Response to Referee 2

General comments:

This is a valuable and timely contribution for the PACE era. The use of quantitative omics is very much a step in the right direction, in my opinion, and the strong and improved correlations observed between quantitative omics and pigment data are reassuring.

Overall, I found the writing and presentation to be of a high standard and the collection of data over multiple years and seasons is a commendable feat.

The authors have done a good job considering environmental gradients and seasonal change, but it feels like an opportunity was missed by not doing a more granular analysis of surface vs. deep communities if the data are in hand.

I found the ecological application 'case studies' to be a mixed bag. On the one hand, Section 3.2.1 that considers ecological assessments shows a lot of promise of combining omics and pigment data to improve understanding of ecological processes. On the other hand, I'm not persuaded by the suggestion of Section 3.2.2 that quantitative omics, as presented in this manuscript, can improve the monitoring and forecasting of harmful algal blooms. But, I hope the authors can change my mind about this. Finally, Section 3.2.3 suggests that quantitative omics may lead to better biogeochemical and metabolic rate estimates. I think this section could be improved by considering further the substantial caveats and current limitations to this potential application.

Finally, the Methods section is missing important details in places, which are described in the Specific comments below.

> We thank the reviewer for their time and constructive feedback. As the reviewer describes these general comments in greater detail below, we have responded to them in addition to their other comments in line.

Specific comments:

Fig. 1 Figure caption; L129-130: Consider adding "regions" after (18S-V9 (blue)"

> We have modified this text to now read: "Relative abundances of different phytoplankton groups using 18S-V4 **rRNA gene** (red), 18S-V9 **rRNA gene** (blue), or transcript (metaT, green) abundances"

Methods:

L137-140: I understand what the authors are trying to say, but this is an awkwardly worded sentence. Perhaps along the lines of "These data represent only a subset of the on-going NOAA-CalCOFI Ocean Genomics (NCOG) time series and are restricted to samples where quantitative approaches for DNA and RNA were employed concurrently with phytoplankton pigments samples (no DNA samples from 2017 and only RNA samples from 2017-2020; James et al., 2022)."

> We have revised this text as suggested (lines 148-151).

LN150-158: The primary productivity section is sparse and could benefit from additional details.

We agree with the reviewer and have made several revisions detailed below to clarify how the primary productivity measurements were performed.

How were the sampling depths determined? Was there a bio-optics CTD cast to determine the light extinction coefficient prior to sampling for ^{14}-C incubations, and the sampling depths were then chosen to match the degree of attenuation of the neutral-density screens used for the deckboard incubation?

> We now describe the sampling depths and light levels on lines 164-167: "seawater was collected from six depths representing 56%, 30%, 10%, 3%, 1%, and 0.3% surface light levels shortly before local apparent noon. Light levels were estimated with a Secchi disk with the assumptions that the 1% light level is three times the Secchi depth and that the extinction coefficient is constant."
>
> This approach is used to maintain consistency within the time series, as this protocol was implemented in 1984, prior to the CalCOFI program's transition to using a CTD rosette.

What was the specific activity, concentration of radioactivity added to the sample bottles, and supplier of the $NaH^{14}CO_3$ used?

> The specific activity varies by cruise depending on the activity provided by the supplier, MP Biomedicals LLC. We have updated the methods text to be more detailed and include this information on lines 168-170: "Bottles were then inoculated with a 200 µL solution containing $NaH^{14}CO_3$ that was prepared by diluting 50 mL of $NaH^{14}CO_3$ (approximately 50-57 µCi $mmol^{-1}$; MP Biomedicals, LLC) with 350 mL of 2.8 mM Na_2CO_3 and then adjusting the pH to ~10 with 1 N NaOH (Fitzwater et al., 1982)."

I believe 'HA' is a Millipore-specific designation for filter type. I suggest mentioning the material of the filter (i.e., mixed cellulose esters (MCE) membrane) for those reader who are not familiar with them.

> We have modified this text as suggested to now read (lines 171-172): "Following incubation, samples were filtered onto 0.45 µm mixed-cellulose ester filters (type HA, Millipore)"

Was the incubation time a constant between seasons? Multiplying by 1.8 to obtain 24 h productivity implies that the incubation time was ~ 13.3 h, but you state also that incubations were performed between local noon and civil twilight, which presumably varies seasonally.

> The reviewer is correct that the incubation time varies seasonally with the length of daylight. As the incubations occur from local noon to civil twilight, the resulting measurement is half light-day productivity. Previous direct comparisons between half light-day productivity incubations and 24-hour incubations on CalCOFI and other regional cruises were made by Eppley (1992). As shown below in Figures 8 and 9 from this publication, the 24 hour values are approximately 1.8 times the half light-day values.

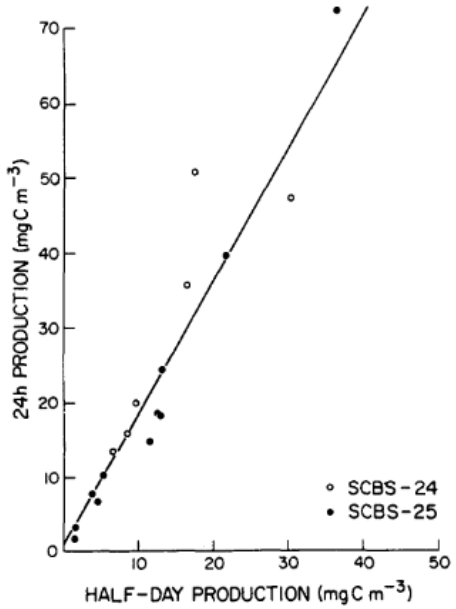

FIG.8. Comparison of half-day photosynthetic production, measured from noon to sunset, with that measured 24 hours on SCBS cruises 24 and 25. The regression line is 24h production = 1.80 (half-day production) + 0.62. The regression explained 91% of the variability, i.e. r^2=0.907.

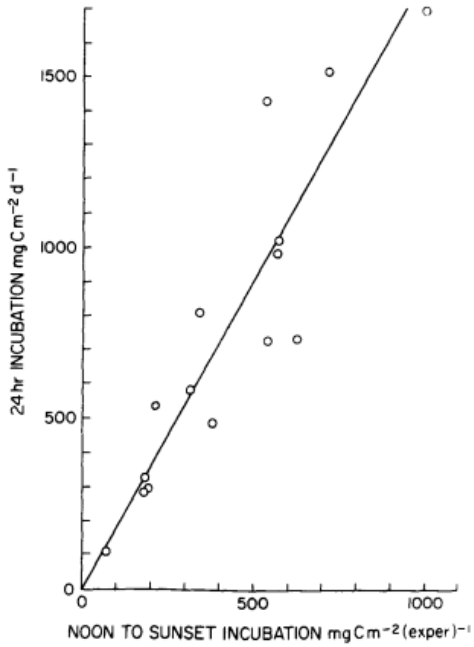

FIG.9. As in Fig.8, except the experiments were done on several CalCOFI cruises and values are per area rather than per volume. The slope of the line is 1.8 from Fig.8.

We have updated this text to clarify this. It now reads on lines 173-176: “As the incubations occurred from local noon until civil twilight, half light-day productivity at each depth was calculated by averaging the two light bottles corrected with the dark uptake bottle. Half light-day productivity was then multiplied by 1.8 to obtain 24 hour productivity as determined by Eppley (1992).”

How closely did the collection depths of productivity and DNA samples match? “Closest” is vague.

> We have updated this text to clarify the differences in depths between productivity and DNA samples (lines 176-178): “When comparing productivity to diversity from DNA, samples from the entire NCOG dataset (2014 to 2020) that were within 20 m of productivity samples were used (*n* = 434). The average vertical distance between DNA and productivity samples was 1.79 m.”

L160-165: Analytical details of the HPLC pigment analyses are entirely missing – you skip directly from sample collection to Phytoclass taxonomic analyses. I suggest adding a citation to the analytical method at the very least.

> We have revised this text to now include details of the HPLC analysis with citations that further detail the method (lines 182-191):
>
> > “Once completed filtering, the filters were carefully folded in half, blotted on a paper towel to remove excess water, and stored in 2 mL cryovials in liquid nitrogen until analysis at the Horn Point Analytical Services Laboratory at the University of Maryland with the HPL method as described in Hooker (2005). Briefly, filters were extracted in 95% acetone and sonicated on ice for 30 s with an output of 40 W. Samples were then clarified by filtering them through a HPLC syringe cartridge filter (0.45 µm) and a glass-fiber prefilter. Extracts were then analyzed with an automated HP 1100 HPLC system with external calibration standards that were either purchased or isolated from naturally occurring sources as described in Van Heukelem and Thomas (2001). The pigments that were measured and used here are Peridinin, 19'-butanoyloxyfucoxanthin, fucoxanthin, neoxanthin, prasinoxanthin, violaxanthin, 19'-hexanoyloxyfucoxanthin, alloxanthin, zeaxanthin, divinyl chlorophyll a, and TChla. The precision including filter extraction and analysis of TChla is estimated to be 5.1%.”

L255: Flow cytometery. It appears that the only flow cytometry data that is presented is for cell abundances for *Prochlorococcus* and *Synechococcus* in Fig. S2. As such, you may wish to consider moving this portion of the methods to the supplementary information. I suggest also that you tailor the description of the methods to explain how *Prochlorococcus* and *Synechococcus* were differentiated and gated.

> Cell abundances from flow cytometry for *Prochlorococcus* are included in Figure 3E and 3F. Since these data are in a main figure, we would prefer to keep the associated methods text in the main Materials and Methods rather than the Supplementary Information. With respect to differentiating *Prochlorococcus* and *Synechococcus*, we also now add a citation for Monger and Landry (1993) which details the flow cytometry methods used here.

L264-266: I’m confused about the use of side- versus forward-scatter. Were forward *and* size-scatter signatures used to estimate the size of *Prochlorococcus* and *Synechococcus*?

> Yes, we now state on lines 301-304: “The optical filter configuration distinguished ***Prochlorococcus*** **and** ***Synechococcus*** **populations** based on chlorophyll a (red fluorescence, 680 nm), phycoerythrin (orange fluorescence, 575 nm), DNA (blue fluorescence, 450 nm), and forward and 90° side-scatter signatures.”

Results and Discussion

LN286: As you're discussing the range here, I suggest reporting the range for Fuco, rather than the maximum concentration, as you don't specify in the previous sentence what pigments were measured at the lowest concentrations (and the log scale used in Fig. 1E makes this difficult to discern by eye).

We have rephrased this text to not only be more specific in reporting the range Fuco and other pigments used here. We now state on lines 327-329:

"Fucoxanthin (Fuco) exhibited the greatest range with concentrations ranging from 0.001 to 6.81 µg L^{-1}. While 19'-hexanoyloxyfucoxanthin (HexFuco) concentrations ranged from 0.005 to 1.170 µg L^{-1}, all other pigment concentrations examined here were always less than 0.697 µg L^{-1}."

L364-372: The authors state that samples were collected from both the near-surface and SCML. It's reasonable to assume that the two communities may differ systematically in terms of both community composition and photophysiology, as the authors discuss. Were samples from these two depths pooled for all analyses? If so, how did the sample numbers differ between the two depths, potentially skewing the aggregated results?

Yes, the samples from all depths were pooled for analysis, reflecting a wide range of conditions and light histories. To clarify that there were similar sample numbers among depth categories, we now state in the Methods on lines 158-159: "For the DNA samples, 219 were from the near-surface (0 – 14 m, mean = 10 m), and 198 were from the SCML (18 – 130 m, mean = 53 m)."

This added text also clarifies that the sampling depths for the SCMLs are wide ranging, since SCMLs in the region generally become deeper from the nearshore environment to the offshore environment along with the nitracline. Considering the amount of variability captured by pooling all samples, we believe that the strength of the correlations is a strength of our analysis; however, we agree with the reviewer that more attention should be paid to surface versus deeper communities as detailed in response to the next comment.

The authors have done a good job considering environmental gradients and seasonal change, but it seems like a missed opportunity to not extend these analyses to surface vs. deep communities if the data are in hand. Have the authors performed the regressions shown in Fig. 2 on low- vs high-light binned samples? Even if they are do not differ from each other, this is useful information and could be included in the supplementary information.

As suggested, we have examined the relationships between pigments and DNA-based abundances for surface and SCML communities separately (Figs. S6 and S9). We also describe these results on lines 417-425:

"To further examine the effects of light and depth on these relationships, separate correlations were performed with the absolute abundances from near-surface (≤ 14 m) or SCML (≥ 18 m) samples. For chlorophytes, cryptophytes, diatoms, dinoflagellates, and prymnesiophytes, the strength of the correlations were similar or higher for both depth categories when separated compared to all samples combined (Figs. 2 and S6). Only pelagophyte correlations were consistently lower when separated, albeit the differences were relatively minor (r = 0.65 – 0.70 versus r = 0.74 - 0.75). Linear regressions also displayed similar results between depth categories,

except for chlorophytes and prymnesiophytes, where the slope of the regressions were 46-49% lower with SCML samples. In these cases, pigment concentrations exhibited a reduced range, where concentrations were elevated wat lower DNA abundances but reached similar concentrations at higher DNA abundances, indicating that pigment concentrations for these taxa are elevated under lower abundance regimes within the SCML."

And on lines 465-467:

"Although separating near-surface and SCML samples showed stronger correlations than all samples together except for all cyanobacteria and Zea (r = 0.55), the relationships were still weaker than with relative abundances (Fig. S9)."

Section 3.2.1 This was a useful illustrative example of how to combine the strengths of 'classical' HPLC data and DNA-based approaches to better understand phytoplankton ecology. I enjoyed reading it.

Thank you.

L500-509: The use of recycled nutrients (a low f-ratio in the case of nitrogen, a low Fe-ratio in the case or iron) should be included as a general strategy.

We agree and have now added a sentence to mention recycled nutrients and nitrogen fixation as means to sustain phytoplankton under oligotrophic conditions. This section now reads (lines 565-569):

"The low productivity and diversity end of unimodal distributions have also been attributed to selective grazing with the dominance of a few slow-growing nutrient specialists (Vallina et al., 2014). As diversity and richness instead remained high, many phytoplankton taxa, particularly dinoflagellates, appear to coexist within low productivity regimes. With low nutrient availability as inferred by deeper nitracline depths (Fig. 4D), diverse phytoplankton taxa within these regimes may be sustained by recycled nutrients including nitrogen and iron as well as nitrogen fixation (Boyd et al., 2017; Zehr and Ward, 2002)."

Section 3.2.2 This section was not nearly as compelling as the previous one. In effect, you argue that if fucoxanthin is detected then there are likely to be *some Pseudo-nitzschia* present, *some* of which *may*produce DA. It's also quite possible for fucoxanthin concentrations to vary independently of *Pseudo-nitzschia*abundance. It's not clear how this improves the current state of knowledge or improves the monitoring and forecasting of harmful algal blooms, even within an intensively sampled region like the CalCOFI survey site.

To clarify how the results of our analysis suggest that Fuco detection may aid HAB monitoring, we now describe that current HAB monitoring for *Pseudo-nitzschia* blooms in the California Current relies on remotely-sensed chlorophyll *a* and reflectance from older sensors. Since Fuco offers greater specificity for diatoms, and if *Pseudo-nitzschia* is regularly a dominant diatom, substituting chlorophyll *a* with Fuco in the model may improve accuracy. This is now stated on lines 601-605:

"In the California Current region, *Pseudo-nitzschia* abundances and DA production are forecasted by the California Harmful Algae Risk Mapping (C-HARM) system, which uses satellite remote-sensing data and a regional ocean circulation model. Specifically, the remote-sensing data used as inputs are chlorophyll *a* concentrations and reflectance at 488 and 555 nm from the S-NPP

NOAA VIIRS instrument. As Fuco offers greater specificity for diatoms, substituting Fuco for chlorophyll *a* may improve model predictions, particularly if *Pseudo-nitzschia* is a dominant diatom overall."

As we further describe in this section, *Pseudo-nitzschia* is indeed one of the most dominant diatoms, and Fuco concentrations were better predictors of *Pseudo-nitzschia* than TChla (Fig. 4). Moreover, increased Fuco concentrations generally aligned with increased expression of a critical domoic acid biosynthesis gene, *dabA*. As a result, we believe that we have shown evidence that Fuco detection may support improvements to HAB forecasts in this region, and potentially others where *Pseudo-nitzschia* is a dominant diatom; however, we recognize that there are important caveats that must be considered, such as the prerequisite that *Pseudo-nitzschia* regularly is a dominant diatom taxa. We also recognize that potential improvements to C-HARM forecasts in the California Current for *Pseudo-nitzschia* or forecasts in other regions would require additional validation. We now describe these caveats on:

- Lines 621-624: "Pending the development and implementation of models for remotely-sensed Fuco concentrations (Kramer et al., 2022), such potential improvements for *Pseudo-nitzschia* forecasts will require validation with *in situ* measurements. Moreover, the utility of remotely-sensed Fuco concentrations for *Pseudo-nitzschia* HAB monitoring would only apply to other regions where *Pseudo-nitzschia* is a dominant diatom."
- Lines 729-730: "for both HAB forecasts and the inference of phytoplankton activity, significant additional validation will be required."

Section 3.2.3 I appreciated the ambition and hope of this section. One thing to note about L587-591: the modules or subsystems of genes that predict a reaction rate are likely to be different between species. As such, it would seem to me that to make use of this correlation, not only would you need to know about abundance and expressed metabolism, but you would need to have a comparable degree of knowledge of your target species as we have for baker's yeast – an extensively studied model organism – to know which gene clusters predict rates. Would this not require intensive lab rate measurements to validate this correlation? Also, the baker's yeast correlation was achieved under steady-state conditions. Trying to accomplish this in a field study? Yikes.

We certainly agree with the reviewer that this section is highly ambitious and has important caveats that must be described. We also agree that extensive additional measurements would be required to connect pigments or transcript abundances to any rates. Our intention in highlighting the study by McCain et al. (2025), which draws relationships between proteins and rates in baker's yeast, is to introduce the concept that proteins or groups of proteins may relate to rates, rather than stating that we have demonstrated any such relationship with pigments or transcripts. We now clarify these caveats on lines 647-649, lines 679-683, and 729-730:

"The absolute quantities of certain proteins have shown promise for inferring rates of nitrite oxidation and carbon fixation (Saito et al., 2020; Roberts et al., 2024), although it is unclear if absolute transcript abundances will be able to serve a similar purpose."

"Although these relatively strong correlations between pigments and transcripts indicates that this application has potential use to infer activity, direct relationships with rates remain to be

demonstrated, and establishing these relationships would require extensive additional validation with field-based studies that integrate these measurements."

"for both HAB forecasts and the inference of phytoplankton activity, significant additional validation will be required."

We believe that with this added text, we have been cautious to not overstate our claims and are upfront about the speculative nature of this discussion. Despite this high uncertainty, we still believe that it is useful to present these relationships and describe these concepts with the data in hand, even if it is ultimately shown in future studies that these connections cannot be made.

Technical corrections:

Fig. 2, column B: Y-axis number format is sometimes in scientific format, sometimes not. I suggest keeping this consistent amongst panels.

As suggested, we have modified this figure to have consistent y-axis labeling.

L481: "The observations of unimodal PDRs have led to hypotheses for the mechanisms that underlying them." Remove 'that' before 'underlying'.

"That" has been removed.

References

Boyd, P. W., Ellwood, M. J., Tagliabue, A., and Twining, B. S.: Biotic and abiotic retention, recycling and remineralization of metals in the ocean, Nat Geosci, 10, 167, 10.1038/ngeo2876, 2017.

Eppley, R. W.: Chlorophyll, photosynthesis and new production in the Southern California Bight, Prog Oceanogr, 30, 117–150, https://doi.org/10.1016/0079-6611(92)90010-W, 1992.

Fitzwater, S. E., Knauer, G. A., and Martin, J. H.: Metal contamination and its effect on primary production measurements1, Limnology and Oceanography, 27, 544–551, 10.4319/lo.1982.27.3.0544, 1982.

Hooker, S. B.: The second SeaWiFS HPLC analysis round-robin experiment (SeaHARRE-2), National Aeronautics and Space Administration, Goddard Space Flight Center2005.

Kramer, S. J., Siegel, D. A., Maritorena, S., and Catlett, D.: Modeling surface ocean phytoplankton pigments from hyperspectral remote sensing reflectance on global scales, Remote Sens Environ, 270, 112879, https://doi.org/10.1016/j.rse.2021.112879, 2022.

McCain, J. S. P., Britten, G. L., Hackett, S. R., Follows, M. J., and Li, G.-W.: Microbial reaction rate estimation using proteins and proteomes, ISME J, wraf018, 10.1093/ismejo/wraf018, 2025.

Monger, B. C. and Landry, M. R.: Flow Cytometric Analysis of Marine Bacteria with Hoechst 33342, Appl Enviro Microbiol, 59, 905–911, doi:10.1128/aem.59.3.905-911.1993, 1993.

Roberts, M. E., Bhatia, M. P., Rowland, E., White, P. L., Waterman, S., Cavaco, M. A., Williams, P., Young, J. N., Spence, J. S., Tremblay, J.-É., Montero-Serrano, J.-C., and Bertrand, E. M.: Rubisco in high Arctic tidewater glacier-marine systems: A new window into phytoplankton dynamics, Limnol Oceanogr, 69, 802–817, https://doi.org/10.1002/lno.12525, 2024.

Saito, M. A., McIlvin, M. R., Moran, D. M., Santoro, A. E., Dupont, C. L., Rafter, P. A., Saunders, J. K., Kaul, D., Lamborg, C. H., and Westley, M.: Abundant nitrite-oxidizing metalloenzymes in the mesopelagic zone of the tropical Pacific Ocean, Nat Geosci, 13, 355–362, 2020.

Vallina, S. M., Follows, M. J., Dutkiewicz, S., Montoya, J. M., Cermeno, P., and Loreau, M.: Global relationship between phytoplankton diversity and productivity in the ocean, Nat Commun, 5, 4299, 10.1038/ncomms5299, 2014.

Van Heukelem, L. and Thomas, C. S.: Computer-assisted high-performance liquid chromatography method development with applications to the isolation and analysis of phytoplankton pigments, Journal of Chromatography A, 910, 31–49, 2001.

Zehr, J. P. and Ward, B. B.: Nitrogen Cycling in the Ocean: New Perspectives on Processes and Paradigms, Applied and Environmental Microbiology, 68, 1015–1024, doi:10.1128/AEM.68.3.1015-1024.2002, 2002.