J. Phycol. **58,** 829–833 (2022)

DOI: 10.1111/jpy.13289

NOTE

EMPIRICAL RELATIONSHIP BETWEEN *NIF*H GENE ABUNDANCE AND DIAZOTROPH CELL CONCENTRATION IN THE NORTH PACIFIC SUBTROPICAL GYRE[1]

Mary R. Gradoville[2,3]

Ocean Sciences Department, University of California Santa Cruz, Santa Cruz, California 95064, USA

Mathilde Dugenne

Department of Oceanography, University of Hawai'i at Mānoa, Honolulu, Hawaii, 96822, USA
Daniel K. Inouye Center for Microbial Oceanography: Research and Education, University of Hawai'i at Mānoa, Honolulu, Hawaii, 96822, USA

Annette M. Hynes

School of Oceanography, University of Washington, Seattle, Washington 98195, USA

Jonathan P. Zehr

Ocean Sciences Department, University of California Santa Cruz, Santa Cruz, California 95064, USA

and Angelicque E. White

Department of Oceanography, University of Hawai'i at Mānoa, Honolulu, Hawaii, 96822, USA
Daniel K. Inouye Center for Microbial Oceanography: Research and Education, University of Hawai'i at Mānoa, Honolulu, Hawaii, 96822, USA

Cyanobacterial N_2-fixing microorganisms (diazotrophs) play a critical role in nitrogen and carbon cycling in the oceans; hence, accurate measurements of diazotroph abundance are imperative for understanding ocean biogeochemistry. Marine diazotroph abundances are often assessed using qPCR of the *nif*H gene, a sensitive, taxa-specific, and time/cost-efficient method. However, the validity of *nif*H abundance as a proxy for cell concentration has recently been questioned. Here, we compare *nif*H gene abundances to cell counts for four diazotroph taxa (*Trichodesmium*, *Crocosphaera*, *Richelia*, and *Calothrix*) on two cruises to the North Pacific Subtropical Gyre, one of the largest habitats for marine diazotrophs. *nif*H:cell relationships were strong and significant for *Crocosphaera*, *Richelia*, and *Calothrix* (*nif*H:cell 1.51–2.58; R^2 = 0.89–0.96) but were not significant for *Trichodesmium*, despite previous studies reporting significant *nif*H:cell relationships for this organism. Limited available data suggest that empirical *nif*H:cell can vary among studies but that relationships are usually significantly linear and >1:1. Our study indicates that *nif*H gene abundance, while not a direct measure of cells, is a useful quantitative proxy for diazotroph abundance.

Key index words: **cyanobacteria; ddPCR; diazotroph; *nif*H; polyploidy; qPCR**

***Abbreviations*: CTD, conductivity temperature depth; DDAs, diazotroph-diatom associations; ddPCR, digital droplet PCR; IFCb, Imaging FlowCytobot; NCDs, non-cyanobacterial diazotrophs; NPSG, North Pacific Subtropical Gyre; UCYN-A, unicellular cyanobacterial group A**

Diazotrophs convert otherwise inert N_2 gas into biologically available ammonia, fueling new production and export in nitrogen-limited surface ocean waters (Karl et al. 1997). Marine diazotrophs are comprised of cyanobacterial taxa, including filamentous *Trichodesmium*, heterocyst-forming diazotroph-diatom associations (DDAs), and unicellular groups such as *Crocosphaera* and UCYN-A, as well as non-cyanobacterial diazotrophs (NCDs), whose contribution to marine N_2 fixation is uncertain (Zehr and Capone 2020).

Efforts to measure and predict marine diazotroph abundances have used both cell- and gene-based approaches (e.g., Luo et al. 2012, Tang and Cassar 2019, Pierella Karlusich et al. 2021), which have

[1]Received 17 June 2022. Accepted 24 September 2022.
[2]Present address: Columbia River Inter-Tribal Fish Commission, Portland, Oregon 97232, USA
[3]Author for correspondence: e-mail rgradoville@critfc.org
Editorial Responsibility: B. Palenik (Associate Editor)

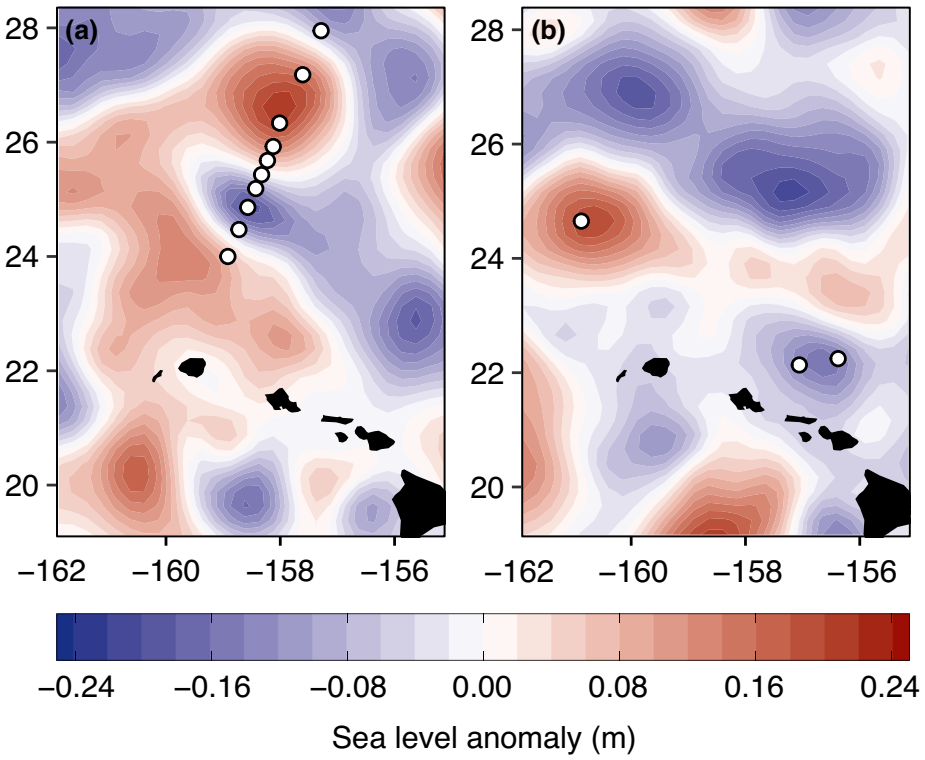

FIG. 1. Sampling locations during the 2017 (a) and 2018 (b) cruises. Locations are superimposed onto satellite sea level anomaly data from June 30, 2017 and April 03, 2018, which were retrieved from the Copernicus Marine and Environment Monitoring Service (http://www.marine.copernicus.eu) and corrected according to Barone et al. (2019).

different advantages and limitations. Cell-based methods, including light and epifluorescence microscopy (White et al. 2018), continuous plankton recording (Olson et al. 2015), and automated flow cytometry (Wilson et al. 2017, Dugenne et al. 2020), have the advantage of directly counting diazotroph cells or filaments. Unfortunately, manual microscope counts are laborious, automated imaging methods have size and volume constraints that exclude some taxa and/or life-stages, and these methods cannot enumerate diazotrophs lacking known phenotypic or fluorometric signatures, such as NCDs. Alternatively, diazotrophs can be quantified using molecular approaches targeting the *nif*H gene, which encodes the iron protein of the nitrogenase enzyme used for N_2 fixation. *nif*H-based approaches have the advantage of being sensitive and phenotype-independent, as well as being relatively high-throughput. The relative ease and affordability of *nif*H qPCR and digital droplet PCR (ddPCR) has led to the increasing use of these techniques, with >10,000 marine measurements to date (10.5281/zenodo.6537451; Tang and Cassar 2019).

Despite the widespread use of *nif*H qPCR/ddPCR measurements, their validity as a proxy for diazotroph abundance has been questioned (Meiler et al. 2022). Diazotroph *nif*H:cell relationships can be confounded by methodological biases in either measurement; for *nif*H, this includes incomplete DNA extraction, suboptimal primer specificity, and uncertainties associated with qPCR, as well as true biological variability (e.g., polyploidy). Empirical gene:cell relationships for field populations of *Trichodesmium* have ranged from <1 to >100 (Hynes 2009, Sargent et al. 2016, White et al. 2018). *Trichodesmium* gene copies per cell also differ between field populations and laboratory cultures and appear to vary among isolates (Hynes 2009, Sargent et al. 2016), though relationships in individual studies have been significantly linear. Unfortunately, *nif*H gene and cell abundances of natural diazotroph assemblages are rarely compared and concerns over potential uncertainties in *nif*H:cell conversion factors have led to suggestions that *nif*H be used merely as a presence/absence metric (Meiler et al. 2022).

Here, we compare *nif*H gene abundances to cell counts for four diazotroph taxa (*Trichodesmium*, *Crocosphaera*, and the DDAs *Richelia* and *Calothrix*) sampled from the surface mixed layer during two cruises to the North Pacific Subtropical Gyre (NPSG). Samples were collected between two mesoscale eddies June–July 2017 (KM1709; Dugenne et al. 2020) and from eddy centers March–April 2018 (FK180310; Gradoville et al. 2021), expeditions which each spanned >200 km (Fig. 1). While limited, this reflects the most geographically extensive field comparison of *nif*H:cell among taxa to date. DNA samples were collected from 15 m (2017) and

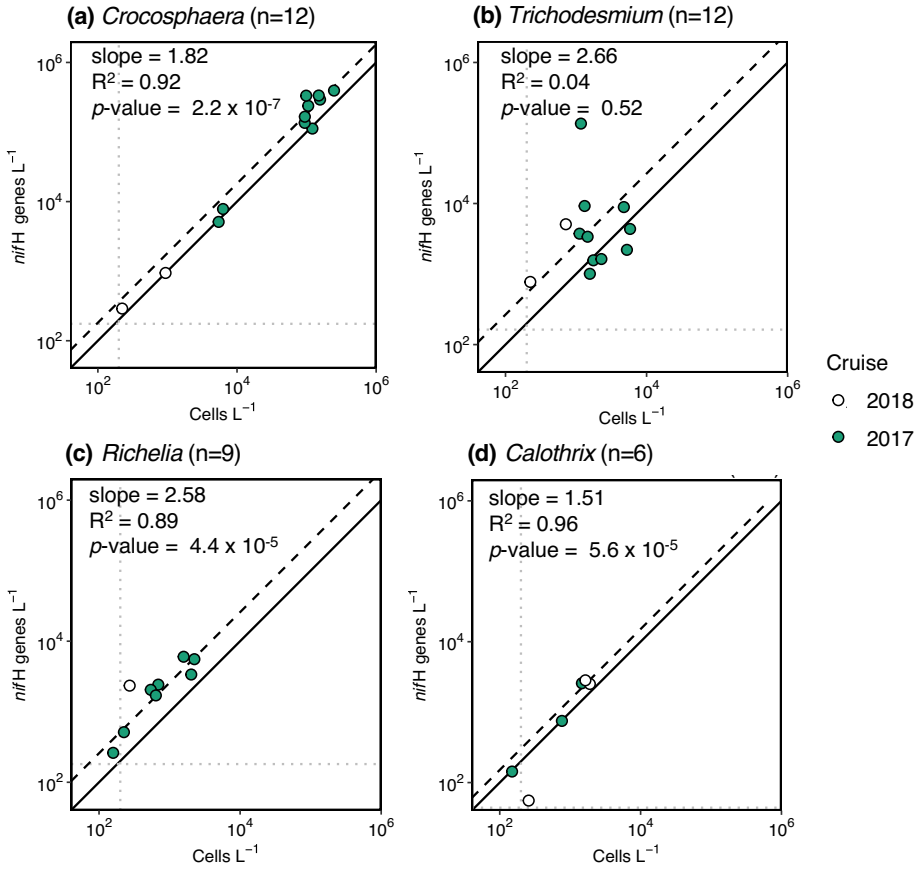

FIG. 2. *nif*H gene abundances versus diazotroph cell concentrations on the 2017 and 2018 cruises. Solid lines show a 1:1 relationship; dashed lines show the slope from a simple linear regression model with a fixed zero intercept. Regression statistics are provided for each subplot. Gray dotted lines reflect detection limits, corresponding to 50 cells · L^{-1} for 20 mL binned samples (~200 cells · L^{-1}) for IFCb measurements and 44–181 *nif*H genes · L^{-1} for ddPCR measurements (Gradoville et al. 2021). Note that data are plotted on a logarithmic scale but regressions were performed using untransformed data. [Color figure can be viewed at wileyonlinelibrary.com]

5 m (2018) by sub-sampling 2–4 L of seawater from Niskin® sampling bottles. Sample preservation, DNA extraction, and *nif*H ddPCR were performed as previously described (Gradoville et al. 2021). Diazotrophs were also enumerated from the ship's uncontaminated underway system (~7 m intake depth) using two types of autonomous flow cytometry: an Imaging FlowCytobot (IFCb, McLane) for taxa >4 μm (*Trichodesmium*, DDAs, and large cell-size *Crocosphaera*; Dugenne et al. 2020) and an autonomous SeaFlow flow cytometer for small cell-size *Crocosphaera* (2–4 μm; Ribalet et al. 2019), at 20 and 3 min intervals, respectively. For filamentous diazotrophs, filaments were enumerated by the IFCb and the number of cells in each individual filament was counted manually using IFCb images. To increase the counting accuracy of rare diazotrophs, data from within 2 h of DNA sample collection were binned to 1.5–2 h measurement windows, representing a cumulative volume of >20 mL from an average of 5.5 ± 1.5 individual IFCb samples. The two *Richelia* ddPCR assays (Het-1 and Het-2) were summed for our comparisons. All abundance data are available at https://doi.org/10.5281/zenodo.6342202.

Diazotroph taxa were detected by ddPCR and autonomous flow cytometry in 39/52 paired measurements. There were strong, significant relationships between *nif*H and cell abundances of *Crocosphaera* and DDAs, with R^2 values from linear regressions ranging from 0.89 to 0.96, while the relationship for *Trichodesmium* was not significant (Fig. 2). *nif*H:cell differed slightly among taxa, with slopes ranging from 1.51 to 2.66 (Fig. 2).

Previous data on in situ *nif*H:cell relationships are sparse. In the most rigorous assessment we are aware of, White et al. (2018) compared cell counts and *nif*H genes of *Trichodesmium* and DDAs at three depths over 2 years of repeat Hawaii Ocean Time-series cruises in the NPSG (51 sample comparisons; White et al. 2018). We reanalyzed these data using the same regression methods as in the present study (Fig. S1 in the Supporting Information) and found that White et al. observed much higher *nif*H:cell (116–971), with a stronger relationship for

Trichodesmium (R^2 = 0.50) and weaker relationships for *Richelia* and *Calothrix* (R^2 = 0.35–0.40).

These contrasting results may be at least partially driven by methodological choices. For cell counts, White et al. (2018) used microscopy after filtering 10 L of seawater onto a 10 μm polycarbonate membrane filter, which may have allowed some filaments to pass through. Since *nif*H measurements were made using 0.2 μm filters, *nif*H:cell might have been artificially inflated. Additionally, DDA counts only included heterocysts, not vegetative cells, which as White et al. (2018) discussed, likely inflated *nif*H:cell by a factor of ~3–5. In our study, *Trichodesmium* cells may have been underestimated due to a 100 μm pre-filter used before IFCb counts, which excluded colonies and perhaps some large filaments, and because the field of view of the IFCb camera prevented accurate counts of filaments longer than 425 μm (representing ~12% of the trichomes counted by microscopy in 2017). In a small comparison on the 2017 cruise, average *Trichodesmium* IFCb counts were 19% lower than microscopy counts using the methods of White et al. (2018; n = 3, data not shown), suggesting that a correction factor of ×1.2 (not applied in the present study) could improve the accuracy of *Trichodesmium* counts via the IFCb. These methodological limitations, as well as the small IFCb sample volume (20 mL), may have contributed to the non-significant *nif*H:cell relationship we observed for *Trichodesmium*, which is well known to be heterogeneous in distribution.

Sampling methodology and *nif*H may have also driven *nif*H:cell differences between studies. Our study used ddPCR, which has greater precision and reproducibility than previous qPCR measurements (Hindson et al. 2013). Observed *nif*H abundances might also vary due to different DNA extraction efficiencies or shifts in community composition influencing primer binding sites. Finally, sampling techniques and ocean patchiness may have influenced *nif*H:cell relationships. The large filtration volumes used by White et al. (2018) surely improved the *nif*H:cell regression fit for *Trichodesmium*, a notoriously patchy organism. In contrast, we observed stronger fits for *Richelia* and *Calothrix*, possibly because our sampling strategy temporally linked cell counts and DNA collection (despite the two sample types being collected from different depths within the mixed layer), while in White et al. (2018), *nif*H and microscopy samples were collected from separate CTD casts with a larger temporal offset. White et al. (2018) did not assess *Crocosphaera*, but our observed *nif*H:cell is ~half of the value reported from a single sample in another NPSG study (3.6; Wilson et al. 2017).

Regardless of whether the different *nif*H:cell relationships observed here and previously were driven by methodological biases or true biological differences, empirical relationships have all been statistically significant, except for *Trichodesmium* in the present study. These significant relationships validate *nif*H as a quantitative abundance proxy. Observed *nif*H:cell likely varies among laboratory groups using different methodologies and possibly also regionally from environmental effects on ploidy (Sargent et al. 2016). The caveats described above should be considered carefully if using *nif*H data to infer cell abundance, just as caveats are considered when using chlorophyll data to infer phytoplankton biomass (Huot et al. 2007).

Ultimately, our limited data suggest that *nif*H is a useful yet imperfect abundance proxy. Since empirical *nif*H:cell is not 1:1, direct cell counts may be preferable when possible and practical. However, not all diazotrophs can be counted using traditional microscopy—this includes UCYN-A, undescribed cyanobacterial diazotrophs (new species continue to be discovered; Schvarcz et al. 2022), and NCDs (Moisander et al. 2017). A recent report of non-N_2-fixing *Trichodesmium* (Delmont 2021) also highlights the need for *nif*H-based methods to complement cell abundance data when the genetic potential for N_2 fixation is in question. We hope that future studies report *nif*H:cell and explore the mechanisms controlling this ratio.

This work was funded by the Simons Foundation (Award # 721256 to AEW and 724220 to JPZ) and expedition funding from the Schmidt Ocean Institute for R/V *Falkor* Cruise FK180310. AMH acknowledges the support of research funds from the Simons Foundation. We thank chief scientists Tara Clemente and Steve Poulos, as well as Katie Watkins-Brandt for support with the IFCb on the 2017 expedition. The authors declare no conflict of interest.

Barone, B., Coenen, A. R., Beckett, S. J., McGillicuddy, D. J., Weitz, J. S. & Karl, D. M. 2019. The ecological and biogeochemical state of the North Pacific Subtropical Gyre is linked to sea surface height. *J. Mar. Res.* 77:215–45.

Delmont, T. O. 2021. Discovery of nondiazotrophic *Trichodesmium* species abundant and widespread in the open ocean. *Proc. Natl. Acad. Sci. U.S.A.* 118:e2112355118.

Dugenne, M., Henderikx Freitas, F., Wilson, S. T., Karl, D. M. & White, A. E. 2020. Life and death of *Crocosphaera* sp. in the Pacific Ocean: Fine scale predator–prey dynamics. *Limnol. Oceanogr.* 65:2603–17.

Gradoville, M. R., Cabello, A. M., Wilson, S. T., Turk-Kubo, K. A., Karl, D. M. & Zehr, J. P. 2021. Light and depth dependency of nitrogen fixation by the non-photosynthetic, symbiotic cyanobacterium UCYN-A. *Environ. Microbiol.* 23:4518–31.

Hindson, C. M., Chevillet, J. R., Briggs, H. A., Gallichotte, E. N., Ruf, I. K., Hindson, B. J., Vessella, R. L. & Tewari, M. 2013. Absolute quantification by droplet digital PCR versus analog real-time PCR. *Nat. Methods* 10:1003–5.

Huot, Y., Babin, M., Bruyant, F., Grob, C., Twardowski, M. & Claustre, H. 2007. Does chlorophyll a provide the best index of phytoplankton biomass for primary productivity studies? *Biogeosci. Discuss.* 4:707–45.

Hynes, A. M. 2009. Diversity of the marine cyanobacterium *Trichodesmium*: characterization of the Woods Hole culture collection and quantification of field populations. Doctoral dissertation, Massachusetts Institute of Technology.

Karl, D., Letelier, R., Tupas, L., Dore, J., Christian, J. & Hebel, D. 1997. The role of nitrogen fixation in biogeochemical cycling in the subtropical North Pacific Ocean. *Nature* 388:533–8.

Luo, Y. W., Doney, S., Anderson, L., Benavides, M., Berman-Frank, I., Bode, A., Bonnet, S., Boström, K., Böttjer, D. & Capone, D. 2012. Database of diazotrophs in global ocean: abundances, biomass and nitrogen fixation rates. *Earth Syst. Sci. Data* 4:47–73.

Meiler, S., Britten, G. L., Dutkiewicz, S., Gradoville, M. R., Moisander, P. H., Jahn, O. & Follows, M. J. 2022. Constraining uncertainties of diazotroph biogeography from *nif*H gene abundance. *Limnol. Oceanogr.* 67:816–29.

Moisander, P. H., Benavides, M., Bonnet, S., Berman-Frank, I., White, A. E. & Riemann, L. 2017. Chasing after non-cyanobacterial nitrogen fixation in marine pelagic environments. *Front. Microbiol.* 8:1736.

Olson, E. M., McGillicuddy, D. J., Flierl, G. R., Davis, C. S., Dyhrman, S. T. & Waterbury, J. B. 2015. Mesoscale eddies and *Trichodesmium* spp. distributions in the southwestern North Atlantic. *J. Geophys. Res. Oceans* 120:4129–50.

Pierella Karlusich, J. J., Pelletier, E., Lombard, F., Carsique, M., Dvorak, E., Colin, S., Picheral, M., Cornejo-Castillo, F. M., Acinas, S. G. & Pepperkok, R. 2021. Global distribution patterns of marine nitrogen-fixers by imaging and molecular methods. *Nat. Commun.* 12:1–18.

Ribalet, F., Berthiaume, C., Hynes, A., Swalwell, J., Carlson, M., Clayton, S., Hennon, G., Poirier, C., Shimabukuro, E., White, A. & Armbrust, E. V. 2019. SeaFlow data v1, high-resolution abundance, size and biomass of small phytoplankton in the North Pacific. *Sci. Data* 6:1–8.

Sargent, E. C., Hitchcock, A., Johansson, S. A., Langlois, R., Moore, C. M., LaRoche, J., Poulton, A. J. & Bibby, T. S. 2016. Evidence for polyploidy in the globally important diazotroph *Trichodesmium*. *FEMS Microbiol. Lett.* 363:fnw244.

Schvarcz, C. R., Wilson, S. T., Caffin, M., Stancheva, R., Li, Q., Turk-Kubo, K. A., White, A. E., Karl, D. M., Zehr, J. P. & Steward, G. F. 2022. Overlooked and widespread pennate diatom-diazotroph symbioses in the sea. *Nat. Commun.* 13: 1–9.

Tang, W. & Cassar, N. 2019. Data-driven modeling of the distribution of diazotrophs in the global ocean. *Geophys. Res. Lett.* 46:12258–69.

White, A. E., Watkins-Brandt, K. S. & Church, M. J. 2018. Temporal variability of *Trichodesmium* spp. and diatom-diazotroph assemblages in the North Pacific Subtropical Gyre. *Front. Mar. Sci.* 5:27.

Wilson, S. T., Aylward, F. O., Ribalet, F., Barone, B., Casey, J. R., Connell, P. E., Eppley, J. M., Ferrón, S., Fitzsimmons, J. N. & Hayes, C. T. 2017. Coordinated regulation of growth, activity and transcription in natural populations of the unicellular nitrogen-fixing cyanobacterium *Crocosphaera*. *Nat. Microbiol.* 2:17118.

Zehr, J. P. & Capone, D. G. 2020. Changing perspectives in marine nitrogen fixation. *Science* 368:eaay9514.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web site:

Figure S1. *nif*H gene abundances versus cell concentrations observed by White et al. (2018). In the original publication, the average *nif*H:cell within the upper 45 m were presented and *Richelia* and *Calothrix* heterocyst abundances (Het-1 + Het-2 + Het-3) were summed. Here, we show data separately for each genus and depth. Regressions were performed using the same method as in the current study (simple linear regression with a fixed zero intercept using all data for which cell and *nif*H abundances were above detection limits).