Response to Reviewers

Editor Comments

Dear Dr. Gold and colleagues,
Your manuscript provides a potentially very interesting data set on the heat wave induced shift in larval fish composition in the California Current system, using a combination of meta-barcoding and morphological determination. The ms received generally favorable reviews, but needs a thorough revision before it can be considered for publication in Science Advances.

Specifically, I encourage the authors to discuss (i) the limitations of the chosen barcoding gene (12S) and the two step PCR procedure (ii) to address other possible causes for the community shift than temperature (iii) address the limitations of using a short time window in spring that restricts the larvae captured to a subset of all possible species (iv) address the issue of having sea surface temperature, while the larvae are often belonging to meso-pelagic species of deeper water layers.

We thank both the editor and reviewers for their constructive feedback and suggestions. We have addressed all reviewer comments thoroughly and hope that formatting the manuscript to *Science Advances* guidelines and including a detailed methods section will address many of the concerns raised here.

discuss (i) the limitations of the chosen barcoding gene (*12S*) and the two-step PCR procedure:

The points below are briefly discussed in the supplement (Lines: S1 125-147). In lieu of specific details regarding reviewer 1's comment regarding the "stochasticity" of the *12S* barcoding gene, we provide a summary of why we employ the MiFish Universal Teleost *12S* primer set with regards to taxonomic specificity and resolution. We further discuss issues of stochasticity and how we targeted the issue in this manuscript, and we discuss two-step PCR as a method to reduce amplification bias associated with sequence indices.

The MiFish Universal Teleost *12S* primers employed in this study are the most rigorously validated fish barcoding gene to date (*1*–*7*). The employed primer set is well benchmarked and provides highly accurate taxonomic assignments for a broad range of fishes (*2*). Given the balance of high specificity and breadth, the research community has largely converged on the MiFish Universal Teleost *12S* primer set as a gold standard for fish metabarcoding (*7*).

There are limitations for all metabarcoding primer sets (*8*), even a "gold standard" like the *16S* rRNA gene marker for prokaryotic sequences struggles with taxonomic assignment accuracy (*9*), especially with short-read sequences. All markers balance specificity [how well target species can be taxonomically resolved] and breadth [range of species across the tree of life that can be amplified] (*10*). Thus we wholeheartedly recognize that there are taxonomic resolution limitations and compromises for the *12S* particular primer set employed in this study (*2*, *11*).

Nonetheless, we have not simply accepted these well-known limitations. Instead, many of the co-authors on this manuscript have devoted substantial time and effort to identify limits and improve the function of this marker through the development of a nearly comprehensive California Current Large Marine Ecosystem *12S* reference database, as well as full factorial cross-validation analysis of bioinformatic approaches (See Gold et al. (*2*)). These efforts dramatically improved the taxonomic resolution of this marker set and identified best practices for taxonomic classification. Furthermore, the results provided in this manuscript demonstrate that the *12S* locus we employed improves species resolution as compared to manual

microscopy methods and is able to successfully recover the vast majority of species observed in each paired jar (Figure S5). In conclusion, we have rigorously benchmarked and validated this barcoding gene for our study system, and as our results demonstrate, we successfully capture important changes in fish assemblage dynamics also revealed through microscopy methods.

One reviewer commented on *CO1* barcoding. Please note that there are no widely used or benchmarked fish-specific *CO1* metabarcoding primer sets due to a) the conserved nature of the locus across the tree of life which results in amplification of a broad array of taxa (*12*, *13*) and b) the mismatch in high throughput sequencing platform length (max is paired-end 300 bp) and rate of *CO1* evolution/accumulation of sequence differences between species (*6*, *14*). Indeed, these shortcomings motivated researchers to develop alternative fish metabarcoding loci targeting *12S* loci for fishes (*7*).

Reviewer 1 mentioned that *CO1* barcoding genes do not have the same level of stochasticity as *12S* MiFish Teleost primer sets. We remain unaware of documented differences in stochasticity of *CO1* and the *12S* metabarcoding loci used here with respect to sequence counts. The reviewer's comment prompted additional literature queries, and the only related reference we could find was Krehenwinkel *et al.* (*15*) which reported that their *12S* marker targeting freshwater benthic macroinvertebrates had lower performance in the ability to reconstruct relative abundance estimates from mock communities. However, the marker they employed is entirely separate and only superficially related to that used in this study for marine fishes. That study is not applicable to this one because the primers a) target different regions of the mtDNA genome, b) have entirely different primer sequences and thus completely different amplification bias profiles, and c) have different fragment lengths impacting PCR dynamics.

To further address the reviewer's concerns, we conducted our own deep dive into the origin and source of variation in amplicon sequence data. These analyses identified stochastic dropouts in which one or more technical replicates have zero reads for a given species are the main driver of variation in this data set. For example, for anchovy we observed an instance of ~1,000 reads, ~800 reads, and 0 reads across three technical PCR replicates from the same DNA extraction. These stochastic drop outs are easily visualized along the x axis in Figure S2. We note that the highest observed species-specific amplicon read proportion associated with a stochastic dropout was 1.3% with the vast majority of such dropouts occurring below 0.5% read proportion within a technical replicate. These results suggest that stochasticity is largely driven by the proportional and absolute abundance of DNA molecules within a sample rather than a specific feature associated with a particular primer set, especially given that dozens of other metabarcoding studies have identified similar patterns (*16–18*). Furthermore, we note that most previous applications of the MiFish primers are from aquatic eDNA applications (*7*). Thus any observed apparent stochasticity, particularly in comparison to microbial and phytoplankton DNA (*19*), is likely a reflection of the relative rarity of fish sequences within a given seawater sample. Here we developed a comprehensive joint Bayesian model that incorporates stochasticity in observed amplicon read counts through a hypergeometric subsampling process (See supplement 2 for full model description). *Thus, we explicitly account for stochasticity in the model through sampling distributions and using the resulting parameters to estimate the uncertainty around our given estimated larvae counts.*

We were also asked to discuss the choice of a two-step PCR amplification method. Overall, this topic is well covered in the literature, including a recent review paper that outlines

a suite of advantages and disadvantages for both one-step and two-step PCR protocols (*20*). Relevant to our choice, previous work has demonstrated that two-step PCR amplification can reduce index-specific amplification biases (*21*, *22*). The inclusion of different indices on PCR primers appears to affect the relative amplification efficiency of each PCR, which is a concern for us given our desire to study an array of targets in an oceanic region over space and time. Granted, there is no clear best practice because other studies have found no effect (*23*, *24*), although we note no study to our knowledge has found more bias from the two-step PCR approach than the single PCR approach (*20*).

(ii) to address other possible causes for the community shift than temperature

We recognize the concerns of the reviewers and have rephrased the manuscript to better characterize the relationship between the shifts in ichthyoplankton assemblages and changes in temperature and other environmental variables (Lines 89-97). Specifically, we expanded the supplemental results sections to highlight site comparisons and acknowledge that the mechanisms underlying the change in fish communities are not definitively known. Nonetheless, our data are compelling in that despite the hundreds of kilometers and distinct oceanographic features associated with each site, we saw similar synchronous changes in fish communities in response to the 2014-2016 Marine Heat Wave. It is important to note that our sample design was chosen to reflect the well-documented four general oceanic regions within the Southern California Bight (See Supplemental Methods), with sites separated by up to 370 km. It was thus not surprising to find significant differences in fish assemblages between these sites as decades of previous CalCOFI research has documented distinct offshore and onshore communities as well as communities associated with the California Current vs. Southern California Counter Current (*25–27*) (See Supplemental Results). Importantly, the changes associated with the warm anomaly were detected here despite the variability introduced by sampling across distinct and persistent ecoregions.

Here we used temperature as a proxy for the Marine Heat Wave, recognizing that a suite of environmental variables including upwelling strength and location, dissolved oxygen, salinity, and nutrient concentrations, also changed significantly in response to the marine heatwave (*28–31*). We also note that many of these factors, particularly salinity, temperature, dissolved oxygen, and nitrogenous nutrients are strongly correlated in the Southern California Bight (*32*). To address the concerns noted, we clarified our observations and presented our results in the context of previous work, highlighting the lack of certainty around temperature being the main driver for these shifts in fish communities.

Clearly, temperature alone cannot explain the dynamics of fish communities in these ecosystems as best exemplified by the breakdown in sardine and anchovy temperature associations established prior to the onset of the Marine Heat Wave (Lines 316-342). However, the Pacific MHW of 2014-2016 itself had a strong climate change driven signature (*27*, *28*, *33–39*) and we found consistent significant differences in fish assemblages before and after the onset of this climate forced phenomenon. Thus, we argue that the occurrence of novel ichthyoplankton assemblages is strongly correlated with the onset and aftermath of this climate change-driven event. We thank the reviewer for ensuring greater clarity and recognition of nuance throughout.

(iii) address the limitations of using a short time window in spring that restricts the larvae captured to a subset of all possible species

We thank the reviewer for this comment and have highlighted why our study focused on Spring sampling in the manuscript (Lines 103-107). Previous research from the study region has found that the majority of species spawn in spring (*27*, *40*). Similarly, the CalCOFI program (https://calcofi.com/) has historically used spring data for their annual California Current Ecosystem Report providing decades of precedent for using spring data to investigate changes in ichthyoplankton assemblages (*41*). However, we acknowledge in the discussion that by limiting our analyses to spring sampling we will not observe species without pelagic larval stages. Likewise, we may miss phenological changes in spawning, although recent work has found little evidence for such trends (*27*, *40*).

(iv) address the issue of having sea surface temperature, while the larvae are often belonging to meso-pelagic species of deeper water layers.

We thank the reviewers for this suggestion. Originally we had conducted the analyses using the average water column depth from 10-100m (here in MWCT) collected from nearly simultaneously conducted CTD casts following the methods of Thompson et al. (*27*) to match species observations and relevant temperature regimes. However, we were equally concerned that a single instantaneous temperature measurement, even if from a more appropriate depth, may not be as reflective of the environment in which the larvae developed. Thus, we adopted the two-month average SST metric (separately obtained from remote sensing) to account for the average life stage of the observed species (See supplemental results).

To address this, we re-ran all analyses using MWCT instead of SST (See Supplemental Results, Lines 320-329). We note that two-month averaged SST and instantaneous MWCT were only 29% correlated. This finding is perhaps unsurprising given the substantial difference in spatial and temporal integration time of these temperature measurements (linear regression, $p<0.01$). Despite these apparent differences, we found nearly identical results in the direction and significance of species–temperature associations as well as temperature-driven variation in fish assemblage dynamics. These results suggest that species-temperature associations and our conclusions are largely robust to temperature metrics. We report the two-month averaged SST in the main manuscript and include MWCT analyses in the supplement (Figures S6-S13).

One reviewer is concerned that the ms itself present highly derived analyses of community shifts. I agree that the author need to do a better job in presenting the fundamental data underlying their analyses, including a clear indication when morphological and when genetic information was used. The authors also need to make good arguments why specific species from the potential collection of 56 species were assessed /presented in depth (cf. Fig 1). This will also entail to move some material /graphs from supplement to the main paper, for example some of the heat maps. Finally, I wonder if it would be possible, though some process of data set randomization, to come up with a null-model on SST vs. species abundance correlations that could happen by chance.
If the authors think they can successfully address all issues raised by the referees, I am happy to receive a revision. This revision will be sent out for re-review, potentially to additional referees.

We thank the editor for their comments and have taken efforts to clarify our writing throughout the manuscript, particularly by including a Methods section that highlights how the data was generated and how they were combined into the joint model. From reading the reviews it is apparent to us that we did not adequately communicate the significance of the joint model or how the model was implemented. We apologize for that mistake as it is ultimately one of the key

novel developments in this paper. Please see the added Methods section in the manuscript (Lines 99-201).

Amplicon sequence data are compositional in nature and thus can only provide relative abundance estimates (*42–44*). By nature, compositional data is inherently limited in its inability to discriminate between 1 anchovy out of 10 observed larvae from 100 anchovy from 1,000 observed larvae. Thus relying on compositional data alone limits our understanding of fish assemblage dynamics. The interpretation of compositional data is further challenged when the resulting amplicon sequencing reads are dependent on co-occuring species within a given sample (*42*). Thus the objective of the joint model was to ground the higher diversity compositional amplicon sequence dataset to the absolute abundance estimates from microscopy-derived larval counts and derive abundance estimates from species only observed in amplicon sequence data. This is done by linking compositional changes in amplicon sequence reads to larvae counts through a mechanistic joint model based on the PCR equation (See supplement 2).

Importantly, to the reviewer's concerns, all analyses presented in the manuscript are conducted on the model estimates of larvae abundance (counts per volume sampled). We argue that without the microscopy counts, amplicon sequences from metabarcoding data are largely uninterpretable for quantitative estimates as observed reads are a function of both the input DNA concentration as well as the amplification efficiency. Thus by linking the abundance estimates from microscopy data to the input DNA concentration, the model becomes identifiable and allows for the estimation of larvae counts for species that cannot be observed/resolved through microscopy.

We note that one reviewer requested comparisons between both methods as is standard practice in many metabarcoding ground-truthing studies including ones conducted by our co-authors (*45*, *46*). To address this we included a heat map of observations between metabarcoding and microscopy counts (Figure S5) as well as general descriptions of the overlap and mismatch of species observed by both methods (See Supplemental Results). However, such comparisons are tangential to our objective in this study: to develop quantitative estimates of ichthyoplankton abundance to investigate how fish assemblage changes in response to the 2014-16 MHW.

Ultimately, our results demonstrate the strong performance of the model and suggest that our compositional PCR-based framework provides a widely applicable tool for deriving quantitative abundance estimates from compositional amplicon sequencing approaches. We believe these results to be of great value to the broader eDNA, microbiome, and metabarcoding fields as they provide a framework for obtaining quantitative estimates from these data streams.

Lastly, we chose to focus our analyses on the 56 species that had sufficient representation across the metabarcoding data set to allow for accurate regression analyses and model convergence. We followed the often-cited "one in ten rule" for the minimum number of parameters to be estimated and thus only included species observed in at least 10 technical PCR replicates (*47*). Although we acknowledge such a threshold is inherently arbitrary, we argue that the other 74 taxa were observed too infrequently across the dataset to reliably track changes in abundances across the samples. Statistical approaches that rely on only using data with a minimum number detection are routine in ecological studies (*48*) and we feel confident that our approaches were consistently applied across all taxa and are statistically sound.

Reviewer: 1

In this study Gold et al., investigates changes in larval fish communities over 23 years using a combination of metabarcoding, morphological identification, and modelling to illuminate potential effects of marine heatwaves on

the tropicalization of CCLME. I think this is an excellent study and illustration of an important, but understudied, phenomena. I think that the study will have wide implications and perspectives in other marine systems. I have some major issues that may be useful for the authors to consider, followed by some minor comments: 1. There are no alternatives discussed for why the larval fish communities may have changed, i.e. what other factors could have caused the observed changes? Climate change and marine heat waves are forwarded as the only explanation for the observed results, whereas, e.g., multiple stressors or climate induced changes in the ecosystem or in predator-prey relationships, may have a similar effects on the fish community? The authors reflect on this in L142-145, and even use "strongly" in that sentence, so I think it would benefit the study to handle the results in a more nuanced manner. I still want to acknowledge that it is inherently difficult to assess the impact of climate change on large scale marine systems.

We thank the reviewer for their comments and agree that the language used in the manuscript should be more nuanced in terms of attribution of impacts of temperature on fish communities. As described above, we focus on temperature as a proxy for the multi-stressor abiotic and resulting biotic changes in the CCLME ecosystem and have rewritten the results and discussion to highlight the inherent difficulty in assessing the impact of climate change on large scale marine ecosystems (Lines 369-371).

2. The use of 12S as a quantitative marker is not so straightforward, as this marker is well-known for its stochasticity when analyzing eDNA samples. This especially, when the study only uses one biological and three technical PCR replicates. Additionally, a two-step PCR protocol was used to amplify/make the libraries which introduce bias in species composition, and likely in the abundance estimates. Although the authors provide a discussion on limitations of the metabarcoding approach (L516-L529), this discussion is focused on the state of the samples, not the nature of the marker. I think that the authors should, at least, provide a discussion of the impact of the 12S stochasticity on their results (in the supplement) and consider if the study would benefit from a validation of the results using occupancy modelling. If possible, the samples could also be analyzed using metabarcoding of COI (that do not have the same issue of stochasticity).

Please see the discussion above where we detail why we are unable to agree with the reviewer's comments about the stochasticity of the *12S* marker set. We have included a suite of analyses and a summary of the stochasticity observed in the dataset within the Supplemental results (Lines S1 363-387).

3. The study uses sea surface temperatures to predict the marine heat waves, but most of the changes in community is observed for mesopelagic and/or species that are not having the surface waters as their main habitat. I think the study would benefit from results, including references in the main body, that show that the MHWs observed in the surface waters, also happens and/or change the termohaline properties of the deeper habitats.

As discussed above, we re-ran all analyses using the mean 10-100 m water column temperature (MWCT) and found nearly identical patterns of assemblage shifts in response to temperature (See supplemental results, Lines S1 320-329). We have also included a greater description of the physical characteristics observed in marine heatwaves within the water column to provide greater context for their effects (Lines 89-97).

Minor inputs L23: Any knowledge on the fish community earlier than 1996? A description would help the reader to understand the impact of MHWs.

Yes, CalCOFI surveys extend back to 1951 and we have included a description of previous fish communities and their responses to changing climatic regimes as highlighted in decades of previous research. We include citations of highly informative manuscripts delving into these results in detail. Lines 49-58

L24: "Increases…", please state what the increases are (biomass, abundance, species…).

We clarified the increases are associated with species abundance (Lines 30).

L65-71: A long and convoluted sentence, please rewrite.

We thank the reviewer and re-wrote this sentence to improve clarity (Lines 160-165).

L73: Please justify the use of "accurate".

Please see Supplement 2 for details of model specificity and performance.

Also, to me, Figure S3 (L74), do not show any correlation between the metabarcoding and the morphological results. I guess that the authors also forgot to add the 1:1 line in Figure S3, although used in figures S1 and S2. Please discuss and clarify this.

Thank you for your comment. Clearly, as pointed out, there is a poor linear correlation between metabarcoding sequence data and morphological counts (Figure S2 [formerly S3]). PCR is an exponential process and not quantitative unless followed at each cycle, and as such the results of metabarcoding data are inherently compositional (*18*, *42–44*). This is especially true for metabarcoding PCR in which multiple species compete within a PCR reactions for general primers with each species having a unique amplification efficiency (*42*). Therefore is is unsurprising that a linear correlation fails to capture the underlying dynamics of the interaction between DNA template abundance and amplification efficiency. We agree that the poor relationship is obvious from Figure S2 and why we initially did not include the 1:1 line, but we have now included this.

Ultimately, Figure S2 serves as justification for the application of our joint modeling efforts. The goal of this study was to move beyond the typical eDNA approach of "let's fit a simple linear regression to the data and hope we find a correlation." As was described in the model description and motivation provided in Supplement 2, the objective of this study is to develop a mechanistic understanding of the underlying process of PCR that explains the observed sequence read patterns so as to be able to successfully leverage morphological counts and metabarcoding data streams to provide a greater understanding of ichthyoplankton assemblages. In essence, our joint model builds off previous amplicon sequencing studies to correct for multiple competing targets with different amplification efficiencies within the PCR process (*42–44*, *49*). Importantly, we treat both observed sequence reads and counts as arising from the true biomass of a given species in a jar linked by the outlined PCR equation in Supplement 2. The results from Figure S3 and S4 demonstrate that employing the joint model substantially decreased variance and increased correlation between predicted and observed abundances. We argue that this approach is a major advancement for metabarcoding and provides an avenue for achieving quantitative amplicon sequencing which is highly relevant to a whole range of fields from microbiome to eDNA.

L88: Should it be SST?

Clarified as SST (Line 237).

L109: Remove the additional space

Removed the additional space (Line 263).

L135: Consider adding the year(s) for this observation for helping the reader

Years of the marine heatwave were added (Line 299).

L145-147: It is unclear to me how the methods of your study accurately can differentiate between prey and predators. Please clarify and describe the approach in M&M.

We clarified the paragraph as follows (Lines 324-340): “Further improving our mechanistic understanding of drivers of fish assemblage dynamics will better inform ecological predictions in the face of extreme ocean events such as MHWs which are likely to increase in frequency and duration under climate change (58). As we demonstrate, a combination of metabarcoding and visual surveys can characterize species across trophic levels (59) and this has the potential to reveal ecological mechanisms. Here, we used metabarcoding to accurately characterize the composition of larval fishes in CalCOFI plankton samples. Future efforts could focus on documenting the phytoplankton and zooplankton assemblages that comprise both larval prey and predators. Several major hypotheses seeking to explain recruitment variability are underpinned by the capacity of young larvae to consume appropriate prey that facilitates faster growth (60). Unfortunately, accurately characterizing the larval prey field has traditionally been difficult as prey are generally too small to be accurately sampled by nets (61). Metabarcoding of water samples from the same locations where larvae are collected, however, can characterize the larval prey field at an unprecedentedly high level of detail. In addition, metabarcoding of the stomachs of larval fishes can then identify actual prey items that were consumed by larvae. Evaluating the larval prey field and gut contents through metabarcoding will help us to finally understand the drivers of recruitment volatility in coastal pelagic and other fishes (5, 53, 62–66).”

L155-156: Please clarify this, the marine biological responses to climate change due to global warming are widely studied and is happening at a greater rate e.g., in the Arctic, than in temperate marine systems.

We thank the reviewer for this comment. We were specifically addressing how biological response to both anchovy dominated forage fish assemblies and ocean warming conditions will be unprecedented given the novelty of this occurrence in the past >70 years of CalCOFI surveys (Lines 298-302). Given that previous anchovy abundance has been associated with cooler temperature regimes, such an occurrence is without precedence and thus our ability to predict biological responses to increases in frequency of MHWs are limited. Updated sentence is as follows (Line 346-350): “Given that conditions comparable to the 2014–2016 MHW are predicted to be more common in the CCLME in the future (1), our results suggest that continued biological responses to both anchovy-dominated forage-fish assemblages and MHW-associated ocean warming conditions are likely to be without modern analog (20).”

L166++: The use of “long-term”. I know it is a matter of definition, but I would use decadal throughout. Also “climate-driven”, I am not convinced that this is the only explanation (see also major-comment 1)

We clarified our language to be more specific and account for nuances throughout. Specifically, we used decadal in place of “long-term.” Please see the above discussion of major comment 1.

L168-171: Please revise following major comments

We clarified our language to address how obtaining quantitative metabarcoding estimates may improve such modeling and prediction analyses as compared presence-absence frequently obtained from metabarcoding datasets (Lines 369). However, we acknowledge that studying climate change impacts on ocean basin scales is inherently challenging as you have keenly pointed out.

L177-, Reference list: Not consistent (e.g., 10., 22.), please revise

Reference formatting issues were addressed (Line 377).

Supplements L326: Replace “other” with “port”

Duly noted (Lines 111-115 & S1 93-99).

L335: Please specify the pressure of the room, e.g. over-pressured. Also, I think a general statement of “anti-contamination” routines was used throughout the whole eDNA lab work, e.g. newly showered/suits/masks/notes on food.

We included a description of the contamination precautions taken (Lines S1 105-111). We note that we did not employ more typical ancient DNA precautions though we were diligent about routinely cleaning surface with 10% bleach and 70% ethanol, separating pre- and post-PCR product and samples, and keeping all food out of the lab. We also note that all lab work was conducted by a vegetarian, reducing accidental contamination.

L340-343: Consider moving this up in the text.

Moved, as requested (Line 103).

L350: Primer-pair?

Sentence is no longer in the manuscript.

L350-362: As the MiFish primers behave stochastic in the PCR, please comment on the effect of your two PCR setup in terms of favorizing specific species and how you avoid the stochasticity to affect the quantitative estimates. Also, see my general comment 2.

Please see the detailed discussion on these points above. If “the MiFish primers behave…” is meant to imply that if these primers behave in a manner that is particularly different from other primers, we were unable to find evidence for such, as detailed above. If instead, this comment was speaking to PCR generically, we agree that there are substantial differences in amplification rates among taxa, including when using the MiFish primers (*43*, *44*, *49*). Indeed, our model approaches this issue head-on and handles this explicitly by assigning a unique value for the amplification efficiency to each taxon (Supplement S2).

If the comment speaks to sporadic MiFish amplification from the same starting material -- whereby a taxon is amplified in one reaction but not in a replicate reaction, this is a well-known phenomenon general to PCR with rare templates (*16*, *17*, *24*), and we document such behavior in this manuscript (see Supplemental Figures S2-S4). This phenomenon adds noise to the observations and limits the accuracy with which we might predict amplicon abundances (particularly rare ones), but it does not fundamentally change the interpretation of our observations or of our model. We mention this outcome in supplement S1 (lines S1 363).

It might also be possible that the reviewer’s comment is about the possibility of differential taxonomic assignment (“in terms of favoring specific species”). If so, there are dozens of previous studies benchmarking and validating the MiFish primer set both *in vitro* and *in silico* (*2*, *5*, *7*), and the MiFish *12S* primer set used in conjunction with the bioinformatics pipeline of the *Anacapa Toolkit* has been rigorously tested for our study system (*2*).

With respect to the comment about 2-step PCR, please see the detailed comments above.

L362-364: Mention the amplification success.

Only filters from 4 jars failed to amplify, and upon further inspection within the archived notes, all these samples had known preservation issues (e.g., preservative dried out, observed mold, etc.). All other DNA extractions successfully amplified. This information is available on Line of supplement S1 (Line S1 163).

L384: Revise, The/Our metabarcoding..

Fixed (Lines S1 162).

L393-396: Excellent that you have reported how you treated index jumping. This is something that is missing in many studies.

We appreciate the compliment and agree that accounting for index jumping is a critical step (Line S1 199).

L450-453: Please clarify, especially "… estimation 3 latent…"…

The paragraph was rewritten as follows to improve clarity (Lines 270): "To explore co-occurrence patterns across species, we fit a generalized linear latent variable model (GLLVM) following the methods of Niku et al. (39), specifically conducting model fitting to determine the best distribution fit as well as number of latent variables to use. The highest performing GLLVM employed 3 latent variables and applied a negative binomial distribution with variational approximation (40) on the joint model predicted larvae count data. We then plotted the correlation matrix of the linear predictors across species with and without incorporating SST in the GLLVM to identify co-occurring species and the effect of SST on co-occurrence patterns."

L485-494: This section appears to hold two redundant statements. Please revise.

We paragraph was rewritten as follows to improve clarity (Lines 304-319): "We also found strong positive and negative co-occurrence patterns through GLLVM analyses across species when controlling for temperature. Specifically, we observed strong negative co-occurrence patterns between fisheries targets like benthic fisheries targets (e.g. Citharichtys sp. sanddabs) and mesopelagic fishes (S27). We also observed strong positive co-occurrence patterns between a suite of benthic species as well as strong positive co-occurrence patterns between a suite of mesopelagic fishes. These results suggest that when controlling for temperature, we observe strong benthic versus pelagic tradeoffs as observed previously (6, 41, 42).

When focusing on co-occurrence patterns without controlling for temperature, we observed strong negative co-occurrence patterns with fisheries targets (North Pacific Hake Merluccius productus) and mesopelagic fishes (S28). Here, temperature explained 19% of the variability among species co-occurrence and was particularly important in driving negative co-occurrence patterns between North Pacific Hake and mesopelagic species. These results suggest that temperature may mediate tradeoffs between fisheries versus southern mesopelagic fish assemblages. Further work exploring the underlying mechanisms of these negative co-occurrence patterns is warranted."

L512: Add "the" before "DNA"

Fixed (Line 346).

L516-529: Please see major comment

Please see the previous responses regarding stochasticity. However, in response to the comment on "validation of the results using occupancy modelling" we note that the implementation of the joint model specifically observes the frequency of stochastic dropouts and uses this as a prior for informing abundance estimates (See supplement 2 for full model implementation and description). In essence, our model is analogous to site occupancy modeling using the pattern of occurrences as well as proportional read counts to inform the abundance of a given species within a given sample. Our results clearly demonstrate that stochasticity is associated with rare molecules, both low morphological counts of species and low proportional sequence read abundance. This point is now mentioned in Lines S1 363 of supplement S1.

L532- : Please clarify the all the formatting issues, including the marking

We thank the reviewer for their comment and have taken better care to address the various formatting issues within supplement S1 references (Line S1 390).

L600-603: Please provide a better caption. This one does not describe the figure adequately.

The improved caption is as follows (Lines 574): "Predicted sequence reads are generated from the joint Bayesian model. The one-to-one line is plotted in red and Pearson correlation coefficient is 0.95. In general, predicted sequence reads track observed sequence reads, and show substantially less variance than observed sequence reads in Figure S3. However, unexpected zeros across multiple technical PCR replicates (stochastic dropouts), deviate notably from expected, low variance results (points along Y-axis). We note that all observed stochastic dropouts occur in less than 2.9% of sample read proportions and less than 9 morphological counts for a given sample. Such dropouts are likely a function of subsampling rare DNA molecules associated with molecular biology processing (See Supplement 2)."

L604-608: As this a key result for supporting the results of the study, please provide a 1:1 line and direct the reader towards what the figure reveal.

As discussed above, we have clarified the text to better explain that this figure is a key justification for the application of the joint Bayesian model; it is not in and of itself supporting the results of the study (Lines 545-556). A line has been added as per the reviewer's request. We did not initially add a 1:1 line because we did not wish to imply that such a linear relationship is expected, which it is not, as detailed above. To restate here briefly, the nature of metabarcoding PCR precludes simple linear correlations between sequence reads and morphological counts. We have tackled this obstacle by application of a mechanistic model that incorporates our understanding of the PCR process (*43*, *44*, *49*). As outlined in detail in Supplement 2, we build off previous amplicon sequencing work to apply a mechanistic compositional PCR model to incorporate both data streams (sequencing and counts) to better predict quantitative abundances of all observed species. As highlighted in Figures S3 and S4, our model provides improved performance over a simple linear correlation.

Ultimately, it is Figure S3 and the resulting predicted morphological counts that underpin our results, supporting the application and performance of our mechanistic model, particularly for abundant ichthyoplankton where the variance of model output is low. Critically, what is not depicted in Figure S3 is the suite of species that could not be identified to species level by microscopy. Our model extends such accuracy in quantitative estimates to dozens of species that cannot be resolved morphologically, improving our quantitative understanding of fish

assemblage dynamics across all samples. To this end, we have mentioned this point in Lines 116-165 and included a further description of the motivation for our modeling efforts and implications throughout the manuscript. We also wrote a new figure caption: “Observed sequencing reads and morphological counts do not follow a clear linear relationship. The one-to-one line is plotted in red and Pearson correlation coefficient is 0.56. This non-linearity is unsurprising given that observed reads are a function of both DNA concentrations (here assumed proportional to morphological counts) as well as species-specific amplification efficiencies (here are unknown) (See Supplement 2). Thus without accounting for species-specific amplification efficiencies within the compositional amplicon data set, we do not expect to observe a clear positive relationship. This apparent lack of relationship depicted here motivated the creation of the mechanistic joint Bayesian model. We also note the occurrence of stochastic dropouts (technical replicates with zero reads) can be observed along the X-axis. We also note that variance is highest at low observed morphological counts.”

L614: Tone down the importance of your results “ Strongly positively…. correlated “. This is SST, not water body temperatures. I have difficulties to understand how you can extrapolate SST to the deeper water layers, see major comment

We have revised the manuscript accordingly and have toned down the importance of our results. To address the concern with SST limitations, we re-ran all analyses using the MWCT as described in detail above (Lines S1 320-329). Although both temperatures are only partially correlated, we found similar results of species changes in response to both temperature metrics resulting in the same ecological conclusions. We thank the reviewer for this comment because the re-analysis further strengthens our conclusions. Please see supplemental results (Figures S6-S13).

Reviewer: 2

Gold et al 's study of larval fish community change over a period >20 years is interesting and valuable in the context of tracking the effects of ocean warming on marine communities. The study is also probably the first one to use the DNA leached into the sample jars to enhance the value of pelagic marine collections from previous years. I like the idea and I agree with the authors' general thrust that this approach for the use of DNA metabarcoding can be a powerful additional tool to monitor ecosystem change. I have, however, struggled a bit to see the link between the evidence presented and the claims made and discussed in this manuscript. I am listing below the issues that I believe should be addressed to warrant publication in Science Advances: - The main manuscript contains highly "derived" evidence, with some combined trends of abundance (making it impossible to tease out morphological vs molecular inference) and their change over time (without considering the differences among the four sampled locations), and some habitat/temperature associations, with a couple of species 'plucked' out of the 56, exemplify correlations with temperature.

We thank the reviewer for their comments and hope to address each of their concerns in turn. We recognize the importance of comparing and contrasting methods and have included a supplemental figure that highlights the co-detection of all species from the surveyed jars (See Figure S5). We agree that the manuscript uses “derived evidence” to make ecological conclusions, and we argue that such approaches are necessitated by the complex nature of compositional amplicon sequencing data (*42–44*, *49*). Compositional data is inherently difficult to interpret as the pattern of observed sequence reads is dependent on the co-occurrence of other species within the sample as detailed extensively in Supplement 2. Thus, on their own, raw observed sequence reads from metabarcoding are very difficult to interpret meaningfully as they are just as much a reflection of 1) who is in the jar, 2) how well did they amplify compared to the other species, and 3) how abundant are they (*42*). The objective of this study was to extract meaningful abundance estimates from metabarcoding data and move past the

simple Venn diagrams and comparisons of amplicon sequence data and manual methods (*45*). We wanted to leverage the strengths of amplicon sequence data and microscopic counts to get a more complete picture of the ecosystem surveyed. Specifically, metabarcoding is able to identify a wide diversity of fishes even if they lack distinguishable morphological characteristics which is particularly useful with microscopic larval fish. However, at best such information provides proportional abundances changes. On the other hand, morphological count data provide absolute abundance estimates. Therefore, using both data streams in the Bayesian joint model allowed us to get the best of both information sources and improve estimated abundances for a wide array of species including those only observed with metabarcoding.

As detailed above, we focused on temperature as a proxy for the onset of the marine heatwave while recognizing a suite of abiotic factors that changed substantially during the event including upwelling strength and duration, dissolved oxygen, and salinity among others (*30*, *31*, *52*). That being said, we have followed the reviewer's advice and taken greater care to present a nuanced analysis of the data without over-extending our findings and highlighting the use of temperature as a proxy for the onset of the MHW. We point to our discussion of the breakdown of the anchovy and sardine dynamics as an example of our recognition of complicated ecosystem dynamics and the need for a mechanistic understanding of drivers as opposed to correlative analyses alone (Lines 314-350 of the current manuscript).

As also discussed above, we have better explained that "without considering the differences among the four sampled locations" was an intentional study design decision. Indeed, the observed patterns with temperature across these diverse regions is cornerstone to our conclusions. We chose 4 sites that represent distinct biogeographic provinces within the greater Southern California Current Ecosystem including the California Current, California Counter Current, and inshore and offshore stations (See Supplemental Methods for full description). Over a half-century of research into these stations have identified distinct fish assemblages associated with these habitats (*27*, *53–55*), and we were wholly unsurprised to find differences in fish communities across >370km distances (See supplemental Results). However, what is telling is that we found synchronous changes in fish assembly dynamics across all four sites that correlate with increases in both SST and MWCT. These results suggest a shared shift in ichthyoplankton communities in response to the marine heatwave that is robust to the temperature metric employed.

As noted above, the 56 taxa chosen were selected based on the need for at least 10 data points across technical PCR replicates to allow for reliable model parameter estimation (*47*). All other taxa that were excluded were observed too infrequently to allow for accurate quantitative estimates. Applying such a signal to noise thresholds is common in ecological data sets (*48*, *56*). We recognize that the Bayesian modeling effort and the multi-tiered analysis is complicated and technical. We appreciate the opportunity to clarify our objectives and methods to better communicate our thinking and analyses. We wish to outright state that we had absolutely no intention of cherry-picking data or "plucking" individual species to tell a preconceived narrative. We strongly adhere to open science standards and have provided all raw data and code used to process the raw data. All data and code will be made publicly available upon acceptance and we have supplied a GitHub link now for review (https://github.com/zjgold/CalCOFI_eDNA). We sincerely hope that our edits have clarified this issue.

(By the way, it wasn't clear to me why the northern lampfish is always qualified as "unique cold variant").

We added a sentence on Line 210-213 to address this point of confusion. From our metabarcoding dataset, we found that two dominant ASVs with contrasting patterns of detections across all samples were both assigned the Northern lampfish (*Stenobrachius leucopsarus*) by the *Anacapa Toolkit* and bioinformatics parameters employed (See Supplemental Methods). These distinct ASVs are indistinguishable via morphology, especially at the ichthyoplankton life stage (*53*). Given the distinct pattern of presence of each ASV across the dataset, we decided to treat them separately in our analyses. One ASV was associated with colder SST and thus we decided to name it the "unique cold variant". We note that Myctophids are a highly understudied group of marine fishes and the taxonomy, cladistics, and phylogenetics of this group are still being actively resolved and this may potentially represent a separate species, although further work to resolve such patterns are clearly warranted (*57*).

The readership needs to see how the novelty of metabarcoding the leached DNA can significantly enhance the conclusions that would be gleaned by morphology alone. The modeling carried out to come up with the abundances is very sophisticated, but a bit of a black box for most readers, so it is quite important that the two patters are compared and contrasted. – linked to the above, I think it is important to also report exactly how many samples were examined (were they 23 (years) x 4 (sites) = 92, plus 2 controls?), how many reads were generated, and how many species detected, for each method. I saw the analysis across technical replicates over the years in the supplementary materials, but I could not find information on variety and homogeneity. One should be able to go and check the number and identity of the species detected by each method over the years. – There are a lot of correlations in the supplementary material, but the interaction between community composition over the years and the four sampling locations is not resolved. Some coefficients for statistical and biological significance differ depending on whether sites are pooled or kept separate in the temporal analysis. Without a clear demonstration of how these variance components compare in magnitude (and I suspect the spatial component accounts for for more than the temporal one), it is difficult to make bold statements on ecological replacement and cascading socio-economic impacts. I noticed the use of MDS, which is one way to comprehensively represent data sets (though not one of my favourites), but the supplementary material only contains one such plot, with combined abundances and combined sites. Providing portrayals for DNA and morphology, separately, and keeping the four sites separate, would help readers better understand the actual scenarios.

We included a paragraph description of the motivation for the sophisticated Bayesian joint model and highlighted the importance of grounding compositional and highly diverse metabarcoding data with the absolute count data obtained from morphology (See Introduction, Line 116-165). The value of the joint method is discussed at length above.

We also included Figure S5 to provide a summary of reads to compare and contrast the detections between morphology and microscopy. This figure allows the reader to readily identify patterns of detection between both methods. Furthermore, we included a general description of the results Line 204-221 that presents the number of samples, number of species, larvae counted, and sequences generated for this study. However, given that raw, untransformed sequencing reads are not readily interpretable on their own, we did not believe providing a summary figure of these results beyond Figure S5 was warranted. However, we note that all raw data, including tables containing all sequence and larvae counts as well as the code to process them, are made available and can be explored in detail more thoroughly outside a word document or figure.

We also point to the previous descriptions of our study design and site selection which address the concerns with variance between sites. Our findings illustrate substantial differences in fish communities across sites, explaining ~3x the variance than temperature. However, we focus on the shared changes in fish assemblage dynamics across the four disparate

biogeographic regions as such patterns demonstrate synchronous ecological responses to the MHW across hundreds of kilometers. Future work investigating the full array of 70 CalCOFI gridded stations sampled quarterly since 1996 will provide the spatial and temporal resolution needed to better elucidate ichthyoplankton assemblage dynamics in response to temperature, salinity, dissolved oxygen, productivity, etc. In light of our limited observational analyses presented here, we clarified our findings making sure to not over-extend the findings of our results.

- I was surprised about the removal of replicates with less than 30,000 reads. It seems like a notably high threshold, and certainly not commonplace in these studies. Why?

We found the Bayesian model to be relatively sensitive to sequencing depth as a function of sampling intensity and thus chose to focus on samples with high sampling effort. We agree that most studies employ more relaxed sequencing depth thresholds. However, we note that we only removed 6 technical replicates with this threshold given the high sequencing depth obtained from using the NextSeq platform.

- what drove the choice of poisson (for morphology) and neg binomial (for metabarcoding) for the modeled distributions? - an interesting aspect is the nature of the DNA obtained: this is coming from animals captured from a bongo net, hence more akin to community DNA than eDNA. This means that the DNA would be expected to derive almost exclusively from the animals captured, so the mismatch of detections between methods should be minimal, but it's never shown and it should, even if the morphological ID was carried out from a different jar. –

We included the following sentence in the introduction to clarify our choice of modeled distributions (Line 160-165): “Microscopy counts were modeled as Poisson-distributed given their relatively small absolute values and low variance (33), and amplicon sequence data were modeled using a Negative Binomial distribution given their relatively high absolute values and high variability among replicates (Figure S2-S4). These statistical distributions are commonly used in models of amplicon and count data, respectively (35–37). A full description of model implementation and assumptions are provided in Supplement 2.”

We thank the reviewer for their insightful comments and agree that compared to aquatic eDNA, mismatch in detections within ethanol-preservative should be minimal. We included a summary of the detections made by both methods in Figure S5. Lines S1 360-362:”Across a total of 4,704 possible detections, 70.2% were non-detections by both methods, 11.2% were detections by both methods, 16.4% were detections only made by eDNA, and 2.1% were detections only made my microscopy.”

Thus, we observed relatively few mismatches between methods. We note that mismatches are driven not only by physical detection by each method, but also by the ability to identify a given sequence or larvae to species level which depends on the resolution of the *12S* locus or presence of distinguishable morphological characteristics respectively. And as highlighted by the reviewer, subsampling between the starboard and port bongo net tows can lead to mismatches, particularly for rare, low abundance species. Here no mismatches occurred when there were at least 10 larvae in a given jar (mean 1.61 larvae).

Conclusions should also consider that ALL the surveys were carried out during a small window of spring, so they cannot fully portray the pelagic fish communities, as the samples will be biased by the larvae present in the water column in that period. Across that geographic gradient, it wouldn't be unusual, across years, to have different bouts and peaks of reproductive outputs, from different species, shifted by several weeks (not to mention the different broadcasting modes); so this is a source of bias that should be addressed.

We thank the reviewer for their insightful comments on the phenology of ichthyoplankton spawning (Line 103-107). We chose to focus on ichthyoplankton assemblages collected in spring following the precedent of decades of research in the Southern California Current system which have highlighted the majority of species spawn in spring within the study region (*27*, *53*, *55*, *58*). Furthermore, the annual state of the California Current report similarly relies on Spring ichthyoplankton assemblages. In addition, recent work has highlighted slight shifts in the phenology of spawning across fish species, but that most have stayed relatively stable even during the MHW(*27*). Therefore, we feel confident that focusing on spring ichthyoplankton assemblages provides a robust sampling of fish assemblages in the Southern California Current system.

References

1. M. Miya, Y. Sato, T. Fukunaga, T. Sado, J. Y. Poulsen, K. Sato, T. Minamoto, S. Yamamoto, H. Yamanaka, H. Araki, M. Kondoh, W. Iwasaki, 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 (2015).
2. Z. Gold, E. E. Curd, K. D. Goodwin, E. S. Choi, B. W. Frable, A. R. Thompson, H. J. Walker, R. S. Burton, D. Kacev, L. D. Martz, P. H. Barber, Improving metabarcoding taxonomic assignment: A case study of fishes in a large marine ecosystem. *Mol. Ecol. Resour.* **21**, 2546–2564 (2021).
3. E. E. Curd, Z. Gold, G. S. Kandlikar, J. Gomer, M. Ogden, T. O'Connell, L. Pipes, T. M. Schweizer, L. Rabichow, M. Lin, B. Shi, P. H. Barber, N. Kraft, R. Wayne, R. S. Meyer, Anacapa Toolkit: An environmental DNA toolkit for processing multilocus metabarcode datasets. *Methods Ecol. Evol.* **10**, 1469–1475 (2019).
4. E. Valsecchi, J. Bylemans, S. J. Goodman, R. Lombardi, I. Carr, L. Castellano, A. Galimberti, P. Galli, Novel universal primers for metabarcoding environmental DNA surveys of marine mammals and other marine vertebrates. *Environ. DNA*. **2**, 460–476 (2020).
5. R. A. Collins, G. Trauzzi, K. M. Maltby, T. I. Gibson, F. C. Ratcliffe, J. Hallam, S. Rainbird, J. Maclaine, P. A. Henderson, D. W. Sims, S. Mariani, M. J. Genner, Meta-Fish-Lib: A generalised, dynamic DNA reference library pipeline for metabarcoding of fishes. *J. Fish Biol.* **99**, 1446–1454 (2021).
6. A. Polanco F., E. Richards, B. Flück, A. Valentini, F. Altermatt, S. Brosse, J. C. Walser, D. Eme, V. Marques, S. Manel, C. Albouy, T. Dejean, L. Pellissier, Comparing the performance of 12S mitochondrial primers for fish environmental DNA across ecosystems. *Environ. DNA*. **3**, 1113–1127 (2021).
7. M. Miya, R. O. Gotoh, T. Sado, MiFish metabarcoding: a high-throughput approach for simultaneous detection of multiple fish species from environmental DNA and other samples. *Fish. Sci.* **86**, 939–970 (2020).
8. K. Deiner, H. M. Bik, E. Mächler, M. Seymour, A. Lacoursière-Roussel, F. Altermatt, S. Creer, I. Bista, D. M. Lodge, N. de Vere, M. E. Pfrender, L. Bernatchez, Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. *Mol. Ecol.* **26**, 5872–5895 (2017).
9. R. C. Edgar, Accuracy of taxonomy prediction for 16S rRNA and fungal ITS sequences. *PeerJ*. **2018**, e4652 (2018).

10. P. Taberlet, A. Bonin, L. Zinger, E. Coissac, *Environmental DNA: For biodiversity research and monitoring* (Oxford University Press, 2018).
11. M. A. Min, P. H. Barber, Z. Gold, MiSebastes: An eDNA metabarcoding primer set for rockfishes (genus Sebastes). *Conserv. Genet. Resour.* **13**, 447–456 (2021).
12. M. Leray, J. Y. Yang, C. P. Meyer, S. C. Mills, N. Agudelo, V. Ranwez, J. T. Boehm, R. J. Machida, A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: Application for characterizing coral reef fish gut contents. *Front. Zool.* **10**, 34 (2013).
13. P. A. Hastings, R. S. Burton, Establishing a DNA Sequence database for the marine fish fauna of California, 5 (2008).
14. B. E. Deagle, S. N. Jarman, E. Coissac, F. Pompanon, P. Taberlet, DNA metabarcoding and the cytochrome c oxidase subunit I marker: Not a perfect match. *Biol. Lett.* **10**, 20140562 (2014).
15. H. Krehenwinkel, M. Wolf, J. Y. Lim, A. J. Rominger, W. B. Simison, R. G. Gillespie, Estimating and mitigating amplification bias in qualitative and quantitative arthropod metabarcoding. *Sci. Rep.* **7**, 1–12 (2017).
16. R. Jiang, W. V. Li, J. J. Li, mbImpute: an accurate and robust imputation method for microbiome data. *Genome Biol.* **22**, 1–27 (2021).
17. J. D. Silverman, K. Roche, S. Mukherjee, L. A. David, Naught all zeros in sequence count data are the same. *Comput. Struct. Biotechnol. J.* **18**, 2789–2798 (2020).
18. J. J. Jos´, J. Egozcue, J. Graffelman, M. I. Ortego, V. Pawlowsky-Glahn, Some thoughts on counts in sequencing studies. *NAR Genomics Bioinforma.* **2**, 1–10 (2020).
19. M. Stat, M. J. Huggett, R. Bernasconi, J. D. Dibattista, T. E. Berry, S. J. Newman, E. S. Harvey, M. Bunce, Ecosystem biomonitoring with eDNA: Metabarcoding across the tree of life in a tropical marine environment. *Sci. Rep.* **7**, 12240 (2017).
20. K. Bohmann, V. Elbrecht, C. Carøe, I. Bista, F. Leese, M. Bunce, D. W. Yu, M. Seymour, A. J. Dumbrell, S. Creer, Strategies for sample labelling and library preparation in DNA metabarcoding studies. *Mol. Ecol. Resour.* (2021), doi:10.1111/1755-0998.13512.
21. J. L. O'donnell, R. P. Kelly, N. C. Lowell, J. A. Port, Indexed PCR primers induce template- Specific bias in Large-Scale DNA sequencing studies. *PLoS One*. **11**, e0148698 (2016).
22. D. M. Gohl, P. Vangay, J. Garbe, A. MacLean, A. Hauge, A. Becker, T. J. Gould, J. B. Clayton, T. J. Johnson, R. Hunter, D. Knights, K. B. Beckman, Systematic improvement of amplicon marker gene methods for increased accuracy in microbiome studies. *Nat. Biotechnol.* **34**, 942–949 (2016).
23. C. Yang, K. Bohmann, X. Wang, W. Cai, N. Wales, Z. Ding, S. Gopalakrishnan, D. W. Yu, Biodiversity Soup II: A bulk-sample metabarcoding pipeline emphasizing error reduction. *Methods Ecol. Evol.* **12**, 1252–1264 (2021).
24. M. Leray, N. Knowlton, Random sampling causes the low reproducibility of rare eukaryotic OTUs in Illumina COI metabarcoding. *PeerJ.* **2017**, e3006 (2017).
25. P. E. Smith, H. G. Moser, Long-term trends and variability in the larvae of Pacific sardine and associated fish species of the California Current region. *Deep. Res. Part II Top. Stud. Oceanogr.* **50**, 2519–2536 (2003).
26. J. M. Nielsen, L. A. Rogers, R. D. Brodeur, A. R. Thompson, T. D. Auth, A. L. Deary, J. T. Duffy-Anderson, M. Galbraith, J. A. Koslow, R. I. Perry, Responses of ichthyoplankton assemblages to the recent marine heatwave and previous climate fluctuations in several

Northeast Pacific marine ecosystems. *Glob. Chang. Biol.* **27**, 506–520 (2021).
27. A. R. Thompson, N. J. Ben-Aderet, N. M. Bowlin, D. Kacev, R. Swalethorp, W. Watson, Putting the Pacific marine heatwave into perspective: The response of larval fish off southern California to unprecedented warming in 2014–2016 relative to the previous 65 years. *Glob. Chang. Biol.* **28**, 1766–1785 (2022).
28. E. C. J. Oliver, S. E. Perkins-Kirkpatrick, N. J. Holbrook, N. L. Bindoff, 9. Anthropogenic and natural influences on record 2016 marine heat waves. *Bull. Am. Meteorol. Soc.* **99**, S44–S48 (2018).
29. C. L. Gentemann, M. R. Fewings, M. García-Reyes, Satellite sea surface temperatures along the West Coast of the United States during the 2014–2016 northeast Pacific marine heat wave. *Geophys. Res. Lett.* **44**, 312–319 (2017).
30. A. S. Ren, D. L. Rudnick, Temperature and salinity extremes from 2014-2019 in the California Current System and its source waters. *Commun. Earth Environ.* **2**, 1–9 (2021).
31. E. D. Weber, T. D. Auth, S. Baumann-Pickering, T. R. Baumgartner, E. P. Bjorkstedt, S. J. Bograd, B. J. Burke, J. L. Cadena-Ramírez, E. A. Daly, M. de la Cruz, State of the California Current 2019–2020: Back to the Future With Marine Heatwaves? *Front. Mar. Sci.*, 1081 (2021).
32. S. McClatchie, A. R. Thompson, S. R. Alin, S. Siedlecki, W. Watson, S. J. Bograd, The influence of Pacific Equatorial Water on fish diversity in the southern California Current System. *J. Geophys. Res. Ocean.* **121**, 6121–6136 (2016).
33. T. L. Frölicher, E. M. Fischer, N. Gruber, Marine heatwaves under global warming. *Nature*. **560**, 360–364 (2018).
34. J. T. Fumo, M. L. Carter, R. E. Flick, L. L. Rasmussen, D. L. Rudnick, S. F. Iacobellis, *J. Geophys. Res. Ocean.*, in press, doi:10.1029/2019JC015674.
35. E. C. J. Oliver, M. G. Donat, M. T. Burrows, P. J. Moore, D. A. Smale, L. V. Alexander, J. A. Benthuysen, M. Feng, A. Sen Gupta, A. J. Hobday, N. J. Holbrook, S. E. Perkins-Kirkpatrick, H. A. Scannell, S. C. Straub, T. Wernberg, Longer and more frequent marine heatwaves over the past century. *Nat. Commun.* **9**, 1–12 (2018).
36. E. C. J. Oliver, J. A. Benthuysen, S. Darmaraki, M. G. Donat, A. J. Hobday, N. J. Holbrook, R. W. Schlegel, A. Sen Gupta, Marine Heatwaves. *Ann. Rev. Mar. Sci.* **13**, 313–342 (2021).
37. N. J. Holbrook, H. A. Scannell, A. Sen Gupta, J. A. Benthuysen, M. Feng, E. C. J. Oliver, L. V. Alexander, M. T. Burrows, M. G. Donat, A. J. Hobday, P. J. Moore, S. E. Perkins-Kirkpatrick, D. A. Smale, S. C. Straub, T. Wernberg, A global assessment of marine heatwaves and their drivers. *Nat. Commun.* **10**, 1–13 (2019).
38. K. X. Ruthrof, D. D. Breshears, J. B. Fontaine, R. H. Froend, G. Matusick, J. Kala, B. P. Miller, P. J. Mitchell, S. K. Wilson, M. van Keulen, N. J. Enright, D. J. Law, T. Wernberg, G. E. S. J. Hardy, Subcontinental heat wave triggers terrestrial and marine, multi-taxa responses. *Sci. Rep.* **8**, 1–9 (2018).
39. C. D. Harvell, D. Montecino-Latorre, J. M. Caldwell, J. M. Burt, K. Bosley, A. Keller, S. F. Heron, A. K. Salomon, L. Lee, O. Pontier, C. Pattengill-Semmens, J. K. Gaydos, Disease epidemic and a marine heat wave are associated with the continental-scale collapse of a pivotal predator (Pycnopodia helianthoides). *Sci. Adv.* **5**, eaau7042 (2019).
40. S. McClatchie, J. Gao, E. J. Drenkard, A. R. Thompson, W. Watson, L. Ciannelli, S. J. Bograd, J. T. Thorson, Interannual and Secular Variability of Larvae of Mesopelagic and Forage Fishes in the Southern California Current System. *J. Geophys. Res. Ocean.* **123**,

6277–6295 (2018).
41. J. N. Heine, California Cooperative Oceanic Fisheries Investigations. Reports. **49**, 264 (2008).
42. G. B. Gloor, J. M. Macklaim, V. Pawlowsky-Glahn, J. J. Egozcue, Microbiome datasets are compositional: And this is not optional. *Front. Microbiol.* **8**, 2224 (2017).
43. M. R. McLaren, A. D. Willis, B. J. Callahan, Consistent and correctable bias in metagenomic sequencing experiments. *Elife*. **8**, e46923 (2019).
44. J. D. Silverman, R. J. Bloom, S. Jiang, H. K. Durand, E. Dallow, S. Mukherjee, L. A. David, Measuring and mitigating PCR bias in microbiota datasets. *PLoS Comput. Biol.* **17**, e1009113 (2021).
45. R. P. Kelly, C. J. Closek, J. L. O'Donnell, J. E. Kralj, A. O. Shelton, J. F. Samhouri, Genetic and manual survey methods yield different and complementary views of an ecosystem. *Front. Mar. Sci.* **3**, 283 (2017).
46. A. O. Shelton, R. P. Kelly, J. L. O'Donnell, L. Park, P. Schwenke, C. Greene, R. A. Henderson, E. M. Beamer, Environmental DNA provides quantitative estimates of a threatened salmon species. *Biol. Conserv.* **237**, 383–391 (2019).
47. F. E. Harrell, K. L. Lee, D. B. Mark, Prognostic/Clinical Prediction Models: Multivariable Prognostic Models: Issues in Developing Models, Evaluating Assumptions and Adequacy, and Measuring and Reducing Errors. *Tutorials Biostat. Stat. Methods Clin. Stud.* **1**, 223–249 (2005).
48. J. A. Royle, W. A. Link, Generalized site occupancy models allowing for false positive and false negative errors. *Ecology*. **87**, 835–841 (2006).
49. R. P. Kelly, A. O. Shelton, R. Gallego, Understanding PCR Processes to Draw Meaningful Conclusions from Environmental DNA Studies. *Sci. Rep.* **9**, 1–14 (2019).
50. K. A. Rose, J. Fiechter, E. N. Curchitser, K. Hedstrom, M. Bernal, S. Creekmore, A. Haynie, S. ichi Ito, S. Lluch-Cota, B. A. Megrey, C. A. Edwards, D. Checkley, T. Koslow, S. McClatchie, F. Werner, A. MacCall, V. Agostini, Demonstration of a fully-coupled end-to-end model for small pelagic fish using sardine and anchovy in the California Current. *Prog. Oceanogr.* **138**, 348–380 (2015).
51. W. J. Sydeman, S. Dedman, M. García-Reyes, S. A. Thompson, J. A. Thayer, A. Bakun, A. D. MacCall, Sixty-five years of northern anchovy population studies in the southern California Current: A review and suggestion for sensible management. *ICES J. Mar. Sci.* **77**, 486–499 (2020).
52. J. A. Santora, N. J. Mantua, I. D. Schroeder, J. C. Field, E. L. Hazen, S. J. Bograd, W. J. Sydeman, B. K. Wells, J. Calambokidis, L. Saez, D. Lawson, K. A. Forney, Habitat compression and ecosystem shifts as potential links between marine heatwave and record whale entanglements. *Nat. Commun.* **11**, 1–12 (2020).
53. H. Moser, R. Charter, P. Smith, D. Ambrose, W. Watson, S. Charter, E. Sandknop, Distributional atlas of fish larvae and eggs in the Southern California Bight region: 1951-1998. *Calif. Coop. Ocean. Fish. Investig. Atlas*. **34**, 1–166 (2001).
54. A. R. Thompson, W. Watson, S. McClatchie, E. D. Weber, Multi-scale sampling to evaluate assemblage dynamics in an oceanic marine reserve. *PLoS One*. **7**, e33131 (2012).
55. H. G. Moser P.E. Smith, and L.E. Eber, Larval fish assemblages in the California Current region, 1954-1960, a period of dynamic environmental change. *CalCOFI Rep.* **28**, 97–127 (1987).
56. T. Chambert, D. S. Pilliod, C. S. Goldberg, H. Doi, T. Takahara, An analytical framework

for estimating aquatic species density from environmental DNA. *Ecol. Evol.* **8**, 3468–3477 (2018).
57. R. P. Martin, E. E. Olson, M. G. Girard, W. L. Smith, M. P. Davis, Light in the darkness: new perspective on lanternfish relationships and classification using genomic and morphological data. *Mol. Phylogenet. Evol.* **121**, 71–85 (2018).
58. H. G. Moser, R. L. Charter, P. E. Smith, D. A. Ambrose, S. R. Charter, C. A. Meyer, E. M. Sandknop, W. Watson, *Distributional atlas of fish larvae and eggs in the California Current region: taxa with 1000 or more total larvae, 1951 through 1984* (Marine Life Research Program, Scripps Institution of Oceanography, 1993), vol. 53.