***Supplementary methodological information***

**Morphological identification**

Seawater of 1 liter was immediately preserved with a mixture of neutralized formaldehyde and acidic Lugol’s iodine solution at a ratio of 6:1. Two methods were applied for the quantitative analysis of phytoplankton communities, (a) the standard method of cell counting under a Zeiss AxioVert 100 inverted microscope [1], preceded by serial sedimentations and final concentration of individuals into settling chambers of 25 ml, and (b) sample filtration (300 ml, 2 replicates) through nitrate cellulose membranes of 0.45 μm mesh [2] followed by transillumination of the filter with immersion oil and observation under a Zeiss Axiolab upright microscope. This approach was adopted due to the observed underestimation of the smaller fraction of phytoplankton (<10 μm) by the Utermöhl method. A graph of the cumulative number of species (y axis) over new fields of vision (x axis) was plotted for 30 random samples to estimate the optimal sample surveying effort. The levelling off of this accumulation curve indicated the minimum number of fields to be counted in the specific magnification to provide a good sample representation. The results for both methods were expressed as number of cells per liter and combined in a single dataset.

**Molecular identification**

Water samples were filtered using low vacuum filtration (<150 mmHg) on 0.2 μm isopore filters (Sartorius Stedim Biotech, Germany). DNA extraction was performed with the MoBio Power Soil kit (MoBio Inc. Carlsbad, CA, USA) following its standard protocol with minor modifications for filters processing. Sequencing of the V2-V3 region of the 18S rRNA gene was performed upon amplification using the primer pair 18S-82F (5′-GAAACTGCGAATGGCTC-3′) [3] and Euk-516r (5′-ACCAGACTTGCCCTCC-3′) for Eukaryotes [4]. Construction of libraries was performed by ‘Genes Diffusion’ company (Lille, France) [5] and amplicons were finally sequenced with Illumina MiSeq PE 2x300 (CNRS-UMR8199, Lille).

Processing of the resulting sequences, i.e. sequence assembly and quality control, was performed with the MOTHUR software (v 1.35) [6]. Only sequences with ≥480 bp, no ambiguous bases and homopolymers shorter than 8 bp were considered for further analysis. These sequences were aligned and classified using the SILVA SSU database (release 119) [7]. Chimeras were removed using the Uchime Software [8]. All sequences were binned into Operational Taxonomic Units (OTUs) and were clustered (average neighbour algorithm) at 97% sequence similarity. Single singletons, that appeared only once in the whole dataset, were removed using MOTHUR (v 1.35). Coverage values were calculated with MOTHUR (v 1.35). The batch of sequences from this study has been submitted in NCBI Short Read Archive under accession code PRJNA515026. Taxonomic classification was assigned using BLAT (Kent 2002) based on the Protist Ribosomal Reference (PR2) curated Database (built on Genbank 203; October 2014), containing 23,003 sequences [9].

For comparisons with the microscopic observations, that mostly took into account unicellular phototrophic algae, only OTUs clustering in protistan groups characterized as autotrophic and/or mixotrophic, that is representatives of Bacillariophyta, Chlorophyta, Haptophyta, Ochrophyta, dinoflaggelates [10] and unicellular Rhodophyta were kept for further analysis. For consistency with microscopic observations, OTUs were assigned to species and genera based on 18S sequences nucleotide similarity with the available PR2 sequences. More specifically, OTUs were classified in species level when 18S nucleotide similarity exceeded 97% and in genus level when the respective nucleotide similarity exceeded 92%. All OTUs assigned to specific species or genera were clustered in order to perform comparisons with microscopic observations.

**Simulation experiment to assess connectivity between sampling sites in the Aegean**

The simulation experiment was based on the surface velocity field (monthly means of magnitude and direction on a 2x2 km grid) for the sampling period (July 2014), by the ALERMO prognostic model [11]. Connectivity (probability of connection and time needed) is estimated for each pair of sampling sites. An imaginary water parcel is released from one site represented by one grid point and is dispersed stochastically following the velocity field. Stochasticity is introduced at each movement of the water parcel between grid points (a) for the direction of movement (± 2 standard deviations around the direction of the mean velocity at the grid point) and (b) for the velocity magnitude (± 2 standard deviations around the mean velocity at the grid point). If the parcel manages to arrive at the target site within a maximum time interval (3 months), the two sites are considered as connected and the time needed is calculated. This stochastic procedure is repeated (1000 times) for each pair of sites and probability of connection and time needed (mean and standard deviation) are estimated.

References

1. Utermohl, H. In press. Zur Ver vollkommung der quantitativen phytoplankton-methodik. *Mitteilung Int. Vereinigung Fuer Theor. unde Amgewandte Limnol.* **9**, 39.

2. Fournier, R. O. 1978 Membrane filtering. In *Monographs on Oceanographic Methods 6: Phytoplankton Manual* (ed A. Sournia), pp. 108–112. Paris: United Nations Educational, Scientific and Cultural Organization.

3. López-García, P., Philippe, H., Gail, F. & Moreira, D. 2003 Autochthonous eukaryotic diversity in hydrothermal sediment and experimental microcolonizers at the Mid-Atlantic Ridge. *Proc. Natl. Acad. Sci.* **100**, 697 LP-702. (doi:10.1073/pnas.0235779100)

4. Amann, R. I., Binder, B. J., Olson, R. J., Chisholm, S. W., Devereux, R. & Stahl, D. A. 1990 Combination of 16S rRNA-targeted oligonucleotide probes with flow cytometry for analyzing mixed microbial populations. *Appl. Environ. Microbiol.* **56**, 1919 LP-1925.

5. Voudanta, E., Kormas, K. A., Monchy, S., Delegrange, A., Vincent, D., Genitsaris, S. & Christaki, U. 2016 Mussel biofiltration effects on attached bacteria and unicellular eukaryotes in fish-rearing seawater. *PeerJ* **4**, e1829. (doi:10.7717/peerj.1829)

6. Schloss, P. D. et al. 2009 Introducing mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing and Comparing Microbial Communities. *Appl. Environ. Microbiol.* **75**, 7537 LP-7541. (doi:10.1128/AEM.01541-09)

7. Pruesse, E., Quast, C., Knittel, K., Fuchs, B. M., Ludwig, W., Peplies, J. & Glöckner, F. O. 2007 SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res.* **35**, 7188–7196. (doi:10.1093/nar/gkm864)

8. Edgar, R. C., Haas, B. J., Clemente, J. C., Quince, C. & Knight, R. 2011 UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* **27**, 2194–2200. (doi:10.1093/bioinformatics/btr381)

9. Guillou, L. et al. 2013 The Protist Ribosomal Reference database (PR2): a catalog of unicellular eukaryote small sub-unit rRNA sequences with curated taxonomy. *Nucleic Acids Res.* **41**, D597–D604. (doi:10.1093/nar/gks1160)

10. Genitsaris, S., Monchy, S., Breton, E., Lecuyer, E. & Christaki, U. 2016 Small-scale variability of protistan planktonic communities relative to environmental pressures and biotic interactions at two adjacent coastal stations . *Mar. Ecol. Prog. Ser.* **548**, 61–75.

11. Mavropoulou, A. M., 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)