Everything is not everywhere: Can marine compartments shape phytoplankton assemblages?

Sofie Spatharis[1&2], Vasiliki Lamprinou[3], Alexandra Meziti[4], Konstantinos A. Kormas[4], Daniel D Danielidis[3†], Evangelia Smeti[5], Daniel L Roelke[6], Rebecca Mancy[2], George Tsirtsis[7]

[1] University of Glasgow, School of Life Sciences, Glasgow, G12 8QQ, Scotland, UK

[2] University of Glasgow, Institute of Biodiversity, Animal Health and Comparative Medicine, Glasgow, G12 8QQ, Scotland, UK

[3] University of Athens, Faculty of Biology, Department of Ecology and Systematics, Panepistimiopolis, Athens 15784, Greece

[4] University of Thessaly, Department of Ichthyology & Aquatic Environment, 3844 46 Volos, Greece

[5] Hellenic Centre for Marine Research, Institute of Marine Biological Resources & Inland Waters, 46.7 km Athens-Souniou Av., 19013, Anavissos, Attica, Greece

[6] Texas A&M University at Galveston, Department of Marine Biology, 200 Seawolf Parkway, Galveston, TX 77554

[7] University of the Aegean, Department of Marine Sciences, University Hill, 81100, Mytilene, Greece

[†] Deceased

Corresponding author: sofie.spatharis@glasgow.ac.uk

Abstract

The idea that “everything is everywhere, but the environment selects” has been seminal in microbial biogeography, and marine phytoplankton is one of the prototypical groups used to illustrate this. The typical argument has been that phytoplankton is ubiquitous, but that distinct assemblages form under environmental selection. It is well established that phytoplankton assemblages vary considerably between coastal ecosystems. However, the relative roles of compartmentalisation of regional seas and site-specific environmental conditions in shaping assemblage structures, have not been specifically examined. We collected data from coastal embayments that fall within two different water compartments within the same regional sea and are characterised by highly localised environmental pressures. We used PCNM and AEM models to partition the effects that spatial structures, environmental conditions and their overlap had on the variation in assemblage composition. Our models explained a high percentage of variation in assemblage composition (59-65%) and showed that spatial structure consistent with marine compartmentalisation played a more important role than local environmental conditions. At least during the study period, surface currents connecting sites within the two compartments failed to generate sufficient dispersal to offset the impact of differences due to compartmentalisation. In other words, our findings suggest that, even for a prototypical cosmopolitan group, everything is not everywhere.

Keywords: phytoplankton, marine realms, biogeography, connectivity, PCNM, Mediterranean, species sorting

Introduction

Planktic organisms, which disperse passively with low energetic costs, comprise the most cosmopolitan species in the marine environment [1,2]. However, due to high growth rates and resultant species turnover, distinct phytoplankton assemblages form rapidly under environmental selection [3]. The joint effect of passive dispersal and environmental selection is conveyed concisely by the statement that 'everything is everywhere, but the environment selects' [4,5]. However, it is unclear at which spatial scale environmental selection occurs. A substantial body of literature has highlighted the role of strong small-scale environmental pressures acting within geographically defined coastal ecosystems (e.g. bays) in driving distinct assemblage composition and diversity [6–8]. However, this leaves open the question of the importance of processes acting at broader scales such as marine compartments within the same regional sea, forming from the movement of water masses originating in marine areas that are biogeographically distinct (e.g. marine realms).

Previous studies have clearly demonstrated that biogeographically defined marine realms (e.g. North Atlantic, Black Sea) can be characterised by distinct assemblage structure with respect to both species richness and composition (e.g. endemic species). This has been demonstrated for taxa with movement limitation, such as benthic organisms, but does not hold for organisms with active movement such as large fish [1]. These marine realms are typically separated by biogeographic barriers, but even when water originating from different marine realms comes into close contact, mixing of waters (and thus biological communities) can also be limited by circulation patterns (e.g. generated as a result of temperature and/or salinity differentials), potentially leading to the formation of broad-scale marine compartments even within the same regional sea [9–11]. This is, for instance, the case in the Aegean Sea which is separated into east (Atlantic origin water) and west (Black sea water) compartments, and also in Sicily strait separated into an eastern and western basin [10,12]. In the case where the water masses originate from sources with a distinct species composition [1], it is possible that compositional differences are maintained inside marine compartments within regional seas, potentially dominating local environmental conditions in shaping phytoplankton assemblages.

The role of physical barriers as drivers of plankton assemblage structure has traditionally been investigated in lakes, salt lakes, rivers and interconnected reservoirs [13–17]. In these systems, the dispersal of organisms is limited or facilitated by processes such as wind [18], directional riverine flow [19,20], pipelines connecting artificial reservoirs [15] and animal transport [21]. In these contexts, evidence suggests that species sorting according to local environmental conditions is the prevailing mechanism driving assemblage composition, whereas dispersal limitation becomes important only at large scales of the order of thousands of kilometres [16,17].

In marine systems, the lack of obvious physical barriers means that passively dispersed microorganisms such as phytoplankton have been traditionally thought to have a ubiquitous presence across the globe [22]. Exceptions occur in coastal areas where strong influences from watersheds create specific, localised environmental conditions that lead to the dominance of specific species and thus differentiation of phytoplankton assemblages [6–8,20]. Complementing this framework is the idea that compartmentalisation of a regional sea could also contribute to the formation of distinct phytoplankton assemblages, due to compositional differences in the assemblages in the marine realm from this the water masses of the compartments originate [23–25]. In addition to being of fundamental importance to community ecology, identifying the relative role of local versus broad-scale processes in shaping assemblage diversity can have important practical implications for conservation and ecosystem services. For instance, in cases where there is an important role of marine compartments, water quality assessment should consider the influence of the broader-scale processes when making inferences about the effect of local natural and/or anthropogenic influences.

Our aim was to test for evidence of compartmentalisation, based on water mass origin in a regional sea, affecting the composition and diversity of phytoplankton assemblages. To address this question, we selected a geographical area in which both local and regional processes operate. Specifically, in the periphery of a regional sea, we sampled from multiple semi-enclosed coastal systems that fall within two different marine compartments. These marine compartments have been defined by previous studies based on hydrodynamics, as well as abiotic and biotic criteria (e.g. nutrients, Chl a). We applied a standard methodology that partitions the effect of environment, spatial structuring and the space-environment overlap in explaining assemblage composition patterns. We expected the role of compartmentalisation to be important when spatial structures that reflect the two marine compartments were able to explain the variation in assemblage composition better than environmental covariates acting at the level of coastal site. Although we expected the influence of dispersal between compartments due to seasonal hydrodynamic circulation to have a weaker effect on assemblage composition than original compositional differences within compartments, we extended our analysis to also account for this effect.

Methodology

Overview

To address the question of whether marine compartmentalisation affects assemblage composition, we sampled from sites within two distinct marine compartments, ensuring that sampling occurred within a short time period to exclude the effects of rapid assemblage turnover. Below, we give a detailed account of the selection of 9 sampling sites, which fall within the two previously identified compartments (Table S2). We also detail our methods for analysing assemblage composition using both morphological and molecular species identification. We then provide a description of the PCNM method which was used to partition variation in assemblage composition into that explained by spatial structures such as marine compartments, environmental conditions in the nine coastal sites, and their overlap. We also describe the complementary AEM method, which considers directional connectivity between pairs of sites and was used to account for possible effects of dispersal due to short term hydrodynamic circulation patterns in the area. Finally, we describe additional statistical tests that were used to confirm whether the effects of compartmentalisation also affected other measures of community structure such as genus richness, evenness, compositional turnover and the abundance of specific genera.

Description of datasets

As sampling sites, we selected 9 gulfs (G1 to G9) in the Aegean Sea, Greece (Fig. 1), characterised by a range of environmental conditions due to differences in hydrology, geomorphology, substrate, terrestrial runoff, and local anthropogenic pressures (see also

Table S1). The sites are set within a polygonal area of 72,600 km^2 and belong to two different large scale marine compartments based on a categorisation that relies on hydrodynamic circulation patterns [10,26], biogeochemical variables such as nutrients, pH, dissolved oxygen, POC e.g. [23–25] and satellite-observed Chl-a [12,27] (see also Table S2). Specifically, based on studies relying on a synthesis of variables of hydrological, climatological and satellite data [12,27], sites G1-G5 can be categorised in the West compartment fed by Black Sea outflow waters, whereas sites G6-G9 can be categorised in the East compartment fed by Atlantic water. Within each site, five stations were sampled at 1m, 5m, and at the Secchi depth (i.e. the depth at which light penetration ceases). Sampling took place during July 2014 and was carried out within 19 days (5 to 24 July) to minimise the effect of temporal variation on assemblage composition. July was selected for sampling as previous studies have shown that physicochemical variables and phytoplankton composition are relatively stable during this summer period, at least for a subset of the coastal areas included in the present study [8,28]; this is in contrast to winter months, during which episodic rainfall events – which at this large scale may not simultaneously affect all sites – add noise that could distort inferences regarding the role of compartmentalisation [8].

At each station and depth, we recorded information on environmental covariates that are necessary for the growth of phytoplankton (i.e. autotrophic and mixotrophic protists) which made up the vast majority of our assemblage composition. Specifically, salinity and temperature were recorded on-site, while 3L seawater samples were collected with a Niskin type sampler for later nutrient measurement (NO_3, NO_2, PO_4, SiO_2, Organic N, Organic P), alongside 2L for species identification (1L for each of two identification methods). Organic nitrogen and phosphorus were strongly correlated with Dissolved Inorganic Nitrogen (DIN) and Dissolved Inorganic Phosphorus (DIP) and were therefore excluded from further analysis. The covariates used in the analysis were thus salinity, temperature, DIN, DIP, and SiO_2.

Our environmental covariates presented significant variation between the different coastal sites (Table S3, Fig. S2) which was a prerequisite to enable the testing of whether phytoplankton composition was affected by site-specific environmental conditions or spatial structuring due to compartmentalisation. The magnitude of variation in nutrient and chl a variables between our coastal sites was considerably higher than that observed by previous studies on the open Aegean waters [29]. Note that any overlap between our environmental covariates and the spatial structures due to compartmentalisation (i.e. spatial autocorrelation) was explicitly accounted for in the PCNM and AEM approaches (see data analysis).

Species identification

Our study focuses on unicellular eukaryotic organisms, either autotrophic or mixotrophic. Although Cyanobacteria play an important role in the oligotrophic Aegean Sea, especially during spring and summer [30], this group was not considered due to its small size range that does not permit a reliable identification and enumeration by microscopy, while it is also undetected by our molecular approach (sequencing the 18S rRNA gene). Species were identified by both morphological and molecular approaches. Although the morphological approach offers a more accurate representation of species abundances, the molecular approach offers a more exhaustive species list - at the level of operational taxonomic unit (OTUs) - as it can capture species that are rare and/or small and thus undetected by microscopy. The samples were analysed morphologically using conventional microscopy techniques, and genetically using the high throughput sequencing platform of Illumina MiSeq 2x300 bp, producing a dataset of species abundance and species read data respectively. Sample size for the morphological dataset was 135 (9 gulfs x 5 stations x 3 depths), whereas the molecular dataset consisted of the same samples except for sites G2 & G3 due to the occurrence of excess mucilage which reduced the filtration efficiency and, consequently, the quality of extracted DNA. Abundance data from the microscopy species identification were transformed with Hellinger transformation to minimise skewness due to low abundances of rare species [31]. To conduct our analyses, we aggregated data to the genus level to minimise any potential biases in the microscopy due to misidentification at the species level. The molecular data were also analysed at this level to allow corresponding inferences from the two approaches.

The molecular method detected 198 genera whereas morphological identification detected 132 (Fig. S1). This higher number genera observed in the molecular approach was mostly because the Illumina MiSeq 2x300 bp high throughput sequencing platform is more sensitive in detecting Haptophyta, Ochrophyta, Chlorophyta and picophytoplankton species (having one dimension <3 μm) that were usually not captured morphologically. Details of the morphological and molecular identification methods are provided in the supplementary methodological information.

Data analysis

To visualise the pattern in assemblage composition of marine phytoplankton with data from the morphological identification approach, we used non-metric multidimensional scaling (nMDS) based on the Bray-Curtis similarity index on Hellinger-transformed species-abundance and presence-absence data. If our samples clustered primarily based on the broader scale of the marine compartment (West and East), and secondarily based on the more local scale of coastal site (G1-G9), this would provide evidence of the effect of compartmentalisation on assemblage composition. To test for these effects statistically, we used PerMANOVA analysis [32]. Finally, we also tested whether the two compartments were different with respect to how homogeneous they were in their species composition (i.e. the similarity of samples within them) using PERMDISP2 analysis [33].

In order to identify specific spatial and environmental factors affecting patterns in assemblage composition, we used the method of Principal Coordinates of Neighbour Matrices (PCNM), a standard method for partitioning the effects of the environment and space in ecological studies e.g. [17,34,35] while taking into account the space-environment overlap [36]. In our context, a higher percentage of variation explained by spatial (PCNM) covariates would indicate the importance of broad-scale spatial structuring whereas a higher percentage of variation explained by the environment would suggest species sorting based on local abiotic factors. According to this method, spatial structures (either fine- or broad-scale) are extracted using solely the geographic coordinates of the sampling stations. The mathematical procedure is based on a Principal Coordinate Analysis of the modified distance matrix of sampling stations [37]. The values of fine- and broad-scale spatial harmonics (similar to the waves used in sound analysis) are recorded for each station resulting to the principal coordinates (i.e. PCNM covariates). Thereafter, Redundancy Analysis (RDA) selects those PCNM and environmental covariates that explain significant variation in assemblage composition [35]. In our study, associations between community composition and large-scale spatial covariates that varied along an east-west gradient would be more indicative of compartmentalisation, whereas associations with spatial covariates that vary over a smaller scale would be indicative of unexplored environmental covariates or biotic interactions [38] (also note that the method focuses on detecting linear rather than non-linear relationships).

The output of the RDA analysis also provides information on the most important genera that are affected by the spatial and environmental covariates. This analysis was conducted using morphological data due to possible issues of interpretation in the molecular data arising because of biases associated with amplification of protists [39–41].

To account for potential transfer of phytoplankton between compartments due to short-term surface circulation, we used the method of Asymmetric Eigenvector Maps (AEM) [42]. This is a method for modelling spatial distributions of species that takes into account between-site connectivity, including asymmetry in connectivity. Connectivity weightings were estimated through a computational experiment that tracked the movement of a theoretical particle (see supplementary methodological material) connecting pairs of sites through surface currents of the Aegean Sea (Table S4). If the inclusion of directional between-site connectivity increased the amount of explained variation by the AEMs (relative to the PCNMs), then this would indicate current-facilitated transfer of organisms acting to increase homogenisation between sites. PCNM and AEM methods were applied separately for each of the three depths for robustness checking of the statistical results.

To further check results regarding spatial structuring in assemblage composition, we tested whether observed differences between sites were also reflected in other measures of assemblage structure. Specifically we focused on: (a) genus richness using the data from the molecular species identification which provided a fuller species identification; (b) assemblage evenness [43] based on our quantitative morphological data; (c) abundances of specific genera that were identified by the RDA analysis as strong drivers of the spatial patterns (i.e. those showing a strong correlation with a PCNM spatial covariate); and (d) abundances of picophytoplankton species that were identified only by the molecular approach. To test for significant differences between pairs of sites, we used Tukey Honest Significant Differences (HSD) test.

We used the 'vegan' v 2.5.3 package in R [44] for running nMDS, PerMANOVA, PERMDISP2, RDA, CCA, PCA, PCNM and variation partitioning, 'packfor' v 0.0.8 [45] for forward selection, and 'AEM' v 0.6 [46] for AEM. The Academo Venn diagram generator free tool was used to visualise the partition of variation into space, environment and their overlap. Ocean Data View [47] was used to visualise the variation of spatial covariates on the Aegean sea maps. Statistical analysis was carried out in R v.3.3.3 [48].

Results

One important condition required to demonstrate an effect of compartmentalisation on assemblage composition would be the observation of greater variation in assemblage composition between the compartments than variation within each compartment. The pattern in genus composition of phytoplankton that our study sought to explain is shown in the NMDS analysis of Figure 2. Stations, based on their genus abundance composition using morphological identification data, showed clustering primarily based on site. A longitudinal spatial separation was also observed, whereby the West compartment G1 -5, formed a broad group of more similar assemblages that was distinct from the East compartment, within which sites G6-9 formed three distinct clusters (Fig. 2a). These clustering patterns were tested statistically using PerMANOVA analysis on morphological data that were Hellinger and presence/absence transformed, and based on sequential addition of compartment, site and depth using 9999 permutations. For Hellinger transformed abundance data, compartment had a significant effect in the clustering of stations ($F_{1,125}=3.901$, $p<0.001$) as well as site ($F_{7,125}=10.839$, $p<0.001$) and depth ($F_{1,125}=0.383$, $p<0.01$) (Fig. 2a). This east-west separation held, although less pronounced, when presence-absence data were used ($F\text{-compartment}_{1,125}=2.088$, $p<0.001$, $F\text{-site}_{7,125}=5.701$, $p<0.001$, $F\text{-depth}_{1,125}=0.442$, $p<0.001$) (Fig. 2b). No differences were observed between the two compartments with respect to their homogeneity in genus composition between sites (for Hellinger-transformed data: $F_{1,133}=0.005$, $p=0.343$; for presence/absence data: $F_{1,133}=0.006$, $p=0.117$).

Regarding the drivers of assemblage composition, the spatially explicit PCNM method explained 59-65% of the variation across the three depths (Fig. 3 & Table 1). In the PCNM analysis, most of this explained variation was accounted for by spatial structures (27%) and environmental covariates that were correlated with these structures spatially (i.e. the overlap between spatial and environmental components was 18-27%), whereas the effect of the environment *per se* was weak (11-14%) (see PCNM results in Fig. 3). The high overlap of the environment with the spatial structures is mainly due to the high correlation of salinity and silicates with the spatial covariate PCNM2 (Pearson correlation coefficient for salinity: -0.732 and silicates: 0.479, p-value<0.001, at 1 m depth). When potential dispersal of plankton cells between sites was accounted for by including connectivity weightings and directionality using the AEM analysis, then the amount of explained variation fell or remained unchanged, while the variation partitioning into spatial, environmental components and their overlap was qualitatively similar to PCNM (see AEM in Fig. 3).

The spatial covariates that contributed significantly and most strongly to the explained variation (RDA, p-value<0.001) were PCNM1 and PCNM2 (Table 1). Specifically, the spatial covariate PCNM1 (Fig. 4a) presented maximum variation along the east-west axis, and was thus aligned with the two marine compartments previously identified in the Aegean Sea. PCNM2 varied only within the East compartment along a SE to NW axis (Fig. 4b), in agreement with the clustering of sites G6-G9. Of the environmental covariates, salinity and silicates were the most important in driving composition (Table 1), although with a high degree of correlation with the spatial covariate PCNM2 as previously mentioned.

The genera that were identified by the RDA analysis (Fig. S3) as the most important in driving the spatial structuring of assemblage composition according to PCNM1 were three typical planktic diatoms, *Bacteriastrum, Leptocylindrus* and *Chaetoceros* (Fig. S4). Specifically, in the G1-5 sites of the West compartment, these species were present at high abundance (e.g. *Chaetoceros* maximum abundance 9,034 cells/L), considering the oligotrophic environment of Eastern Mediterranean, but were absent or present only at very low levels in the G6-9 sites of the East compartment.

The separation of assemblage composition between the West and East marine compartments, identified by the PCNM analysis, was supported by further analysis of the patterns of genus richness as measured by the molecular approach. Pairwise comparisons of molecular genus richness between sites G1-5 of the West compartment and G6-9 of the East compartment based on an LSD test were significantly different ($p<0.05$) at both 1m and 5m depths, showing lower genus richness in the G1-5 sites (Fig. S5A & Table S5). However, consistent differences between sites of the two compartments were not observed at Secchi depth, or when using morphological genus richness data (Fig. S5B). Having accounted for depth, the different compartments also explained a significant amount of variation in the evenness of assemblages (using the morphological approach) (ANOVA, $F_{1,124}=10.26$, $p<0.01$) which was not collinear with the variation explained by sites (ANOVA, $F_{8,124}=11.10$, $p<0.001$). However, these differences in evenness were not consistent among the sites of the two compartments (Fig. S5C).

From a total of 198 genera identified by the molecular approach, 32% were not shared between the West and East compartments (28% of 132 genera using the morphological approach). Specifically, a higher number of genera unique to a single compartment was observed in the East compartment (36 genera), compared to the West compartment (27 genera) (see also Fig. S6). This observation was consistent between the two identification methods and across four main phyla present in our samples (Baccilariophyta, Miozoa, Ochrophyta, Haptophyta), with the exception of Chlorophyta, for which more genera were detected in the West compartment compared to the East (Fig. S7 top panel). Eukaryotic picophytoplankton species, which were only detected by the molecular approach, were present in both compartments and their abundance (based on reads) did not show a consistent dominance of this group in either West or East compartment (Fig. S7).

Discussion

Our findings provide evidence that marine compartments fed by water and associated assemblages from distinct biogeographic marine realms can affect the composition and diversity of phytoplankton assemblages, even within the same regional sea. Our findings also suggest that although assemblages are shaped by environmental conditions at the level of coastal site, this effect is insufficient to explain broader scale patterns in assemblage composition and diversity. Our findings emphasise the importance of large-scale biogeographical processes in driving patterns of diversity and composition, at least for certain periods of the year (summer, in our case), and confirm the suggestion by [49] that not all protists are ubiquitous. Although, compartmentalisation in the Aegean Sea has been previously based on satellite data of primary productivity [12,27], this is the first time that the existence of these compartments has been confirmed by more complex measures such as genus composition and richness. Previous studies in community ecology, following similar methodological approaches may have missed the effect of marine compartmentalisation [34,35,50] because the spatial extent of studies did not incorporate different marine compartments or marine realms.

In our study, compartmentalisation (PCNM1) and environmental variation overlapping with a finer scale spatial structure (PCNM2) were able to explain a large proportion (>50%) of the variation in assemblage composition. Specifically, the role of spatial structuring, as expressed by the PCNM1 gradient, explained most of the variation in assemblage composition and reflected the two marine compartments in the Aegean Sea [10] that are formed by the general hydrodynamic circulation pattern. The East compartment is characterised by a northward current of Atlantic origin, flowing from the Levantine Sea and then along the East coastline of the Aegean. The West compartment is fed by light, brackish water from the Black Sea flowing along the North and West coastline [51,52]. The distinct phytoplankton assemblage composition that we identified in our study between East and West reflects the previous reports of compartmentalisation and was attributed to the fact that almost one third of the genera was not shared between the two marine compartments, and that some genera that had high abundances in the West compartment were absent from the East. Additionally, this finding seems to support a previous synthesis of plant and animal data showing that the Black Sea, from which water in the West compartment originates, is quite distinct and characterised by high endemicity, compared to the N Atlantic Ocean, from which water in the East compartment originates [1]. The fact that quantitative data explained almost three times more variation than presence-absence data could also suggest that the differences in physicochemical conditions may result in species sorting at the compartment level. Although during winter, nutrient-rich freshwater inputs might increase forcing at the level of local site, the influence of the western marine compartment might nevertheless also become more prevalent, given that Black Sea (its origin) receives important freshwater inputs from main European Rivers (e.g. Danube, Dnieper, Dniester, and Don).

Also linked to the origin of these water masses might be the observed difference in molecular genus richness between the two compartments. It is a general principle in biogeography that larger geographical areas (or water masses) are more species diverse as they provide the scope for more niches [53]. Following this principle, the higher genus richness observed in the East compartment is expected given that the origin of water flowing into it, is the much larger N Atlantic Ocean, which is much more species diverse than the Black Sea water which flows into the West compartment [1]. This observation is also confirmed by a previous comparison of copepod diversity between the South Aegean Sea (part of the East compartment) and the Black Sea water masses [54]. We conducted additional analysis to check whether this distinction could be attributed to higher levels of environmental heterogeneity in the East compartment (island sites) than the West compartment (mainland sites). However, our analysis showed that the two compartments were not different in the level of heterogeneity within them in relation to the most critical environmental conditions for phytoplankton (Figure S8). We also observed no significant difference between the two compartments in their levels of assemblage heterogeneity.

Interestingly, our study also shows that picophytoplankton species, which can be even less constrained by dispersal limitation than microphytoplankton due to their considerably smaller sizes, did not follow the east-west compartmentalisation. Regarding the whole phytoplankton assemblage, differences between compartments were also not picked up by the measure of evenness; this is expected because phytoplankton evenness is known to be driven primarily by the dominance of species that respond rapidly to nutrient enrichment commonly occurring at a local scale within both compartments [55].

Our study indicated another smaller-scale spatial pattern (expressed by the spatial covariate PCNM2), operating within the East compartment (sites G6-G9), in the neighbourhood of the frontal area isolating the warm, saline Levantine Surface Water from the cold, low-salinity Black Sea Water [52]. This finding seems to resolve previous ambiguity concerning the association of Lemnos Island with a specific compartment (see Table S2), at least with respect to phytoplankton assemblage composition. Specifically, our salinity and temperature covariates revealed a separation of the northern-most island sites G6-7 of Lemnos Island from the Lesvos Island sites, in agreement with some studies associating this island with the Western ecoregion [10,23,24,26]. However, the gradient dictated by PCNM2 did not offset the broader-scale pattern in genus composition, according to which stations G6-9 all belong to the East compartment, in agreement with studies based on climatology and satellite data [12,27]. Taken together, these findings suggest that the sea between Lemnos (sites G6-7) and Lesvos (sites G8-9) islands represents the meeting point of the two distinct water masses of Black Sea and Atlantic Ocean water; through this exchange of phytoplankton species occurs, primarily following the direction of the eastern upwards current bearing Atlantic water.

The fact that salinity and silicates were amongst the most important environmental drivers of assemblage composition appears to be inconsistent with localised coastal monitoring studies in which sampling within annual cycles reveal nitrogen and phosphorous as the main drivers of compositional turnover in temperate oligotrophic waters [56–58]. This apparent contradiction could be a result of spatial and temporal scales. Specifically, phytoplankton composition responds in a more pronounced manner to temporal variability of nutrients (e.g. due to freshwater inputs) than to salinity, because of opportunistic species such as *Pseudo-nitzschia* spp. that form high biomass blooms within short periods of time [57]. However, our results show that over larger spatial scales, the effect of nitrogen and phosphorus appear to diminish, and salinity and silicates are the main drivers of assemblage composition. Consistent with our findings, an important role for salinity and silicates in microeukaryotic assemblage structure has been shown in the East China Sea, over a large geographic area of hundreds of kilometres [34] comparable to the scale employed in our study.

In our study area, surface currents are affected by seasonal variability in the intensity of Black Sea water inflow, which generates short-term surface circulation patterns in the Aegean Sea (e.g. gyre formations). Although such surface currents were present during the study period (July) it seems that these were either not as strong, or did not operate long enough, to cause the dispersal of phytoplankton genera between sites. This was demonstrated by the inclusion of connectivity between sites introduced with the AEM model, which did not improve the explained variation in assemblage structure. This is consistent with previous studies showing that short-term hydrodynamic patterns in the area do not affect the main hydrodynamic features of the northward Eastern current and southward Western current [59,60], creating the separation between the two marine compartments. It seems therefore that stochastic processes such as dispersal are not intense enough to generate homogenisation of assemblage composition and thus override the separation of assemblages due to larger-scale compartmentalisation in the area.

Conclusions

Our findings suggest that the effect of regionalisation on assemblage composition and diversity should be considered carefully when investigating drivers of assemblage structure. In contrast to the dominant paradigm in ecology, according to which diversity and composition of phytoplankton assemblages have been thought to be strongly driven by local environmental conditions that arise due to inflow from watersheds, our findings indicate that circulation patterns which lead to broad-scale compartmentalisation might also have an important impact on the diversity and composition of phytoplankton assemblages in coastal systems. This effect may nevertheless depend on the strength of surface circulation affecting dispersal process between compartments.

In addition to the impact of this finding for our understanding of ecosystems, the management implications are also critical because neglecting this factor while prioritising sites for conservation (e.g. for the purposes of directives such as Natura Network or Water Framework), might result in the misdirection of management resources (e.g. if there are differences in resilience/resistance to perturbations). If sufficient care is not taken, differences in composition and diversity between sites within a common regional sea might be erroneously attributed to localised anthropogenic pressures (e.g. eutrophication) rather than processes operating at much broader scales, such as marine compartmentalisation. Given that the diversity and composition of phytoplankton fuels higher trophic levels [61], any shifts in marine compartments due to e.g. climate change may have important knock-on effects on the ecosystem goods and services provided by coastal ecosystems.

Acknowledgement

This research was conducted within the research project 'ECOGENE: The relative role of niche and neutral mechanisms in controlling phytoplankton genetic and morphological diversity' (Code Number 4691), implemented within the framework of the Action ARISTEIA II of the

Operational Program 'Education and Lifelong Learning' (Action's Beneficiary: General Secretariat for Research and Technology), and co-financed by the European Social Fund (ESF) and the Greek State. RM gratefully acknowledges funding from the Leckie Fellowship.

References

1. Costello MJ, Tsai P, Wong PS, Cheung AKL, Basher Z, Chaudhary C. 2017 Marine biogeographic realms and species endemicity. *Nat. Commun.* **8**, 1–9. (doi:10.1038/s41467-017-01121-2)
2. Cermeño P, de Vargas C, Abrantes F, Falkowski PG. 2010 Phytoplankton biogeography and community stability in the ocean. *PLoS One* **5**. (doi:10.1371/journal.pone.0010037)
3. Sommer U. 1989 The role of competition for resources in phytoplankton succession. In *Plankton ecology* (ed U Sommer), pp. 57–106. New-York: Springer-Verlag.
4. de Wit R, Bouvier T. 2006 'Everything is everywhere, but, the environment selects'; what did Baas Becking and Beijerinck really say? *Environ. Microbiol.* **8**, 755–758. (doi:10.1111/j.1462-2920.2006.01017.x)
5. Baas-Becking LGM. 1934 *Geobiologie of inleiding tot de milieukunde*. (doi:10.1007/s10825-011-0356-9)
6. Justic D, Rabalais N, Turner RE, Dıaz RJ. 2009 Global change and eutrophication of coastal waters. *ICES J. Mar. Sci.* , 1528–1537.
7. Roelke DL, Li H-P, Miller-DeBoer CJ, Gable GM, Davis SE. 2017 Regional shifts in phytoplankton succession and primary productivity in the San Antonio Bay System (USA) in response to diminished freshwater inflows. *Mar. Freshw. Res.* **68**, 131–145.
8. Spatharis S, Tsirtsis G, Danielidis DB, Chi TD, Mouillot D. 2007 Effects of pulsed nutrient inputs on phytoplankton assemblage structure and blooms in an enclosed coastal area. *Estuar. Coast. Shelf Sci.* **73**. (doi:10.1016/j.ecss.2007.03.016)
9. Gaines SD, Gaylord B, Gerber LR, Hastings PA, Kinlan BP. 2005 The ecological consequences of dispersal in the sea. *Oceanography* **20**, 90–99.
10. Berline LO, Rammou AM, Doglioli A, Molcard A, Petrenko A. 2014 A connectivity-based Eco-regionalization method of the mediterranean sea. *PLoS One* **9**, 3–11. (doi:10.1371/journal.pone.0111978)

11. Planque B, Lazure P, Jegou AM. 2006 Typology of hydrological structures modelled and observed over the Bay of Biscay shelf. *Sci. Mar.* **70**, 43–50. (doi:10.3989/scimar.2006.70s143)

12. Ayata SD *et al.* 2018 Regionalisation of the Mediterranean basin, a MERMEX synthesis. *Prog. Oceanogr.* **163**, 7–20. (doi:10.1016/j.pocean.2017.09.016)

13. Liu J, Soininen J, Han BP, Declerck SAJ. 2013 Effects of connectivity, dispersal directionality and functional traits on the metacommunity structure of river benthic diatoms. *J. Biogeogr.* **40**, 2238–2248. (doi:10.1111/jbi.12160)

14. Soininen J, Jamoneau A, Rosebery J, Passy SI. 2016 Global patterns and drivers of species and trait composition in diatoms. *Glob. Ecol. Biogeogr.* **25**, 940–950. (doi:10.1111/geb.12452)

15. Hu R, Duan X, Peng L, Han B, Naselli-Flores L. 2017 Phytoplankton assemblages in a complex system of interconnected reservoirs: the role of water transport in dispersal. *Hydrobiologia* **800**, 17–30. (doi:10.1007/s10750-017-3146-y)

16. Xiao LJ, Zhu Y, Yang Y, Lin Q, Han BP, Padisák J. 2018 Species-based classification reveals spatial processes of phytoplankton meta-communities better than functional group approaches: a case study from three freshwater lake regions in China. *Hydrobiologia* **811**, 313–324. (doi:10.1007/s10750-017-3502-y)

17. Yang Y, Hu R, Lin Q, Hou J, Liu Y, Han BP, Naselli-Flores L. 2018 Spatial structure and β-diversity of phytoplankton in Tibetan Plateau lakes: nestedness or replacement? *Hydrobiologia* **808**, 301–314. (doi:10.1007/s10750-017-3433-7)

18. Horváth Z, Vad CF, Ptacnik R. 2016 Wind dispersal results in a gradient of dispersal limitation and environmental match among discrete aquatic habitats. , 726–732. (doi:10.1111/ecog.01685)

19. Alpine AE, Cloern JE. 1992 Trophic interactions and direct physical effects control phytoplankton biomass and production in an estuary. *Limnol. Oceanogr.* **37**, 946–955. (doi:10.2307/2837841)

20. Roelke DL, Li H, Hayden NJ, Miller CJ. 2013 Co-occurring and opposing freshwater inflow effects on phytoplankton biomass , productivity and community composition of Galveston Bay , USA. (doi:10.3354/meps10182)

21. Kristiansen J. 1996 Dispersal of freshwater algae - A review. *Hydrobiologia* **336**, 151–157. (doi:10.1007/BF00010829)

22. BJ F. 2002 Global dispersal of free-living microbial eukaryote species. *Science (80-.).* **296**, 1061–1063.

23. Mayot N, D'Ortenzio F, D'Alcalà MR, Lavigne H, Claustre H. 2016 Interannual variability of the Mediterranean trophic regimes from ocean color satellites. *Biogeosciences* **13**, 1901–1917. (doi:10.5194/bg-13-1901-2016)

24. Palmiéri J. 2014 Modélisation biogéochimique de la mer Méditerranée avec le modèle régional couplé NEMO-MED12/PISCES. Université Versailles-Saint-Quentin.

25. Reygondeau, G., Irisson, J.-O., Ayata, S., Gasparini, S., Benedetti, F., Albouy, C., Hattab, T., Guieu, C., Koubbi P. 2014 Definition of the Mediterranean eco-regions and maps of potential pressures in these eco-regions.

26. Rossi V, Lòpez C. 2014 Auxiliary Material for " Hydrodynamic provinces and oceanic connectivity from a transport network help designing marine reserves " Geophysical Research Letters , 2014. , 7122. (doi:10.1002/2014GL059540.Received)

27. Reygondeau G, Guieu C, Benedetti F, Irisson JO, Ayata SD, Gasparini S, Koubbi P. 2017 Biogeochemical regions of the Mediterranean Sea: An objective multidimensional and multivariate environmental approach. *Prog. Oceanogr.* **151**, 138–148. (doi:10.1016/j.pocean.2016.11.001)

28. Dokulil MT, Naselli-Flores L, Chorus I, Padisák J. 2003 Equilibrium/steady-state concept in phytoplankton ecology. *Hydrobiologia* **502**, 395–403.

29. Ignatiades L. 2005 Scaling the trophic status of the Aegean Sea, eastern Mediterranean. *J. Sea Res.* **54**, 51–57. (doi:10.1016/j.seares.2005.02.010)

30. Christaki U, Van Wambeke F, Dolan JR. 1999 Nanoflagellates (mixotrophs, heterotrophs and autotrophs) in the oligotrophic eastern Mediterranean: standing stocks, bacterivory and relationships with bacterial production. *Mar. Ecol. Prog. Ser.* **181**, 297–307.

31. Legendre P, Gallagher ED. 2001 Ecologically meaningful transformations for ordination of species data. *Oecologia* **129**, 271–280. (doi:10.1007/s004420100716)

32. Anderson MJ. 2001 A new method for non-parametric multivariate analysis of variance. *Austral Ecol.* **26**, 32–46. (doi:10.1111/j.1442-9993.2001.01070.pp.x)

33. Anderson MJ, Ellingsen KE, McArdle BH. 2006 Multivariate dispersion as a measure of beta diversity. *Ecol. Lett.* **9**, 683–693. (doi:10.1111/j.1461-0248.2006.00926.x)

34. Zhang H, Huang X, Huang L, Bao F, Xiong S, Wang K, Zhang D. 2018 Microeukaryotic

biogeography in the typical subtropical coastal waters with multiple environmental gradients. *Sci. Total Environ.* **635**, 618–628. (doi:10.1016/j.scitotenv.2018.04.142)

35. Wu N, Qu Y, Guse B, Makarevičiūtė K, To S, Riis T, Fohrer N. 2018 Hydrological and environmental variables outperform spatial factors in structuring species, trait composition, and beta diversity of pelagic algae. *Ecol. Evol.* **8**, 2947–2961. (doi:10.1002/ece3.3903)
36. Borcard D, Legendre P, Avois-Jacquet C, Tuomisto H. 2004 Dissecting the spatial structure of ecological data at multiple scales. *Ecology* **85**, 1826–1832. (doi:10.1890/03-3111)
37. Borcard D, Legendre P. 2002 All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. *Ecol. Modell.* (doi:10.1016/S0304-3800(01)00501-4)
38. Monti D, Pepin P, Legendre P, Maranger R, Blanchet FG. 2010 Modelling the effect of directional spatial ecological processes at different scales. *Oecologia* **166**, 357–368. (doi:10.1007/s00442-010-1867-y)
39. Peters L, Spatharis S, Dario MA, Dwyer T, Roca IJT, Kintner A, Kanstad-Hanssen Ø, Llewellyn MS, Praebel K. 2018 Environmental DNA: A new low-cost monitoring tool for pathogens in salmonid aquaculture. *Front. Microbiol.* **9**. (doi:10.3389/fmicb.2018.03009)
40. Mäki A, Salmi P, Mikkonen A, Kremp A, Tiirola M. 2017 Sample Preservation, DNA or RNA extraction and data analysis for high-throughput phytoplankton community sequencing. *Front. Microbiol.* **8**, 1–13. (doi:10.3389/fmicb.2017.01848)
41. Domaizon I, Kahlert M, Vasselon V, Rimet F, Bouchez A. 2017 Application of high-throughput sequencing (HTS) metabarcoding to diatom biomonitoring: Do DNA extraction methods matter? *Freshw. Sci.* **36**, 162–177. (doi:10.1086/690649)
42. Blanchet FG, Legendre P, Borcard D. 2008 Modelling directional spatial processes in ecological data. *Ecol. Modell.* **215**, 325–336. (doi:10.1016/j.ecolmodel.2008.04.001)
43. Pielou EC. 1966 No Title The measurement of diversity in different types of biological collections. *J. Theor. Biol.* **13**, 131–44.
44. Oksanen AJ *et al.* 2013 Package ‘ vegan ’. , 0–291.
45. Dray S. 2013 packfor: forward selection with permutation 0.0-8/r100., (Canoco p.46). – R package.

46. Blanchet FG. 2010 AEM: tools to construct asymmetric eigenvector maps (AEM) spatial variables.

47. Schlitzer R. 2018 Ocean Data View, odv.awi.de.

48. R Core Team. 2017 R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing.

49. Fontaneto D, Hortal J. 2013 At least some protist species are not ubiquitous. *Mol. Ecol.* **22**, 5053–5055. (doi:10.1111/mec.12507)

50. Lovejoy C, Potvin M. 2011 Microbial eukaryotic distribution in a dynamic Beaufort Sea and the Arctic Ocean. *J. Plankton Res.* **33**, 431–444. (doi:10.1093/plankt/fbq124)

51. Stergiou KI, Christou ED, Georgopoulos D, Zenetos A, Souvermezoglou C. 1997 The Hellenic Seas: physics, chemistry, biology and fisheries. *Oceanogr. Mar. Biol.* **35**, 415–538.

52. Zervakis V, Georgopoulos D, Karageorgis AP, Theocharis A. 2004 On the response of the Aegean Sea to climatic variability: A review. *Int. J. Climatol.* **24**, 1845–1858. (doi:10.1002/joc.1108)

53. Costello MJ, Chaudhary C. 2017 Marine Biodiversity, Biogeography, Deep-Sea Gradients, and Conservation. *Curr. Biol.* **27**, R511–R527. (doi:10.1016/j.cub.2017.04.060)

54. Zervoudaki S, Nielsen TG, Christou ED, Siokou-Frangou I. 2006 Zooplankton distribution and diversity in a frontal area of the Aegean Sea. *Mar. Biol. Res.* **2**, 149–168. (doi:10.1080/17451000600702037)

55. Spatharis S, Tsirtsis G. 2010 Ecological quality scales based on phytoplankton for the implementation of Water Framework Directive in the Eastern Mediterranean. *Ecol. Indic.* **10**. (doi:10.1016/j.ecolind.2010.01.005)

56. Siokou-Frangou I, Christaki U, Mazzocchi MG, Montresor M, Ribera D'Alcala M, Vaque D, Zingone A. 2010 Plankton in the open mediterranean Sea: A review. *Biogeosciences* **7**, 1543–1586. (doi:10.5194/bg-7-1543-2010)

57. Spatharis S, Danielidis DB, Tsirtsis G. 2007 Recurrent Pseudo-nitzschia calliantha (Bacillariophyceae) and Alexandrium insuetum (Dinophyceae) winter blooms induced by agricultural runoff. *Harmful Algae* **6**. (doi:10.1016/j.hal.2007.04.006)

58. Tamvakis A, Miritzis J, Tsirtsis G, Spyropoulou A, Spatharis S. 2012 Effects of meteorological forcing on coastal eutrophication: Modeling with model trees. *Estuar.*

Coast. Shelf Sci. **115**. (doi:10.1016/j.ecss.2012.09.003)

59. Tzali M, Sofianos S, Mantziafou A, Skliris N. 2010 Modelling the impact of Black Sea Water inflow on the North Aegean Sea hydrodynamics. *Ocean Dyn.* **60**, 585–596. (doi:10.1007/s10236-010-0277-3)

60. Mavropoulou AM, Mantziafou A, Jarosz E, Sofianos S. 2016 The influence of Black Sea Water inflow and its synoptic time-scale variability in the North Aegean Sea hydrodynamics. *Ocean Dyn.* **66**, 195–206. (doi:10.1007/s10236-016-0923-5)

61. Marzetz V, Koussoroplis AM, Martin-Creuzburg D, Striebel M, Wacker A. 2017 Linking primary producer diversity and food quality effects on herbivores: A biochemical perspective. *Sci. Rep.* **7**, 1–9. (doi:10.1038/s41598-017-11183-3)

62. Olson DB, Kourafalou VH, Johns WE, Samuels G, Veneziani M. 2007 Aegean Surface Circulation from a Satellite-Tracked Drifter Array. *J. Phys. Oceanogr.* **37**, 1898–1917. (doi:10.1175/JPO3028.1)

Figure captions

Figure 1. Map of the Mediterranean Sea showing the location of Aegean Sea (a), and close up on the Aegean Sea, showing the location of the 9 coastal sites (b). The minimum distance between two stations (within a site) was 0.5 Km and the maximum was 356 Km. Blue arrows indicate the south-westerly flow of Black Sea water across the Northern and Western Aegean coastline, whereas red arrows indicate the flow of water of Atlantic origin northwards along the Eastern Aegean coastline [62].

Figure 2. Pairwise similarities between 135 samples (9 sites x 5 stations x 3 depths), quantified by the Bray-Curtis similarity index and visualised by nMDS analysis. In panel a, similarity is based on Hellinger-transformed genus abundance data and in panel b, on presence-absence data (both based on morphological identification). Samples show clustering by marine compartment (polygons) and within these polygons clustering occurs by site (shapes). The grey star, which indicates the centroid point of all samples within the East compartment, does not coincide with the black star, which is the centroid point within the West compartment.

Figure 3. Partitioning of explained variation in assemblage composition (using morphologically identified, transformed genus abundance data). White circles indicate the variation explained by the spatial (PCNM) covariates, grey circles indicate the variation explained by environmental covariates (S, T, DIN, DIP, SiO_2), and their intersection indicates the explained variation in the environment that is correlated (i.e. overlaps) with the spatial covariates. Also shown are the total explained and residual variation. All reported values represent percentages.

Figure 4. Spatial variation of the PCNM variables 1 & 2 superimposed on the geographical locations of the 9 sampling sites. Dashed line indicates the separation of the Aegean Sea into an East and a West marine compartment according to [12]. The colour scale indicates the range of values taken by the spatial covariates PCNM1 & 2 in the context of this analysis. The variability of spatial covariate PCNM 1 (top panel), which best explains the assemblage composition data, confirms this separation as the directionality of change occurs along the west-to-east axis. PCNM 2 (bottom panel), which was the second most significant spatial

covariate in explaining assemblage composition, varies only across the G5 to G9 sites, along a SE to NW gradient.

Tables

Table 1. The percentage of overall variation in phytoplankton assemblage composition that is explained by spatial structures, the environmental covariates and their overlap using the PCNM method (column 3). Results are shown for each of three depths for morphologically identified Hellinger-transformed genus abundance data and presence-absence genus data. Also shown is the contribution to explained variation (as percentages) due to each of the four significant and top ranked environmental and spatial (PCNM) covariates (RDA analysis, p-value<0.001) (columns 4-7).

Data type	**Depth**	**Total explained variation %**	**% contribution of covariates to explained variation**			
			Covariate 1	Covariate 2	Covariate 3	Covariate 4
Abundance	1m	65	24 (Salinity)	22 (PCNM2)	15 (PCNM1)	13 (SiO_2)
Abundance	5m	65	22 (PCNM2)	17 (Salinity)	16 (PCNM1)	13 (SiO_2)
Abundance	max	59	16 (SiO_2)	14 (PCNM1)	13 (PCNM2)	12 (Salinity)
Presence-absence	1m	25	9 (PCNM2)	9 (Salinity)	6 (PCNM1)	6 (SiO_2)
Presence-absence	5m	25	9 (PCNM2)	7 (SiO_2)	6 (PCNM1)	5 (Salinity)
Presence-absence	max	23	8 (PCNM2)	7 (SiO_2)	7 (PCNM1)	5 (Salinity)