Check for updates

ARTICLE   OPEN

Influence of short and long term processes on SAR11 communities in open ocean and coastal systems

Luis M. Bolaños [1✉], Karen Tait[2], Paul J. Somerfield [2], Rachel J. Parsons [3], Stephen J. Giovannoni [4], Timothy Smyth [2] and Ben Temperton [1✉]

SAR11 bacteria dominate the surface ocean and are major players in converting fixed carbon back to atmospheric carbon dioxide. The SAR11 clade is comprised of niche-specialized ecotypes that display distinctive spatiotemporal transitions. We analyzed SAR11 ecotype seasonality in two long-term 16S rRNA amplicon time series representing different North Atlantic regimes: the Sargasso Sea (subtropical ocean-gyre; BATS) and the temperate coastal Western English Channel (WEC). Using phylogenetically resolved amplicon sequence variants (ASVs), we evaluated seasonal environmental constraints on SAR11 ecotype periodicity. Despite large differences in temperature and nutrient availability between the two sites, at both SAR11 succession was defined by summer and winter clusters of ASVs. The summer cluster was dominated by ecotype Ia.3 in both sites. Winter clusters were dominated by ecotypes Ib and IIa.A at BATS and Ia.1 and IIa.B at WEC. A 2-year weekly analysis within the WEC time series showed that the response of SAR11 communities to short-term environmental fluctuations was variable. In 2016, community shifts were abrupt and synchronized to environmental shifts. However, in 2015, changes were gradual and decoupled from environmental fluctuations, likely due to increased mixing from strong winds. We demonstrate that interannual weather variability disturb the pace of SAR11 seasonal progression.

INTRODUCTION

SAR11 are heterotrophic, free-living planktonic alpha-proteobacteria with streamlined genomes and small cell sizes (~0.04 μm³) [1–4]. SAR11 are globally distributed throughout the surface oceans and account for around 25% of the plankton cells in this layer of the water column [5–7]. At a broad phylogenetic level, the SAR11 order (Pelagibacterales [8]) is divided into subclades that are congruent with spatial and temporal distributions defined by environmental factors such as depth, latitude, and season [9–12]. Therefore, these subclades have been defined as ecotypes [11]. Genomic studies have determined clear phylogenetic subclade boundaries within the SAR11 clade and around half of these lack a cultivable representative [13]. While fundamentally similar in phenotype and core genome conservation, SAR11 ecotypes display differences in many characteristics that have adaptive and geochemical significance. Variations among SAR11 ecotypes in phosphorous compound metabolism [14], carbon utilization [15, 16], production of gaseous compounds [17], and utilization rates of volatile organic compounds [18] have important implications for ocean [19, 20] and atmosphere [21] geochemistry.

16S rRNA surveys across oceanic environmental gradients have shown that SAR11 communities differ by favouring the most adapted ecotypes [22, 23]. Time series in temperate regions of the ocean revealed that SAR11 communities are sensitive to seasonal changes, which determine the successional progression of dominant ecotypes [11, 24–28]. Specifically, long-term time series have shown that SAR11 dynamics display a highly consistent ecotype periodicity through seasons [27, 29]. Short time-scale variation (days to weeks) of SAR11 using high frequency sampling suggests that SAR11 communities gradually transition through the seasons [30], yet the influence of short-term environmental fluctuations on SAR11 variability remains unexplored.

Despite advances in understanding of SAR11 seasonality, biogeography, distribution in the water column and the environmental drivers shaping its composition, the analysis and interpretation of this globally dominant bacteria remain limited to specific geographic locations of isolated time-series observatories or snapshots of transects with restricted or no temporal resolution. Furthermore, throughout the evolution of sequencing technologies, different molecular methodologies have been preferentially adopted or continuously updated by different projects. This increases the difficulty of compiling a highly resolved model of SAR11 global ecology and evolution. In addition, challenges to sustain and fund ocean time series observations and shipboard collections lead to sampling discontinuity, disconnected databases, lack of standardized sampling, among other complications [31]. Data integration has been proposed as an important area to develop to overcome logistic challenges, improve predictive capacity and enhance the monitoring accuracy of climate change effects in the oceans [31].

In the North Atlantic western subtropical gyre, the Bermuda Atlantic Time-series Study (BATS) has collected oceanographic and biological data since 1988 [32]. 16S rRNA surveys have been consistently generated from monthly sampling [12, 30]. The

[1]School of Biosciences, University of Exeter, Exeter, UK. [2]Plymouth Marine Laboratory, Plymouth, UK. [3]Bermuda Institute of Ocean Sciences, St. George's, Bermuda. [4]Department of Microbiology, Oregon State University, Corvallis, OR, USA. ✉email: l.bolanos-avellaneda@exeter.ac.uk; b.temperton@exeter.ac.uk

Western English Channel (WEC), Station L4, located off the southwestern coast of the United Kingdom is one of the longest oceanographic time series in the world [33]. Since the year 2000, samples from the WEC have been collected weekly for multiple microbial molecular studies [34–37]. These time series provide a robust framework to quantify the influence of environmental factors on SAR11 composition across long and short time scales, and identify similarities and differences between temperate coastal waters and the subtropical open ocean.

Improvement of taxonomic resolution of SAR11 and other microorganisms on amplicon surveys with phylogenetic placement methods have been shown to increase the accuracy of microbial dynamics in the water column and geographical patterns in large-scale ocean provinces [22, 29, 38–40]. However, data integration to compare these surveys has remained unexplored. The PhyloAssigner pipeline [29], based on pplacer [41], was developed to encourage consistent taxonomies, permit comparisons of non-overlapping partial rRNA sequences, and optimize the taxonomic assignment of new ASVs. The use of a phylogenetic tree based on a curated full-length 16S rRNA multiple sequence alignment ensures a consistent taxonomy to analyze datasets. This unified method avoids rearrangements among internal nodes that would occur if multiple independent trees were inferred from different sets of shorter sequences within the gene (for example, different amplicon hypervariable regions of the 16S rRNA gene). The PhyloAssigner pipeline addresses pplacer uncertainties by implementing a Last Common Ancestor analysis [29], which moves uncertain placements to lower and more reliable levels in the phylogenetic tree. None of these features addresses specific primer bias or DNA extraction efficiencies of different methodologies, but they address substantial sources of phylogenetic error that have impeded comparisons of different amplicon data sets. By comparing disparate multi-annual amplicon datasets of highly resolved time series under the same phylogenetic placement framework we aim to provide more accurate parameters of the drivers of SAR11 communities' spatiotemporal distribution, evaluate the impacts of the variability of these drivers on SAR11 and analyze whether significant changes have occurred in the last decades.

We analyzed the SAR11 fraction of the 16S rRNA surveys retrieved from datasets spanning more than 27 years since 1991 in BATS and seven continuous years (2012–2018) in the WEC. SAR11 amplicon sequence variants (ASV) were phylogenetically classified using the PhyloAssigner pipeline [29] and a curated phylogenetic database of full-length SAR11 16S rRNA gene sequences [22]. In both time series, we evaluated the persistence of SAR11 ecotype succession and the influences of the environment on the annual progression of SAR11 composition. Finally, we leveraged the near-complete, weekly sampling within a two-year period in the WEC to evaluate the short-term environmental influences within the annual progression. The results provide a panoramic overview of SAR11 responses to interannual fluctuations in the surface ocean.

MATERIALS AND METHODS
Sample collection, DNA extraction, 16S rRNA library preparation, and sequencing
Bermuda Atlantic time-series. Depth profiles were collected monthly at BATS (31°40′N, 64°10′W). Surface samples (1–5 m) for the periods 1991–2002 [29] and 2016–2018 were analyzed in this study. Microbial biomass from 4 L of water was collected on 47 mm, 0.22 μm pore Supor filters (Sigma-Aldrich, St. Louis, MO, USA) from 1991 to 2002 and on Sterivex filters of the same composition from 2016 to 2018. DNA of microbial biomass collected in the two time periods was extracted using the same phenol:chloroform protocol [42–44]. 16S rRNA V1-V2 amplicons from 1991 to 2002 were generated with the primers 27FB (5′-AGRGTTYGATYMTGGCTCAG-3′) [44] and 338RPL (5′-GCWGCCWCCCGTAGGWGT-3′) [29, 45]. Amplicon libraries were sequenced using Roche 454 (Branford, CT, USA) FLX technology. 454 FLX sequences were retrieved from iMicrobe (CAM_P_0000950). 16S rRNA V1-V2 amplicons from samples collected between 2016 and 2018 were generated with the

primers 27F (5′-AGAGTTTGATCNTGGCTCAG-3) [46] and 338RPL (5′-GCWGCCWCCCGTAGGWGT-3′) [29, 45]. Amplicon libraries (2016–2018) were prepared with the Nextera XT Kit (Illumina Inc.). Libraries were sequenced using the MiSeq platform (v.2; 2 × 250) by the Centre for Genome Research and Biocomputing (Oregon State University).

Western English Channel time-series. Surface water samples (1–5 m) were collected weekly at the station L4 (50°15′N, 4°13′W) as part of the Western Channel Observatory time series. Five litres of water were filtered through 0.22 μm pore Sterivex filters. DNA was extracted using a Qiagen AllPrep DNA/RNA Micro Kit. 16S rRNA V4-V5 amplicons were generated with the primers 515fB (5′-GTGYCAGCMGCCGCGGTAA-3′) and 806rB (5′-GGAC TACNVGGGTWTCTAAT-3′). Amplicon libraries were generated with the Nextera XT Kit (Illumina Inc.). Libraries were sequenced using the MiSeq platform (v.3; 2 × 300) by NU-OMICS (Northumbria University, U.K.). A list of samples is provided as Table S1.

Sequence processing and taxonomic classification
Surface 454 FLX amplicon datasets from BATS spanning from 1991 to 2002 [29] and the MiSeq Illumina datasets (2015–2018 BATS and 2012–2018 WEC) were quality filtered, dereplicated, and merged with Dada2 v1.18 [47] and phyloseq v1.34 [48] R packages (see Supplementary methods). SAR11 sequences were extracted, aligned, and phylogenetically placed on a 16S rRNA full-length custom tree [22] with Phyloassigner v6.166 [29]. For V4-V5 SAR11 amplicons, oligotyping [49] was used to further discern between Ia.1 and Ia.3 ecotypes [50]. The result of the highly resolved phylogenetic placement was used to generate an updated taxonomic file.

SAR11 composition and seasonality analysis
Relative contribution for each of the SAR11 ecotypes was obtained by collapsing the amplicon counts of the ASVs based on their phylogenetic classification. Subsequent analysis (ordinations and hierarchical clustering) were based on Bray-Curtis dissimilarity obtained from rarefied counts of ASVs. These ecological measurements were calculated in R v4.0 [51] using phyloseq v1.34 [48] and vegan v2.5 [52]. ggplot2 v3.3 package [53] was used to plot the figures. Averaged Bray–Curtis dissimilarities from same time lengths (days) were used to generate time-decay curves. A periodic spline (df = 4 and regular interval = 366 days) was fitted on the averaged Bray–Curtis dissimilarities using the splines2 v4.6 package [54]. Seasonality, defined as the periodic changes of each ASV, was determined with a Fisher G-test [55] implemented in GeneCycle R package [56] based on individual relative contributions. To classify the significant periodic ASVs we created harmonic linear models defined by a seasonal cycle mirroring day-length peaking in midwinter determined by the function $Xc = \cos 2\pi t$, where $t = \frac{d}{365}$ being d the number of days since the winter solstice; or peaking at other times of the year determined by the function $Xs = \sin 2\pi t$ where $t = \frac{d}{365}$ being d the number of days since 1st January. Figures were edited in Inkscape2 (www.inkscape.org) for esthetics.

Short-term community progression and influence of environmental variables on SAR11
To determine whether SAR11 ASV community changes mirror the environmental changes within short-time scales, we performed a wavelet coherence analysis with R package WaveletComp [57] between Bray–Curtis dissimilarities and environmental Euclidean distances on the weekly WEC sampling from April 2015 to April 2017 (see Supplementary methods). To achieve fixed equidistant sampling dates (every 7 days), we reassigned the date to the closest Monday, as it is when most of the sampling took place, and linearly interpolated the missing data. The data of 17 environmental variables were standardized to values between 0 and 1, based on the minimum and maximum values of each variable, using the formula: $X\prime = \frac{x - \min(x)}{\max(x) - \min(x)}$. These standardized values were used to generate Euclidean distances between subsequent samples representing the change in the environment between weeks.

RESULTS
Decreasing primary productivity is negatively correlated with the increasing temperature observed in open ocean, but not coastal site
BATS is located in the western North Atlantic subtropical gyre (31°40′N, 64°10′W). At this open ocean site, the annual cycle of the
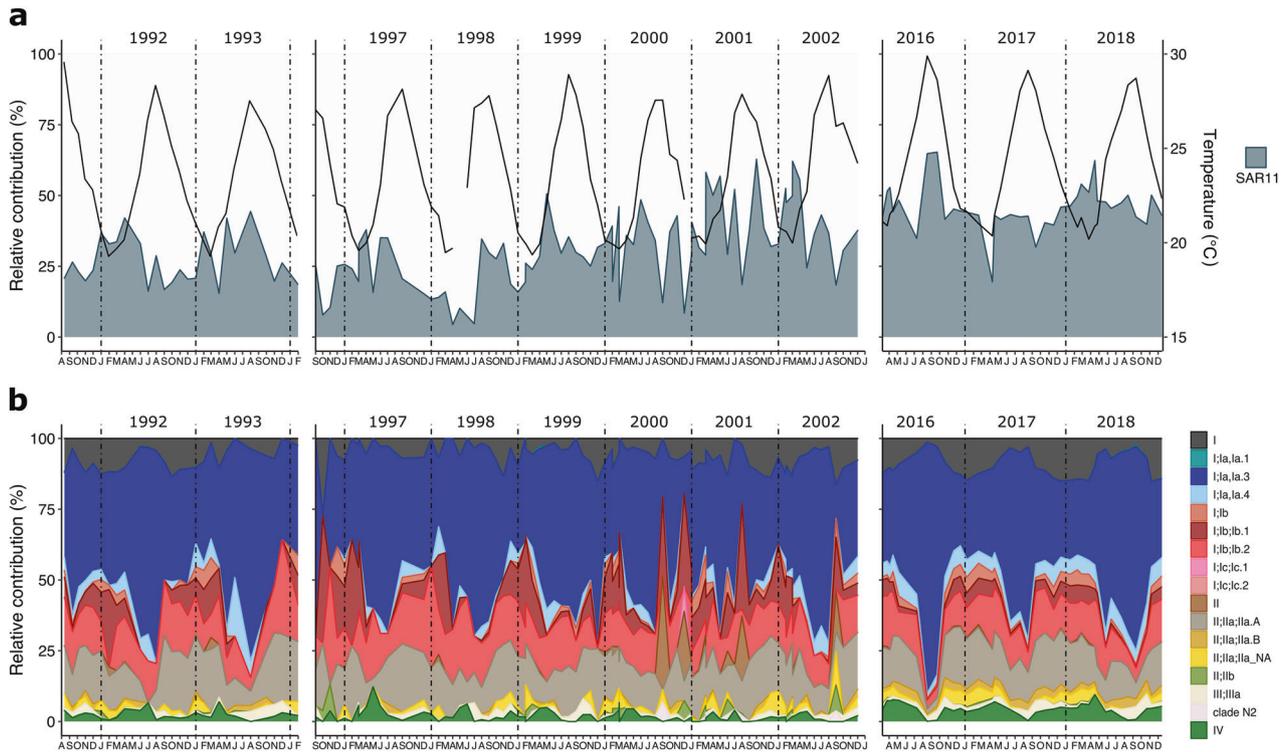
Fig. 1 Monthly relative contribution of surface SAR11 ecotypes through multiannual 16S rRNA surveys at the Bermuda Atlantic Time Series. a Monthly relative contribution of the SAR11 fraction to the total of the amplicon dataset. The black line represents temperature through the sampled years. **b** Monthly relative contribution of ecotypes to the total SAR11 fraction. Samples from 1991–1994 and 1996–2002 were sequenced using 454 FLX technology, while 2016–2018 were sequenced using the Illumina MiSeq platform.

water column is characterized by a strong vertical thermal stratification in summer and the deepening of the mixed layer in winter (150–300 m) [32]. In the years covered by this study, surface water temperature ranged from 19.27 °C to 29.58 °C (Fig. S1). Nitrogen and phosphate nutrients were at or close to analytical detection limits (30 nmol kg^{-1} before 2004 and 1 nmol kg^{-1} after 2004) [58] throughout the sampling, reflecting the oligotrophic nature of the gyre. Chlorophyll concentration showed clear peaks in the winter-spring transitions. Over the course of the time series, spanning 27 years, the peak chlorophyll concentration at 5 m decreased while minimum annual temperatures have increased (Figs. S1, S2). This observation is concurrent with the reduction of primary production and phytoplankton community shift in the last decade at BATS [59]. The L4 coastal station (50°15′N, 4°13′W) is located in the Western English Channel. During the 7 years analyzed in this study, surface water temperature ranged between 7.62 °C and 19.16 °C (Fig. S3). Nitrate and nitrite, silicate, and phosphate concentrations peaked in winter and were depleted during summer. Ammonium concentrations varied throughout the year. Chlorophyll concentration displayed multiple annual peaks, with the highest occurring in spring and did not show a correlation with long-term minimum temperature changes as in BATS (Fig. S2).

SAR11 ecotype seasonality is persistent through multiannual time series in the Sargasso Sea and Western English Channel
At BATS, SAR11 comprised between 4.4% and 65.4% of the community (Fig. 1a). Sequenced with 454 FLX, SAR11 relative contribution from 1991 to 2004 at BATS showed a broad distribution of values (29.7 ± 12.8%). However, the period from 2016 to 2018 (Illumina; MiSeq) displayed a constant high contribution with more than 85% samples >40% SAR11, with a narrower distribution (45.9 ± 8.5%). This difference could be explained by the sequencing saturation achieved by the higher

coverage of Illumina datasets compared to those generated by 454 FLX. SAR11 relative contribution was more variable within the WEC time series (Illumina; MiSeq), ranging from 0.04% to 53.9% (15.7 ± 9.7%) (Fig. 2a) with a decrease in SAR11 contribution during summer months (June, July, August). This result supports previous observations of SAR11 seasonality at WEC [35, 60].

SAR11 Ia ecotypes dominate the surface of both locations, comprising 49% of the total at BATS (Fig. 1b, S4) and 71.9% at WEC (Fig. 2b, S5). Ecotype Ia.3 ("warm water" [12]) dominates at BATS, showing an oscillatory progression reaching a maximum during the stratified summer and decreasing to the lowest contributions in the deepest mixing of the water column in winter, thereafter, returning to an increasing progression. Unexpectedly, given the comparatively low seasonal temperatures (BATS maximum 29.6 °C, minimum 19.3 °C; WEC maximum 19.2 °C, minimum 7.7 °C), "warm water" ecotype Ia.3 ASVs were also highly abundant in the WEC, comprising 35.8% of SAR11 sequences, and displayed a similar annual pattern to that observed in BATS. Ecotype Ia.1, prominent in cold regions [12], was negligible at BATS as expected (Fig. 2b, S6). At WEC, Ia.1 made up 35.9% of SAR11 sequences and displayed an opposite oscillation pattern to Ia.3, with an increasing progression starting in summer (June–July), reaching a peak in early winter (December–January). Ecotype IIa.B was the third most abundant in WEC (18.8%). It displayed a similar annual progression to Ia.1, reaching a maximum peak in winter. Other ecotypes that increased their relative contribution during winter were Ib, II, and IIb. Clade IV displayed low contributions with periodic peaks after summer. At BATS, the progression to winter was dominated by Ib ecotypes and IIa.A (instead of IIa.B at WEC). Additionally, Ia.4, IV, non-determined I and IIa accompanied it. These differences in winter dominant ecotypes between the two time series might be related to adaptations to the different temperature ranges and nutrient concentrations intrinsic to each environment (Figs. S1, S2, S3).
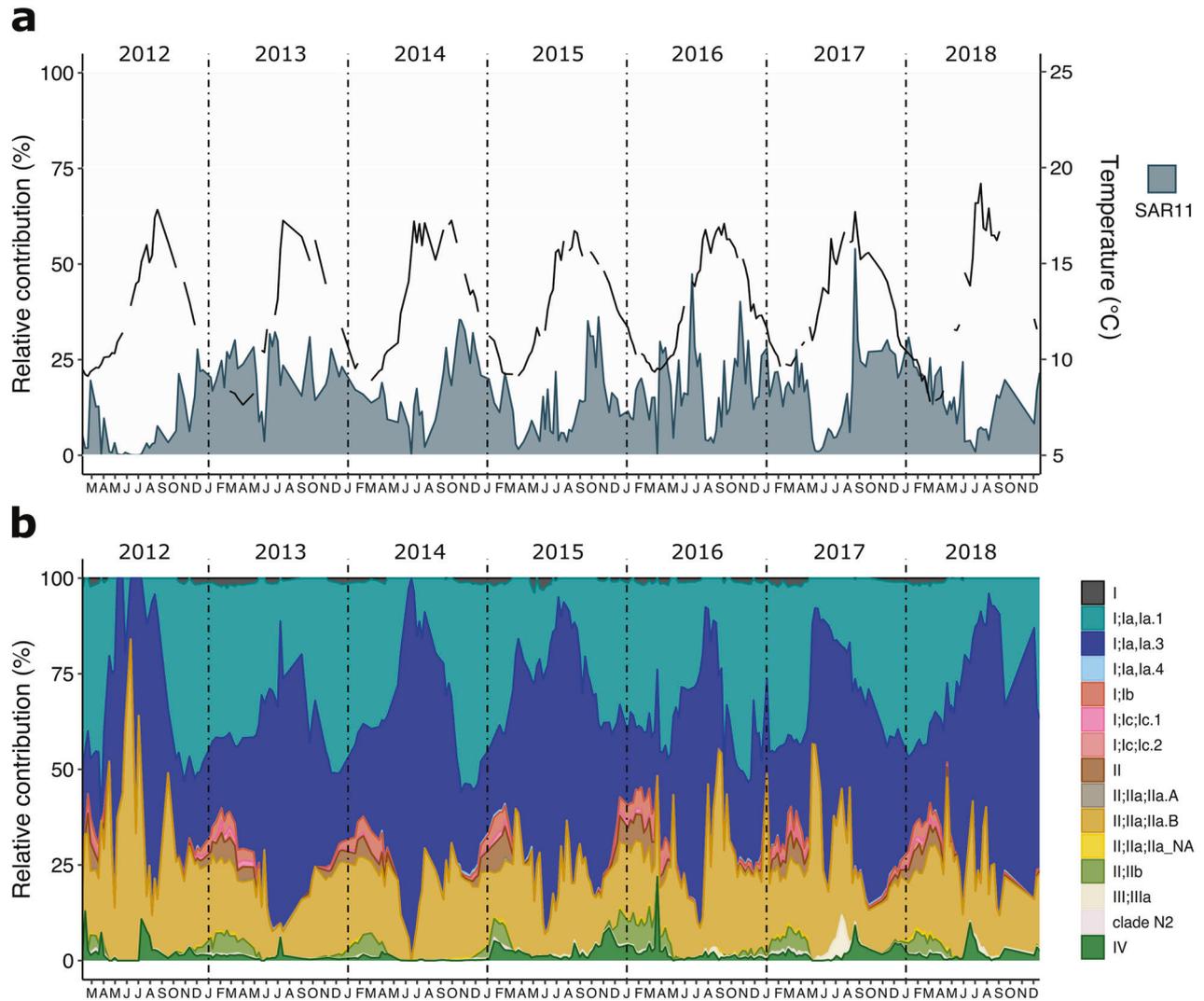
Fig. 2 **Monthly relative contribution of surface SAR11 ecotypes through a multiannual 16S rRNA survey at station L4 from the Western English Channel. a** Monthly relative contribution of the SAR11 fraction to the total of the amplicon dataset. The black line represents temperature through the sampled years. **b** Monthly relative contribution of ecotypes to the total SAR11 fraction.

SAR11 communities showed a recurrent annual cycle in both locations (Fig. 3). Pairwise comparison of Bray–Curtis dissimilarities of rarefied SAR11 ASVs over time showed a periodic annual pattern from 1991 to 2002 in the 454 FLX data (Fig. 3a). This pattern has a short amplitude in community differences, most likely as a result of reduced coverage of the community. The 2016–2018 datasets (Illumina) showed a consistent sinusoidal pattern with a wavelength of 1 year (Fig. 3b). In the WEC, a similar 1-year sinusoidal pattern was also evident, with greater clarity provided by the weekly sampling. Community turnover was maximized approximately every 180 days in the WEC data.

SAR11 seasonality is configured in an annual two-state pattern
A constrained ordination of the SAR11 communities was generated for BATS and WEC based on rarefied Bray-Curtis dissimilarities to determine whether SAR11 annual periodicity follows a strict seasonal progression. The constrained ordination revealed that the most important factors explaining the variance of the communities are temperature, oxygen, and nutrient concentrations (Fig. 4). In both locations, summer and winter SAR11 composition represented the most divergent communities (Fig. 4). Spring and autumn were not characterized by a distinctive

set of ASVs but composed mainly of increasing and decreasing contributions of summer and winter SAR11 ASVs, as if these were transition stages between summer and winter clusters. The samples in the constrained ordination form two clusters that are not determined by astronomical seasons (Fig. 4). Measured environmental variables at BATS and the WEC explained 63% (adjusted 43%) and 60% (adjusted 57%) of the variance in SAR11 communities, respectively. To evaluate whether sequencing methodology introduced significant biases beyond the difference in the coverage, an MDS ordination coupled with an ADONIS test was performed (Fig S7). When testing for the effect of samples sequencing technology origin, the difference was significant (Pr(>F) = 0.001). However, the clustering of the 454 FLX samples is consistent with that of the MiSeq samples, suggesting that this difference may be largely attributed to the difference in the sampling coverage.

Hierarchical clustering of the samples based on the Bray–Curtis dissimilarities of SAR11 ASVs corroborated the "warm" and "cold" separation described in the ordination (Fig. 4) in both BATS and WEC data (Figs. S8, S9). Furthermore, hierarchical clustering uncovered a more detailed pattern of the "warm" and "cold" clusters at BATS. While warm samples clustered together regardless their sequencing origin, the cold cluster split further
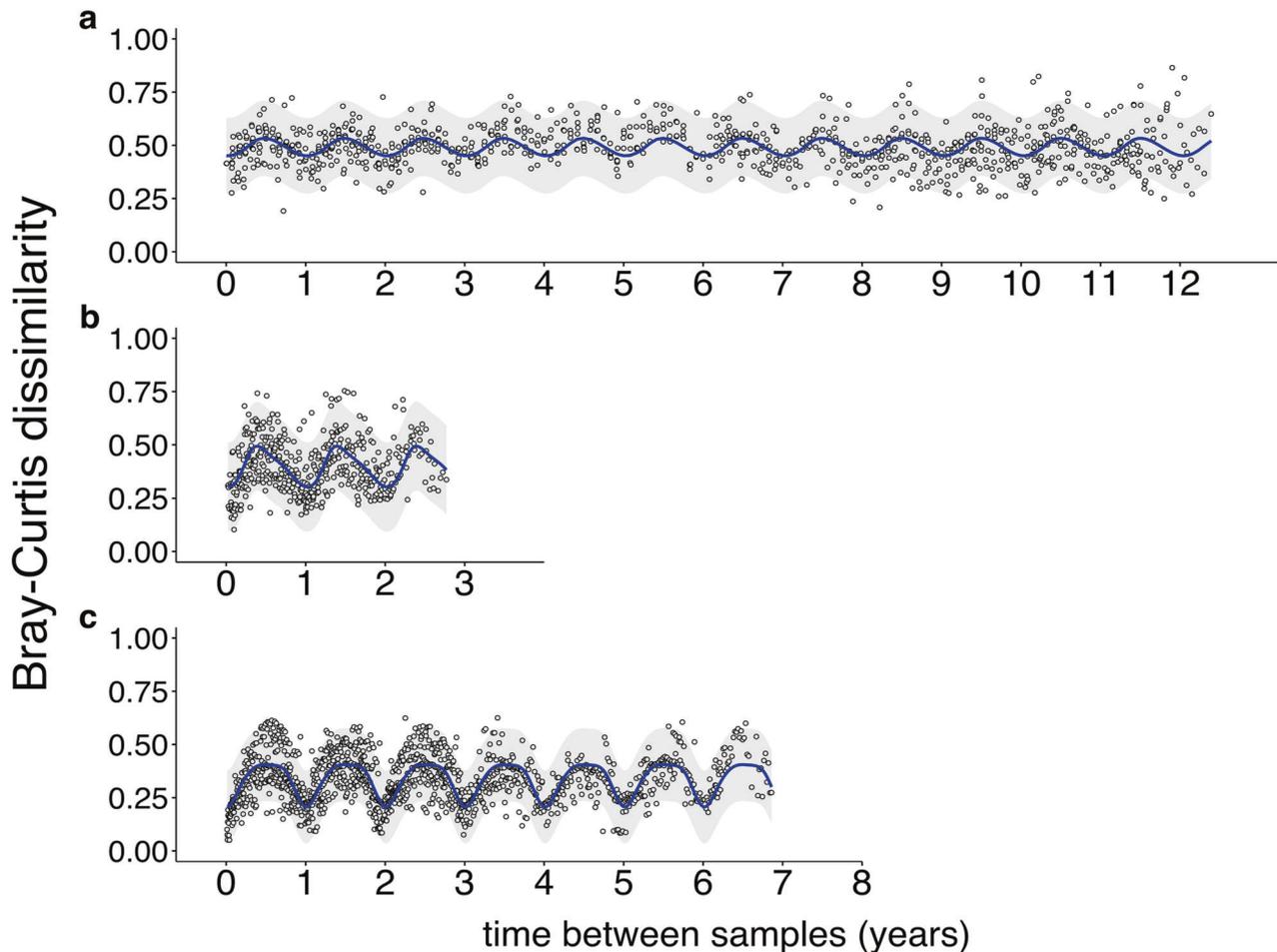
Fig. 3 Inter-annual patterns of Bray–Curtis dissimilarity between SAR11 ASV communities. Pairwise SAR11 community dissimilarity was estimated using the ASV rarefied datasets from BATS and WEC. Bray-Curtis dissimilarities were averaged to generate a one-to-one relationship between dissimilarity and time distance (time datapoint). **a** Average Bray–Curtis dissimilarities of BATS monthly samples from 1991 to 1994 and 1996 to 2002 sequenced using 454 FLX technology. **b** Average Bray–Curtis dissimilarities of BATS monthly samples from 2016 to 2018 sequenced using the Illumina MiSeq platform. **c** Average Bray–Curtis dissimilarities of WEC weekly samples from 2012 to 2018. The blue curve in the three panels represents a periodic spline fit (df = 4) on the Bray-Curtis dissimilarities estimated on a regular interval of 366 days.

into two consistent groups characterized for being mostly dominated either by 454 FLX or MiSeq samples. Considering that no important biases beyond the sequencing coverage were found between MiSeq and 454 FLX datasets (Figs. S6b, S6c, S7) and that minimal temperatures in winter have been increasing steadily in BATS since the year 2000 (Figs. S1, S2), a correlation may exist between increasing temperature and the clustering pattern of "cold" samples. Changes at the ASV level (within the cold ecotypes) might have taken place in the last 30 years at BATS.

A linear regression model was used to classify significant periodic ASVs from BATS (2016–2018) and WEC. Two broad seasonal patterns emerged: ASVs with a peak in relative abundance in summer and ASVs peaking in winter. Both seasonal patterns were consistently correlated with the SAR11 ecotype classification of the ASVs (Fig. 5). Abundant (>0.5%) "warm water" Ia.3 ASVs in both locations, dominate the summer peak pattern. Other ASVs that followed a summer peak pattern at BATS included a single ASV from subclade Ib.2 and multiple low-abundance ASVs from subclade IIIa. In the WEC, only one other ASV belonging to ecotype IV showed a summer peak. ASVs that peaked in winter at BATS included those belonging to Ib.2, IIa, IIa.A and IV. These subclades were not represented in the ASVs peaking in winter at the WEC, which instead comprised subclades Ia.1, Ib, II, IIa.B, and IIb. Thus, while the summer communities at both sites are broadly

similar, the ecological niches at both sites in winter select different SAR11 taxa. Despite having similar abundance peaks, some ASVs had variable peaking times within the season. These may be the result of short-term variability within the seasonal progression of SAR11. These results confirm the consistency of SAR11 ecotype seasonality throughout the multiannual surveys and provide evidence supporting a two-state pattern seasonality.

Interannual weather variability influences SAR11 short term β diversity and its response to ocean physical and chemical changes

A high-resolution, near complete sampling regime in the WEC from April 2015 to April 2017 (Fig. S10) allowed us to analyze how changes in SAR11 β diversity between weekly subsequent samples are correlated to short-term environmental fluctuations (Fig. 6). We generated consecutive Euclidean distances based on 17 environmental measurements: six weather and eleven physico-chemical variables from the ocean surface (Fig. 6a). The Bray–Curtis dissimilarities from consecutive samples showed that during the year 2015, SAR11 communities did not show drastic transitions from week to week (Fig. 6b). In contrast, in 2016 sharp β diversity transitions occurred (Fig. 6b). These steep differentiations were sharpest in the first half of summer and in the transition from autumn to winter. Environmental weekly fluctuations were
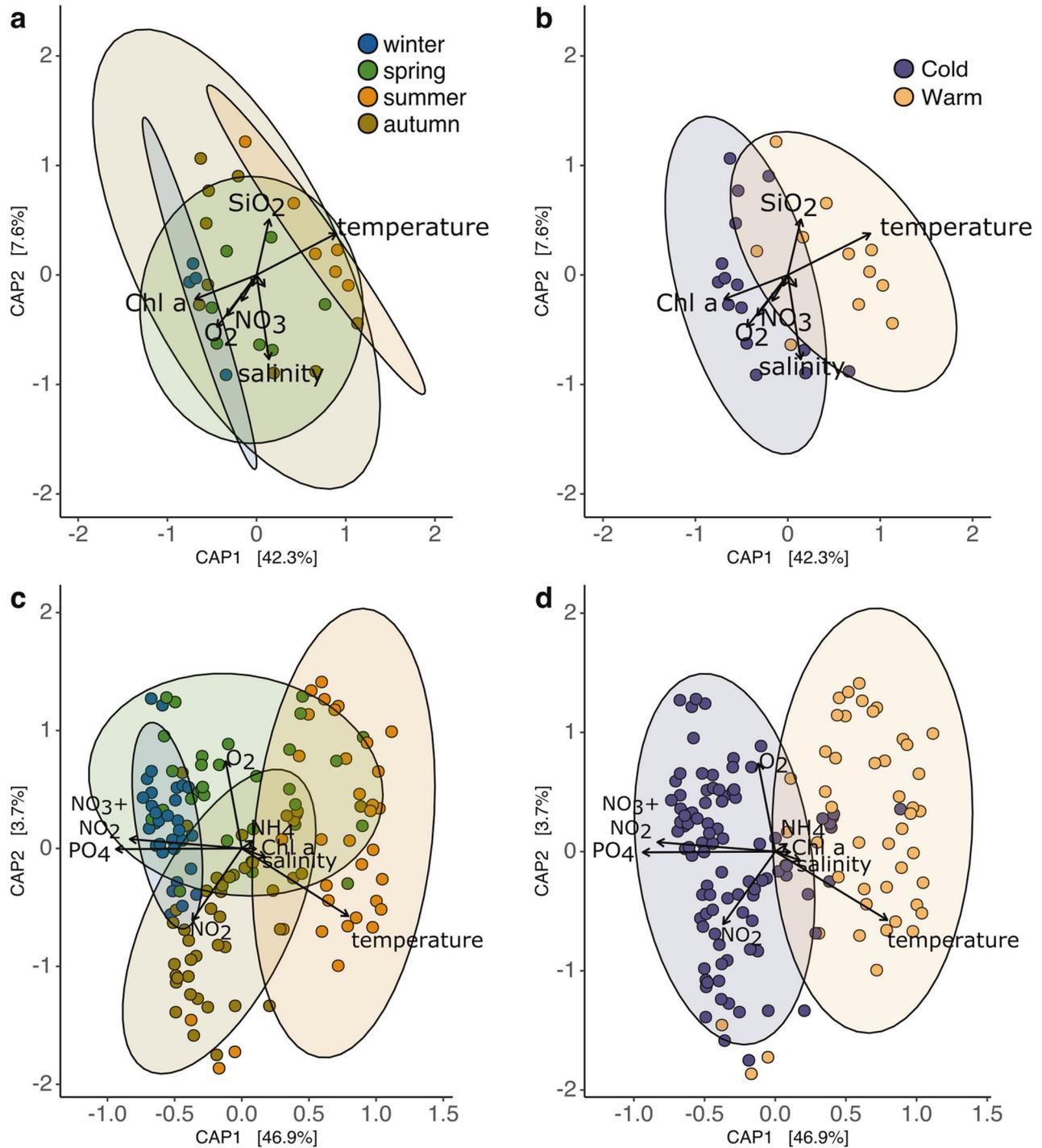
Fig. 4 Canonical Analysis of Principle Coordinates (CAP) of SAR11 ASV profiles constrained by physical and chemical environmental parameters. Ordinations were generated with the Bray–Curtis dissimilarities between samples. Bray–Curtis dissimilarities were estimated using rarefied ASV counts. Axis x and y represent the first and second constrained components, respectively. The percentages between brackets are the fraction of the variance explained by each component under the linear model of the explanatory variables. Overlaid arrows represent the explanatory variables, the arrow's length is proportional to the variation explained and its orientation shows the direction in which the variable increases. **a** BATS monthly samples from 2016 to 2018 color-coded by season. **b** BATS monthly samples from 2016 to 2018 color-coded by cold (22 September–1st May) and warm clusters (2nd May–21 September). **c** WEC weekly samples color-coded by season (**d**) WEC weekly samples color-coded by cold and warm clusters (as in **b**).

variable in both years, however, the changes in 2016 clearly display similar progressions to those of the SAR11 Bray–Curtis dissimilarities (Fig. 6b). Wavelet coherence analysis corroborated that from the vernal equinox of 2016, Euclidean and Bray–Curtis distances were in phase with a significant oscillation period of

16 weeks. In-phase oscillations decoupled after the vernal equinox of 2017 (Fig. S11).

Significantly different environmental variables between the uncoupled and coupled phases (Table S2) were maximum and weekly average wind speed and river flow measurements
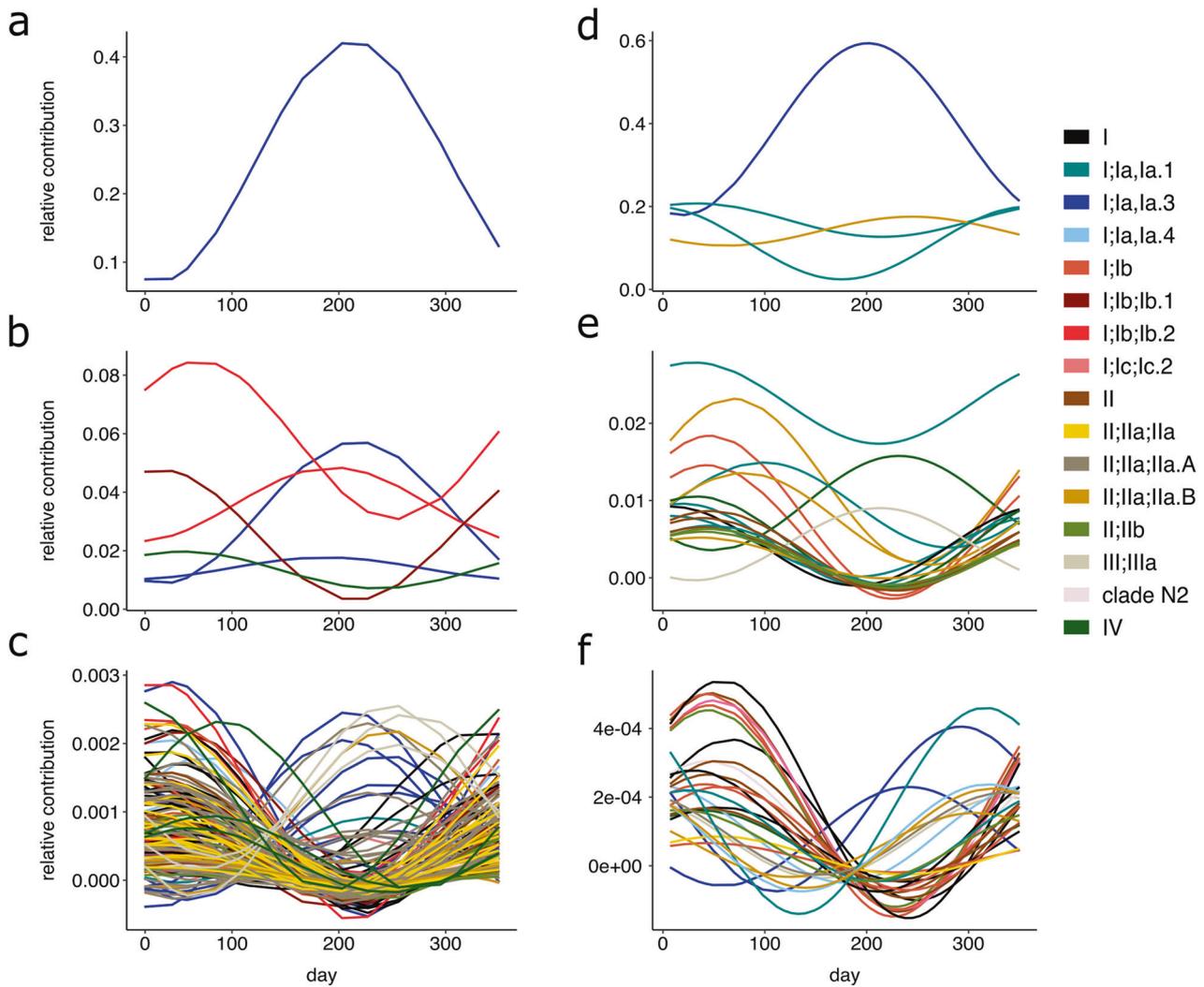
Fig. 5 Annual abundance trends of SAR11 ASVs in the Bermuda Atlantic Time-Series and the Western English Channel. A harmonic linear regression model was used to determine significant seasonal trends ($p < 0.05$) of individual ASVs. Two main patterns of ASVs with seasonal contributions peaking in winter and summer were identified. The Bermuda Atlantic Time-Series ASVs are shown in **a**–**c**. **a** ASVs with a total relative contribution greater than 5%, **b** ASVs with a total relative contribution between 0.5 and 5% and **c** ASVs with a total relative contribution less than 0.5%. The Western English Channel ASVs are shown in **d**–**f**. **d** ASVs with a total relative contribution greater than 10%, **e** ASVs with a total relative contribution between 0.5 and 10% and **f** ASVs with a total relative contribution less than 0.5%. ASV-adjusted model curves are color-coded by ecotype. Day 0 corresponds to the 1st of January.

($p < 0.05$). Net heat flux, which has previously been shown to drive changes within the whole microbial communities of the WEC [61] was not significantly different between the phases of SAR11 communities ($p > 0.05$; Table S2, Fig. S12). These differences suggest that when unsettled weather conditions occur as in 2015, the surface layer of the ocean (~0–5 m) is well-mixed, homogenizing the physicochemical transitions of the water which in turn causes a gradual seasonal progression of SAR11. Contrarily, in calmer weather conditions, as in 2016, physicochemical seasonal transitions of the surface water are sharper thus reflecting the short-term response of SAR11 and its progression within the annual cycle.

DISCUSSION
SAR11 bacteria have been the focus of oceanographic molecular surveys because of their abundance and pivotal role in ocean biogeochemistry [23, 27, 29]. However, these studies have not been directly comparable due to different sampling periods, sampling frequency, molecular data generation protocols, and data analysis. In this study, using a highly defined phylogenetic placement of SAR11 sequences, we analyzed for the first time in a comparable manner SAR11 dynamics of two multiannual molecular surveys in the North Atlantic: BATS and WEC.

A major difference between the molecular surveys was the set of primers used, which amplify different 16S rRNA hypervariable regions. Because of the different conservation levels of the 16S rRNA hypervariable between region V4-V5 (WEC) and V1-V2 (BATS) [62–64] (Fig. S13), we cannot draw conclusive comparisons regarding the total number of SAR11 ASVs identified, perceived diversity and the proportion of seasonal ASVs predicted in each dataset. Furthermore, the data generated with 454 FLX technology has an output of approximately one order of magnitude less than the 2016–2018 Illumina dataset (0.45 Gb 454FLX vs ~7.5 Gb MiSeq). Our results show that the general trends described from the Illumina data are observed in the 454 FLX and no significant biases exist in the generation of the amplicons by itself but rather differences are related to the inherent coverage limitations of each technology. By analyzing these datasets using the same phylogenetic backbone, we overcame the inherent differences
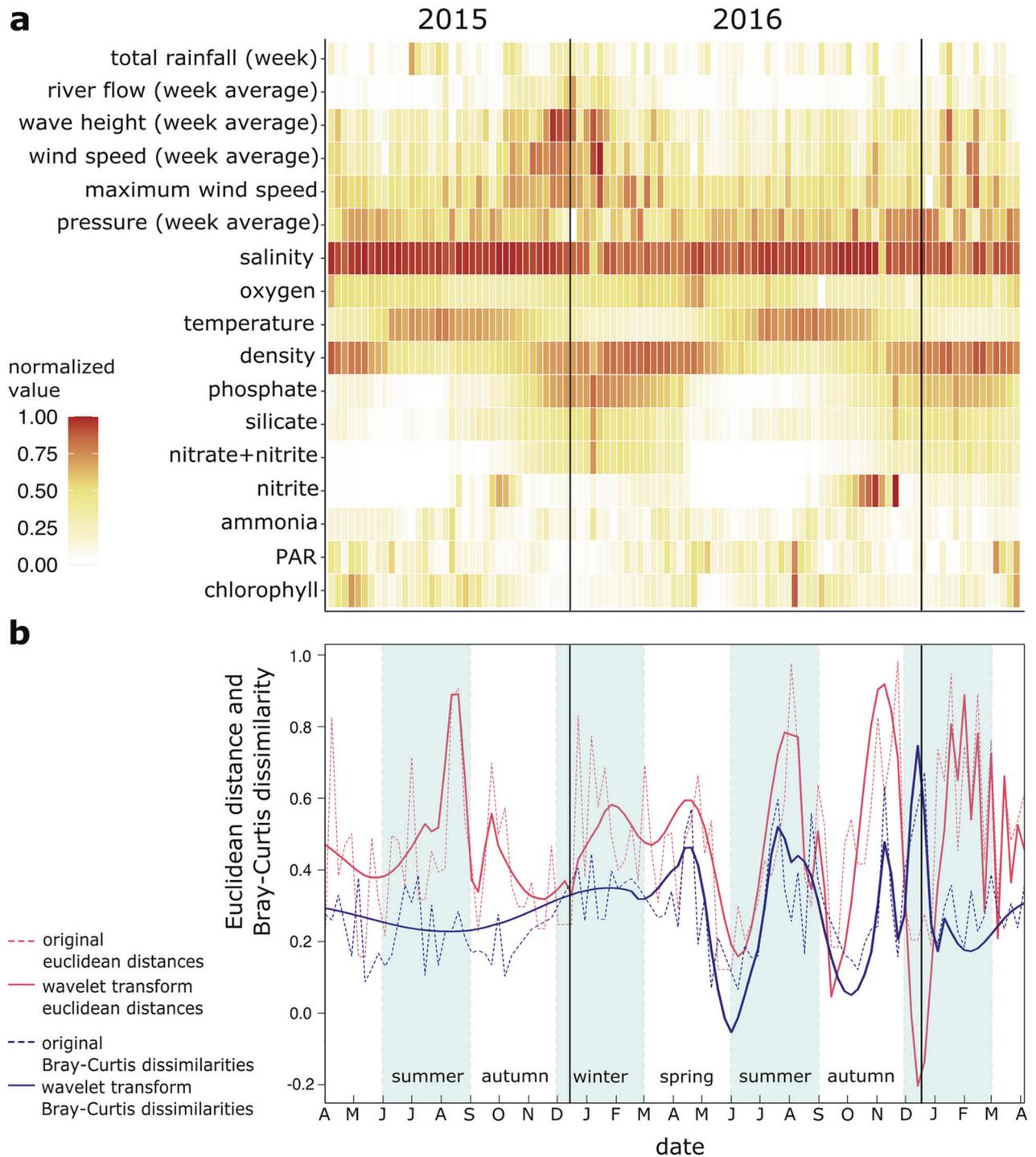
Fig. 6 Response of SAR11 community to environmental heterogeneity on a weekly scale at the Western English Channel. a Normalized values (0–1) of the 17 weather and oceanographic environmental parameters used to estimate the Euclidean distance between weekly consecutive samples. **b** Wavelet transforms (solid red line) of the weekly consecutive Euclidean distances representing the environmental changes and Bray–Curtis dissimilarities of the SAR11 populations (solid blue line). The original Euclidean distances and Bray-Curtis dissimilarities are shown as dashed lines, red and blue respectively. **a** and **b** These are aligned in the x-axis to represent the same week timepoint.

between the time series to provide a panoramic view of SAR11 long-term composition seasonality and the environmental constraints that define it.

Our results showed that SAR11 ecotype seasonality has been consistent in both locations throughout the studied years. We confirm that temperature and nutrients are strong predictors of

SAR11 ecotype geographical distribution and seasonality [3, 12, 22]. However, the specific thresholds that shape SAR11 seasonality and biogeography may be transitional and influenced by other factors such as turbulent mixing, light and biological interactions (viral predation or co-occurring organisms). For example, the warm ecotype Ia.3, thought to be constrained to oligotrophic

environments [65], is a seasonal component at WEC, located at latitude 50°N with high nutrient concentrations and temperatures as low as 7 °C. This result may be explained by genomospecies within the ecotype with specific adaptations to different environments [13, 66] or by horizontal water transport that disperse planktonic microorganisms stirring its distribution across biogeographical regions [66–68]. Whether the Gulf Current constantly seeds the WEC with warm-adapted bacteria remains to be investigated. In contrast, the comparison between WEC and BATS suggests that cold ecotype Ia.1 [69], is restricted to a maximum temperature of 20 °C, therefore hardly retrieved from BATS (minimum temperature 19.27 °C). Cultivable representatives, HTCC1062 (Ia.1) and HTCC1072 (Ia.3), have optimal growth temperatures of 16 °C and 23 °C, respectively [70]. However, the presence of Ia.3 through a larger range of temperatures than Ia.1 suggests a wider landscape of genetic adaptability or phenotypic plasticity.

Ecotypes from clade II also appear to have specific limits of temperature and nutrients shaping their seasonality and occurrence at BATS and WEC. While subclade IIa.A peaked at BATS in winter and early spring, subclades IIa.B and IIb did in the WEC. In contrast, a recent study from the Blanes Bay Microbial Observatory (BBMO), a time series at the coastal oligotrophic NW Mediterranean Sea, showed no seasonal trends of SAR11 clade II ASVs [71]. To date, no single cultures of SAR11 II subclades have been reported, limiting our capacity to experimentally verify optimal growth temperatures. SAR11 IIa.A are prominently found in the mesopelagic and oxygen minimum zones of open oceans [72, 73], which partly explains their presence in winter at BATS surface, when deep mixing occurs. SAR11 IIa.B has been reported to be more abundant in specific regions such as the Mediterranean Sea [13], which is a more closely related regime to the Western English Channel. Additionally, it has been shown that overall SAR11 clade II were abundant and diverse contributors of the particle-associated fraction in the multiannual San-Pedro Ocean Time series (SPOTS) [74], adding a previously unexplored link between SAR11 and organic matter that may also contribute as a source of variability that need to be further studied. The contrasting seasonality of subclade II ecotypes at BBMO, WEC, and BATS and the differential abundance in size fractions at SPOTS suggest that environmental factors, such as weather conditions, horizontal transport, or the abundance of particles may expand or constrain this ecotype seasonality and distribution.

Weekly SAR11 transitions were uncoupled from the environmental transitions from April 2015 to March 2016 in the WEC. In contrast, from March 2016 to April 2017 transitions were coupled. Significant differences between these years were wind speed and river flow, which we assumed created a more turbulent surface; particularly record-breaking weather in winter 2015–2016 (flooding, high winds, and temperature) across the United Kingdom [75]. We hypothesize that these unsettled weather conditions kept a constant mixing of the surface creating a more homogeneous system and blurring the sharp seasonal transitions that occur in more stable years. Annual differences in wind and river input are common throughout the time series (Fig. S14) suggesting that coupling and uncoupling of SAR11 β diversity to water conditions may be a common phenomenon.

We demonstrate that surface ocean SAR11 populations are influenced by short- and long-time changes in environmental drivers at two ocean sites that differ dramatically in average productivity and temperature. Patterns of seasonal succession were fundamentally similar between the two sites, but major differences in ecotype dominance were observed that foreshadow changes in ocean ecology likely to occur as ocean temperature increases. Extreme weather events, which are also predicted to increase [76], affect SAR11 surface ecotype progression. Given SAR11 ecotypes utilize different carbon sources and release different gaseous compounds, the result of these changes may have hitherto unknown impacts on ocean and atmospheric processes.

REFERENCES

1. Thrash JC, Boyd A, Huggett MJ, Grote J, Carini P, Yoder RJ, et al. Phylogenomic evidence for a common ancestor of mitochondria and the SAR11 clade. Sci Rep. 2011;1:9.
2. Ferla MP, Thrash JC, Giovannoni SJ, Patrick WM. New rRNA gene-based phylogenies of the alphaproteobacteria provide perspective on major groups, mitochondrial ancestry and phylogenetic instability. PLoS One. 2013;8:e83383.
3. Giovannoni SJ. SAR11 bacteria: the most abundant plankton in the oceans. Annu Rev Mar Sci. 2017;9:231–55.
4. Zhao X, Schwartz CL, Pierson J, Giovannoni SJ, McIntosh RJ, Nicastro D. Three-dimensional structure of the ultraoligotrophic marine bacterium "Candidatus pelagibacter ubique". Appl Environ Microbiol. 2017;83:807–16.
5. Giovannoni SJ, DeLong EF, Schmidt TM, Pace NR. Tangential flow filtration and preliminary phylogenetic analysis of marine picoplankton. Appl Environ Microbiol. 1990;56:4.
6. Morris RM, Rappé MS, Connon SA, Vergin KL, Siebold WA, Carlson CA, et al. SAR11 clade dominates ocean surface bacterioplankton communities. Nature. 2002;420:806–10.
7. Rappé MS, Connon SA, Vergin KL, Giovannoni SJ. Cultivation of the ubiquitous SAR11 marine bacterioplankton clade. Nature. 2002;418:630–3.
8. Grote J, Thrash JC, Huggett MJ, Landry ZC, Carini P, Giovannoni SJ, et al. Streamlining and core genome conservation among highly divergent members of the SAR11 clade. mBio. 2012;3:e00252–12.
9. Field KG, Gordon D, Wright T, Rappé M, Urback E, Vergin K, et al. Diversity and depth-specific distribution of SAR11 cluster rRNA genes from marine planktonic bacteria. Appl Environ Microbiol. 1997;63:63–70.
10. Suzuki MT, Beja O, Taylor LT, DeLong EF. Phylogenetic analysis of ribosomal RNA operons from uncultivated coastal marine bacterioplankton. Environ Microbiol. 2001;3:323–31.
11. Carlson CA, Morris R, Parsons R, Treusch AH, Giovannoni SJ, Vergin K. Seasonal dynamics of SAR11 populations in the euphotic and mesopelagic zones of the northwestern Sargasso Sea. ISME J. 2009;3:283–95.
12. Brown MV, Lauro FM, DeMaere MZ, Muir L, Wilkins D, Thomas T, et al. Global biogeography of SAR11 marine bacteria. Mol Syst Biol. 2012;8:595.
13. Haro-Moreno JM, Rodriguez-Valera F, Rosselli R, Martinez-Hernandez F, Roda-Garcia JJ, Gomez ML, et al. Ecogenomics of the SAR11 clade. Environ Microbiol. 2020;22:1748–63.
14. Carini P, White AE, Campbell EO, Giovannoni SJ. Methane production by phosphate-starved SAR11 chemoheterotrophic marine bacteria. Nat Commun. 2014;5:4346.
15. Sun J, Steindler L, Thrash JC, Halsey KH, Smith DP, Carter AE, et al. One carbon metabolism in SAR11 Pelagic marine bacteria. PLoS One. 2011;6:e23973.
16. Schwalbach MS, Tripp HJ, Steindler L, Smith DP, Giovannoni SJ. The presence of the glycolysis operon in SAR11 genomes is positively correlated with ocean productivity. Environ Microbiol. 2010;12:490–500.
17. Sun J, Todd JD, Thrash JC, Qian Y, Qian MC, Temperton B, et al. The abundant marine bacterium Pelagibacter simultaneously catabolizes dimethylsulfoniopropionate to the gases dimethyl sulfide and methanethiol. Nat Microbiol. 2016;1:16065.
18. Halsey KH, Giovannoni SJ, Graus M, Zhao Y, Landry Z, Thrash JC, et al. Biological cycling of volatile organic carbon by phytoplankton and bacterioplankton: VOC cycling by marine plankton. Limnol Oceanogr. 2017;62:2650–61.
19. Carlson CA, Giovannoni SJ, Hansell DA, Goldberg SJ, Parsons R, Vergin K. Interactions among dissolved organic carbon, microbial processes, and community structure in the mesopelagic zone of the northwestern Sargasso Sea. Limnol Oceanogr. 2004;49:1073–83.
20. Wagner S, Schubotz F, Kaiser K, Hallmann C, Waska H, Rossel PE, et al. Soothsaying DOM: a current perspective on the future of oceanic dissolved organic carbon. Front Mar Sci. 2020;7:341.
21. Quinn PK, Bates TS. The case against climate regulation via oceanic phytoplankton sulphur emissions. Nature. 2011;480:51–6.
22. Bolaños LM, Choi CJ, Worden AZ, Baetge N, Carlson CA, Giovannoni S. Seasonality of the microbial community composition in the North Atlantic. Front Mar Sci. 2021;8:624164.
23. Tucker SJ, Freel KC, Monaghan EA, Sullivan CES, Ramfelt O, Rii YM, et al. Spatial and temporal dynamics of SAR11 marine bacteria across a nearshore to offshore transect in the tropical Pacific Ocean. PeerJ. 2021;9:e12274.

24. Giovannoni SJ, Vergin KL. Seasonality in ocean microbial communities. Science. 2012;335:671–6.

25. Eren AM, Maignien L, Sul WJ, Murphy LG, Grim SL, Morrison HG, et al. Oligotyping: differentiating between closely related microbial taxa using 16S RRNA gene data. Methods Ecol Evol. 2013;4:1111–9.

26. Vergin K, Done B, Carlson C, Giovannoni S. Spatiotemporal distributions of rare bacterioplankton populations indicate adaptive strategies in the oligotrophic ocean. Aquat Microb Ecol. 2013;71:1–13.

27. Salter I, Galand PE, Fagervold SK, Lebaron P, Obernosterer I, Oliver MJ, et al. Seasonal dynamics of active SAR11 ecotypes in the oligotrophic Northwest Mediterranean Sea. ISME J. 2015;9:347–60.

28. Ortmann AC, Santos TTL. Spatial and temporal patterns in the Pelagibacteraceae across an estuarine gradient. FEMS Microbiol Ecol. 2016;92:fiw133.

29. Vergin KL, Beszteri B, Monier A, Cameron Thrash J, Temperton B, Treusch AH, et al. High-resolution SAR11 ecotype dynamics at the Bermuda Atlantic Time-series Study site by phylogenetic placement of pyrosequences. ISME J. 2013;7:1322–32.

30. Needham DM, Fichot EB, Wang E, Berdjeb L, Cram JA, Fichot CG, et al. Dynamics and interactions of highly resolved marine plankton via automated high-frequency sampling. ISME J. 2018;12:2417–32.

31. Benway HM, Lorenzoni L, White AE, Fiedler B, Levine NM, Nicholson DP, et al. Ocean time series observations of changing marine ecosystems: an era of integration, synthesis, and societal applications. Front Mar Sci. 2019;12:6–393.

32. Steinberg DK, Carlson CA, Bates NR, Johnson RJ, Michaels AF, Knap AH. Overview of the US JGOFS Bermuda Atlantic Time-series Study (BATS): a decade-scale look at ocean biology and biogeochemistry. Deep Sea Res Part II Top Stud Oceanogr. 2001;48:1405–47.

33. Southward AJ, Langmead O, Hardman-Mountford NJ, Aiken J, Boalch GT, Dando PR, et al. Long-term oceanographic and ecological research in the Western English Channel. In: Advances in marine biology. Elsevier. 2005;47:1–105.

34. Gilbert JA, Field D, Swift P, Newbold L, Oliver A, Smyth T, et al. The seasonal structure of microbial communities in the Western English Channel. Environ Microbiol. 2009;11:3132–9.

35. Gilbert JA, Steele JA, Caporaso JG, Steinbrück L, Reeder J, Temperton B, et al. Defining seasonal marine microbial community dynamics. ISME J. 2012;6:298–308.

36. Caporaso JG, Paszkiewicz K, Field D, Knight R, Gilbert JA. The Western English Channel contains a persistent microbial seed bank. ISME J. 2012;6:1089–93.

37. Warwick-Dugdale J, Solonenko N, Moore K, Chittick L, Gregory AC, Allen MJ, et al. Long-read viral metagenomics captures abundant and microdiverse viral populations and their niche-defining genomic islands. PeerJ. 2019;7:e6800.

38. Vergin KL, Done B, Carlson CA, Giovannoni SJ. Spatiotemporal distributions of rare bacterioplankton populations indicate adaptive strategies in the oligotrophic ocean. Aquat Microb Ecol. 2013;71:1–3.

39. Choi CJ, Jimenez V, Needham DM, Poirier C, Bachy C, Alexander H, et al. Seasonal and geographical transitions in eukaryotic phytoplankton community structure in the Atlantic and Pacific Oceans. Front Microbiol. 2020;11:542372.

40. Bolaños LM, Karp-Boss L, Choi CJ, Worden AZ, Graff JR, Haëntjens N, et al. Small phytoplankton dominate western North Atlantic biomass. ISME J. 2020;14:1663–74.

41. Matsen FA, Kodner RB, Armbrust E. pplacer: linear time maximum-likelihood and Bayesian phylogenetic placement of sequences onto a fixed reference tree. BMC Bioinform. 2010;11:1–6.

42. Treusch AH, Vergin KL, Finlay LA, Donatz MG, Burton RM, Carlson CA, et al. Seasonality and vertical structure of microbial communities in an ocean gyre. ISME J. 2009;3:1148–63.

43. Giovannoni SJ, Rappe MS, Vergin KL, Adair NL. 16S rRNA genes reveal stratified open ocean bacterioplankton populations related to the Green Non-Sulfur bacteria. Proc Natl Acad Sci. 1996;93:7979–84.

44. Morris RM, Vergin KL, Cho J-C, Rappé MS, Carlson CA, Giovannoni SJ. Temporal and spatial response of bacterioplankton lineages to annual convective overturn at the Bermuda Atlantic Time-series Study site. Limnol Oceanogr. 2005;50:1687–96.

45. Daims H, Brühl A, Amann R, Schleifer K-H, Wagner M. The domain-specific probe EUB338 is insufficient for the detection of all bacteria: development and evaluation of a more comprehensive probe set. Syst Appl Microbiol. 1999;22:434–44.

46. Lane DJ. Nucleic acid techniques in bacterial systematics. In: Nucleic acid techniques in bacterial systematics. New York: Wiley; p. 115–75.

47. Callahan BJ, McMurdie PJ, Holmes SP. Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. ISME J. 2017;11:2639–43.

48. McMurdie PJ, Holmes S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. PLoS One. 2013;8:e61217.

49. Eren AM, Borisy GG, Huse SM, Mark Welch JL. Oligotyping analysis of the human oral microbiome. Proc Natl Acad Sci. 2014;111:E2875–84.

50. Buchholz HH, Michelsen ML, Bolaños LM, Browne E, Allen MJ, Temperton B. Efficient dilution-to-extinction isolation of novel virus–host model systems for fastidious heterotrophic bacteria. ISME J. 2021;15:1585–98.

51. R Core Team. R: a language and environment for statistical computing. Vienna, Austria; https://www.R-project.org/

52. Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, et al. Package "vegan".

53. Wickham H. ggplot2: ggplot2. Wiley Interdiscip Rev Comput Stat. 2011;3:180–5.

54. Wang W, Yan J. Shape-restricted regression splines with R package splines2. J Data Sci. 2021;19:498–517.

55. Auladell A, Sánchez P, Sánchez O, Gasol JM, Ferrera I. Long-term seasonal and interannual variability of marine aerobic anoxygenic photoheterotrophic bacteria. ISME J. 2019;13:1975–87.

56. Ahdesmaki M, Fokianos K, Strimmer K, Ahdesmaki MM. Package 'GeneCycle' 2015.

57. Roesch A, Schmidbauer H and Roesch MA. Package 'WaveletComp.' 2014.

58. Lomas MW, Bates NR, Johnson RJ, Knap AH, Steinberg DK, Carlson CA. Two decades and counting: 24-years of sustained open ocean biogeochemical measurements in the Sargasso Sea. Deep Sea Res Part II Top Stud Oceanogr. 2013;93:16–32.

59. Lomas MW, Bates NR, Johnson RJ, Steinberg DK, Tanioka T. Adaptive carbon export response to warming in the Sargasso Sea. Nature Commun. 2022;13:1–0.

60. Sargeant SL, Murrell JC, Nightingale PD, Dixon JL. Basin-scale variability of microbial methanol uptake in the Atlantic Ocean. Biogeosciences. 2018;15:5155–67.

61. Smyth TJ, Allen I, Atkinson A, Bruun JT, Harmer RA, Pingree RD, et al. Ocean net heat flux influences seasonal to interannual patterns of plankton abundance. PLoS One. 2014;9:e98709.

62. Van de Peer Y. A quantitative map of nucleotide substitution rates in bacterial rRNA. Nucleic Acids Res. 1996;24:3381–91.

63. Baker GC, Smith JJ, Cowan DA. Review and re-analysis of domain-specific 16S primers. J Microbiol Methods. 2003;55:541–55.

64. Vasileiadis S, Puglisi E, Arena M, Cappa F, Cocconcelli PS, Trevisan M. Soil bacterial diversity screening using single 16S rRNA gene V regions coupled with multi-million read generating sequencing technologies. PLoS ONE. 2012;7:e42671.

65. Stingl U, Tripp HJ, Giovannoni SJ. Improvements of high-throughput culturing yielded novel SAR11 strains and other abundant marine bacteria from the Oregon coast and the Bermuda Atlantic Time-series study site. ISME J. 2007;1:361–71.

66. Delmont TO, Kiefl E, Kilinc O, Esen OC, Uysal I, Rappé MS, et al. Single-amino acid variants reveal evolutionary processes that shape the biogeography of a global SAR11 subclade. eLife. 2019;8:e46497.

67. Lévy M, Jahn O, Dutkiewicz S, Follows MJ, d'Ovidio F. The dynamical landscape of marine phytoplankton diversity. J R Soc Interface. 2015;12:20150481.

68. Hellweger FL, van Sebille E, Calfee BC, Chandler JW, Zinser ER, Swan BK, et al. The role of ocean currents in the temperature selection of plankton: insights from an individual-based model. PLoS ONE. 2016;11:e0167010.

69. Giovannoni SJ, Tripp HJ, Givan S, Podar M, Vergin KL, Baptista D, et al. Genome streamlining in a cosmopolitan oceanic bacterium. Science. 2005;309:1242–5.

70. Brown SN, Giovannoni S, Cho JC. Polyphasic taxonomy of marine bacteria from the SAR11 group Ia: Pelagibacter ubiquis (strain HTCC1062) & Pelagibacter bermudensis (strain HTCC7211). Oregon State University; 2012.

71. Auladell A, Barberán A, Logares R, Garcés E, Gasol JM, Ferrera I. Seasonal niche differentiation among closely related marine bacteria. ISME J. 2022;16:178–89.

72. Tsementzi D, Wu J, Deutsch S, Nath S, Rodriguez-R LM, Burns AS, et al. SAR11 bacteria linked to ocean anoxia and nitrogen loss. Nature. 2016;536:179–83.

73. Ruiz-Perez CA, Bertagnolli AD, Tsementzi D, Woyke T, Stewart FJ, Konstantinidis KT. Description of Candidatus Mesopelagibacter carboxydoxydans and Candidatus Anoxipelagibacter denitrificans: nitrate-reducing SAR11 genera that dominate mesopelagic and anoxic marine zones. Syst Appl Microbiol. 2021;44:126185.

74. Yeh YC, Fuhrman JA. Contrasting diversity patterns of prokaryotes and protists over time and depth at the San-Pedro Ocean Time series. ISME Commun. 2022;13:1–12.

75. McCarthy M, Spillane S, Walsh S, Kendon M. The meteorology of the exceptional winter of 2015/2016 across the UK and Ireland. Weather. 2016;71:305–13.

76. Met Office. UK Climate Projections: Headline Findings. 2021.

ACKNOWLEDGEMENTS

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/s43705-022-00198-1.

Correspondence and requests for materials should be addressed to Luis M. Bolaños or Ben Temperton.

Reprints and permission information is available at http://www.nature.com/reprints

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