LETTER

doi:10.1038/nature14488

Interaction and signalling between a cosmopolitan phytoplankton and associated bacteria

S. A. Amin[1,2], L. R. Hmelo[3], H. M. van Tol[1], B. P. Durham[4], L. T. Carlson[1], K. R. Heal[1], R. L. Morales[1], C. T. Berthiaume[1], M. S. Parker[1], B. Djunaedi[1], A. E. Ingalls[1], M. R. Parsek[3], M. A. Moran[5] & E. V. Armbrust[1]

Interactions between primary producers and bacteria impact the physiology of both partners, alter the chemistry of their environment, and shape ecosystem diversity[1,2]. In marine ecosystems, these interactions are difficult to study partly because the major photosynthetic organisms are microscopic, unicellular phytoplankton[3]. Coastal phytoplankton communities are dominated by diatoms, which generate approximately 40% of marine primary production and form the base of many marine food webs[4]. Diatoms co-occur with specific bacterial taxa[3], but the mechanisms of potential interactions are mostly unknown. Here we tease apart a bacterial consortium associated with a globally distributed diatom and find that a *Sulfitobacter* species promotes diatom cell division via secretion of the hormone indole-3-acetic acid, synthesized by the bacterium using both diatom-secreted and endogenous tryptophan. Indole-3-acetic acid and tryptophan serve as signalling molecules that are part of a complex exchange of nutrients, including diatom-excreted organosulfur molecules and bacterial-excreted ammonia. The potential prevalence of this mode of signalling in the oceans is corroborated by metabolite and metatranscriptome analyses that show widespread indole-3-acetic acid production by *Sulfitobacter*-related bacteria, particularly in coastal environments. Our study expands on the emerging recognition that marine microbial communities are part of tightly connected networks by providing evidence that these interactions are mediated through production and exchange of infochemicals.

In terrestrial systems, interactions between photosynthetic organisms and bacteria occur primarily within the rhizosphere, a region surrounding plant roots in which gradients of released molecules support distinct microbial communities, enhancing the growth of some bacteria while restricting the growth of others[5]. In aquatic systems, similar interactions were proposed over 40 years ago to occur within the phycosphere, a rhizosphere analogue[6]. Today, theoretical and empirical studies confirm that phytoplankton are surrounded by a diffusive boundary layer in which secreted molecules accumulate in excess of bulk seawater concentrations[7,8], enhancing the potential for bacterial detection of, and communication and interaction with, algal cells[3].

Marine diatoms commonly co-occur with members of the Proteobacteria and Bacteroidetes in laboratory cultures and some natural blooms[3]. To identify mechanisms underlying specific interactions, we isolated 49 cultivable bacterial strains co-occurring with four isolates of the coastal diatom *Pseudo-nitzschia multiseries* originating from the Pacific and the Atlantic Oceans (Extended Data Table 1). We focus on *P. multiseries*, a diatom with a publicly available draft genome, because of its ubiquitous distribution in coastal ecosystems, ecological importance as a harmful alga[9], and relatively large size (~50 μm).

Bacteria affiliated with the *Sulfitobacter*, *Hyphomonas*, *Marinobacter*, *Limnobacter*, and *Croceibacter* were among isolated bacteria and displayed more than 97% identity in 16S rRNA sequences regardless of the originating *P. multiseries* culture (Extended Data Fig. 1 and Supplementary Information Table 1). The potential impact of these bacteria on host physiology was examined by first curing *P. multiseries* PC9 of bacteria via antibiotic treatment[10]. The specific growth rate of *P. multiseries* PC9 was not significantly

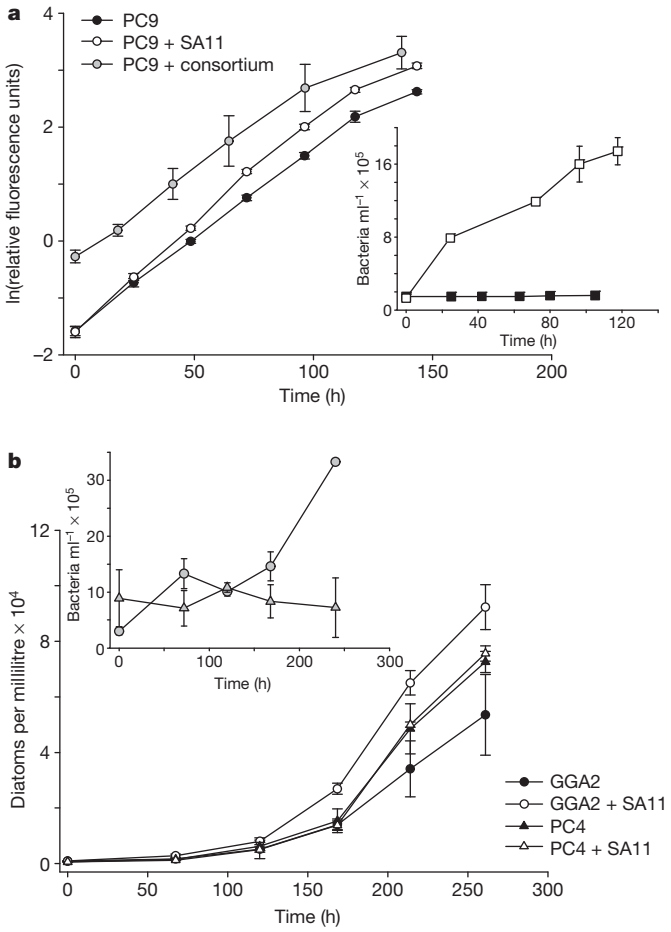


Figure 1 | Growth characteristics of the *P. multiseries*–*Sulfitobacter* sp. SA11 co-culture. **a**, Growth of axenic *P. multiseries* PC9, PC9 with SA11, and PC9 with the bacterial consortium as monitored by relative chlorophyll *a* fluorescence. Inset: cell concentration of SA11 grown without PC9 (filled squares) or with PC9 (open squares). Error bars, s.d. of triplicate cultures. Axenic versus co-culture with SA11 growth experiments were replicated five times. **b**, Abundance of axenic *P. multiseries* GGA2, GGA2 with SA11, axenic *P. multiseries* PC4, and PC4 with SA11. Inset: cell concentration of SA11 grown with GGA2 (circles) or with PC4 (triangles). Error bars, s.d. of triplicate cultures.

[1]School of Oceanography, University of Washington, Seattle, Washington 98195, USA. [2]Chemistry Faculty, New York University Abu Dhabi, PO Box 129188, Abu Dhabi, United Arab Emirates. [3]Department of Microbiology, University of Washington, Seattle, Washington 98195, USA. [4]Department of Microbiology, University of Georgia, Athens, Georgia 30602, USA. [5]Department of Marine Science, University of Georgia, Athens, Georgia 30602, USA.

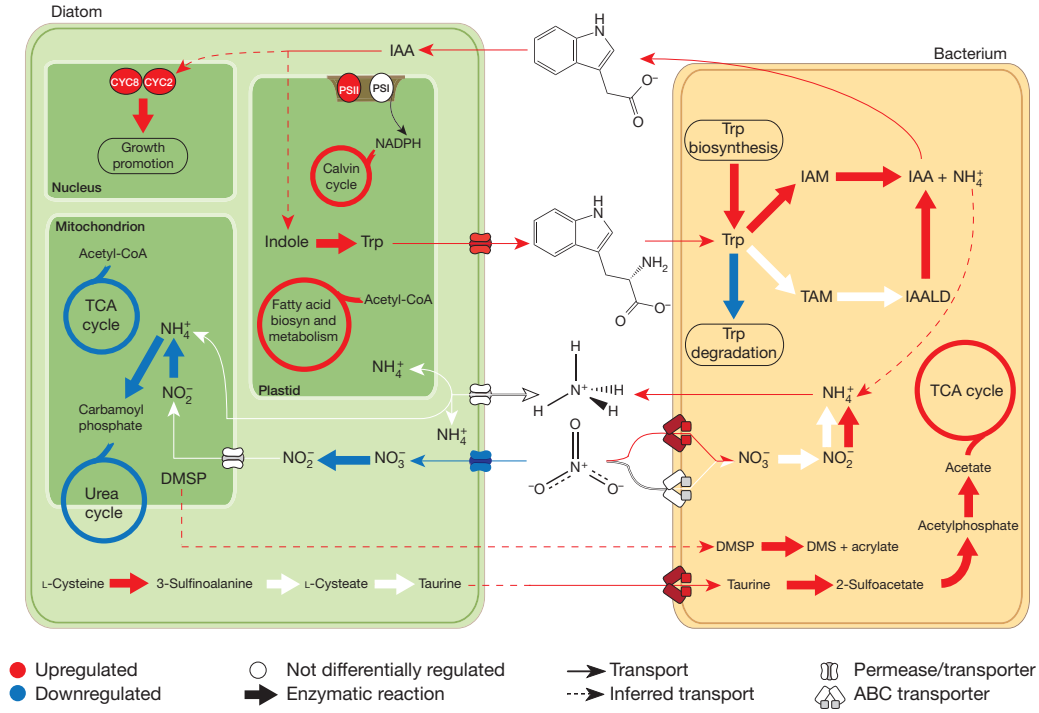


Figure 2 | Model of *P. multiseries*–*Sulfitobacter* interactions based on transcriptomic and targeted metabolite analyses. Molecules with a structure indicate detection in the co-culture supernatant. Genes/transporters/metabolic cycles are shown as upregulated (red), downregulated (blue), or not differentially regulated (white) in co-culture relative to monocultures. Metabolic cycles were assigned an expression pattern if at least one gene specific for the cycle was differentially expressed and no others were regulated in the opposite direction. Supplementary Information Tables 1 and 2 list fold-expression and statistical significance based on triplicate biological experiments. IAA potentially regulates expression of two cyclins that typically regulate the cell cycle[30]. Trp, tryptophan; DMS, dimethyl sulfide; PSI, PSII, photosystem I, II; CYC2, CYC8, cyclins 2 and 8; IAALD, indole-3-acetaldehyde.

affected by removal of its bacterial consortium in the short term (μ_{axenic} = 0.75 ± 0.03; $\mu_{consortium}$ = 0.80 ± 0.10 d^{-1}) (Fig. 1a). Over the longer term (>18 months), the growth rate decreased significantly (to $\mu_{axenic} < 0.3\ d^{-1}$), implying dependence on bacteria[11]. Within 7 months of curing it of bacteria, *P. multiseries* PC9 was co-cultured with individual bacterial strains in a synthetic seawater medium lacking added organic carbon[12], ensuring bacterial growth was dependent upon diatom released organic molecules. Growth rates of the cured diatom were unaffected when co-cultured with *Marinobacter* or *Limnobacter* strains, whereas a *Croceibacter* strain was lethal. Four *Sulfitobacter* strains significantly enhanced the specific growth rate of the diatom by 18–35% despite use of a medium optimized to support axenic diatom growth[12] (Fig. 1a, Extended Data Fig. 2a and Extended Data Table 2). Co-culture with two *Phaeobacter* strains closely related to *Sulfitobacter* did not enhance diatom growth (Extended Data Table 2). Together, these results indicate that the *Sulfitobacter* strains produce a diatom growth-altering factor. A single strain of *Sulfitobacter* (SA11) was chosen for further study.

The growth effect of SA11 appears remarkably specific. No growth enhancement was observed when SA11 was co-cultured with another diatom *Thalassiosira pseudonana* or with two of the four strains of *P. multiseries* (Fig. 1b, Extended Data Table 3 and Extended Data Fig. 2b). SA11 cell numbers increased over time in co-culture with responsive *P. multiseries* strains (PC9, GGA2), but not with a non-responsive strain (PC4) (Fig. 1 insets), implying the diatom somehow modulates SA11 growth. To identify pathways involved in the interaction, we generated a draft genome sequence for SA11 and used a combination of comparative whole-cell transcriptomics and targeted metabolite analyses of the partners grown either in isolation or in co-culture.

In co-culture, *P. multiseries* PC9 provided SA11 with the organic carbon necessary for growth, as evidenced by increased SA11 cell numbers (Fig. 1a inset). Transcriptome changes implicated diatom-produced taurine, a sulfonated intracellular metabolite previously identified in several *Pseudo-nitzschia* species[9]. *P. multiseries* increased transcription of cysteine dioxygenase (*cdo*), the enzyme that catalyses the first step in biosynthesis of taurine from L-cysteine and whose activity is correlated with intracellular taurine concentrations[13]. SA11 increased abundance of transcripts required for taurine uptake (*tauABC*) and catabolism to acetate (*tpa*, *xsc*, *ackA*), which can feed into the TCA cycle[14]. SA11 appears particularly responsive to diatom-produced organosulfur molecules as transcripts associated with dimethylsulfoniopropionate (DMSP) lyase (*dddL*) also increased (Fig. 2 and Supplementary Information Tables 2 and 3), suggesting degradation of DMSP to acrylate and dimethylsulfide[15]. *P. multiseries* also increased the abundance of transcripts associated with photosystem II (*psbB*), light-harvesting proteins (*LHCA4*, *LHCF4*), fucoxanthin, and enzymes in the Calvin cycle while reducing most transcripts associated with genes in the TCA cycle (Fig. 2 and Supplementary Information Table 2). These observations suggest decreased respiration and increased photosynthesis and carbon fixation, perhaps fuelling carbon excretion to SA11.

Symbiotic interactions, such as terrestrial plant–microbe interactions, commonly involve exchange of reduced nitrogen[5]. The sole source of added nitrogen in the growth medium was nitrate[12]. However, in co-culture, *P. multiseries* decreased abundance of transcripts associated with nitrate transport (*NRT1*, *NRT2*) and reduction to ammonia (*NR*, *NiR*) while SA11 increased the abundance of transcripts associated with nitrate uptake (*nitTABC*) and reduction to ammonia (*nirB*) (Fig. 2 and Supplementary Information Tables 2 and 3). Significantly more ammonium was detected in the co-culture medium than in the medium blank or when the diatom was cultured alone, indicating that SA11 released a fraction of its imported nitrate into the media as ammonium (Extended Data Fig. 3a). Together, these results suggest that, in co-culture, SA11 increases nitrate uptake and ammonium release and that *P. multiseries* preferentially utilizes bacterial-derived ammonium for growth, rather than exogenous nitrate.

Addition of NH_4Cl to axenic *P. multiseries* had no impact on growth (μ: 0.57 ± 0.03 d^{-1} vs 0.55 ± 0.02 d^{-1}), indicating that although reduced nitrogen was essential to the interaction, another molecule was responsible for the growth effect.

Tryptophan and related derivatives are common signalling molecules in the marine environment[16,17]. In co-culture, *P. multiseries* increased transcript abundance for the conversion of indole to tryptophan and for a putative tryptophan/tyrosine permease (Fig. 2 and Supplementary Information Table 2), suggesting increased biosynthesis and export of tryptophan. Tryptophan was detected in the growth media after harvesting cells when *P. multiseries* was maintained alone (448 ± 106 pM) or in co-culture (202 ± 20 pM) (Extended Data Fig. 3b). A reduced tryptophan concentration in the co-culture suggests that SA11 may be importing diatom-released tryptophan. Increased transcript abundance for endogenous tryptophan biosynthesis and decreased transcript abundance for tryptophan degradation by SA11 (Fig. 2 and Supplementary Information Table 3) suggest that SA11 increased utilization of extra- and intracellular tryptophan in co-culture.

Co-culture with *P. multiseries* triggered SA11 to increase transcripts associated with the indole-3-acetamide (IAM) and tryptamine (TAM) pathways (Fig. 2 and Supplementary Information Table 3) that convert tryptophan to indole-3-acetic acid (IAA), an endogenous plant hormone that is also produced and excreted by rhizobia to skew symbiotic plant development[18]. IAA was detected in the growth medium when SA11 was maintained alone or in co-culture (Extended Data Fig. 3c). Assuming a constant rate of production and release of IAA by SA11, the concentration in the co-culture (6.1 ± 0.4 pM) was significantly lower than predicted (540 pM) (see Methods), implying that *P. multiseries* takes up a minimum of 5 amol IAA per cell each day. We confirmed that *P. multiseries* is responsive to a narrow range of synthetic IAA (50–100 nM) added either once during the growth cycle (Extended Data Table 4) or as multiple 50 nM additions over 8 days (Extended Data Fig. 4). The growth enhancement by SA11 (19–35%) versus IAA (~10%) suggests that SA11 produces other molecules that further enhance *P. multiseries* growth. Furthermore, the difference in orders of magnitude in concentration of synthetic versus bacterial IAA required to stimulate a diatom response (nanomolar vs picomolar) reiterates the potential importance of phycosphere interactions where local concentrations of IAA are significantly higher than bulk concentrations in the media, an observation consistent with previous work on diffusive boundary layers[7,19].

To explore the potential prevalence of these interactions in natural populations, we looked for evidence of bacterial production of IAA by performing targeted metabolite analysis on seawater from the surface and chlorophyll maxima at five stations from different regions of the North Pacific Ocean (Extended Data Fig. 5). Since IAA has no clear metabolic role in bacteria[18], production and excretion of IAA in natural populations might indicate bacterial manipulation of responsive phytoplankton similar to rhizobia[5]. Extracellular IAA (1.5–383 pM) was detected in all samples, with the highest concentrations detected in coastal sites with high phytoplankton abundance (Fig. 3a and Extended Data Fig. 3c). The range of measured IAA concentrations in the environment is comparable to that measured in our co-culture, indicating that environmental IAA could elicit a response from diatoms associated with IAA-producing bacteria.

The *Roseobacter* clade, to which SA11 belongs, is among the most ubiquitous lineages observed with phytoplankton[3,20], and active IAA production by this group could impact diverse phytoplankton species, many of which have been shown to respond to synthetic IAA[21,22]. To determine whether the *Roseobacter* produced IAA in the field, we examined metatranscriptomic data sets for transcripts associated with the three IAA biosynthetic pathways found in publicly available *Roseobacter* genomes—the indole-3-acetonitrile (IAN), IAM and TAM pathways (Extended Data Fig. 6). We analysed transcripts collected at two coastal stations in the North Pacific Ocean and from three publicly available metatranscriptome data sets from Monterey Bay, the California coast, and station ALOHA in the North Pacific Gyre

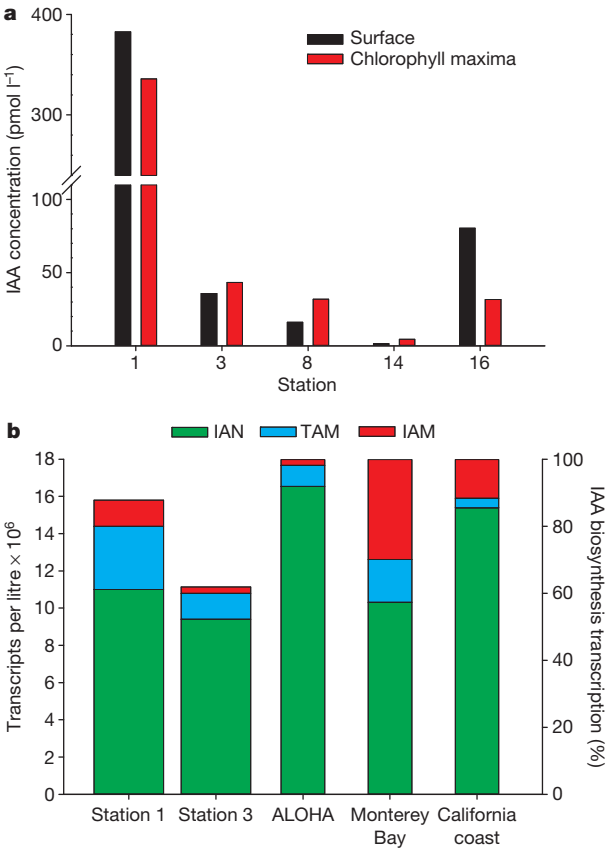


Figure 3 | Detection of IAA and IAA biosynthesis in the marine environment. a, IAA concentrations at five stations in the North Pacific Ocean from surface (black) and chlorophyll maxima (red) waters. **b**, Abundance of transcripts from the three IAA biosynthetic pathways present in the *Roseobacter*. Thick bars represent transcripts per litre associated with any gene in the pathways calculated on the basis of an internal standard; thin bars represent percentage IAA biosynthesis transcription contributed by each pathway for data sets in which no internal standard information was available. Genes used in each pathway are in Extended Data Fig. 6. ALOHA coincides with station 16 in Fig. 3a.

(Extended Data Fig. 5). In all data sets, the three IAA biosynthetic pathways were actively transcribed with an average abundance of 10^7 l^{-1} (~0.01% of total transcripts). Transcripts associated with the IAN pathway dominated all data sets, with TAM and IAM transcripts contributing 10–40% of total IAA transcription (Fig. 3b). Although SA11 also possesses the IAN pathway, no IAN transcripts were detected in our laboratory experiments, suggesting complex regulatory processes control different pathways. Our results thus present a lower limit on IAA biosynthesis as other bacterial taxa probably also produce IAA, and other IAA biosynthesis pathways (Extended Data Fig. 6) may be active.

The *P. multiseries*–*Sulfitobacter* model system developed here demonstrates the complexity of microbial interactions, potentially occurring within a phycosphere that concentrates hydrophobic signalling molecules and persists despite seawater turbulence[23–25]. Tryptophan secretion by *P. multiseries* may attract a wide range of bacteria, but only bacteria that can convert tryptophan to IAA could create a positive feedback loop between diatom tryptophan and bacterial IAA (Fig. 2). Accumulation of IAA to significantly higher local concentration around algal cells relative to seawater[26] would ensure that IAA producers residing within the phycosphere could skew the growth of algae whereas distant bacteria would not. Exchange of essential molecules such as ammonia and organosulfur compounds would further enhance synergy. Additional signalling molecules between bacteria and diatoms and among bacteria are probably key to recognizing

and sustaining beneficial partners and excluding cheaters. Such added specificity could explain the different responses of closely related *P. multiseries* strains to SA11 and is reminiscent of legume–rhizobia interaction specificity achieved through multiple signalling molecules[27]. In this context, signalling may distinguish between organisms with a long history of association and organisms with latent capacity for interaction[28].

Besides diatoms, several unicellular green algal lineages and cyanobacteria have also been shown to respond to synthetic IAA[21,22,29]. Direct detection of algal responses to IAA in the field is not yet possible as the genetic basis for algal responses to IAA remains unknown[21]. Further work is needed to characterize these genetic elements. The interactions described here illustrate how bacterial influence on phytoplankton physiology may be linked to the global carbon cycle and algal bloom formation, and probably affect ecosystem functioning.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 14 October 2014; accepted 23 April 2015.
Published online 27 May 2015.

1. Azam, F. & Malfatti, F. Microbial structuring of marine ecosystems. *Nature Rev. Microbiol.* **5,** 782–791 (2007).
2. Little, A. E. F., Robinson, C. J., Peterson, S. B., Raffa, K. F. & Handelsman, J. Rules of engagement: interspecies interactions that regulate microbial communities. *Annu. Rev. Microbiol.* **62,** 375–401 (2008).
3. Amin, S. A., Parker, M. S. & Armbrust, E. V. Interactions between diatoms and bacteria. *Microbiol. Mol. Biol. Rev.* **76,** 667–684 (2012).
4. Yool, A. & Tyrrell, T. Role of diatoms in regulating the ocean's silicon cycle. *Glob. Biogeochem. Cycles* **17,** 1103–1124 (2003).
5. Jones, K. M., Kobayashi, H., Davies, B. W., Taga, M. E. & Walker, G. C. How rhizobial symbionts invade plants: the *Sinorhizobium–Medicago* model. *Nature Rev. Microbiol.* **5,** 619–633 (2007).
6. Bell, W. & Mitchell, R. Chemotactic and growth responses of marine bacteria to algal extracellular products. *Biol. Bull.* **143,** 265–277 (1972).
7. Jonsson, P. R., Pavia, H. & Toth, G. Formation of harmful algal blooms cannot be explained by allelopathic interactions. *Proc. Natl Acad. Sci. USA* **106,** 11177–11182 (2009).
8. Karp-Boss, L., Boss, E. & Jumars, P. A. Nutrient fluxes to planktonic osmotrophs in the presence of fluid motion. *Oceanogr. Mar. Biol. Annu. Rev.* **34,** 71–107 (1996).
9. Smith, G. J., Ladizinsky, N. & Miller, P. E. in *Harmful Algal Blooms 2000* (eds Hallegraeff, G. M., Blackburn, S. I., Bolch, C. J. S. & Lewis, R. J.) 324–327 (UNESCO, 2000).
10. Shishlyannikov, S. M. *et al.* A procedure for establishing an axenic culture of the diatom *Synedra acus* subsp. *radians* (Kütz.) Skabibitsch. from Lake Baikal. *Limnol. Oceanogr. Methods* **9,** 478–484 (2011).
11. Windler, M. *et al.* Influence of bacteria on cell size development and morphology of cultivated diatoms. *Phycol. Res.* **62,** 269–281 (2014).
12. Price, N. M. *et al.* Preparation and chemistry of the artificial algal culture medium Aquil. *Biol. Oceanogr.* **6,** 443–461 (1988).
13. Stipanuk, M. H. & Dominy, J. E. Surprising insights that aren't so surprising in the modeling of sulfur amino acid metabolism. *Amino Acids* **30,** 251–256 (2006).
14. Cook, A. M. & Denger, K. Metabolism of taurine in microorganisms: a primer in molecular biodiversity? *Adv. Exp. Med. Biol.* **583,** 3–13 (2006).
15. Curson, A. R. J., Rogers, R., Todd, J. D., Brearley, C. A. & Johnston, A. W. B. Molecular genetic analysis of a dimethylsulfoniopropionate lyase that liberates the climate-changing gas dimethylsulfide in several marine α-proteobacteria and *Rhodobacter sphaeroides*. *Environ. Microbiol.* **10,** 757–767 (2008).
16. Riffell, J. A., Krug, P. J. & Zimmer, R. K. The ecological and evolutionary consequences of sperm chemoattraction. *Proc. Natl Acad. Sci. USA* **101,** 4501–4506 (2004).
17. Yambe, H. *et al.* L-Kynurenine, an amino acid identified as a sex pheromone in the urine of ovulated female masu salmon. *Proc. Natl Acad. Sci. USA* **103,** 15370–15374 (2006).
18. Spaepen, S. & Vanderleyden, J. Auxin and plant-microbe interactions. *Cold Spring Harb. Perspect. Biol.* **3,** a001438 (2011).
19. Stocker, R. Marine microbes see a sea of gradients. *Science* **338,** 628–633 (2012).
20. Geng, H. & Belas, R. Molecular mechanisms underlying roseobacter–phytoplankton symbioses. *Curr. Opin. Biotechnol.* **21,** 332–338 (2010).
21. Lau, S., Shao, N., Bock, R., Jürgens, G. & De Smet, I. Auxin signaling in algal lineages: fact or myth? *Trends Plant Sci.* **14,** 182–188 (2009).
22. Suzuki, T., Nakasato, K., Shapiro, S., Pomati, F. & Neilan, B. A. Effects of synthetic local anaesthetics on the growth of the cyanobacterium *Synechococcus leopoliensis*. *J. Appl. Phycol.* **16,** 145–152 (2004).
23. Durham, W. M. *et al.* Turbulence drives microscale patches of motile phytoplankton. *Nature Commun.* **4,** 2148 (2013).
24. Stocker, R. & Seymour, J. R. Ecology and physics of bacterial chemotaxis in the ocean. *Microbiol. Mol. Biol. Rev.* **76,** 792–812 (2012).
25. Taylor, J. R. & Stocker, R. Trade-offs of chemotactic foraging in turbulent water. *Science* **338,** 675–679 (2012).
26. Guasto, J. S., Rusconi, R. & Stocker, R. Fluid mechanics of planktonic microorganisms. *Annu. Rev. Fluid Mech.* **44,** 373–400 (2012).
27. Wang, D., Yang, S., Tang, F. & Zhu, H. Symbiosis specificity in the legume-rhizobial mutualism. *Cell. Microbiol.* **14,** 334–342 (2012).
28. Hom, E. F. Y. & Murray, A. W. Niche engineering demonstrates a latent capacity for fungal-algal mutualism. *Science* **345,** 94–98 (2014).
29. De Smet, I. *et al.* Unraveling the evolution of auxin signaling. *Plant Physiol.* **155,** 209–221 (2011).
30. Roudier, F. *et al.* The *Medicago* species A2-type cyclin is auxin regulated and involved in meristem formation but dispensable for endoreduplication-associated developmental programs. *Plant Physiol.* **131,** 1091–1103 (2003).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank U. John for the antibiotic recipe, L. Gram for providing *Phaeobacter* strains, J. Tsai for preparing the SA11 DNA for sequencing, the captain and crew of the R/V *Kilo Moana* and the R/V *Thomas G. Thompson* for help during the cruises, D. French for Extended Data Fig. 5, and T. Chiang for discussions. This work was supported in part by Gordon and Betty Moore Foundation grant GBMF3776 to E.V.A, and support from National Science Foundation (NSF) award OCE-1228770 to A.E.I., OCE-1205233 to E.V.A., and OCE-1342694 to M.A.M. S.A.A. was partly supported by a NSF/National Institutes of Health Pacific Northwest Consortium postdoctoral fellowship. H.M.V. was partly supported by a Natural Sciences and Engineering Research Council of Canada Postgraduate Scholarship – Master's grant.

Author Contributions S.A.A., L.R.H., B.D., B.P.D. and H.V.T. conducted experiments; B.P.D. isolated and prepared bacterial transcriptomic and metatranscriptomic samples; R.L.M. isolated all other nucleic acids and prepared libraries for sequencing; C.T.B. quality-trimmed sequenced data, assembled the SA11 genome and translated the metatranscriptome; M.S.P. quality-trimmed and quantified the PC9 transcriptomes; L.T.C. and K.R.H. collected environmental metabolome samples and performed MS analyses; S.A.A., L.R.H., M.R.P., A.E.I., M.A.M., and E.V.A. designed experiments; S.A.A. and L.T.C. analysed the data. All authors were involved in manuscript writing.

Author Information The data reported in this paper are presented in Supplementary Information and archived at the following databases: 16S rDNA sequences, GenBank accession numbers KM033232–KM033280; transcriptomes, Gene Expression Omnibus accession number GSE65189; metatranscriptomes, Sequence Read Archive accession number PRJNA272345; SA11 genome, Integrated Microbial Genomes (IMG) submission 11682. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to E.V.A. (armbrust@uw.edu) or S.A.A. (sa132@nyu.edu).

METHODS

No statistical methods were used to predetermine sample size.

Diatom growth and axenic culture generation. Milli-Q water (>18.2 MΩ cm) was used for all synthetic seawater media preparations. *T. pseudonana* CCMP1335 was acquired from National Center for Marine Algae and Microbiota. *P. multiseries* strains were isolated from seawater samples collected from locations indicated in Extended Data Table 1 and were identified using Automated Ribosomal Intergenic Spacer Analysis (ARISA) according to ref. 31. Cultures were initially grown in f/2 medium[32] and were acclimated and further maintained for all experiments in the synthetic seawater medium, Aquil[12]. All cultures were grown in 13 °C in a 16 h light/8 h dark diurnal cycle (80 μE m^{-2} s^{-1}) in semi-continuous batch cultures[33] with an initial cell density ~8,000–10,000 cells per millilitre for non-axenic cultures and ~2,000–4,000 cells per millitre for axenic cultures. Diatom growth was monitored by measuring *in vivo* fluorescence using a 10-AU fluorometer (Turner Designs) or by counting cells using a Sedgwick-rafter (Wildlife Supply Company). Growth rates were estimated by measuring *in vivo* chlorophyll *a* fluorescence (relative fluorescence units) or cell counts. Specific growth rates (μ) were calculated from the linear regression of the natural log of *in vivo* fluorescence or cell counts versus time during the exponential growth phase of cultures. Standard deviation of μ was calculated from μ values from biological replicates ($n = 3$ unless otherwise indicated) over the exponential growth period. Percentage growth enhancement was calculated as the difference between $\mu_{co\text{-}culture}$ and μ_{axenic} divided by $\mu_{co\text{-}culture}$.

Axenic cultures were generated by adopting the protocol from ref. 10 with minor modifications as described below. Approximately 25 ml of a mid-exponential phase growing diatom culture was gravity filtered onto 0.65 μm pore-size polycarbonate membrane filter (Millipore). Cells were quickly rinsed with sterile Aquil media. Using sterile tweezers, the filter was carefully removed from the filtration unit and washed for ~1 min in sterile media containing 20 μg ml^{-1} Triton X-100 detergent to remove surface-attached bacteria. The filter was discarded after re-suspension of cells by gentle shaking in sterile detergent-free media. Cells were again gravity filtered onto a fresh 0.65 μm pore-size polycarbonate membrane filter and rinsed with sterile media. Subsequently, cells were washed off the filter by gentle shaking into sterile media containing a suite of antibiotics (per millilitre: 50 μg streptomycin, 67 μg gentamycin, 20 μg ciprofloxacin, 2.2 μg chloramphenicol, and 100 μg ampicillin). Cells were incubated in antibiotic-containing media for 24–48 h under regular growth conditions. Finally, 0.5–1.0 ml of antibiotics-treated cells were transferred to antibiotic-free media. Cultures were regularly monitored (every four or five transfers, ~1 month) for bacterial contamination by checking for bacterial growth in Zobell marine broth[34] in addition to using Sybr Green I (Invitrogen) staining and epifluorescence microscopy (Nikon Eclipse 80i) as described previously[35]. Bacterial contamination was observed only once over the course of ~18 months for PC9 owing to human error; the culture was discarded and fresh axenic cultures were prepared as above.

Bacterial growth, isolation, and classification. Bacteria were typically grown on marine agar plates (per litre: 5 g peptone, 0.5 g yeast extract, 15 g agar, and 750 ml seawater) incubated at 20 °C in the dark or in marine broth[34] at 30 °C with shaking at 150 r.p.m. Bacterial growth was measured by counting colony-forming units or by using a Guava EasyCyte Plus flowcytometer (Millipore) after cells were stained with Sybr Green I stain.

Bacteria were isolated from late-exponential phase growing *P. multiseries* cultures by serially diluting 0.5 ml aliquots of culture into sterile Aquil. Diluted aliquots were then plated onto agar plates containing, per litre of seawater, 15 g agar and 2 g of a carbon source (peptone and yeast extract, succinate, glucose, CAS amino acids, or only background organic carbon in seawater). Plates were incubated at room temperature in the dark and morphologically different bacterial colonies were isolated and stored in 15% glycerol stocks at −80 °C for future experiments.

To identify isolated bacteria, isolates were grown from single colonies in marine broth overnight and cells were centrifuged at 13,000*g* for 2 min. The supernatants were removed and DNA was extracted using a DNA Blood & Tissue kit (Qiagen) according to the manufacturer's instructions. Using universal 16S rDNA primers (27F, 1492R), 16S rDNA from all bacterial isolates was amplified using a Taq DNA polymerase kit (Apex). The temperature profile for PCR consisted of an initial incubation at 94 °C for 3 min, followed by 32 cycles of 94 °C for 30 s, 55 °C for 1 min and 72 °C for 2 min, and a final extension step at 72 °C for 20 min. Amplified product was cleaned using a High Pure PCR Product Purification Kit (Roche). Purified PCR products were sequenced using Sanger technology (Genewiz).

Sequences were quality trimmed using Sequencher 4.6 (Gene Codes) and initially aligned using ClustalW as implemented in Mega 5.2.2 (ref. 36). The alignment was refined using NAST (http://greengenes.lbl.gov). Phylogenetic inference of the masked alignment was based on maximum likelihood, using the JTT model with bootstrap support of 100 replicates as implemented in Mega 5.2.2 (ref. 36). Sequences were deposited in GenBank under accession numbers KM033232–KM033280.

Co-culture experiments. Because we observed long term (>18 months) growth rate differences between axenic and non-axenic *P. multiseries*, all co-culture experiments (including the transcriptome and metabolite analyses experiments) were conducted within 7 months of curing the diatom of bacteria. All experiments were conducted in Aquil[12]. Bacteria were plated freshly before each experiment on marine agar and were grown from single colonies in marine broth overnight (30 °C, 150 r.p.m.). Cells were centrifuged (3,500*g* for 5 min), washed twice with sterile Aquil, and diluted to a stock cell density of ~1×10^8 cells per millilitre with sterile Aquil. This stock was used to inoculate the freshly prepared diatom culture to achieve a final bacterial cell density of ~1×10^5–2×10^5 cells per millilitre. Diatoms were inoculated from an early to mid-exponential phase growing culture into fresh media to an initial diatom cell density of ~2,000–4,000 cells per millilitre to achieve an ~50:1 bacteria:diatom ratio. Diatom and bacterial growth were measured as described above. For experiments where *Sulfitobacter* sp. strain SA11 was grown alone, Aquil was supplemented with 11 mM glucose as the sole carbon source except for the transcriptome experiment (see below), where only 1 μM glucose was used.

For the transcriptome experiments, axenic *P. multiseries* strain PC9 was used to inoculate 2 l sterile polycarbonate bottles. Treatments consisted of (1) PC9 and SA11 co-culture, (2) axenic PC9, and (3) SA11 supplemented with 1 μM glucose. All treatments were in triplicate. Growth rates for PC9 were $\mu_{axenic} = 0.75 \pm 0.03\ d^{-1}$ and $\mu_{co\text{-}culture} = 0.94 \pm 0.04\ d^{-1}$. Growth rates for SA11 were $\mu_{glucose} = 0.45 \pm 0.01\ d^{-1}$ and $\mu_{co\text{-}culture} = 0.46 \pm 0.02\ d^{-1}$. Cells were harvested at mid-exponential growth (96 h after inoculation for all treatments) by filtering the culture through a 3 μm pore-size polycarbonate filter to capture the majority of diatom cells (this step was skipped for treatment 3), followed by filtration through 0.22 μm pore-size polycarbonate filter to capture bacteria. Final cell densities at time of harvesting were $PC9_{axenic}$ ~3.5×10^4 cells per millilitre, $PC9_{co\text{-}culture}$ ~5.8×10^4 cells per millilitre, $SA11_{co\text{-}culture}$ ~1.7×10^6 cells per millilitre, and $SA11_{glucose}$ ~3×10^6 cells per millilitre. Filters were immediately flash frozen in liquid nitrogen and later stored at −80 °C. Flow-through media were used for targeted metabolite analyses (see below).

IAA addition experiments. IAA (Sigma-Aldrich) was dissolved in Milli-Q water and sterilized using syringe filtration. *P. multiseries* strain GGA2 was inoculated into fresh Aquil media supplemented with different concentrations of IAA or equivalent volumes of Milli-Q water (control). For IAA additions in Extended Data Fig. 5, IAA or Milli-Q water was added every 2 days as described to simulate a continuum of active IAA concentrations in the media.

***P. multiseries* genome.** A publically available genome of strain CLN-47 was used for our analysis (http://genome.jgi-psf.org/Psemu1/Psemu1.home.html).

SA11 genome. All nucleic acid quantifications were measured using a Qubit Fluorometer (Invitrogen; Life Technologies). DNA was extracted using a Qiagen DNA Blood and Tissue kit according to the manufacturer's instructions. Ten micrograms of DNA were sheared to 2–3 kb using a Hydroshear (Genomic Solutions) with a standard shearing assembly. To prepare the DNA for SOLiD 60 base pair (bp) × 60 bp mate-pair sequencing, we used a unique protocol combining different steps from the SOLiD 3, SOLiD 4, and SOLiD 5500 mate-paired protocols (R.M. & E.V.A., manuscript in preparation). The library was attached to beads by emulsion PCR, which was done at the Life Technologies Research and Development Unit. For sequencing, 50 million library-containing beads were deposited onto one spot of an eight-spot slide and run on a SOLiD 4 Next Generation Sequencer (Life Technologies). In addition to SOLiD sequencing, we used Ion Torrent to improve the genome assembly. DNA was prepared as described above. Ion Torrent library preparations were performed according to the manufacturer's instructions.

The SA11 genome was assembled using a combination of fragment Ion Torrent reads and 50 × 50 mate-paired SOLiD colorspace reads with an insert size of 2,200 bp (s.d. 800 bp). Fastq_quality_trimmer from a FASTX Toolkit was used to quality trim and filter Ion Torrent reads with parameters '-Q33 -l 50 -t 14'. SOLiD reads were screened for PCR duplication artefacts using fastq_nodup from SEAStAR[37] with parameters '-d 2–no_prefix' and were trimmed and filtered on the basis of quality score using trimfastq from SEAStAR with parameters '-p.75 -l 34 -e 3.0–add_len–no_prefix'. Contigs were created with the *de novo* assembly tool Newbler. The contigs were organized into scaffolds using graph_ops from SEAStAR on the basis of mate-pair connections identified by a BWA alignment of SOLiD reads to Ion Torrent contigs. Contigs were also created for a *de novo* assembly of SOLiD reads using Velvet with a kmer size of 31, coverage cutoff of 35, expected coverage of 200, insert size of 2,200, insert size standard deviation of 800, scaffolding disabled, and a minimum contig length of 100. These contigs were used to fill gaps between scaffolded Ion-Torrent-based Newbler contigs where possible.

The genome can be publicly accessed through IMG (http://img.jgi.doe.gov; submission 11682).

PC9 transcriptome. Total RNA was isolated from 3-μm filters using a ToTALLY RNA Total RNA Isolation Kit (Ambion; Life Technologies), and messenger RNA (mRNA) was purified from the total RNA using a MicroPoly(A)Purist Kit (Ambion; Life Technologies). SOLiD 75 bp × 35 bp paired-end libraries were generated from ~500 ng of mRNA from each replicate and treatment using a SOLiD Total RNA-Seq Kit (Life Technologies) with the gel option according to the manufacturer's instructions. The libraries were attached to beads in-house by emulsion PCR according to the SOLiD manual. For sequencing, 700 million library-containing beads were deposited onto a full slide and run on a SOLiD 4 Next Generation Sequencer (Life Technologies).

The SOLiD reads were trimmed on the basis of quality score using custom in-house software[37]. Trimmed reads with a length shorter than 28 colorspace transitions were removed. Filtered reads were then aligned to the *P. multiseries* draft genome gene catalogue transcripts provided by JGI (Psemu1_GeneCatalog_transcripts_20111011.nt.fasta) using BWA (version 0.5.8) allowing for two mismatches in a seed length of 18, and up to four mismatches across an entire read. Anti-sense reads were removed and counts for SOLiD reads aligning to gene catalogue transcripts were calculated from the resulting SAM alignment file[38] using SEAStAR[37]. SEAStAR counts were then analysed with edgeR[39] for differential expression and significance testing using Benjamini Hochberg multiple testing corrections. Count tables were then merged with KEGG annotations provided by JGI. Using a false discovery rate cutoff of 0.05, 2,143 genes were differentially expressed in PC9 out of ~19,703 gene models.

SA11 transcriptome. RNA was isolated from 0.22-μm filters using an RNeasy Mini Kit along with RNAprotect Bacteria Reagent (Qiagen) according to the manufacturer's instructions using 1 mg ml^{-1} lysozyme solution (Fisher) to lyse the cells. Total RNA was treated for DNA contamination using two successive treatments with Turbo-DNase (Ambion; Life Technologies) and cleaned/concentrated with an RNeasy MinElute Cleanup Kit (Qiagen). Ribosomal RNAs (rRNAs) were selectively removed using a subtractive hybridization protocol[40] with biotinylated rRNA probes specific to the organism(s) in each sample (for example, 16S and 23S for SA11 mono- and co-cultures or eukaryotic 18S and 28S for co-cultures). Subtracting diatom rRNAs was essential as the first filtration through 3-μm pore-size filter (see above) did not completely remove all diatom cells. Probe-bound RNAs were removed with strepdavidin-coated magnetic beads (New England Biolabs). rRNA-depleted samples were then linearly amplified using a MessageAmp II-Bacteria kit (Ambion; Life Technologies).

Amplified mRNA was then converted into complementary DNA (cDNA) using a SuperScript III First-Strand Synthesis System for RT–PCR kit (Ambion; Life Technologies) and a NEBNext mRNA Second Strand Synthesis Module (New England Biolabs) according to the manufacturer's instructions. SOLiD 75 bp × 35 bp paired-end libraries were generated from 1 μg of cDNA using a SOLiD Fragment Library Construction Kit according to the manufacturer's instructions. The libraries were attached to beads in-house by emulsion PCR according to the SOLiD manual. For sequencing, 450 million library-containing beads were deposited onto three lanes of a flow cell and run on a SOLiD 5500 Next Generation Sequencer (Life Technologies).

SOLiD reads were trimmed and filtered to remove low-quality or low information content sequence using trimfastq from SEAStAR with settings 'trimfastq -z -s–add_len -l 30 -p.9 -e 3.0'. Trimmed reads were aligned to IMG-derived SA11 reference genome using BWA with settings '-k 2 -n.001 -l 18 -t 8 –c' for the aln subcommand and settings '-n 100' for the samse subcommand. Resulting SAM files were processed with ref_select and graph_ops from SEAStAR to get per gene read counts. These counts were processed with the R package edgeR to identify SA11 genes with significant differential expression between the two conditions. Using a false discovery rate cutoff of 0.001 and a fold-change ≥2, 2,620 genes were differentially expressed in SA11 out of ~5,281 open reading frames.

Culture media targeted metabolite analysis. Ammonium and nitrate concentrations were analysed using an AAII autoanalyzer system (Technicon). To detect and quantify IAA and tryptophan, culture supernatants were acidified to pH ~3.5 using concentrated formic acid (Fisher, Baker Analyzed). Supernatants were then passed through conditioned solid-phase extraction (SPE) HLB columns (Waters) at a flow rate of ~5 ml min^{-1} to bind organic molecules. The columns were washed thoroughly with Milli-Q water and eluted with methanol (ultra-performance liquid chromatography (UPLC) grade, Fisher) according to the manufacturer's instructions. Eluted samples were dried under a stream of nitrogen gas and were then frozen at −80 °C for later analysis. Before analysis, samples were dissolved in water.

Mass spectrometry. Tryptophan and IAA were purchased (Sigma-Aldrich) as well as labelled standards for IAA (indole-2,4,5,6,7-d_5-3-acetic acid) and tryptophan (L-tryptophan-2,3,3-d_3) (CDN Isotopes). To verify the identity and quantify IAA and tryptophan in all standards and culture samples, an ultra-performance liquid chromatography–electrospray ionization–tandem mass spectrometry (UPLC–ESI–MS) method adapted from ref. 41 was used, with some changes. A quadrupole time-of-flight (Waters Xevo G2-S QTOF) mass spectrometer was used for verification of compound identity. For the quadrupole time-of-flight, a cone voltage of 2 V and collision energy of 28 V were used when running IAA, and a cone voltage of 32 V and collision energy of 20 V were used when running tryptophan. Compound quantitation was done using selected reaction monitoring (SRM) on a Waters Xevo TQ-S triple quadrupole mass spectrometer. The same UPLC method was used for all analyses. Both retention times and SRM transition masses were used to target compounds in environmental samples run only using the SRM method (Supplementary Information Table 4). The UPLC method used a Waters Acquity UPLC system equipped with a Waters Acquity UPLC BEH C18 column (1.7 μm, 2.1 mm × 50 mm) at 30 °C with the mobile phase consisting of 0.1% formic acid in water (solvent A) and methanol (solvent B). A linear gradient with a flow rate of 0.2 ml min^{-1} was used from 0 to 7 min (5% B to 90% B), followed by 2 min at 90% B, and 3 min linear gradient back to 5% B to re-equilibrate the column. Both mass spectrometers were configured with positive ion ESI with source conditions as follows: capillary voltage 0.5 kV, source temperature 130 °C, desolvation temperature 550 °C, cone gas flow at 150 l h^{-1}, and desolvation gas at 1,000 l h^{-1}.

Analysis of the retention time and the accurate mass of the molecular ion (176.0706) and fragment ion (130.0656) of an authentic standard of IAA confirmed its presence in a monoculture of SA11 (Supplementary Information Table 4). Fragmentation of IAA was induced by a program that ramped collision energy from 20 to 30 V. To quantify tryptophan and IAA in cultures and environmental samples, a positive ion mode SRM method was used that monitored the following transitions: 176.02 → 130.07 (IAA) (Supplementary Information Table 4) and 188.07 → 118.05 (tryptophan). For SRM, the same cone voltage and collision energy were used as the quadrupole time-of-flight.

For quantification of IAA and tryptophan in culture media samples, percentage recovery during SPE was determined for each molecule in Aquil by acidifying and passing different batches of Aquil through SPE HLB columns and treating them the same way as culture media described above. Three batches were spiked with labelled IAA-d_5 and tryptophan-d_3 before SPE treatment while the rest were spiked after SPE. Extracts were then dried, re-dissolved as described above, and analysed on a Xevo TQ-S triple quadrupole mass spectrometer. Percentage recovery was determined on the basis of the peak area of labelled compounds in treatments spiked before and after SPE (70 ± 16% for IAA; 32 ± 12% for tryptophan). Reported concentrations were corrected for percentage recovery during SPE. An additional batch of Aquil media was extracted by SPE and used to construct a standard curve for each molecule to determine the linear range of the detector ($R^2_{IAA} = 0.999$, $R^2_{trp} = 0.98$). Samples were diluted such that they were in the range of the standard curve concentrations (0–100 nM for IAA and 0–25 nM for tryptophan) and quantified using isotope-labelled internal standards. Internal standard spikes were also within the standard curve concentrations. Concentrations calculated from the standard curve and from the internal standards were generally similar. Reported concentrations are from the internal standard calculation. No IAA contamination was found in blank Aquil. Traces of (<1 pM) tryptophan were detected in Aquil but were significantly lower than all measured concentration in cultures. We attempted to detect taurine in the culture media but were not able to do so, presumably because of poor SPE recovery and inappropriate UPLC chromatography column type (C18).

SA11 IAA production rate. In co-culture, the amount of IAA detected is probably lower than what is produced by SA11 if we presume there was active removal of IAA from solution by the diatom. Because the cell density and the length of the exponential growth phase of SA11 differ in co-culture and monoculture, a direct comparison between IAA concentrations in both treatments is not informative. To calculate the production rate of IAA by SA11, three sets of triplicate cultures of Aquil supplemented with 11 mM glucose were inoculated with SA11. Each set of cultures was harvested after 2, 3, and 4 days of growth. Organic molecules were extracted from the media after removing cells, and IAA was quantified as described in the Mass Spectrometry section. IAA production per cell per day was calculated from the three sets of cultures that were harvested on days 2, 3, and 4. Using the known SA11 cell density in co-culture, we calculated the minimum expected concentration of IAA in co-culture (540 pM). This concentration served as a lower limit on the expected *in situ* IAA concentration that would result from the measured production rate in monoculture since our transcriptome data showed that IAA biosynthesis increased in the co-culture.

Environmental metatranscriptomics and targeted metabolite analysis. Samples were collected in May 2012 at stations 1 and 3 along Line P and in August 2013 at stations 1 (48.6965° N, 126.0387° W), 3 (48.8168° N, 128.6648° W), 8 (49.9872° N, 144.8077° W), 14 (27.3462° N, 152.6717° W), and 16 (22.7603°

N, 158.0003° W) for metatranscriptomes and targeted metabolite analysis, respectively. Seawater samples were collected from the surface using a conductivity–temperature–depth (CTD) rosette equipped with Niskin bottles (20 l).

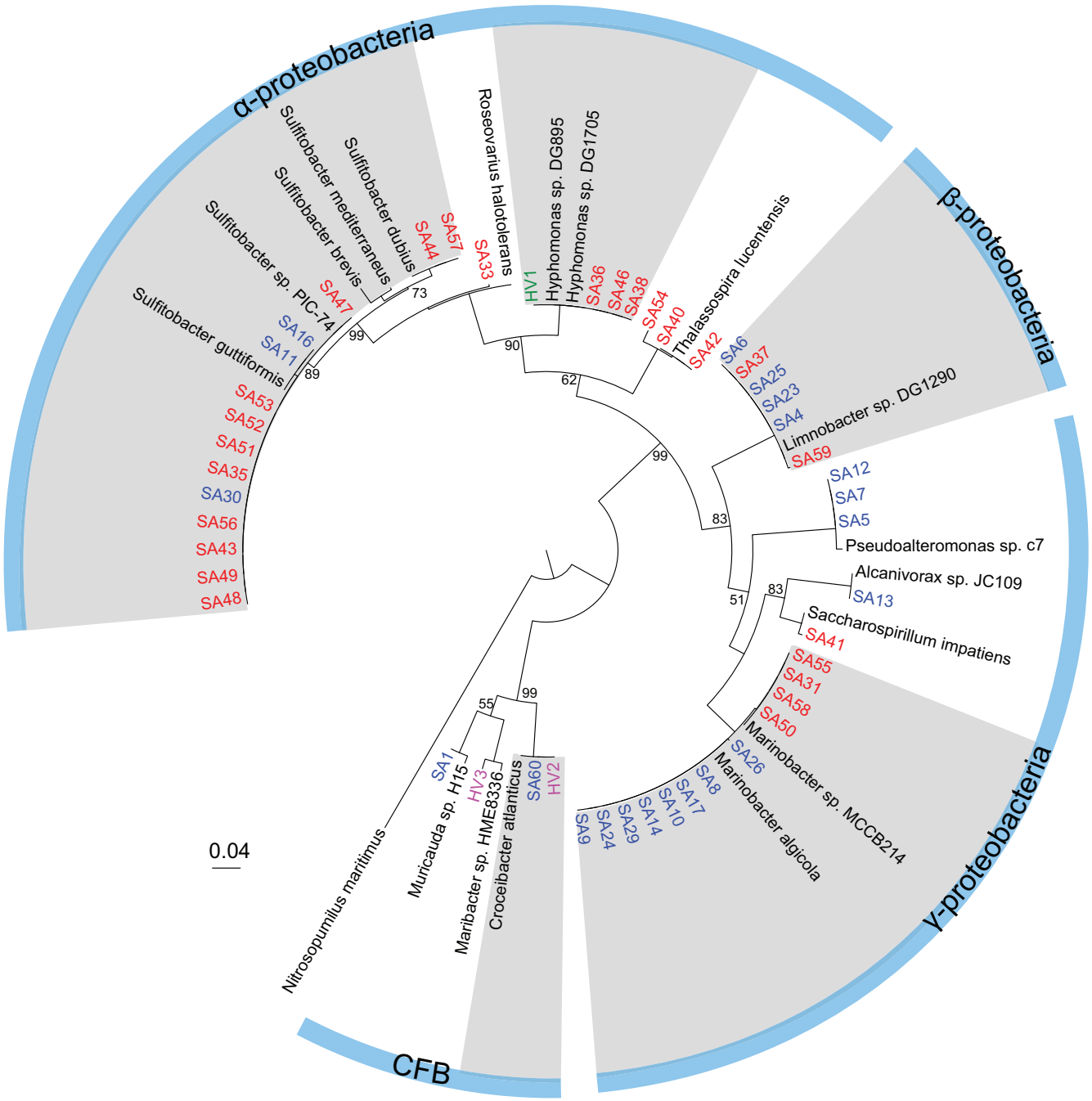
For RNA, cells were collected by sequential filtration on a Nitex screen of 53 μm pore-size to remove large particles, 142 mm 2.0 μm pore-size polycarbonate filters (mostly eukaryotic), and 142 mm 0.2 μm pore-size Supor filters (mostly prokaryotic). Results shown are from the combined eukaryotic and prokaryotic size fractions. Filters were flash frozen in liquid nitrogen and subsequently stored at −80 °C until processing.

RNA extraction and DNA removal were performed as previously described for environmental metatranscriptomics[42,43] with the following modifications: lysis in 10 ml of Ambion lysis buffer (AM8540G) + 0.5 ml each of 0.5 and 0.1 zirconia beads. rRNAs were selectively removed using a subtractive hybridization method[40] with biotinylated rRNA probes specific to the samples (that is, bacterial and archaeal 16S and 23S and eukaryotic 18S and 23S). Probe-bound RNAs were removed with strepdavidin-coated magnetic beads (New England Biolabs). rRNA-depleted samples were then linearly amplified using a MessageAmp II-Bacteria kit (Ambion; Life Technologies). Amplified mRNAs were then converted into cDNAs for Illumina sequencing using a Superscript III First-Strand Synthesis System (Invitrogen; Life Technologies) followed by the NEBNext mRNA Second Strand Synthesis Module (New England Biolabs). cDNAs were then purified using a QIAquick PCR purification kit (Qiagen) followed by ethanol precipitation. Purified cDNAs were sheared to ~200–250 bp fragments and HiSeq libraries (Illumina) were constructed for paired-end (2 × 150) sequencing using an Illumina HiSeq 2500 platform. After sequencing, paired-end Illumina reads were joined using a PANDAseq assembler[44], and paired reads were trimmed using FASTX-Toolkit (http://hannonlab.cshl.edu/fastx_toolkit/).

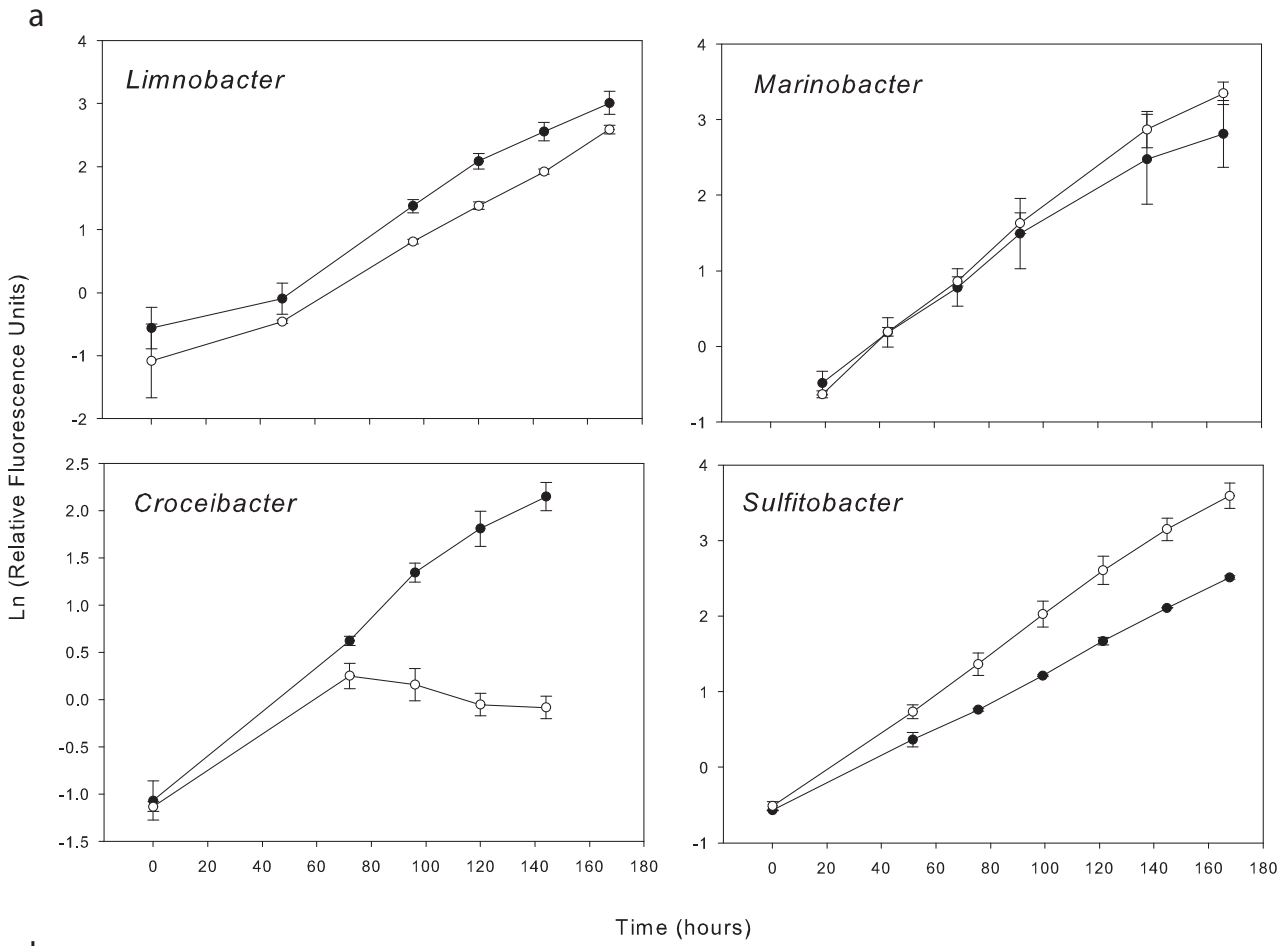
For the metatranscriptome analyses, rhizobial IAA biosynthesis proteins with experimentally verified functions were identified and used to identify homologues in SA11 and from Roseobase (http://www.roseobase.org) using BLASTp. Proteins with no homologues in Roseobase or SA11 were not included in the analysis (Extended Data Fig. 6). In addition, indole-3-acetaldehyde (IAAld) dehydrogenase, a protein commonly annotated as aldehyde dehydrogenase in rhizobia, was not included in the analysis because of the presence of several homologues within each *Roseobacter* genome that are probably involved in other pathways besides IAA biosynthesis. Therefore, our analysis in Fig. 3b probably represents an underrepresentation of IAA biosynthesis transcripts in the North Pacific. On average, each station had an estimated 3.2×10^{11} transcripts per litre based on the recovery of the internal standard reads after sequencing, suggesting IAA transcripts recovered in our analyses represent ~0.01% of the total transcripts. The *Roseobacter* reference sequences identified above were used as the query for tBLASTn searches to identify transcripts representing *Roseobacter*-clade IAA biosynthesis genes in our North Pacific metatranscriptomes and three publicly available metatranscriptomes from the North Pacific Gyre, Monterey Bay (California), and the California Coastal system (NCBI accession numbers PRJNA244754, PRJNA183166, and PRJNA268385, respectively). Only reads with ≥60% sequence identity and ≥140 bp of the read length aligning to the query were included in our final analysis. Transcript concentrations in seawater for the North Pacific metatranscriptomes were calculated on the basis of the recovery of the internal standard reads. Percentage IAA biosynthesis transcription was calculated by dividing the number of reads from each pathway by the total. The lack of complete metadata for the public data sets prevented the calculation of accurate transcripts per litre.

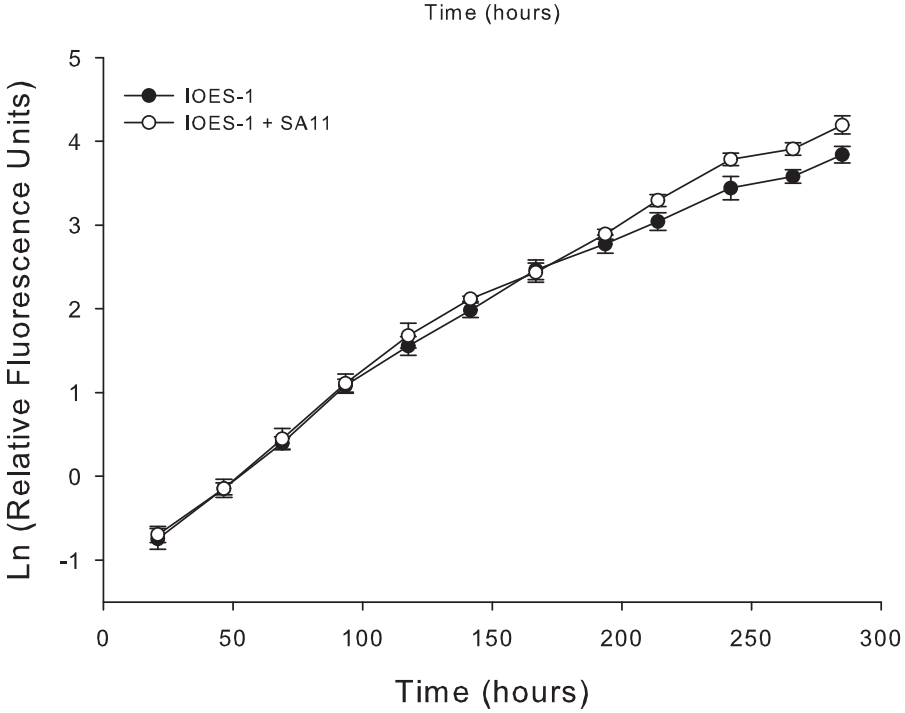
For targeted metabolite analysis, samples were collected and treated as described in ref. 45. Standard curves for IAA and percentage recovery were determined as described for Aquil, except seawater from station ALOHA (station 16) was used as matrix. Because seawater samples were mainly collected and processed for targeted vitamin B detection[45], recovery of IAA was poor but relatively consistent (32 ± 7%). Concentrations detected were corrected using this percentage recovery. Detection and quantification were conducted as described in the Mass Spectrometry section.

31. Hubbard, K. A., Rocap, G. & Armbrust, E. Inter-and intraspecific community structure within the diatom genus *Pseudo-nitzschia* (Bacillariophyceae). *J. Phycol.* **44,** 637–649 (2008).
32. Guillard, R. R. L. in *Culture of Marine Invertebrate Animals* (Smith, W. L. & Chanley, M. H. eds) 29–60 (Springer US, 1975).
33. Brand, L. E., Guillard, R. R. L. & Murphy, L. S. A method for the rapid and precise determination of acclimated phytoplankton reproduction rates. *J. Plankton Res.* **3,** 193–201 (1981).
34. ZoBell, C. E. Studies on marine bacteria. I. The cultural requirements of heterotrophic aerobes. *J. Mar. Res.* **4,** 42–75 (1941).
35. Lunau, M., Lemke, A., Walther, K., Martens-Habbena, W. & Simon, M. An improved method for counting bacteria from sediments and turbid environments by epifluorescence microscopy. *Environ. Microbiol.* **7,** 961–968 (2005).
36. Tamura, K. *et al.* Mega5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* **28,** 2731–2739 (2011).
37. Iverson, V. *et al.* Untangling genomes from metagenomes: revealing an uncultured class of marine Euryarchaeota. *Science* **335,** 587–590 (2012).
38. Li, H. *et al.* The sequence alignment/map format and SAMtools. *Bioinformatics* **25,** 2078–2079 (2009).
39. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26,** 139–140 (2009).
40. Stewart, F. J., Ottesen, E. A. & Delong, E. F. Development and quantitative analyses of a universal rRNA-subtraction protocol for microbial metatranscriptomics. *ISME J.* **4,** 896–907 (2010).
41. Matsuda, F., Miyazawa, H., Wakasa, K. & Miyagawa, H. Quantification of indole-3-acetic acid and amino acid conjugates in rice by liquid chromatography-electrospray ionization-tandem mass spectrometry. *Biosci. Biotechnol. Biochem.* **69,** 778–783 (2005).
42. Gifford, S. M., Sharma, S., Rinta-Kanto, J. M. & Moran, M. A. Quantitative analysis of a deeply sequenced marine microbial metatranscriptome. *ISME J.* **5,** 461–472 (2010).
43. Poretsky, R. S. *et al.* Comparative day/night metatranscriptomic analysis of microbial communities in the North Pacific subtropical gyre. *Environ. Microbiol.* **11,** 1358–1375 (2009).
44. Masella, A. P., Bartram, A. K., Truszkowski, J. M., Brown, D. G. & Neufeld, J. D. PANDAseq: PAired-eND Assembler for Illumina sequences. *BMC Bioinform.* **13,** 31 (2012).
45. Heal, K. R. *et al.* Determination of four forms of vitamin B12 and other B vitamins in seawater by liquid chromatography/tandem mass spectrometry. *Rapid Commun. Mass Spectrom.* **28,** 2398–2404 (2014).

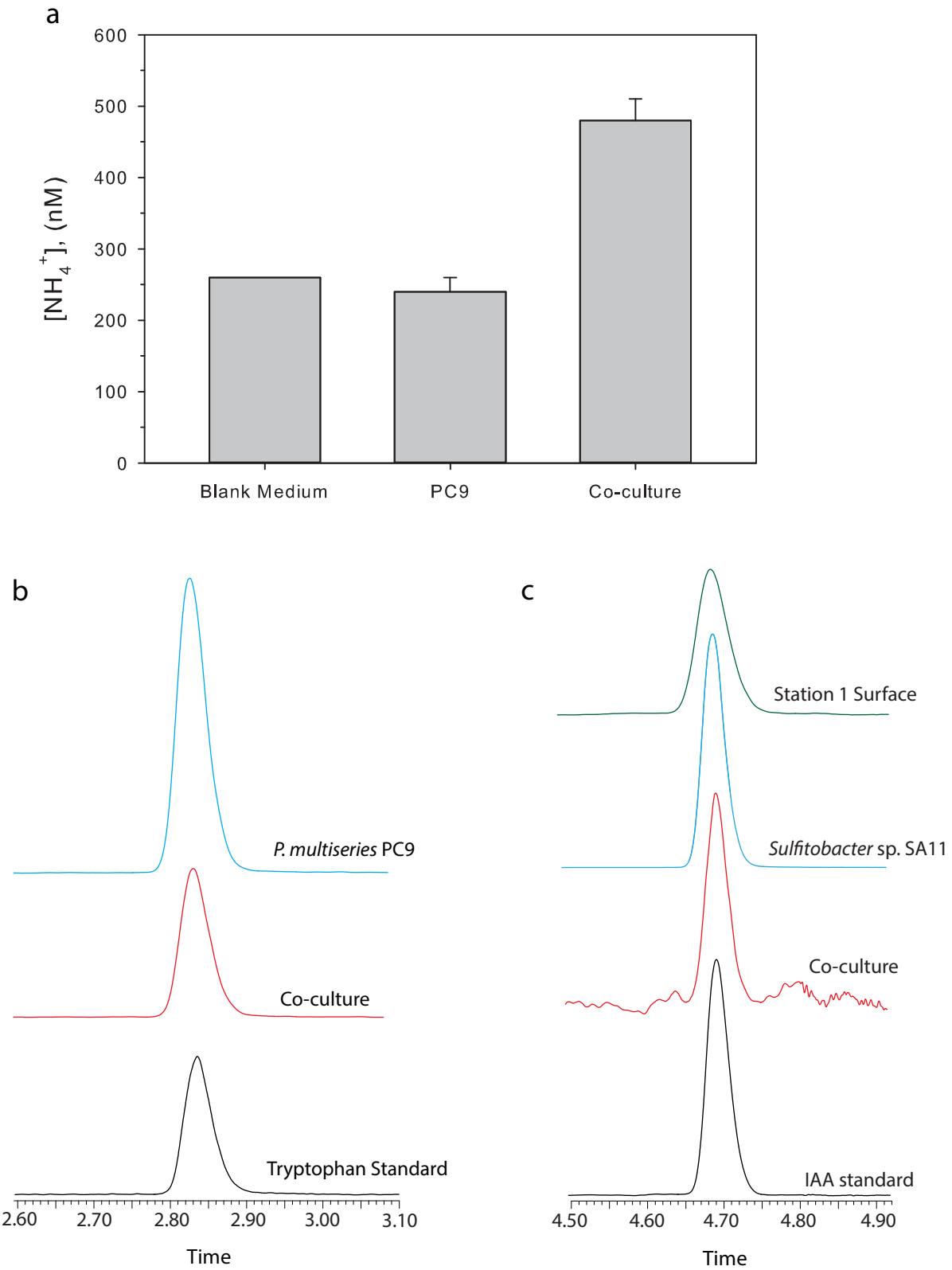


Extended Data Figure 1 | Phylogeny of *P. multiseries*-associated bacteria. Maximum likelihood tree showing the 16S rRNA phylogeny of all bacterial strains cultivated from *P. multiseries* isolates. Colour of bacterial strain designation indicates which isolate of *P. multiseries* a bacterial strain originated from: red, PC9; blue, PnCLNN-17; green, PC4; magenta, GGA2 (see Extended Data Table 1). Genera/clades that were considered to be associated with *P. multiseries* (contained two or more isolates from different diatom cultures with >99% 16S rRNA identity) are highlighted in grey. Bootstrap values greater than 50 are indicated at the branch points. Detailed information about each isolate is provided in Supplementary Information Table 1.

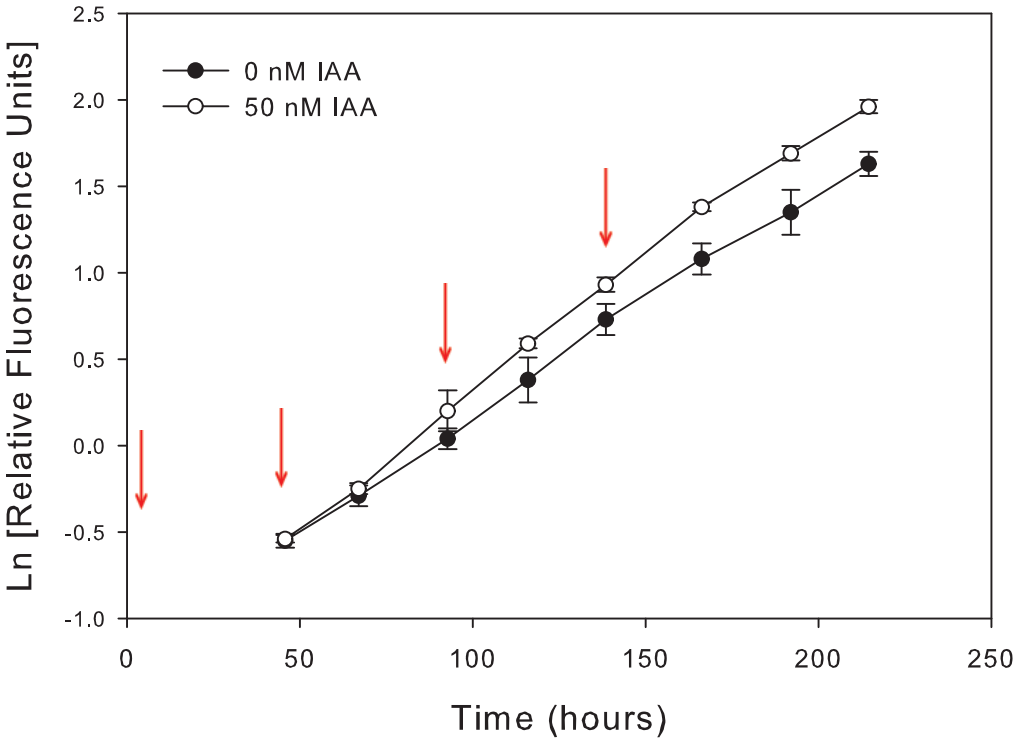




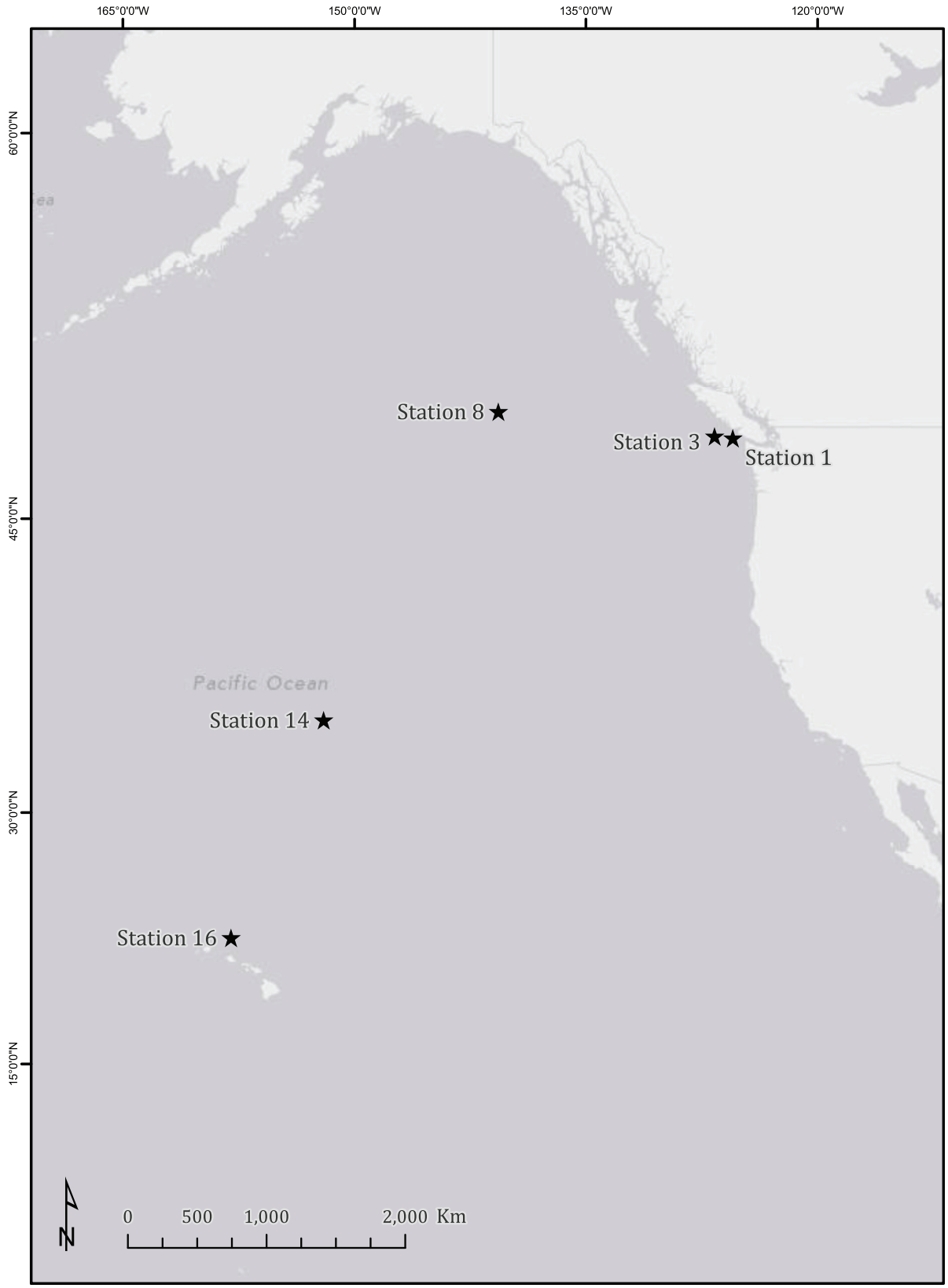
Extended Data Figure 2 | Effect of select bacterial strains on growth of *P. multiseries*. **a**, Growth of *P. multiseries* PC9 in the presence of different representative bacteria from its consortium (open circles) relative to axenic growth (filled circles). Bacterial representatives (*Limnobacter*, SA37; *Marinobacter*, SA14; *Croceibacter*, SA60; *Sulfitobacter*, SA52; see Extended Data Table 2) were inoculated at ~1 × 10^5 cells per millilitre relative to ~4,000 cells per millilitre axenic PC9. Error bars, s.d. from triplicate cultures. **b**, Growth of *P. multiseries* IOES-1 in axenic culture or with SA11. Error bars, s.d. from four replicates.



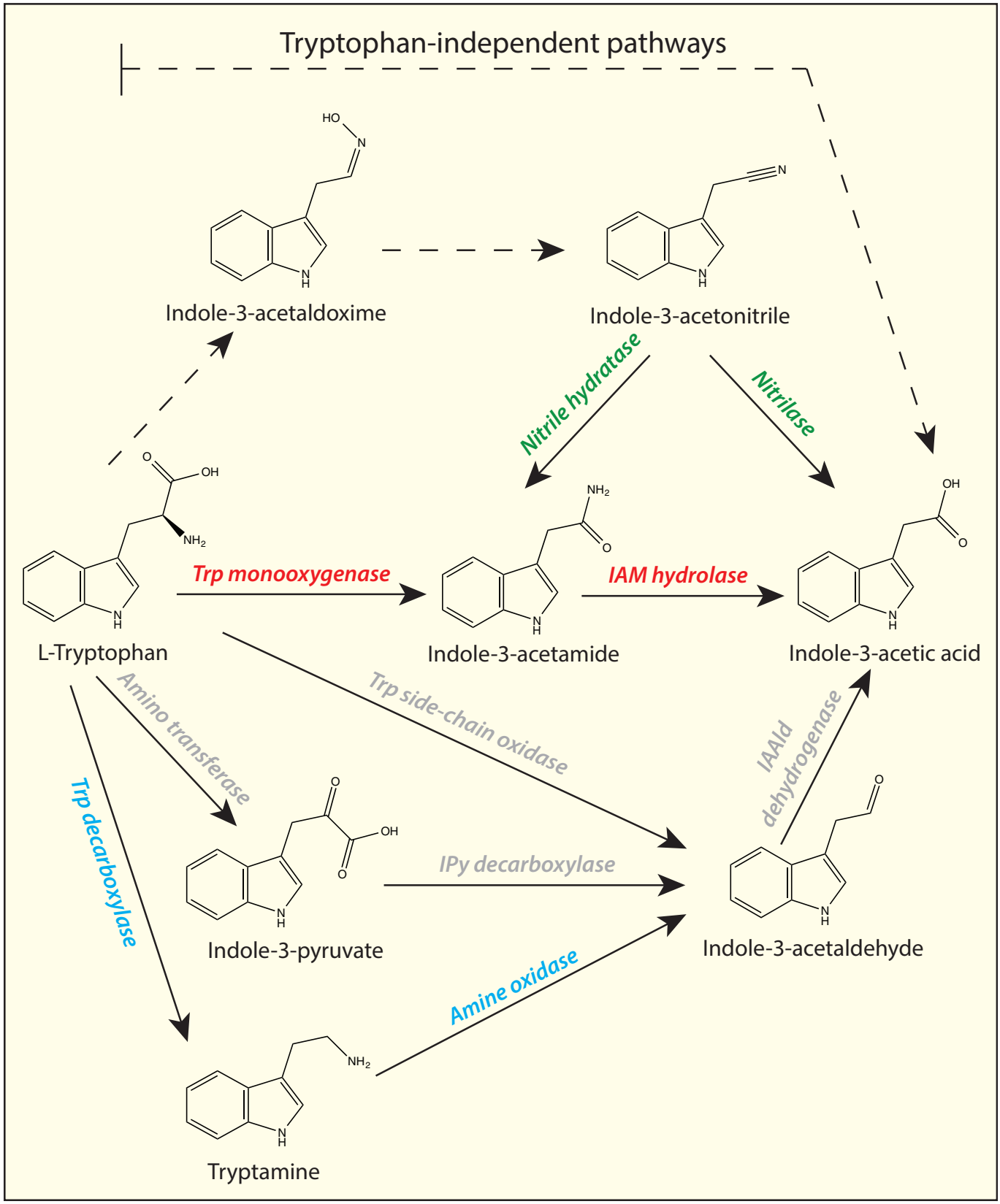
Extended Data Figure 3 | Select metabolite analyses from the *P. multiseries*–*Sulfitobacter* sp. SA11 co-culture and the environment. **a**, Dissolved ammonium concentrations in a medium blank, in axenic *P. multiseries* PC9, and in PC9 with SA11 (co-culture). Error bars, the range from duplicate supernatants. **b**, UPLC–ESI–MS/MS chromatograms of tryptophan in axenic PC9 or co-culture supernatants. Tryptophan was detected in positive ion mode by SRM from *m*/*z* 188 to 118. A 500 pM tryptophan standard is shown for retention time comparison. Tryptophan concentrations in the diatom monoculture and co-culture were 448 ± 106 pM and 202 ± 20 pM, respectively. **c**, UPLC–ESI–MS/MS chromatograms of IAA from surface water at station 1, SA11, and co-culture (with PC9) supernatants. IAA was detected in positive ion mode by SRM from *m*/*z* 176 to 130. A 0.5 pM IAA standard is shown for retention time comparison. IAA concentrations in the co-culture and SA11 monoculture were 6.1 ± 0.4 pM and 540 ± 280 pM, respectively.



Extended Data Figure 4 | Effect of multiple exogenous IAA additions on *P. multiseries* GGA2. Axenic GGA2 was grown in synthetic seawater media and 50 nM IAA was added at times indicated by the red arrows. Error bars, s.d. from six cultures.



Extended Data Figure 5 | Map of stations in the North Pacific Ocean where seawater samples were collected. Surface and chlorophyll maximum waters were collected for targeted metabolite analysis (all stations indicated) and metatranscriptomics (stations 1 and 3). Station 8 coincides with historic station PAPA and station 16 coincides with station ALOHA. The different stations exhibit dramatic differences in chemical and physical characteristics. For example, stations 1 and 3 are nutrient-rich coastal sites, station 8 is iron-limited, and stations 14 and 16 reside within the North Pacific Gyre and are oligotrophic. The map was created with Esri ArcGIS and Esri ArcMap 10.1 software.



Extended Data Figure 6 | IAA biosynthesis pathways in bacteria examined in the North Pacific Ocean metatranscriptomes. IAA biosynthesis in bacteria is divided into tryptophan-dependent and -independent pathways. Known bacterial enzymes involved in IAA biosynthesis all belong to the former (italic names). Dotted arrows represent biosynthetic steps with no known enzymes in bacteria[18]. Enzyme names are coloured according to the different pathways present in *Roseobacter* genomes: green, IAN pathway; red, IAM pathway; cyan, TAM pathway. Grey enzyme names were not included in our analysis because either no homologues were found in *Roseobacter* genomes or, in the case of IAAld dehydrogenase (belonging to the aldehyde dehydrogenase family), the presence of multiple homologues within a given genome that were involved in multiple pathways not related to IAA biosynthesis prevented our ability to decide on a reliable query for blast analysis. IAAld, indole-3-acetaldehyde; IPy, indole-3-pyruvate. This figure was modified from ref. 18.

Extended Data Table 1 | Diatom species and isolates used in this study

Species	Strain/Isolate name	Origin of Isolation	Isolation Date
Pseudo-nitzschia multiseries			
	PnCLNN-17[§]	Bay of Fundy, Canada	2007
	PC9[§]	Penn Cove, WA	2010
	PC4	Penn Cove, WA	2010
	GGA2	Golden Gardens, WA	2010
	IOES-1	East Sound, WA	2010
Thalassiosira pseudonana			
	CCMP1335	Long Island, NY	1958

[§] Strains from which most bacteria were isolated.

Extended Data Table 2 | Specific growth rate promotion of *P. multiseries* isolate PC9 in co-culture with different bacteria

Bacterial genus	Isolate Name	μ_{axenic} ± s.d.	$\mu_{coculture}$ ± s.d.	% change in μ
Sulfitobacter				
	SA11	0.69±0.03	1.06±0.05	35
	SA30	0.59±0.02	0.84±0.06	30
	SA44	0.40±0.01	0.49±0.01	18
	SA52	0.46±0.02	0.62±0.01	26
Phaeobacter				
	GS35	0.46±0.02	0.52±0.05	11
	GS36	0.46±0.02	0.49±0.01	6
Limnobacter				
	SA23	0.64±0.02	0.66±0.08	3
	SA37	0.59±0.02	0.54±0.09	-8
Marinobacter				
	SA14	0.64±0.1	0.70±0.06	8
Croceibacter				
	SA60	0.69±0.03	§	§

§ Growth rate could not be calculated, as the bacterium was algicidal.

Standard deviation values were calculated from biological triplicates.

Extended Data Table 3 | Specific growth rate promotion of different diatoms in co-culture with *Sulfitobacter* sp. SA11

Species	Culture name	μ_{axenic} ± s.d.	$\mu_{coculture}$ ± s.d.	% change in μ
Pseudo-nitzschia multiseries	PC9*	0.69±0.03	1.06±0.05	35
		0.59±0.02	0.87±0.03	32
		0.70±0.01	0.95±0.05	26
		0.75±0.03	0.94±0.04	20
		0.72±0.01	0.89±0.02	19
	GGA2	0.53±0.01	0.70±0.02	24
	PC4	0.47±0.02	0.48±0.01	2
	IOES-1	0.55±0.03	0.58±0.02	5
Thalassiosira pseudonana	CCMP1335	0.98±0.00	0.98±0.01	0

* Growth rate change ranged from 19-35% over five separate experiments.

Standard deviation values were calculated from biological triplicates except for IOES-1 (n = 4).

Extended Data Table 4 | The effect of single IAA additions on the growth of *P. multiseries* GGA2

IAA concentration	Mean Growth rate ± s.e.
0 nM	0.53 ± 0.01
1 nM	0.53 ± 0.01
50 nM	0.58 ± 0.01*
100 nM	0.58 ± 0.01*
250 nM	0.52 ± 0.02
10 μM	§

* Indicates statistically significant growth rate enhancement relative to 0 nM IAA.
§ Growth rate could not be calculated, as this concentration was inhibitory.

Standard error was calculated from $n = 6$ cultures.