



OPEN

## Seasonal bacterial niche structures and chemolithoautotrophic ecotypes in a North Atlantic fjord

Eric J. Raes<sup>1,2</sup>✉, Jennifer Tolman<sup>1</sup>, Dhvani Desai<sup>1,3</sup>, Jenni-Marie Ratten<sup>1</sup>, Jackie Zorz<sup>4</sup>, Brent M. Robicheau<sup>1</sup>, Diana Haider<sup>1,5</sup> & Julie LaRoche<sup>1</sup>

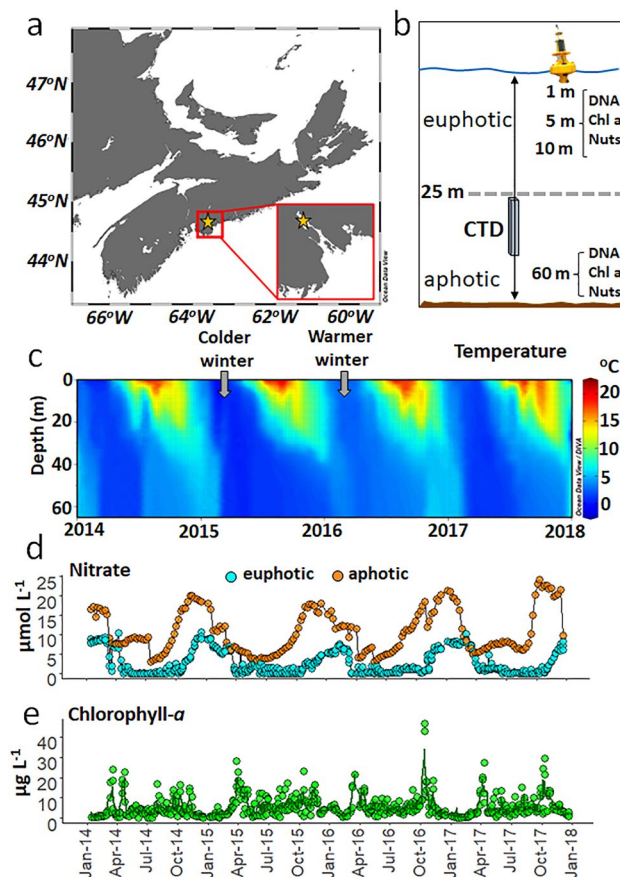
Quantifying the temporal change of bacterial communities is essential to understanding how both natural and anthropogenic pressures impact the functions of coastal marine ecosystems. Here we use weekly microbial DNA sampling across four years to show that bacterial phyla have distinct seasonal niches, with a richness peak in winter (i.e., an inverse relationship with daylength). Our results suggest that seasonal fluctuations, rather than the kinetic energy or resource hypotheses, dominated the pattern of bacterial diversity. These findings supplement those from global analyses which lack temporal replication and present few data from winter months in polar and temperate regions. Centered log-ratio transformed data provided new insights into the seasonal niche partitioning of conditionally rare phyla, such as Modulibacteria, Verrucomicrobiota, Synergistota, Deinococcota, and Fermentibacterota. These patterns could not be identified using the standard practice of ASV generation followed by rarefaction. Our study provides evidence that five globally relevant ecotypes of chemolithoautotrophic bacteria from the SUP05 lineage comprise a significant functional group with varying seasonal dominance patterns in the Bedford Basin.

Coastal regions and estuaries are the interface between land and ocean, and as biogeochemical hot spots they are responsible for ~ 19% of global oceanic net primary productivity<sup>1,2</sup>. However, along with all biomes on earth, coastal ecosystems are undergoing changes at different temporal scales both naturally and from anthropogenic pressures<sup>3,4</sup>. Measuring temporal change in these ecosystems is an issue of ongoing concern, given the valuable ecological, cultural and economic services they provide<sup>1,5</sup>. Monitoring programs that generate multi-disciplinary time-series datasets have been instrumental in detecting temporal patterns of change and seasonal shifts/cycles, alongside variability in environmental parameters and biological communities<sup>6,7</sup>. Furthermore, high-frequency sampling has allowed quantification of the impacts of short- and long-term stressors on coastal ecosystems specifically<sup>8–10</sup> and marine ecosystems in general<sup>11,12</sup>.

At the base of the marine food web, pro- and eukaryotic microorganisms sustain, inject and control the fluxes of energy and nutrients, and provide the organic matter for all higher trophic levels<sup>13</sup>. Despite considerable data gaps in certain coastal areas (e.g., the Atlantic South American coast and most of the Indian Ocean), the continued international effort in managing coastal time-series is delivering valuable mechanistic insights into the seasonal trends, periodicity, and phenology of marine prokaryotes<sup>9,14–20</sup>. This concerted effort highlights the need to identify seasonal patterns and changes in prokaryotic diversity as these will provide baselines to disturbances and long-term trends. Furthermore, while marine prokaryotes respond to external changes, they also play a crucial role in shaping their environment through the microbial loop<sup>21</sup>; any changes in their community structure could thus have important ramifications for biota higher up the food chain<sup>22,23</sup>.

In this study we present findings from a 4-year time series of bacterial 16S ribosomal RNA gene (rRNA) metabarcoding data in Bedford Basin, a temperate coastal embayment in the Northwest Atlantic Ocean. Samples were collected weekly from four depths congruent with physical and biochemical metadata. Traditionally, monitoring occurs at monthly or quarterly intervals, which may be insufficient to capture important transitions and disruptions to the community such as transient bloom periods or severe weather events (i.e., shorter-lived community dynamics<sup>24,25</sup>). The goal of this study was therefore to use high-resolution weekly sampling to describe the yearly reoccurring cycles, and hence in situ dynamics, of prokaryotic communities, and to describe important metabolic pathways. First, we calculated taxonomic diversity indices across time to test whether day length, rather

<sup>1</sup>Department of Biology, Dalhousie University, Halifax, NS B3H 4R2, Canada. <sup>2</sup>Flourishing Oceans, Mindereroo Foundation, Broadway, WA 6009, Australia. <sup>3</sup>Department of Pharmacology, Dalhousie University, Halifax, NS B3H 4R2, Canada. <sup>4</sup>Department of Geoscience, University of Calgary, Calgary, AB T2N 1N4, Canada. <sup>5</sup>Faculty of Computer Science, Dalhousie University, Halifax, NS B3H 4R2, Canada. ✉email: Julie.LaRoche@dal.ca

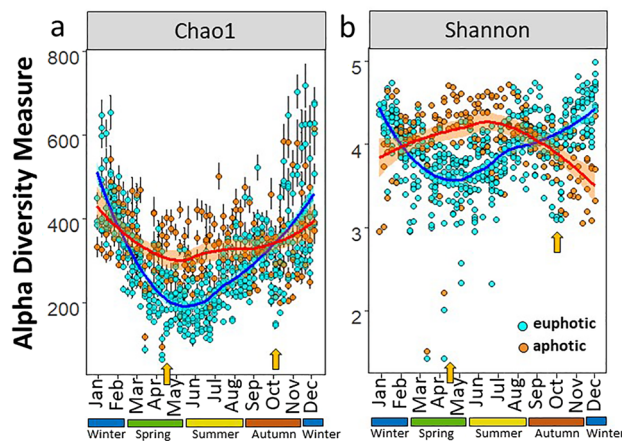


**Figure 1.** Our sampling site was located in Nova Scotia, Canada (a), and specifically within the Bedford Basin [inset of a; sampling site (star) is also known as HLO = 44.6936 latitude, -63.6403 longitude]. (b) Schematic representation of variables and conditions measured during this study, included: sampling depths for DNA, chlorophyll-*a* and concentrations for dissolved nutrients (nuts; which included  $\text{PO}_4^-$ ,  $\text{NO}_3^-$ ,  $\text{NO}_2^-$ ,  $\text{NH}_4^+$  and Si) along with CTD profiles. Weekly measurements from 2014 to 2017 included: temperature profiles (c), nitrate concentrations in the euphotic (1, 5 and 10 m; cyan circles) and in the aphotic zone (60 m; orange circles; d), and chlorophyll-*a* concentrations in the euphotic zone (the green line is the average for the three depths 1, 5 and 10 m; e). Grey arrows denote colder and warmer winters for 2015 and 2016, respectively, as detailed by Haas et al.<sup>32</sup>. Image credit for yellow buoy; Robert Brazzell.

than temperature and nutrients alone, was the environmental variable most effectively describing changes in prokaryotic alpha and beta diversities over an annual cycle. Next, we aimed to identify whether there were marked seasonal changes in abundance for different phyla, and if these results could be further supported by metabolic pathways inferred from PICRUSt2 analyses of 16S rRNA gene sequences<sup>26</sup>. Finally, we tested whether different ecotypes within one of the diverse prokaryotic clades observed at our site exhibited seasonality over an annual cycle<sup>27</sup>. In this regard, we investigated the chemolithoautotrophic bacterial SUP05 clade, which comprise a significant functional group in suboxic ( $< 100 \mu\text{mol O}_2 \text{ L}^{-1}$ ) dark (aphotic) coastal waters<sup>28</sup>, such as the Bedford Basin.

## Results and discussion

**Seasonal changes in a North Atlantic embayment.** The Bedford Basin is a coastal embayment with a maximum depth of 71 m and is connected via a narrow (500 m) and shallow (20 m) sill to the Scotian Shelf in the northwest Atlantic (NWA; Fig. 1a,b). Lower density surface waters flow out to the NWA, while deeper, slightly more saline waters flow into the basin over the sill<sup>29</sup>. Lowest salinity values during our study (25.40) were measured at the surface, and highest values (31.58) near the bottom. The average water column salinity was  $30.09 \pm 0.88$  and was not significantly different between years from 2014 to 2017 (Wilcoxon tests  $p > 0.1$ ; Supplementary Table 1). The euphotic depth is shallower than 30 m year-round, and like the Scotian Shelf, the basin is subject to distinct temperate seasonal cycles<sup>30</sup> (Fig. 1d). As is typical of other northern temperate fjords, the Bedford Basin also experiences seasonally defined water column stratification accompanied by spring (~ April) and autumn (~ September) phytoplankton blooms (Li<sup>31</sup> and Fig. 1c,e). This seasonal cycling also influences the available absolute macronutrient (N and P) concentrations as well as their stoichiometry, with a depletion of N relative to P in summer (Supplementary Fig. 1). Furthermore, the suboxic conditions that develop in the deep water during late autumn ( $< 100 \mu\text{mol O}_2 \text{ L}^{-1}$ ; below 30 m and Supplementary Fig. 2) are generally dissipated by



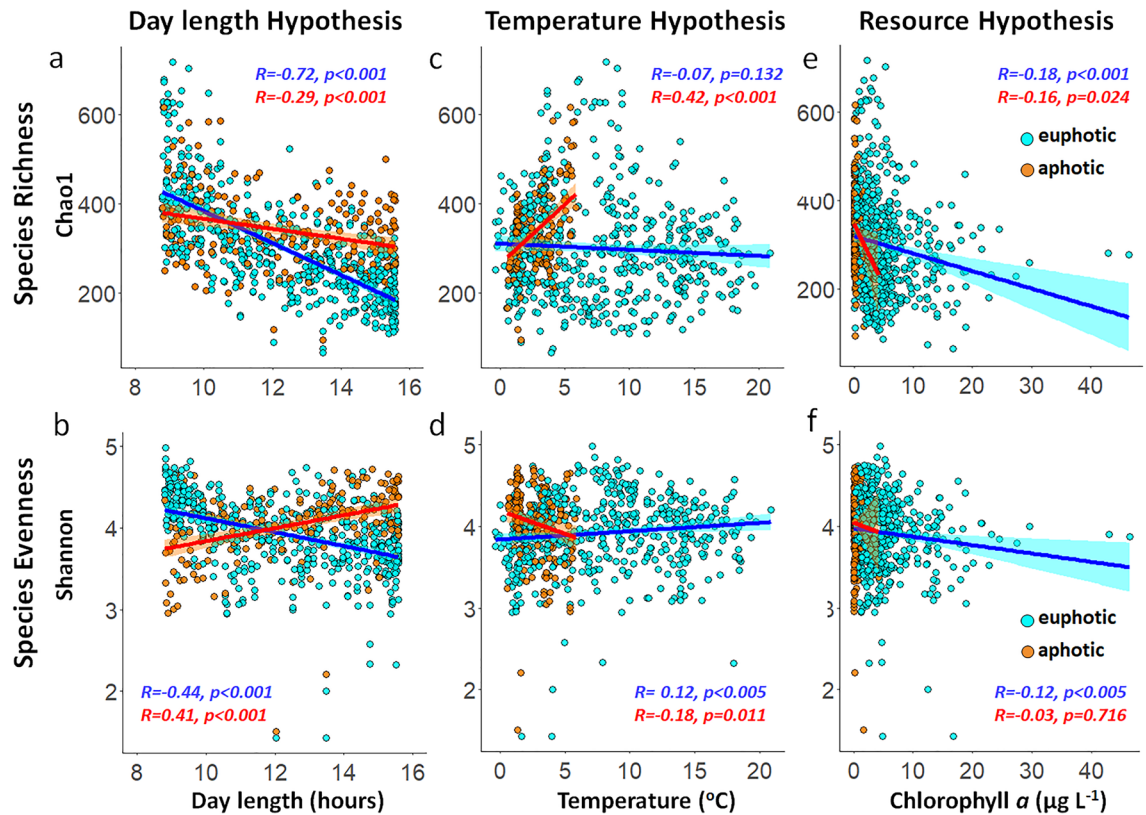
**Figure 2.** Seasonal trends for alpha diversity indices in the Bedford Basin from 2014 to 2017. Chao1 diversity (a; with standard errors as defined for the Chao1 model for estimating richness), and Shannon diversity (b). Orange arrows indicate low alpha diversity during the spring and autumn blooms. Samples from the euphotic (1, 5 and 10 m; cyan circles) and aphotic zone (60 m; orange circles) are colour coded. Loess regression lines are fitted for each zone and the inflection points on the regression lines show a minimum in April/May and in October for both indices.

the convective mixing of the entire water column that results from colder temperatures, replenishing dissolved oxygen in the deep water and nutrients in the surface waters.

**Alpha diversity peaks in winter and Beta diversity follows a resilient multi-annual cycle.** Across the four years, both Chao1 (species richness) and Shannon (evenness) diversity indices for prokaryotes showed a clear seasonal trend in the upper 10 m of the water column, with higher diversities in the colder, nutrient-rich, winter months (Fig. 2). Lowest alpha diversity was found during early summer (Chao1) and the productive spring bloom (Shannon; Fig. 2). Changes in alpha diversity at 60 m depth, although still statistically significant, were less pronounced compared to the surface waters (quadratic fits in the aphotic zone had lower  $R^2$ ; see Supplementary Table 2 and Supplementary Fig. 3). This aligns with the previously observed lack of global diversity gradient in the aphotic pelagic zone<sup>33</sup>. The weak seasonal trend in the aphotic waters could, in our study, be best explained by an increased stability of the deep water compared to the surface; i.e., smaller seasonal variations in temperature and minimal in situ autotrophic production which have been suggested to be the two main variables driving changes in richness trends (i.e., the kinetic energy<sup>33</sup> and resource hypotheses<sup>34</sup>). Possible explanations for the high diversity in the euphotic zone during the colder winter months are (1) mixing of the water column will introduce species to the surface that usually reside at depth<sup>32</sup> and (2) an increased range of available nutrients (i.e.,  $\text{NO}_3^-$ ,  $\text{PO}_4^{3-}$  and Si) in the euphotic waters will support more ecological niches and therefore more species (resource heterogeneity). The lower diversity during the bloom periods could be related to the proportional increase of fast growing specialists such as Flavobacteria, which thrive in coastal marine waters during phytoplankton blooms<sup>35</sup>. Their temporal niche advantage is supported by their ability to express specific degradation and transport genes for algal-derived (exo)-polysaccharides<sup>36</sup>. However, further metabolic rate measurements are needed to test whether the lower prokaryotic diversity during the productive months leads to a lower functional diversity<sup>37</sup>.

Across the 4-year study, our results highlight a seasonal reversal in prokaryotic diversity in the upper water column relative to productive and warmer waters in summer. Similar peaks in diversity in winter have been observed in the temperate latitudes at a regional scale in the English Channel<sup>24</sup>, in polar regions such as the Fram Strait<sup>38</sup> and on a global scale<sup>39</sup>. The seasonal patterns observed in our data remained present at higher rarefaction depths (to 20,000 reads) confirming the robustness of these trends (Supplementary Fig. 3). All seasonal comparisons in the euphotic zone, with the exception of spring and summer for Chao1 and autumn and winter for the Shannon index, were significantly different from each other (Wilcoxon tests  $p < 0.05$ ; Supplementary Tables 3–4). Similar seasonal trends also remained regardless of rarefaction depths for the rare microbiome (prokaryotes which comprised  $< 1\%$  of the community; Supplementary Fig. 4).

Spatial patterns in prokaryotic richness and diversity have mainly been explained by two hypotheses: the kinetic energy hypothesis and the resource hypothesis. The former postulates that higher temperatures increase the metabolic rate of pathways, resulting in higher speciation rates and ultimately in higher alpha diversities (Sunagawa et al.<sup>40</sup> and Ibarbalz et al.<sup>33</sup>, and reviewed by Brown<sup>41</sup>). The latter suggests that more energy production will support a higher numbers of species<sup>34,42</sup>. The observation that day length may overrule the kinetic energy hypothesis (Fig. 3) is in contrast to the findings from a number of studies showing that temperature appears to be the strongest positive predictor of prokaryotic diversity in the global ocean (i.e., see Fuhrman et al.<sup>43</sup>, Ibarbalz et al.<sup>33</sup> and Mittelbach et al.<sup>34</sup>). Furthermore, the seasonal trends in our time-series data do not support the productivity/resource hypothesis which show that autotrophic productivity is positively correlated with prokaryotic richness (i.e., see Mittelbach et al.<sup>34</sup> and Raes et al.<sup>42,44</sup>; and Fig. 3). Instead, day length had the



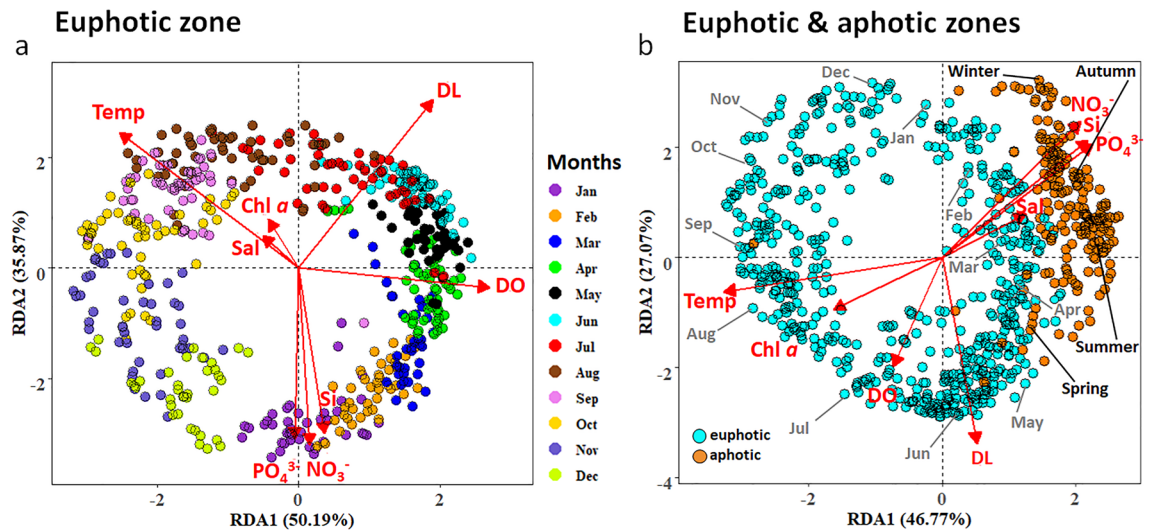
**Figure 3.** Alpha diversity trends in the Bedford Basin from 2014 to 2017. Species richness (Chao1) and species evenness (Shannon diversity) were correlated with day length (a,b), temperature (c,d), and chl-*a* concentrations as a proxy for biomass production (e,f). Linear regression lines are shown in blue for the euphotic zone and in red for the aphotic zone. Pearson correlation coefficients are indicated as *R* on the plots with their associated *p*-value. The ASV table was rarefied to 5000 reads prior to analyses, and seasonal patterns remained present independent of rarefaction depths (Supplementary Fig. 3).

strongest correlation coefficients for both species richness and evenness in the euphotic zone (Fig. 3). Multiple regression models (with temperature, chl-*a* and day length fitted last) also supported these findings and showed that day length contributed the largest proportion to the models in the euphotic zone (Supplementary Fig. 5). For completion we note that the richness peak between 5° and 10° C in the euphotic and aphotic zone (Fig. 3c,d) coincided with the winter months (Supplementary Fig. 6).

Redundancy analyses (RDAs) confirmed that day length had the strongest correlative power to explain variance in the prokaryotic community, similar to the results from Gilbert et al.<sup>24</sup> for the L4 time series. In our study, day length alone explained 91% of the variance in the prokaryotic community (Supplementary Table 5). Overall, regression fitting of single explanatory variables including temperature,  $\text{NO}_3^-$ ,  $\text{PO}_4^-$ , Si and DO were also all significantly correlated with the prokaryotic community composition ( $r^2$ : 0.83, 0.71, 0.67, 0.62, 0.51 respectively; all  $p < 0.001$ ; Supplementary Table 5). Both chl-*a* concentrations and salinity also revealed significant correlations with the prokaryotic community composition, but each parameter contributed <8% of the variance ( $r^2$ : 0.07, 0.04 respectively; all  $p < 0.001$ ; Supplementary Table 5). RDA showed that beta diversity was largely explained by changes in day length (similar to alpha diversity), and that seasonality again seemed to overrule both the kinetic energy and the productivity/resource hypotheses. Partial Mantel tests were used in an attempt to disentangle the effects of day length and temperature on the prokaryotic beta diversity. The Mantel correlation statistics between beta diversity and temperature, controlled by day length ( $r$ : 0.494,  $p < 0.001$ ), and between beta diversity and day length, controlled by temperature ( $r$ : 0.437,  $p < 0.001$ ) were very similar. These results corroborate the findings from our RDA analyses where both temperature and day length revealed highly similar  $r^2$  values (Fig. 4a and Supplementary Table 5). Day length and temperature covary and the results from both the Mantel test and RDA analyses highlight the difficulty of disentangling these factors. While alpha diversity revealed a significant negative relationship with day length, the main driver for the beta-diversity pattern remains unresolved.

The constrained (RDA; Fig. 4a) and unconstrained (PCA; Supplementary Fig. 7) prokaryotic community analyses in the euphotic zone displayed a clear seasonal pattern, and no significant differences were found within the euphotic zone (ANOSIM, all  $p$ -values > 0.05 between samples from 1, 5 and 10 m, nested within sampling month, across all four years; see Supplementary Table 6). Samples from the aphotic zone (60 m), however, were significantly different from the euphotic zone (ANOSIM  $p < 0.05$ ; see Supplementary Table 6). The communities in the aphotic and euphotic zones were more similar in winter, when a well-mixed water column (Fig. 1c) reduced the proportion of unique ASVs in the aphotic zone by half compared to summer (Supplementary Fig. 8).





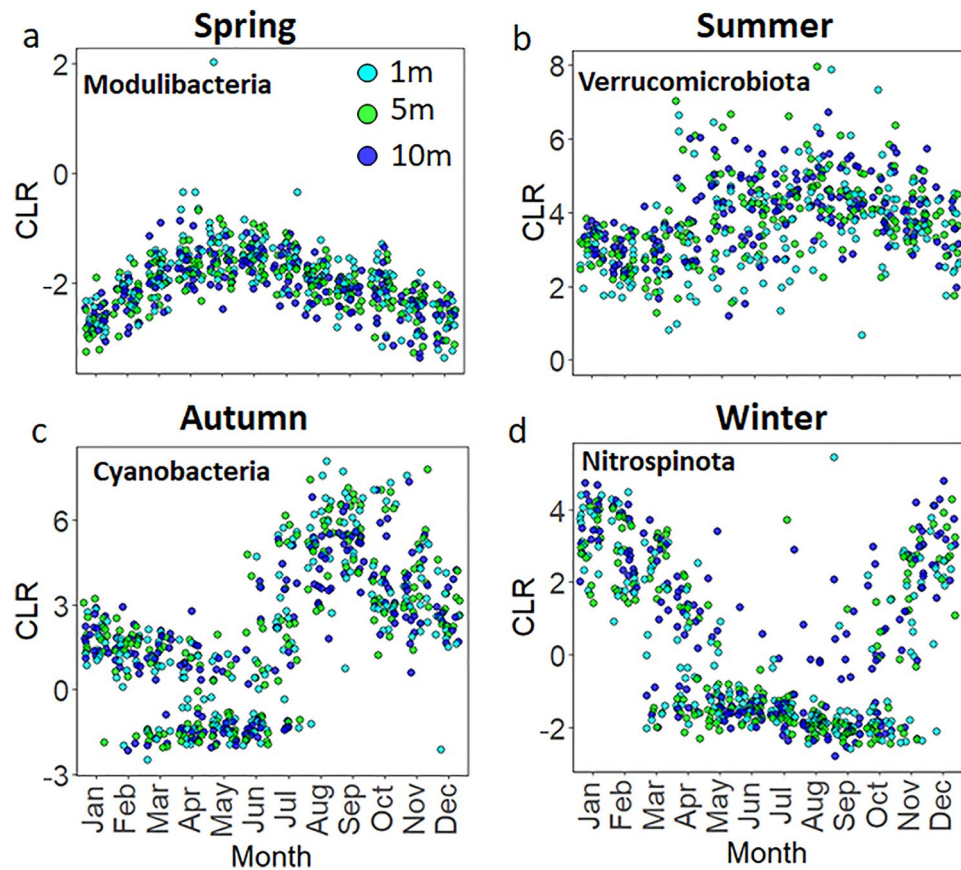
**Figure 4.** Redundancy analyses (RDA) to characterize the specific abiotic variables in the water column which exerted the largest influence on the prokaryotic community composition in euphotic zone only (a) and across all four depths (b). The ASV data were CLR transformed and environmental parameters were standardized with the ‘standardize’ function using decostand from the Vegan Package<sup>45</sup>. Months are coded with different colours (a) and sampling depths are colour coded in cyan (euphotic zone; 1, 5, 10 m) and orange (aphotic zone; 60 m; b). Abbreviated months are shown in grey, and seasons are highlighted in black for the aphotic zone (b). Vector abbreviations are temperature (temp), salinity (sal), chl *a* (chlorophyll-*a*), dissolved oxygen (DO), day length (DL), nitrate ( $\text{NO}_3^-$ ), silicate (Si), phosphate ( $\text{PO}_4^{3-}$ ).

The increase in  $\text{NH}_4^+$  concentrations in the aphotic zone from spring to autumn suggest the remineralisation of organic matter (Supplementary Fig. 2b and Haas et al.<sup>32</sup>). The export of the particle attached surface prokaryotic community to the deeper waters during the productive months may contribute to higher alpha diversities in summer in the aphotic zone. Sinking particles and faecal pellets are known to vertically connect prokaryotic communities in the water column<sup>46</sup>, and suggest that sinking particle-attached bacteria can continue to degrade the exported organic matter in the aphotic zone.

Monthly differences, within a sampling year were significantly greater than interannual differences (i.e., in the euphotic zone ANOSIM Global R-values averaged  $0.796 \pm 0.045$  ( $\pm$  sd) and Fig. 4a) between months across all years, while the interannual Global R-value was 0.091; Supplementary Table 7 and 8). The lower global R-values between years provide evidence that the seasonal cycle was consistent multi-annually (a resilient seasonal cycle; Fig. 4). These findings of a resilient seasonal cycle reaffirm the strong cyclic patterns previously recorded in the Bedford Basin by El-Swais et al.<sup>47</sup>, who investigated prokaryotic seasonality by targeting the V5 region of the 16S rRNA gene (100 bp) from one millilitre of formalin-fixed seawater using 45 samples spanning six years (2005–2010). Our study, however, used weekly samples ( $n = 792$ ) across four years, and the 16S rRNA gene V4-V5 region (412 bp) was amplified from 500 ml of filtered seawater, with sequences clustered at an ASV level rather than a 90% identity cut-off as done by El-Swais et al.<sup>47</sup>. The longer amplicon sequence combined with the ASV analyses enabled a more in-depth description of the temporal changes in the prokaryotic community and improved taxonomic assignment. Clear examples are the higher variability in winter (probably due to a higher alpha diversity or an artefact due to the lower concentrations of particulate organic matter (biomass) relative to spring and autumn; Supplementary Fig. 9), and the findings of temporal trends of conditionally rare taxa and ecotypes as shown below.

As discussed in Haas et al.<sup>32</sup>, notable events across our 4-years of weekly observations included: (i) a significantly colder winter period in 2015 (Wilcoxon test  $p < 0.05$ ); with lowest water temperatures measuring  $-0.35$  °C in the upper 10 m (Fig. 1c), resulting in stronger winter mixing relative to the other years; (ii) higher minimum temperatures in the winter of 2016 ( $1.51$  °C in the upper 10 m), resulting in reduced winter mixing relative to the other years (Fig. 1c); and (iii) short-lived (~3 to 4 weeks) intrusion events detected as relatively higher oxygenated waters which interrupt the stratified period (below 30 m; Supplementary Fig. 2). While these intrusions can impact the N-cycling fluxes mediated by ammonium-oxidising microorganisms in the deeper waters, they do not impact the vertical stratification of the water column<sup>32</sup>. Although we were able to detect these events as significant environmental disturbances, the recurring beta diversity patterns (Fig. 4) suggest that neither the intrusion events, nor aberrant colder winter temperatures impacted the prokaryotic community assemblages.

In summary, our results and those of others<sup>18,24</sup> support the premise that seasonal fluctuations dominate the pattern of prokaryotic diversity in temperate regions<sup>39</sup>. Although global ocean surveys (e.g., TARA Oceans expeditions) have observed that highest prokaryotic diversity occurs near equatorial latitudes<sup>33</sup>, the lack of seasonal observations, compounded by the difficulties of sampling during winter months in polar and temperate regions, bring into question our current perception of global microbial diversity patterns. Recent studies utilizing time-series data have shown that species richness is negatively correlated with day length over an annual cycle with a



**Figure 5.** Seasonal changes in bacterial phyla in the euphotic zone (1, 5 and 10 m) in Bedford Basin. Significance patterns and temporal niches were assessed with a permutation test (“multipatt” function from the Indicspecies package<sup>52</sup>; see Methods and Supplementary Table 9). Due to the compositional nature of the sequence data, the changes are visualized with CLR values rather than relative abundances.

highest richness peak in winter<sup>24,38</sup>. The finding that the relationship between temperature and diversity changes from a regional to a global scale re-emphasises the need to adjust our sampling designs and analyses by seasons if we want to truly understand the drivers of marine prokaryotic richness and its influence on the metabolic potential of the community<sup>48</sup>. These findings are especially important for temperate and high latitude regions which are undersampled in winter, whereas tropical and equatorial oceanic zones with modest variations in day length are consequently expected to be less impacted by seasonal change.

**Seasonal niche partitioning of bacterial phyla and functional pathways.** The increased recognition of the compositional nature of microbiome datasets<sup>49</sup> encouraged us to use the centred log-ratio (CLR) transformation to explore relative changes and seasonal trends in the prokaryotic community. Using CLR transformed data, we identified several bacterial phyla that have temporally defined niches in the euphotic and aphotic zones (Fig. 5; Supplementary Fig. 10 and Supplementary Tables 9 and 10). We also present the change in relative abundances of certain bacterial genera using rarefied data in contrast to the CLR transformed data (Supplementary Fig. 11). While the main seasonal trends remain the same for the abundant bacterial phyla (i.e., similar trends are seen for the Cyanobacteria and Nitrospinota in autumn and winter respectively when the data are rarefied or CLR transformed; Supplementary Fig. 11), the CLR transformed data revealed new insights into the seasonal niche partitioning of phyla with low relative abundances, thereby adding a new level of resolution for rarer taxa previously unavailable<sup>47</sup>. Clear examples, in spring and summer, were revealed for the bacterial phyla Modulibacteria and Verrucomicrobiota (Fig. 5), and Synergistota, Deinococcota, and Fermentibacterota (Supplementary Table 9). However, when the ASV table was rarefied and expressed as relative abundances, the above phyla (which contributed <0.1% of the relative bacterial abundance) either did not show any seasonal trends or were removed during rarefaction (Supplementary Fig. 12). Bacteria present in low relative abundance can play significant roles in the C and N-cycles; Musat et al.<sup>50</sup>, for example, demonstrated that an anaerobic phototrophic bacteria representing 0.3% of the community contributed a large fraction of the total C-fixation and  $\text{NH}_4^+$  assimilation rates (up to 70 and 40% respectively). CLR transformed data revealed the cyclic distribution pattern of conditionally rare taxa (or “microbial dark matter”<sup>51</sup>) that may be seasonally important in the Bedford Basin, although their specific roles in biogeochemical cycling would require additional studies.

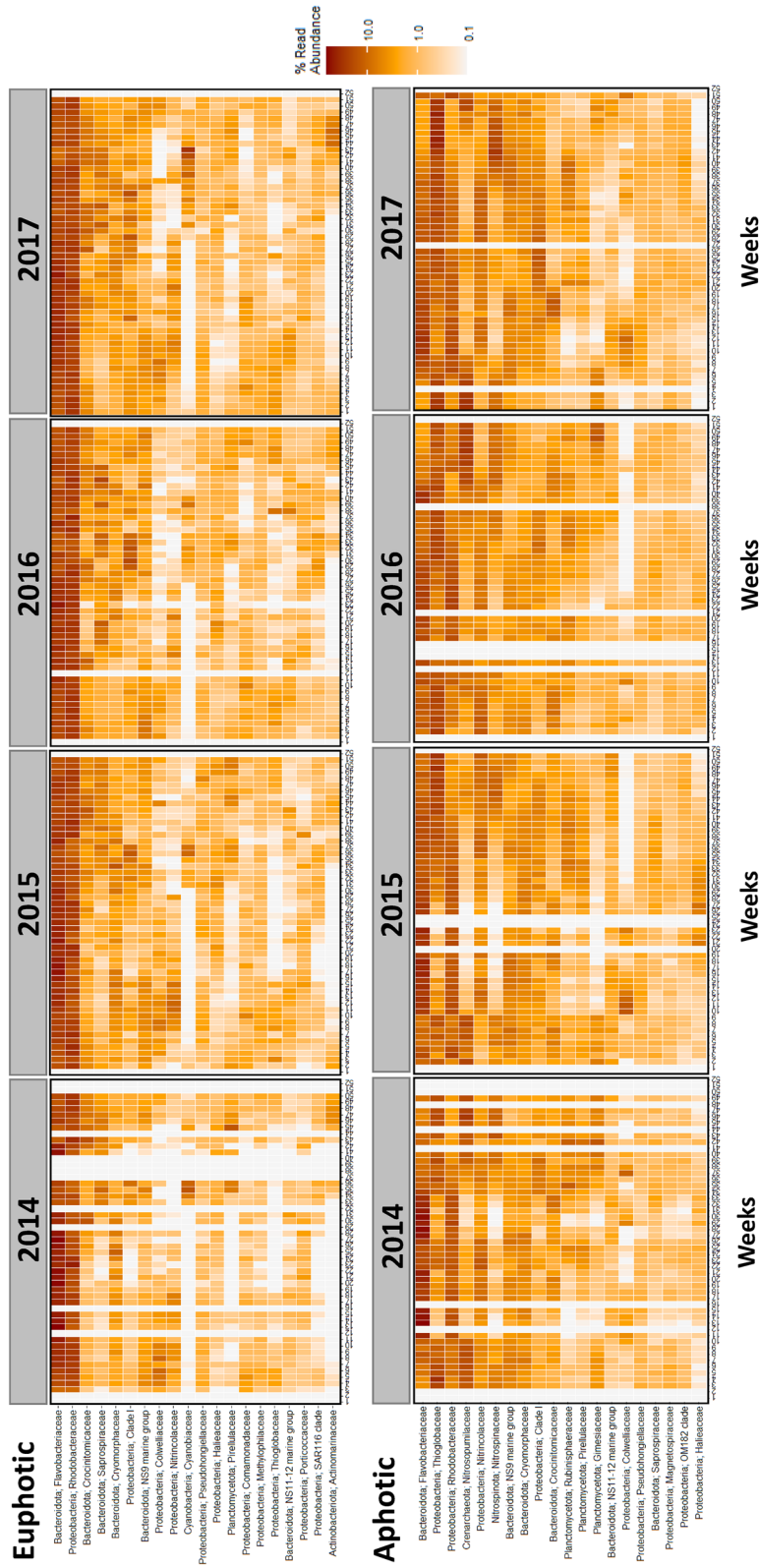
A detailed overview of seasonal trends of the top 20 most abundant taxa is shown in Fig. 6. Flavobacteriaceae and Rhodobacteraceae clearly dominated the euphotic zone throughout the year in terms of relative abundances. Crocinitomicaceae and Saprospiraceae revealed higher abundance in winter, whereas the long absences preceding seasonal blooms in Cyanobiaceae, Pirellulaceae, Thioglobaceae, SAR116 clade and Actinomarinaceae are noteworthy to mention (Fig. 6). In general, our results broadly mirror previous studies<sup>47</sup> that showed the association of the Flavobacteria such as *Polaribacter* and *Ulvibacter* with the spring bloom (Supplementary Fig. 13), and ammonia-oxidizing species as indicator ASVs in winter months<sup>32</sup>. A similar seasonal succession from Flavobacteria in spring to nitrifying organisms in winter has recently been shown in the Arctic and subarctic regions of the Atlantic Ocean<sup>38</sup>. Diminishing ice cover and the advection of warmer Atlantic water (referred to as Atlantification) have been hypothesised to trigger a shift in phytoplankton diversity and a proportional increase in Flavobacteria at high latitudes. Prolonged summer stratification and reduced wintertime convection due to Atlantification have been further shown to weaken deep-water renewal resulting in a decrease in associated “microbial recyclers”, such as nitrifying Archaea and Bacteria in the Arctic<sup>38</sup>. These microbial recyclers play an important role in the global N-cycle, and in the Bedford Basin, winter mixing was identified as a major variable controlling the abundance and diversity of nitrifying organisms<sup>32</sup>. The seasonal changes in photoautotrophic production (with low alpha diversity) and heterotrophic recycling (with high alpha diversity) periods in our study complement the observations from the Arctic Ocean<sup>38</sup>. The interconnection between nutrient depletion and autotrophic biomass production in spring, and nutrient recycling and renewal in winter is a natural cycle which can be impacted by increasing ocean temperatures which prolong the stratification of the mixed layer<sup>53</sup>. Time-series data, such as those used in this study, will allow a better characterization of how a warmer ocean will impact the biogeochemical fluxes mediated by micro-organisms in the North Atlantic but also in Arctic and subarctic regions, where similar seasonal trends have been observed<sup>38</sup>.

Although the results from metabolic inference-based prediction tools such as PICRUSt2<sup>26</sup> should be cautiously interpreted<sup>54</sup>, the approach has provided valuable mechanistic insights into the large-scale functional ecology of microbes in open oceanic ecosystems<sup>48</sup>. Here we used inferred metabolic predictions to show the positive correlation between seasonal spring blooms and prokaryotic lipid biosynthesis (Fig. 7a). Polysaccharide degradation pathways showed a positive correlation with both the spring and autumn blooms in the photic zone across four years (Fig. 7b). A decreasing trend with depth was noted with the highest number of predicted lipid biosynthesis and polysaccharide degradation pathways in the surface waters.

The positive relationships between lipid biosynthesis, the breakdown of organic matter, and the spring bloom offer the opportunity to track energy production in the Bedford Basin, and potentially other coastal embayments. Our results show that community-level metabolic data can be used as a qualitative tool for understanding changes in ecosystem functions. These insights, along with the temporal connectivity patterns of pro- and eukaryotic organisms, will allow us to generate new hypotheses regarding how local environmental perturbations, whether natural or anthropogenic, could impact the top-down and bottom-up controls in marine food webs<sup>55</sup>. This will open the possibility to test how changes at the lower microbially-driven trophic levels can propagate to the higher trophic levels and, how taxonomic changes could impact the fluxes of macro- and micronutrients mediated by the prokaryotic community.

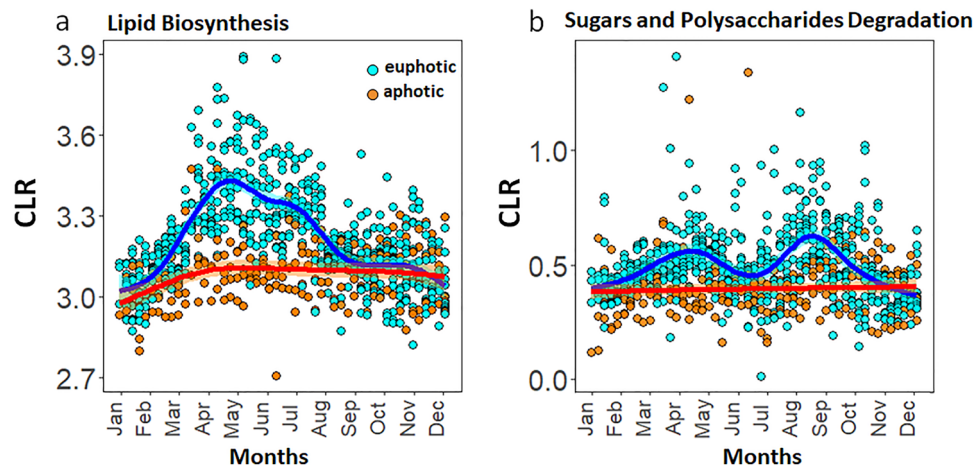
**Chemolithoautotrophy and seasonal ecotypes.** The majority of the metabolic predictions including lipid biosynthesis and polysaccharides-degradation pathways revealed the clearest change across seasons in the euphotic zone. While metabolic pathways also showed seasonal trends in the aphotic zone (Supplementary Table 13), we note two examples which illustrate the limitations in using PICRUSt2 to infer the metabolic potential of marine microbial communities. The first is the absence of nitrification pathways, which have been shown to be important in the Bedford Basin<sup>32</sup>, and the second is the absence of chemolithoautotrophy pathways. In the Bedford Basin, we noted a significant increase in the relative abundance of the chemolithoautotrophic SUP05 clade in the euphotic zone in winter (Fig. 8a), while overall highest relative abundances were observed throughout the year in the deeper suboxic waters of the basin in winter (up to 25% of the relative bacterial abundance). Members from the SUP05 clade are known chemolithoautotrophs and execute important roles in energy production in marine ecosystems including suboxic fjords<sup>28</sup>, deep ocean hydrothermal plumes<sup>56,57</sup>, and coastal shelf systems such as those off the West coast of Africa<sup>58</sup>.

The SUP05 clade of the Bedford Basin contained five distinct ecotypes (ASVs). Three out of the five ecotypes differed from one another by a single nucleotide within the 16S V4-V5 region (i.e., SUP05\_ASV1, SUP05\_ASV2 and SUP05\_ASV5). SUP05\_ASV3 and SUP05\_ASV4 also differed from one another by only one 16S V4-V5 nucleotide, but showed a difference of up to 18 nucleotides for the same gene region with the annually-persistent SUP05\_ASV1 (Fig. 8b-c and Supplementary File 1). One ecotype (SUP05\_ASV5) showed a clear temporal niche in autumn and two ecotypes (SUP05\_ASV3 and SUP05\_ASV4) occupied a niche in winter (Fig. 8b). The persistence of SUP05\_ASV1 through the spring and summer months could suggest that this ecotype is a mixotroph and competes with phytoplankton for light and nutrients. The potential for autotrophy in the aphotic zone is likely, as members of the SUP05 clade are known to contain large RuBisCo subunit genes (*rbcl*; a key enzyme during C-fixation; Swan et al.<sup>59</sup>). The co-occurrence of dissimilatory sulfur oxidation (adenosine-5'-phosphosulfate reductase alpha (*aprA*), reverse dissimilatory sulfite reductase (*rdsrA*) and autotrophic carbon fixation (RuBisCO) genes in members of the SUP05 clade suggest that these bacteria oxidise sulfur to support the autotrophic fixation of C<sup>59</sup>. In the Bedford Basin, Taguchi & Platt<sup>60</sup> showed that annual dark C-fixation in the whole water column contributes up to 25% (50 g C m<sup>-2</sup> year<sup>-1</sup>) of the estimated annual production, with highest dark C-fixation rates in winter<sup>61</sup>. We propose that the chemolithoautotrophic bacteria from the SUP05 lineage might be a prime candidate for the dark C-fixation rates measured 44 years ago by Taguchi & Platt<sup>60</sup>.

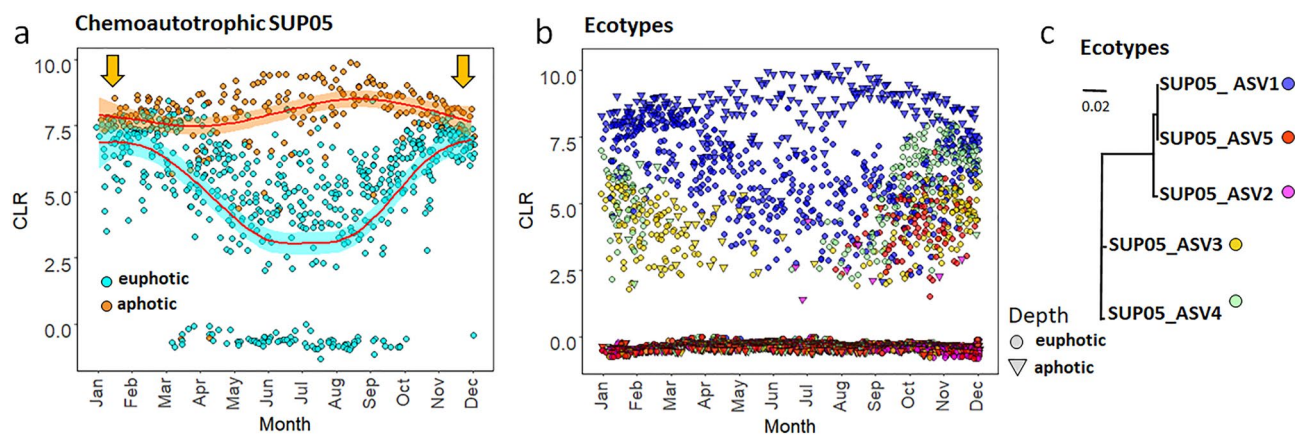


**Figure 6.** Temporal change in the Bedford Basin. Top panel shows the euphotic zone and the bottom panel the aphotic zone. ASVs were aggregated at a Family level and the Phylum level is added in front. The top 20 abundant taxa are shown and the colour scale is the % read abundance on a log10 scale. Note: Samples with a read depth < 5000 were removed, resulting in empty columns.





**Figure 7.** Seasonal trends in lipid biosynthesis pathways (associated with the spring bloom) and Sugars and Polysaccharides Degradation (associated with the spring and autumn blooms) in the Bedford basin from 2014 to 2017. Temporal pathways were identified using a permutation test (“multipatt” function from the Indicspecies package<sup>52</sup>; see “Methods” and Supplementary Table 12). Generalized additive mode smoothings are shown in blue and red (a,b).



**Figure 8.** Seasonal trends for the chemoautotrophic SUP05 lineage in the Bedford Basin from 2014–2017. The SUP05 reads are conglomerated at a genus level (a) to highlight the differences between the euphotic and aphotic zones, while (b) highlights the different SUP05 ecotypes from non-conglomerated reads. Winter mixing occurs during Nov–Feb shown by orange arrows on (a). Generalized additive mode smoothings are shown in red (a), note: analysis of CLR data reveals how ASVs behave relative to the per-sample average. Sampling depths are presented by different symbols on (b), and ecotypes are shown in different colours. The red arrow (b) highlights the temporal niche for ecotype SUP05\_ASV5. Neighbor-Joining tree with a Tamura-Nei genetic distance model for the five SUP05 ecotypes (c).

The metabolic flexibility of the SUP05 clade and the persistence of the SUP05\_ASV1 ecotype specifically in summer could also be due to (1) an increased ability to use a larger range of reduced organic and inorganic sulfur at sub-micromolar concentrations, and/or (2) the ability to store a larger proportion of intracellular sulfur when it respire oxygen<sup>62,63</sup>. While SUP05\_ASV1 remained the dominant ecotype throughout the year, SUP05\_ASV3 and SUP05\_ASV4 appeared to benefit from the reduction of light and oxygen<sup>32</sup> during the colder autumn and winter months. This temporal niche separation highlights that the SUP05 ecotypes in the Bedford Basin have adapted specific metabolic variations<sup>27</sup> which allow them to compete and/or coexist with their close relatives, as has been seen for the *Pelagibacter* SAR11 and the cyanobacteria *Prochlorococcus* in the open ocean<sup>64,65</sup>. Overall, these findings reaffirm the importance for considering dark C-fixation and sulfur cycling pathways during winter in coastal embayments in the NWA.

## Conclusions

Weekly metabarcoding sampling across four years revealed clear temporal niche structures with a richness peak in winter. Our results from a time series at a mid to high latitude suggest that seasonal variations, integrated here in a deterministic daylength parameter, had a dominant effect on shaping the pattern of bacterial diversity, overriding the contribution of kinetic energy or available resources. Our findings indicate that in temperate to

high latitude regions, the measured diversity must be interpreted in the context of an increasingly strong seasonal cycle at high latitude where highest diversity is observed at the shortest daylength period. Day length as a variable is an integration of seasonal processes which include, but are not limited to, net heat flux, stratification, nutrient depletion by primary producers, nutrient regeneration by benthic processes, and the release of POC and DOC from cell lysis and senescence. Altogether these processes lead to a smooth recurring pattern of increases and decreases in diversity not driven by temperature (kinetic energy hypothesis) per se or reflected in chl-*a* directly (resource hypothesis).

Furthermore, using the CLR transformation, we presented novel insights into the seasonal niche partitioning of conditionally rare phyla. In our study we provided evidence that five ecotypes of chemolithoautotrophic bacteria from the SUP05 lineage might be prime candidates for the dark C-fixation rates measured by Taguchi and Platt<sup>60</sup>. Insights into the seasonal changes of indicator species that perform distinct functions, such as ammonia oxidisers, sulfur reducers, bacterial primary producers such as Cyanobacteria, and those that form a close relationship with polysaccharide degradation, will allow us to further explore the correlative relationships of inferred functions. The inferred functions, we argue, have the potential to reveal new insights into how changes in bacterial diversity relate to energy production (in our study lipid biosynthesis pathways correlated with the productive spring bloom). Ultimately, we showed that metabarcoding data can be used to establish baselines of microbial diversity that can then be used to assess how natural and anthropogenic stressors impact the base of marine food webs.

## Materials and methods

**The Bedford Basin: sampling & oceanographic data.** Bedford Basin is an estuary that forms the inner portion of Halifax Harbour, Nova Scotia, Canada. The Bedford area is known as Kwipek (Head of the Tide) to the Mi'kmaq First Nation people. Weekly water samples were collected at the Compass Buoy Station (HL0; 44.6936, -63.6403; Fig. 1a,b) in Bedford Basin, in association with the Bedford Basin Monitoring Program (<https://www.bio.gc.ca/science/monitoring-monitorage/bbmp-pobb/bbmp-pobb-en.php>). Using Niskin bottles, discrete samples were taken at depths of 1, 5, 10, and 60 m (Fig. 1b). The sampling depths in our study were chosen to align with the long-term oceanographic Bedford Basin Monitoring Program which has been collecting weekly physical, chemical and biological data since 1992. Seawater samples were maintained in the dark at cool temperatures and transported to nearby Dalhousie University for processing. Seawater (500 mL) from each depth was filtered onto 0.2 µm polycarbonate Isopore filters (Millipore, Ireland) with a mesh prefilter (160 µm in 2014–2015; 330 µm in 2016–2017). Filters were stored at -80 °C. CTD profiles performed by BIO and the CERC.OCEAN group at Dalhousie University recorded environmental measurements in conjunction with the collection of water samples for DNA.

**DNA extraction and sequencing.** DNA was extracted from filtered cells using the DNeasy Plant Mini Kit (Qiagen, Germany) and the modified protocol described in Zorz et al.<sup>66</sup>. In brief, the lysis protocol was enhanced with a 5-min incubation with 50 µL lysozyme (5 mg/mL; Fisher BioReagents, UK) and a one hour incubation at 52 °C with Buffer AP1 and 45 µL Proteinase K (20 mg/mL; Fisher BioReagents, UK). At the conclusion of the manufacturer's extraction protocol, DNA was eluted twice in 50 µL AE Buffer (Qiagen, Germany) and quantified with the NanoDrop 2000c (Thermo Scientific, US). The V4–V5 region of the 16S ribosomal RNA was sequenced using the Illumina MiSeq (300 + 300 paired-end sequencing) at the Integrated Microbiome Resource (IMR; Dalhousie University, Halifax, NS); libraries were prepared by the IMR as in Comeau et al.<sup>67</sup> using primers 515FB = GTGYCAGCMGCCGCGGTAA and 926R = CCGYCAATTYMTTTRAGTTT<sup>68,69</sup>.

**Amplicon sequence variants (ASVs).** Raw Illumina reads were processed using the Microbiome Helper amplicon pipeline<sup>67</sup>. In brief, primer sequences were removed using *cutadapt*<sup>70</sup> and paired-end reads were stitched together using PEAR<sup>71</sup>. These sequences were imported into the QIIME2 2019.7 environment<sup>72</sup> and low-quality reads were removed; the remaining reads were denoised into Amplicon Sequence Variants (ASVs) using deblur<sup>73</sup> with a trim length of 350 bp. ASVs with sequence read frequencies less than [0.001 × mean sample depth] were attributed to sequencer bleed-through and removed. Taxonomy was assigned using the assignTaxonomy command in DADA2<sup>74</sup> which uses a full-length 16S Naïve Bayes-trained classifier based on the SILVA database (release 138<sup>75,76</sup>). Sequences which were assigned to chloroplasts and mitochondria were removed prior to the analyses.

**PICRUSt2.** The software PICRUSt2 (version 2.4.1b<sup>26,77</sup>) was used with default settings to infer the functional potential of the microbial communities across the 4 year time series study. Chloroplasts and mitochondrial sequences were removed from the data set prior to analyses. The average Nearest Sequenced Taxon Index (NSTI) score, based on 793 samples (covering the 4 depths), was 0.155 ± 0.046 (± SD). Pathways with < 10 reads were removed from the data set. The final predicted metabolic pathway abundance data were centered log-ratio (CLR) transformed to plot seasonal trends.

**Statistical analyses.** The Phyloseq (v.1.32.0; McMurdie and Holmes<sup>78</sup>), and microbiome (v. 1.7.21<sup>79</sup>) packages were used to analyse, visualize and plot the microbiome and physical and bio-chemical metadata using R (v4.0.2; Team<sup>80</sup>) in RStudio (v.1.3.1093). Plots were generated in R using ggplot2 or Ocean Data View version 4.6.5 unless stated otherwise<sup>81,82</sup>. Day lengths were calculated using the Geosphere package (v.1.5.10<sup>83</sup>) in R.

Statistical tests were conducted using the Vegan package version 2.5–6<sup>45</sup>. The Pearson's correlation coefficient was used as a statistical measure to test the strength of the linear relationships between our paired data and to compare our findings to the results from Fuhrman et al. (2008). The plyr package (v.1.8.4) was used to calculate

means and summarise the data<sup>84</sup>. The rarefaction curves and diversity measures were calculated using Phyloseq<sup>78</sup> (v.1.32.0). Alpha diversity trends did not change when ASV tables were subsampled to different sampling depths of 5000, 8000, 10,000 and 20,000 sequences per sample (Supplementary Figs. 3 and 4). Multiple regression models were conducted using the Vegan package<sup>45</sup> (v.2.5–6). The relaimpo package with the calc.relimp function was used to calculate the relative importance of each parameter to the model (v. 2.2.5; Groemping and Matthias<sup>85</sup>). Analysis of similarities (ANOSIM) were used to test whether we could identify statistical differences between the four depths, the years and different seasons using the PRIMER-e version 7.0.17<sup>86</sup>. Prior to ANOSIM analyses, the community matrices were CLR transformed and Aitchison distances were generated using the Vegan package in R<sup>45</sup>.

Redundancy analyses (RDAs; using CLR transformed ASV tables) were used to characterize the specific abiotic variables in the water column which exerted the largest influence on the prokaryotic community compositions. Environmental parameters were standardized with the ‘standardize’ function (variables were scaled to zero mean and unit variance) using decostand from the Vegan Package<sup>45</sup> and significant ( $p < 0.05$ ) environmental parameters were derived using the ‘envfit’ function in Vegan and overlaid as vectors to identify multiple explanatory variables between the estuarine zones<sup>45</sup>. The ‘indicpecies’ package (ver. 1.7.8;<sup>52</sup>) with the ‘multipatt’ function and 9999 permutations was used to identify prokaryotic indicators within each month. ASV tables were agglomerated at a Phylum level (using Phyloseq (v.1.32.0) and the abundance tables were then CLR transformed. Pearson’s phi coefficients of association<sup>87</sup> were calculated to determine significant indicator pathways across the different months. The phi coefficient was corrected using the function = “r.g.” to accommodate for the fact that some months had more sampling points than others<sup>88</sup>. To account for compositionality of sequencing data<sup>49</sup>, the trends in seasonal changes of the prokaryotic indicators were plotted using CLR values rather than relative abundances. The heatmap was produced using the *ampvis2* R package.

### Data availability

All physical and biogeochemical water column data are available through the Bedford Basin Monitoring Program (<https://www.bio.gc.ca/science/monitoring-monitorage/bbmp-pobb/bbmp-pobb-en.php>). Sequence data are archived at the NCBI under BioProject ID PRJNA785606. Sequences for the five SUP05 ecotypes are available under GenBank accession numbers MZ890602, MZ890603, MZ890604, MZ890605 and MZ890606. Scripts to reproduce the figures along with the ASV table, metadata and taxonomy are available at [https://github.com/EricRaes/Bedford\\_Basin\\_Time\\_Series](https://github.com/EricRaes/Bedford_Basin_Time_Series).

Received: 20 March 2022; Accepted: 25 August 2022

Published online: 12 September 2022

### References - citation 47 and 54 are the same

- Cloern, J. E., Foster, S. Q. & Kleckner, A. E. Phytoplankton primary production in the world’s estuarine-coastal ecosystems. *Biogeosciences* **11**, 2477–2501 (2014).
- Field, C. B., Behrenfeld, M. J., Randerson, J. T. & Falkowski, P. Primary production of the biosphere: Integrating terrestrial and oceanic components. *Science* (80-) **281**, 237–240 (1998).
- Gervais, C. R., Champion, C. & Pecl, G. T. Species on the move around the Australian coastline: A continental scale review of climate-driven species redistribution in marine systems. *Glob. Chang. Biol.* **685**, 171–181 (2021).
- Scanes, E., Scanes, P. R. & Ross, P. M. Climate change rapidly warms and acidifies Australian estuaries. *Nat. Commun.* **11**, 1–11 (2020).
- Rodrigues, J. G. *et al.* Marine and coastal cultural ecosystem services: knowledge gaps and research priorities. *One Ecosyst.* **2** (2017).
- O’Brien, T. D., Lorenzoni, L., Isensee, K. & Valdés, L. What are marine ecological time series telling us about the ocean. A status report. *IOC Tech. Ser.* **129**, 1–297 (2017).
- Ajani, P. A., Davies, C. H., Eriksen, R. S. & Richardson, A. J. Global warming impacts micro-phytoplankton at a long-term Pacific Ocean Coastal Station. *Front. Mar. Sci.* **7**, 878 (2020).
- Wiltshire, K. H. *et al.* Helgoland roads, North Sea: 45 years of change. *Estuaries Coasts* **33**, 295–310 (2010).
- Benway, H. M. *et al.* Ocean time series observations of changing marine ecosystems: An era of integration, synthesis, and societal applications. *Front. Mar. Sci.* **6**, 393 (2019).
- Wilson, J. M., Chamberlain, E. J., Erazo, N., Carter, M. L. & Bowman, J. S. Recurrent microbial community types driven by nearshore and seasonal processes in coastal Southern California. *Environ. Microbiol.* **23**, 3225 (2021).
- Keeling, C. D. *et al.* Atmospheric carbon dioxide variations at Mauna Loa observatory, Hawaii. *Tellus* **28**, 538–551 (1976).
- Doney, S. C., Fabry, V. J., Feely, R. A. & Kleyvas, J. A. Ocean acidification: The other CO<sub>2</sub> problem. *Ann. Rev. Mar. Sci.* **1**, 169–192 (2009).
- Falkowski, P. G. Evolution of the nitrogen cycle and its influence on the biological sequestration of CO<sub>2</sub> in the ocean. *Nature* **387**, 272–275 (1997).
- Brown, M. V. *et al.* Systematic, continental scale temporal monitoring of marine pelagic microbiota by the Australian Marine Microbial Biodiversity Initiative. *Sci. Data* **5**, 180130 (2018).
- Buttigieg, P. L. *et al.* Marine microbes in 4D—Using time series observation to assess the dynamics of the ocean microbiome and its links to ocean health. *Curr. Opin. Microbiol.* **43**, 169–185 (2018).
- Chow, C.-E.T. *et al.* Temporal variability and coherence of euphotic zone bacterial communities over a decade in the southern California Bight. *ISME J.* **7**, 2259–2273 (2013).
- Krabberød, A. K. *et al.* Long-term patterns of an interconnected core marine microbiota. *bioRxiv* 2021.03.18.435965. <https://doi.org/10.1101/2021.03.18.435965> (2021).
- Lambert, S. *et al.* Rhythmicity of coastal marine picoeukaryotes, bacteria and archaea despite irregular environmental perturbations. *ISME J.* **13**, 388–401 (2019).
- Auladell, A. *et al.* Seasonal niche differentiation among closely related marine bacteria. *ISME J.* <https://doi.org/10.1038/s41396-021-01053-2> (2021).
- Robicheau, B. M., Tolman, J., Bertrand, E. M. & LaRoche, J. Highly-resolved interannual phytoplankton community dynamics of the coastal Northwest Atlantic. *ISME Commun.* **2**(1), 1–12 (2022).
- Legendre, L., Rivkin, R. B., Weinbauer, M. G., Guidi, L. & Uitz, J. The microbial carbon pump concept: Potential biogeochemical significance in the globally changing ocean. *Prog. Oceanogr.* **134**, 432–450 (2015).

22. Hutchins, D. A. & Fu, F. Microorganisms and ocean global change. *Nat. Microbiol.* **2**, 1–11 (2017).
23. Gross, T., Rudolf, L., Levin, S. A. & Dieckmann, U. Generalized models reveal stabilizing factors in food webs. *Science (80-)*. **325**, 747–750 (2009).
24. Gilbert, J. A. *et al.* Defining seasonal marine microbial community dynamics. *ISME J.* **6**, 298–308 (2012).
25. Karl, D. M. & Church, M. J. Microbial oceanography and the Hawaii Ocean time-series programme. *Nat. Rev. Microbiol.* **12**, 699–713 (2014).
26. Douglas, G. M. *et al.* PICRUSt2 for prediction of metagenome functions. *Nat. Biotechnol.* **38**, 685–688 (2020).
27. Pachiadaki, M. G. *et al.* Charting the complexity of the marine microbiome through single-cell genomics. *Cell* **179**, 1623–1635 (2019).
28. Walsh, D. A. *et al.* Metagenome of a versatile chemolithoautotroph from expanding oceanic dead zones. *Science (80-)*. **326**, 578–582 (2009).
29. Shan, S., Sheng, J., Thompson, K. R. & Greenberg, D. A. Simulating the three-dimensional circulation and hydrography of Halifax Harbour using a multi-nested coastal ocean circulation model. *Ocean Dyn.* **61**, 951–976 (2011).
30. Petrie, B. & Yeats, P. Simple models of the circulation, dissolved metals, suspended solids and nutrients in Halifax Harbour. *Water Qual. Res. J.* **25**, 325–350 (1990).
31. WK, W. L. *The State of Phytoplankton and Bacterioplankton at the Compass Buoy Station: Bedford Basin Monitoring Program 1992–2013.* (Fisheries and Oceans Canada = Pêches et Océans Canada, 2014).
32. Haas, S. *et al.* Physical mixing in coastal waters controls and decouples nitrification via biomass dilution. *Proc. Natl. Acad. Sci.* **118**, e2004877118 (2021).
33. Ibarbalz, F. M. *et al.* Global trends in marine plankton diversity across kingdoms of life. *Cell* **179**, 1084–1097.e21 (2019).
34. Mittelbach, G. G. *et al.* What is the observed relationship between species richness and productivity?. *Ecology* **82**, 2381–2396 (2001).
35. Pernthaler, J. Competition and niche separation of pelagic bacteria in freshwater habitats. *Environ. Microbiol.* **19**, 2133–2150 (2017).
36. Teeling, H. *et al.* Substrate-controlled succession of marine bacterioplankton populations induced by a phytoplankton bloom. *Science (80-)* **336**, 608–611 (2012).
37. Vallina, S. M. *et al.* Global relationship between phytoplankton diversity and productivity in the ocean. *Nat. Commun.* **5**, 4299 (2014).
38. Wietz, M. *et al.* The polar night shift: Annual dynamics and drivers of microbial community structure in the Arctic Ocean. *bioRxiv* 2021.04.08.436999. <https://doi.org/10.1101/2021.04.08.436999> (2021).
39. Ladau, J. *et al.* Global marine bacterial diversity peaks at high latitudes in winter. *ISME J.* **7**, 1669–1677 (2013).
40. Sunagawa, S. *et al.* Structure and function of the global ocean microbiome. *Science (80-)*. **348**, 1261359 (2015).
41. Brown, J. H. Why are there so many species in the tropics?. *J. Biogeogr.* **41**, 8–22 (2014).
42. Raes, E. J. *et al.* Oceanographic boundaries constrain microbial diversity gradients in the South Pacific Ocean. *Proc. Natl. Acad. Sci.* **115**, E8266–E8275 (2018).
43. Fuhrman, J. A. *et al.* A latitudinal diversity gradient in planktonic marine bacteria. *Proc. Natl. Acad. Sci.* **105**, 7774–7778 (2008).
44. Raes, E. J., Bodrossy, L., van de Kamp, J., Bissett, A. & Waite, A. M. Marine bacterial richness increases towards higher latitudes in the eastern Indian Ocean. *Limnol. Oceanogr. Lett.* **3**, 10–19 (2018).
45. Oksanen, J. *et al.* The vegan package. *Commun. Ecol. Packag.* **10**, 719 (2007).
46. Mestre, M. *et al.* Sinking particles promote vertical connectivity in the ocean microbiome. *Proc. Natl. Acad. Sci.* **115**, E6799–E6807 (2018).
47. El-Swais, H., Dunn, K. A., Bielawski, J. P., Li, W. K. W. & Walsh, D. A. Seasonal assemblages and short-lived blooms in coastal north-west Atlantic Ocean bacterioplankton. *Environ. Microbiol.* **17**, 3642–3661 (2015).
48. Raes, E. J. *et al.* Metabolic pathways inferred from a bacterial marker gene illuminate ecological changes across South Pacific frontal boundaries. *Nat. Commun.* **12**, 2213 (2021).
49. Gloor, G. B., Macklaim, J. M., Pawlowsky-Glahn, V. & Egozcue, J. J. Microbiome datasets are compositional: And this is not optional. *Front. Microbiol.* **8**, 2224 (2017).
50. Musat, N. *et al.* A single-cell view on the ecophysiology of anaerobic phototrophic bacteria. *Proc. Natl. Acad. Sci.* **105**, 17861–17866 (2008).
51. Wong, H. L., MacLeod, F. I., White, R. A., Visscher, P. T. & Burns, B. P. Microbial dark matter filling the niche in hypersaline microbial mats. *Microbiome* **8**, 1–14 (2020).
52. De Cáceres, M. How to use the indicpecies package (ver. 1.7.1). *R Proj.* **2**, 29 (2013).
53. Hood, R. R. *et al.* Pelagic functional group modeling: Progress, challenges and prospects. *Deep Sea Res. Part II Top. Stud. Oceanogr.* **53**, 459–512 (2006).
54. Sun, S., Jones, R. B. & Fodor, A. A. Inference-based accuracy of metagenome prediction tools varies across sample types and functional categories. *Microbiome* **8**, 1–9 (2020).
55. Lynam, C. P. *et al.* Interaction between top-down and bottom-up control in marine food webs. *Proc. Natl. Acad. Sci.* **114**, 1952–1957 (2017).
56. Zhou, Z. *et al.* Gammaproteobacteria mediating utilization of methyl-, sulfur- and petroleum organic compounds in deep ocean hydrothermal plumes. *ISME J.* **14**, 3136–3148 (2020).
57. Dede, B. *et al.* Niche differentiation of sulfur-oxidizing bacteria (SUP05) in submarine hydrothermal plumes. *ISME J.* **16**(6), 1479–1490 (2022).
58. Lavik, G. *et al.* Detoxification of sulphidic African shelf waters by blooming chemolithotrophs. *Nature* **457**, 581–584 (2009).
59. Swan, B. K. *et al.* Potential for chemolithoautotrophy among ubiquitous bacteria lineages in the dark ocean. *Science (80-)* **333**, 1296–1300 (2011).
60. Taguchi, S. & Platt, T. Assimilation of <sup>14</sup>CO<sub>2</sub> in the dark compared to phytoplankton production in a small coastal inlet. *Estuar. Coast. Mar. Sci.* **5**, 679–684 (1977).
61. Platt, T. & Irwin, B. *Phytoplankton Production and Nutrients in Bedford Basin, 1969–1970.* (1971).
62. Vega, S. *et al.* Morphological plasticity in a sulfur-oxidizing marine bacterium from the SUP05 clade enhances dark carbon fixation. *MBio* **10**, e00216–e219 (2021).
63. Mattes, T. E., Ingalls, A. E., Burke, S. & Morris, R. M. Metabolic flexibility of SUP05 under low DO growth conditions. *Environ. Microbiol.* **23**, 2823 (2020).
64. Brown, M. V. *et al.* Global biogeography of SAR11 marine bacteria. *Mol. Syst. Biol.* **8**, 595 (2012).
65. Martiny, A. C., Coleman, M. L. & Chisholm, S. W. Phosphate acquisition genes in Prochlorococcus ecotypes: Evidence for genome-wide adaptation. *Proc. Natl. Acad. Sci.* **103**, 12552–12557 (2006).
66. Zorz, J. *et al.* Drivers of regional bacterial community structure and diversity in the Northwest Atlantic Ocean. *Front. Microbiol.* **10**, 281 (2019).
67. Comeau, A. M., Douglas, G. M. & Langille, M. G. I. Microbiome helper: A custom and streamlined workflow for microbiome research. *MSystems* **2**, e00127 (2017).
68. Parada, A. E., Needham, D. M. & Fuhrman, J. A. Every base matters: Assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environ. Microbiol.* **18**, 1403–1414 (2016).



69. Walters, W. *et al.* Improved bacterial 16S rRNA gene (V4 and V4–5) and fungal internal transcribed spacer marker gene primers for microbial community surveys. *mSystems* **1**, e00009–15 (2016).
70. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10 (2011).
71. Zhang, J., Kobert, K., Flouri, T. & Stamatakis, A. PEAR: A fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics* **30**, 614–620 (2014).
72. Bolyen, E. *et al.* Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* **37**, 852–857 (2019).
73. Amir, A. *et al.* Deblur rapidly resolves single-nucleotide community sequence patterns. *MSystems* **2**, 191–16 (2017).
74. Callahan, B. J. *et al.* DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581–583 (2016).
75. Wang, Q., Garrity, G. M., Tiedje, J. M. & Cole, J. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.* **73**, 5261–5267 (2007).
76. Quast, C. *et al.* The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res.* **41**, D590–D596 (2012).
77. Langille, M. G. I. *et al.* Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nat. Biotechnol.* **31**, 814–821 (2013).
78. McMurdie, P. J. & Holmes, S. phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS ONE* **8**, e61217 (2013).
79. Lahti L. & Shetty, S.A. *Tools for Microbiome Analysis in R. Microbiome Package Version 1.7.21. R/Bioconductor* <http://microbiome.github.com/microbiome>. (2017).
80. Team, R. C. R: *A Language and Environment for Statistical Computing*. (2013).
81. Wickham, H. ggplot2. *Wiley Interdiscip. Rev. Comput. Stat.* **3**, 180–185 (2011).
82. Schlitzer, R. *Ocean Data View. 2018*. Available *adv. awi*. (2015).
83. Hijmans, R. J., Williams, E., Vennes, C. & Hijmans, M. R. J. Package 'geosphere'. in *Spherical Trigonometry*. Vol. 1 (2017).
84. Wickham, H. The split-apply-combine strategy for data analysis. *J. Stat. Softw.* **40**, 1–29 (2011).
85. Groemping, U. & Matthias, L. Package 'relaimpo'. (2021).
86. Clarke, K. R. & Gorley, R. N. *Primer. Prim. Plymouth* (2006).
87. Chytrý, M., Tichý, L., Holt, J. & Botta-Dukát, Z. Determination of diagnostic species with statistical fidelity measures. *J. Veg. Sci.* **13**, 79–90 (2002).
88. Tichý, L. & Chytrý, M. Statistical determination of diagnostic species for site groups of unequal size. *J. Veg. Sci.* **17**, 809–818 (2006).

## Acknowledgements

Financial support for this work was provided through the Ocean Frontiers Institute (OFI) via an International Postdoctoral Scholarship to EJR, as well as OFI research funding awarded to JLR, which then also provided support to DH and JT. The Natural Sciences and Engineering Research Council of Canada (NSERC) also provided funding in the form of an NSERC Discovery grant to JLR and an NSERC CGS-D scholarship to BMR. Additionally, BMR was funded through a Nova Scotia Graduate Scholarship and a Killam Predoctoral award. We thank W.K.W. Li, A. Cogswell, the Bedford Institute of Oceanography, and Fisheries and Oceans Canada (DFO) for starting and maintaining the time series. We also thank the numerous undergraduate/graduate students and lab personnel over the years for their weekly stamina during the four years of field and lab work (particularly, I. Luddington, C. Willis, S. Lau, M. Stewart, H. Blanchard, S. Rose, and C. Mackie). The Coastal Environmental Observation Technology and Research team led by R. Davis and supported by the Marine Environmental Observation, Prediction and Response Network (MEOPAR) provided valuable data, and D. Atamanchuk and K. Patterson helped to make them available. Lastly, we thank D. Wallace and the CERC-OCEAN team for providing nutrient analyses.

## Author contributions

E.J.R. analysed the data and wrote the manuscript; J.L.R. designed the research. J.T. and D.D. were responsible for the bioinformatic analyses. J.M.R., J.Z., B.M.R., J.T. performed sampling and E.J.R., J.T., D.D., B.M.R., D.H., and J.L.R. provided ecological interpretations. All authors contributed to the final data interpretation and write up of the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-19165-w>.

**Correspondence** and requests for materials should be addressed to E.J.R.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022