabarcoding was conducted on the ethanol in which port side samples
100 were stored; consequently, we expected the contents of the paired samples to differ slightly as a
101 function of sampling stochasticity.

102 *Metabarcoding Collection Isolation, Amplification, and Sequencing*

103 Prior to filtration, the ethanol-preserved samples were inverted three times and let rest for
104 30 minutes to resuspend and homogenize samples in the preservative. Filtration of ethanol from
105 the port-side bongo samples was conducted in a pre-PCR clean room at the NOAA Southwest
106 Fisheries Science Center within a biological safety cabinet in July 2019. The pre-PCR room had
107 no previous post-PCR work conducted within and all surfaces and equipment were sterilized
108 frequently with 10% bleach and 70% ethanol. The pre-PCR clean room was at ambient pressure
109 and reasonable precautions to limit contamination were conducted including only wearing clean
110 clothes that have not been exposed to labs with PCR product, no food brought into the lab, and
111 gloves were exchanged regularly.

112 Ethanol preservative was filtered using a vacuum filtration manifold with Nalgene
113 Analytical Test Filter Funnels (Thermofisher Scientific, Waltham, MA, USA) with the
114 manufacturer's 0.45 µm filters replaced with 0.2 µm Durapore PVDF filters (Sigma Aldrich, St.
115 Louis, MO, USA) using sterile forceps. Up to 125 mL of ethanol was then transferred from the
116 preserved jars into the filter funnels using a 10 mL pipette, carefully avoiding any sample
117 contents and thus preserving CalCOFI specimens for future research and analysis. Sample jars
118 were refilled using freshly prepared tris-buffered ethanol before being returned to the collection
119 archive. We included two negative controls to test for lab contamination by filtering 125 mL of
120 molecular grade water. Filters were stored at -20°C before DNA extraction.

121 Filters were extracted using the standard Qiagen DNAeasy Kit (Qiagen Inc., Valencia,
122 CA, USA) in a pre-PCR molecular lab. Extracted DNA was amplified using the MiFish
123 Universal Teleost primer sets to capture fish diversity (67).

124 Here, we highlight our decision to utilize the MiFish Universal Teleost *12S* primers. First,
125 these primers have been rigorously validated for fish barcoding (22, 23, 29, 30, 79–81) and
126 shown to provide accurate taxonomic assignments for a broad range of fishes (22). We recognize
127 that there are limitations for this, and indeed all, metabarcoding primer sets (82) which are forced
128 to balance specificity [how well target species can be taxonomically resolved] against breadth
129 [range of species across the tree of life that can be amplified] (83). Even a “gold standard” like
130 the *16S* rRNA gene marker for prokaryotic sequences struggles with taxonomic assignment
131 accuracy (84), especially with short-read sequences. Although taxonomic resolution limitations
132 and compromises remain for the *12S* target (22, 85), the taxonomic resolution has been improved
133 and best practices for taxonomic classification have been identified through the development of a

134 nearly comprehensive California Current Large Marine Ecosystem *12S* reference database along
135 with a full factorial cross-validation analysis of bioinformatic approaches (22).

136 Second, there are no widely used or benchmarked *COI* metabarcoding primer sets for
137 fish applications although *COI* barcoding is a common barcoding target. This is because a) the
138 conserved nature of the locus across the tree of life which results in amplification of a broad
139 array of taxa (86, 87), and b) the mismatch in high throughput sequencing platform length (max
140 is paired-end 300 bp) and rate of *COI* evolution/accumulation of sequence differences between
141 species (81, 88). In fact, these shortcomings were the original motivation for researchers to
142 develop alternative fish metabarcoding loci targeting *12S* loci for fishes (23). Together, the
143 research community has largely converged on the MiFish Universal Teleost *12S* primer set as
144 standard practice for fish metabarcoding given its balance of high specificity and breadth (23).
145 Thus we feel confident that the MiFish Universal Teleost *12S* primer set was an appropriate
146 choice for metabarcoding here.

147 Each metabarcoding extraction was subsampled for three PCR reactions using the MiFish
148 *12S* primer set. PCR amplification for the MiFish primer set was conducted following the
149 thermocycler profile of Curd *et al.* (30). MiFish PCR reactions had 25 µL reaction volume
150 containing 12.5 µL QIAGEN Multiplex Taq PCR 2x Master Mix (Qiagen Inc., Valencia, CA,
151 USA), 6.5 µL of molecular grade water, 2.5 µL of each primer (2 µmol/L), and 1 µL DNA
152 extraction. MiFish PCR thermocycling employed a touchdown profile with an initial
153 denaturation at 95°C for 15 min to activate the DNA polymerase, followed by 13 cycles of a 30s
154 denaturation at 94°C, a 30s annealing that started at 69.5°C and then decreased by 1.5°C for each
155 subsequent cycle (last cycle was 50°C), finishing with a 1 min extension at 72°C. This initial

156 touchdown profile was followed by 35 additional cycles using identical parameters except a
157 constant annealing temperature of 50°C and ending with a final extension at 72°C for 10 min.

158 Two non-native non-marine vertebrates, American alligator (*Alligator mississippiensis*)
159 and dromedary camel (*Camelus dromedarius*), were purchased at a local market and used as
160 positive controls. For all positive controls, tissues were extracted using the Qiagen Blood and
161 Tissue kit following the manufacturer's instructions. All PCR products were visualized via
162 electrophoresis on 2% agarose gels to ensure amplification success and correct product size.
163 Only filters from four jars failed to amplify, and upon further inspection within the archived
164 notes, all these samples had known preservation issues (e.g., preservative dried out, observed
165 mold, etc.). All other DNA extractions successfully amplified.

166 We prepared libraries following the methods of Curd *et al.* using a two-step PCR
167 amplification method with one final pool per primer set. Previous work indicated that two-step
168 PCR amplification can reduce amplification biases (89, 90) perhaps introduced by the inclusion
169 of various indices during one-step PCR procedures. Variations in the relative amplification
170 efficiency of each PCR is a concern here given the desire to study an array of targets in an
171 oceanic region over space and time. Overall, there are review papers available that outline the
172 advantages and disadvantages for one-step and two-step PCR protocols (91).

173 Prior to the second indexing PCR reaction, PCR samples from the first reaction were
174 cleaned using the Serapure magnetic bead protocol. We quantified bead-cleaned samples with
175 the Quant-iT™ broad range dsDNA Assay Kit (Thermofisher Scientific, Waltham, MA, USA)
176 on a Victor3 plate reader (Perkin Elmer Waltham, MA, USA). We indexed the sample libraries
177 using unique combinations of the Nextera Index A, B, C, and D Kit (Illumina, San Diego, CA,
178 USA) and KAPA HiFi HotStart Ready Mix (Kapa Biosystems, Sigma Aldrich, St. Louis, MO,

179 USA). Indexing was performed with a second PCR using a 25 µL reaction mixture containing
180 12.5 µL of Kapa HiFi Hotstart Ready mix, 1.25 µL of index primers, 10 ng of template DNA to
181 ensure equal copy number, and the remaining volume was filled using molecular grade water
182 depending on cleaned PCR product concentration. Index thermocycling parameters were:
183 denaturation at 95°C for 5 min, 5 cycles of denaturation at 98°C for 20 sec, annealing at 56°C for
184 30 sec, extension at 72°C for 3 min, followed by a final extension at 72°C for 5 min. To confirm
185 successful PCR and correct product size, we electrophoresed PCR products on 2% agarose gels.
186 We then bead cleaned and quantified DNA concentration, as described above so that we could
187 pool samples so as to have equal copy number for each unique library. Pooled libraries were
188 sequenced on an Illumina NextSeq PE 2x150 at UCLA Technology Center for Genomics and
189 Bioinformatics.

190 *Bioinformatics*

191 The resulting metabarcoding data were processed using the *Anacapa Toolkit* to conduct
192 quality control, amplicon sequence variant (ASV) parsing, and taxonomic assignment using user-
193 generated custom reference databases. We processed sequences using default parameters except
194 using a Q score cutoff of 30 and assigned taxonomy using *CRUX*-generated metabarcode
195 specific reference databases (22). The MiFish sequencing data was assigned taxonomy using the
196 California fish specific reference database and a bootstrap confidence cutoff score of 60
197 following Gold *et al.* (22).

198 The two resulting raw ASV community tables were decontaminated following Kelly *et al.*
199 (92). First, only merged paired reads that occurred at least twice (e.g., no singletons) were
200 retained. Second, we estimated index hopping between samples by calculating the proportion of
201 sequences within the positive control samples and then subtracting reads from each sample by

202 the sample read depth multiplied by the proportion of reads observed in the positive controls.
203 Third, we discarded technical replicates with fewer than 30,000 reads. Fourth, we calculated
204 Bray-Curtis dissimilarities between technical PCR replicates and fit a skewed beta distribution
205 ($a= 0.6$, $b= 9.5$). We then removed all replicates with greater than 95% probability of belonging
206 to the beta distribution. Resulting tables were then combined into a final ASV community table
207 in *R*.

208 *Microscopy Identification of Ichthyoplankton*

209 Plankton samples were processed at the NOAA Southwest Fisheries Science Center
210 ichthyoplankton laboratory. From each plankton sample, fish larvae were sorted and identified
211 through microscopy to the lowest practical taxon (64, 66). Most taxa were identified to species
212 although some were only characterized to genus or family level (See
213 `larval_counts_20210305.csv`). The number of larvae per species per jar, total abundance of
214 filtered ichthyoplankton, and proportion of jar sorted were recorded.

215 *Estimating Abundance*

216 We estimated the abundance of ichthyoplankton in each jar using a novel joint Bayesian
217 hierarchical model described in detail in Supplement 2.

218 *Environmental Covariates*

219 We specifically examined the relationship of ichthyoplankton communities to sea surface
220 temperatures (SST). Two month prior mean SSTs were obtained using the *rerddapXtracto*
221 package (93) in *R* to collect PathFinder Ver 5.3 monthly remotely sensed composites. To
222 calculate two-month prior means we first obtained monthly composites from April 1995 to April
223 2019 for each station. We then averaged across monthly composite sea surface temperatures

224 ignoring any missing values. Prior two month sea surface temperatures were chosen given the
225 average age of spring larvae (94) (Figure 4).

226 We then investigated the relationship of ichthyoplankton assemblages against sea surface
227 temperature as a proxy for a multitude of environmental shifts associated with the MHW. In
228 addition, we specifically characterized fish that are uniquely present or absent before (1996-
229 2013) and after the 2014-16 Marine Heatwave (2014-2019). Analyses were repeated using the
230 mean average water column temperature (MWCT) obtained from nearly simultaneously
231 conducted CTD rosette deployments (15). The mean water column temperature was averaged
232 across 10 to 100m depth, where the majority of ichthyoplankton reside, following Thompson et
233 al. (15).

234 *Data Analysis*

235 After model estimation, we calculated mean abundance estimates (larvae counts per
236 standardized volume towed) per species per station per year. To explore species-specific sea
237 surface temperature (SST) relationships, we fit a Bayesian generalized linear model using log
238 (abundance) as the response variable and SST ($^{\circ}\text{C}$) as a continuous predictor variable. Models
239 were implemented for each species using Stan as implemented in R (75). We then summarized
240 the affinity between each species and SST by calculating a T-statistic based on each species'
241 estimated coefficients (mean slope/standard deviation). We further plotted the estimated slope
242 for each “species grouping” by habitat associations derived from previous CalCOFI research
243 (habitat_association_to_check_art.csv)(63). We summed total log(abundance) per habitat
244 association per station per year and fit a Bayesian generalized linear model using log (total
245 abundance) as the response variable and SST ($^{\circ}\text{C}$) as a continuous predictor variable.

246 We repeated the above analyses using a Bayesian binomial model using presence as the
247 response variable and SST ($^{\circ}\text{C}$) as a continuous predictor variable across the data set to explore
248 occurrence relationships with temperature and identify warm- and cool- associated taxa. We set a
249 threshold of presence/absence based on the model using a threshold of < 0.01 larvae per
250 standardized volume to be considered absent within a station.

251 We further explored species occurrence and abundance relationships with SST by fitting
252 the above Bayesian generalized linear models with station as a random effect. In addition, we
253 repeated all of the above analyses using MWCT instead of SST.

254 To explore how fish assemblages change over time we plotted a heatmap of observed
255 abundance summed across stations each year. Chronological clustering was conducted across
256 years using Bray Curtis dissimilarities of abundances using a K of 8 using the package *rioja* in R
257 (95) and a dendrogram of years was constructed using the *ggdenro* package (96). Similarly,
258 hierarchical clustering was conducted across species using Bray Curtis dissimilarities of
259 abundances using a K of 6. To further explore fish assemblage changes NMDS Ordination of
260 Bray-Curtis dissimilarities were calculated from estimated abundances of each year summed
261 across stations as implemented by the *metaMDS* function from *vegan* in R (97). The above
262 analyses were also conducted with station separated as well as each station on it's own. To
263 investigate the relative effect of year, SST, and station to the explained variance in fish
264 assemblages across the ata set, we ran a PERMANOVA on Bray-Curtis dissimilarities using the
265 following model: ~ Year + SST + station.

266 We visualized anchovy and sardine abundance over time by calculating the median log
267 (abundance) of each species per station per year. We then plotted the log (median) abundance of

268 each of the four stations while error bars represent the 95% confidence intervals observed for a
269 given species at a given station in that year.

270 To explore co-occurrence patterns across species, we fit a generalized linear latent
271 variable model (GLLVM) following the methods of Niku et al. (98), specifically conducting
272 model fitting to determine the best distribution fit as well as number of latent variables to use.

273 The highest performing GLLVM employed 3 latent variables and applied a negative binomial
274 distribution with variational approximation (99) on the joint model predicted larvae count data.

275 We then plotted the correlation matrix of the linear predictors across species with and without
276 incorporating SST in the GLLVM to identify co-occurring species and the effect of SST on co-
277 occurrence patterns.

278 To evaluate the effect of the marine heatwave (MHW) on CCLME fishes we compared
279 estimated species abundances before the MHW (1996-2013), to both during and after the MHW
280 (2014-2019), at each station respectively. We first calculated the mean abundance for each
281 species at each station for each model run. We then subtracted the means for each model run to
282 evaluate changes in MHW abundance per species per station per model run. We then calculated a
283 95% CI of change in MHW abundance per species to identify which species were significantly
284 different before vs. during and after the MHW at each station.

285 All data and code to conduct analyses and generate all figures are available on GitHub
286 (https://github.com/zjgold/CalCOFI_eDNA) and associated Google Drive link
287 (https://drive.google.com/drive/folders/12cU9mY_CWoro-x6Hgh_pgv_66zZEzm1h?usp=sharing) [will be replaced with a Dryad repository upon
289 acceptance].

290

291 **Results**

292 *Fish Assemblage Structure*

293 We observed substantial changes in fish assemblage structure across stations, time, and
294 temperature sampled (NMDS stress =0.03) (Figure S11-26). station explained the greatest
295 observed variance (12%) which is unsurprising given the intentionally chosen distinct
296 biogeographic characteristics of each station (PERMANOVA p <0.05). However, despite the >
297 370km distance between stations, we captured significant synchronous changes in fish
298 assemblage dynamics in response to year (2.4%) and temperature (4.6%) (PERMANOVA p
299 <0.05). In particular, we observed strong clustering of the post MHW period from 2017-2019,
300 the 2005 El Niño and the 1998 El Niño along with southern mesopelagic species. Both 2014 and
301 2016 were distinct from other years and associated with a suite of mesopelagic species, although
302 the MHW itself was not strongly clustered largely due to the differential onset and
303 characterization of the warming event within the region (5).

304 We also found strong positive and negative co-occurrence patterns through GLLVM
305 analyses across species when controlling for temperature. Specifically, we observed strong
306 negative co-occurrence patterns between fisheries targets like benthic fisheries targets (e.g.
307 *Citharichthys* sp. sanddabs) and mesopelagic fishes (S27). We also observed strong positive co-
308 occurrence patterns between a suite of benthic species as well as strong positive co-occurrence
309 patterns between a suite of mesopelagic fishes. These results suggest that when controlling for
310 temperature, we observe strong benthic versus pelagic tradeoffs as observed previously (13, 20,
311 62).

312 When focusing on co-occurrence patterns without controlling for temperature, we
313 observed strong negative co-occurrence patterns with fisheries targets (North Pacific Hake

314 *Merluccius productus*) and mesopelagic fishes (S28). Here, temperature explained 19% of the
315 variability among species co-occurrence and was particularly important in driving negative co-
316 occurrence patterns between North Pacific Hake and mesopelagic species. These results suggest
317 that temperature may mediate tradeoffs between fisheries versus southern mesopelagic fish
318 assemblages. Further work exploring the underlying mechanisms of these negative co-
319 occurrence patterns is warranted.

320 *Analysis Using Mean Water Column Temperature*

321 We repeated the analyses presented in the main manuscript with mean water column
322 temperature (MWCT) as opposed to two-month SST (Figures 1-3). Two-month averaged SST
323 and instantaneous MWCT were only 29% correlated (linear regression, $p < 0.01$). This finding is
324 perhaps unsurprising given the substantial difference in spatial and temporal integration time of
325 these temperature measurements. Despite these apparent differences, we found nearly identical
326 results in the direction and significance of species–temperature associations as well as
327 temperature-driven variation in fish assemblage dynamics (Figures S6–S13). These results
328 suggest that species–temperature associations and our conclusions are largely robust to
329 temperature metrics.

330 *Metabarcoding Signal Appears Stable in the Ethanol-Preserved Samples*

331 For each station-species combination, if metabarcoding signals appear auto-correlated in
332 time -- that is, if one year's metabarcoding signal is correlated with the previous year's signal --
333 then we require a time-series model that incorporates such autocorrelation into the error
334 structure. If, by contrast, years appear independent of one another, we can treat model variation
335 as time-independent and therefore treat each data point as being independent. We observe no

336 such correlation (mean = -0.014, standard deviation = 0.35) and so we treat all observations as
337 independent of one another.

338 In further investigating the question of whether these samples can be considered time-
339 independent, we considered whether or not older samples might have less metabarcoding signal
340 due to sample degradation. If the metabarcoding signal were degrading overtime in the preserved
341 samples, we would expect several parameters to change as a function of sample age: (1) a
342 decrease in precision with which we observe amplicon abundance, (2) a decrease in richness of
343 species detected, and (3) a decrease in the confidence in posterior estimates of larval abundances
344 from our joint Bayesian model. We test for these effects in turn.

345 First, among triplicate PCR reactions, we might expect degraded DNA to behave more
346 stochastically than non-degraded DNA, such that technical replicates would yield increasingly
347 divergent amplicon abundances with greater degradation. Here, we measure the precision of our
348 estimates with the coefficient of variation (CV) of species-specific amplicons across three
349 technical replicates. An increase in CV with the age of the sample would signal degradation, but
350 we saw no such trend (Figure S29). Second, rare amplicons often make up a large fraction of
351 metabarcoding datasets, and because of their rarity, these often show up stochastically across
352 replicates or sequenced samples. If older DNA samples were degraded, we would expect fewer
353 of these rare species, and by extension, fewer species overall. We saw no such effect (linear
354 regression $p > 0.5$; linear mixed effect model failed convergence). Third, we might expect -- if
355 DNA were degrading -- that such degradation would impair our ability to estimate the larval
356 abundance of each species in older samples. Again, we saw no evidence of this effect (Figure
357 S30-31).

358 *Overlap in Species Detections*

359 The maximum observed morphological counts in which metabarcoding failed to detect a given
360 taxa was 9 (mean = 1.61). Across a total of 4,704 possible detections, 70.2% were non-detections
361 by both methods, 11.2% were detections by both methods, 16.4% were detections only made by
362 metabarcoding, and 2.1% were detections only made my microscopy (Figure S4).

363 *Stochasticity in Metabarcoding Data*

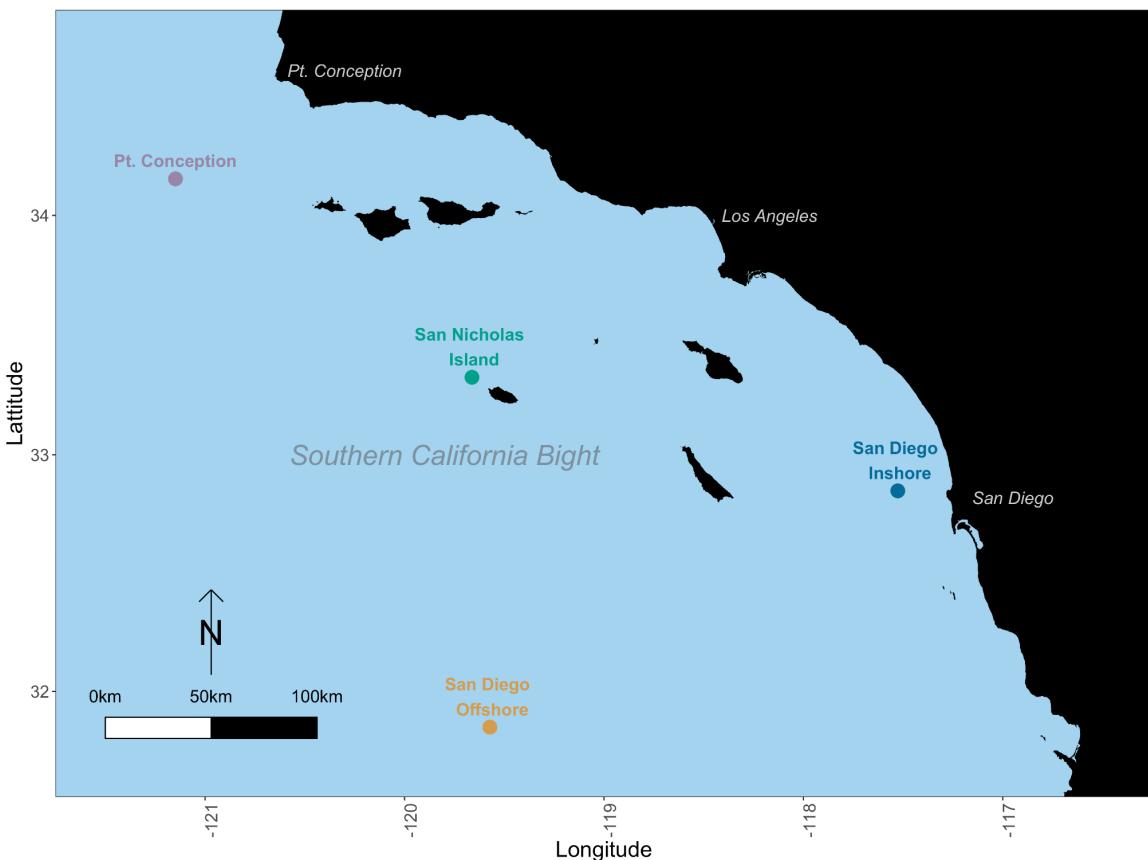
364 We conducted a deep dive into the origin and source of variation in amplicon sequence data.
365 These analyses identified stochastic dropouts in whereby a taxon is amplified in one PCR
366 reaction but not in a replicate PCR reaction as a main driver of variation in this data set. This is a
367 well-known phenomenon general to PCR with rare templates (69, 100, 101) and we document
368 such behavior in this manuscript (Figures S2-S4). For example, for *Symbolophorus*
369 *californiensis* we observed an instance of 3,897 reads, 165 reads, and 0 reads across three
370 technical PCR replicates with sample read depths of 132,731, 196,260, 55,400 from the same
371 DNA extraction. These stochastic dropouts are easily visualized along the X axis in Figure S2.
372 We note that the highest observed species-specific amplicon sample read proportion associated
373 with a stochastic dropout was 2.9% (3,897 /132,731) with the vast majority of such dropouts
374 occurring below 0.03% read proportion within a technical replicate. These results suggest that
375 stochasticity is largely driven by the abundance of DNA molecules within a sample rather than a
376 specific feature associated with a particular primer set, especially given that dozens of other
377 metabarcoding studies have identified similar patterns (69, 70, 100).

378 This phenomenon of stochastic dropouts adds noise to the observations and limits the
379 accuracy with which we might predict amplicon abundances (particularly rare ones). This is
380 best visualized by the noise near the origin at Figures S3 & S4. To address this, we developed a
381 comprehensive joint Bayesian model that incorporates stochasticity in observed amplicon read

382 counts through a multinomial subsampling process (See supplement 2 for full model
383 description). Thus, we explicitly account for stochasticity in the model through sampling
384 distributions and using the resulting parameters to estimate the uncertainty around our given
385 estimated larvae counts. Ultimately, such noise in the dataset does not fundamentally change
386 the interpretation of our observations or of our model but serves to limit our confidence in the
387 abundance of rare targets, a persistent problem in community ecology (102).

388

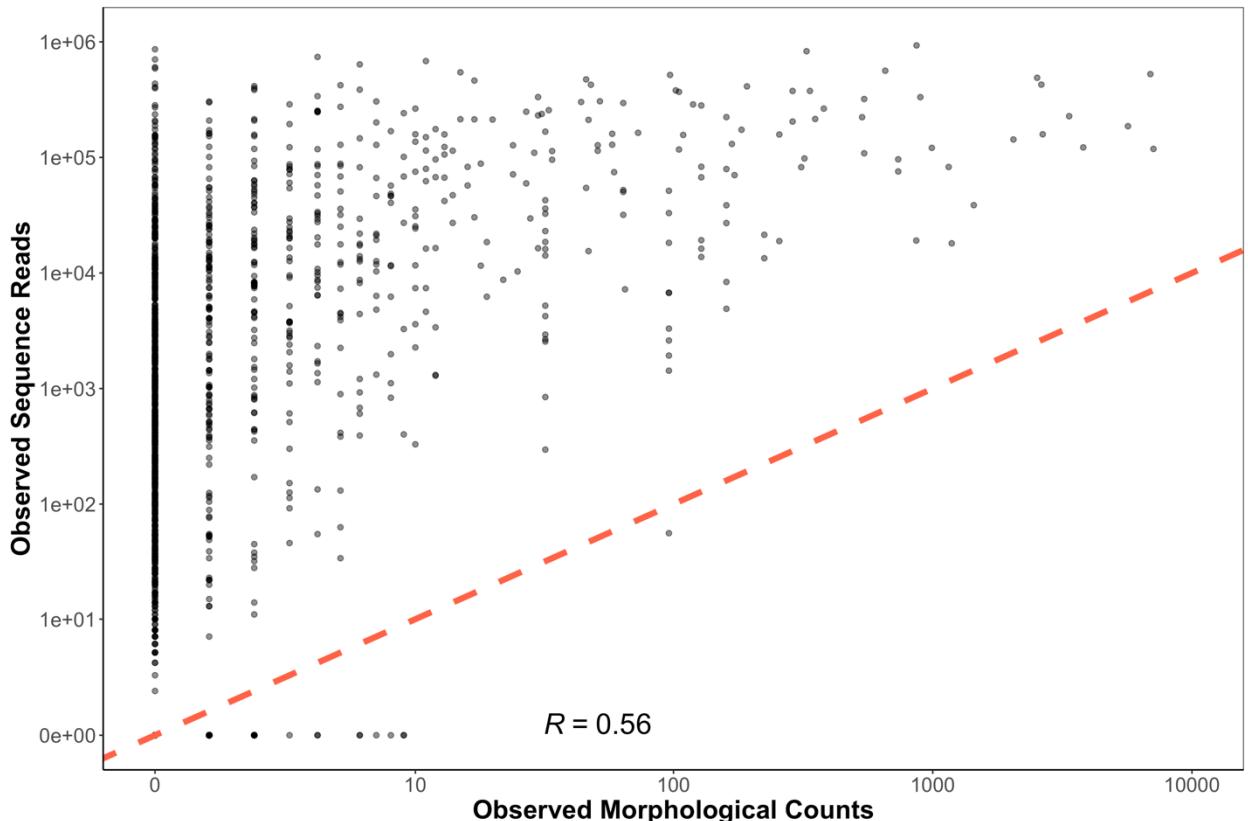
389 **Supplemental Figures**



390

391 **Figure S1. station Map**

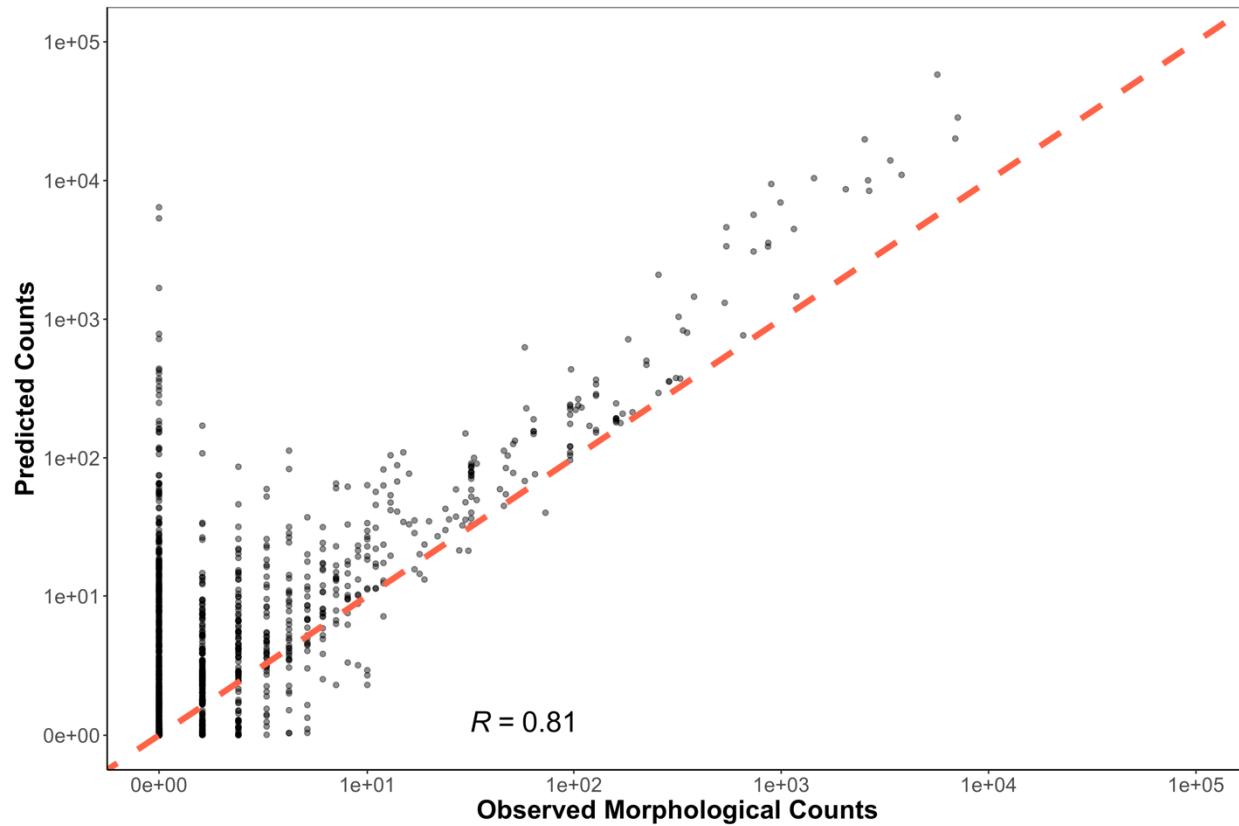
392 Ichthyoplankton samples were collected from four stations with distinct biogeographic
393 characteristics.



394

395 Figure S2. Observed Sequence Reads vs. Observed Morphological Counts

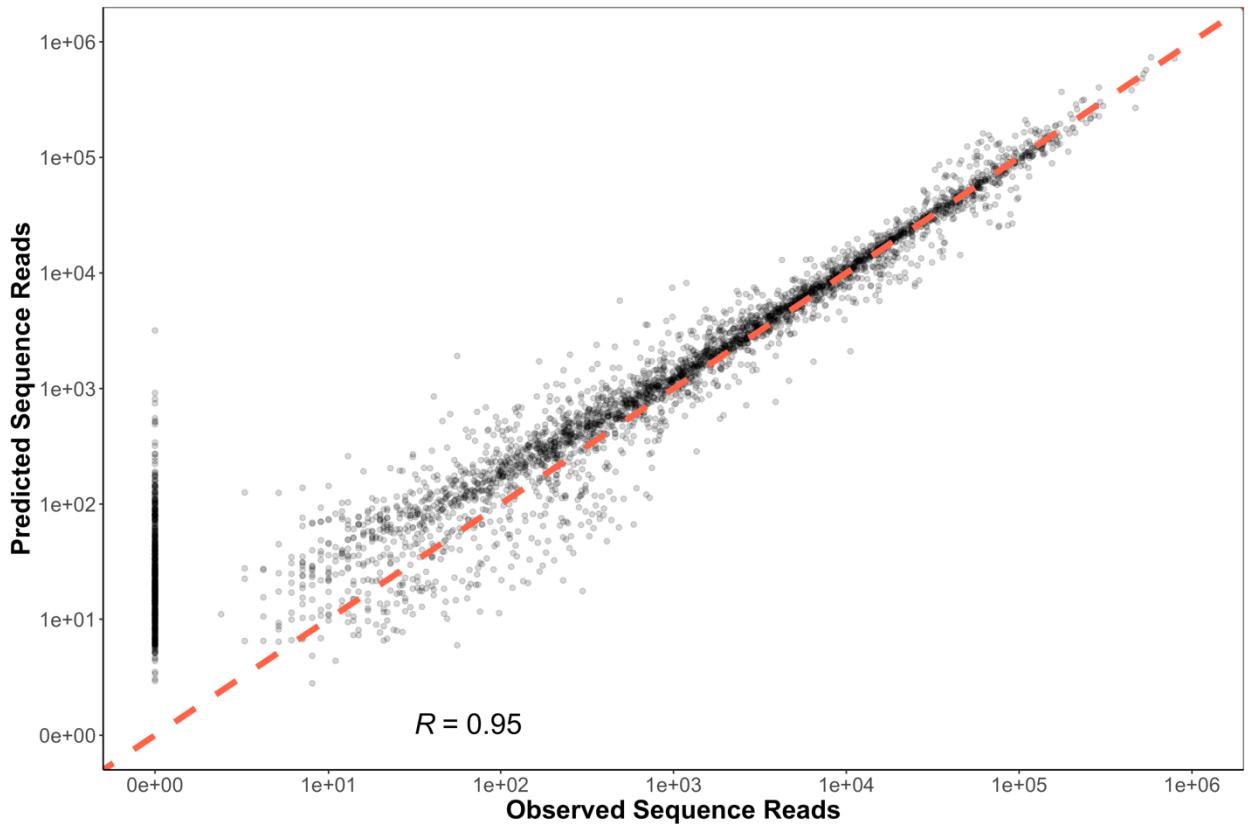
396 Observed sequencing reads and morphological counts do not follow a clear linear
 397 relationship. The one-to-one line is plotted in red and Pearson correlation coefficient is
 398 0.56. This non-linearity is unsurprising given that observed reads are a function of both
 399 DNA concentrations (here assumed proportional to morphological counts) as well as
 400 species-specific amplification efficiencies (here are unknown) (See Supplement 2). Thus
 401 without accounting for species-specific amplification efficiencies within the compositional
 402 amplicon data set, we do not expect to observe a clear positive relationship. This apparent
 403 lack of relationship depicted here motivated the creation of the mechanistic joint Bayesian
 404 model. The occurrence of stochastic dropouts (technical replicates with zero reads) can be
 405 observed along the X-axis. Variance is highest at low observed morphological counts.



406

407 **Figure S3. Predicted Counts vs. Observed Morphological Counts**

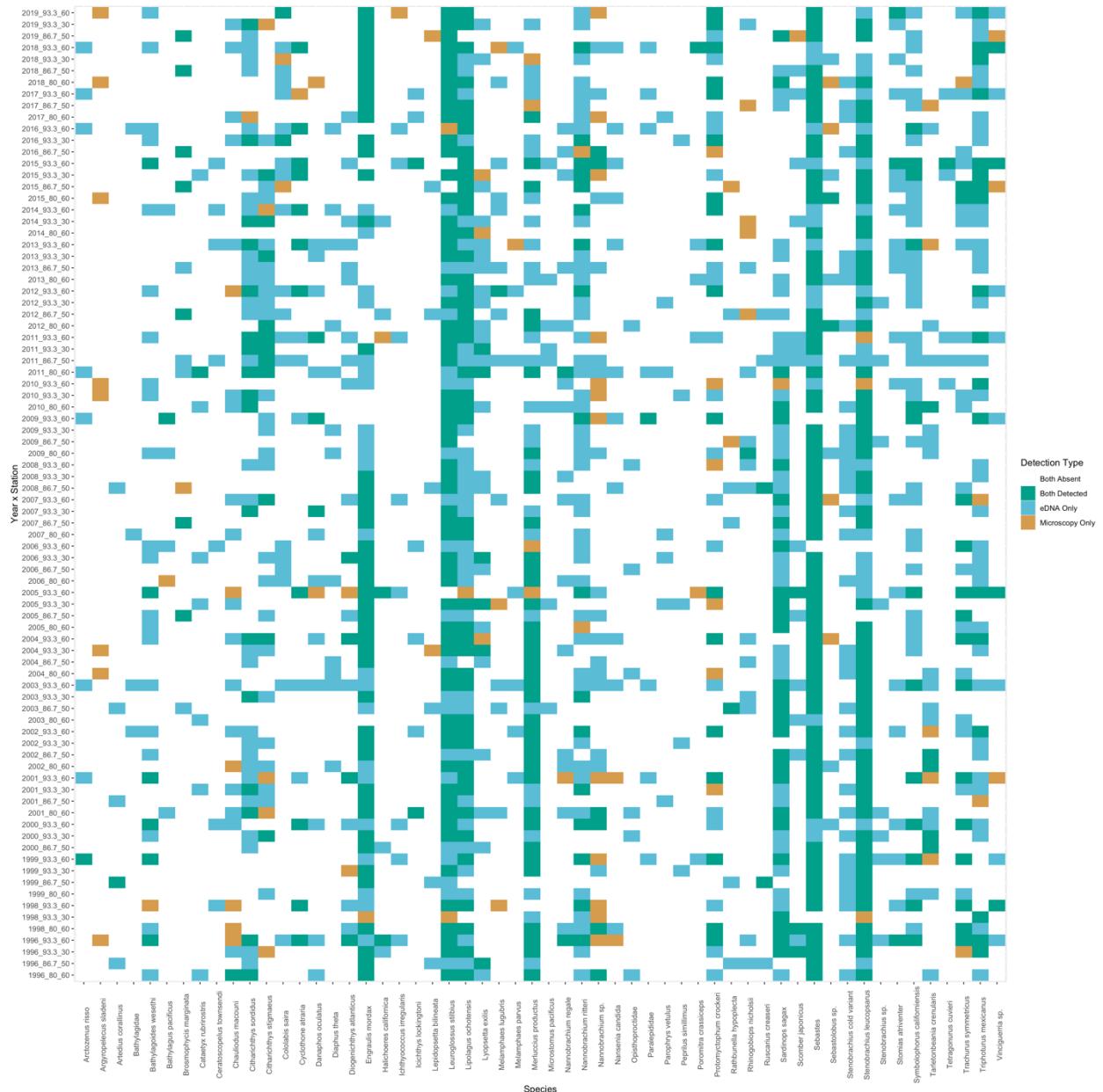
408 Predicted counts are generated from the joint Bayesian model. The one-to-one line is
 409 plotted in red and Pearson correlation coefficient is 0.81. Variance in predicted counts
 410 markedly decreases with higher observed morphological counts. We note variance in
 411 predicted counts is substantially less than that between observed reads and morphological
 412 counts Figure S3.



413

414 Figure S4. Predicted Sequence Reads vs. Observed Sequence Reads

415 Predicted sequence reads are generated from the joint Bayesian model. The one-to-one
 416 line is plotted in red and Pearson correlation coefficient is 0.95. In general, predicted
 417 sequence reads track observed sequence reads, and show substantially less variance than
 418 observed sequence reads in Figure S3. However, unexpected zeros across multiple
 419 technical PCR replicates (stochastic dropouts), deviate notably from expected, low
 420 variance results (points along Y-axis). We note that all observed stochastic dropouts occur
 421 in less than 2.9% of sample read proportions and less than 9 morphological counts for a
 422 given sample. Such dropouts are likely a function of subsampling rare DNA molecules
 423 associated with molecular biology processing (See Supplement 2).



424

425

Figure S5. Co-detection of Taxa By Metabarcoding and Microscopy

426

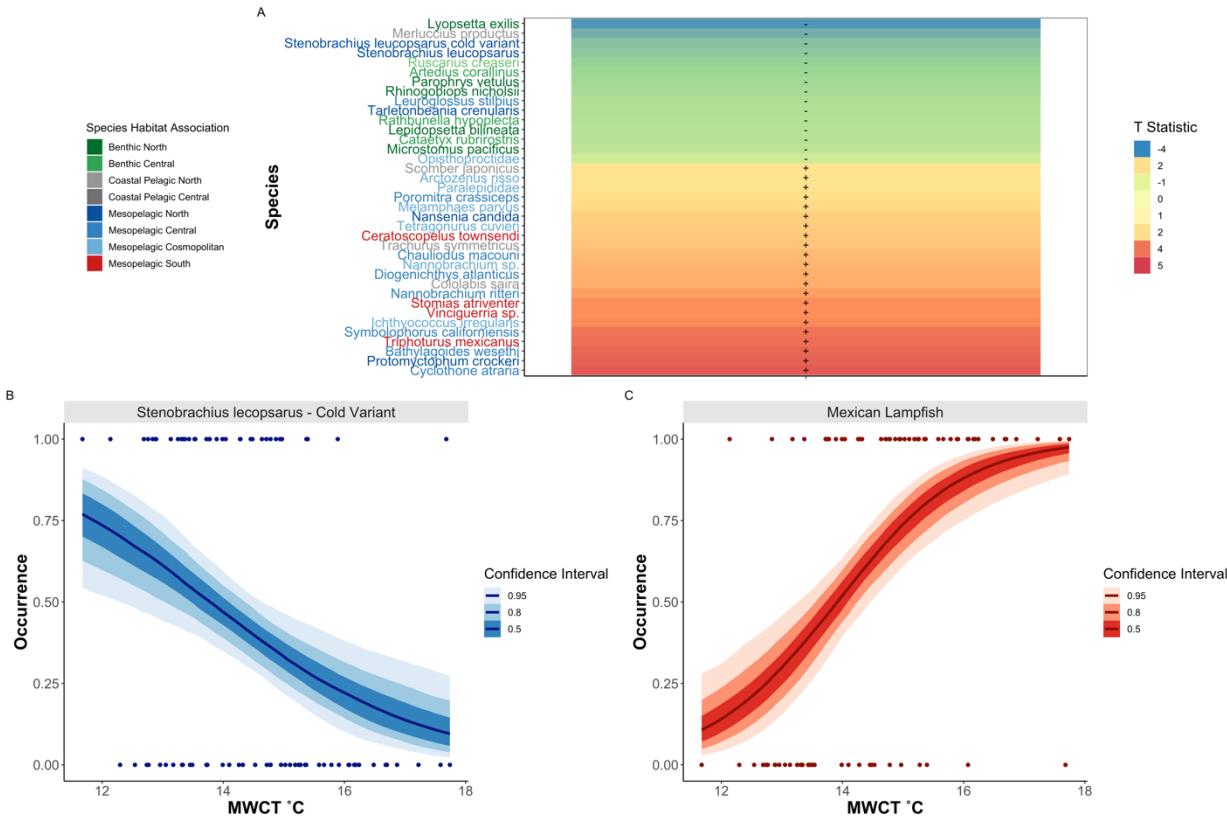
Of the 56 taxa used for modeling efforts (Supplemental Methods), both metabarcoding

427

and microscopy detected 46 taxa, with nine detected only by metabarcoding and one

428

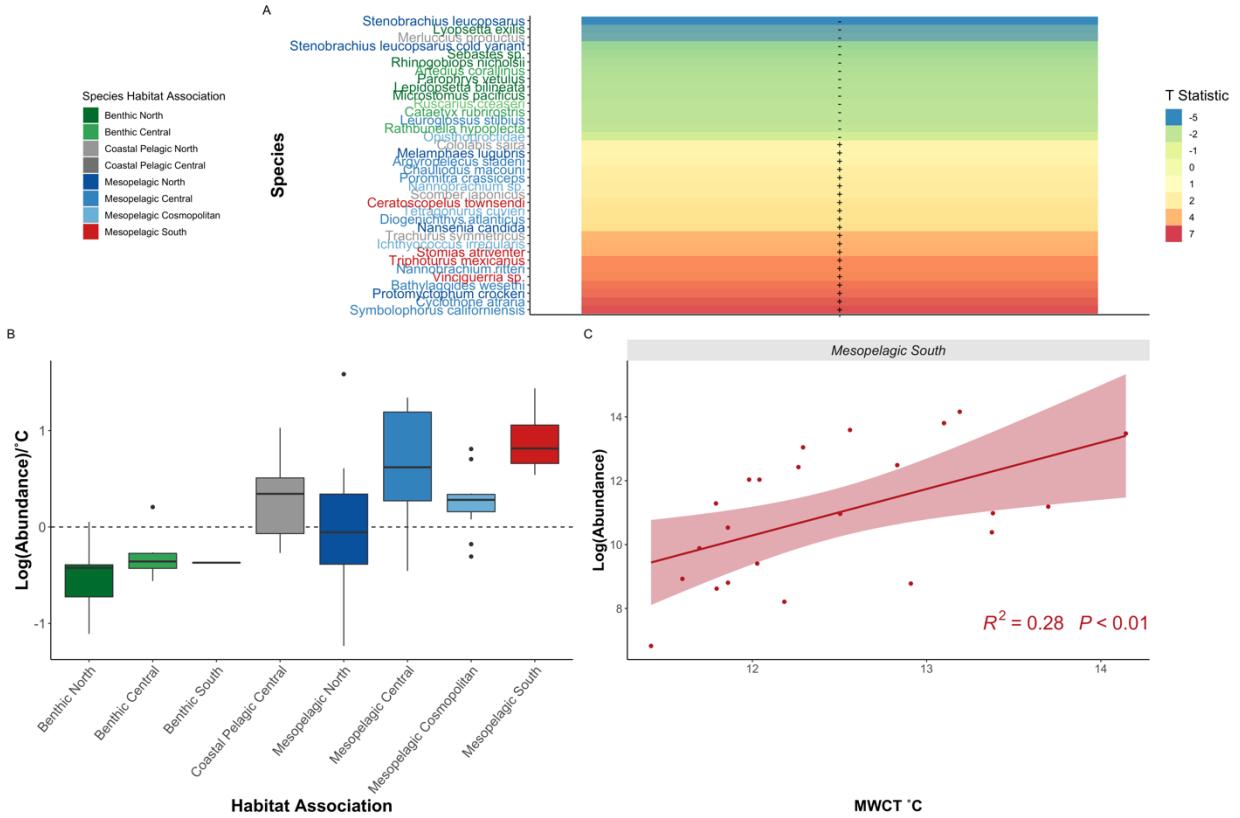
detected only by microscopy.



429

430 **Figure S6. Temperature Associations in Fish Species**

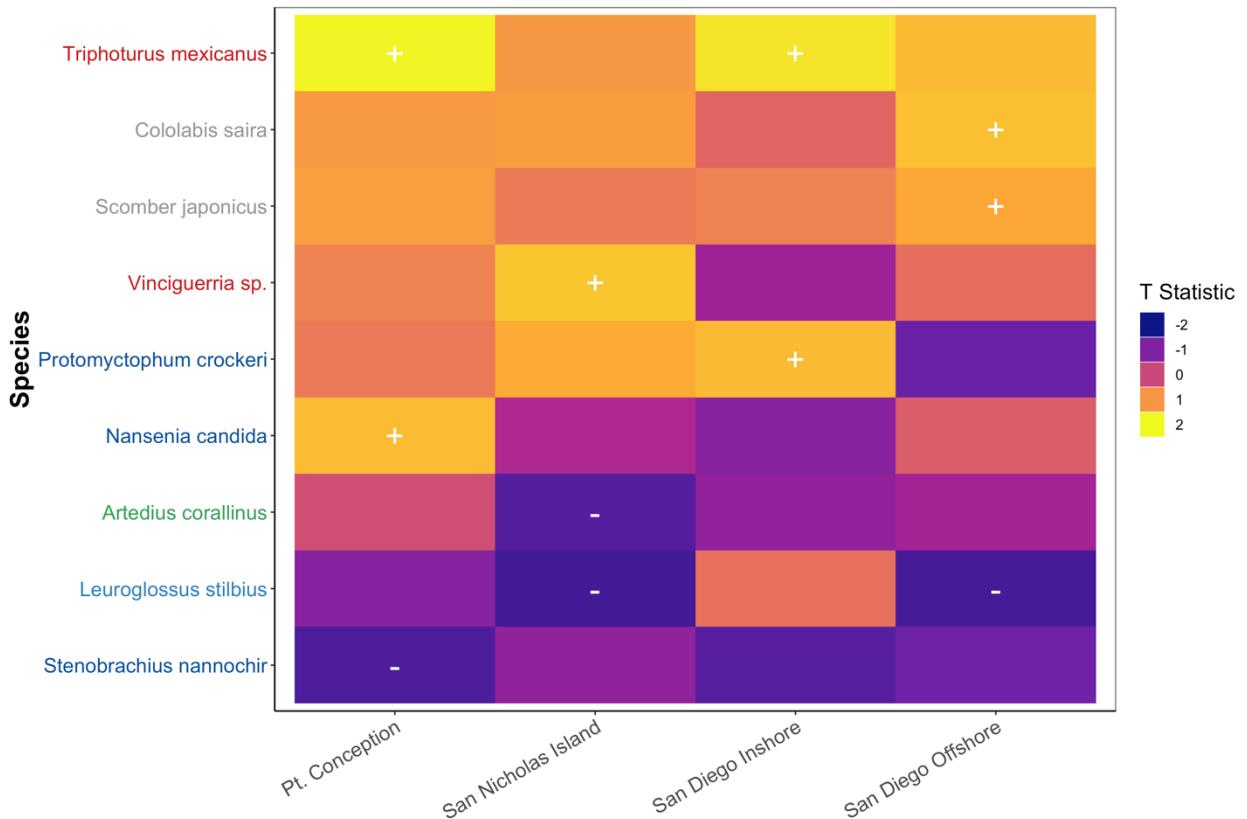
431 Changes in species occurrence patterns in response to MWCT, with Southern
 432 Mesopelagic species increasing in prevalence with elevated temperature (A). T statistic
 433 (slope coefficient/ standard error) from generalized binomial mixed model was calculated
 434 for each species across all stations. Only species with significantly different slopes (95%
 435 CI greater or less than zero) are plotted. Importantly, metabarcoding identified cold
 436 associated variants of the Northern Lanternfish (*Stenobrachius leucopsarus*) which
 437 cannot be morphologically identified (B) as well as warm-associated species like the
 438 Mexican Lampfish (*Triphoturus mexicanus*) (C).



439

440 **Figure S7. Southern Oceanic Species Drive Fish Community Shifts**

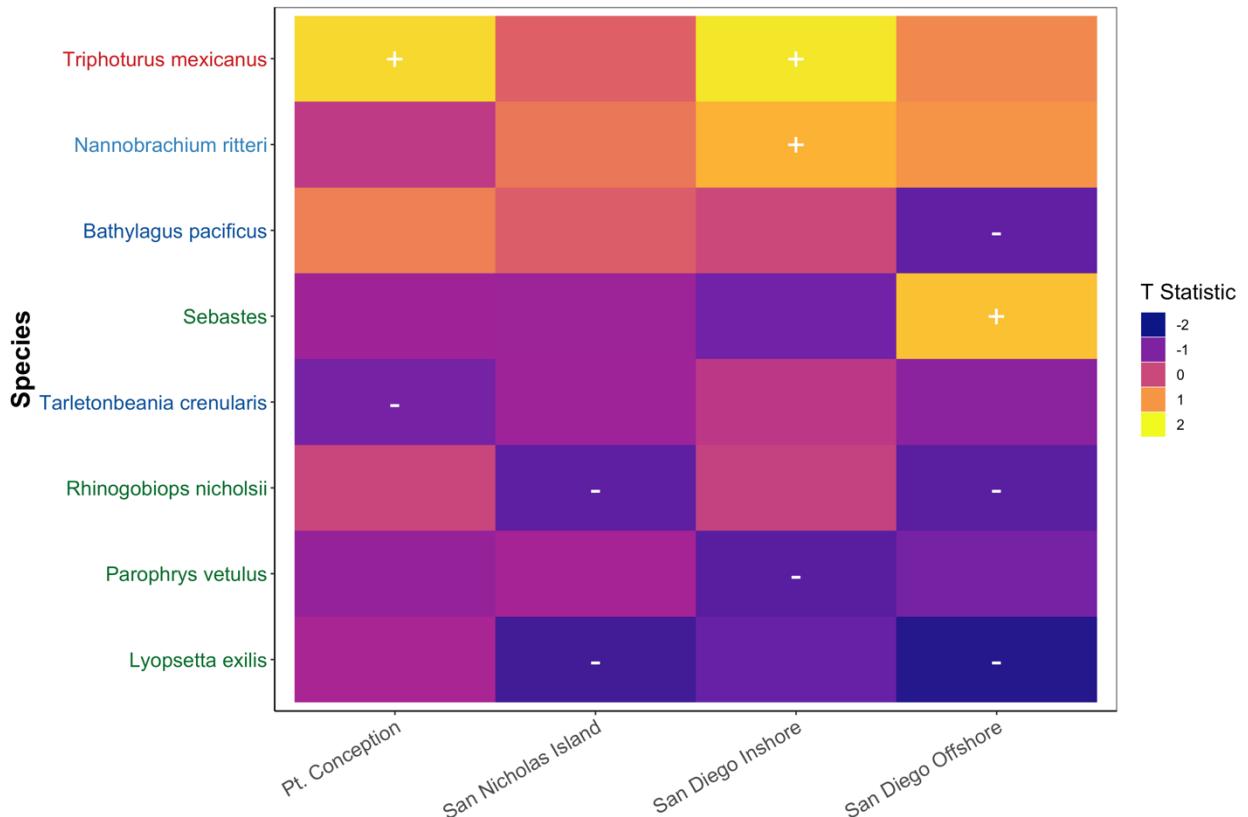
441 We capture changes in species biomass in response to MWCT, with Southern
 442 Mesopelagic species increasing in abundance with elevated temperature (A). T statistic
 443 (slope coefficient/standard error) from generalized linear models were calculated for each
 444 species across all stations. Only species with significantly different slopes (95% CI
 445 greater or less than zero) are plotted. Southern mesopelagic fishes were associated with
 446 increased temperature as indicated by the boxplots of all species-specific slopes from
 447 generalized linear models (B) and by the aggregated abundance relationship (C). In
 448 contrast, benthic species, as well as Northern Hake and Pacific Sardine abundances, were
 449 correlated with cooler temperatures.



450

451 **Figure S8. Significant Species Occurrence and SST Correlations at each Station**

452 Occurrence of the Mexican lampfish (*Triphoturus mexicanus*) was positively correlated
 453 with increased SST at the three northernmost stations. Generalized binomial mixed model
 454 of occurrence versus SST was calculated for each species at each station. Only species
 455 with significantly different slopes (95% CI greater or less than zero) are plotted. Colors
 456 correspond to T statistic (slope coefficient/ standard error). Colors of species names
 457 correspond to habitat associations described in Figure 2.



458

459

Figure S9. Significant Species Occurrence and MWCT Correlations at each Station

460

Occurrence of the Mexican lampfish (*Triphoturus mexicanus*) was positively correlated

461

with increased MWCT at the three northernmost stations. Generalized binomial mixed

462

model of occurrence versus MWCT was calculated for each species at each station. Only

463

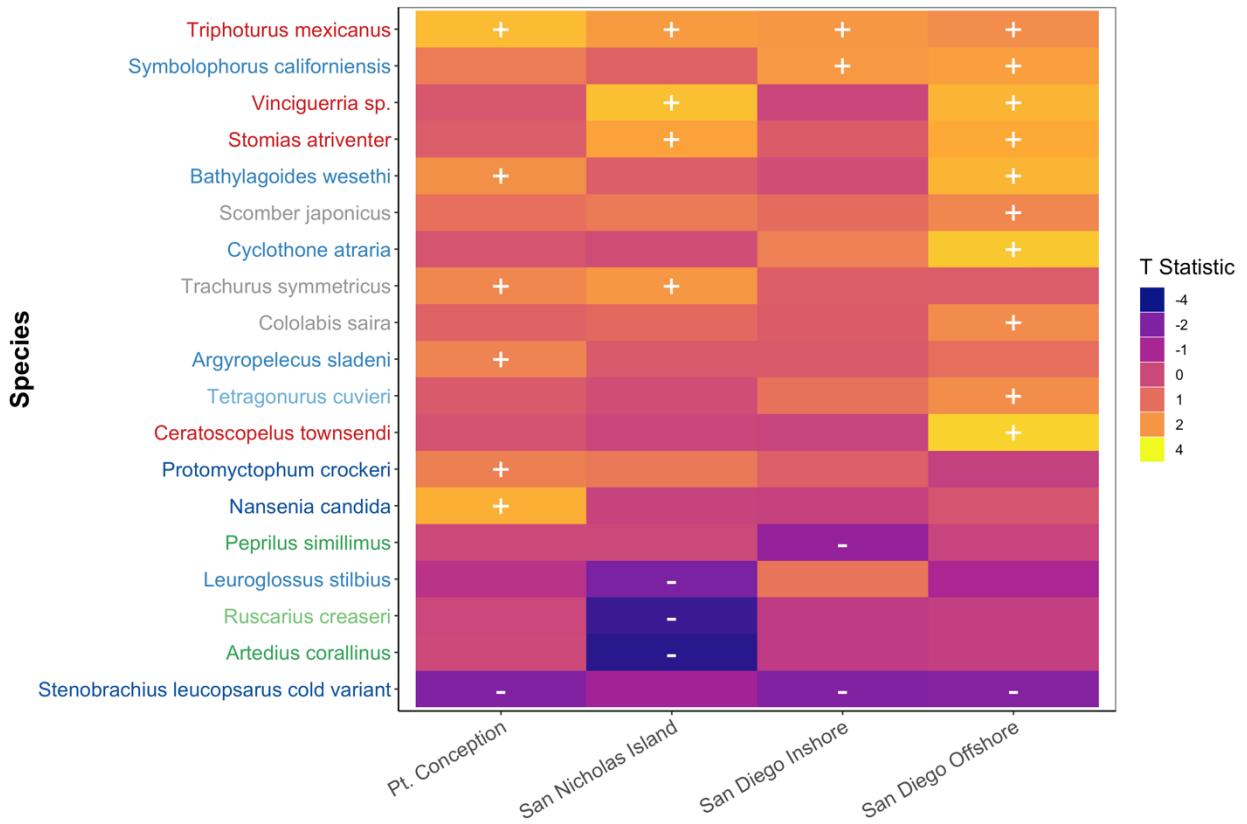
species with significantly different slopes (95% CI greater or less than zero) are plotted.

464

Colors correspond to T_statistic (slope coefficient/ standard error). Colors of species

465

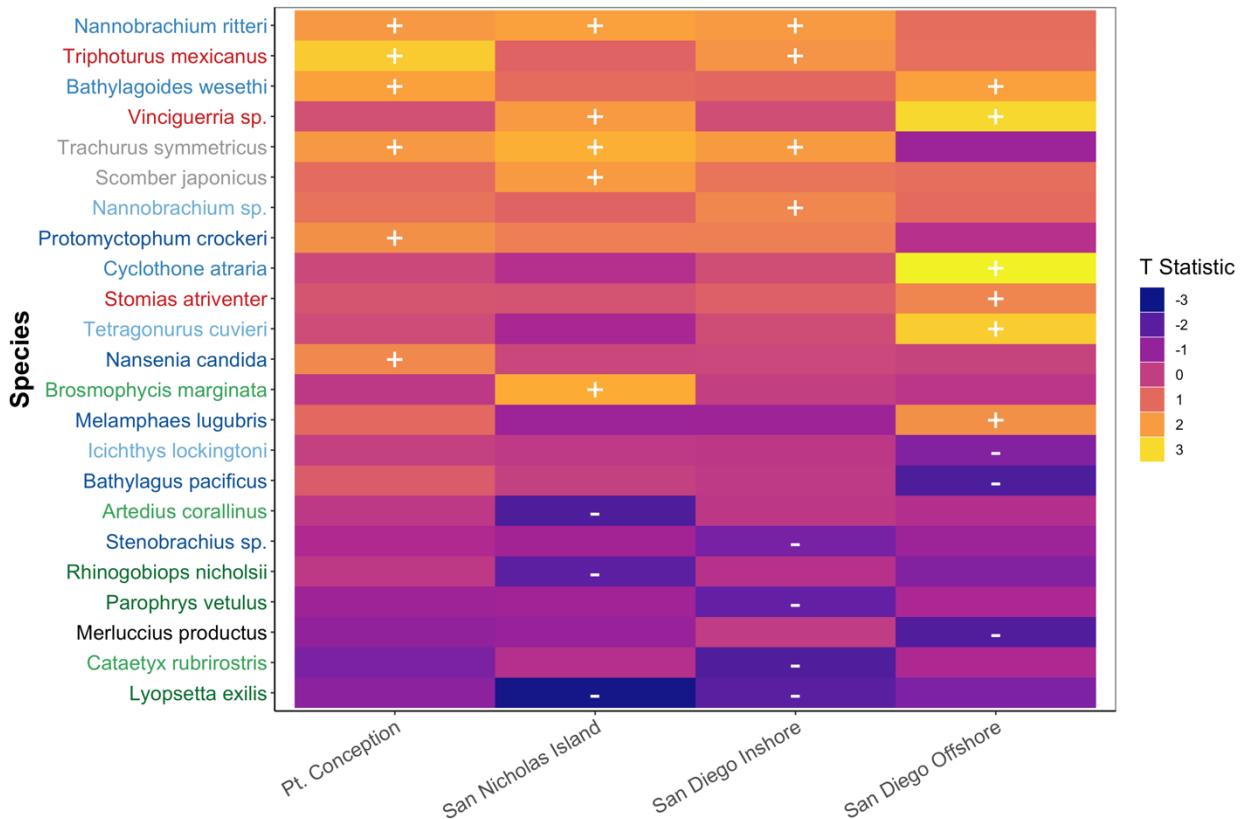
names correspond to habitat associations described in Figure 2.



466

467 **Figure S10. Significant Species Abundance and SST Correlations at each station**

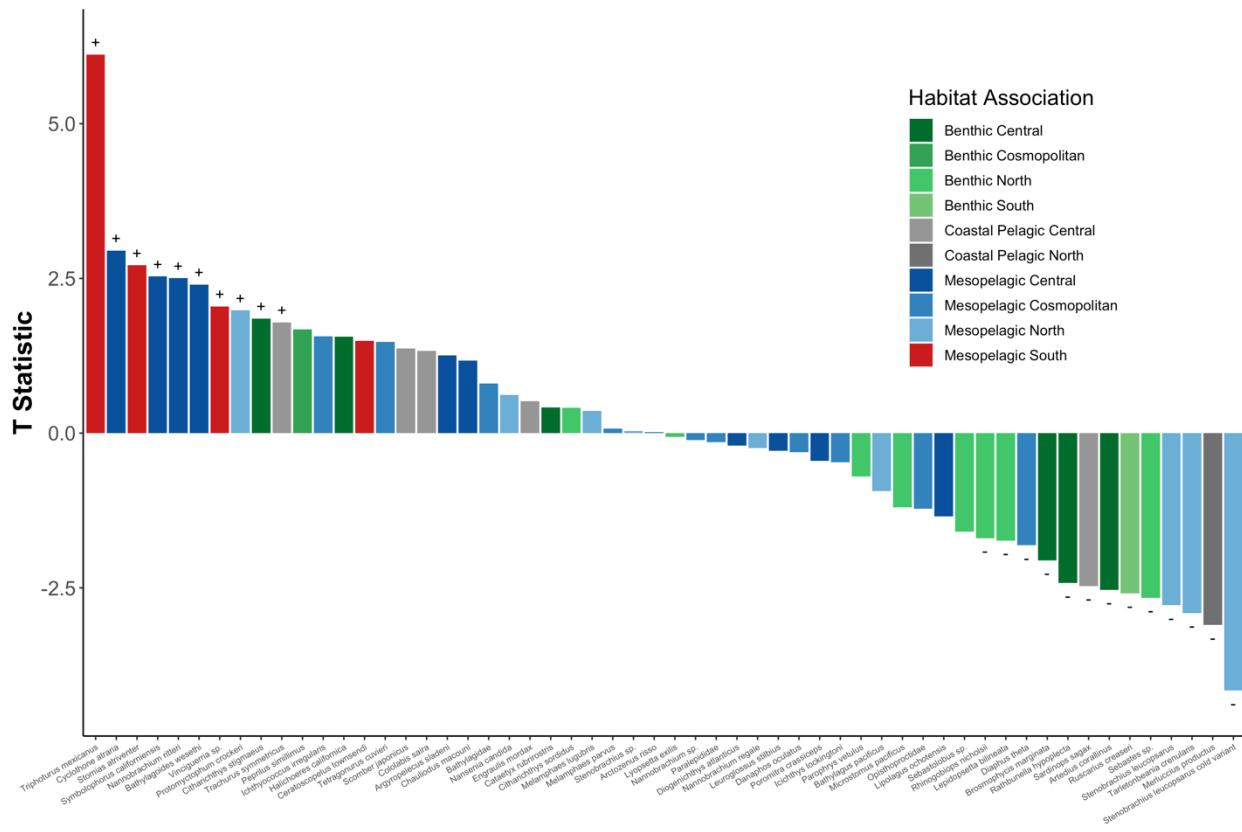
468 The abundance of the Mexican lampfish (*Triphoturus mexicanus*) was positively
 469 associated with increased SST at all stations. Likewise, the abundance of suite of
 470 mesopelagic species including *Vinciguerra* sp., *Symbolophorus californiensis*, and
 471 *Stomias atriventer* among others increased with warmer SST. Generalized linear mixed
 472 model of log (abundance) versus SST was calculated for each species at each station. Only
 473 species with significantly different slopes (95% CI greater or less than zero) are plotted.
 474 Colors correspond to T statistic (slope coefficient/ standard error). Colors of species
 475 names correspond to habitat associations described in Figure 2.



476

477 **Figure S11. Significant Species Abundance and MWCT Correlations at each station**

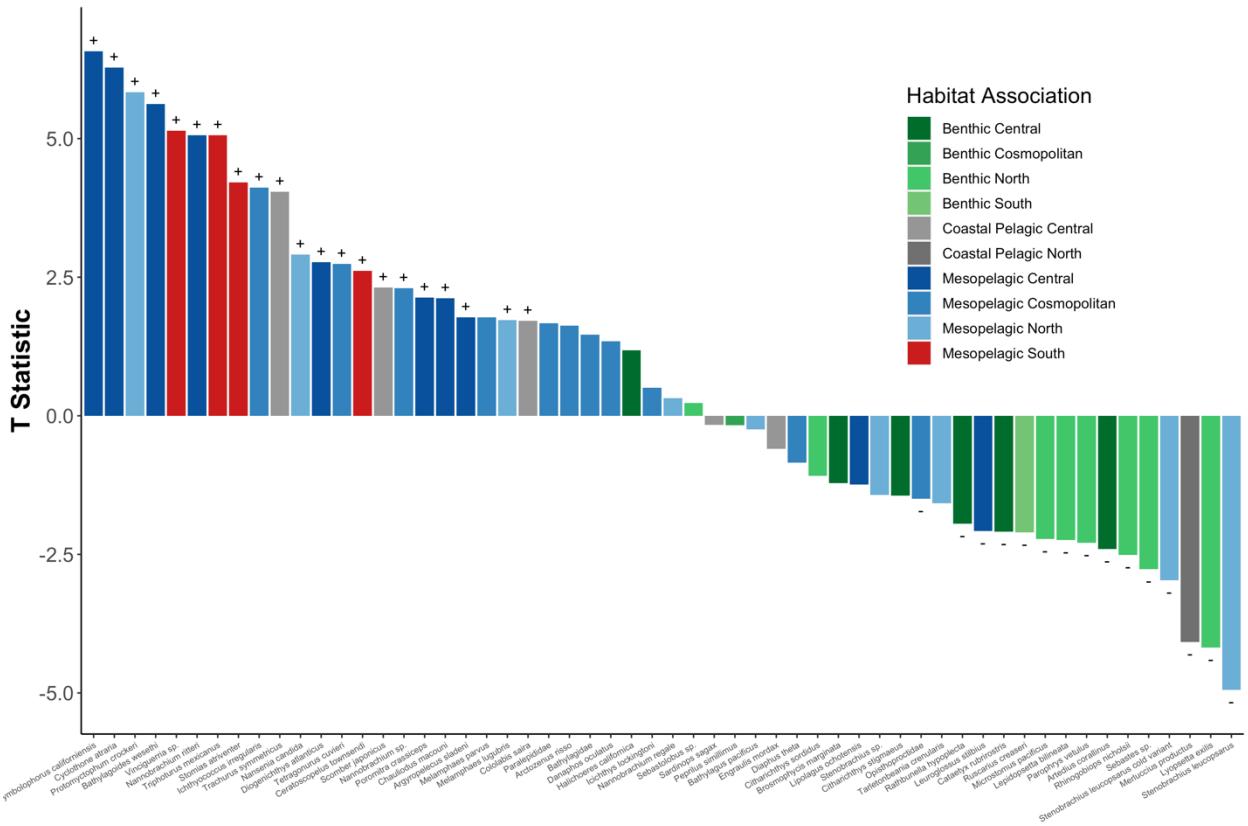
478 The abundance of the Mexican lampfish (*Triphoturus mexicanus*) was positively
 479 associated with increased MWCT at all stations. Likewise, the abundance of suite of
 480 mesopelagic species including *Vinciguerra* sp., *Symbolophorus californiensis*, and
 481 *Stomias atriventer* among others increased with warmer SST. Generalized linear mixed
 482 model of log (abundance) versus MWCT was calculated for each species at each station.
 483 Only species with significantly different slopes (95% CI greater or less than zero) are
 484 plotted. Colors correspond to T statistic (slope coefficient/ standard error). Colors of
 485 species names correspond to habitat associations described in Figure 2.



486

487 **Figure S12. Bar Plot of Significant Species Abundance and SST Correlations Across**
 488 **All stations**

489 The abundance of southern mesopelagic species increased with warmer temperatures
 490 while fisheries targets like Pacific Sardine (*Sardinops sagax*) and North Pacific Hake
 491 (*Merluccius productus*) decreased. Generalized linear mixed model of log (abundance)
 492 versus SST was calculated for each species across all stations. Only species with
 493 significantly different slopes (95% CI greater or less than zero) are plotted. Colors
 494 correspond to T statistic (slope coefficient/ standard error).



495

496

Figure S13. Bar Plot of Significant Species Abundance and MWCT Correlations

497

Across All stations

498

The abundance of southern mesopelagic species increased with warmer temperatures

499

while fisheries targets like Pacific Sardine (*Sardinops sagax*) and North Pacific Hake

500

(*Merluccius productus*) decreased. Generalized linear mixed model of log (abundance)

501

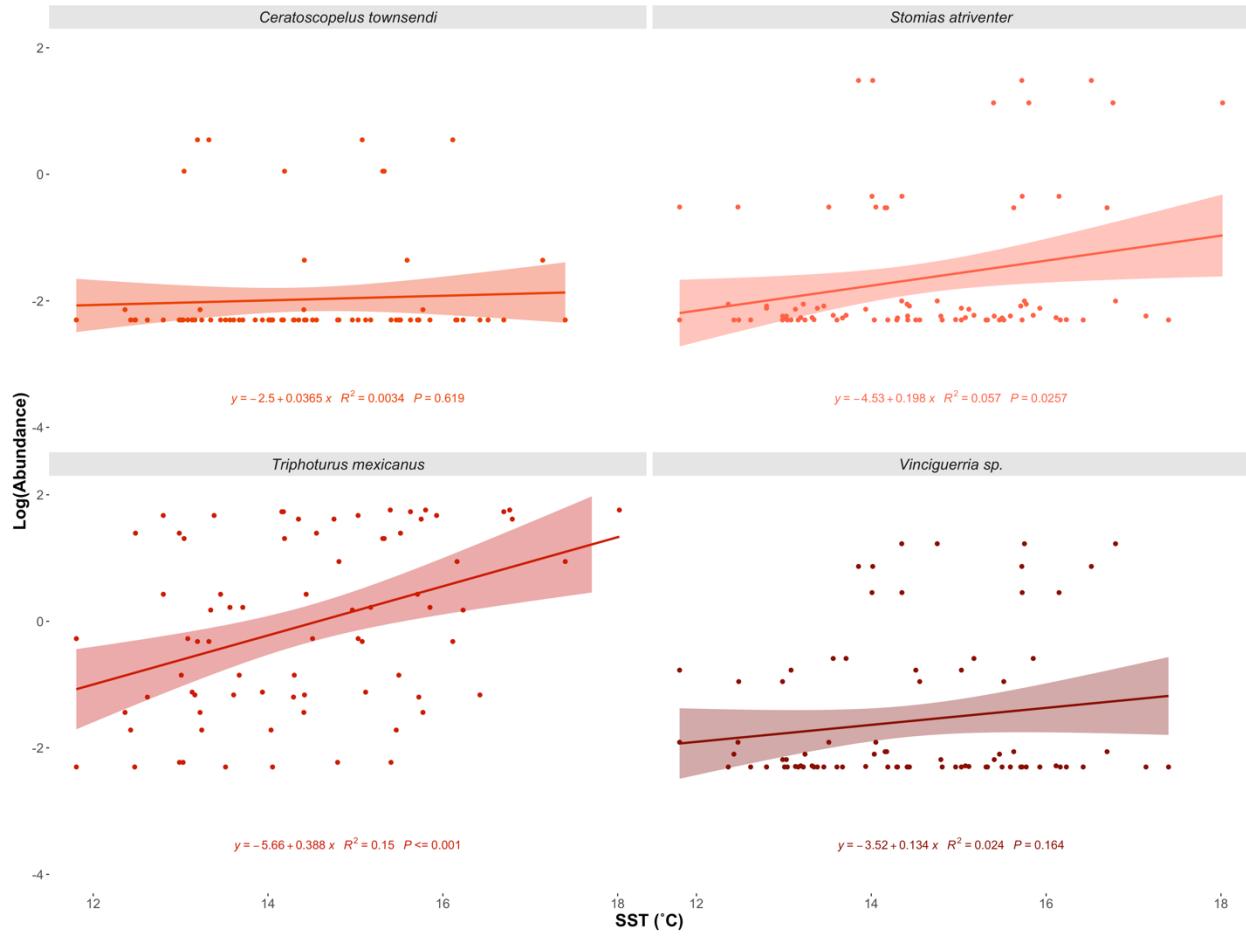
versus SST was calculated for each species across all stations. Only species with

502

significantly different slopes (95% CI greater or less than zero) are plotted. Colors

503

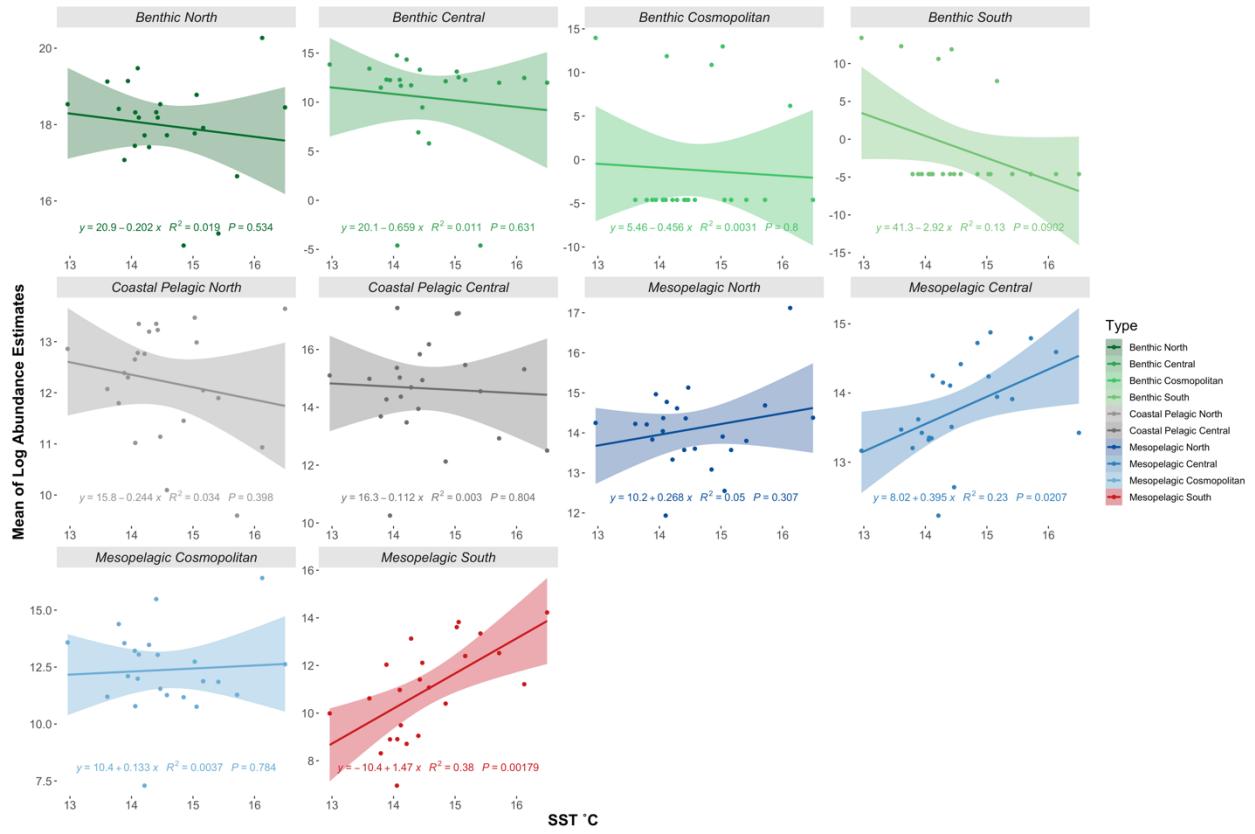
correspond to T statistic (slope coefficient/ standard error).



504

505 **Figure S14. Increased Abundance of Southern Mesopelagic Species with Higher
506 SST**

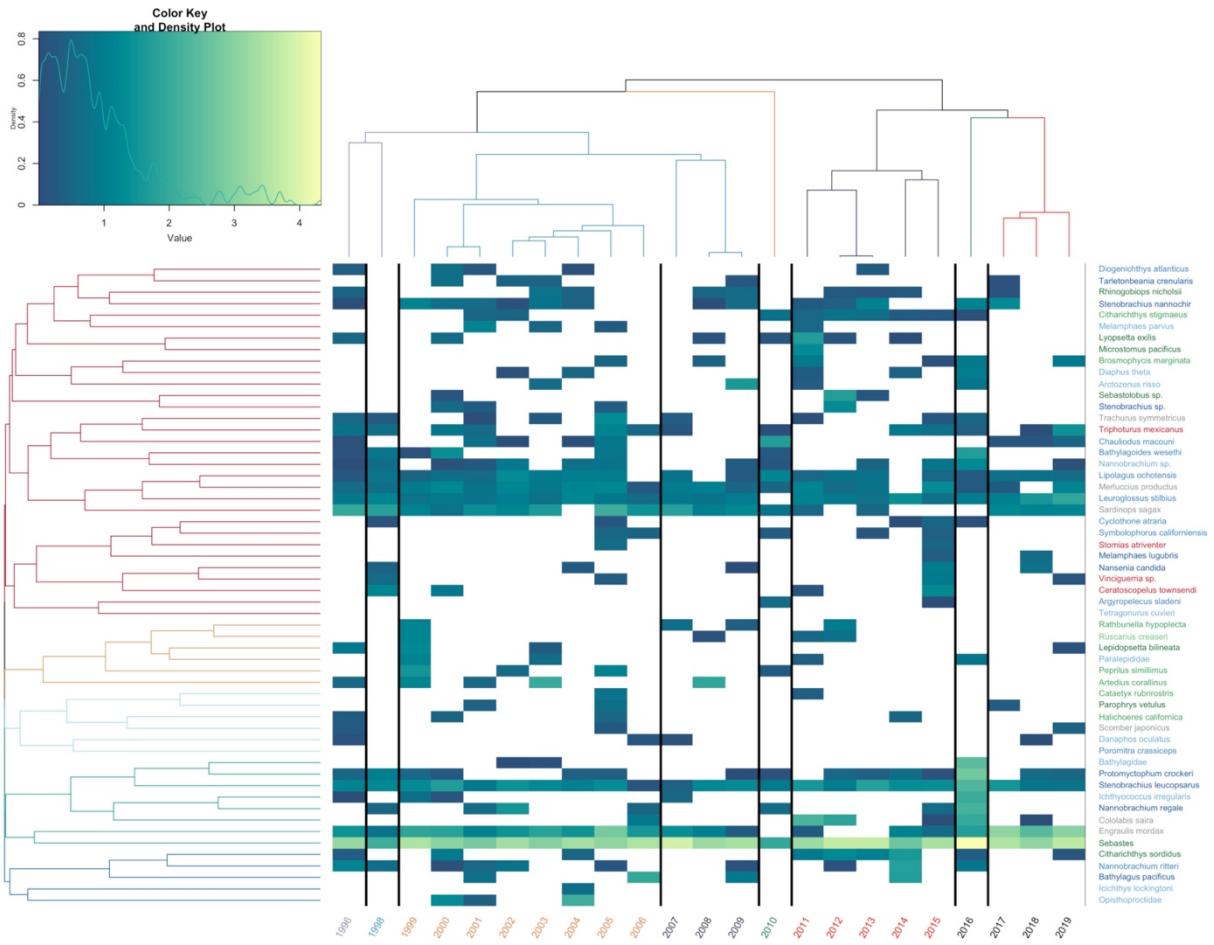
507 Species specific regressions of log (Abundance) vs. SST (°C). Three of the four southern
508 mesopelagic species had significant positive associations between abundance and SST.
509 Although not significant across all stations, the abundance of *Ceratoscopelus townsendi*
510 significantly increased with higher SST at the San Diego Offshore station (Figure S7).



511

512 **Figure S15. Increased Abundance of Southern and Central Mesopelagic Species with
513 Higher SST**

514 Abundances of each habitat association were summed at each station. Habitat association
515 regressions were fit using sum total log (Abundance) vs. SST (°C).

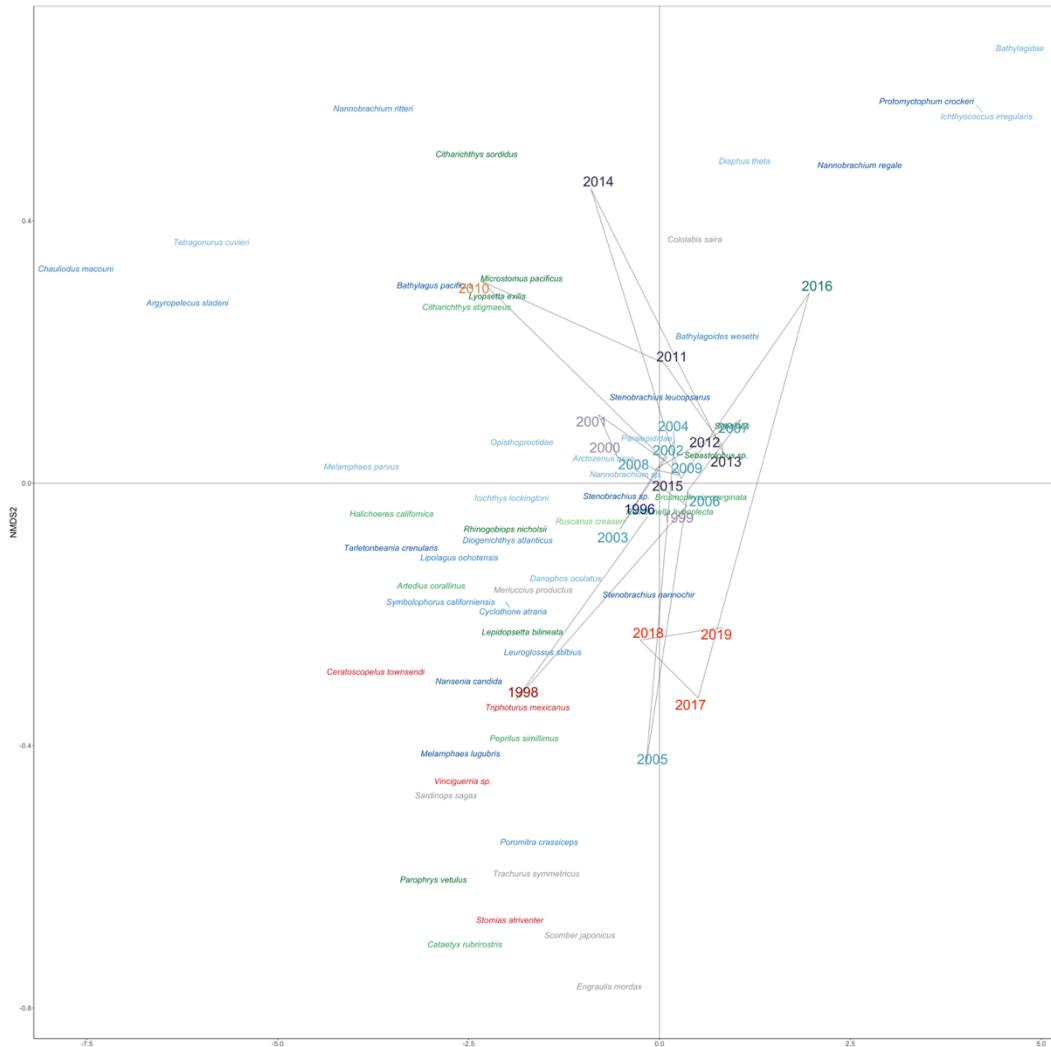


516

517

Figure S16. Heat Map of Abundances Over Time

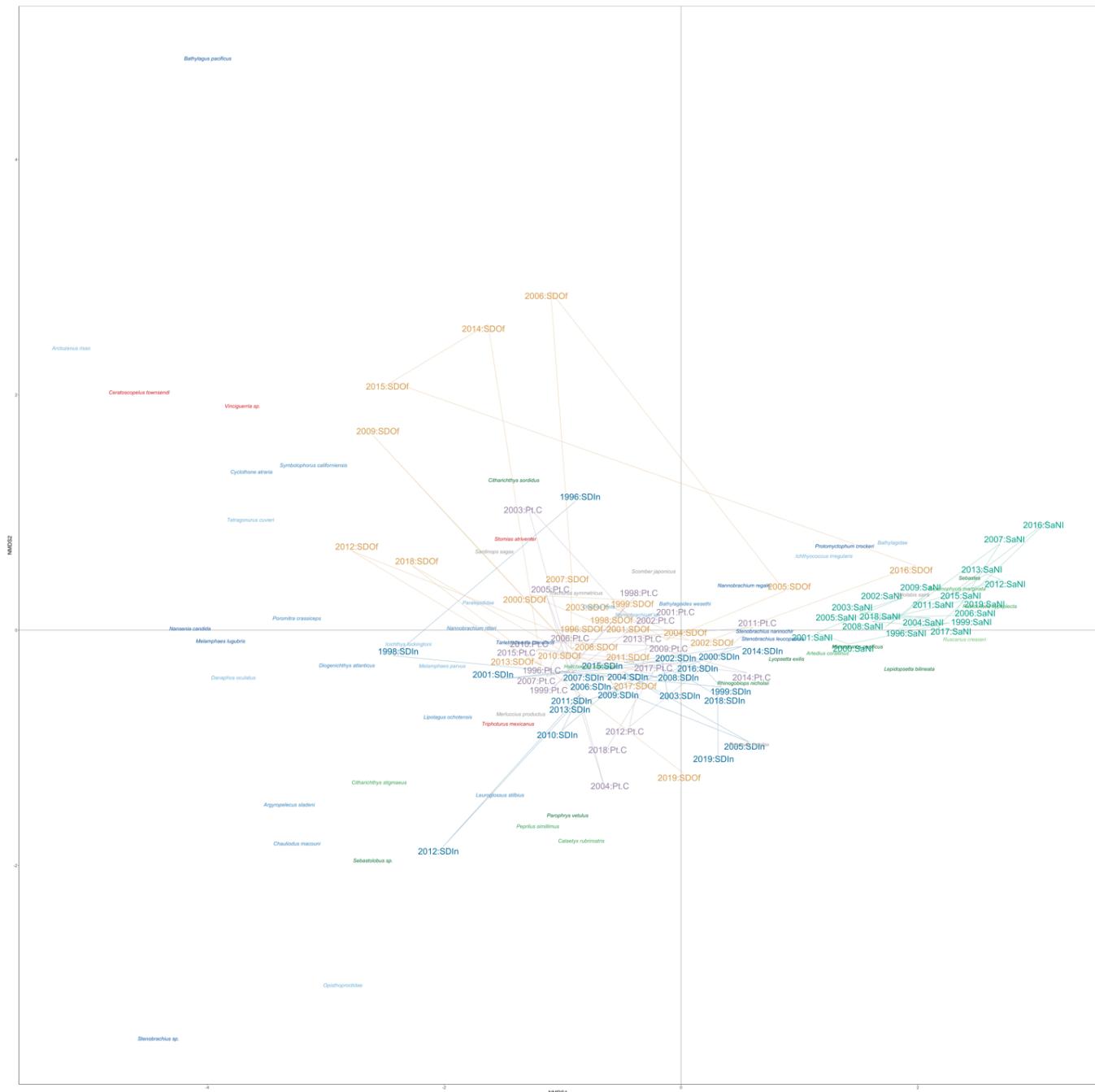
518 Northern Anchovy (*Engraulis mordax*) and rockfishes *Sebastodes* sp. dominated predicted
 519 counts. Estimated abundance of each year, averaged across stations, plotted over time.
 520 Years are color coded by chronological clustering. Species are grouped by hierarchical
 521 clustering. Lighter colors indicate higher abundance, white is a lack of detection. Species
 522 are color coded by habitat association matching Figure 1.



523

524 **Figure S17. NMDS Ordination of Species and Years**

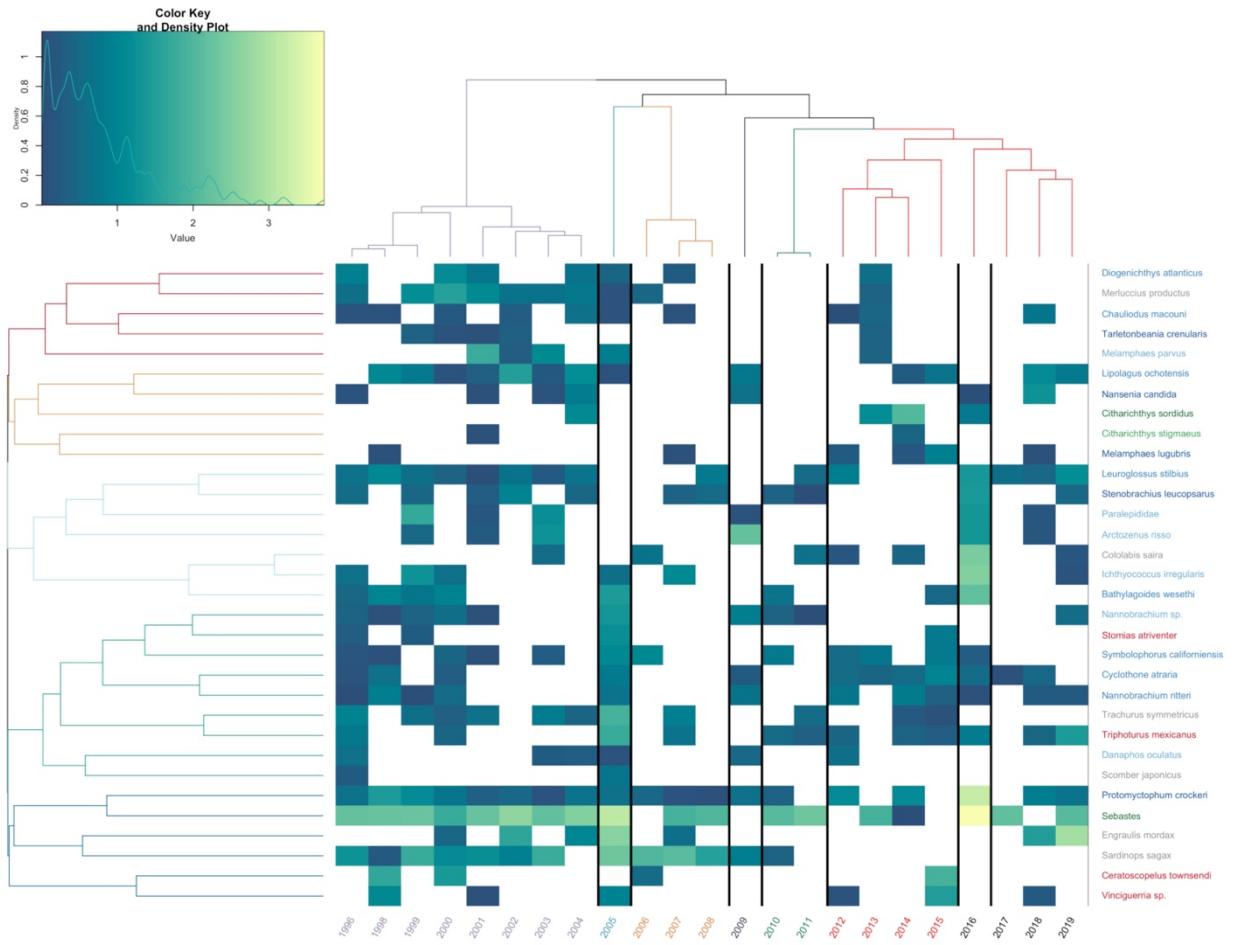
525 Fish assemblages changed across time with southern mesopelagics clustering with the
 526 1998 and 2005 El Niños as well as 2017-2019 after the peak of the marine heat wave.
 527 NMDS Ordination of Bray-Curtis dissimilarities calculated from summed abundance of
 528 each year averaged across stations. Marine heatwave patterns are obscured by differential
 529 onset and receding of the warming event across stations. Years are color coded by
 530 chronological clustering ($k=8$). Species are color coded by habitat association matching
 531 Figure 2.



532

533 **Figure S18. NMDS Ordination of Species and Samples**

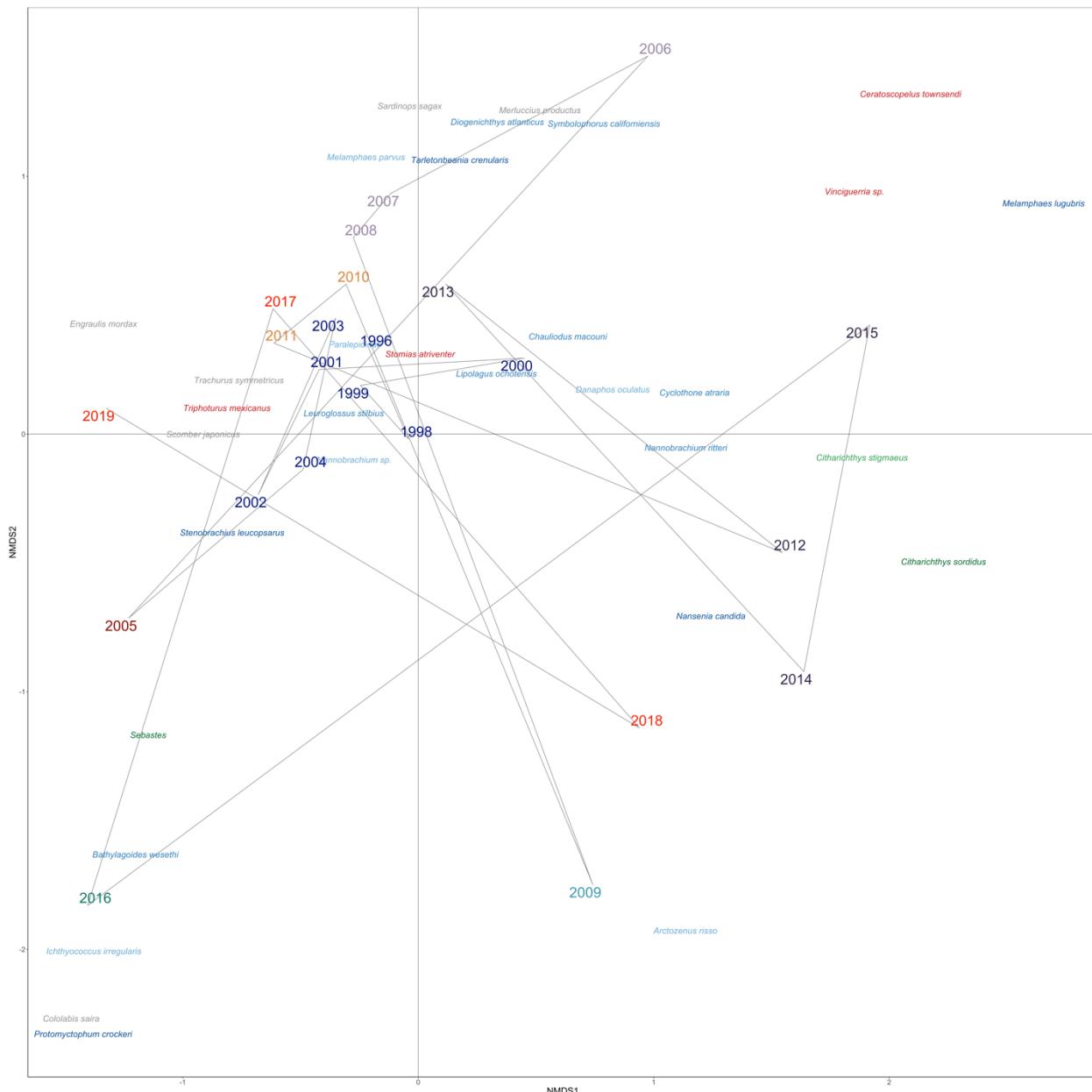
534 Fish assemblages were strongly structured by stations, particularly the San Nicholas Island
 535 station which had the highest abundance of *Sebastes* sp. NMDS Ordination of Bray-Curtis
 536 dissimilarities calculated from the abundance of each year and station. Samples are color
 537 coded by station. Species are color coded by habitat association matching Figure 2.



538

539 **Figure S19. Heat Map of San Diego Offshore Abundances Over Time**

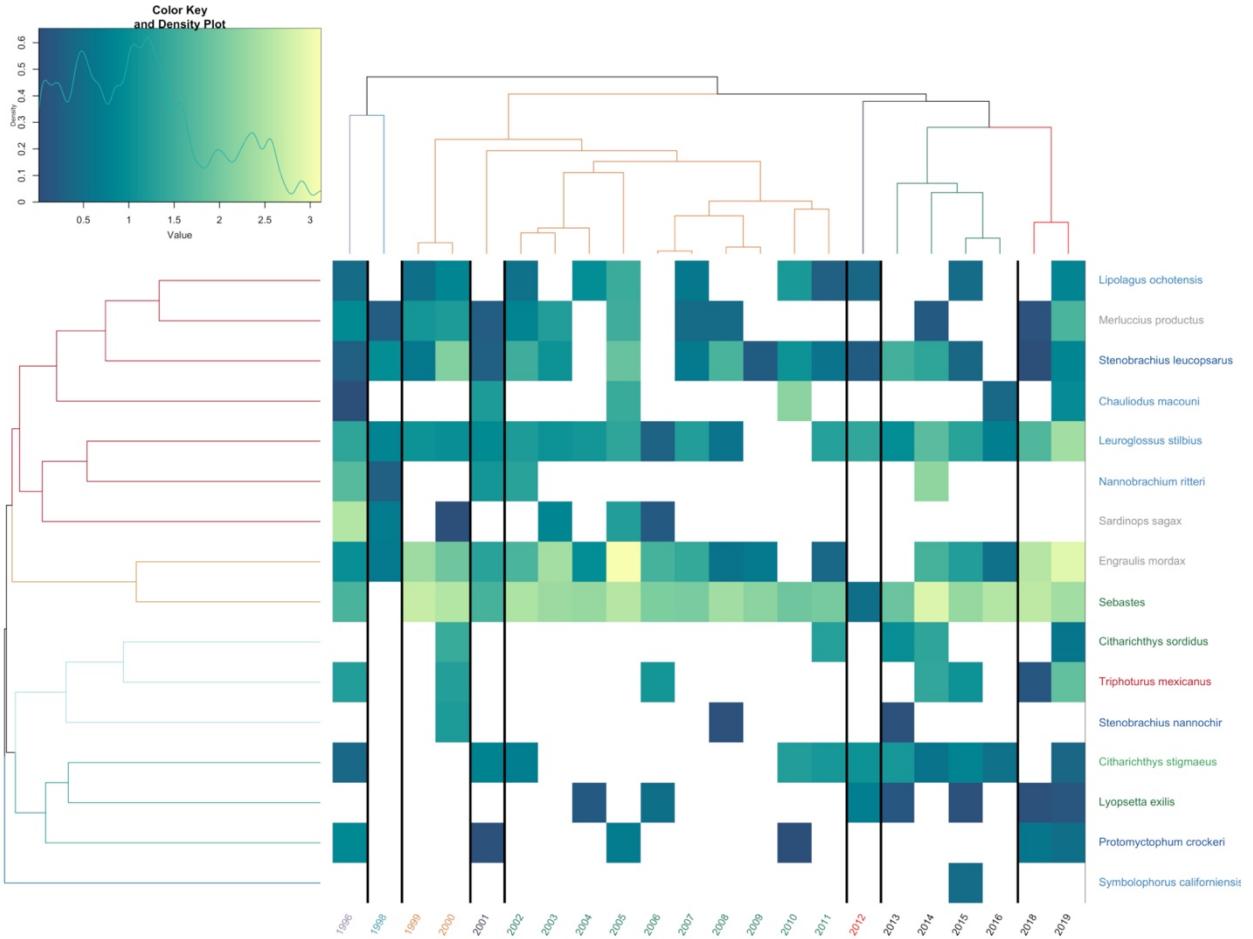
540 Estimated abundance of each year plotted over time. Years are color coded by
 541 chronological clustering. Species are grouped by hierarchical clustering. Lighter colors
 542 indicate higher abundance, white is a lack of detection. Species are color coded by habitat
 543 association matching Figure 1.



544

545 **Figure S20. NMDS Ordination of San Diego Offshore Species and Years**

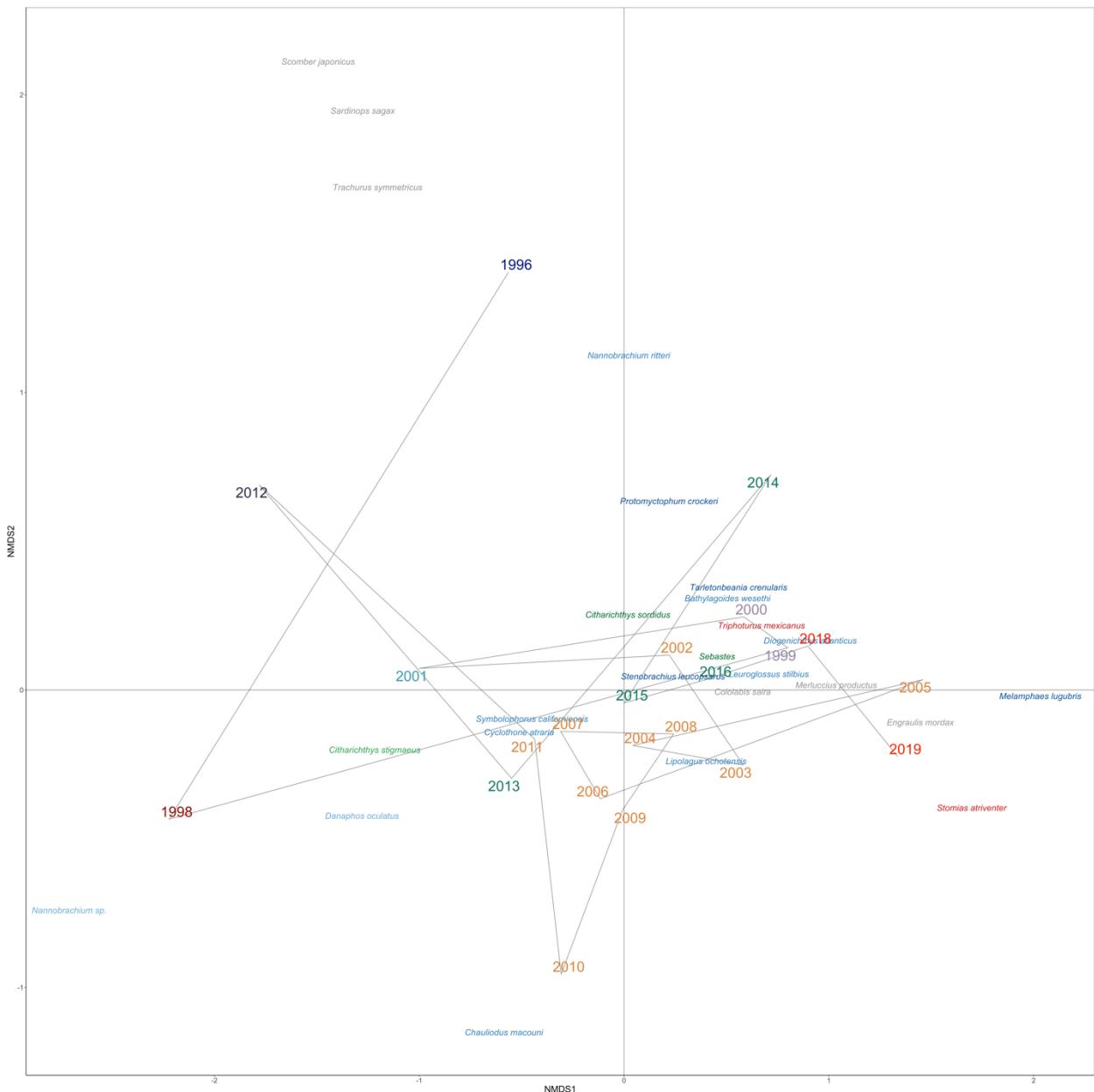
546 NMDS Ordination of Bray-Curtis dissimilarities calculated from abundance of each year.
 547 Years are color coded by chronological clustering (k =8). Species are color coded by
 548 habitat association matching Figure 2.



549

550 **Figure S21. Heat Map of San Diego Inshore Abundances Over Time**

551 Estimated abundance of each year plotted over time. Years are color coded by
 552 chronological clustering. Species are grouped by hierarchical clustering. Lighter colors
 553 indicate higher abundance, white is a lack of detection. Species are color coded by habitat
 554 association matching Figure 1.

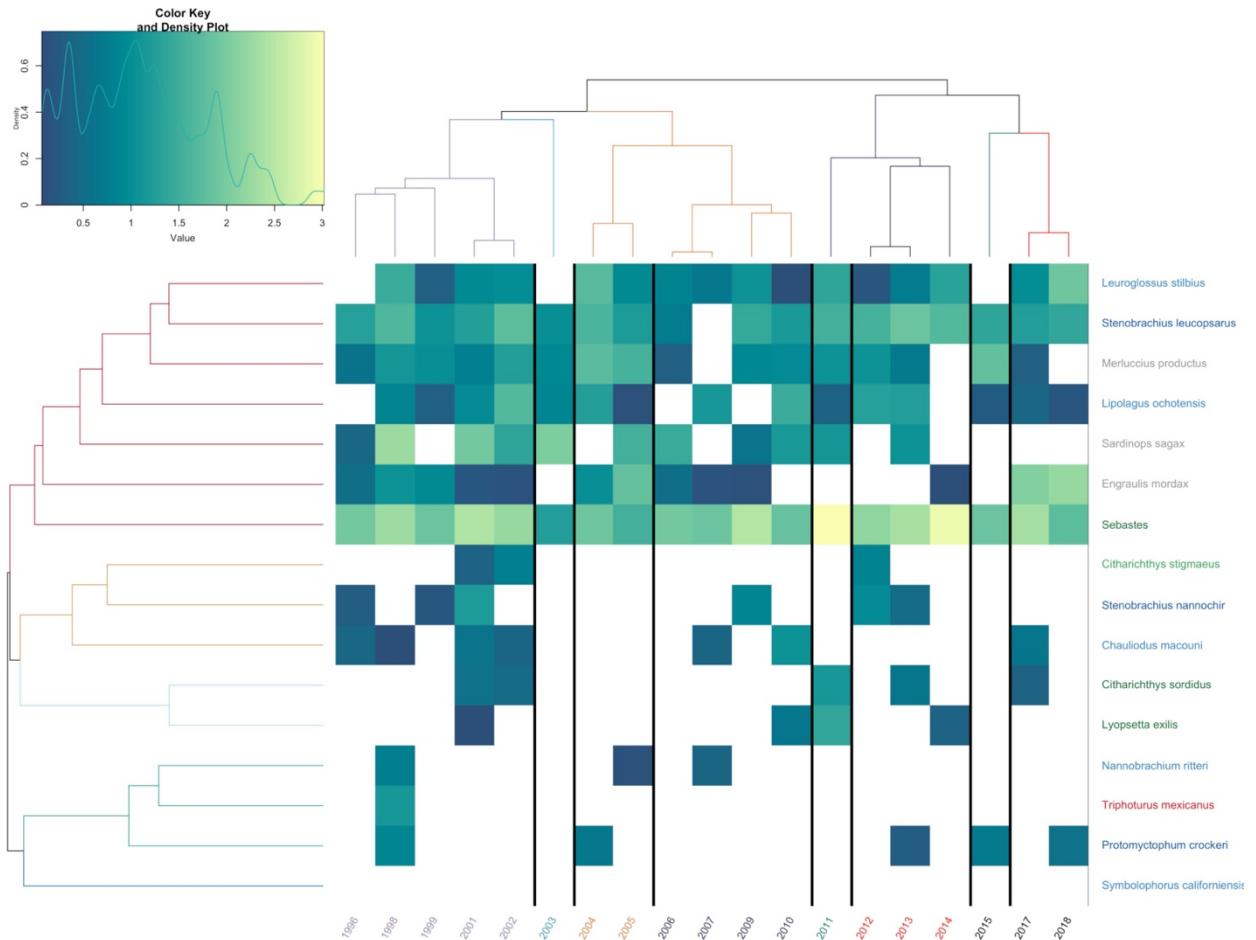


555

556 **Figure S22. NMDS Ordination of San Diego Inshore Species and Years**

557 NMDS Ordination of Bray-Curtis dissimilarities calculated from abundance of each year.

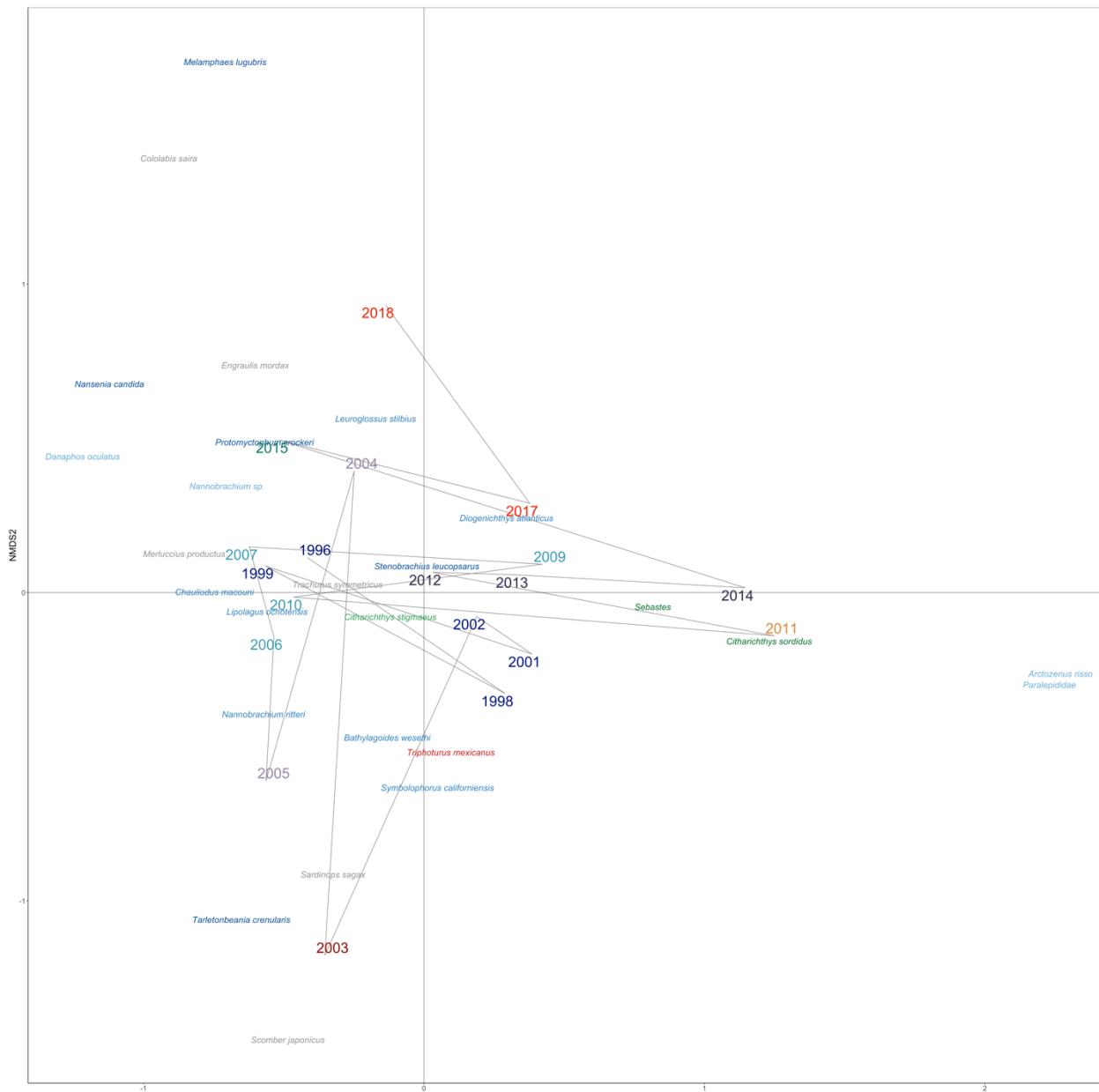
558 Years are color coded by chronological clustering (k =8). Species are color coded by
559 habitat association matching Figure 2.



560

561 **Figure S23. Heat Map of Pt. Conception Abundances Over Time**

562 Estimated abundance of each year plotted over time. Years are color coded by
 563 chronological clustering. Species are grouped by hierarchical clustering. Lighter colors
 564 indicate higher abundance, white is a lack of detection. Species are color coded by habitat
 565 association matching Figure 1.

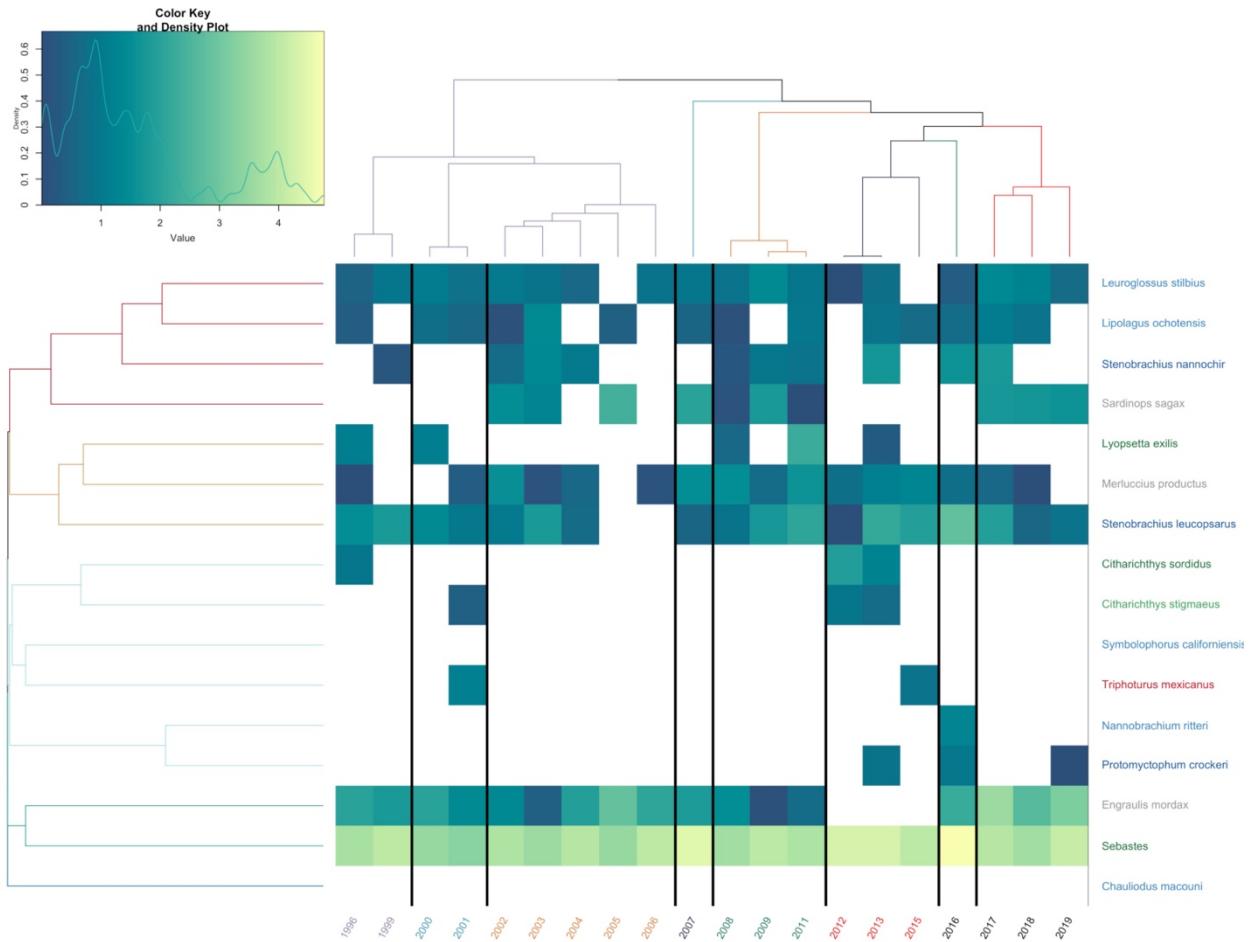


566

567 **Figure S24. NMDS Ordination of Pt. Conception Species and Years**

568 NMDS Ordination of Bray-Curtis dissimilarities calculated from abundance of each year.

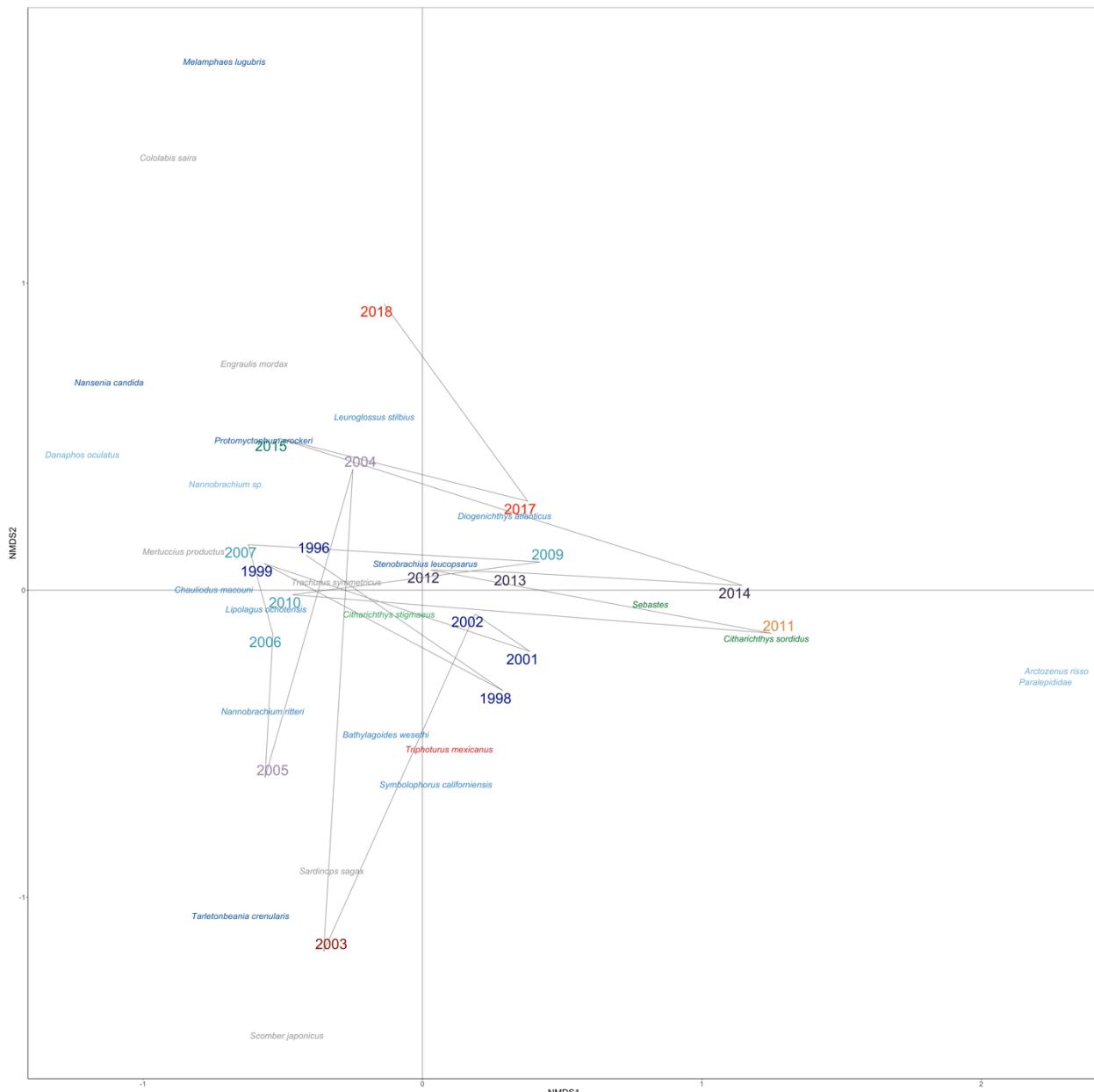
569 Years are color coded by chronological clustering ($k = 8$). Species are color coded by
570 habitat association matching Figure 2.



571

572 **Figure S25. Heat Map of San Nicholas Island Abundances Over Time**

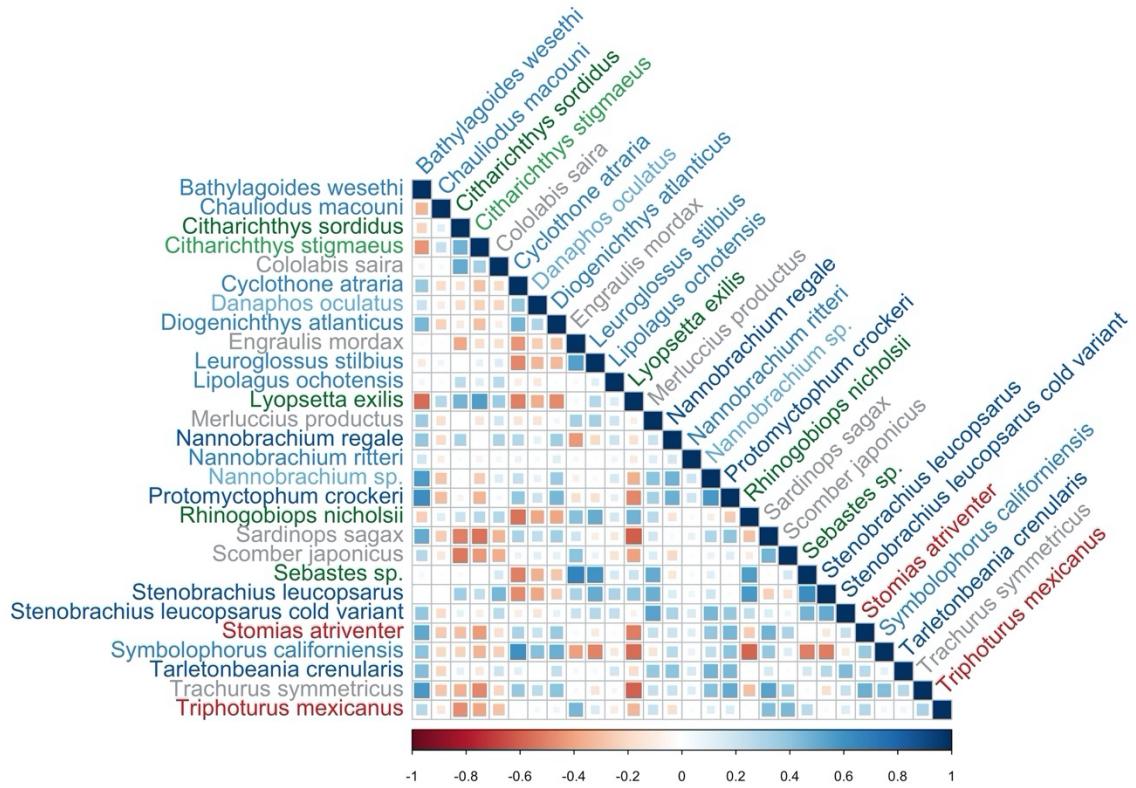
573 Estimated abundance of each year plotted over time. Years are color coded by
 574 chronological clustering. Species are grouped by hierarchical clustering. Lighter colors
 575 indicate higher abundance, white is a lack of detection. Species are color coded by habitat
 576 association matching Figure 1.



577

578 **Figure S26. NMDS Ordination of San Nicholas Island Species and Years**

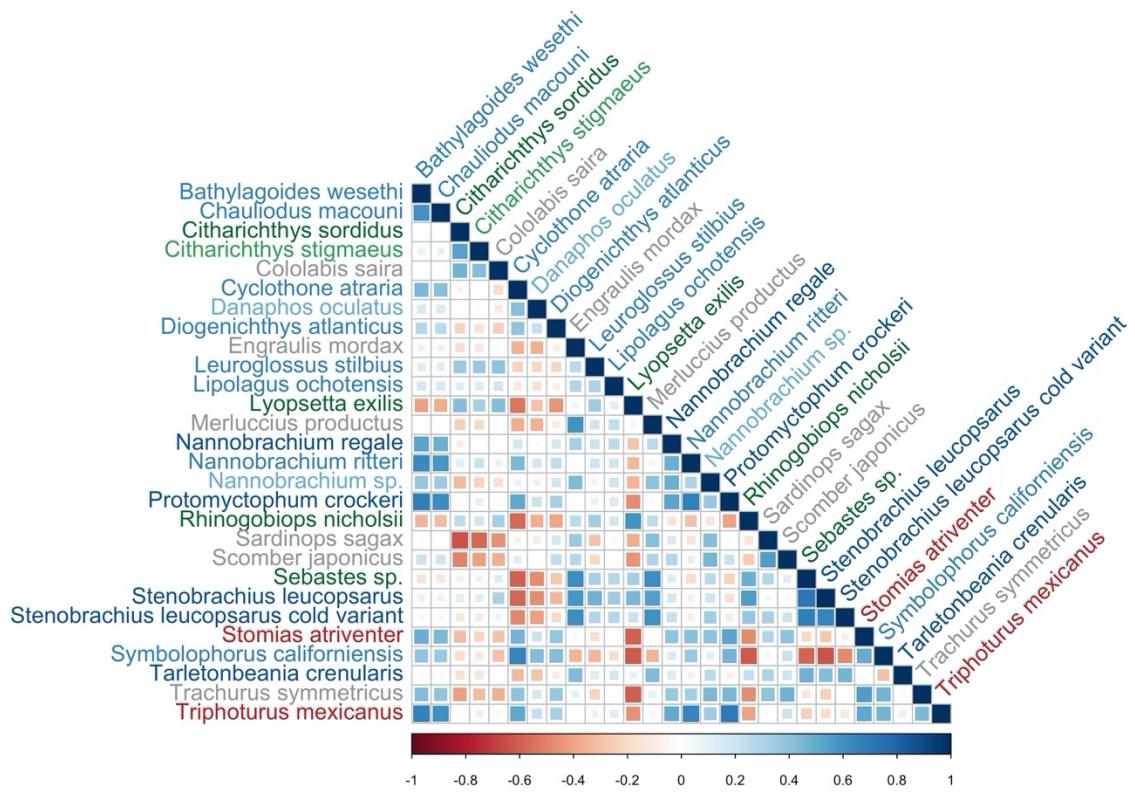
579 NMDS Ordination of Bray-Curtis dissimilarities calculated from abundance of each year.
 580 Years are color coded by chronological clustering ($k = 8$). Species are color coded by
 581 habitat association matching Figure 2.



582

583 **Figure S27. Co-occurrence Patterns of Species Controlling for SST**

584 Benthic fisheries targets like sanddabs were negatively correlated with mesopelagic
 585 species when controlling for SST. These results capture a strong benthic vs. pelagic
 586 community dynamics when controlling for SST. We plot the correlation matrix of the
 587 generalized linear latent variable model predictors across species. Species in blue correlate
 588 are positively correlated with each other while species in red are negatively correlated.
 589 Strength of association is visualized by size of square. Species are color coded by habitat
 590 association matching Figure 2.



591

592

Figure S28. Co-occurrence Patterns of Species

593

Fisheries targets like North Pacific Hake were negatively correlated with mesopelagic

594

species. We plot the correlation matrix of the generalized linear latent variable model

595

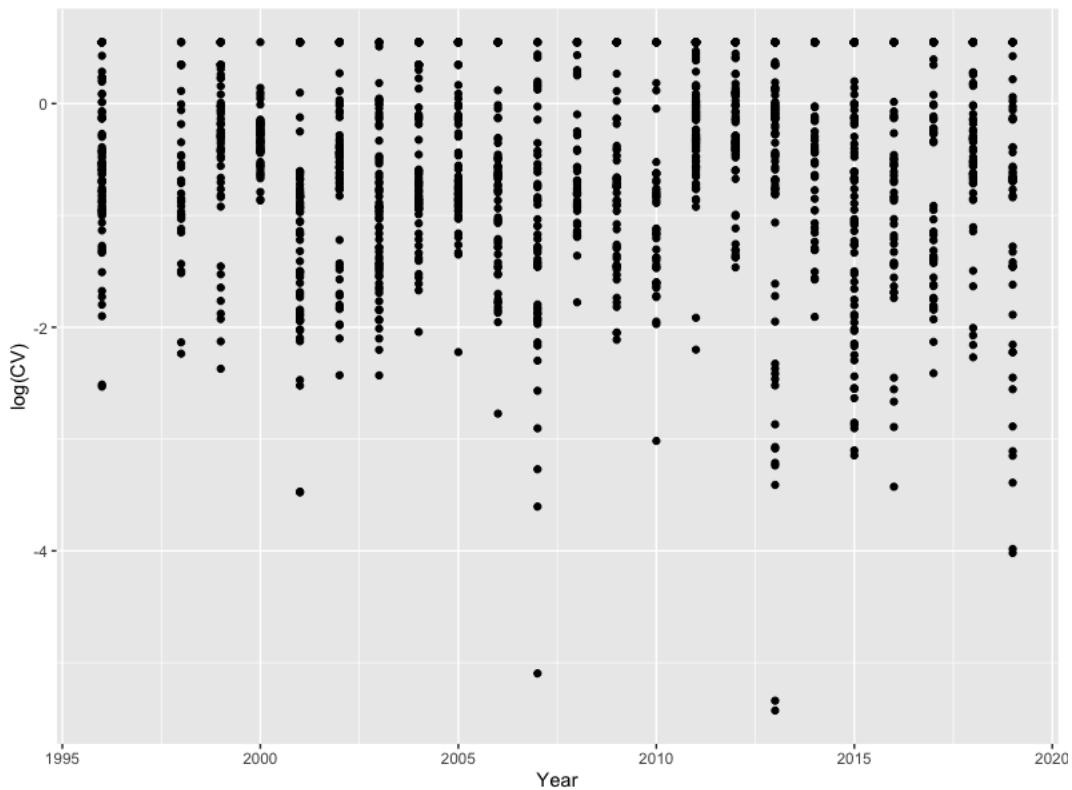
predictors across species. Species in blue correlate are positively correlated with each

596

other while species in red are negatively correlated. Strength of association is visualized

597

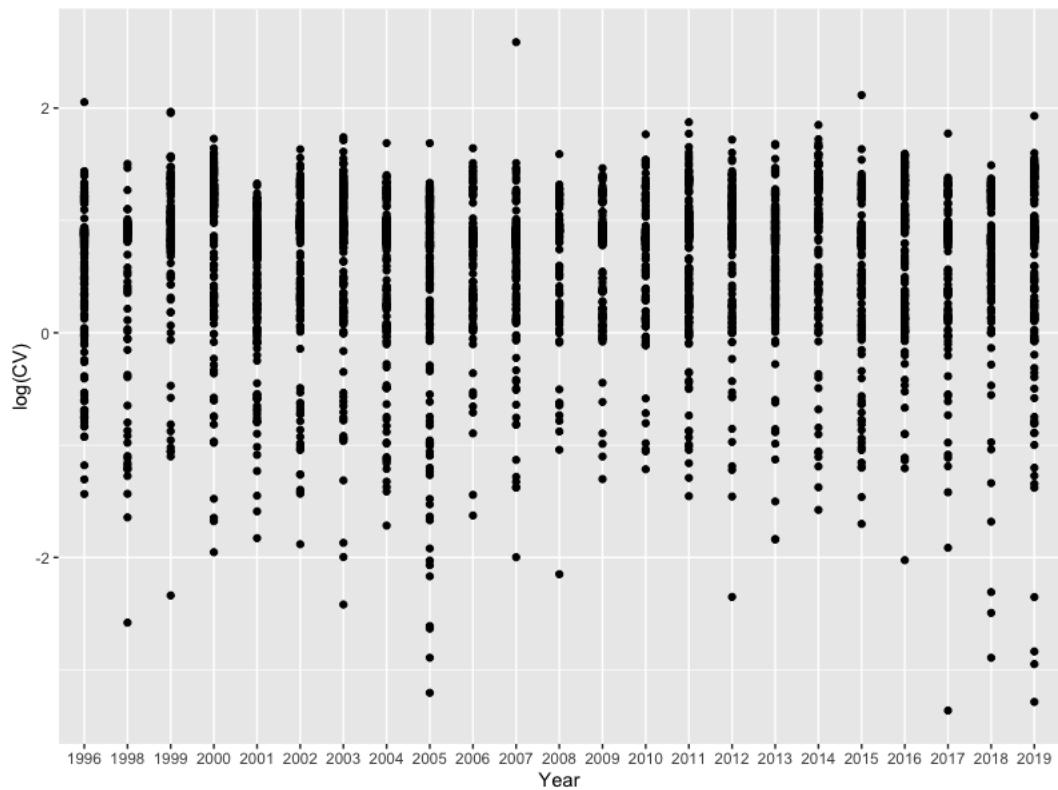
by size of square. Species are color coded by habitat association matching Figure 2.



598

599 **Figure S29. Stable Precision of Amplicon Abundance Over Time**

600 Here we measure the coefficient of variation (CV) of species-specific amplicons across
601 three technical replicates. An increase in CV with the age of the sample would signal
602 degradation; we see no such trend.



603

604

Figure S30. Stable Precision of Abundance Estimates Over Time

605

606

Coefficient of variation of model estimates of abundance over time. We observe no evidence of change in precision over time.

Supplement S2

Gold et al. 2021

Model description of morphology and metabarcoding analysis of CALCOFI samples

Motivating models.

We use metabarcoding approaches outlined in Supplement S1 to produce amplicons from ethanol-preserved ichthyoplankton samples. Briefly, we generated amplicons using the MiFish 12S Universal Teleost primer set (23) on DNA extractions derived using the Qiagen DNeasy Blood and Tissue kit filtered ethanol-preservative. Each amplicon library was sequenced separately on an Illumina NextSeq. See Supplement S1 for full details.

We estimate that, for any species i , the number of sequenced amplicons is proportional to the fraction of DNA from that species in the PCR template (53,55,56). We note that here we use i to represent species, but can be more generalized to represent “molecular targets” or “amplicon sequence variants.” The amplicons produced during a PCR reaction are governed by the efficiency parameter α_i , which is characteristic of the interaction between the particular primer set and each species being amplified. Thus, for any species i , the number of amplicons should be directly related to the efficiency of amplification and the starting concentration of DNA template such that

$$A_i = c_i(1 + \alpha_i)^{N_{PCR}} \quad (1)$$

where A_i is amplicon abundance, c_i is the true number of DNA copies in the reaction attributable to species i , α_i is the amplification efficiency (bounded on $(0, 1)$), and N_{PCR} is a known constant (an integer giving the number of PCR cycles used in the reaction). If we could perfectly observe amplicons, the above equation would be all that we needed to relate the amplicons we observe and the biological value of interest, c_i , the true number of template DNA copies. Unfortunately standard PCR and sequencing technology does not allow for such direct observation. Due to N_{PCR} being a large number and α_i typically being not close to 0, the number of amplicons expected for any species with $c_i > 0$ is very, very large (e.g. with $c_i = 2$, $\alpha_i = 0.75$, and $N_{PCR} = 36$, $A_i = 1.12 \times 10^9$) and there are typically many species being amplified simultaneously, producing 10^{10} or more DNA copies in a single reaction. The actual number driven primarily by the α values among species and N_{PCR} . This model assumes that PCR amplification has not approached saturation and therefore the PCR is still amplifying exponentially. We, and others (53,55,56), argue this assumption is valid because 1) the total concentration of DNA within a filtered ethanol sample is low ($<1\text{ng}/\mu\text{L}$) and 2) the PCR reagents are supplied in excess and therefore are unlikely to be saturating the PCR. However, future models could be developed to account for a saturating PCR curve (68).

DNA sequencing machines do not read all of the copies from such a reaction; they read only a small fraction of the reads (on the order of 10^6 to 10^7 reads) (54,70). This subsampling changes what in eq. 1 appears to be a single-species process – each species being amplified independently – into a multi-species process; the number of amplicons observed for species i will depend upon both the amplicons produced for species $i = 1$ and the amplicons produced for species $i = 2, 3, \dots, I$ in the same reaction. Observations of amplicons are thus compositional data and need to be analyzed as such.

Models for compositional data

We want to retain the data-generating structure from eq.1 as much as possible, so we develop a model for a single sample with many species. As above, let i index species with $i = 1, 2, \dots, I$ and then we can write a deterministic equation for the number of amplicons observed in log-space as

$$\log(A_i) = \log(c_i) + N_{PCR} \log(1 + \alpha_i) + \log(\eta) \quad (2)$$

where the only new term is η which represents the proportion of reads observed from a given sampling run. Note that in this formulation η is a single value shared across all species and serves to scale the number of amplicons observed. Additionally we can rewrite the number of DNA copies in terms of proportional number of counts, $\log(\beta_i) = \log(c_i) - \log(\sum_i c_i)$. Note that the second term in this equation is a sum of the counts across all species, and so is a single shared value for all species. As such it can be absorbed into the value η that scales the overall abundance,

$$\log(A_i) = \log(\beta_i) + N_{PCR} \log(1 + \alpha_i) + \log(\eta) \quad (3)$$

This equation is appealing because it provides a process-oriented description of the biology of interest (the β s), a term for how PCR modifies our observation of the true abundance ($N_{PCR} \log(1 + \alpha_i)$), and a term for how subsampling of DNA reads in the sequencing machine will adjust the number of amplicons observed $\log(\eta)$. This third term also links all of the single-species components to produce a multi-species model. It is important to note that while both eq. 2 and 3 use the term η , the interpretation and plausible range of this parameter are distinct in the two equations. In eq. 2, $0 < \eta \leq 1$, while in eq. 3 η is not constrained to be less than 1 ($\eta > 0$).

In practice, PCR and subsampling are not deterministic but random processes (70). Furthermore, we are rarely interested in results from a single sample but rather multiple samples collected across sites j and times t . We let λ_{ijtk} be the expected number of amplicons observed, with k indexing unique PCR reactions to account for the fact that there may be multiple individual PCR reactions for a single collected sample,

$$\log(\lambda_{ijtk}) = \log(\beta_{ijt}) + N_{PCR} \log(1 + \alpha_i) + \log(\eta_{jtk}) \quad (4)$$

Importantly, α_i is assumed to be constant for each species among all sites, times, and PCR reactions (53,55,56). We incorporate stochasticity by allowing the number of observed amplicons to vary from the deterministic mean by specifying the observed values as following a negative binomial distribution,

$$Y_{ijk} \sim \text{NegativeBinomial}(\lambda_{ijtk}, \phi) \quad (5)$$

$$\phi = \exp[\tau_0 + \tau_1 \log(\lambda_{ijtk})] \quad (6)$$

where the expected value and variance of Y_{ijk} are $E[Y_{ijk}] = \lambda_{ijtk}$, and $Var[Y_{ijk}] = \lambda_{ijtk} + \frac{\lambda_{ijtk}^2}{\phi}$, respectively.

Note that we allow for the scale parameter ϕ to vary with the predicted mean, this allows for the amount of dispersion in the negative binomial to shift with changing predictions. For our datasets, this allows dispersion to be large when λ is small and decrease as λ increases.

By itself, this model has substantial identifiability problems; in the absence of additional information, it is not possible to estimate the β and α parameters. In the next section we discuss how to integrate data from amplicon sequencing as well as other data sources to make the parameters identifiable.

Application to the CalCOFI Dataset

At each CalCOFI station, an oblique bongo-net tow is conducted from 210 m to the surface with the starboard side preserved in buffered formaldehyde and the port side preserved in buffered ethanol-preserved (See Supplemental S1 for more details)(12). Manual counts of ichthyoplankton were quantified using microscopy to identify species abundance from formaldehyde-preserved samples. Metabarcoding was conducted on the ethanol preserved side; consequently, we expect the contents of the paired samples to differ slightly as a function of sampling stochasticity. Counts of larvae/juveniles were done once for each jar.

The manual counts provide extra information that enable us to estimate the confounded parameters from the metabarcoding. Specifically, for each sampled station, we have two sets of observed data: 1) counts of larval/juvenile fishes for each taxon from the formaldehyde jars (Z_{ijt} ; indexes as above) and 2) counts of amplicons for each taxon from ethanol jars (Y_{ijtk}). These observed data arise from a common (but unobserved) biomass for each species at each station-year combination (γ_{ijt} ; a latent (unobserved) variable).

We assume that the data are counts for each species in each sample, Z_{ijt} , derived from the true density of each species γ_{ijt} , the fraction of total specimens counted in each vial, P_{jt} , and the volume of water filtered for that station relative to a standard volume, V_{jt} ; $V_{jt} \approx 1$ for most samples, $V_{jt} < 1$ indicates a smaller volume of water was sampled.

$$Z_{ijt} \sim \text{Poisson}(\theta_{ijt}) \quad (7)$$

$$\log(\theta_{ijt}) = \log(\gamma_{ijt}) + \log(P_{jt}) + \log(V_{jt}) \quad (8)$$

We consider β_{ijt} to be the true proportion of biomass at a given station-year for each taxon i , $\beta_i = \frac{\gamma_{ijt}}{\sum_i \gamma_{ijt}}$.

Joint Model for Counts and Amplicons

To combine our information from the manual counts and metabarcoding, we need to recognize that our observations (Y_{ijtk} and Z_{ijt}) are linked together by a common variable (γ_{ijt}) and thus we can model them jointly (71). We represent the amplification process using equation 5 and 6 above (amplicons were sequenced in triplicate reactions for each jar). The manual count are modeled as in equations 7 and 8.

Our model assumes the fraction of template DNA in each sample is proportional to the counts of individual species' larvae in each paired jar (53,55,56). Moreover, we assume that in each sample there is template DNA from species that are uncounted, unidentifiable, or otherwise unobserved (69,70). In practice, this DNA might derive from stochastic sampling between each side of the bongo net, excreted waste, stray tissue, cells, or microscopic genetic material extracted along with the visible larvae.

Dealing with the fact that not all methods see the same species

The above is sufficient if all of the species identified by morphological counts are identical to the species identified by the genetic methods. But this is often not the case; some larvae are not separable to species based on morphology and some species are not separable to species based on a single genetic primer. Furthermore, some species do not amplify at all in the PCR ($\alpha_i \approx 0$). To accommodate non-overlapping sets of species among sampling methods we introduce a new variable, γ_{Mijt} , which specifies the true (M is for "main") density of species i at site j and time t . We assume that there is a mapping between this main density and the density observed by each sampling method. Specifically, we assume the species in the main list maps uniquely onto no more than one taxonomic group in each observation method, but multiple main species can map onto a single group for each observation method. For example, if the observation of larval counts identified as a specimen as *Sebastes* sp., we assume this may map onto one or more specific taxa (e.g., *Sebastes paucispinis*) in the main list, but conversely, *Sebastes paucispinis* on the main list may not map to more than one entity identified by each observation method.

We can construct a mapping matrix, \mathbf{M}_{MS} , that transforms the species in the main list, γ_M (a vector of length I_M , the number of true species in the sample) into the species grouping observed by sampling method

S , γ_S (a vector of length I_S , the number of groups observed by method S). We drop the j and t subscript because this mapping does not depend on the identity of the community being sampled. Then,

$$\gamma_S = \mathbf{M}_{MS} \gamma_M \quad (9)$$

\mathbf{M}_{MS} is a I_S by I_M matrix.

For example, if there are four species in the true community and method S only observes three groups, the matrix \mathbf{M}_{MS} could look like this

$$\mathbf{M}_{MS} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (10)$$

This might happen if species 2 and species 4 (columns 2 and 4, respectively) were from the same genus and the PCR primer from method S can only resolve those two species at the genus level. To provide a further example, take an invented community of four species with $\gamma_M = \{1, 15, 6, 7\}$ individuals in the community. The true community as observed through method S would be

$$\gamma_S = \mathbf{M}_{MS} \gamma_M \quad (11)$$

$$\begin{bmatrix} 1 \\ 22 \\ 6 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 15 \\ 6 \\ 7 \end{bmatrix} \quad (12)$$

and so γ_S is a linear combination of the true community. Of course there is no requirement that elements of γ_M be integers, but that makes the above example easy and transparent.

It is easy to incorporate this added complexity into the models in the previous section. If we assign method S to be manual counts and W to be the Mifish PCR primer, we need to construct a main list of species to define γ_M and build two mapping matrices, \mathbf{M}_{MS} and \mathbf{M}_{MW} that determine which species or species-groups are observed by each method. We can then add an additional subscript for each additional method and use the same form as above. For example,

$$\log(\theta_{Sijt}) = \log(\gamma_{Sijt}) + \log(P_{Sjt}) + \log(V_{Sjt}) \quad (13)$$

$$\log(\lambda_{Wijtk}) = \log(\beta_{Wijt}) + N_{W,PCR} \log(1 + \alpha_{Wi}) + \eta_{Wjtk} \quad (14)$$

And with additional sampling methods, we can make different mappings from the true abundance to the observations of each method.

Estimation and Identifiability

We developed and fit the above model in a Bayesian framework using the Stan language, as implemented in *RStan* (75, 76). All code is available as supplementary material. Table S2.1 provide prior distributions used in the model.

We ran five MCMC chains with 1,000 warmup and 4,000 sampling iterations. We retained every other MCMC sample. We initiated each chain at randomly determined starting values. The model converged ($\hat{R} < 1.01$; Gelman-Rubin diagnostics) and had no divergent transitions. We performed standard posterior predictive checks to assess model fit.

Table S2.1: Prior and parameter descriptions for the Stan Model.

Parameter & Prior	Description
$\alpha_i \sim Beta(1, 1)$	Amplification efficiency for species i
$\log(\gamma_{Mijt}) \sim Normal(0, 4)$	True biomass of each species at each site-year
$\log(\eta_{jtk}) \sim Normal(-4, 4)$	Estimated offset for each PCR reaction at each site-year
$\tau_0 \sim Normal(0, 2)$	Negative Binomial shape parameter intercept
$\tau_1 \sim Normal(0, 2)$	Negative Binomial shape parameter slope