environmental microbiology

Environmental Microbiology (2021) **00**(00), 00–00 doi:10.1111/1462-2920.15424

Revealing environmentally driven population dynamics of an Arctic diatom using a novel microsatellite PoolSeq barcoding approach

Klara K. E. Wolf,[1*] **Clara J. M. Hoppe**,[1] **Florian Leese**,[2] **Martina Weiss**,[2] **Björn Rost**,[1,3] **Stefan Neuhaus**,[1] **Thilo Gross**,[1,4,5] **Nancy Kühne**[1] **and Uwe John**[1,5]

[1]*Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, Germany.*

[2]*Faculty of Biology, Aquatic Ecosystem Research, University of Duisburg-Essen, Essen, Germany.*

[3]*University of Bremen, FB2, Bremen, Germany.*

[4]*University of Oldenburg, ICBM, Oldenburg, Germany.*

[5]*Helmholtz Institute for Functional Marine Biodiversity, Oldenburg, Germany.*

Summary

Ecological stability under environmental change is determined by both interspecific and intraspecific processes. Particularly for planktonic microorganisms, it is challenging to follow intraspecific dynamics over space and time. We propose a new method, microsatellite PoolSeq barcoding (MPB), for tracing allele frequency changes in protist populations. We successfully applied this method to experimental community incubations and field samples of the diatom *Thalassiosira hyalina* from the Arctic, a rapidly changing ecosystem. Validation of the method found compelling accuracy in comparison with established genotyping approaches within different diversity contexts. In experimental and environmental samples, we show that MPB can detect meaningful patterns of population dynamics, resolving allelic stability and shifts within a key diatom species in response to experimental treatments as well as different bloom phases and years. Through our novel MPB approach, we produced a large dataset of populations at different time-points and locations with comparably little effort. Results like this can add insights into the roles of selection and plasticity in natural protist populations under stable experimental but also variable field conditions. Especially for organisms where genotype sampling remains challenging, MPB holds great potential to efficiently resolve eco-evolutionary dynamics and to assess the mechanisms and limits of resilience to environmental stressors.

Introduction

Phytoplankton responses to ongoing and future environmental changes will significantly impact earth system processes at many levels. These primary producers are the photosynthetic base of marine food webs and responsible for half of the global oxygen production (Field *et al*., 1998). In the Arctic, climate change is progressing at a much faster rate than the global average (Larsen *et al*., 2014), providing a natural laboratory to study organismal responses. Generally, phytoplankton are considered relatively resistant to rapid environmental changes owing to their large census population sizes and short generation times (Lynch *et al*., 1991; Collins *et al*., 2014; Rengefors *et al*., 2017). However, even minor changes in competitive ability at the population level and species shifts within communities have the potential to impact higher trophic levels, biogeochemical cycling, and ecosystem functioning (Rost *et al*., 2008; Hillebrand and Matthiessen, 2009; Boyd *et al*., 2018).

Adjustments to environmental change can occur by physiological, ecological, and evolutionary mechanisms. Organisms have the capacity to modify their phenotypes within certain intrinsic limits by a plastic response to surrounding conditions (West-Eberhard, 2003). The range of conditions in which an organism can dwell defines its plasticity. Furthermore, the genotypic composition of a population can be adjusted by selection for fitter individuals, which is referred to as strain or genotype sorting (Becks *et al*., 2010; Collins *et al*., 2014; Scheinin *et al*., 2015). Individuals with an adaptive advantage under changed environmental conditions can emerge by new mutations, by immigration from other locations, or are already present within the standing genetic stock of a population. Large effective population sizes host a

Received 7 April, 2020; revised 1 February, 2021; accepted 2 February, 2021. *For correspondence. E-mail: klara.wolf@awi.de; Tel. (+49) 471 4831-1450; Fax +49(471)4831-1149.

greater genetic diversity and thus a larger potential to adjust more rapidly to occurring change. This is why large and highly diverse populations, as found in phytoplankton, are considered to be particularly resistant to environmental variation and change through fast evolutionary adaptation by selection within a large standing stock of phenotypes (Yachi and Loreau, 1999; Bernhardt and Leslie, 2013).

The high intraspecific diversity of phytoplankton (Godhe and Rynearson, 2017) makes the dynamics and selective processes within populations also methodologically difficult to resolve. Because classic sample sizes of isolated genotypes (up to a few hundred) are rarely sufficient to detect patterns within the usually huge genotypic diversity, few studies have found evidence for genotype sorting in marine environments (but see Scheinin *et al*., 2015; Ruggiero *et al*., 2017). A major challenge in population genetic studies is that individuals can only be distinguished via highly polymorphic loci and not via conventional 'DNA barcoding' gene markers. Furthermore, traditional approaches to resolve population composition, such as microsatellites, cannot be directly applied to diverse natural community samples. Instead, they must be performed separately for a number of tediously isolated strains. Researchers are therefore caught in a practical tradeoff between a higher resolution per population (i.e. number of isolates) and a higher frequency of temporal or spatial sampling. A recently established method is pooled sequencing (PoolSeq; Futschik and Schlötterer, 2010), which so far has only been applied to multicellular and larger eukaryotes. In this method, single nucleotide polymorphisms (SNPs) of populations are analysed using a shotgun or targeted approach by mixing strains, tissues, or DNA of conspecifics for sequencing. Based on this technique, more targeted approaches for genotyping have also been developed ('Genotyping-by-Sequencing', e.g. Narum *et al*., 2013; Vartia *et al*., 2016). While PoolSeq allows the analysis of many individuals simultaneously and yields results at the level of nucleotide sequences instead of relying on length-polymorphisms, it still requires samples that contain genotypes of only the target species (Ferretti *et al*. 2013), which is challenging to achieve in natural samples of planktonic protists.

Due to these methodological constraints, little is known about the complex population structures and dynamics of phytoplankton. Despite the usual absence of visible clonal dominance (but see Ruggiero *et al*., 2017), allele frequency data revealed that phytoplankton populations are often distinct, even in adjacent and frequently intermixing locations (Rynearson and Armbrust, 2004; Medlin, 2007; Gaebler-Schwarz *et al*., 2015). While some studies suggest that subpopulations in the same place can coexist and replace each other throughout a season (Rynearson *et al*., 2006; Saravanan and Godhe, 2010; Erdner *et al*., 2011) or that population structure develops over time (Tammilehto *et al*., 2017), others have shown that populations can remain relatively stable across timescales of decades or centuries (Härnström *et al*., 2011). From the available evidence, it appears increasingly likely that population structure is much more influenced by selection according to ambient conditions (i.e. local adaptation) than by genetic drift (Sjöqvist *et al*., 2015; Godhe *et al*., 2016). This existing knowledge is mainly based on natural populations in which selection environments are highly variable and sampling procedures are tedious, allowing only snapshots of their dynamic state without elucidating underlying processes.

In order to find answers about the role of intraspecific shifts in diverse community experiments and natural environments, our study had the aim of developing and validating a new 'microsatellite PoolSeq barcoding' (MPB) method. This method directly assesses intraspecific variation at microsatellite loci for an entire phytoplankton population from bulk community samples without previous strain isolation. We first compared the results of established methods based on fragment analysis with results of MPB, by applying both methods to simplified experimental populations as well as entire phytoplankton communities of the diatom *Thalassiosira hyalina* in manipulation experiments. In a next step, we tested MPB on a larger set of experimental and *in situ* samples of Arctic spring bloom phytoplankton over two consecutive years in order to assess whether the method can indeed help us to resolve intraspecific shifts in field populations at full diversity.

Materials and methods

To develop the MPB method and to assess its resolution of allelic composition in comparison to conventional microsatellite analyses, four sets of samples were used (Fig. 1): (i) 'community experiment' samples of Arctic phytoplankton obtained by Hoppe *et al.* (2018b) from incubations under *present day*, *future* and *acidification* selection regimes (temperature and pCO_2 perturbations, see next paragraph for details); (ii) single-strain 'monocultures' (=single genotypes) obtained by Wolf *et al*. (2019) from the same experimental communities; (iii) 'multi-strain' cultures consisting of six genotypes with a known composition, as assessed by allele-specific qPCR (Wolf *et al*., 2019). To test the application of MPB in further contexts at natural diversity, we analysed additionally (iv) 'natural community' samples of Arctic diatom populations acquired in spring seasons 2016 and 2017 in an Arctic fjord.

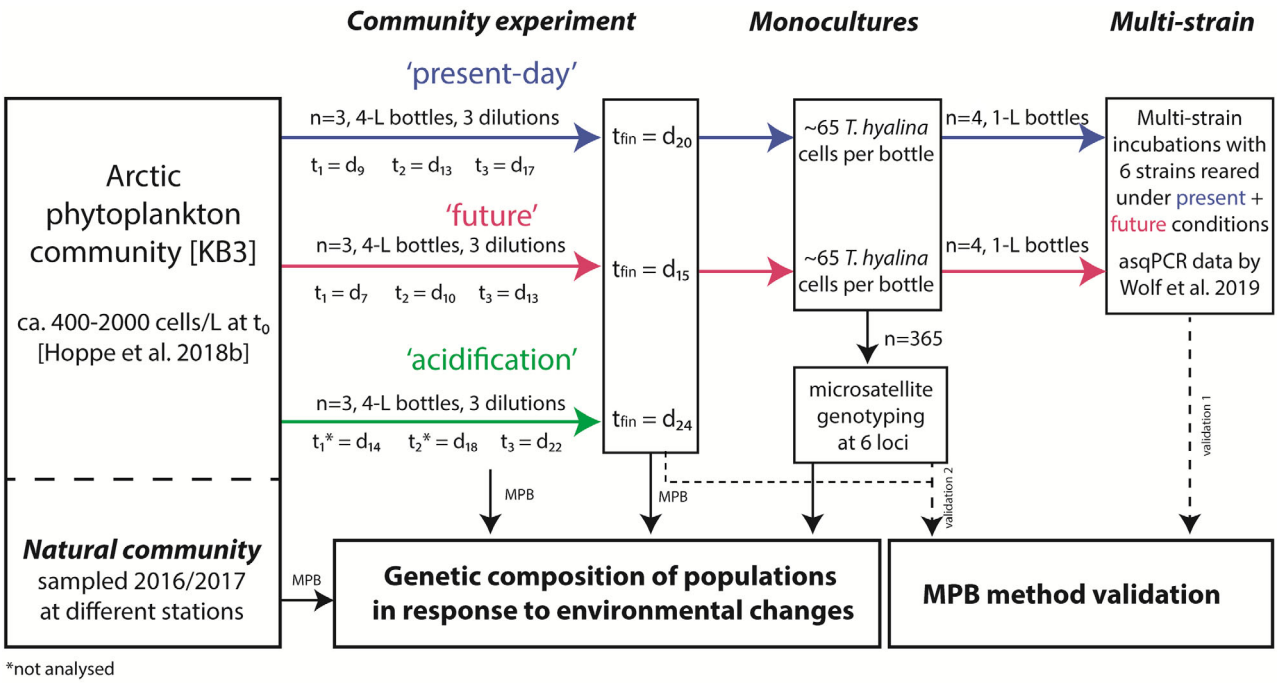

Fig. 1. Overview of the four data sets (bold and italic) generated in this study to validate the MPB method (dashed black arrows) and to assess changes in genetic composition of *T. hyalina* populations in response to temperature and CO_2 treatments (solid black arrows). Samples from 'Natural communities' and the 'Community experiment' confirmed the applicability of MPB within whole communities. Treatment conditions for the experimental incubations were: 'present-day' (blue): 2°C 400 µatm pCO_2, 'future' (red): 7°C 1000 µatm pCO_2, 'acidification' (green): 2°C 1000 µatm pCO_2.

Community experiment samples

Details of the community incubation experiment using Arctic phytoplankton assemblages are described in the study by Hoppe *et al*. (2018b), where incubations are referred to as KFb_1 and KFb_2. In brief, the experiment was initiated in mid-April 2016 with a natural Arctic phytoplankton spring community from the Kongsfjord (Svalbard, 78°55'N, 11°56'E) by gently pumping seawater from a depth of 23–25 m into 4-L polycarbonate bottles. Bottles were incubated in triplicates in temperature-controlled chambers and aerated via continuous bubbling of air with the target partial pressures of CO_2 (pCO_2) for each treatment: *present-day* (1.8 ± 0.1°C and 324 ± 12 µatm pCO_2), *acidification* (1.8 ± 0.1°C and 966 ± 50 µatm pCO_2) and *future* conditions (6.8 ± 0.4°C and 1078 ± 16 µatm pCO_2). Cultures were exposed to continuous light at 50 ± 2 µmol photons m^{-2} s^{-1} under non-limiting nutrient concentrations. To avoid pH drift and to increase the effect of genotype sorting, three dilutions were performed at a ratio of 1:25. Depending on the growth rate of the respective community, the experiments ran for 16–22 days, equivalent to 27–31 generations according to bulk community growth rates of 1.3–1.7 divisions per day (k, estimated by nutrient drawdown, data not shown). Filter samples were collected during each dilution (*t*1–*t*3) and the final time-point of the community experiment (*tfin*) for *present day* and *future* treatments. For the *acidification* treatment, only filter samples from the last dilution (*t*3) and the final time-point (*tfin*) were analysed. All samples for MPB were mixed thoroughly and filtered (filtration volume 300–500 ml) on 10 µm PC filters (Whatman Nucleopore), which were stored at −80°C until further analysis.

Monocultures

At the final time-point of the community experiment, 55–65 single cells of the diatom *T. hyalina* (Grunow; Gran, 1897) were isolated from each triplicate bottle of the *present day* and *future* treatments (yielding 365 strains in total). Cells were picked manually under a light microscope and washed three times in sterile seawater. Single-cell isolation of each strain was repeated after 10–14 days of growth in 48-well plates at 6.8°C and 50 µmol photons m^{-2} s^{-1} in sterile nutrient-enriched seawater. Each of the resulting strains was checked microscopically for contamination with other algal species before being grown as 250 ml monocultures at 3°C and 5–10 µmol photons m^{-2} s^{-1}. Samples were taken as described above and used for traditional microsatellite genotyping as well as for MPB validation (see validation 2).

Multi-strain samples

Wolf *et al*. (2019) performed a multi-strain experiment in 2017, based on six of the above described monocultures. They chose three clones selected under the *present-day* and three under the *future* scenario and incubated them together under both conditions ($n = 4$ replicates each). For the current study, samples were taken at two time-points (i.e. dilutions, *t*1 and *t*2) and at the final time-point (*tfin*). Strain composition was originally assessed using allele-specific qPCR by Wolf *et al*. (2019). These data were used for MPB validation on the same samples (see validation 1).

Natural community

Field samples of natural populations were obtained from 2016 and 2017 spring blooms in the Kongsfjord (Svalbard). These samples were collected regularly (every 2–6 days) at the mid-fjord station KB3 and occasionally at neighbouring stations in the fjord (KB2 and KB5, see Supporting Information S3 + S4). Sampling took place between 29 April and 16 May 2016 as well as between 18 April and 26 May, 2017. The filtration volume consisted of 1–2 l seawater, and samples were treated as described above.

Microsatellite genotyping of monocultures

The detailed methods are described in the Supporting information of Wolf *et al*. (2019). In brief, DNA was extracted using the NucleoSpin Plant II Kit (Macherey-Nagel GmbH, Duren, Germany). Six microsatellite primers (Supporting information S1) were applied to DNA samples in equimolar concentrations with fluorescent markers (FAM, HEX, and AT) as single or multiplex PCR, using the Type-it Microsatellite PCR Kit (Qiagen, Hilden, Germany) and a thermal cycler (Mastercycler Nexus Gradient; Eppendorf, Hamburg, Germany) under the following reaction conditions: 5 min at 94°C, 30 cycles of 30 s at 94°C, 90 s at 57°C, and 40 s at 72°C, final elongation step at 72°C for 10 min. Fragment analysis was performed by capillary electrophoresis using the 3130xl Genetic Analyzer (Applied Biosystems, Foster City, CA, USA), and sizes were assigned relative to a GeneScan ROX standard (Thermo Fisher, Waltham, MA, USA). Microsatellite alleles were scored using Genemapper (version 4; Applied Biosystems). From samples measured repeatedly throughout all runs, an error rate in allele length assignment of 2% was calculated. Arlequin (version 3.5.2.2; Excoffier and Lischer, 2010) was used to calculate expected and observed heterozygosity (H_e and H_o), to assess deviations from Hardy–Weinberg equilibrium within each experimental bottle, and to determine population differentiation (F_{ST}) among pairs of bottles. Population structure was also tested by a Bayesian clustering analysis using STRUCTURE (version 2.3.4; Pritchard *et al*., 2000) without a population prior and STRUCTURE HARVESTER (version 0.6.94) (Earl and vonHoldt, 2012) in combination with CLUMPP (version 1.1.2) (Jakobsson and Rosenberg, 2007) to determine the optimal number of clusters, following the Evanno method (Evanno *et al*., 2005). Linkage disequilibrium (LD) was calculated using LIAN (Haubold and Hudson, 2000), and significance was assessed after sequential Bonferroni correction. The probability (p_{sex}) of encountering the same genotype twice in the same population (i.e. bottle) was estimated using the programme GenClone2 (Arnaud-Haond and Belkhir, 2007).

Microsatellite PoolSeq barcoding analysis

In order to provide an overview over the entire procedure, a workflow of the steps from laboratory to bioinformatic analysis is shown in Fig. 2. Two primer sets, ThKF3 and ThKF7, were selected for MPB based on their relatively high allelic richness and small fragment size in the prior fragment analysis of 365 monocultures (ThKF3: 180–270 bp, 24 alleles; ThKF7: 200–300, 14 alleles; Wolf *et al*. (2019)). In a first PCR, triplicates of all MPB samples (i.e. community experiment, multi-strain, and natural community samples) were run per locus using 10 ng of sample DNA (settings as described above). Triplicate PCR products were pooled and visualized on a 1.5% agarose gel. Bands at the approximate size range of microsatellite sequences (150–300 and 150–350 bp for ThKF3 and ThKF7, respectively) were excised and purified using the NucleoSpin Gel and PCR Clean-up Kit (Macherey-Nagel, Düren, Germany). In a second index PCR, dual indices and adapters were attached (Index PCR, Nextera XT Index Kit; Illumina, San Diego, CA, USA) and samples were purified using AMPure XP Beads (Beckman Coulter, Brea, CA, USA). Libraries were validated using the Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). The final DNA libraries were pooled at equimolar ratios and sequenced using the MiSeq System with the v3 2×300 bp Paired-end Kit (Illumina).

Demultiplexing and FASTQ sequence generation were performed using the 'Generate FASTQ' workflow of MiSeq Reporter. About 11 million raw amplicons for primer set ThKF3 and about 12 million raw amplicons for primer set ThKF7 were obtained in all MPB samples. Amplicon contingency tables were constructed for each primer set using an in-house but modified metabarcoding analysis pipeline (available upon request). Trimmomatic (version 0.38; Bolger *et al*., 2014) was used to crop reads to a length of 275 bp and to truncate the pre-trimmed

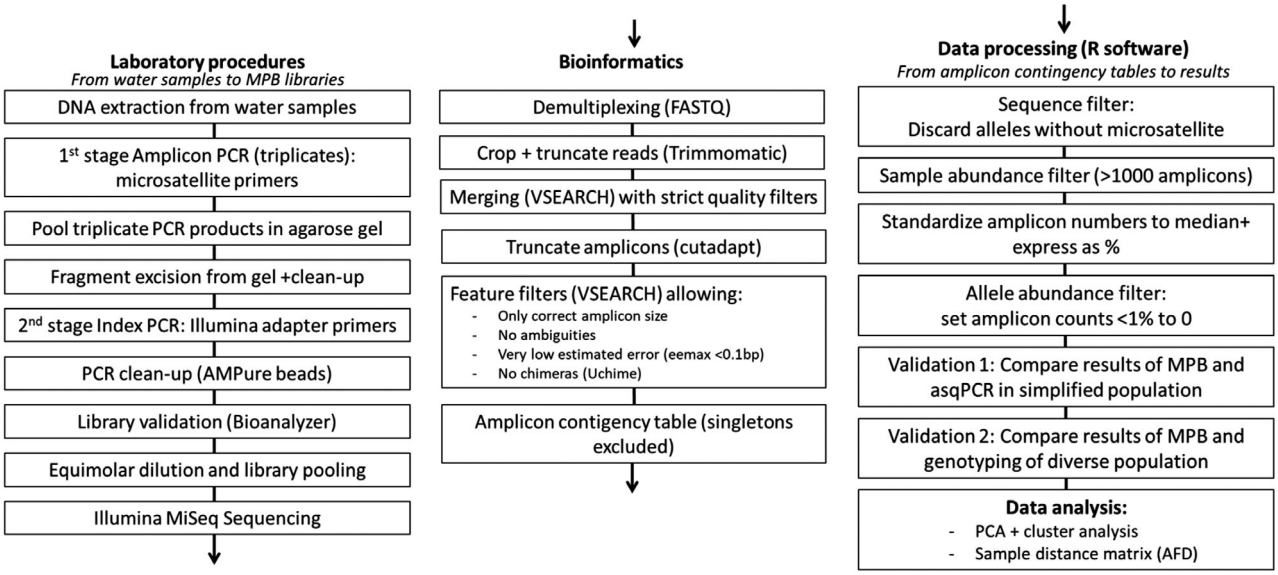

Fig. 2. Workflow of sample preparation and bioinformatic analysis for MPB.

reads at the base position at which an averaged Q-score in a sliding-window of length 3 dropped below 8 (scanned from the 5′-end to the 3′-end). The paired-end reads were merged using VSEARCH (version 2.3.0; Rognes *et al*., 2016), allowing a maximum of five mismatched bases while requiring a minimum overlap length of 50 bp. Sequences that could not be merged unambiguously were discarded. As most amplicons were expected to be shorter than the remaining read length, VSEARCH parameters were adjusted to allow the merging of staggered reads and the removal of possibly existing non-overlapping artificial amplicon segments outside the target sequence. Target-flanking sequence segments with 100% forward and reverse primer matches were truncated from the amplicons using cutadapt (version 1.9; Martin, 2011) and amplicons were only kept in the sequence pool if segments for both the forward and the reverse primers were found during truncation. The remaining sequences were further selected by applying a feature filter (VSEARCH; version 2.3.0). Sequences were discarded if (i) they were longer than 320 bp or shorter than 120 bp; (ii) they carried any base ambiguity; or (iii) the expected number of miscalled bases per sequence (i.e. the sum of all base error probabilities = eemax) exceeded 0.1. Each sample was independently checked for chimeric sequences using VSEARCH (version 2.3.0) and the UCHIME (Edgar *et al*., 2011) algorithm in de-novo mode, and predicted chimeras (as identified within non-repetitive regions of the sequences) were removed from the sample files.

All sequences were pooled and amplicon contingency tables were created for both primer sets using a script adapted from https://github.com/torognes/swarm/wiki/Working-with-several-samples (retrieved: December 2018). Alleles represented by only one amplicon (i.e. singletons) were excluded. For the primer-set ThKF3, about 6 million reads, and for the primer-set ThKF7, about 7.5 million reads passed all filtering procedures. Strict quality filters and short amplicon lengths guaranteed the reliability of the sequence information. Only one sample for the primer ThKF3 from the field bloom in 2016 (KB3_t12) did not pass the quality filtering (requiring >1000 sequences to remain after all sequence filtering steps) and was excluded. For the primer ThKF7, some samples did not yield sufficient PCR products and were excluded (all samples of the experimental *acidification* incubation and eight field bloom samples, seven of them from 2017).

The amplicon tables for both primer sets were further processed using the R software (version 3.5.1). Sample libraries were normalized to the median of total read numbers in all samples. Allele frequencies were expressed as percentages of the total reads per sample, and different alleles with the same lengths were assigned unique names (e.g. 214.1 and 214.2). A difference in sequence but not in length is caused by single-nucleotide polymorphisms (SNPs), which can be present within the flanking as well as the repeat regions of microsatellites (Viruel *et al*., 2018). Alleles not containing at least two repeats of the target microsatellite sequence were discarded (‘sequence filter’), which resulted in only minor loss of reads (less than 1% for both loci). To exclude sequencing errors, alleles with frequencies <1% in the whole sample were set to 0 (‘allele abundance filter’). This resulted in an average loss of 7.5% (ThKF7) to 10% (ThKF3) of all reads in samples for validation

1 (multi-strain) and 18% (ThKF7) to 23% (ThKF3) of all reads in samples for validation 2 (diverse community experiment), but substantially reduced the error probability. All remaining amplicon sequences were checked for the expected microsatellite repeat structure and compared with those of the fragment analysis from the same experimental incubations in two validation steps.

Validation of MPB-derived allele frequencies

Within this study, MPB produced pooled allele frequency tables without knowledge about specimen number (no or only approximated cell counts for most samples exist) and identity of alleles, which is why most commonly evaluated parameters in population genetics cannot be directly compared. However, the relative allele frequencies provide a straightforward measure for population differentiation. Allele frequencies obtained by the MPB method for both primers were validated using two sample types. In validation 1, the allele frequencies of multi-strain samples previously analysed by allele-specific qPCR (Wolf *et al*., 2019), and therefore containing a known composition of genotypes, were compared with those of the same multi-strain samples analysed by MPB. Since fragment analyses only yield information on allele length, all alleles in the MPB analysis that had the same number of base pairs (even those containing point mutations) were summed for this comparison. The allele frequencies (%) of each allele in all samples from the fragment analysis and the MPB analysis were compared by linear regression to obtain Pearson's *r* and significance levels and by calculating the mean deviation (i.e. error) between analysis types for each allele length. For additional comparison, principal component analyses (PCAs) of both datasets were performed and visualized in ordination plots. In validation 2, the same comparison based on allele lengths was performed for samples from the genotyped *present-day* and *future* community incubations. In this case, however, the allele frequencies of the fragment analysis contained only a limited subsample of genotypes from a highly diverse population, which introduces stochastic effects.

Data analysis of MPB datasets

The final MPB results of both primer sets were evaluated using principal component analyses (PCA) in the software R for each set of population samples, using different allele versions (including point mutations) and their relative frequency as variables. The results were visualized in ordination plots for the first two principal component axes. Additional PCAs were performed for the same samples based on a merged dataset of alleles and their frequency of both loci (see Supporting Information Fig. S5 + S6). Here, samples yielding amplicons for only one primer were omitted.

Pairwise comparison of samples was obtained by allele frequency differentiation (AFD) calculated following Berner (2019) after normalizing MPB allele frequencies to 100% for each sample. These AFD-values served as an alternative to pairwise F_{ST} values applicable for the data structure of MPB frequency tables. In parallel to F_{ST}, AFD values range from 0 to 1, with 0 describing identical samples.

Samples were clustered using Laplacian spectral clustering (Nadler *et al*. 2005), on the basis of similarities computed as the inverse of AFD values and unnormalized Laplacian matrices. Clusters were defined up to the largest normalized spectral gap, which was considered valid above a value of 0.1, and is referred to in the text below. The greater the normalized spectral gap between clusters, the more clearly they are separated. Additionally, the obtained clusters were statistically tested by permutational multivariate analysis of variance (PERMANOVA) analysis (R-package RVAideMemoire, v0.9-5) on the basis of the allele frequency tables.

Population genetic simulation of community experiment data without selection

To generate a null model of allele frequency changes over time under the community experiment settings assuming genetic drift without selection, genetic data were simulated using Nemo 2.3 (Guillaume and Rougemont, 2006). Simulation settings, including allele numbers, initial frequencies, fecundity, bottlenecks, and final sampling dates, were chosen to best match the real experimental settings (see Supporting Information S8 for details). After the simulation, allele frequencies were calculated for each population and a PCA similar to that described above was performed. To evaluate whether the PCA results from the experiments differ significantly from the simulated data and thus whether genetic drift could be the sole driver of differentiation, a Mann–Whitney U-test was used to compare PC1 values between treatment groups (*present day*, *acidification*, and *future*) of simulated values and experimental data.

Results and discussion

Validation of the MPB method

Since the use of microsatellite markers in pooled analyses of complex environmental samples is a novel approach, we tested the reliability of MPB in samples with different degrees of diversity. The first validation step (validation 1) using samples from artificial populations consisting of six strains of *T. hyalina* (multi-strain)

supported the reliability of the method (Fig. 3A + B, Table 1(a)): The same alleles were found and the correlation between allele frequencies from MPB, and allele-specific qPCR (asqPCR) was very compelling and strongly significant for both microsatellite loci in samples from the multi-strain experiment (Fig. 3A + B; Pearson's $r > 0.99$; $p < 0.001$). This demonstrates that allele identification and quantification were highly accurate. The mean allele frequency error was 0.95 $\pm$ 0.8% for alleles of ThKF3 and 1.3 $\pm$ 0.4% for ThKF7. The error size did not depend on allele length (see Table 1(a)). Furthermore, when visualized in a PCA, the same samples of multi-strain samples measured by asqPCR and MPB were congruent with each other and with the temporal development previously observed (Fig. 4). For both methods, the final composition of populations showed significant differences in line with the two treatments (PERMANOVA for both loci by MPB $p = 0.001$, by asqPCR $p = 0.03$) and the temporal population development described in the study by Wolf *et al*. (2019). We therefore conclude that MPB is a promising approach for the semi-quantitative and qualitative evaluation of allele frequencies in a low-diversity setup, comparable to allele-specific qPCR (Meyer *et al*., 2006; John *et al*., 2015; Minter *et al*., 2015; Wolf *et al*., 2019).

In the second validation step (validation 2), the MPB results for the community experiment were compared with allele frequencies obtained with the same two primers by classical genotyping of monocultures from single-cell isolates of the same populations (*tfin*; see Fig. 1). This set of samples contained a much higher inter- and intra-specific diversity, which comprises a higher number of possible error sources for the PCR and sequencing procedure. These potential errors might include masking effects by environmental DNA fragments, unknown primer sites in the environmental DNA, but also potential stutter effects from the PCR, depending on the microsatellite itself, as well as a size bias during sequencing (Illumina sequencing having a bias towards smaller fragments). The potential impact of stuttering was resolved during the validation 2 procedure, since genotyping results free of stuttered reads were used for validation of MPB results. In this validation step, the MPB method showed a larger deviation from, but still a high degree of concordance in allele frequencies with the genotyping results (Fig. 3C and D: Pearson's $r = 0.87$;

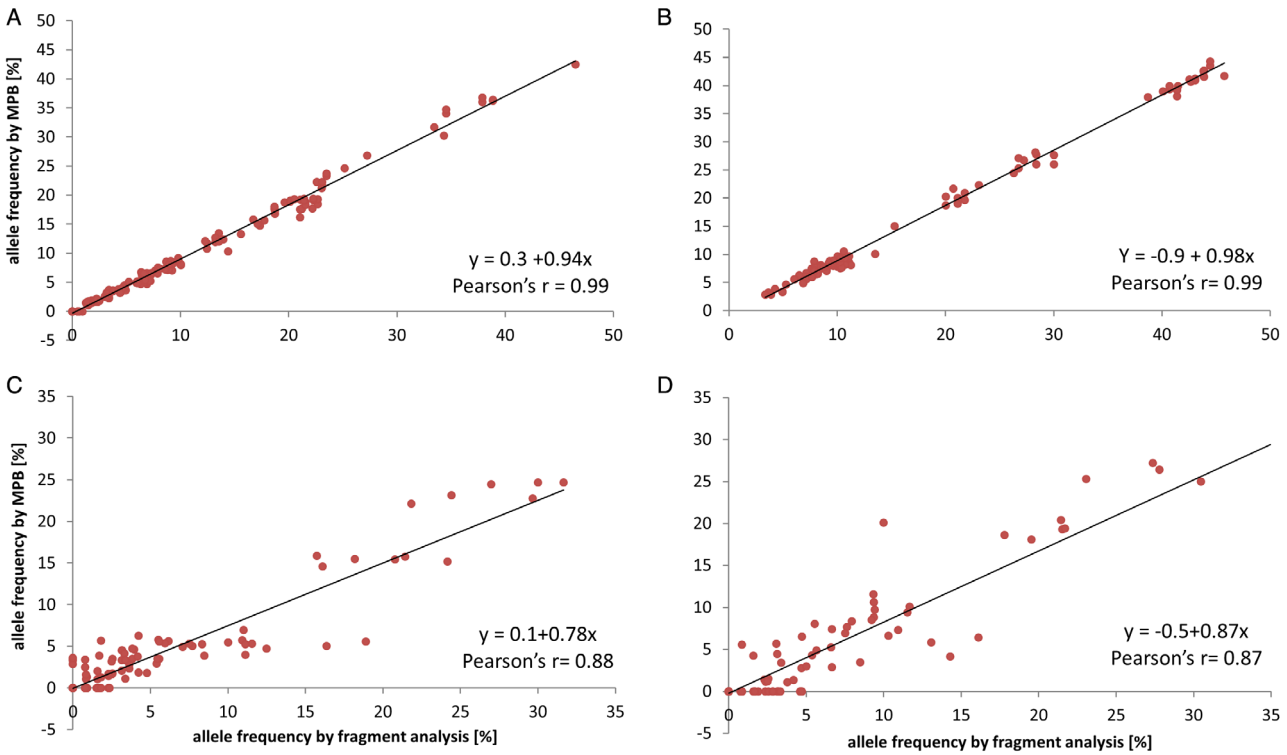

Fig. 3. Validation of relative allele frequencies for each allele as measured by fragment analysis (x-axis) and by MPB (y-axis) using the primers ThKF3 (A and C) and ThKF7 (B and D). Pearson's correlation coefficient (r) was used to evaluate the linear regression (black line = best fit). Validation 1 for primers ThKF3 (A) and ThKF7 (B): allele frequencies as measured by MPB and allele-specific qPCR in 18 multi-strain samples from artificial populations with limited diversity (six strains of *T. hyalina*). Table 1a shows the detected error for each allele separately. Validation 2 for primers ThKF3 (C) and ThKF7 (D): allele frequencies of each allele as measured by MPB and genotyping in six community samples of phytoplankton assemblages including natural *T. hyalina* populations at the final time-point of the community experiment. Genotype analyses were based on 55–65 strains per sample. The error therefore includes inaccuracies of MPB as well as the subsampling bias of the genotyping.

Table 1. (a) Mean error in the relative frequency as well as mean relative abundance of each allele (defined by length) per sample at ThKF3 and ThKF7 using MPB, as estimated in validation 1 with multi-strain samples of known composition (asqPCR). (b) Allelic richness of alleles for primers ThKF3 and ThKF7 for each subgroup of data as estimated by MPB analysis.

(a)	Mean error (%)	Mean allele abundance (%)
ThKF3		
205	0.29	1.05
208	2.65	12.73
211	0.54	4.02
214	1.38	30.85
217	0.12	1.70
220	1.26	9.07
223	0.89	12.79
226	1.79	17.40
229	0.32	5.42
232	0.35	0.88
235	0.91	4.07
ThKF7		
233	1.89	10.37
239	1.71	42.37
245	0.91	9.37
251	1.32	24.08
257	0.9	8.11
269	1.03	5.69

(b)	# total alleles	# unique length	# homoplastic	% homoplastic
ThKF3				
Exp16	24	13	11	46
Exp16 fragm	24	24	-	-
Bloom 16	27	14	13	48
Bloom 17	23	12	11	48
Bloom 16 + 17	26	13	13	50
Total	27	14	13	48
ThKF7				
Exp16	20	11	9	45
Exp16 fragm	14	14	-	-
Bloom 16	28	16	12	43
Bloom 17	42	18	24	57
Bloom 16 + 17	45	18	27	60
Total	47	18	29	62

#total alleles: total number of detected alleles; #unique alleles: number of alleles of differing length; #homoplastic: the number of homoplastic alleles (same length, different sequence), % homoplastic: percentage of homoplastic alleles with the total number. Subgroups: Exp16 = community incubation 2016; Exp16 fragm = alleles identified by fragment analysis; Bloom = natural spring community samples.

$p < 0.001$ for both primers). The mean frequency error per allele was 1.6 $\pm$ 1.1% for ThKF3 and 2.3 $\pm$ 1.4% for ThKF7. The intensity of stutters per allele had been quantified during genotyping as 5%–25% in ThKF3 and close to none in ThKF7 (Wolf *et al*., 2019). As this was not reflected in the mean error of the two loci, stuttered reads may not have been the main source of error. It may, however, partly explain the regression slope below 1 in validation 2, as smaller alleles may be slightly overrepresented due to the inclusion of stutter artefacts. Overall, we consider these discrepancies in validation 2 very small, given the potential technical PCR and sequencing errors of the MPB approach (see above), as well as subsampling biases for classical genotyping (analysis of only ~65 isolates out of >500 000 cells per bottle). When defined by allele length, the allelic richness was generally lower in MPB results compared to the fragment analyses of separate genotypes (Table 1(b)). These 'lost' alleles were very low in frequency (<5%) and did probably not appear in the MPB results because of the statistical effects of overrepresentation in the subsampled genotypes combined with the strict abundance filters applied to MPB amplicon tables (allele frequencies below 1% were omitted). Because of their low frequencies, this discrepancy did not have a large effect on overall results as described by the low error rates.

A clear advantage of MPB over classical fragment analysis-based methods is the resolution of otherwise cryptic point mutations, i.e., alleles that are identical by size but not in base-pair sequence (Estoup *et al*., 2002). Sequence variants of the same size in microsatellites and their flanking regions can have important implications for the type of mutation models used (Kosman and Jokela, 2019) and can lead to misinterpretation of traditional population data based only on length evaluation, as has been shown in several publications and various organisms (Germain-Aubrey *et al*., 2016; Vartia *et al*., 2016). A large number of such homoplastic alleles were also detected in our samples (e.g. several alleles of 214 bp; see arrows in Figs. 5 and 6). It needs to be noted, however, that homoplastic alleles could not be validated in the same way as length-variant alleles, since variants of the same length are not resolved in traditional fragment analysis. This being said, 48% (ThKF3) and 62% (ThKF7) of the allelic diversity detected by MPB consisted of such homoplastic alleles (Table 1(b)). Although stuttering may slightly inflate this percentage, such values are in line with estimates of homoplasy in other plant species analysed by established methods (Šarhanová *et al*., 2018; Viruel *et al*., 2018). This additional diversity can have considerable implications, as illustrated, for example, in the field populations of 2017. Including homoplastic alleles, locus ThKF7 had a much larger allelic richness in 2017 than in populations of the previous year (Table 1(b)), and a much higher number of different alleles than would be estimated by analyses based on allele length only. As discussed below, this also allowed a differentiation among samples of 2017, which was not resolved with the other primer-set (ThKF3). When the datasets of both loci were combined and analysed together; however, a large fraction of the differentiation between samples from 2016 was lost (see

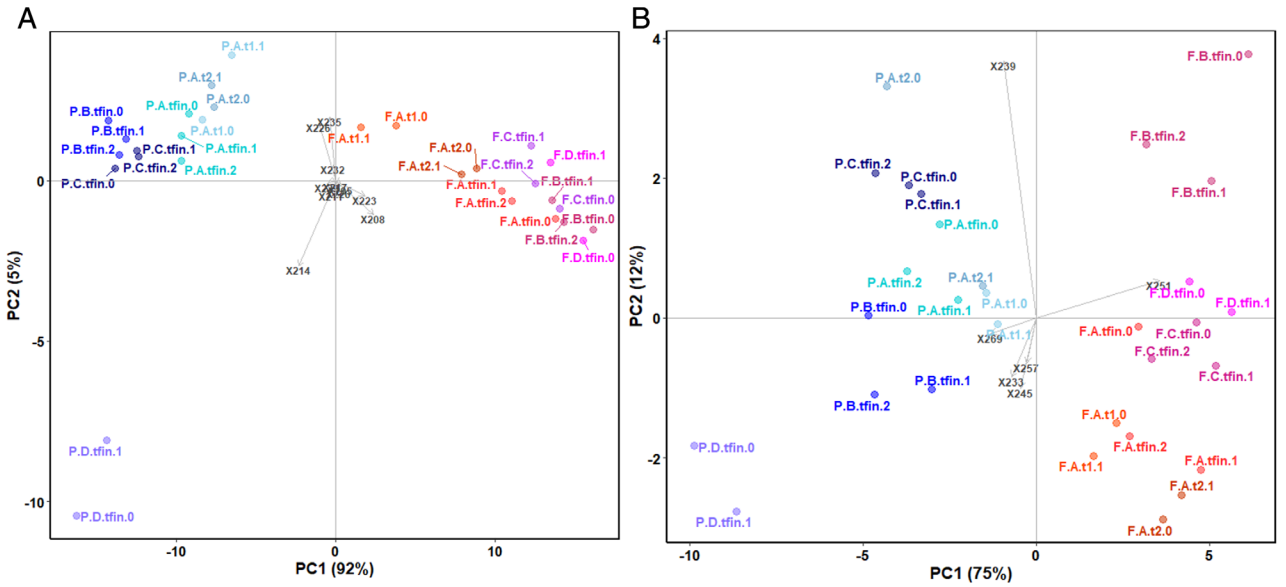

Fig. 4. PCA of allele frequencies in multi-strain experiment samples in validation 1 as measured by MPB and allele-specific qPCR (asqPCA) of microsatellite loci ThKF3 (A) and ThKF7 (B). Samples of the final sampling time-point of all four replicate bottles (A–D) of the present-day treatment ('P') are in shown in blue tones, those of the future treatment ('F') in red tones. The method by which the result was obtained is marked by the suffix of each label (0 = asqPCR; 1 or 2 = MPB). For bottles A in both treatments, the time-points t1 and t2 are shown as well. The present-day replicate bottle D appears as an outlier because it was initiated with only five instead of six genotypes. Final population composition was significantly different (PERMANOVA) between the two treatments when measured at both loci with asqPCR (p = 0.03) and MPB (p = 0.001).

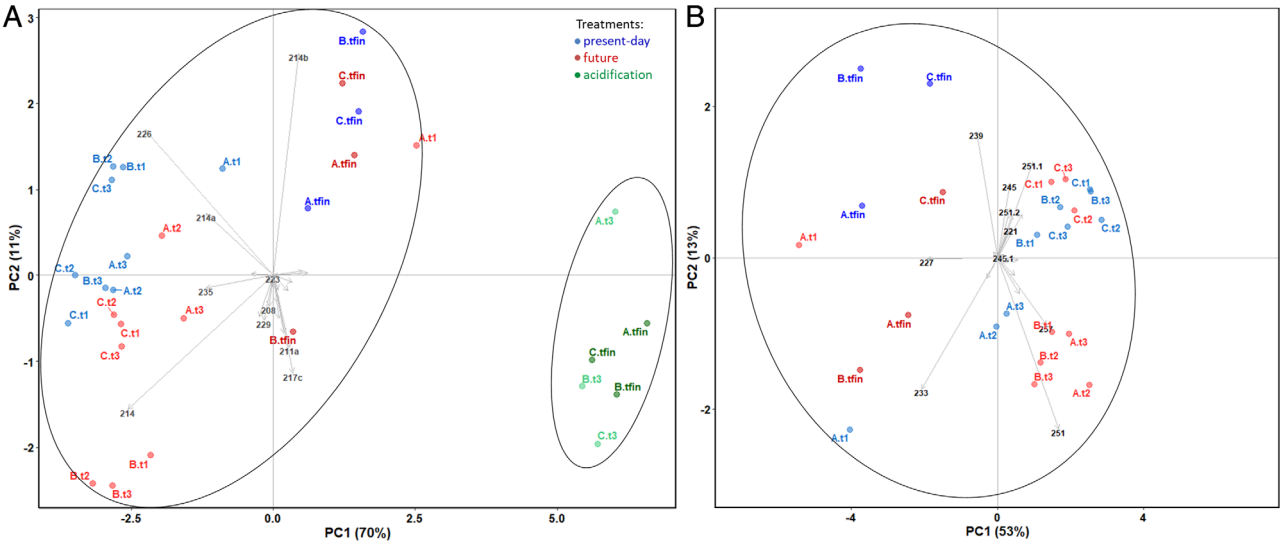

Fig. 5. PCA of relative allele frequency data for *T. hyalina* obtained from the community experiment and analysed by MPB using the primer ThKF3 (A) and ThKF7 (B). Three replicate incubations exposed to the present-day (P: 2°C and 400 µatm pCO_2; blue) and future treatments (F: 7°C and 1000 µatm pCO_2; red) were sampled at four time-points (t1, t2, t3 (light colour) and tfin (dark colour)). The three replicate incubations for the acidification treatment (A: 2°C and 1000 µatm pCO_2; green) were analysed only at time-points *t*3 and *tf*in and did not produce results with primer-set ThKF7. Black numbers identify alleles (by their lengths), which were used as variables for PCA. Only the 20 most abundant alleles are identified here. Different allele versions of the same lengths (homoplasy) have unique names (e.g. 214 and 214a). Ellipses mark sample clusters as identified by spectral clustering.

Supporting Information S6). This could be attributed to the fact that the most abundant (but largely undifferentiated) alleles of locus ThKF7 partly mask the resolution provided by ThKF3 for the samples from 2016. Thus, unlike in traditional microsatellite genotyping, the information from several loci cannot be necessarily merged in MPB and is better evaluated separately before being discussed in concert.

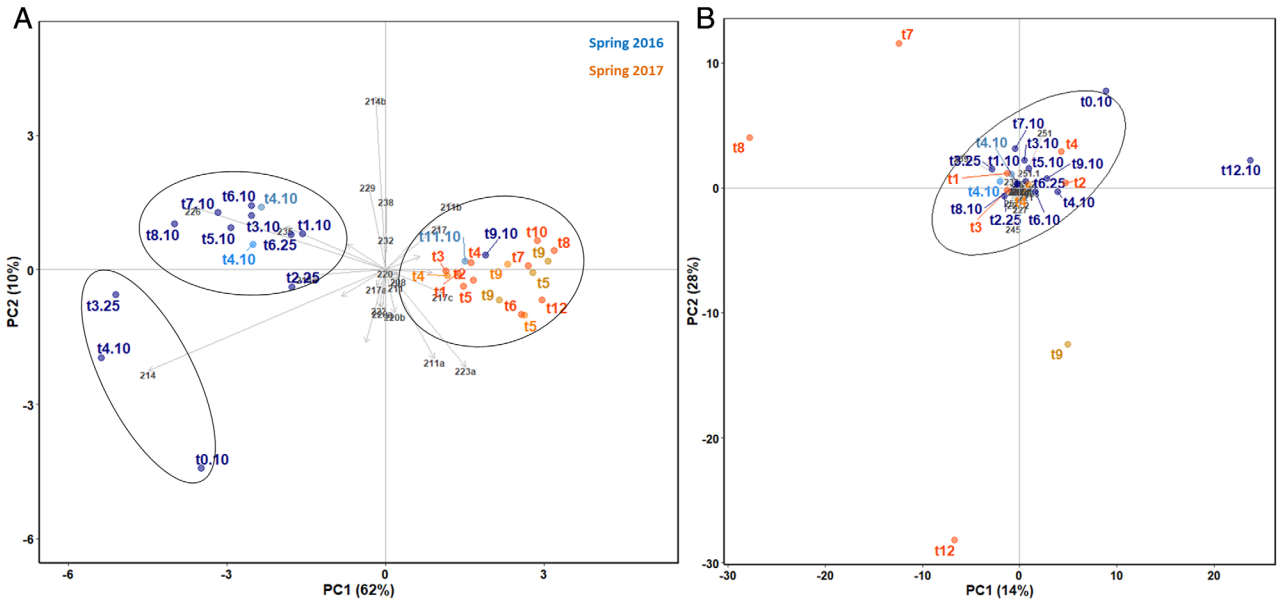

Fig. 6. PCA of relative allele frequency data for *T. hyalina* obtained by MPB using the primer ThKF3 (A) and ThKF7 (B) on natural community samples from the Kongsfjord in spring 2016 (blue) and 2017 (orange). Samples for 2016 were taken at 10 m or 25 m (denoted as 0.10 and 0.25) between 29 April and 16 May 2016; samples for 2017 were all obtained at 25 m between 18 April and 26 May 2017. The sampling stations throughout the Kongsfjord are marked by colour variation [KB3 (dark blue/orange); KB2 and KB5 (light blue/orange)]. The exact date and location for each field sample are provided in the Supporting Information S3. Black numbers identify alleles (by their lengths), which were used as variables for PCA. Only the 20 most abundant alleles are identified here. Different allele versions of the same lengths (homoplasy) have unique names (e.g. 214 and 214a). Ellipses mark sample clusters as identified by spectral clustering.

The congruence of results obtained by the two methods and the accurate technical validation with multi-strain and diverse samples demonstrate that MPB is a versatile method. When compared to established population genetic methods, it cannot provide individual genotype characteristics (especially heterozygosity) but is capable of correctly estimating allele frequencies and thus differences in populations. Next to traditional genotyping, allele frequencies assessed by MPB probably deviate most because of two aspects: they contain a larger error concerning absolute allele numbers and frequencies due to the inclusion of stuttered reads, but they contain less error in the relative allele frequencies since subsampling bias of samples is much smaller. It is therefore more meaningful to interpret MPB results relative to each other than in absolute terms. Since MPB is based on the entire range of genotypes present in a naturally diverse sample and takes homoplastic alleles into account, it potentially provides a higher resolution than fragment analysis of tediously isolated subsamples.

Application of MPB to diverse environmental samples

We applied MPB to resolve intraspecific patterns in *T. hyalina* populations in natural, highly diverse communities to an extent that would not have been feasible using traditional methodology. The pooled microsatellite analysis by MPB allows relative intercomparisons of a high number of populations without genotyping individual isolates. By removing the bias of subsampling from large and highly diverse populations and by adding temporal resolution to the measured allele frequencies, we gained valuable information about selection in Arctic diatom populations in the laboratory and under natural field conditions.

Experimental community incubations. Investigations of single strains of *T. hyalina* isolated from community incubations as well as a simplified multi-strain experiment revealed large intraspecific phenotypic differences and that strong selection among diatom genotypes can take place on very short timescales (Wolf *et al*., 2019). Resolving such dynamics in populations with high or natural diversity, however, remains challenging. We here show that MBP is an innovative approach to fill this gap. In our dataset of natural communities exposed to either *present day* or *future* conditions (i.e. warming plus ocean acidification), both classical microsatellite genotyping of single isolates and the novel MPB approach suggested no directional changes or apparent differences between populations (Fig. 5A + B, Supporting Information S7). MPB-based analysis of a third treatment (i.e. *acidification* without warming), which was not genotyped due to logistical and resource limitations, allowed us to reveal that a

Table 2. Summary statistics for microsatellite genotyping using six microsatellite markers for six incubation bottles in the community incubation experiment. Detailed microsatellite characteristics are provided in the Supporting Information Table S1.

	Run time (days)	# approx. generations	# genotyped	# unique MLG (% diversity)	H_O	H_E	F_{ST} and LD	Present day A	Present day B	Present day C	Future A	Future B	Future C
Present day A	20	27	64	64 (100%)	0.642	0.667	Present day A	NA	-	-	-	-	*
Present day B	20	27	65	63 (96.9%)	0.623	0.669	Present day B	0.00193	NA	-	-	-	-
Present day C	20	27	63	57 (90.5%)	0.633	0.653	Present day C	–0.00188	0.00131	NA	-	-	-
Future A	15	31	55	53 (96.4%)	0.645	0.682	Future A	0.00373	–0.00077	0.00259	NA	-	-
Future B	15	31	58	56 (96.6%)	0.601	0.633	Future B	0.00221	0.00192	–0.00003	0.00578	NA	-
Future C	15	31	60	60 (100%)	0.663	0.681	Future C	0.00677*	0.00255	0.00189	0.00205	0.00129	NA
Present day strains	**20**	**27**	**192**	**184 (95.8%)**	**0.632**	**0.663**							
Future strains	**15**	**31**	**173**	**169 (97.6%)**	**0.636**	**0.665**							
Total			**365**	**351 (96.2%)**	**0.634**	**0.664**							

For each bottle are given: run time of the incubation experiment, approximate number of generations as calculated from assemblage growth rate, number of genotyped individuals, number of multilocus genotypes (MLG) and genotypic diversity in %, observed (H_O) and expected (H_E) heterozygosity. Sum or average values are marked in bold. The right part of the table shows pairwise F_{ST} values (lower diagonal) and significant linkage disequilibrium (LD; upper diagonal) at a significance level (*$p < 0.05$). None of the differences in heterozygosity nor LD within bottles were significant.

directional allelic shift had taken place under those conditions.

For genotyping, a total of 365 strains were isolated from six incubation bottles of the final time-point under *present* and *future* conditions (*tfin*; Fig. 1; Table 2). Among all analysed single strains using six microsatellite loci, only seven multi-locus genotypes were found twice, indicating an overall genotypic diversity of 96%, which is in line with most studies based on single cell isolation (Godhe and Rynearson, 2017). Within each bottle from the community experiment, genotypic diversity varied between 90% and 100%. Among the few identical genotypes, all but one pair originated from the same replicate bottle, potentially indicating weak selective processes. This was supported by the low probability of encountering the same genotype twice within a bottle ($p_{sex} < 0.001$ in all cases, analysed via GenClone2). Still, in view of the maintained diversity, strong clonal dominance within the bottles is unlikely. F_{ST} values for pairs of populations were generally low (Table 2), and no population subdivision was detected by a Bayesian clustering analysis (best cluster estimate judging from Δ(K) was K = 2, within which samples from all subpopulations/bottles were evenly distributed, see Supporting Information S7), suggesting no substantial differentiation between populations in these two treatments at the final time-point.

Absolute AFD-values based on genotyping data were generally higher (i.e. less similar) than those based on MPB data (Supporting Information Table S2). This makes sense in view of distorting effects caused by subsampling of a small number of analysed genotypes (~60 per bottle) compared to several thousand genotypes potentially included in each MPB sample: If two incidents of a rare allele are found during genotyping, it will be assumed to represent 1:60 of the population (i.e. 1.6%), while two incidents out of 50.000 MPB amplicons will only amount to 1:50.000 (i.e. 0.002%) and populations will still appear more similar. This assumption was supported by a simulation of 60 subsamples per population based on MPB data, which then yielded similar absolute values as the genotyping (data not shown).

The MPB method was successfully applied to the same experimental communities for two loci (ThKF3 and 7). For the *present day* and the *future* treatment, both primer pairs alone as well as combined yielded a similar outcome as the genotyping, yielding no obvious pattern (Fig. 5 and Supporting Information S5). Cluster analysis (normalized spectral gaps below 0.1) and PERMANOVA analysis (ThKF3 $p = 0.51$, ThKF7 $p = 0.27$) detected no differentiation or cluster formation in the allelic composition of *T. hyalina* populations in accordance with these two treatments, neither over time nor between final time-points for the different incubations (*t*1-*tfin*; Fig. 5). In

parallel to F_{ST} values, the AFD values were similarly low between all bottles (*tfin*) in the present day and future treatment, regardless whether calculations were based on data including (Supporting Information Table S2) or excluding homoplastic alleles (data not shown), which suggests very low differentiation between those populations.

Within a well-controlled and non-variable environment in the laboratory, selection would be expected to be directional (Morrissey and Hadfield, 2012). The observed allelic stability could be indicative of two ecological scenarios: First, the treatment conditions may not have exceeded the phenotypic plasticity of most genotypes, and the selection pressures were thus insufficient to induce population shifts. This is consistent with the high plasticity observed before in *T. hyalina* and other key species of this ecosystem (Pančić *et al*., 2015; Hoppe *et al*., 2018a; Wolf *et al*., 2018), as well as with stable productivity and species composition in the same community incubation (Hoppe *et al*., 2018b). Alternatively, either the characteristics under selection or the microsatellite alleles may be widely and evenly distributed throughout the population, and the selected genetic variants may therefore not be associated with a single genomic background ('soft sweep', Weigand and Leese, 2018). In this case, intraspecific selection would remain undetected on the basis of a neutral marker, such as the applied microsatellites, but can nevertheless provide clues on the evolution of the investigated populations (such as frequency of recombination or exchange with other populations).

Since bulk field sampling for MPB is very simple compared to genotyping individual isolates, we used the opportunity to apply our barcoding approach also to a third experimental treatment. This additional dataset revealed that there are limits to the described population stability under certain conditions. All samples in the *acidification* treatment (1000 μatm pCO_2 and 1.8°C) formed a distinct cluster separate from all other community samples in the PCA (Fig. 5A, normalized spectral gap: 0.15, PERMANOVA: $p = 0.003$) and total population size decreased (Supporting Information Fig. S9). Accordingly, AFD values between bottles of this treatment and the other two were significantly higher (Supporting Information Table S2; *t* test: $t(11) = -9.03$, $p \leq 0.001$), i.e. populations were more different. Here, the plasticity of *T. hyalina* populations was apparently insufficient to sustain their competitive ability. This finding is in line with the frequently observed pattern that negative effects of ocean acidification manifest already at lower pCO_2 levels when combined with low temperatures (Sett *et al*., 2014; Hoppe *et al*., 2018a; Wolf *et al*., 2018). Although a decrease in population size adds stochasticity to the MPB results, it is still likely that the remaining individuals at the final time-point were a subset of genotypes with high fitness under the applied conditions. Directional selection underlying this strain sorting is furthermore supported by the results of the genetic simulation based on metadata from the experiments assuming no selection (Supporting Information S8): In contrast to our experimental results, the different treatments and replicates in the simulation were randomly distributed, which makes it unlikely that the observed allelic shift was caused by coincidence or genetic drift. Thus, the application of our novel method allowed us to reveal potential climate change driven changes in the population structure of a marine diatom under some but not all experimental scenarios, and without tedious isolation of individuals from all treatments.

Natural populations in in situ spring bloom communities. The application of MPB in field samples could drastically increase our understanding of *in situ* population dynamics, which thus far remain largely under-sampled. We here show that the method can also be applied to field samples containing the full diversity of marine spring bloom protist communities. The analysis of samples collected throughout spring blooms in 2016 and 2017 from a Svalbard fjord resolved meaningful population differentiation with regard to environmental factors for a sample number almost impossible to investigate by traditional genotyping. Here, each of the two primer pairs yielded different pieces of information about the two bloom seasons.

A PCA of the results of ThKF3 revealed clear differentiation in allele frequency distributions between the spring blooms of 2016 and 2017, but only few allelic shifts within the peak of one bloom season (Fig. 6A). Cluster analysis grouped the majority of samples of the same year together (normalized spectral gap: 0.15), which was confirmed by permanova analysis ($p = 0.003$). A few exceptions exist in 2016, where the early as well as the latest samples deviate from the main cluster. Interestingly, the late samples lie within the 2017 cluster (*t*9 and *t*11) and were obtained during the end-phase of the diatom dominated bloom (mid-May), when nutrient concentrations declined (Z.T. Smoła, J.M. Wiktor, C.M. Hoppe, F. Cottier, M. Greenacre, I. Salter, et al., in preparation). Samples of 2017 appeared very homogenous when analysed through primer ThKF3 and formed a single cluster. Analysis with primer ThKF7 (Fig. 6B) showed only one cluster of samples before and during the bloom peaks of both years [which occurred in early of May (∼*t*5), see C.J.M. Hoppe, K.K.E. Wolf, F. Cottier, E. Leu, and B. Rost (submitted to Global Biogeochem Cycles)]. Samples from the late bloom phase of 2017, however, when nitrate concentrations declined and conditions became less favourable for early bloom forming diatoms like *T. hyalina*, clearly diverged from this main cluster, spreading out into different directions in the PCA

(normalized spectral gap between main cluster and outliers: 0.32). This could be a distortion of MPB results due to much smaller *T. hyalina* populations or may suggest that the method detected non-directional genetic drift during the final phase of the bloom.

It is certainly possible that genetic marker resolution of both loci may have been too low to detect more subtle shifts within the peak phase of the blooms. However, population differentiation between early and late bloom phases and (especially with primer ThKF3) also between years (Fig. 6A), were clearly detectable. This dataset illustrates how the application of MPB in environmental contexts, for example, though the feasibility of high temporal resolution, can resolve intraspecific shifts in response to complex environments, even without previous knowledge of when and where exactly they will occur.

MPB as a tool for population genetics

While MPB opens new avenues for increased sampling capacities and resolution, especially in planktonic organisms, it does have some limitations compared to traditional methodology: In parallel to PoolSeq approaches, the analysis is based on the entire pool of present alleles of the target loci, but in contrast to PoolSeq, this pool contains not an artificial collection of individuals but an entire natural community sample. Although this removes the bias of subsampling only a small, potentially unrepresentative fraction of individuals, alleles cannot be attributed to specific genotypes. Therefore, information on haplotypes and thus some commonly used population genetic measures cannot be retrieved (e.g. estimates of genotypic diversity, linkage of loci and heterozygosity levels). Furthermore, PCR and sequencing biases are possible and some erroneous sequences may be included, depending on the susceptibility of the locus to PCR stutter at hand and the frequency distribution of alleles (stutters of low-frequency alleles will be removed by the frequency filter). These potential errors need to be considered and contained during analysis (see below).

Using sequencing instead of fragment analyses allows the analysis of alleles based on their genetic code instead of their length, which is increasingly recognized as being crucial for a realistic assessment of genetic diversity (e.g. Estoup *et al*., 2002; Germain-Aubrey *et al*., 2016). At the same time, the potential influence of sequencing errors needs to be considered (especially for homoplastic alleles). In this study, these were minimized by choosing relatively short target sequences and by applying strict quality control filters. These strict filters include discarding sequences with any primer mismatches, large merging overlap, very low maximum expected error threshold (eemax), and very importantly, removal of low frequency amplicons, which are most likely to contain errors. Microsatellite loci for MPB are best chosen to be as short as possible in order to decrease PCR bias and error probabilities during sequencing as well as allowing improved quality control during reverse complementation by a large overlap. They should also contain a high signal-to-noise ratio with small stutter susceptibility, which is more likely in longer repeat patterns (minimum 3*n* repeats).

Our data illustrate that the current knowledge about populations remains a patchwork picture of the methodological perspectives and even the specific loci at hand: the results of both primer pairs (ThKF3 and ThKF7) provide consistent yet distinct aspects of information about the investigated field samples. Since analysis is based on a multidimensional approach, the information of one locus can mask the other, which is why data of several loci should not necessarily be merged but first examined separately. The analysis and interpretation of high diversity datasets can thus not just follow a ready-made protocol but needs to be adjusted and evaluated on an individual basis.

Finally, as for all methods based on neutral marker genes of non-model species, we cannot know where our target region is located in relation to genes under selection or how evenly its alleles are distributed throughout the population. In many ways, these advantages and disadvantages are similar to those discussed for well-established pool-sequencing approaches (Futschik and Schlötterer, 2010; Schlötterer *et al*., 2014; Vartia *et al*., 2016).

Conclusions

Our results show that MPB provides a simple protocol for pooled analysis of relative microsatellite allele frequencies at high accuracy in simplified laboratory populations as well as in experimental and environmental samples of natural diversity. The four most prominent benefits of MPB are (i) resource efficiency, (ii) easy applicability to bulk field samples, which in turn (iii) avoids subsampling bias, and (iv) the potential to resolve homoplasy. These benefits allow for the analysis of population dynamics and selection processes within natural populations of non-model species at high temporal resolution, which, until now, represents a major challenge in ecology and evolution research (Weigand and Leese, 2018). While MPB provides limited information on absolute characteristics of single populations as a snapshot, it provides the possibility to compare many populations from different environments and follow their development even over longer periods of time at high resolution, e.g. within experimental setups, seasonal or interannual variations or expected incidents. Its advantages are therefore

strongest when linked to eco-evolutionary hypotheses in controlled or closely observed environments, where population shifts can be correlated with treatments or events. Our new MPB approach also revealed that loci can vary in their homoplastic variability and their information content throughout the seasonal development of a population. Especially in unicellular planktonic organisms, MPB can facilitate highly improved temporal and spatial resolution of intraspecific patterns in complex natural communities without single specimen isolation. The mechanisms that drive phytoplankton population dynamics or stability in highly productive systems (such as those of *T. hyalina* in the coastal Arctic) are still not well understood. Applying MPB can help to elucidate temporal and spatial population dynamic processes in natural settings as well as experimental scenarios and can thus reveal evolutionary mechanisms at work that shape phytoplankton populations in their complex environment.

Acknowledgements

The authors thank the AWIPEV station teams of 2016 and 2017 for their help during fieldwork on Svalbard and are grateful for the financial support of the Norwegian Research Council through an Arctic Field grant. The authors thank Laura Wischnewski and Mandy Kiel for their help with community incubations and strain isolations and Annegret Müller and Alek Bolte for technical assistance with molecular tools. Daniel Scholz, Eva Leu, Finlo Cottier and the team of the project 'FAABulous' kindly provided us with environmental data of their mooring in the Kongsfjord. The authors are also grateful to Mathias Wegner for constructive criticism on methodology and differentiation measures as well as to the anonymous reviewers of this manuscript, whose feedback has greatly improved its quality.

References

Arnaud-Haond, S., and Belkhir, K. (2007) Genclone: a computer program to analyse genotypic data, test for clonality and describe spatial clonal organization. *Mol Ecol Notes* **7**: 15–17.

Becks, L., Ellner, S.P., Jones, L.E., and Hairston, N.G., Jr. (2010) Reduction of adaptive genetic diversity radically alters eco-evolutionary community dynamics. *Ecol Lett* **13**: 989–997.

Berner, D. (2019) Allele frequency difference AFD–an intuitive alternative to FST for quantifying genetic population differentiation. *Genes* **10**: 308.

Bernhardt, J.R., and Leslie, H.M. (2013) Resilience to climate change in coastal marine ecosystems. *Annu Rev Mar Sci* **5**: 371–392.

Bolger, A.M., Lohse, M., and Usadel, B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**: 2114–2120.

Boyd, P.W., Collins, S., Dupont, S., Fabricius, K., Gattuso, J.-P., Havenhand, J., *et al*. (2018) Experimental strategies to assess the biological ramifications of multiple drivers of global ocean change—a review. *Glob Chang Biol* **24**: 2239–2261.

Collins, S., Rost, B., and Rynearson, T.A. (2014) Evolutionary potential of marine phytoplankton under ocean acidification. *Evolution Appl* **7**: 140–155.

Earl, D.A., and vonHoldt, B.M. (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv Genet Resou* **4**: 359–361.

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

Erdner, D.L., Richlen, M., McCauley, L.A.R., and Anderson, D.M. (2011) Diversity and dynamics of a widespread bloom of the toxic Dinoflagellate *Alexandrium fundyense*. *PLoS One* **6**: e22965.

Estoup, A., Jarne, P., and Cornuet, J.-M. (2002) Homoplasy and mutation model at microsatellite loci and their consequences for population genetics analysis. *Mol Ecol* **11**: 1591–1604.

Evanno, G., Regnaut, S., and Goudet, J. (2005) Detecting the number of clusters of individuals using the software structure: a simulation study. *Mol Ecol* **14**: 2611–2620.

Excoffier, L., and Lischer, H.E.L. (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and windows. *Mol Ecol Resour* **10**: 564–567.

Ferretti, L., Ramos-Onsins, S.E., and Pérez-Enciso, M. (2013) Population genomics from pool sequencing. *Mol Ecol* **22**: 5561–5576.

Field, C.B., Behrenfeld, M.J., Randerson, J.T., and Falkowski, P. (1998) Primary production of the biosphere: integrating terrestrial and oceanic components. *Science* **281**: 237–240.

Futschik, A., and Schlötterer, C. (2010) The next generation of molecular markers from massively parallel sequencing of pooled DNA samples. *Genetics* **186**: 207–218.

Gaebler-Schwarz, G., Medlin, L.K., and Leese, F. (2015) A puzzle with many pieces: the genetic structure and diversity of *Phaeocystis Antarctica* Karsten (Prymnesiophyta) AU - Gäbler-Schwarz. *S Eur J Phycol* **50**: 112–124.

Germain-Aubrey, C.C., Nelson, C., Soltis, D.E., Soltis, P.S., and Gitzendanner, M.A. (2016) Are microsatellite fragment lengths useful for population-level studies? The case of *Polygala lewtonii* (Polygalaceae). *Appl Plant Sci* **4**: 1500115.

Godhe, A., and Rynearson, T. (2017) The role of intraspecific variation in the ecological and evolutionary success of diatoms in changing environments. *Philos Trans Roy Soc B* **372**: 20160399.

Godhe, A., Sjöqvist, C., Sildever, S., Sefbom, J., Harðardóttir, S., Bertos-Fortis, M., *et al*. (2016) Physical barriers and environmental gradients cause spatial and temporal genetic differentiation of an extensive algal bloom. *J Biogeogr* **43**: 1130–1142.

Gran, H.H. (1897) Bacillariaceen vom Kleinen Karajakfjord. *Bibliotheca Botanica* **42**: 1–24.

Guillaume, F., and Rougemont, J. (2006) Nemo: an evolutionary and population genetics programming framework. *Bioinformatics* **22**: 2556–2557.

Härnström, K., Ellegaard, M., Andersen, T.J., and Godhe, A. (2011) Hundred years of genetic structure in a sediment

revived diatom population. *Proc Natl Acad Sci* **108**: 4252–4257.

Haubold, B., and Hudson, R.R. (2000) LIAN 3.0: detecting linkage disequilibrium in multilocus data. *Bioinformatics* **16**: 847–849.

Hillebrand, H., and Matthiessen, B. (2009) Biodiversity in a complex world: consolidation and progress in functional biodiversity research. *Ecol Lett* **12**: 1405–1419.

Hoppe, C.J.M., Flintrop, C.M., and Rost, B. (2018a) The Arctic picoeukaryote *Micromonas pusilla* benefits synergistically from warming and ocean acidification. *Biogeosci Discuss* **15**: 4353–4365.

Hoppe, C.J.M., Wolf, K.K.E., Schuback, N., Tortell, P.D., and Rost, B. (2018b) Compensation of ocean acidification effects in Arctic phytoplankton assemblages. *Nat Clim Change* **8**: 529–533.

Jakobsson, M., and Rosenberg, N.A. (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **23**: 1801–1806.

John, U., Tillmann, U., Hülskötter, J., Alpermann, T.J., Wohlrab, S., and Van de Waal, D.B. (2015) Intraspecific facilitation by allelochemical mediated grazing protection within a toxigenic dinoflagellate population. *Proc Roy Soc London B* **282**: 20141268.

Larsen, J.N., Anisimov, O.A., Constable, A., Hollowed, A.B., Maynard, N., Prestrud, P., *et al*. (2014) Polar regions. In *Climate Change 2014: Impacts, Adaptation, and Vulnerability Part B: Regional Aspects Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge: Cambridge University Press.

Kosman, E., and Jokela, J. (2019) Dissimilarity of individual microsatellite profiles under different mutation models: empirical approach. *Ecol Evol* **9**: 4038–4054.

Lynch, M., Gabriel, W., and Wood, A.M. (1991) Adaptive and demographic responses of plankton populations to environmental change. *Limnol Oceanogr* **36**: 1301–1312.

Martin, M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J* **17**: 10–12.

Medlin, L.K. (2007) If everything is everywhere, do they share a common gene pool? *Gene* **406**: 180–183.

Meyer, J.R., Ellner, S.P., Hairston, N.G., Jones, L.E., and Yoshida, T. (2006) Prey evolution on the time scale of predator–prey dynamics revealed by allele-specific quantitative PCR. *Proc Natl Acad Sci* **103**: 10690–10695.

Minter, E.J.A., Lowe, C.D., Brockhurst, M.A., and Watts, P. C. (2015) A rapid and cost-effective quantitative microsatellite genotyping protocol to estimate intraspecific competition in protist microcosm experiments. *Methods Ecol Evol* **6**: 315–323.

Morrissey, M.B., and Hadfield, J.D. (2012) Directional selection in temporally replicated studies is remarkably consistent. *Evolution* **66**: 435–442.

Nadler, B., Lafon, S., Kevrekidis, I., and Coifman, R. (2005) Diffusion maps, spectral clustering and eigenfunctions of Fokker-Planck operators. *Adv Neural Inform Process Syst* **18**: 955–962.

Narum, S.R., Buerkle, C.A., Davey, J.W., Miller, M.R., and Hohenlohe, P.A. (2013) Genotyping-by-sequencing in ecological and conservation genomics. *Mol Ecol* **22**: 2841–2847.

Pančić, M., Hansen, P.J., Tammilehto, A., and Lundholm, N. (2015) Resilience to temperature and pH changes in a future climate change scenario in six strains of the polar diatom *Fragilariopsis cylindrus*. *Biogeosciences* **12**: 4235–4244.

Pritchard, J.K., Stephens, M., and Donnelly, P. (2000) Inference of population structure using multilocus genotype data. *Genetics* **155**: 945–959.

Rengefors, K., Kremp, A., Reusch, T.B., and Wood, A.M. (2017) Genetic diversity and evolution in eukaryotic phytoplankton: revelations from population genetic studies. *J Plankton Res* **39**: 165–179.

Rognes, T., Flouri, T., Nichols, B., Quince, C., and Mahé, F. (2016) VSEARCH: a versatile open source tool for metagenomics. *PeerJ* **4**: e2584.

Rost, B., Zondervan, I., and Wolf-Gladrow, D. (2008) Sensitivity of phytoplankton to future changes in ocean carbonate chemistry: current knowledge, contradictions and research directions. *Mar Ecol Prog Ser* **373**: 227–237.

Ruggiero, M.V., D'Alelio, D., Ferrante, M.I., Santoro, M., Vitale, L., Procaccini, G., and Montresor, M. (2017) Clonal expansion behind a marine diatom bloom. *ISME J* **12**: 463–472.

Rynearson, T., Newton, J., and Armbrust, E. (2006) Spring bloom development, genetic variation, and population succession in the planktonic diatom *Ditylum brightwellii*. *Limnol Oceanogr* **51**: 1249–1261.

Rynearson, T.A., and Armbrust, V.E. (2004) Genetic differentiation among populations of the planktonic marine diatom *Ditylum brightwellii* (Bacillariophyceae). *J Phycol* **40**: 34–43.

Saravanan, V., and Godhe, A. (2010) Genetic heterogeneity and physiological variation among seasonally separated clones of *Skeletonema marinoi* (Bacillariophyceae) in the Gullmar Fjord, Sweden. *Eur J Phycol* **45**: 177–190.

Šarhanová, P., Pfanzelt, S., Brandt, R., Himmelbach, A., and Blattner, F.R. (2018) SSR-seq: genotyping of microsatellites using next-generation sequencing reveals higher level of polymorphism as compared to traditional fragment size scoring. *Ecol Evol* **8**: 10817–10833.

Scheinin, M., Riebesell, U., Rynearson, T.A., Lohbeck, K.T., and Collins, S. (2015) Experimental evolution gone wild. *J Roy Soc Interface* **12**: 20150056.

Schlötterer, C., Kofler, R., Versace, E., Tobler, R., and Franssen, S.U. (2014) Combining experimental evolution with next-generation sequencing: a powerful tool to study adaptation from standing genetic variation. *Heredity* **114**: 431.

Sett, S., Bach, L.T., Schulz, K.G., Koch-Klavsen, S., Lebrato, M., and Riebesell, U. (2014) Temperature modulates coccolithophorid sensitivity of growth, photosynthesis and calcification to increasing seawater CO_2. *PLoS ONE* **9**: e88308.

Sjöqvist, C., Godhe, A., Jonsson, P.R., Sundqvist, L., and Kremp, A. (2015) Local adaptation and oceanographic connectivity patterns explain genetic differentiation of a marine diatom across the North Sea–Baltic Sea salinity gradient. *Mol Ecol* **24**: 2871–2885.

Tammilehto, A., Watts, P.C., and Lundholm, N. (2017) Isolation by time during an Arctic phytoplankton spring bloom. *J Eukaryot Microbiol* **64**: 248–256.

Vartia, S., Villanueva-Cañas, J.L., Finarelli, J., Farrell, E.D., Collins, P.C., Hughes, G.M., *et al*. (2016) A novel method

of microsatellite genotyping-by-sequencing using individual combinatorial barcoding. *R Soc Open Sci* **3**: 150565.
Viruel, J., Haguenauer, A., Juin, M., Mirleau, F., Bouteiller, D., Boudagher-Kharrat, M., *et al*. (2018) Advances in genotyping microsatellite markers through sequencing and consequences of scoring methods for *Ceratonia siliqua* (Leguminosae). *Appl Plant Sci* **6**: e01201.
Weigand, H., and Leese, F. (2018) Detecting signatures of positive selection in non-model species using genomic data. *Zool J Linn Soc* **184**: 528–583.
West-Eberhard, M.J. (2003) *Developmental Plasticity and Evolution*. Oxford: Oxford University Press.
Wolf, K.K.E., Hoppe, C.J.M., and Rost, B. (2018) Resilience by diversity: large intraspecific differences in climate change responses of an Arctic diatom. *Limnol Oceanogr* **63**: 397–411.
Wolf, K.K.E., Romanelli, E., Rost, B., John, U., Collins, S., Weigand, H., and Hoppe, C.J. (2019) Company matters: the presence of other genotypes alters traits and intraspecific selection in an Arctic diatom under climate change. *Glob Chang Biol* **25**: 2869–2884.
Yachi, S., and Loreau, M. (1999) Biodiversity and ecosystem productivity in a fluctuating environment: the insurance hypothesis. *Proc Natl Acad Sci U S A* **96**: 1463–1468.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Appendix S1: Supplementary Information.