



OPEN

## Artificial neural network analysis of microbial diversity in the central and southern Adriatic Sea

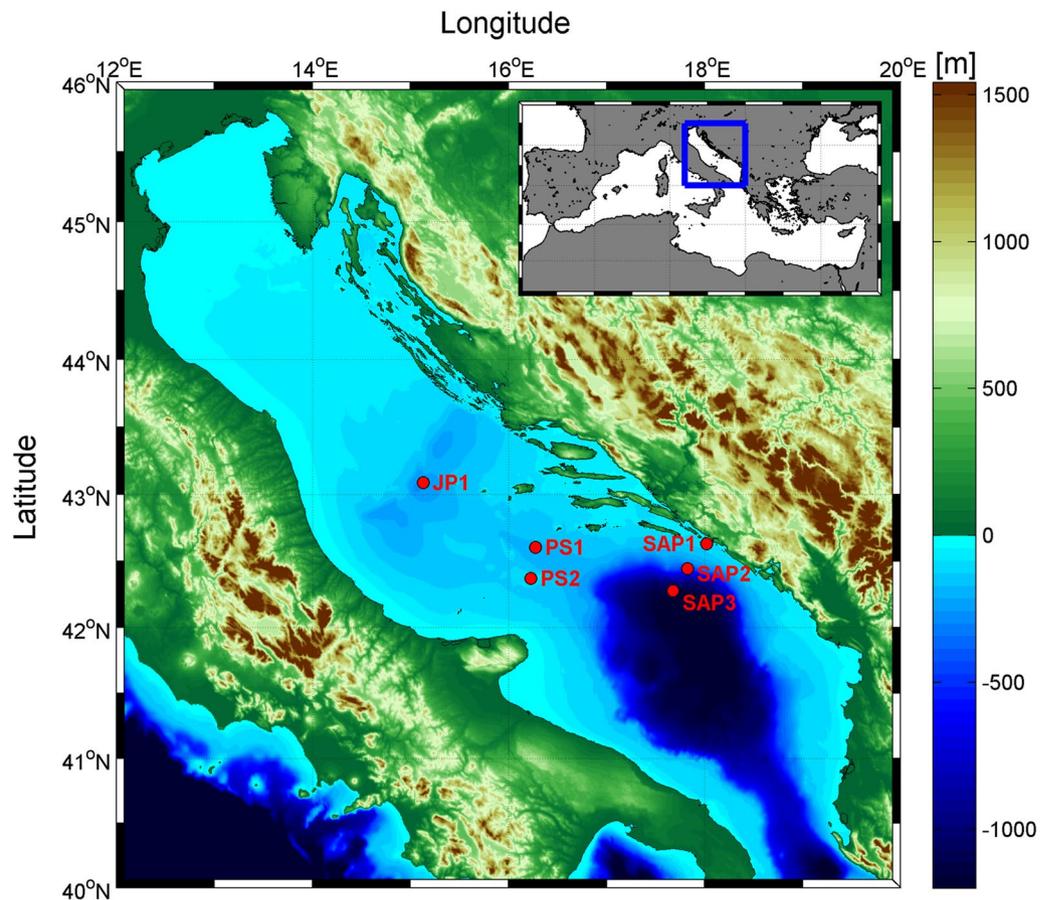
Danijela Šantić<sup>1</sup>✉, Kasia Piwosz<sup>2</sup>, Frano Matić<sup>1</sup>, Ana Vrdoljak Tomaš<sup>1</sup>, Jasna Arapov<sup>1</sup>, Jason Lawrence Dean<sup>3</sup>, Mladen Šolić<sup>1</sup>, Michal Koblížek<sup>3,4</sup>, Grozdan Kušpilić<sup>1</sup> & Stefanija Šestanović<sup>1</sup>

Bacteria are an active and diverse component of pelagic communities. The identification of main factors governing microbial diversity and spatial distribution requires advanced mathematical analyses. Here, the bacterial community composition was analysed, along with a depth profile, in the open Adriatic Sea using amplicon sequencing of bacterial 16S rRNA and the Neural gas algorithm. The performed analysis classified the sample into four best matching units representing heterogenic patterns of the bacterial community composition. The observed parameters were more differentiated by depth than by area, with temperature and identified salinity as important environmental variables. The highest diversity was observed at the deep chlorophyll maximum, while bacterial abundance and production peaked in the upper layers. The most of the identified genera belonged to Proteobacteria, with uncultured AEGEAN-169 and SAR116 lineages being dominant Alphaproteobacteria, and OM60 (NOR5) and SAR86 being dominant Gammaproteobacteria. Marine *Synechococcus* and *Cyanobium*-related species were predominant in the shallow layer, while *Prochlorococcus* MIT 9313 formed a higher portion below 50 m depth. Bacteroidota were represented mostly by uncultured lineages (NS4, NS5 and NS9 marine lineages). In contrast, Actinobacteriota were dominated by a candidatus genus *Ca. Actinomarina*. A large contribution of Nitrospinae was evident at the deepest investigated layer. Our results document that neural network analysis of environmental data may provide a novel insight into factors affecting picoplankton in the open sea environment.

Microorganisms drive the biogeochemical fluxes in marine environments. Since various microbial processes are conducted by different species, knowledge of bacterial community composition (BCC) is essential to understand the functioning of a particular ecosystem [e.g.<sup>1</sup>]. Knowing that the BCC is closely associated with environmental conditions, many studies conducted in different marine ecosystems, including the Mediterranean Sea, have linked these characteristics using different scales<sup>2–8</sup> and highlight the importance of interdependencies, particularly between free-living bacteria in the upper layers of the epipelagic zone<sup>9</sup>.

In our previous work we focused on the role of heterotrophic bacteria, picoautotrophs and aerobic anoxygenic phototrophic (AAP) bacteria in the production and transfer of biomass and energy through the microbial food web in the coastal and open Adriatic Sea<sup>10–16</sup>. The Adriatic Sea is a highly dynamic environment, where the picoplankton community is exposed to sudden physio-chemical changes<sup>14</sup>. Its open waters are generally phosphorus (P) and nitrogen (N) limited<sup>17</sup>, however they can occasionally be enriched with nutrients and organic substances from advected water masses, leading to changes in the food web structure<sup>10,14</sup>. Indeed, increased microbial diversity was found to be associated with deep winter convection in the South Adriatic Sea<sup>18,19</sup>. The herbivorous food web dominates during such periods of the nitrate-rich mixed water column, but it gradually changes to a multivorous food web, which in turn changes to the typical microbial food web during the stratified period<sup>20</sup>. Moreover, unusually high salinity values for this area have been recorded since 2016, which have been shown to adversely affect the picoplankton community<sup>21</sup>. In addition to the bacterial community research mentioned above, the recent use of Self-organizing map analysis has elucidated the response of the microbial food web to changes in seawater temperature<sup>10</sup>, as mentioned previously. Different artificial neural network methods have already been used successfully for several biological applications<sup>10,22–24</sup>.

<sup>1</sup>Institute of Oceanography and Fisheries, Šetalište Ivana Meštrovića 63, POB 500, 21000 Split, Croatia. <sup>2</sup>National Marine Fisheries Research Institute, Kołłątaja 1, 81-332 Gdynia, Poland. <sup>3</sup>Centre Algatech, Institute of Microbiology of the Czech Acad. Sci., 379 81 Třeboň, Czech Republic. <sup>4</sup>University of South Bohemia, Faculty of Science, Braníšovská 1760, Ceske Budejovice, Czech Republic. ✉email: segvic@izor.hr



**Figure 1.** (A) Map of the Mediterranean Sea. (B) Study area and sampling stations in the Jabuka Pit (JP1), the Palagraža Sill (PS1, PS2) and the South Adriatic Pit (SAP1, SAP2, SAP3). The software MATLAB, version 7.10.0 (R2018), Natick, Massachusetts: The MathWorks Inc. (2018) (<https://www.mathworks.com/>) was used to generate the figure.

We propose that the picoplankton community display a heterogenic response to changes associated with different environmental factors in the open sea areas of the central and southern Adriatic Sea, throughout the stratified water column. To test this hypothesis, we used a Neural gas algorithm to identify bacterial community response in terms of abundance and community composition. The survey was conducted in three open sea areas, the South Adriatic Pit and, for the first time, the Jabuka Pit and Palagraža Sill. The novelty of this study is community characterisation and its relation with environmental data using the Neural gas algorithm<sup>25</sup>. The Neural gas algorithm, without the use of prior knowledge about the topological structure of data, quantified the manifold by distributing neural units homogeneously over the input space<sup>26</sup>.

## Material and methods

**Sampling sites.** The Adriatic Sea is an elongated semi-enclosed basin of the Eastern Mediterranean Sea, connected with the Ionian Sea through the Strait of Otranto [e.g.<sup>18,27</sup>]. According to its morphology and bathymetry, it is divided into three sub-basins (northern, central and southern). Jabuka Pit (JP) is located in the central Adriatic, with depths up to approximately 270 m. Palagraža Sill (PS) (170 m deep) separates the Jabuka Pit from the much deeper South Adriatic Pit (SAP) (1250 m deep).

In total, 43 samples (from the surface to the bottom) were collected from 6 sites located in the area of Jabuka Pit (JP1), Palagraža Sill (PS1, PS2) and the South Adriatic Pit (SAP1, SAP2, SAP3) during September/October 2016 (29.09.-03.10) (Fig. 1).

**Environmental parameters.** Temperature and salinity were measured using a SeaBird 25 CTD profiler with accuracy  $\pm 0.01$  °C and  $\pm 0.02$ , respectively. Concentrations of inorganic nutrients were determined using the autoanalyser-modified method<sup>28</sup>. Detection limits were  $0.001 \mu\text{mol L}^{-1}$  for  $\text{NO}_2^-$ ,  $0.01 \mu\text{mol L}^{-1}$  for  $\text{NO}_3^-$ ,  $0.0002 \text{ mg L}^{-1}$  for  $\text{NH}_4^+$  and  $0.02 \mu\text{mol L}^{-1}$  for soluble reactive phosphorus (SRP).

**Criteria for nutrient limitation.** Nutrient limitation at the studied sites was assessed based on stoichiometry defined in<sup>29</sup>. Microorganisms in samples with concentrations of SRP  $< 0.1 \mu\text{mol L}^{-1}$ , and with N/P ratio  $> 22$

and Si/P ratio > 22 were considered to be limited by phosphorus. The criteria for nitrogen limitation was total inorganic nitrogen (TIN) < 1  $\mu\text{mol L}^{-1}$ <sup>29</sup>.

**Microbial analysis.** *Sample collection, DNA extraction and sequencing.* Samples for DNA sequencing were taken at various depths based on physico-chemical parameters: euphotic zone, oxygen concentration maximum, chlorophyll *a* maximum and deep layers. These conditions were different for each station: SAP1: 40 m; SAP2: 5 m, 20 m and 33 m; SAP3: 5 m, 20 m, 78 m and 100 m; JP1: 0 m, 5 m, 10 m, 20 m, 30 m, 50 m, 66 m and 75 m; PS1: 5 m, 25 m and 65 m; and PS2: 5 m, 16 m and 50 m. 0.5–1 L of seawater was filtered through 0.22  $\mu\text{m}$  polycarbonate filters (Merck Millipore, USA). The filters were flash-frozen in liquid nitrogen and stored at  $-80^\circ\text{C}$  for < 1 month. DNA was extracted using the Power Water DNA isolation kit (MO BIO Laboratories, Inc., Carlsbad, CA, USA).

The V3–V4 region of the 16S rRNA gene was amplified using 341F–785R primers<sup>30</sup> and Phusion HotStart II High Fidelity PCR MasterMix (Thermo Scientific). PCR conditions were the following: initial denaturation for 3 min at  $98^\circ\text{C}$ , 25 cycles of  $98^\circ\text{C}$  for 10 s,  $60^\circ\text{C}$  for 20 s,  $72^\circ\text{C}$  for 20 s, final elongation at  $72^\circ\text{C}$  for 5 min. The amplicons were purified from the gel using Wizard SV Gel and PCR clean system (Promega) and quantified with Qubit dsDNA HS assay. Library preparation (2  $\times$  250 bp) and sequencing on a MiSeq Illumina were performed by the Genomic Service of the Universitat Pompeu Fabra, Barcelona, Spain. The number of raw reads per sample ranged from 74,051 to 193,832 (mean  $\pm$  st. dev. 100,931  $\pm$  24,420).

*Sequence analysis.* The quality of reads was evaluated using FastQC v0.11.7. The sequences of the primers were trimmed using cutadapt v1.16<sup>31</sup> and subsequent analyses were carried out in the R/Bioconductor environment using the dada2 v1.6 package<sup>32</sup>. Low quality sequences were filtered out with the filterAndTrim function (maxN = 0 (no N allowed), maxEE = c(2, 2), truncQ = 2, rm.phix = TRUE). After quality filtering and denoising, the number of reads per sample ranged from 70,750 to 187,751 (mean  $\pm$  st. dev. 96,216  $\pm$  23,563). After error learning and sequence variants interference, the sequences were merged, and the amplicon sequence variant (ASV) table was produced. Chimeric sequences were removed using the default removeBimeraDenovo function with the “consensus” method. The taxonomic assignment was performed with the naive Bayesian classifier method using the SILVA 132 database<sup>33,34</sup>. The final ASV table contained from 68,483 to 185,399 reads per sample (mean  $\pm$  st. dev. 93,317  $\pm$  23,087). The total number of observed ASVs was 4102, and it ranged from 391 to 1426 in specific samples. Removal of non-bacterial ASVs, ASVs classified as chloroplasts or mitochondrion, and those seen less than three times in at least 20% of the samples reduced the number of ASVs from 4102 to 714 (252–454 per sample), and the number of reads to 1,712,027. The phyloseq v1.30 package<sup>35</sup> was used for this operation. Figures displaying characteristic genus and phylum contribution of the bacterial community composition were generated in R statistical software<sup>36</sup> (<https://cran.r-project.org/>, v. 3.6.3), using ggplot2 v. 3.3.3 package<sup>37</sup>, and they were assembled in Inkscape (<https://inkscape.org/>, v. 0.92).

*Flow cytometry.* Picoplankton abundance was determined using flow cytometry<sup>38</sup>. For the count of autotrophic cells, 2 mL of sample were fixed with glutaraldehyde (0.5% final concentration) and stored at  $-80^\circ\text{C}$  until analysis (5–10 days). Autotrophic cells were classified into three groups: *Synechococcus*, *Prochlorococcus* and picoeukaryotes, distinguished according to light scattering, red emission of cellular chlorophyll content and orange emission of phycoerythrin-rich cells. Samples for bacteria and heterotrophic nanoflagellate (HNF) abundance were fixed with formaldehyde (2% final concentration) and stored at  $4^\circ\text{C}$  until analysis (5–10 days). They were stained with Sybr Green-I, and three groups, namely, high nucleic acid content (HNA) bacteria, low nucleic acid content (LNA) bacteria and HNF were determined using flow cytometry<sup>38</sup>. The samples were processed on a Beckman Coulter EPICS XL-MCL flow cytometer with a high flow rate from 1 to 1.2  $\mu\text{L s}^{-1}$ , and the analysed volume was calculated according to the acquisition time.

*Epifluorescence microscopy.* Aerobic Anoxygenic Phototrophs (AAP) were fixed with formaldehyde (0.5% final concentration) and stored at  $-80^\circ\text{C}$ . AAP cells were collected on 0.2  $\mu\text{m}$  polycarbonate filters by filtration and stained with 4',6-diamidino-2-phenylindole (DAPI, 1  $\mu\text{g mL}^{-1}$  final concentration) using a 3:1 mixture of Citifluor AF1 and Vectashield after drying<sup>39</sup>. AAP were counted using an Olympus BX51 microscope equipped with an Olympus UPlanSApo 100 $\times$ /1.40 OIL, IR objective and software for image analysis (CellSens). The microscope was equipped with a Hg Lamp U-LH100H6 for excitation. Fluorescent images were taken using an XM10-IR camera. Three epifluorescent filter sets were used: DAPI, IR and red. The chlorophyll *a* signal was subtracted from the IR image to obtain a net count of AAP cells.

*Bacterial production.* Bacterial cell production was estimated by measuring the incorporation of  $^3\text{H}$ -thymidine into bacterial DNA<sup>40</sup>, which was added to 10 mL samples at a final concentration of 10 nmol (specific activity: 86 Ci mmol<sup>-1</sup>). Triplicate samples and a formaldehyde-killed absorption control (0.5% final concentration) were incubated for 1 h. The incubations were stopped with formaldehyde (0.5% final concentration). The thymidine samples were extracted with ice-cold trichloroacetic acid (TCA). The TCA-insoluble fraction was collected on 0.2  $\mu\text{m}$  pore size polycarbonate filters. Samples were treated with Cocktail Filter Count (Perkin Elmer) before analysis on a liquid scintillation analyser TriCarb 4910 TR (Perkin Elmer).

**Statistical analysis.** *Calculation of diversity and evenness.* The PRIMER 6 software package<sup>41</sup> was used to calculate diversity indices: Shannon (*H'*) and Pielou's (*J'*), based on the untransformed and not rarefied ASV table.

**Neural gas.** Neural gas is an artificial neural network trained by unsupervised learning<sup>26,42</sup>. The method reduces the dimensionality of the data space to a certain number of neurons (Best-Matching Units, BMUs) distributed in the data space like a gas. During the learning process in a predefined number of adaptation steps together with the size of the adaptation step and neighbourhood impact range, the neural gas algorithm quantifies the manifold by distributing the BMUs over the relevant part of the input data space and tries to minimize the representation error. As a result, the connections between BMUs are weak without a predefined topological structure in the manifold. The lack of a fixed topological structure is the main difference from the commonly used Self-organizing maps that maximize the similarity between the BMUs and the measured data<sup>10,20,22</sup>. Weak connections between BMUs make neural gas a weak smoothing algorithm and more suitable for detecting anomalies and outlier data.

During the adaptation process, the neural gas was trained with 500 training epochs, with default values for initial step size (0.5) and an initial decay constant (2.5), following Martinetz et al.<sup>25</sup>. The number of BMUs was chosen to be 5 using a coefficient of variation of the SSIntra quantization error<sup>43–45</sup>. The data are scattered in PC1-PC2 space, making Principal Component Analysis (PCA) unsuitable for our analysis (Supplementary Fig. S1). The SOM-BMU distribution are always grouped in the origin and form more/less regular rectangular distribution, missing to model the data distant from centre. The Neural gas spreads enough to include all the data. As a result, the one extreme data is modelled with one BMU.

Two different NG analyses were conducted, picoplankton community (PIC) analysis and BCC. In the PIC-BMU analysis, the elements of the data vectors were separately normalized biotic parameters (total prokaryotes, *Synechococcus*, *Prochlorococcus*, picoeukaryotes, heterotrophic nanoflagellates, aerobic anoxygenic phototrophs, high and low nucleic acid bacteria, bacterial production), while in the BCC-BMU analysis they were separately normalized relative abundances of genera (bacterial community composition) obtained from sequencing. Both models were run using 22 different data vectors as sampling stations. The results of the analysis were PIC-BMU/BCC-BMU neural gas clusters. SOM Toolbox version 2.0 for MATLAB used in this study was developed by E. Alhoniemi, J. Himberg, J. Parhankangas, and J. Vesanto at Helsinki University of Technology, Finland, and is available at <http://www.cis.hut.fi/projects/somtoolbox>.

## Results

**Physico-chemical conditions.** Sampling was performed at 6 stations representing the physical and chemical characteristics of the investigated area (Supplementary Table S1). Thermohaline properties were the result of horizontal advection of above-average salinities driven by a North Ionian cyclonic gyre controlled by the Adriatic Ionian Bimodal Oscillating System<sup>46</sup>. September and the whole summer of 2016 was characterized by extremely high temperatures, and precipitation in the climatologic expected range. A cyclone with a cold front followed by a strong Bora wind passed over the Adriatic a week before the cruise, in the period between the 16th and 20th of September 2016. Heat and mass exchange in the air-sea boundary layer were responsible for the characteristic vertical thermohaline profiles measured in late summer. Over the investigated area, the mixed layer depth located between 20 and 25 m was horizontally homogenous. The coldest water mass (temperature 12.94 °C, salinity 38.68) was located at the bottom of Jabuka Pit.

**Abundance of bacteria, autotrophic picoplankton and AAP.** Bacterial abundances ranged between 0.05 and  $0.46 \times 10^6$  cell mL<sup>-1</sup> in all three areas, with a slightly higher average value in Jabuka Pit ( $0.31 \times 10^6$  cell mL<sup>-1</sup>). The bacterial abundances were the highest in the upper layers down to the 50 m deep layer and showed a decreasing trend towards the bottom (Supplementary Table S2). The portion of HNA bacteria ranged from 37.8 to 73.12% (on average 51.27%), with the prevalence of HNA over the LNA group below the epipelagic layer.

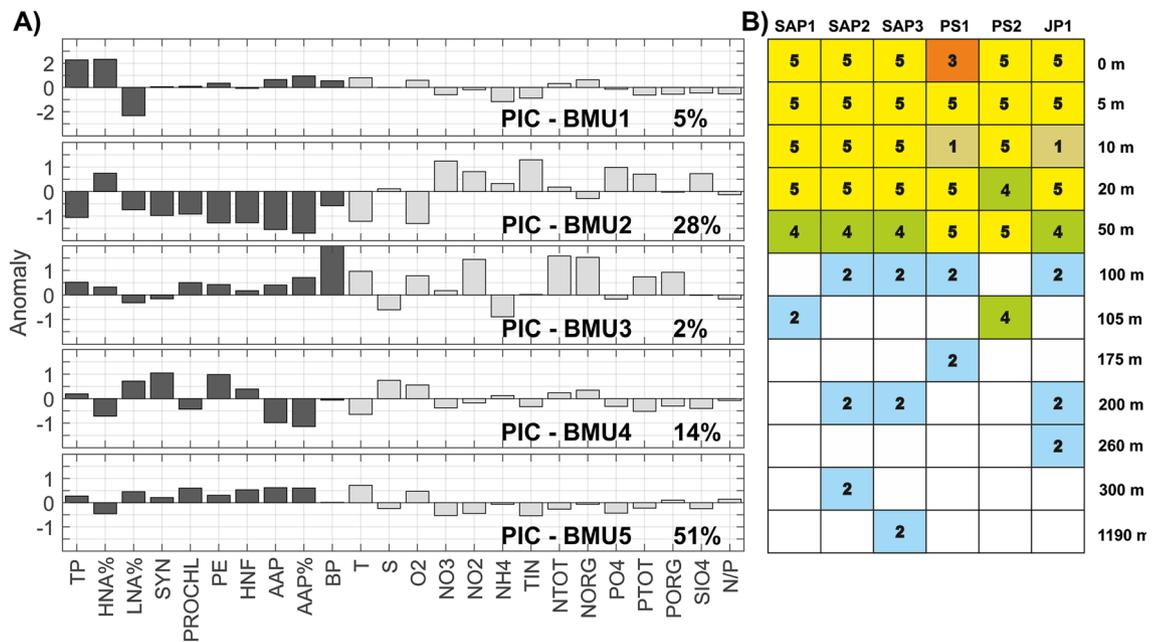
Marine *Synechococcus* dominated the autotrophic picoplankton community with abundances ranging from 0.08 to  $23.86 \times 10^3$  cell mL<sup>-1</sup>. The presence of *Prochlorococcus* cells was also detected in all samples in a range from a few cells to  $1.33 \times 10^3$  cell mL<sup>-1</sup>. Picoeukaryotes also showed a similar range from a few cells to  $0.83 \times 10^3$  cell mL<sup>-1</sup>. The highest abundances of picophytoplankton were measured in the upper 50 m, with the exception of the Palagruža Sill (PS) area, where an increase in abundance was observed at 100 m depth. Bacterial production ranged from  $0.2 \times 10^4$  to  $0.36 \times 10^4$  cell mL<sup>-1</sup> h<sup>-1</sup>, with increased values in the shallow layers and a mostly uniform vertical distribution in the water column (Supplementary Table S2).

AAP bacteria abundance ranged from  $0.9 \times 10^3$  to  $22.3 \times 10^3$  cell mL<sup>-1</sup>, thus constituting 0.42% to 6.83% of the bacteria. Their highest average contribution was observed in the South Adriatic Pit (4.11%), while on the vertical scale, their highest contribution was observed in the upper 20 m of the seawater (see Supplementary Table S2).

**Relationship between the picoplankton community and environmental parameters.** Based on biological characteristics (total prokaryotes, *Synechococcus*, *Prochlorococcus*, picoeukaryotes, heterotrophic nanoflagellates, aerobic anoxygenic phototrophs, high and low nucleic acid bacteria, bacterial production), we distinguished five picoplanktonic clusters (PIC-BMUs) and then searched for explanations of the observed patterns (Fig. 2A,B). The mean values of biological and physico-chemical parameters for each cluster are shown in Table 1.

PIC-BMU1 described a very rare pattern, found in only two samples from 10 m depth in Palagruža Sill and Jabuka Pit. They were characterised by the highest abundances of total prokaryotes with a dominance of HNA and elevated AAP abundance. These samples were unique in terms of hydrological parameters, as they represented an N-limited environment (TIN < 1 µmol L<sup>-1</sup>, TIN/SRP < 10, Si/TIN > 1), where the water temperature was high (22.02 °C).

PIC-BMU2 described 28% of the samples. This cluster included the picoplankton community from the water column at and below 100 m and is characterised by the dominance of HNA in total prokaryote abundance and



**Figure 2.** (A) Bar plot representation of biotic (black) and abiotic (grey) parameters for neural gas best-matching units (picoplankton-PIC-BMUs) with relative frequency appearance for each neuron. TP-total prokaryotes, SYN-*Synechococcus*, PROCHL-*Prochlorococcus*, PE-picoeukaryotes, HNF-heterotrophic nanoflagellates, AAP-aerobic anoxygenic phototrophs, AAP%-portion of AAP, HNA% percentage of high nucleic acid content bacteria, LNA%-percentage of low nucleic acid content bacteria-LNA%, BP-bacterial production. (B) Water column distribution of Neural gas best-matching units (BMU, labels with numbers, and stained with a different colour for clearance, coloured non-labelled squares shows clarity) for measuring stations (SAP1-3, PS1-2 and JP1). The software MATLAB, version 7.10.0 (R2018). Natick, Massachusetts: The MathWorks Inc. (2018) (<https://www.mathworks.com/>) was used to generate the figure.

a decrease of all other picoplankton parameters. This layer describes the lowest seawater temperature (14.42 °C) and the highest concentrations of nitrates, nitrites, ammonium ions, silicates and SRP.

PIC-BMU3 included only one surface sample from Palagruža Sill with the highest values of bacterial production and AAP portion. The measured temperature was 22.62 °C and the highest values of N-organic compounds were detected in this sample. This sample was P-limited.

PIC-BMU4 described the pattern mostly from 50 m, with the highest contribution of LNA to total prokaryote abundance, the highest abundances of *Synechococcus* and picoeukaryotes, and the lowest AAP abundance. These samples were collected below the well-developed thermocline in the area of deep chlorophyll maximum (DCM) and were characterised by a P-limited environment with elevated salinity values compared to the other BMU clusters.

PIC-BMU5 described the most frequent pattern in our samples, the frequency can be attributed to the sampling effort rather than to certain general features of the sampling area. It grouped the samples from the first 50 m with high values for all measured biological parameters, except for HNA contribution to total prokaryote abundance. This layer is characterised by a temperature of approximately 21.70 °C and P limitation.

The results of the above analyses (Fig. 2A,B) document that all observed parameters were differentiated by depth rather than location. For at least 81% of the data, the negative impact of salinity is visible through the opposite direction of its anomaly values compared to most other picoplankton variables (heterotrophic bacteria, HNA%, LNA%, *Synechococcus*, *Prochlorococcus*, picoeukaryotes, AAP, bacterial production). Finally, our results suggest that an increase in temperature had a positive impact primarily in terms of high picoplankton abundances and bacterial productivity, given that the anomaly values of total prokaryotes, bacterial production and temperature display the same direction.

**Bacterial community composition.** Proteobacteria (mainly Alpha- and Gammaproteobacteria) and Cyanobacteria were the most abundant phyla in all samples, followed by Bacteroidota and Actinobacteriota. The changes between sites and depths at phylum level were minimal, with higher relative abundances of Planctomycota and Myxococcota observed at Jabuka Pit at 75 m, and of Nitrospinota and Myxococcota in the South Adriatic Pit at 100 m (Fig. 3A). The changes in bacterial communities were more conspicuous with depth rather than between the different basins, as will be shown below at the genus level for the most abundant phyla.

‘Ca.’ Actinomarina and uncultured Microbacteriaceae (Actinobacteriota) co-dominated in the upper layers down to about 20 m (Fig. 3B). The relative abundance of Microbacteriaceae substantially decreased in deeper layers, where it was replaced either by ‘Ca.’ Actinomarina or by uncultured Microtrichaceae from the Sva0996 lineage. The latter dominated in the deepest layers at the South Adriatic Pit (Fig. 3B).

	Neural gas clusters				
	PIC-BMU1	PIC-BMU2	PIC-BMU3	PIC-BMU4	PIC-BMU5
TP ( $10^6$ cell $\text{mL}^{-1}$ )	0.45 ± 0.02	0.14 ± 0.06	0.28	0.25 ± 0.09	0.26 ± 0.06
HNA (%)	65.39 ± 11.17	54.05 ± 5.54	51	43.62 ± 6.26	45.46 ± 3.38
LNA (%)	34.61 ± 11.17	45.95 ± 5.54	49	56.38 ± 6.23	54.54 ± 3.38
SYN ( $10^3$ cell $\text{mL}^{-1}$ )	6.77 ± 2.01	0.68 ± 0.62	5.51	12.58 ± 4.07	7.68 ± 5.80
PROCHL ( $10^3$ cell $\text{mL}^{-1}$ )	1.24 ± 0.31	0.36 ± 0.21	1.59	0.78 ± 0.64	1.68 ± 0.81
PE ( $10^3$ cell $\text{mL}^{-1}$ )	0.85 ± 0.08	0.27 ± 0.25	0.87	1.07 ± 0.20	0.83 ± 0.20
HNF ( $10^3$ cell $\text{mL}^{-1}$ )	0.86 ± 0.33	0.33 ± 0.17	0.97	1.06 ± 0.49	1.13 ± 0.24
AAP ( $10^4$ cell $\text{mL}^{-1}$ )	1.53 ± 0.31	0.18 ± 0.17	1.38	0.53 ± 0.42	1.51 ± 0.36
AAP (%)	5.79 ± 0.07	0.83 ± 0.66	5.34	1.89 ± 0.91	5.15 ± 0.73
BP ( $10^4$ cell $\text{h}^{-1}$ $\text{mL}^{-1}$ )	0.10 ± 0.05	0.04 ± 0.01	0.36	0.06 ± 0.03	0.07 ± 0.02
Temperature ( $^{\circ}\text{C}$ )	22.03 ± 0.62	14.43 ± 1.02	22.62	16.57 ± 1.37	21.70 ± 2.16
Salinity	38.81 ± 0.15	38.82 ± 0.10	38.75	38.87 ± 0.05	38.78 ± 0.08
O <sub>2</sub> (%)	113.44 ± 5.42	96.21 ± 2.84	115.16	113.12 ± 11.56	112.35 ± 4.05
NO <sub>3</sub> <sup>-</sup> ( $\mu\text{mol L}^{-1}$ )	0.21 ± 0	1.77 ± 0.93	0.86	0.40 ± 0.47	00.20.27 ±
NO <sub>2</sub> <sup>-</sup> ( $\mu\text{mol L}^{-1}$ )	0.06 ± 0.01	0.12 ± 0.09	0.16	0.06 ± 0.02	0.05 ± 0.03
NH <sub>4</sub> <sup>+</sup> ( $\mu\text{mol L}^{-1}$ )	0.12 ± 0.10	0.48 ± 0.23	0.19	0.44 ± 0.24	0.39 ± 0.25
TIN ( $\mu\text{mol L}^{-1}$ )	0.38 ± 0.09	2.37 ± 0.84	1.21	0.89 ± 0.39	0.70 ± 0.36
N tot ( $\mu\text{mol L}^{-1}$ )	8.48 ± 3.22	8.13 ± 1.40	11.59	8.26 ± 3.16	7.02 ± 2.62
N org ( $\mu\text{mol L}^{-1}$ )	8.09 ± 0.04	5.75 ± 0.06	10.38	7.37 ± 0.03	6.32 ± 0.02
SRP ( $\mu\text{mol L}^{-1}$ )	0.05 ± 0.03	0.11 ± 0.08	0.05	0.04 ± 0.02	0.04 ± 0.05
P tot ( $\mu\text{mol L}^{-1}$ )	0.13 ± 0.07	0.22 ± 0.07	0.22	0.14 ± 0.04	0.16 ± 0.06
P org ( $\mu\text{mol L}^{-1}$ )	0.08 ± 0.06	0.11 ± 1.29	0.17	0.10 ± 0.22	0.12 ± 0.32
SiO <sub>4</sub> <sup>-</sup> ( $\mu\text{mol L}^{-1}$ )	0.91 ± 0.06	1.85 ± 1.29	1.27 ± 0.01	0.96 ± 0.22	1.08 ± 0.32

**Table 1.** Characteristics of biological (abundances of total prokaryotes-TP, *Synechococcus*-SYN, *Prochlorococcus*-PROCHL, picoeukaryotes-PE, heterotrophic nanoflagellates-HNF, aerobic anoxygenic phototrophs(AAP); contributions-(%) of AAP, High nucleic acid content bacteria-HNA and Low nucleic acid content bacteria-LNA%; and bacterial production-BP) and environmental factors in the sampling terms assigned to the neural gas clusters. Mean values ± standard deviation from all the samples cluster in a PIC-BMU are displayed.

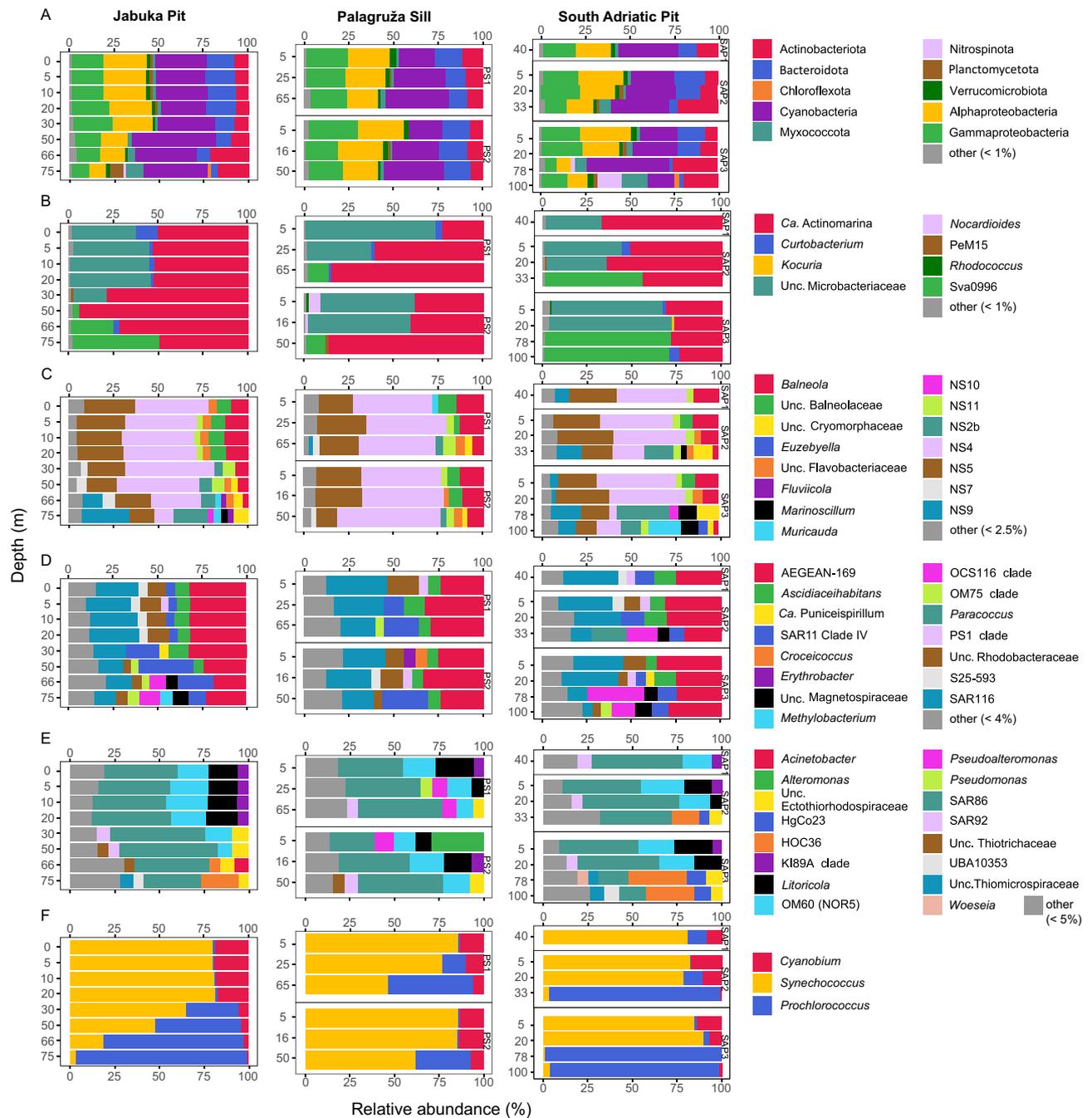
Uncultured NS4 and NS5 lineages of Bacteroidetes co-dominated up to approximately 50 m depth, with minor contributions of genus *Balneola* and uncultured Balneolaceae, Cryomorphaceae and, especially in the Jabuka Pit, Flavobacteriaceae (Fig. 3C). Uncultured lineages NS2b and NS9 became relatively more abundant in deeper layers, together with genera *Marinoscillum* and *Muricauda*.

Proteobacteria was the most diverse phylum in terms of the number of detected genera. Thus, Alphaproteobacteria and Gammaproteobacteria are shown separately (Fig. 3D,E, respectively). The members of Alphaproteobacteria with the highest relative abundance in the euphotic zone (~20 m) in all basins were affiliated with uncultured AEGEAN-169 and SAR116 lineages, with minor contributions of *Ascidiaecihabitans*, SAR11 clade IV, uncultured Rhodobacteraceae and S25–593 lineage. Uncultured Magnetospiraceae, *Methylobacterium* and OCS116 clade were more abundant in deeper layers (Fig. 3D). Moreover, 29 of genera with a relative abundance < 4% contributed 10–25% of all reads in all the samples.

The most abundant Gammaproteobacterial lineages in the upper layers down to about 50 m were OM60 (NOR5) and SAR86, with a minor contribution of KI89A clade and genus *Litoricola* (Fig. 3E). The exception was 5 m depth at station PS2, Palagruža Sill, where a higher relative abundance of *Alteromonas* was observed. Higher relative abundances of uncultured HOC36 and HgCo23lineages, and Ectothiorhodospiraceae and Thiomicrospiraceae were observed below 75 m. Moreover, a substantial proportion of Gammaproteobacterial reads consisted of numerous other genera with individual contributions < 5% (Fig. 3E). Cyanobacteria were represented by only three genera, with distinct distribution at different depths. Marine *Synechococcus* and *Cyanobium*-like sequences dominated in the euphotic zone, while *Prochlorococcus* in deeper waters (Fig. 3F).

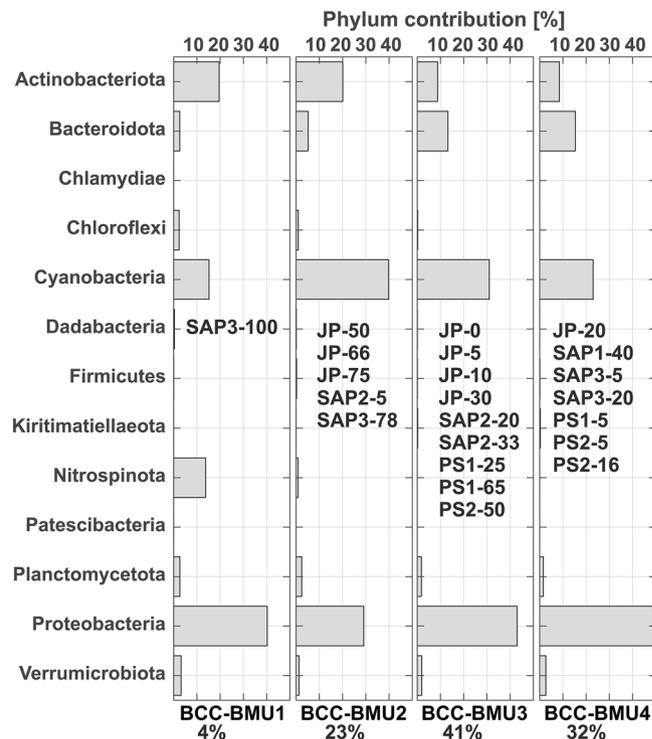
The neural gas analysis grouped the samples into four best matching units (BCC-BMU1, BCC-BMU2, BCC-BMU3, BCC-BMU4) representing heterogenic patterns of bacterial community composition that differed already at the phylum level (Fig. 4). Moreover, genera that showed higher importance in the samples were not always important for delineation of the BCC-BMU. The samples are grouped by depth rather than by basin.

BCC-BMU1 included a single sample from station SAP3 from 100 m depth. It was unique because of the significant contribution of Nitrospina (LS-NOB), the high contribution of Actinobacteriota, and the lower contribution of Cyanobacteria and Bacteroidota (Fig. 4). The differences were even more pronounced at genus level (Figs. 5, 6). Actinobacteriota were dominated by *Curtobacterium* and the uncultured Sva0996 lineage. An interesting pattern was observed for Bacteroidota that, despite the lower relative abundance, show higher diversity at the genus level, with a relatively equal contribution of all genera, e.g. *Mauricauda*, *Marinoscillum* or



**Figure 3.** Bacterial community composition (BCC) in the Jabuka Pit (JP), Palagruža Sill (PS) and South Adriatic Pit (SAP). For the most abundant phyla or classes BCC as the genus level is shown. (A) Relative contribution of the bacterial phyla and proteobacterial classes (B) Genus contribution in the Actinobacteriota (C) Genus contribution in the Bacteroidota (D) Genus contribution in the Alphaproteobacteria (E) Genus contribution in the Gammaproteobacteria (F) Genus contribution in the Cyanobacteria. Category 'other' groups taxa with the relative abundances below the threshold given in the parentheses. The figure was generated in R statistical software<sup>35</sup> (<https://cran.r-project.org/>, v. 3.6.3), using ggplot2 v. 3.3.3 package<sup>36</sup>, and they were assembled in Inkscape (<https://inkscape.org/>, v. 0.92).

*Euzebyella*. Similarly, Alphaproteobacteria were very diverse in this BMU, with equal contributions of *Erythrobacter*, *Maricaulis*, *Martella*, uncultured OM75 clade, *Tistilia* and *Tistrella*. In contrast, Gammaproteobacteria were less diverse in this deep-water BMU than in the other units, with a higher contribution of uncultured UP05 lineage and genus *Woeseia* (Fig. 5). Planctomycetota were dominated by the uncultured JL-ENTP-F27 lineage, with a minor contribution of *Rubripirellula*, while Verrucomicrobiota was dominated by *Roseibacillus* with a minor contribution of the uncultured SCGC\_AAA164-Eo4 lineage. *Prochlorococcus* was the dominant genus of Cyanobacteria.

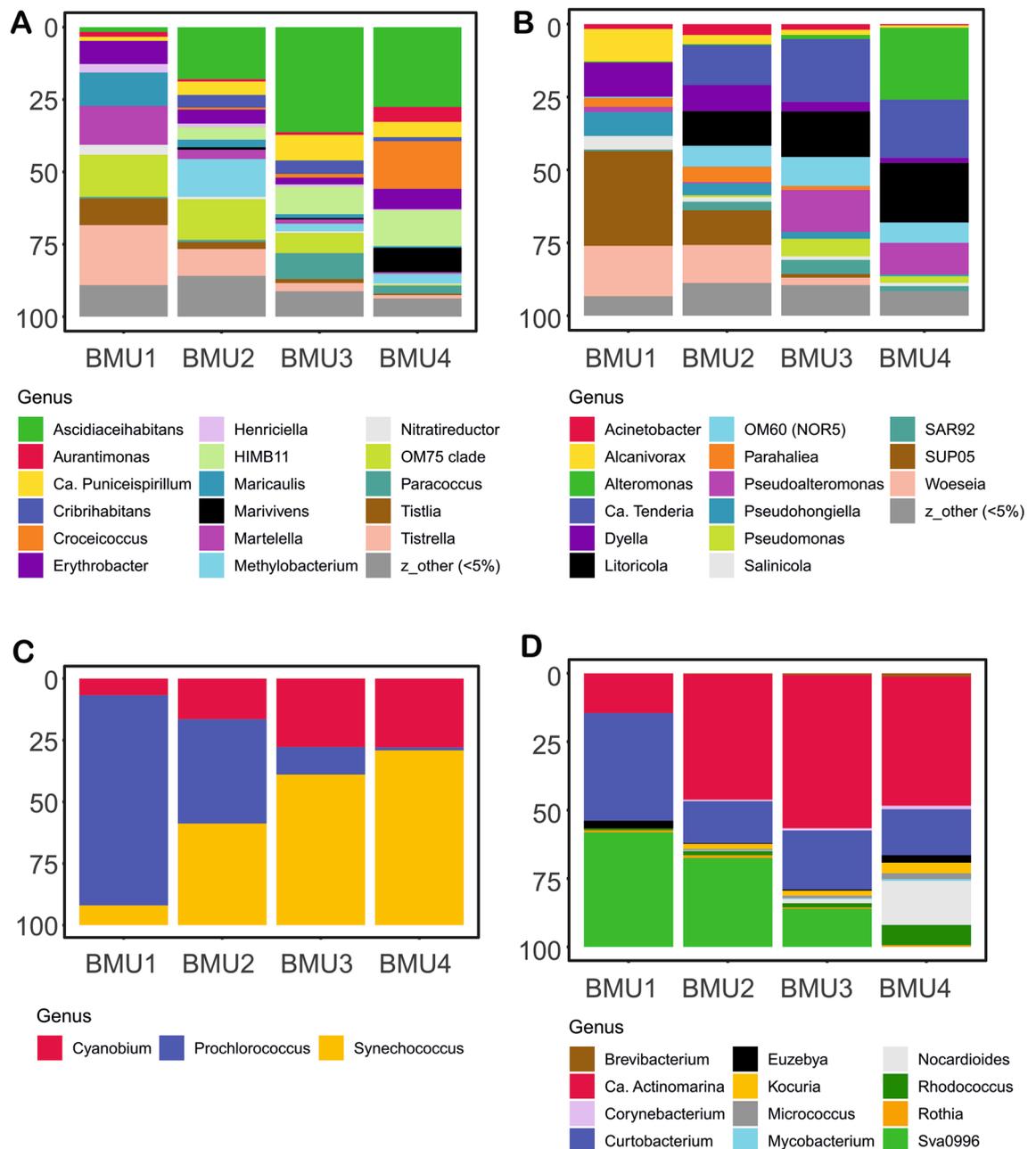


**Figure 4.** Characteristic phylum contribution (bacterial community composition-BCC-BMUs) modelled using Neural gas. Samples associated with a particular BMU are shown in the panels. The software MATLAB, version 7.10.0 (R2018), Natick, Massachusetts: The MathWorks Inc. (2018) (<https://www.mathworks.com/>) was used to generate the figure.

BCC-BMU2 represents 23% of samples, mostly from the DCM layers (JP-66, JP-75, SAP3-78). This BMU appears transitional between the BCC-BMU1 and the other BCC-BMUs. This layer had the highest proportion of Cyanobacteria, represented in relatively equal proportions by *Synechococcus* and *Prochlorococcus* (Fig. 5). The contribution of Actinobacteriota was similar to that of the BCC-BMU1 (Fig. 5), but the relative abundance of *Curtobacterium* and the uncultured Sva0996 lineage decreased, and *Ca. Actinomarina* contributed almost 50% to this phylum. Bacteroidota were dominated by uncultured NS lineages and genus *Balneola* (Fig. 5). Within Alphaproteobacteria, there was a higher contribution of *Asciadiaceihabitans* and *Methylobacterium*, but no genus was dominant. A similar pattern was noticed within Gammaproteobacteria, which showed a relatively equal contribution of all the genera, e.g. *Ca. Tenderia*, *Dyella*, *Litoricola*, SUP05 lineage or *Woeseia*, and Verrucomicrobiota, where *Coraliomargarita* and *Lentimonas* co-dominated with *Roseibacillus* and the SCGC\_AAA164-E04 lineage. In contrast, the composition of Planctomycetota was very similar to that of BCC-BMU1 (Fig. 6).

BCC-BMU3 describes the most frequent pattern, grouping the samples from the surface layer down to 65 m depth, regardless of the area. It is characterised by the dominance of Proteobacteria and Cyanobacteria followed by higher contributions of Bacteroidota and Actinobacteriota (Fig. 4). *Ca. Actinomarina* and *Curtobacterium* dominated Actinobacteriota, with a minor contribution of the uncultured Sva0996 lineage (Fig. 5). The proportion of genus *Balneola* and uncultured NS4 and NS5 lineages of Bacteroidota further increased in this BCC-BMU. *Asciadiaceihabitans* contributed most to Alphaproteobacteria, with higher proportions of *Ca. Puniceispirillum*, uncultured HIMB11 and OM75 clades, and *Paracoccus*. Within Gammaproteobacteria no genus dominated and the contribution of e.g. *Ca. Tenderia*, *Litoricola*, *Pseudoalteromonas*, and OM60/NOR5 lineage were similar (Fig. 5). The composition of Planctomycetota was very different from that of the BCC-BMUs described in the deeper samples, with a clear dominance of *Urania-1B-19* and CL005 lineages. Marine *Synechococcus* was the dominant Cyanobacterium, followed by *Cyanobium* and *Prochlorococcus* (Fig. 5).

Finally, BCC-BMU4 describes 32% of the observed samples from surface layers down to 40 m. This pattern is similar to that of the BCC-BMU3, but with a lower contribution of Cyanobacteria and a higher contribution of Proteobacteria and Bacteroidota (Fig. 4). The composition of most phyla at a finer taxonomic level was very similar to that of BCC-BMU3, but with lower proportions of the genera that dominated in BCC-BMU1 and BCC-BMU2 (Figs. 5, 6). Actinobacteriota were dominated by *Ca. Actinomarina* and *Curtobacterium*, but the uncultured Sva0996 lineage was absent, while *Nocardioideis* and *Rhodococcus* showed increased proportions. The composition of Bacteroidota was almost identical to that of BCC-BMU3, with a visibly lower proportion of NS2b lineage. Alphaproteobacteria showed higher proportions of *Aurantimonas*, *Croceicoccus*, *Erythrobacter* and *Marivivens*, with Gammaproteobacteria of *Alteromonas* and *Litoricola* displayed the same pattern (Figs. 5, 6). A further increase of *Urania-1B-19* lineage (Planctomycetota), *Lentimonas*, *Coraliomargarita* (Verrucomicrobiota) and *Synechococcus* (Cyanobacteria) was observed compared to BCC-BMU3 (Figs. 5, 6).

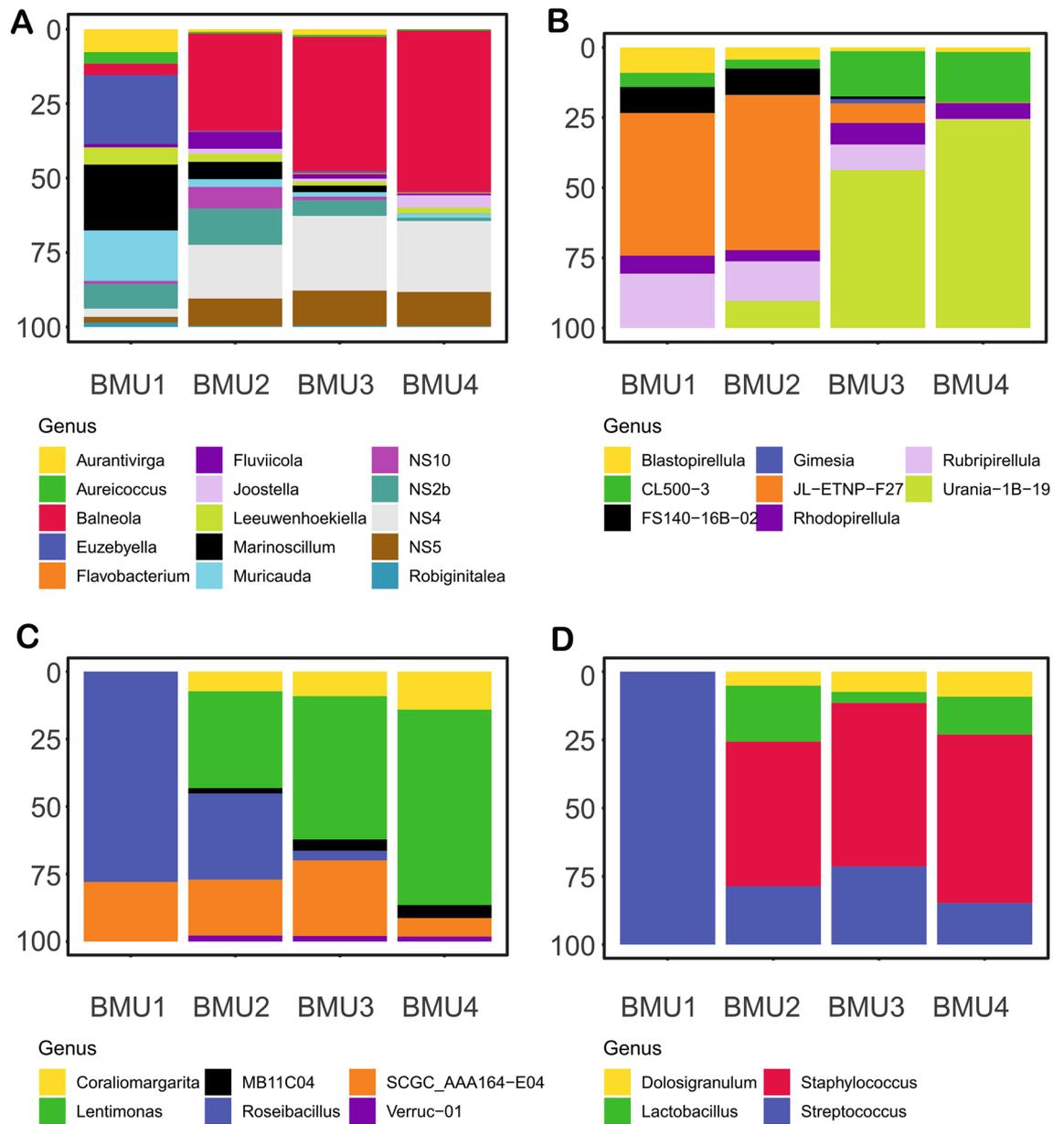


**Figure 5.** Characteristic genus contribution (A) Alphaproteobacteria (B) Gammaproteobacteria (C) Cyanobacteria (D) Actinobacteriota. The figure was generated in R statistical software<sup>35</sup> (<https://cran.r-project.org/>, v. 3.6.3), using ggplot2 v. 3.3.3 package<sup>36</sup>, and they were assembled in Inkscape (<https://inkscape.org/>, v. 0.92).

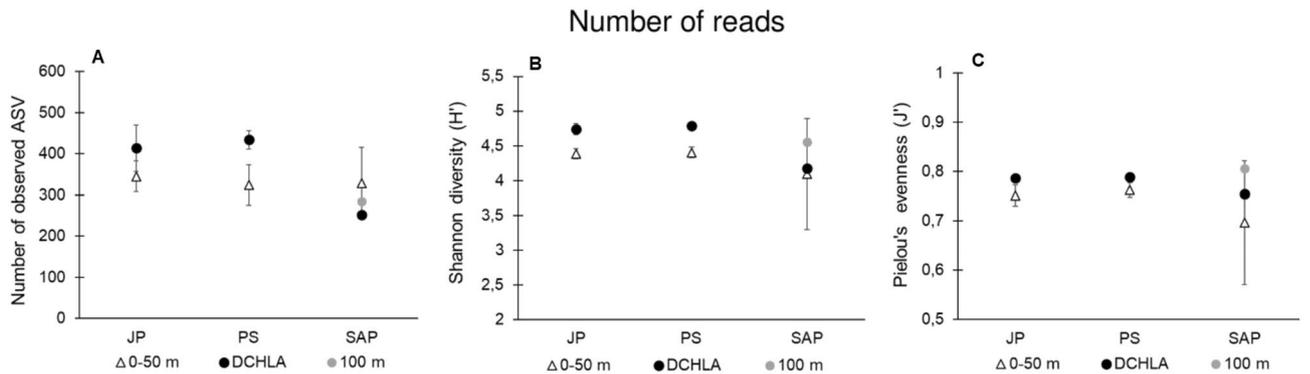
**Bacterial community diversity patterns.** The rarefaction analysis indicated that this sequencing depth was sufficient to describe the diversity of bacteria in the investigated areas of the Adriatic Sea (Fig. S2). The average values of observed ASV numbers, Shannon diversity index ( $H'$ ) and Pielou's evenness ( $J'$ ) showed the highest values at the DCM depth in the Jabuka Pit and Palagruža Sill, while at SAP the highest diversity and evenness was determined at 100 m depth and the highest number of ASVs was observed in the surface layer (0–50 m) (Fig. 7). All diversity indexes correlated with Chl  $a$  concentrations (approximated with fluorescence measurements):  $H'$  (Spearman correlation:  $r=0.45$ ,  $n=22$ ,  $P<0.05$ ),  $J'$  ( $r=0.21$ ,  $n=22$ ,  $p<0.05$ ) and number of ASVs ( $r=0.56$ ,  $n=22$ ,  $p<0.05$ ), indicating the relationship between bacterial diversity and phytoplankton biomass.

## Discussion

Environmental conditions affect both the rate of microbial processes and the composition of microbial communities in aquatic environments<sup>47–49</sup>. However, detailed knowledge about these relationships still lags behind data available for larger organisms, such as zooplankton or even microphytoplankton. One of the reasons is that the



**Figure 6.** Characteristic genus contribution (A) Bacteroidota (B) Planctomycetota (C) Verrucomicrobia (D) Firmicutes. The figure was generated in R statistical software<sup>35</sup> (<https://cran.r-project.org/>, v. 3.6.3), using ggplot2 v. 3.3.3 package<sup>36</sup>, and they were assembled in Inkscape (<https://inkscape.org/>, v. 0.92).



**Figure 7.** (A) Number of observed ASV. (B) Values of Shannon diversity index and (C) Values of Pielou's evenness index at Jabuka Pit (JP), Palagruža Sill (PS) and South Adriatic Pit (SAP). Average values  $\pm$  standard deviations (error bars) from layers are shown.

enormous bacterial diversity can be adequately described only with high-throughput sequencing methods that have been used in microbial ecology research for only about 15 years. Although huge progress has been made in describing microbial communities in the ocean<sup>50</sup>, bacterial diversity in many smaller basins is less known.

Here we present the first data of bacterial community composition in three open water basins of the oligotrophic Adriatic Sea, Jabuka Pit, Palagruža Sill and South Adriatic Pit. The data was obtained from a survey conducted in late summer 2016, and is associated with environmental parameters. During this survey the average values of SRP were  $< 0.1 \mu\text{mol L}^{-1}$  (N/P and Si/P were  $> 22$ ), while total inorganic nitrate was  $< 1 \mu\text{mol L}^{-1}$  (N/P was  $< 10$ ), suggesting that this area was generally phosphorus and nitrogen-limited. These values are typical of this region<sup>17,29</sup>, indicating that our results are applicable to wider contexts than this snap-shot study. Moreover, it is well-known that under such nutrient-poor conditions in marine ecosystems the microbial food web becomes the dominant pathway for carbon flux<sup>51</sup>. An important novelty of our study is that we used the Neural gas method to group samples with similar picoplankton characteristics, which subsequently allowed us to identify the most important environmental factors affecting their distribution. The algorithm is suitable for detecting anomalies and outlier data; in this survey, we had an unequal number of samples from the different basins. Thus, by using Neural gas method, we overcame the limitations of our research.

During thermal stratification and oligotrophic conditions, values of bacterial abundances and production were typical of the open Adriatic<sup>11,12,14,15</sup>. LNA bacteria were predominant in the P-limited epipelagic environment, which reflects their adaptation to nutrient-poor conditions<sup>52</sup>. Similarly, a general prevalence of *Synechococcus* cells within the picoautotrophic community, especially in the shallow epipelagic layer, was noted for P-depleted environments<sup>11,13,14,53,54</sup>. In contrast, HNA bacteria dominated in the upper N-limited layer and also in the deeper waters enriched with nutrients, which agrees with previous observations from the oligotrophic open sea station during the summer<sup>55</sup> and from the open Adriatic Sea<sup>11,14,15</sup>. Moreover, our results suggest that temperature increases have a positive impact on the picoplankton community in this P and N limited open sea area. On a seasonal scale, it was determined that the bacterial community was under a positive influence of temperature<sup>10</sup>. We also observed a negative relationship between picoplankton and salinity in at least 81% of the dataset, as discovered recently in the central Adriatic Sea<sup>21</sup>. Furthermore, our analysis of relationships between picoplankton and environmental characteristics pointed out that all observed parameters were more differentiated by depth than by area (Fig. 2B). The highest values of bacterial abundance and production were distributed in the shallow epipelagic layer, while the highest diversity was found in the deeper layer (Fig. 7), thus confirming the findings of previous research<sup>1,56</sup>. Korlević et al.<sup>19</sup> have linked the increase in diversity with deep convection mixing of the seawater in the Adriatic Sea, which transports nutrients to the euphotic layer and triggers phytoplankton blooms. On the other hand, higher bacterial diversity was also shown to correlate with lower bacterial abundance and production<sup>19,57</sup>, as observed here (Fig. 6).

According to sequence data, *Prochlorococcus* was present in high proportions below 50 m depth (Fig. 3F), which may be attributed to the presence of a low light-adapted ecotype<sup>58</sup>. However, flow cytometry data confirms its prevalence only in sample SAP3 at 100 m, but absolute abundances of *Synechococcus* exceeded that of *Prochlorococcus* in most samples (Table S2). An interesting observation is the presence of *Cyanobium*-related sequences. *Cyanobium* was originally established as an exclusively freshwater genus (based on the reference strain *C. gracile* PCC 6307 and its relatives), but recently several marine representatives of the genus were also isolated. It is also important to note that the primers used in this study poorly cover the SAR11 clade<sup>30</sup>, which dominates bacterial communities in the seas and oceans<sup>59</sup>, contributing up to 50% in most oligotrophic areas<sup>60</sup>. This clade is also dominant in the Mediterranean Sea<sup>61</sup> and the southern Adriatic sub-basin<sup>62</sup>. Therefore, it is highly likely that the proportion of SAR11 clade is underestimated in our study.

To link the changes in the composition of the bacterial community and environmental features we combined the results of 16S rDNA amplicon sequencing with Neural gas analysis. Recently, it has been shown that although amplicon sequencing may not accurately reflect the relative abundance of specific phylotypes, it allows for reliable ecological interpretations at the community level<sup>63</sup>. Our results clearly showed that depth was the key factor differentiating bacterial community composition in all investigated areas: Jabuka Pit, Palagruža Sill and South Adriatic Pit. Numerous studies have shown that bacterial community structure changes with depth in different oligotrophic systems, such as the Pacific and Atlantic<sup>64</sup>, as well as the Mediterranean Sea<sup>65,66</sup> suggesting that vertical stratification is a very important factor<sup>65,67</sup>.

Neural gas analysis revealed four diversity patterns of bacterial community composition (BCC-BMUs) during the summer stratification.

The first pattern (BCC-BMU1) was found at 100 m depth, the coldest investigated layer, not limited by N or P (PIC-BMU2). Although it included only a single sequenced sample, it is possible that similar bacterial communities would have been found in other deep-water ( $> 100$  m) samples of the PCI-BMU2. The majority of reads were attributed to the uncultured SUP05 lineage of Gammaproteobacteria (Fig. 5). These chemoautotrophic sulfur-oxidizing microbes are ubiquitous in diverse marine minimum oxygen zones<sup>68,69</sup>. Moreover, what makes this pattern special with regards to the shallower layers is the higher contribution of phylum Nitrospina, with the majority of reads attributed to the LS-NOB lineage. Nitrospinae uses the energy provided by nitrite oxidation to assimilate bicarbonate, which is a vital function for dark-ocean carbon fixation<sup>70</sup>. This pattern also includes increased relative abundances of Chloroflexota (Fig. 3A), represented by the uncultured SAR202 lineage. These heterotrophic bacteria are an important component of bathypelagic bacterial communities, and as potential sulfite-oxidizers may play a specific role in the sulfur cycle<sup>71</sup>. Another important uncultured heterotroph in this BCC-BMU belonged to the marine group Sva0996 of Actinobacteriota (Fig. 5). They display a close relationship with organic matter<sup>72</sup> and have an important role in dissolved organic nitrogen cycling in the ocean<sup>73</sup>. Based on the total prokaryote numbers in PIC-BMU1, these groups did not make up abundant populations, despite nutrient concentrations, which could be due to low temperature (Table 1).

The second pattern of bacterial community composition (BCC-BMU2) was found in the DCM layer. This layer was characterised by an increased abundance of cyanobacteria (Fig. 2, Table S2). In terms of bacterial community composition, a higher proportion of lineages that probably participated in the decomposition of organic matter secreted at these depths by cyanobacteria were observed, e.g. *Ca. Actinomarina* and Sva0996 of Actinobacteriota, *Balaneola* and uncultured lineages NS2b, NS4 and NS5 of Bacteroidota, and uncultured Planctomycetota from the JL-ETNP-F27 lineage (Fig. 6). In general, this BCC-BMU was intermediate between the BCC-BMU1 associated with cold, nutrient-rich waters (PIC-BMU2) and the BCC-BMU3 that is characteristic of P and N limited waters around the thermocline (PIC-BMU4 and 5).

Two patterns, BCC-BMU3 and BCC-BMU4, were associated mainly with shallow epipelagic layers (PCI-BMU 4 and 5). BCC-BMU3 was intermediate between BCC-BMU2 and BCC-BMU4, which may result from the fact that BCC-BMU3 seems to be related with the thermocline, as seawater temperatures > 15.6 °C were recorded, while BCC-BMU4 includes mainly samples above the thermocline with seawater temperatures > 18 °C (Table S1). In general, the dominant bacteria in these BMUs are known to be associated with summer phytoplankton blooms or to degrade labile organic matter, e.g. alphaproteobacterium *Asciidiaceihabitans*<sup>74</sup> and Gammaproteobacterial *Pseudoalteromonas* and *Pseudomonas* (Fig. 5). However, the genus *Alteromonas* (Gammaproteobacteria), also an important bacterium for the degradation of organic matter, was important only in BCC-BMU4.

Other phyla that substantially contributed to relative abundances in these BMUs were Bacteroidota and Actinobacteriota (Fig. 4). Phylum Bacteroidota is known to represent one of the most abundant bacterial groups in coastal environments, exhibiting a free-living or particle-attached lifestyle<sup>61,75,76</sup>. They are believed to play a major role in the degradation of organic matter in the investigated area. Members of this phylum were also abundant in the southern Adriatic sub-basin, with more operational taxonomic units at the station with the highest content of total suspended matter and particulate organic carbon<sup>62</sup>. Furthermore, this phylum showed the highest contribution in the shallow epipelagic layer, which agrees with<sup>77</sup>, who showed that relative abundances of Bacteroidota decrease with depth. We showed that this overall decrease was connected to change in their composition at finer taxonomic level, from the dominance of uncultured marine NS clusters to an increase of cultivated genera in the deepest layers, as already evidenced in the Mediterranean Sea<sup>78</sup>.

Noteworthy, the genera that are important for delineating different BCC-BMUs were not necessarily those with the highest relative abundance in the samples, but rather those whose contribution to bacterial communities differed most. The most conspicuous example is the alphaproteobacterial AEGEAN-169 lineage and gammaproteobacteria SAR86 clade, which contributed to all the samples (Fig. 3), but were not identified as important in any of the BCC-BMUs (Fig. 5). This may be due either to the high plasticity of such lineages, or to their functional diversity that cannot be recognized at the level of 16S rRNA genes.

In general, our results indicate relationships between environmental factors, picoplankton abundance, production and bacterial community composition, pointing to the importance of the species sorting processes in the open Adriatic Sea, as commonly observed in aquatic environments<sup>79</sup>. However, some exceptions were observed. For instance, environmental conditions were rather unique at 10 m depth in Jabuka Pit, where N-limited conditions could result in increased abundances of prokaryotes, HNA and AAP bacteria (Fig. 2, Table 1). Nevertheless, BCC composition in this sample did not differ substantially from the other samples in Jabuka Pit or BCC-BMU3 (Fig. 3), thus indicating the importance of other ecological variables in shaping bacterial communities in the area.

## Conclusions

The picoplankton community displayed heterogenic response to different environmental factors in the open sea areas of the central and southern Adriatic Sea, throughout the stratified water column.

The use of Neural gas analysis allowed us to determine associations between selected environmental parameters and observed picoplankton variables, bacterial community structure and profiles.

Our results provide a novel insight into factors affecting picoplankton in this area. They revealed that picoplankton and bacterial community composition were more associated with the depth rather than investigated area. Furthermore, an increase in temperature correlated positively with the picoplankton community (heterotrophic bacteria, HNA%, LNA%, *Synechococcus*, *Prochlorococcus*, picoeukaryotes, AAP, bacterial production) in this N and P limited environment. BCC showed the dominance of Proteobacteria and Cyanobacteria, followed by Bacteroidota and Actinobacteriota phyla, whereas changes at finer taxonomic levels were related to distinct environmental conditions at different depths. Future research should focus on high-frequency sampling of bacterial community composition, together with environmental parameters, to obtain a clearer representation of ecological patterns.

## Data availability

The data supporting the findings of this study are available on request from the corresponding author Danijela Šantić, segvic@izor.hr, Institute of Oceanography and Fisheries.

Received: 20 November 2020; Accepted: 17 May 2021

Published online: 27 May 2021

## References

- Pommier, T. *et al.* Spatial patterns of bacterial richness and evenness in the NW Mediterranean Sea explored by pyrosequencing of the 16S rRNA. *Aquat. Microb. Ecol.* **61**, 221–233. <https://doi.org/10.3354/ame01484> (2010).
- Pinhassi, J. *et al.* Seasonal changes in bacterioplankton nutrient limitation and their effects on bacterial community composition in the NW Mediterranean Sea. *Aquat. Microb. Ecol.* **44**(3), 241–252. <https://doi.org/10.3354/ame044241> (2006).

3. Sánchez, O., Koblížek, M., Gasol, J. M. & Ferrera, I. Effects of grazing, phosphorus and light on the growth rates of major bacterioplankton taxa in the coastal NW Mediterranean. *Environ. Microbiol. Rep.* **9**, 300–309. <https://doi.org/10.1111/1758-2229.12535> (2017).
4. Cram, J. A. *et al.* Seasonal and interannual variability of the marine bacterioplankton community throughout the water column over ten years. *ISME J.* **9**, 563–580. <https://doi.org/10.1038/ismej.2014.153> (2014).
5. Ferrera, I. *et al.* Evaluation of alternative high-throughput sequencing methodologies for the monitoring of marine picoplanktonic biodiversity based on rRNA gene amplicons. *Front. Mar. Sci.* **3**, 147. <https://doi.org/10.3389/fmars.2016.00147> (2016).
6. Walsh, E. A. *et al.* Bacterial diversity and community composition from seasurface to seafloor. *ISME J.* **10**(4), 979–989. <https://doi.org/10.1038/ismej.2015.175> (2016).
7. Zhou, J. *et al.* Distribution patterns of microbial community structure along a 7000-mile latitudinal transect from the Mediterranean Sea across the Atlantic Ocean to the Brazilian Coastal Sea. *Microb. Ecol.* **76**(3), 592–609. <https://doi.org/10.1007/s00248-018-1150-z> (2018).
8. Haber, M. *et al.* Microbial communities in an ultra-oligotrophic sea are more affected by season than by distance from shore. *Preprint at bioRxiv* <https://doi.org/10.1101/2020.04.17.044305> (2020).
9. Milici, M. *et al.* Co-occurrence analysis of microbial taxa in the Atlantic ocean reveals high connectivity in the free-living bacterioplankton. *Front. Microbiol.* **7**, 649. <https://doi.org/10.3389/fmicb.2016.00649> (2016).
10. Šolić, M. *et al.* Spatio-temporal reproducibility of the microbial food web structure associated with the change in temperature: Long-term observations in the Adriatic Sea. *Prog. Oceanogr.* **161**, 87–101. <https://doi.org/10.1016/j.poccean.2018.02.003> (2018).
11. Šantić, D., Krstulović, N., Šolić, M., Ordulj, M. & Kušpilić, G. Dynamics of prokaryotic picoplankton community in the central and southern Adriatic Sea (Croatia). *Helgol. Mar. Res.* **67**, 471. <https://doi.org/10.1007/s10152-012-0336-x> (2013).
12. Batistić, M. *et al.* Biological evidence of a winter convection event in the South Adriatic: A phytoplankton maximum in the aphotic zone. *Cont. Shelf Res.* **44**, 57–71. <https://doi.org/10.1016/j.csr.2011.01.004> (2012).
13. Šantić, D. *et al.* Dynamics of picoplankton community from coastal waters to the open sea in the Central Adriatic. *Mediterr. Mar. Sci.* **15**, 179–188. <https://doi.org/10.12681/mms.701> (2014).
14. Šantić, D. *et al.* Picoplankton distribution and activity in the deep waters of the southern Adriatic Sea. *Water* **11**, 1655. <https://doi.org/10.3390/w11081655> (2019).
15. Šilović, T. *et al.* Picoplankton distribution influenced by thermohaline circulation in the southern Adriatic. *Cont. Shelf Res.* **155**, 21–33. <https://doi.org/10.1016/j.csr.2018.01.007> (2018).
16. Vrdoljak Tomaš, A. *et al.* Dynamics of Aerobic Anoxygenic Phototrophs along the trophic gradient in the central Adriatic Sea. *Deep Sea Res. Pt II* **164**, 112–121. <https://doi.org/10.1016/j.dsr2.2019.06.001> (2019).
17. Vilbić, I., Matijević, S., Šepić, J. & Kušpilić, G. Changes in the Adriatic oceanographic properties induced by the Eastern Mediterranean Transient. *Biogeosciences* **9**, 2085–2097. <https://doi.org/10.5194/bg-9-2085-2012> (2012).
18. Yari, S., Kovačević, V., Cardin, V., Gačić, M. & Bryden, H. L. Direct estimate of water, heat, and salt transport through the Strait of Otranto. *J. Geophys. Res. Oceans* <https://doi.org/10.1029/2012JC007936> (2012).
19. Korlević, M., Ristova, P. P., Garić, R., Amann, R. & Orlić, S. Bacterial diversity in the South Adriatic Sea during a strong, deep winter convection year. *Appl. Environ. Microbiol.* **81**, 1715–1726. <https://doi.org/10.1128/AEM.03410-14> (2015).
20. Šolić, M. *et al.* Impact of water column stability dynamics on the succession of plankton food web types in the offshore area of the Adriatic Sea. *J. Sea Res.* **158**, 101860. <https://doi.org/10.1016/j.seares.2020.101860> (2020).
21. Beg Paklar, G. *et al.* Record-breaking salinities in the middle Adriatic during summer 2017 and concurrent changes in the microbial food web. *Prog. Oceanogr.* **185**, 102345. <https://doi.org/10.1016/j.poccean.2020.102345> (2020).
22. Bandelj, V. *et al.* Analysis of multitrophic plankton assemblages in the Lagoon of Venice. *Mar. Ecol. Prog. Ser.* **368**, 23–40. <https://doi.org/10.3354/meps07565> (2008).
23. Mazzocchi, M. G. *et al.* Regional and seasonal characteristics of epipelagic mesozooplankton in the Mediterranean Sea based on an artificial neural network analysis. *J. Mar. Syst.* **135**, 64–80. <https://doi.org/10.1016/j.jmarsys.2013.04.009> (2014).
24. Ninčević Gladan, Ž. *et al.* The relationship between toxic phytoplankton species occurrence and environmental and meteorological factors along the Eastern Adriatic coast. *Harmful Algae* **92**, 101745. <https://doi.org/10.1016/j.hal.2020.101745> (2020).
25. Martinetz, T. M., Berkovich, S. G. & Schulten, K. J. “Neural-gas” network for vector quantization and its application to time-series prediction. *IEEE Trans. Neural Netw.* **4**(4), 558–569 (1993).
26. Martinetz T. & Schulten K. A “Neural Gas” Network Learns Topologies. *Artificial Neural Networks*, 397–402 (Elsevier, 1991).
27. Cushman-Roisin, B., Gačić, M., Poulain, P. M. & Artegiani, A. (Eds.) *Physical Oceanography of the Adriatic Sea: Past, Present and Future* (Springer, 2013).
28. Grasshoff, K. *Methods of Seawater Analysis* 307 (Verlag Chemie, 1976).
29. Justić, D., Rabalais, N. N., Turner, R. E. & Dortch, Q. Changes in nutrient structure of river-dominated coastal waters: Stoichiometric nutrient balance and its consequences. *Estuar. Coast. Shelf Sci.* **40**, 339–356. [https://doi.org/10.1016/S0272-7714\(05\)80014-9](https://doi.org/10.1016/S0272-7714(05)80014-9) (1995).
30. Klindworth, A. *et al.* Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res.* **41**(1), e1. <https://doi.org/10.1093/nar/gks808> (2013).
31. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. J.* **17**(1), 10–12. <https://doi.org/10.14806/ej.17.1.200> (2011).
32. Callahan, B. J., Sankaran, K., Fukuyama, J. A., McMurdie, P. J. & Holmes, S. P. Bioconductor workflow for microbiome data analysis: From raw reads to community analyses. *F1000Research* **5**, 1492. <https://doi.org/10.12688/f1000research.8986.2> (2016).
33. Wang, Q., Garrity, G. M., Tiedje, J. M. & Cole, J. R. Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.* **73**, 5261–5267. <https://doi.org/10.1128/AEM.00062-07> (2007).
34. Pruesse, E. *et al.* SILVA: A comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res.* **35**(21), 7188–7196. <https://doi.org/10.1093/nar/gkm864> (2007).
35. McMurdie, P. J. & Holmes, S. phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS ONE* **8**, e61217. <https://doi.org/10.1371/journal.pone.0061217> (2013).
36. R Core Team. R: A Language and Environment for Statistical Computing. <https://www.R-project.org> (2015).
37. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis* (Springer, 2009).
38. Gasol, J. M. & Morán, X. A. G. Flow Cytometric determination of microbial abundances and its use to obtain indices of community structure and relative activity (2015). In *Hydrocarbon and Lipid Microbiology Protocols*. Springer Protocols Handbooks (McGenity, T. *et al.*) 159–187 (Springer, 2017).
39. Mašin, M. *et al.* Seasonal changes and diversity of aerobic anoxygenic phototrophs in the Baltic Sea. *Aquat. Microb. Ecol.* **45**, 247–254. <https://doi.org/10.3354/ame045247> (2006).
40. Fuhrman, J. A. & Azam, F. Thymidine incorporation as a measure of heterotrophic bacterioplankton production in marine surface waters: Evaluation and field results. *Mar. Biol.* **66**, 109–120. <https://doi.org/10.1007/BF00397184> (1982).
41. Clarke, K. & Gorley, R. *Primer v6. User Manual/Tutorial* (Plymouth Routine in Multivariate Ecological Research, 2006).
42. Kohonen, T. Self-organized information of topologically correct features maps. *Biol. Cybern.* **43**, 59–69. <https://doi.org/10.1007/BF00337288> (1982).
43. de Bolt, E., Cottrell, M. & Verleysen, M. Statistical tools to assess the reliability of self-organising maps. *Neural Netw.* **15**, 967–978. [https://doi.org/10.1016/S0893-6080\(02\)00071-0](https://doi.org/10.1016/S0893-6080(02)00071-0) (2002).

44. Matic, F., Kalinić, H. & Vilibić, I. Interpreting Self-Organizing Map errors in the classification of ocean patterns. *Comput. Geosci.* **119**, 9–17. <https://doi.org/10.1016/j.cageo.2018.06.006> (2018).
45. Zhuang, Z. *et al.* Variability of Kuroshio surface axis northeast of Taiwan island derived from satellite altimeter data. *Remote Sens.* **12**, 1059. <https://doi.org/10.3390/rs12071059> (2020).
46. Matic, F. *et al.* Oscillating Adriatic temperature and salinity regimes mapped using the Self-Organizing Maps method. *Cont. Shelf Res.* **132**, 11–18. <https://doi.org/10.1016/j.csr.2016.11.006> (2017).
47. Gołębiewski, M., Calkiewicz, J., Creer, S. & Pivosz, K. Tideless estuaries in brackish seas as a possible freshwater-marine transition zones for bacteria—The case study of the Vistula river estuary. *Environ. Microbiol. Rep.* **9**, 129–143. <https://doi.org/10.1111/1758-2229.12509> (2017).
48. Ameryk, A., Podgorska, B. & Witek, Z. The dependence between bacterial production and environmental conditions in the Gulf of Gdansk. *Oceanologia* **47**, 27–45 (2005).
49. Langenheder, S., Lindström, E. S. & Tranvik, L. J. Weak coupling between community composition and functioning of aquatic bacteria. *Limnol. Oceanogr.* **50**, 957–967. <https://doi.org/10.4319/lo.2005.50.3.0957> (2005).
50. Sunagawa, S. *et al.* Structure and function of the global ocean microbiome. *Science* **348**, 1261359. <https://doi.org/10.1126/science.1261359> (2015).
51. Azam, F. & Malfatti, F. Microbial structuring of marine ecosystems. *Nat. Rev. Microbiol.* **5**, 782–791. <https://doi.org/10.1038/nrmicr.01747> (2007).
52. Mary, I. *et al.* SAR11 dominance among metabolically active low nucleic acid bacterioplankton in surface waters along an Atlantic meridional transect. *Aquat. Microb. Ecol.* **45**, 107–113. <https://doi.org/10.3354/ame045107> (2006).
53. Martiny, A. C., Kathuria, S. & Berube, P. M. Widespread metabolic potential for nitrite and nitrate assimilation among *Prochlorococcus* ecotypes. *PNAS* **106**, 10787–10792. <https://doi.org/10.1073/pnas.0902532106> (2009).
54. Cerino, F. *et al.* Spatial and temporal variability of pico-, nano- and microphytoplankton in the offshore waters of the southern Adriatic Sea (Mediterranean Sea). *Cont. Shelf Res.* **44**, 94–105. <https://doi.org/10.1016/j.csr.2011.06.006> (2012).
55. Sherr, E. B., Sherr, B. F. & Longnecker, K. Distribution of bacterial abundance and cell-specific nucleic acid content in the Northeast Pacific Ocean. *Deep Sea Res. Pt. I* (53), 713–725. <https://doi.org/10.1016/j.dsr.2006.02.001> (2006).
56. Ghiglione, J. F., Conan, P. & Pujó-Pay, M. Diversity of total and active free-living vs. particle-attached bacteria in the euphotic zone of the NW Mediterranean Sea. *FEMS Microbiol. Lett.* **299**, 9–21. <https://doi.org/10.1111/j.1574-6968.2009.01694.x> (2009).
57. Mittelbach, G. G. *et al.* What is the observed relationship between species richness and productivity?. *Ecology* **82**(9), 2381–2396. [https://doi.org/10.1890/0012-9658\(2001\)082\[2381:WITORB\]2.0.CO;2](https://doi.org/10.1890/0012-9658(2001)082[2381:WITORB]2.0.CO;2) (2001).
58. Moore, L. R., Post, A. F., Rocap, G. & Chisholm, S. W. Utilization of different nitrogen sources by the marine cyanobacteria *Prochlorococcus* and *Synechococcus*. *Limnol. Oceanogr.* **47**, 989–996. <https://doi.org/10.4319/lo.2002.47.4.0989> (2002).
59. Giovannoni, S. J. SAR11 bacteria: The most abundant plankton in the oceans. *Ann. Rev. Mar. Sci.* **9**, 231–255. <https://doi.org/10.1146/annurev-marine-010814-015934> (2017).
60. Carlson, C. A. *et al.* Seasonal dynamics of SAR11 populations in the euphotic and mesopelagic zones of the northwestern Sargasso Sea. *ISME J.* **3**, 283–295. <https://doi.org/10.1038/ismej.2008.117> (2008).
61. Alonso-Sáez, L. & Gasol, J. M. Seasonal variations in the contributions of different bacterial groups to the uptake of low-molecular-weight compounds in northwestern Mediterranean coastal waters. *Appl. Environ. Microbiol.* **73**, 3528–3535. <https://doi.org/10.1128/AEM.02627-06> (2007).
62. Quero, G. M. & Luna, G. M. Diversity of rare and abundant bacteria in surface waters of the Southern Adriatic Sea. *Mar. Genom.* **17**, 9–15. <https://doi.org/10.1016/j.margen.2014.04.002> (2014).
63. Pivosz, K. *et al.* Bacterial and eukaryotic small-subunit amplicon data do not provide a quantitative picture of microbial communities, but they are reliable in the context of ecological interpretations. *mSphere* **5**, e00052–e120. <https://doi.org/10.1128/mSphere.00052-20> (2020).
64. Lee, S. & Fuhrman, J. A. Spatial and temporal variation of natural bacterioplankton assemblages studied by total genomic DNA cross-hybridization. *Limnol. Oceanogr.* **36**(7), 1277–1287. <https://doi.org/10.4319/lo.1991.36.7.1277> (1991).
65. Acinas, S. G., Rodríguez-Valera, F. & Pedrós-Alió, C. Spatial and temporal variation in marine bacterioplankton diversity as shown by RFLP fingerprinting of PCR amplified 16S rDNA. *FEMS Microbiol. Ecol.* **24**(1), 27–40. <https://doi.org/10.1111/j.1574-6941.1997.tb00420.x> (1997).
66. Moeseneder, M. M., Winter, C. & Herndl, G. J. Horizontal and vertical complexity of attached and free-living bacteria of the eastern Mediterranean Sea, determined by 16S rDNA and 16S rRNA fingerprints. *Limnol. Oceanogr.* **46**, 95–107. <https://doi.org/10.4319/lo.2001.46.1.0095> (2001).
67. Ghiglione, J. F. *et al.* Role of environmental factors for the vertical distribution (0–1000 m) of marine bacterial communities in the NW Mediterranean Sea. *Biogeosci. Discuss.* **5**(3), 2131–2164 (2008).
68. Spietz, R. L. *et al.* Heterotrophic carbon metabolism and energy acquisition in *Candidatus* Thioglobus singularis strain PS1, a member of the SUP05 clade of marine *Gamma*proteobacteria. *Environ. Microbiol.* **21**(7), 2391–2401. <https://doi.org/10.1111/1462-2920.146232391-2401> (2019).
69. Walsh, D. A. *et al.* Metagenome of a versatile chemolithoautotroph from expanding oceanic dead zones. *Science* **326**(5952), 578–582. <https://doi.org/10.1126/science.1175309> (2009).
70. Ngugi, D. *et al.* Diversification and niche adaptations of Nitrospina-like bacteria in the polyextreme interfaces of Red Sea brines. *ISME J.* **10**, 1383–1399. <https://doi.org/10.1038/ismej.2015.214> (2016).
71. Mehrshad, M., Rodríguez-Valera, F., Amoozegar, M. A., López-García, P. & Ghai, R. The enigmatic SAR202 cluster up close: Shedding light on a globally distributed dark ocean lineage involved in sulfur cycling. *ISME J.* **12**, 655–668. <https://doi.org/10.1038/s41396-017-0009-5> (2018).
72. Wang, Y. *et al.* Bacterial community structure in the Bohai Strait provides insights into organic matter niche partitioning. *Continent. Shelf Res.* **169**, 46–54. <https://doi.org/10.1016/j.csr.2018.08.009> (2018).
73. Orsi, W. D. *et al.* Diverse, uncultivated bacteria and archaea underlying the cycling of dissolved protein in the ocean. *ISME J.* **10**, 2158–2173. <https://doi.org/10.1038/ismej.2016.20> (2016).
74. Chafee, M. *et al.* Recurrent patterns of microdiversity in a temperate coastal marine environment. *ISME J.* **12**, 237–252. <https://doi.org/10.1038/ismej.2017.165> (2018).
75. Kirchman, D. L. The ecology of Cytophaga-Flavobacteria in aquatic environments. *FEMS Microbiol. Ecol.* **39**, 91–100. <https://doi.org/10.1111/j.1574-6941.2002.tb00910.x> (2002).
76. Alonso, C., Warnecke, F., Amann, R. & Pernthaler, J. High local and global diversity of Flavobacteria in marine plankton. *Environ. Microbiol.* **9**(5), 1253–1266. <https://doi.org/10.1111/j.1462-2920.2007.01244> (2007).
77. Díez-Vives, C., Gasol, J. M. & Acinas, S. G. Spatial and temporal variability among marine Bacteroidetes populations in the NW Mediterranean Sea. *Syst. Appl. Microbiol.* **37**, 68–78. <https://doi.org/10.1016/j.syapm.2013.08.006> (2014).
78. Díez-Vives, C. *et al.* Delineation of ecologically distinct units of marine Bacteroidetes in the Northwestern Mediterranean Sea. *Mol. Ecol.* **28**(11), 2846–2859. <https://doi.org/10.1111/mec.15068> (2019).
79. Langenheder, S., Bulling, M. T., Solan, M. T. & Prosser, J. I. Bacterial biodiversity-ecosystem functioning relations are modified by environmental complexity. *PLoS ONE* **5**(5), e10834. <https://doi.org/10.1371/journal.pone.0010834> (2010).

## Acknowledgements

We thank Ivan Vučić, Hrvoje Mihanović, PhD and the crew of R/V BIOS DVA.

## Author contributions

D.Š. was responsible for conceptualization, methodology, sampling and data collection, formal analysis, validation, visualisation, writing and funding acquisition; K.P., F.M. was responsible for methodology, sampling and data collection, formal analysis, validation, visualisation writing, funding acquisition; A.V.T. was responsible for methodology, formal analysis, visualisation, writing; J.A. was responsible for visualisation, writing; J.L.D. was responsible for sampling and data collection, formal analysis, proofreading; M.Š. was responsible for conceptualization; M.K. was responsible for conceptualization, funding acquisition, writing; G.K., S.Š. was responsible for formal analysis.

## Funding

This work was supported by the Adriatic monitoring plan in 2016 (central and southern Adriatic area), “Ecology of the Aerobic Anoxygenic Phototrophs in the Adriatic Sea- ADRISAAP” (grant number UIP-2019-04-8401), “Aerobic anoxygenic phototrophs: their role in carbon utilization under light in freshwater ecosystems—AAPs rule!” (grant number 18-14095Y) from the Czech Science Foundation awarded to K.P., and “Middle Adriatic Upwelling and Downwelling- MAUD” (grant number IP-2018-01-9849).

## 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/s41598-021-90863-7>.

**Correspondence** and requests for materials should be addressed to D.Š.

**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) 2021