ORIGINAL ARTICLE

Physical barriers and environmental gradients cause spatial and temporal genetic differentiation of an extensive algal bloom

Anna Godhe[1]*, Conny Sjöqvist[2,3], Sirje Sildever[4], Josefin Sefbom[1], Sara Harðardóttir[5], Mireia Bertos-Fortis[6], Carina Bunse[6], Susanna Gross[1], Emma Johansson[1], Per R. Jonsson[7], Saghar Khandan[8], Catherine Legrand[6], Inga Lips[4], Nina Lundholm[5], Karin E. Rengefors[8], Ingrid Sassenhagen[8], Sanna Suikkanen[2], Lisa Sundqvist[1] and Anke Kremp[2]

[1]*Department of Marine Sciences, University of Gothenburg, Gothenburg SE 405 30, Sweden,* [2]*Finnish Environmental Institute/Marine Research Centre, Helsinki 00560, Finland,* [3]*Environmental and Marine Biology, Åbo Akademi University, Turku 20520, Finland,* [4]*Marine Systems Institute, Tallinn University of Technology, Tallinn 12618, Estonia,* [5]*Natural History Museum of Denmark, University of Copenhagen, Copenhagen K DK-1307, Denmark,* [6]*Centre for Ecology and Evolution in Microbial model Systems - EEMiS, Department of Biology and Environmental Science, Linnæus University, Kalmar SE 391 82, Sweden,* [7]*Department of Marine Sciences - Tjärnö, University of Gothenburg, Strömstad SE 452 96, Sweden,* [8]*Aquatic Ecology, Lund University, Lund SE 223 62, Sweden*

*Correspondence: Anna Godhe, Department of Marine Sciences, University of Gothenburg, Box 461, SE 405 30 Gothenburg, Sweden.
E-mail: anna.godhe@gu.se

ABSTRACT

Aim To test if a phytoplankton bloom is panmictic, or whether geographical and environmental factors cause spatial and temporal genetic structure.

Location Baltic Sea.

Method During four cruises, we isolated clonal strains of the diatom *Skeletonema marinoi* from 9 to 10 stations along a 1132 km transect and analysed the genetic structure using eight microsatellites. Using *F*-statistics and Bayesian clustering analysis we determined if samples were significantly differentiated. A seascape approach was applied to examine correlations between gene flow and oceanographic connectivity, and combined partial Mantel test and RDA based variation partitioning to investigate associations with environmental gradients.

Results The bloom was initiated during the second half of March in the southern and the northern- parts of the transect, and later propagated offshore. By mid-April the bloom declined in the south, whereas high phytoplankton biomass was recorded northward. We found two significantly differentiated populations along the transect. Genotypes were significantly isolated by distance and by the south–north salinity gradient, which illustrated that the effects of distance and environment were confounded. The gene flow among the sampled stations was significantly correlated with oceanographic connectivity. The depletion of silica during the progression of the bloom was related to a temporal population genetic shift.

Main conclusions A phytoplankton bloom may propagate as a continuous cascade and yet be genetically structured over both spatial and temporal scales. The Baltic Sea spring bloom displayed strong spatial structure driven by oceanographic connectivity and geographical distance, which was enhanced by the pronounced salinity gradient. Temporal transition of conditions important for growth may induce genetic shifts and different phenotypic strategies, which serve to maintain the bloom over longer periods.

Keywords
adaptation, environmental gradient, gene flow, genetic structure, isolation by distance, population, seascape, *Skeletonema*

INTRODUCTION

Algal blooms occur in aquatic systems whenever grazing is low and favourable hydrographic conditions support a quick increase in cell abundance. Often only one or a few phytoplankton species are involved, which rapidly grow to cell densities that may be visible as discoloration of the water (Smayda & Reynolds, 2001). Blooms can be local phenomena in closed embayments or extend over vast ocean basins, and may last from a few days to many weeks (Edwards *et al.*,

2006; Kremp *et al.*, 2009). Despite the predominant asexual mode of propagation during a phytoplankton bloom, extensive genotypic diversity has been observed throughout bloom periods (Tesson *et al.*, 2014).

Although it is well known that phytoplankton populations may display large genotypic diversity (Rynearson *et al.*, 2006; Erdner *et al.*, 2011), information on spatial and temporal population genetic dynamics in blooms is still sparse. A bloom that is spread out in time and space may consist of one genetic population that propagates over vast areas or, alternatively, of several successive or spatially differentiated populations each adapted to local hydrographic conditions of the variable hydrographic environment. The propagation of a bloom in spatial and temporal dimensions can potentially be explained by a seascape approach, as marine population connectivity depends on oceanographic currents combined with characteristics of dispersive life history stages (Kenchington *et al.*, 2006). Correlations between empirical genetic data and high-resolution oceanographic connectivity models have shown good agreement for marine meroplanktonic (Serra *et al.*, 2010; Sunday *et al.*, 2014) and holoplanktonic organisms, including phytoplankton (Casabianca *et al.*, 2012). Additionally, the environment may affect the genetic structure of a phytoplankton bloom. Phytoplankton populations have been shown to contain high phenotypic diversity, for example, in their response to environmental parameters or gradients (Alpermann *et al.*, 2009; Sjöqvist *et al.*, 2015), indicating that standing genetic variation exists for selection to act upon. Moreover, phenotypic adaptation has been documented from multi-generation lab experiments (Collins & Bell, 2004; Lohbeck *et al.*, 2012). Environmental selective forces can thus act on strains, and generate populations adapted to a specific geographical or temporal niche of the bloom. Recent studies have reported changes in genetic structure of phytoplankton populations and discussed these results in context of changing environmental conditions (Rynearson & Armbrust, 2005; Richlen *et al.*, 2012). However, statistical evidence for a relationship between environmental conditions and population level spatial and temporal dynamics of blooms of planktonic protists is lacking.

In most parts of the Baltic Sea, the spring bloom develops during March and April with vertical stratification, and lasts until late April to May when nutrients are depleted and sedimentation of cells occurs. The bloom commences when the depth of the upper mixed layer becomes shallow enough to coincide with the euphotic zone, and remains stable for a long enough period to allow entrained seed populations to proliferate to bloom concentrations (Wasmund *et al.*, 1998). Vertical stability in spring can result from both thermal and salinity stratification (Smetacek & Passow, 1990). Different hypotheses on the spatial and temporal expansion of the bloom across the Baltic basins have been proposed, including a progression from south to north due to a northward delay in warming of surface waters (Jansson, 1978), as well as a synchronized mosaic-like development due to local salinity stratification including movement of water masses from shallow stratified southern and northern coastal areas to the deep central off-shore regions (Kahru & Nömmann, 1990; Stipa, 2002). Transport of water masses is implied by both scenarios, and suggests a high degree of connectivity in the Baltic spring bloom, which could support a panmictic phytoplankton community.

In this study we investigated whether an extensively distributed basin-wide algal bloom like the Baltic spring bloom is panmictic, as expected by large scale transport processes underlying the suggested bloom expansion scenarios, or whether it is genetically structured in time and space due to temporal and spatial changes in physical and environmental factors. To address these questions we analysed the genetic structure and gene flow of one of the dominant spring diatom species, *Skeletonema marinoi* Sarno et Zingone, during four cruises in March and April 2013 together with a variety of hydrographic and hydrochemical parameters. Additionally, an oceanographic connectivity model was applied to assess the relationship between oceanographic transport and gene flow. The chain-forming marine diatom *S. marinoi* occurs all year round in the Baltic Sea region, but reaches its highest abundance during the spring bloom (Saravanan & Godhe, 2010). It forms resting stages which sediment to the sea floor, and these propagules may survive for years in the sediment and thereby re-seed the planktonic population (McQuoid *et al.*, 2002). Provided a plentiful nutrient supply, the cells exhibit a growth rate of one division per day (Taylor *et al.*, 2009). The predominant means of propagation is through vegetative division, but auxospore formation and sexual reproduction has been documented from Baltic Sea populations (Godhe *et al.*, 2014).

MATERIALS AND METHODS

Sample collection and environmental parameters

All material and data were collected while on board the cargo vessel Finnmaid during four cruises in March and April 2013 (cruise A March 4–7; cruise B March 19–22; cruise C April 4–7; cruise D April 16–19). Water samples were collected during a 48 hour time period for each cruise. The shipping route stretched across a 1132 km transect in the Baltic Sea (Fig. 1, Table S1 in Appendix S1) from Travemünde (Germany, 53.9667° N, 10.8667° E) to Helsinki (Finland, 60.1708° N, 24.9375° E). As a 'ship of opportunity', Finnmaid is equipped with a ferry box, which is an automated measuring system recording chlorophyll fluorescence, salinity, temperature and fluorescence of coloured dissolved organic matter (fCDOM) while moving (Rantajärvi, 2003). Water from 8 m depth is pumped into the system every 100–200 m. Additional water samples are taken by an automatic sampling carousel at 24 fixed stations along the transect for spectrophotometric chlorophyll *a* (Chl *a*) and inorganic nutrients (NO_3-NO_2, PO_4, SiO_2) analyses (Grasshoff *et al.*, 1983).

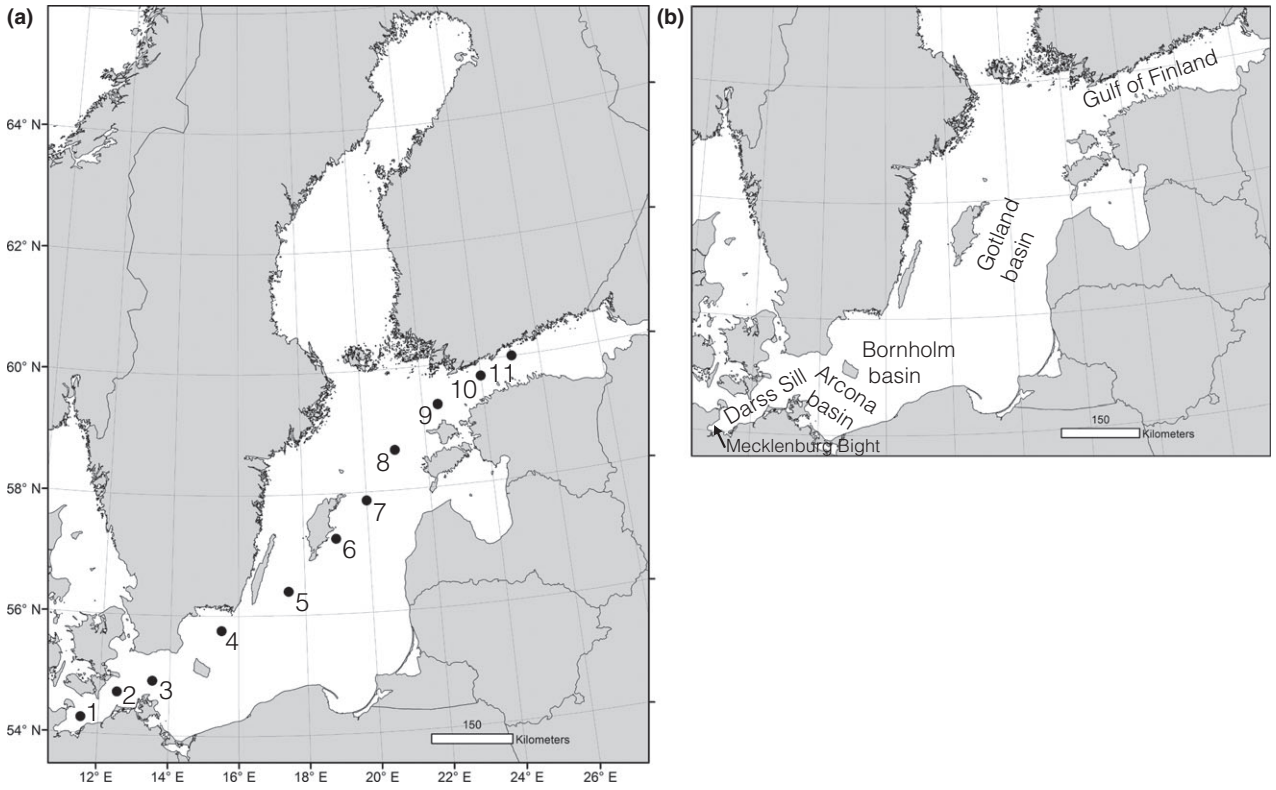

Figure 1 Map of Baltic Sea. (a) Four cruises were conducted on board MS Finnmaid along a south-west–north-east transect during the period March 4 to April 19, 2013. Eleven fixed sampling locations (1–11) were selected for *Skeletonema marinoi* isolation and water sampling. (b) Baltic Sea basins mentioned in the text.

For this study, 9–10 fixed locations per cruise were sampled for *S. marinoi* isolation (Fig. 1a, Table 1). Water was collected either automatically from the sampling carousel or manually from a tap (up to 120 L) depending on the *Skeletonema* abundance. The sample was concentrated by sieving the water through a 10 μm plankton net. An additional non-concentrated water sample was collected and fixed with Lugol's solution for phytoplankton species identification and enumeration. Samples were settled in a 25 mL Utermöhl chamber (Utermöhl, 1958) and analysed under an inverted microscope Leica DM IL Bio at magnifications 200–400X.

Correlation between hydrographic variables, geographical locations and *Skeletonema* abundance was assessed by Spearman correlation using SPSS 21 (IBM Statistics, Armonk, NY, USA).

Culturing, DNA extraction and genotyping

Individual chains of *S. marinoi* were isolated immediately on board by micropipetting. After several washing steps the cells were transferred into separate wells of 24 well NUNC tissue culture plates containing 1 mL of f/2 medium (Guillard, 1975) prepared from local seawater. The plates were incubated at 5 °C, 50 μmol photons m^{-2} s^{-1}. After transfer to the laboratory, plates were regularly monitored for cell growth. When growth was confirmed, the contents were transferred to 50 mL NUNC flasks. Dense cultures were filtered onto 3 μm pore size filters (Ø 25 mm, Versapor®-3000; Pall Corporation, Cortland, NY, USA) for DNA extraction.

Genomic DNA was extracted following a CTAB-based protocol (Kooistra *et al.*, 2003). Eight microsatellite loci (Almany *et al.*, 2009) were amplified by PCR as described in Godhe & Härnström (2010). The products were analysed in an ABI 3730 (Applied Biosystems) and allele sizes were assigned relative to the internal standard (GS600LIZ). Allele sizes for the individual loci were determined and processed using GeneMapper (ABI Prism®GeneMapper™Software 3.0).

Population differentiation and structure

Genepop 4.0.7 (Raymond & Rousset, 1995) was used to estimate deviations from Hardy–Weinberg equilibrium (HWE, 10,000 Markov chain dememorizations, 20 batches and 5000 iterations per batch) of each locus, the inbreeding coefficient F_{IS}, and genotypic linkage disequilibrium between pairs of loci in each sample (10,000 dememorizations, 100 batches and 5000 iterations per batch). Levels of statistical significance were adjusted according to sequential Bonferroni correction (Rice, 1989). Identical eight-loci genotypes were identified in Microsatellite Tools for Excel (Park, 2001). The microsatellite data set was analysed for null alleles, stuttering, and large allele drop out by means of 1000 randomizations using MicroChecker 2.2.3. Null allele frequencies cannot be

Table 1 Estimates of genetic diversity among *Skeletonema marinoi* samples at 8 microsatellite loci collected along a transect in the Baltic Sea during spring bloom of 2013. Cruise A March 4–7; Cruise B March 19–22; Cruise C April 4–7, Cruise D April 16–19. For station locations see Fig 1 and Appendix S1 Table S1. *I*, total number of isolates; *N*, total numbers of genotypes; *G/N*, proportion of unique genotypes; H_E, H_O, expected and observed heterozygosity; *HD,* heterozygote deficiency (*indicates significant HD $P < 0.05$), F_{IS} inbreeding coefficient; LD_{prop} proportion of loci which deviate significantly from linkage equilibrium.

Cruise/Station	1	2	3	4	5	6	7	8	9	10	11
A											
I	26	28	21	26	22	33	32	33	27	38	†
N	20	24	16	21	16	23	28	28	20	18	
G/N	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	
H_E	0.70	0.66	0.65	0.62	0.67	0.69	0.68	0.70	0.64	0.68	
H_O	0.42	0.29	0.41	0.26	0.31	0.25	0.31	0.24	0.33	0.25	
HD	*	*	*	*	*	*	*	*	*	*	
F_{IS}	0.39	0.55	0.39	0.58	0.52	0.65	0.55	0.66	0.48	0.63	
LD_{prop}	0	0.107	0.036	0.036	0.0	0.071	0	0	0.036	0	
B											
I	27	28	20	17	17	‡	‡	22	24	12	19
N	20	16	11	11	8			16	14	4	16
G/N	1.0	1.0	1.0	1.0	1.0			1.0	1.0	1.0	1.0
H_E	0.67	0.74	0.66	0.57	0.52			0.68	0.68	0.67	0.66
H_O	0.40	0.39	0.34	0.25	0.21			0.37	0.33	0.36	0.33
HD	*	*	*	*	*			*	*	*	*
F_{IS}	0.40	0.48	0.46	0.57	0.59			0.46	0.55	0.55	0.49
LD_{prop}	0	0.071	0	0	0			0	0	§	0.036
C											
I	34	36	39	31	29	16	21	28	11	32	
N	16	19	23	13	18	6	7	3	3	8	
G/N	1.0	1.0	1.0	1.0	1.0	0.7	1.0	1.0	1.0	1.0	
H_E	0.68	0.63	0.68	0.60	0.68	0.62	0.61	0.53	0.66	0.68	
H_O	0.34	0.35	0.30	0.35	0.31	0.29	0.20	0.29	0.71	0.25	
HD	*	*	*	*	*	*	*	*	0.65	*	
F_{IS}	0.51	0.46	0.53	0.46	0.52	0.54	0.68	0.52	−0.13	0.65	
LD_{prop}	0	0	0	0.036	0	0	0	§	§	0	
D											
I	19	15	20	26	32	‡	30	14	14	14	
N	18	14	20	21	28		26	13	12	13	
G/N	1.0	1.0	1.0	1.0	1.0		1.0	1.0	1.0	1.0	
H_E	0.63	0.67	0.65	0.64	0.64		0.65	0.68	0.69	0.68	
H_O	0.34	0.34	0.31	0.35	0.30		0.25	0.24	0.24	0.19	
HD	*	*	*	*	*		*	*	*	*	
F_{IS}	0.45	0.44	0.51	0.45	0.53		0.62	0.64	0.62	0.72	
LD_{prop}	0	0	0	0	0.107		0.036	0	0	0	

†*Skeletonema marinoi* strains were only isolated from station 11 during Cruise B.

‡Due to severe wind conditions the ship passed west of the island Gotland on the south (cruise B and D) and north (cruise B) bound journey.

§Cannot be estimated due to low sample size.

accurately estimated in non-HWE loci unless the rate of inbreeding is known (van Oosterhout *et al.*, 2006). Despite susceptibility of heterozygote deficiency in some microsatellite loci (Godhe & Härnström, 2010), and no prior knowledge of the proportion of asexually reproducing individuals, we calculated null allele frequencies according to Brookfield (1996). Hence, we could exclude that heterozygote deficiency in any locus was biased at particular sampling sites.

We investigated temporal and spatial genetic differentiation between all pairs of cruises and sample locations, respectively, by calculating pairwise F_{ST} using Genetix 4.05 (Belkhir, 2004) with 10,000 permutations. Bonferroni adjustment was used to calculate *P*-values from all multiple tests (Rice, 1989). Bayesian analysis, as implemented in Structure 2.3.4 (Pritchard *et al.*, 2000; Falush *et al.*, 2003), was used to gain further insight into the gene flow between sampling stations pooled from all cruises, at each individual cruise, and at each separate sampling station over the four cruises. We assessed the number of potential clusters (*K*) among the 11 stations with pooled genotypes from all cruises, from 9 to 10 stations per individual cruise, and from 3 to 4 time points per sampling station. Each analyses consisted of five different runs at each *K* ranging from 1 to 15 for all cruises, 1–10 for individual cruises, and 1–4 for the single stations over time respectively. We calculated the estimated posterior log probability of the data, L(*K*), and the

stability of assignment patterns across runs. For each run, a burn-in period of 50,000 generations and 300,000 Markov chain Monte Carlo replications were applied. We used a model of admixture with the assumption of independent allele frequencies. An ad hoc statistic, ΔK, was calculated on the basis of the rates of change in the log likelihood of data between consecutive K values (Evanno *et al.*, 2005).

Isolation by distance and environment

Isolation by distance (IBD) analysis from matrices of genetic ($F_{ST}/1-F_{ST}$) versus Euclidean distances (Log_e of kilometres) was performed in Genepop (Raymond & Rousset, 1995). Euclidean distances were measured as linear distances between pairs of sites. Genetic isolation by salinity differences was estimated by analyses of matrices of genetic ($F_{ST}/1-F_{ST}$) versus pairwise salinity differences (Log_{10} ΔPSU). We tested for significant correlations by using Mantel test in R using the 'vegan' package Oksanen *et al.*, 2011). Statistical significance was assessed by 999 random permutations.

Gene flow and oceanographic connectivity

A new approach was used to calculate directional relative migration. The approach is a directional extension to measures of genetic differentiation and is based on a pool of migrants that is defined for each combination of two samples in pairwise comparison. The pool of migrants is calculated as the geometric means of the frequencies of the respective alleles in the two populations. The pool is then compared to each of the two populations to calculate directional genetic differentiation. Directional differentiation can subsequently be used to calculate directional relative migration. The method is further explained in (Sundqvist *et al.*, 2013). Relative migration rates were calculated using Josts'D (Jost, 2008) as a measure of genetic differentiation. Calculations were performed using the function divMigrate from the R-package 'diversity' (Keenan *et al.*, 2013).

We estimated oceanographic connectivity between the sampling sites with a biophysical model, where velocity fields from an ocean circulation model were combined with a particle-tracking routine to simulate drift trajectories at two different depth intervals representing the dispersal of diatoms. The dispersal of diatoms was simulated using the Lagrangian trajectory model TRACMASS (Döös, 1995). It is a particle-tracking model that calculates transport of particles using temporal and spatial interpolation of flow-field data from the BaltiX circulation model (see Text S1 in Appendix S2) with a time step of 15 min. Each sample site was represented by 15 grid cells closest to the locations given in Table S1 in Appendix S1. Particles were released on the 15th day of March or April over an 8-year period and allowed to drift in surface (0–2 m) or sub-surface (10–12 m) water for 20, or 30 days. Connectivity among the sampling sites was estimated by calculating the proportion of particles released from site i that ended up in site j. In total, the connectivity estimated among the sites was based on 1.98 million released particles.

We tested for significant correlations (Pearson's product-moment correlation coefficient) between estimated migration and oceanographic dispersal probability by using a variant of the Mantel test (Mantel, 1967) but including the full matrix allowing for asymmetric migration and dispersal. Statistical significance ($P < 0.05$) was assessed by 5000 random permutations. The matrix of directional migration consisted of pooled data from all stations during all four cruises (A–D), except stations 6 and 11 which were excluded due to low sample size (< 30 genotyped individuals). The oceanographic connectivity matrices represented (1) cells dispersed in the month of March or April; (2) cells dispersed in surface water (0–2 m) or sub-surface water (10–12 m); (3) cells drifting for 20 or 30 days. Additionally, we tested for significant correlations between the matrix of estimated migration versus an average of all dispersal matrices.

Genetic differentiation and environmental factors

We performed environmental association analyses to test for significant correlations between the genetic variation and the hydrographic variables estimated for the four cruises, that is, Chl *a*, *S. marinoi* abundance, temperature, salinity, fCDOM, NO_2-NO_3, PO_4, SiO_2. In all the analyses, we omitted stations with fewer than 10 genotyped strains. The environmental association analysis was performed using partial Mantel tests and redundancy analysis (RDA). One partial Mantel test per cruise was performed between all pairs of environmental variables and the genetic distance, F_{ST}. All hydrographic variables were $\log_{10}(x + 1)$ transformed before analyses using the software passage (Rosenberg & Anderson, 2011).

RDA analyses, one for each cruise and one for all cruises together, was used to disentangle the relative contribution of independent variables (biological-hydrographic and spatial components) in driving the genetic structure (dependent variable) of the Baltic Sea spring bloom. For the processing of the genetic data, the spatial data into a spatial matrix (S) and the environmental data into an environmental matrix E, see Text S2 in Appendix S2. We performed RDA analysis according to the double stop criterion (Blanchet *et al.*, 2008), which means that the forward selection is done only when the global model (containing all variables) is significant. For the global RDA, we reported adjusted regression coefficients of multiple determinations (R^2_{adj}). In addition to the global RDA, we calculated variance partitioning (Borcard *et al.*, 1992) to estimate the variance uniquely explained by the environment (E|S) or by space (S|E). Variance decomposition was performed if either the spatial or the environmental model were significant (Borcard *et al.*, 2011). In the partial RDA models, co-variables were included only if significant in either the spatial or the environmental model. All RDA analyses were performed in R (v3.1.2) using the rda and varpart functions of the 'vegan' package (Oksanen *et al.*,

2011). In addition, the forward.sel function of the 'packfor' package was used (Dray *et al.*, 2006).

RESULTS

Hydrography

The Baltic Sea spring bloom of 2013 developed during the second half of March in the south-west end of the transect, off the German coast, and simultaneously in the north-east end of the transect in the Gulf of Finland (Figs 1b & 2a), and later propagated offshore. By mid-April the spring bloom declined in the south-west, whereas high chl *a* fluorescence, was recorded northwards along the transect (Fig. 2a). The spatio-temporal abundance of *S. marinoi* was significantly correlated (Spearman's rho = 0.84, $P < 0.001$) to the chl *a* fluorescence (Fig. 2b, Table S2 in Appendix S1), and constituted up to 33% of the recorded taxa based on microscopic counts converted to biomass (data not shown). The sea surface temperature was homogenous along the transects during the three-first cruises, but in mid-April the temperature had increased in the southern part of the transect (Fig. 2c). Salinity and fCDOM concentration displayed a distinct south–north gradient, and small temporal changes (Fig. 2d,e). Concentrations of inorganic nutrients (NO_3-NO_2, PO_4, SiO_4) were initially high in the coastal areas, and lower in the Baltic proper. By mid-April, the nutrients were depleted, or markedly reduced, at all stations (Fig. 2f–h). Several hydrographic variables co-varied and were significantly correlated with latitude (Table S2 in Appendix S1).

Skeletonema population structure

Out of 611 individuals, two identical eight-loci genotypes were identified at station 6 during the third cruise (April 4–7, Table 1). At all stations the sampled populations were genetically diverse and we observed no increase in clonal structure as the bloom progressed. We observed significant heterozygote deficiency ($P < 0.05$), high $F_{IS,}$ and low proportion of loci deviating from linkage equilibrium in all samples during the four cruises, except from station 9 during cruise C (Table 1). However, from this station we were only able to genotype three individuals. All eight microsatellite loci were polymorphic. Locus S.mar3 was the least variable while S.mar5 and S.mar6 were the most variable (Table S3 in Appendix S1). There was no evidence for large allele drop out or stuttering effects using MicroChecker. Based on the method Brookfield 1, estimates of null alleles frequency were low in S.mar2, S.mar3 and S.mar6; moderate in S.mar4 and S.mar8; and highest in S.mar1, S.mar5 and S.mar7. There was no significant correlation between samples and potential null allele frequencies (2-tailed paired samples *t*-test, $P > 0.05$), and all loci were used in subsequent calculations of genetic differentiation and gene flow.

Spatial genetic differentiation displayed pronounced genetic structure in the south-west to north-east direction based on the Bayesian clustering method and pairwise F_{ST}. The calculated ΔK indicated that two clusters best explained the uppermost hierarchical level of the genetic structuring among the 11 stations (Fig. 3). The pairwise F_{ST} displayed that the southernmost stations (1–3), and the stations 4–5, respectively, were genetically differentiated from all other stations. Among stations 6–11, no or low degree of genetic differentiation was observed (Table S4 in Appendix S1). Besides the spatial structuring, there was temporal genetic differentiation among all paired cruises, except cruise A and B (Table S5 in Appendix S1). The genotypes in stations 1–3 were during all four cruises assigned to one cluster, and strains from stations 5–11 to a separate cluster (Fig. S1 a–d in Appendix S3). When testing for genetic differentiation at each sampling location over time (4 cruises), station 1–3 and 5–10 yielded one or two clusters. In the analyses in which two clusters were identified, the clusters were mixed across all individuals isolated from any cruise (data not shown). Contrary, the strains isolated from sampling station 4 formed two distinct clusters. The first two cruises were assigned to one cluster, and the two last to a second cluster (Fig. S1e in Appendix S3).

Based on the results from the Bayesian clustering analyses and the pairwise F_{ST}, we rejected the model of panmixia for the Baltic Sea spring bloom.

Genetic versus geographical and environmental distance

The Mantel test between matrices of genetic (F_{ST}) and geographical distances (km) was significant ($R = 0.79$, $P < 0.001$) and suggests significant isolation by distance. However, the genetic distance was also significantly correlated with the south–north salinity gradient ($R = 0.57$, $P < 0.001$, Fig. S2 a-b in Appendix S3), which illustrates that the effects of distance and environment were confounded and could not easily be disentangled.

Correlation of oceanographic connectivity and gene flow

Significant correlations between spatio-temporal migration patterns of *S. marinoi* and oceanographic connectivity ranged between 0.24 and 0.50 (Table 2). The analyses between the matrices of relative directional migration pattern from the different sampling stations during the four cruises, and the matrices of oceanographic connectivity (Tables S6 and S7 in Appendix S1) accentuated the importance of oceanographic circulation pattern for the genetic structuring of *S. marinoi*.

Correlation of environmental factors and genetic distance

The partial Mantel test identified a set of environmental variables that significantly co-varied with changes in population structure over the four different cruises (Table S8 in Appendix S1), but significant variables associated varied

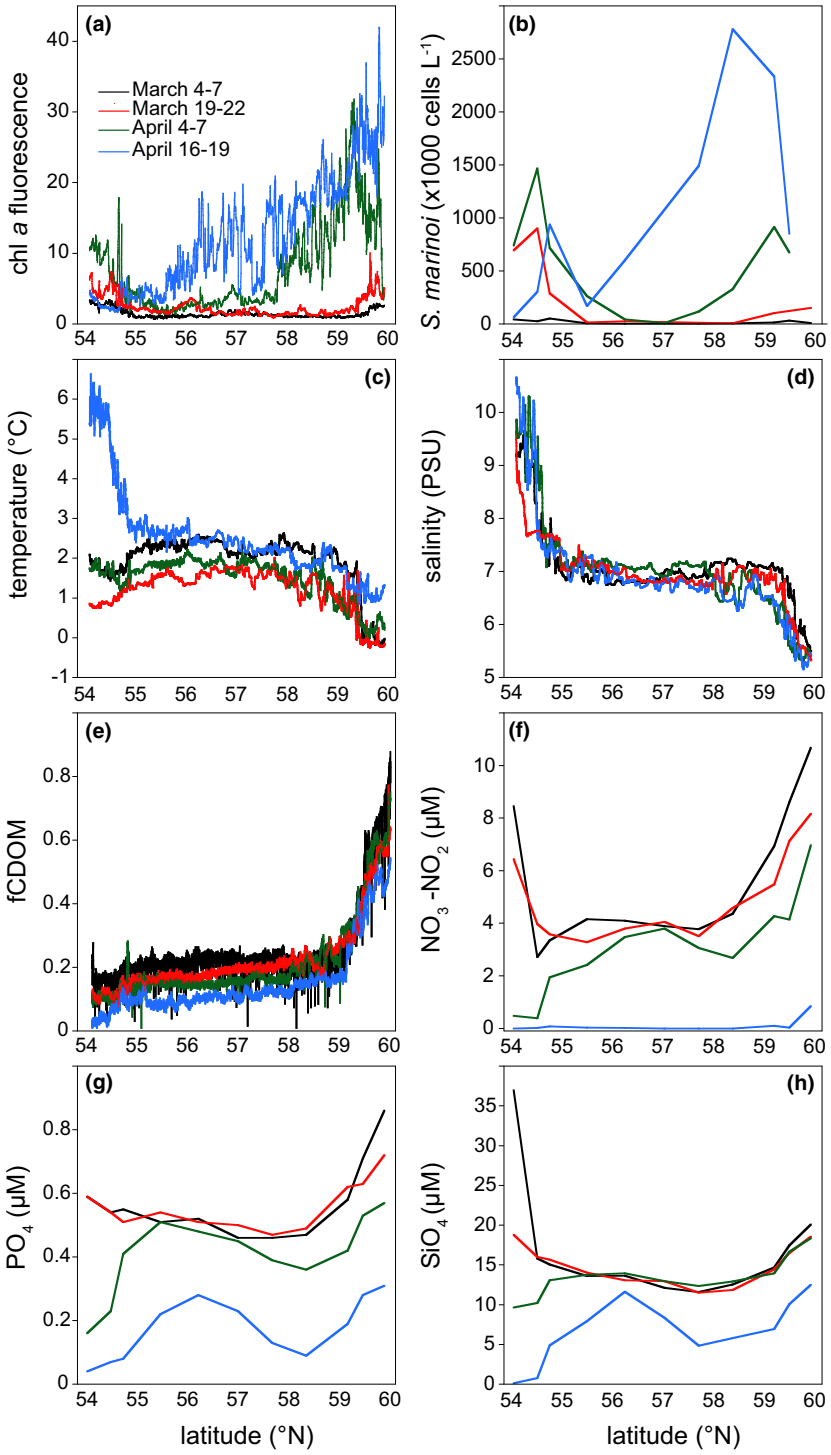

Figure 2 (a–h) Hydrographic variables along the latitudinal transect at each of the four cruises on a south–north transect in the Baltic Sea (March 4–7, March 19–22, April 4–7, April 16–19, 2013) at 8 m depth. (a) Phytoplankton biomass estimated by relative chl *a* fluorescence. (b) *Skeletonema marinoi* abundance ($\times 1000$ cells L^{-1}) (c) Temperature (°C). (d) Salinity (PSU). (e) Dissolved organic matter estimated as relative fCDOM. (f) NO_3-NO_2 concentration (µM). (g) PO_4 concentration (µM). (h) SiO_4 concentration (µM). Data in (a) and (c–e) are registered by automated ferry-box system with spatial resolution of 100–200 m along the transect of each cruise. *S. marinoi* abundance (b) is based on microscopic enumerations from the water samples collected at each of the 9–10 stations per cruise. Data in (f–h) were obtained from water samples collected by an automatic sampler at 24 fixed stations per cruise.

among the cruises. Salinity emerged as an important variable associated with the genetic structure during cruise A and D. SiO_4 concentration was significantly correlated with the F_{ST} during the two last cruises.

The global RDA was significant for the first and the last cruise (Cruise A and D), and for the merged data set including all stations from all cruises (Table 3). Salinity, *Skeletonema* cell density and the spatial matrix (PCNM) ($P < 0.05$) were included in the forward selection for Cruise A. The environmental (salinity and *Skeletonema* cell density) and the spatial matrices explained ~ 30% each of the variability in the genetic data ($P < 0.01$). The E|S and the S|E models for cruise A were not statistically significant, but the population differentiation with a split between southern (higher salinity) and northern (lower salinity) population was evident (Fig. S3a in Appendix S3). For Cruise D, *Skeletonema* cell density and the PCNM were included in the forward selection. Using the variation partitioning and conditioned analyses, the S matrix in

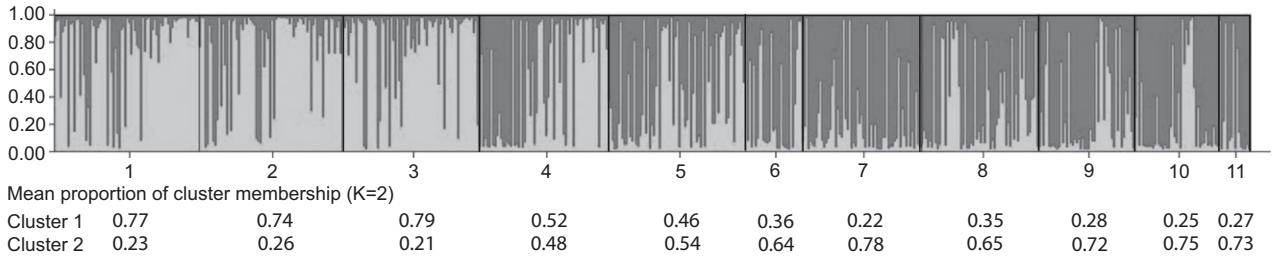

Figure 3 Genetic clustering of *Skeletonema marinoi* based on eight microsatellite loci estimated by Bayesian analyses implemented in STRUCTURE. Samples 1–11 are groups of strains established from fixed locations pooled from four cruises in the Baltic Sea conducted during the period March to April, 2013. Assignment of 611 individuals to $K = 2$ genetically distinguishable clusters. Each individual is represented by a vertical bar coloured according to the assigned group. Mean proportion of cluster membership for each sampling station is given below.

Table 2 Correlations between directional relative migration of *Skeletonema marinoi* and oceanographic connectivity in the Baltic Sea. Directional migration from all stations during all four cruises (Cruise A–D) was tested against connectivity matrices by Mantel test. Connectivity matrices tested were an average matrix based on the 2 months (March and April), the two depths (0–2 m and 10–12 m), and dispersal time of 20 and 30 days; March 0–2 and 10–12 m at dispersal time 20 and 30 days; April 0–2 m and 10–12 m at dispersal time 20 and 30 days. All connectivity matrices were based on particles realized from 15 grid cells per station. Correlation coefficients are given as R.

Migration matrix	Connectivity matrix	Correlation	*P*
Cruise A–D	Average months, depths, dispersal time	0.50	< 0.001
	March, 0–2 m, 20 days	0.15	0.13
	March, 10–12 m, 20 days	0.42	< 0.001
	March, 0–2 m, 30 days	0.24	0.046
	March, 10–12 m, 30 days	0.29	0.02
	April, 0–2 m, 20 days	0.26	0.02
	April, 10–12 m, 20 days	0.31	0.02
	April, 0–2 m, 30 days	0.25	0.047
	April, 10–12 m, 30 days	0.43	0.001

Cruise D explained ~12% of the variability in the genetic data. The E|S model was not statistically significant, but a split between southern (lower cell density) and northern (higher cell density) population was obvious (Table 3, Fig. S3b in Appendix S3). For the merged data set, SiO_4 concentration and PCNM were significant in the forward selection. The E|S model (SiO_4) and the S|E model were statistically significant. The genotypes isolated during the first two cruises grouped together and were separated from the genotypes isolated during the last two cruises. The results indicate that the lower silica concentrations towards the end of the spring bloom selected for different genotypes at all stations (Table 3, Fig. S3c,d in Appendix S3).

DISCUSSION

Until recently it was considered that marine planktonic organisms are made up of vast populations with unlimited dispersal potential due to ocean circulation. Here we show that the *Skeletonema marinoi* spring bloom within the Baltic Sea does not consist of a single panmictic population despite the uninterrupted stretch of several ocean basins. We detected two significantly differentiated populations delimited in space and separated by correlated physical and environmental dispersal barriers. We found a strong signal of genetic isolation by geographical distance. Lack of oceanographic connectivity obstructing gene flow between distant locations, and environmental gradients (salinity) along the Baltic Sea basins, appear to exert a selective pressure that accentuates genetic differentiation. Our data indicate that the spatial population structure remained largely stable throughout the bloom period. However, the depletion of silica during the progression of the bloom was related to a shift in the genetic structure. In response to environmental changes, local genotypes may be selected for and thus cause temporal differentiation.

Provided a significant isolation-by-distance effect we expect that samples collected along a transect consist of a continuum of distributed individuals over space, in which populations in remote locations become differentiated by isolation due to restricted mating probability (Wright, 1943). A significant correlation between geographical and genetic distance was indeed indicated in our data at the larger scale; however, at smaller distances this trend was not as pronounced. In addition, the circulation pattern (not necessarily associated with distance) and the associated gradient of hydrographic variables, essential for microalgal growth, also correlated significantly with the genetic structure. We therefore argue that the basis for the observed genetic structure is more complex than purely a function of distance. The transect was divided into two genetically differentiated populations in which the end locations were genetically distant, but the most important spatial barrier for gene flow was located between sample locations 3 and 5. This dispersal barrier may not be geographically fixed, but rather dynamic and dependent on seasonal circulation, for example, cyclonic gyres in the Arkona and Bornholm Basins during early spring (Leppäranta & Myrberg, 2009). At sampling station 4, two genetically separated populations over the four cruises were

Table 3 Results of the global RDA, RDA with forward selection, and variation partitioning analyses estimating the effect of environmental (E) and spatial (S) components on the genetic distance of *Skeletonema marinoi* populations in the Baltic Sea. E|S, environment controlled for space; S|E, space controlled for environment; sal, salinity; cel, *Skeletonema* cell density; sil, SiO_4 concentration; lat, latitude. Significant *P*-values are in bold. Forward selection and variation partitioning were not performed when the global model (either space or environment) was not significant.

	Cruise A		Cruise B		Cruise C		Cruise D		Cruise A+B+C+D	
	Mar 4–7		Mar 19–22		Apr 4–7		Apr 16–19		Mar 4–Apr 19	
Global RDA	E	S	E	S	E	S	E	S	E	S
R^2	**0.361**	**0.300**	0.138	0.002	0.221	0.061	0.255	**0.409**	**0.142**	**0.337**
P-value	**0.001**	**0.004**	0.275	0.456	0.299	0.160	0.098	**0.001**	**0.021**	**0.001**
Forward selection	E	S	E	S	E	S	E	S	E	S
(variables included)	(sal, cel)	(lat)	–	–	–	–	(cel)	(lat)	(sil)	(lat)
R^2	**0.304**	**0.299**	–	–	–	–	**0.230**	**0.376**	**0.125**	**0.308**
P-value	**0.008**	**0.006**	–	–	–	–	**0.019**	**0.002**	**0.004**	**0.001**
Variation partitioning	E|S	S|E	E|S	S|E	E|S	S|E	E|S	S|E	E|S	S|E
R^2	0.038	0.034	–	–	–	–	−0.023	**0.123**	**0.053**	**0.236**
P-value	0.277	0.309	–	–	–	–	0.816	**0.025**	**0.004**	**0.001**

identified. The dispersal model (averaged through 1995–2002, Fig. 4) supported this, as trajectories seeded from stations 3 and 5 were not continuous. In addition, the oceanographic matrices did not support retention of cells at station 4 as in the other stations, which prevents the build-up of a local buffering seed bank (Tahvanainen *et al.*, 2012). Hence, our data indicate two stable populations; one in the south and another in the north, and in-between a transition zone which is seeded by the two populations.

Oceanographic circulation is undoubtedly important for dispersal and gene flow in marine planktonic organisms (White *et al.*, 2010). We applied a robust biophysical dispersal model based on 8 years of monitoring data, which captures the inter-annual fluctuations in velocities. The high resolution of the BaltiX model permitted us to investigate drift at different depths, duration, and release periods. Thus, we could identify the dispersal conditions, which best correlated with calculated migration data. In the current and previous studies we have found that oceanographic connectivity explains up to 60% of the *S. marinoi* population genetic structure (Godhe *et al.*, 2013; Sjöqvist *et al.*, 2015). This indicates an important impact of oceanographic circulation on gene flow, but also that other factors are important for structuring a phytoplankton bloom. For instance, the Darss Sill (depth 18 m, Fig. 1b) constitutes a topographic and oceanographic borderline between Mecklenburg Bight and the Baltic Proper (HELCOM, 1988). Sampling location 1 and 2 are positioned to the west of the sill, whereas station 3 is positioned east of it. According to the biophysical model the dispersal potential traversing the sill was lower than between the two westernmost stations. Nevertheless, the gene flow appeared not to be affected by the constraints imposed by the sill.

The association between the genetic distances and environmental variables highlights that local hydrographic regimes are also important drivers of population genetic structure, in addition to spatial distribution and oceanographic circulation. Especially salinity was significantly correlated with latitude and disentangling the effects was problematic. The salinity gradient across the Baltic Sea explained the genetic differentiation observed, but only when it was not controlled for the spatial component. Space by itself was not a strong explanatory factor either within this data set. Genetic structure was explained with statistical confidence only when both salinity and space were included. The explicit effect of latitude per se on genetic population structure has no comprehensive biological meaning. It tells us that some factor(s) that vary across the latitudinal gradient is (are) involved in driving the differentiation. One such factor could be local adaptation to a specific salinity regime of the Baltic Sea, which could restrict gene flow and enhance the correlation between genetic and geographical distance.

Another important spatial aspect is the presence of local seed banks of *S. marinoi* and the strong link between the benthic and pelagic assemblages that supports local recruitment of vegetative cells (Godhe & Härnström, 2010). Oceanographic connectivity estimates, in fact indicate local seeding along the investigated transect. Local shallow coastal sites that are well connected to the sampling sites appear as important sources that may seed the spring bloom (Fig. 4). On the contrary, resuspension and seeding from the deeper areas of the Bornholm and Gotland Basins (100–200 m deep) is unlikely as the permanent pycnocline at 60–80 m depth prevents mixing of water and transportation of cells to the euphotic layer. Further, the deep Baltic Sea basins are permanently anoxic (Diaz & Rosenberg, 2008), which prevent resting stage germination (Lundholm *et al.*, 2011).

When all cruises were analysed by RDA and variation partitioning together, a pattern emerged where the populations from the two-first cruises in March grouped together separate from the last two cruises in April. Thus, genetic differences between populations may emerge over short temporal

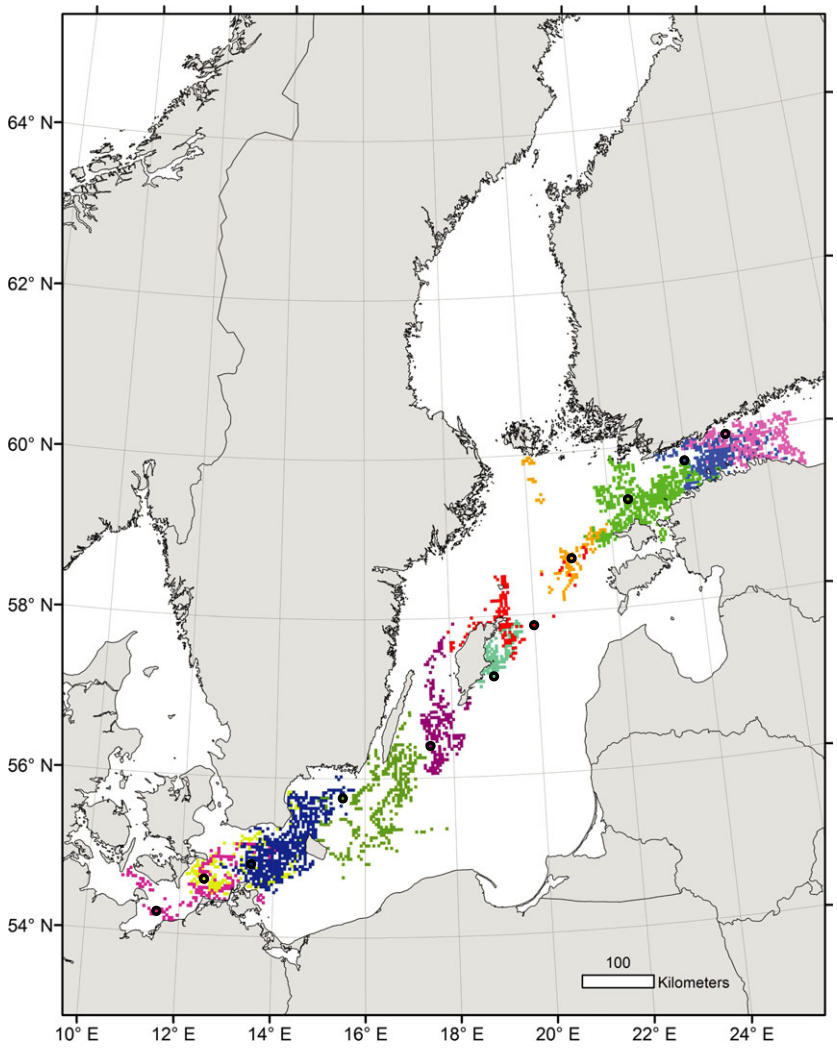

Figure 4 Map showing results from a simulation of diatom dispersal within the studied region in the Baltic Sea based on current velocities from the BaltiX circulation model. The colour-coded grid cells show the sources of simulated algae that end up at the 11 stations sampled during the four cruises (one colour for each station; 1-pink, 2-yellow, 3-blue, 4 green, 5-magenta, 6-turquoise, 7-red, 8- orange, 9-green, 10-blue, 11 pink). A grid cell is colour-coded if at least one trajectory from that grid cell ended up in the 15 grid cells representing each sample station. The simulation showed in the map represents release in April averaged over the years 1995–2002, and a dispersal duration of 20 days with a drift depth of 0–2 m.

scales, that is, a few weeks. This pattern was correlated with the depletion of silica during the course of the bloom. Silica concentrations were on average 50% lower during the last cruise (8.0 μM) compared to the first cruise (16.1 μM). Growth of diatoms, and the extent of blooms can be dependent on ambient silica concentration (Egge & Aksnes, 1992). Our data suggest that the Baltic Sea spring bloom consists of strains with different growth optima, and as the bloom progresses and the hydrographic conditions change there may be a shift towards genotypes with specific physiological adaptations to low silica concentrations. Such shifts in genotype composition imply the presence of diverse phenotypes and large standing genetic variation in the Baltic *S. marinoi* populations. Indeed, high genotypic diversity was recorded at every sampling station during each of the four cruises. Hence, the large genetic diversity observed might be the result of fluctuating selection and subsequent integration of different genotypes as the bloom progressed. Nevertheless, as we have no data on phenotypic traits from the isolated genotypes, such as silica requirements, this hypothesis remains to be tested.

We assumed neutrality of the microsatellites used for identifying the spatio-temporal genetic structure, and therefore we did not expect to see strong effects of the environment. Accordingly, the overall outcome of the RDA analyses was that the environment, that is, spatial and temporal gradients of salinity and silica, had a significant but weaker effect compared to oceanographic and spatial aspects. However, differential selection can provide genetic differentiation also in neutral loci by physical linkage to selected loci (Thibert-Plante & Hendry, 2010). In organisms with low recombination rates, which is applicable for diatoms with primarily vegetative propagation (D'Alelio *et al.*, 2010), a strong selection pressure may yield even greater effects on linked loci (Nosil *et al.*, 2009). Also, selection at certain locally adaptive loci has the effect that immigrants, which lack suitable alleles, are selected against and that this reduces realized gene flow and add to differentiation also in neutral loci (Nosil *et al.*, 2009).

Niche partitioning was recently demonstrated by single-cell genomics of the prokaryote *Prochlorococcus* in which a set of subpopulations had differential fitness and changed their rel-

ative abundance with changing environmental conditions (Kashtan *et al.*, 2014). Based on our discovery of the temporal genetic shift during the Baltic Sea diatom spring bloom, we anticipate that diatom blooms might consist of a sequence of short-lived subpopulations. If so, each subpopulation may be well adapted to particular hydrographic bloom conditions and occupy a distinct temporal-hydrographic niche, which maintain the bloom over longer periods.

ACKNOWLEDGEMENTS

We acknowledge Finnlines for providing space during the cruises, Petri Manula and Pia Varmanen, SYKE, for technical support and nutrient analyses. Dr Riina Klais, University of Tartu, assisted in planning the study, and provided statistical support. The research was supported by grants from Nordforsk Research Network, Swedish Research Council FORMAS (215-2010-751), Marine Strategic grant ECOCHANGE, Academy of Finland (256074, 251564), Oscar and Lilli Lamms Minne, Estonian Ministry of Education and Research (IUT 19-6), Centre for Ecology and Evolution in Microbial Model Systems at Linnaeus University, the Linnaeus Centre for Marine Evolutionary Biology at the University of Gothenburg through Swedish Research Council and FORMAS, Danish Research Council (DFF–1323-00258).

REFERENCES

Almany, G.R., Arruda, M.P.D., Arthofer, W. *et al.* (2009) Permanent genetic resources added to molecular ecology resources database 1 May 2009-31 July 2009. *Molecular Ecology Resources*, **9**, 1460–1466.

Alpermann, T., Beszteri, B., John, U., Tillmann, U. & Cembella, A.D. (2009) Implications of life-history transitions on the population genetic structure of the toxigenic marine dinoflagellate *Alexandrium tamarense*. *Molecular Ecology*, **18**, 2122–2133.

Belkhir, K. (2004) *GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations*. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier, France.

Blanchet, F.G., Legendre, P. & Borcard, D. (2008) Forward selection of explanatory variables. *Ecology*, **89**, 2623–2632.

Borcard, D., Legendre, P. & Drapeau, P. (1992) Partialling out the spatial component of ecological variation. *Ecology*, **73**, 1045–1055.

Borcard, D., Gillet, F. & Legendre, P. (2011) *Numerical ecology with R*. Springer, New York, Dordrecht, London, Heidelberg.

Brookfield, J.F.Y. (1996) A simple new method for estimating null allele frequency from heterozygote deficiency. *Molecular Ecology*, **5**, 453–455.

Casabianca, S., Penna, A., Pecchioli, E., Jordi, A., Basterretxea, G. & Vernesi, C. (2012) Population genetic structure and connectivity of the harmful dinoflagellate *Alexandrium minutum* in the Mediterranean Sea. *Proceedings of the Royal Society B: Biological Sciences*, **279**, 129–138.

Collins, S. & Bell, G. (2004) Phenotypic consequences of 1000 generations of selection at elevated CO_2 in a green algae. *Nature*, **431**, 566–569.

D'Alelio, D., d'Alcala, M., Dubroca, L., Sarno, D., Zingone, A. & Montresor, M. (2010) The time for sex: a biennial life cycle in a marine planktonic diatom. *Limnology and Oceanography*, **55**, 106–114.

Diaz, R. & Rosenberg, R. (2008) Spreading dead zones and consequences for marine ecosystems. *Science*, **321**, 926–929.

Döös, K. (1995) Inter-ocean exchange of water masses. *Journal of Geophysical Research-Oceans*, **100**, 13499–13514.

Dray, S., Legendre, P. & Peres-Neto, P.R. (2006) Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). *Ecological Modelling*, **196**, 483–493.

Edwards, M., Johns, D.G., Leterme, S.C., Svendsen, E. & Richardson, A.J. (2006) Regional climate change and harmful algal blooms in the northeast Atlantic. *Limnology and Oceanography*, **51**, 820–829.

Egge, J.K. & Aksnes, D.L. (1992) Silicate as regulating nutrient in phytoplankton competition. *Marine Ecology Progress Series*, **83**, 281–289.

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

Evanno, G., Regnaut, S. & Goudet, J. (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, **14**, 2611–2620.

Falush, D., Stephens, M. & Pritchard, J. (2003) Inference of population genetic structure: extensions to linked loci and correlated allele frequencies. *Genetics*, **164**, 1567–1587.

Godhe, A. & Härnström, K. (2010) Linking the planktonic and benthic habitat: genetic structure of the marine diatom *Skeletonema marinoi*. *Molecular Ecology*, **19**, 4478–4490.

Godhe, A., Egardt, J., Kleinhans, D., Sundqvist, L., Hordoir, R. & Jonsson, P.R. (2013) Seascape analysis reveals regional gene flow patterns among populations of a marine planktonic diatom. *Proceedings of the Royal Society B: Biological Sciences*, **280**, 20131599.

Godhe, A., Kremp, A. & Montresor, M. (2014) Genetic and microscopic evidence for sexual reproduction in the centric diatom *Skeletonema marinoi*. *Protist*, **165**, 401–416.

Grasshoff, K., Erhardt, M. & Kremling, K. (1983) *Methods of seawater analyses*. Verlag Chemie, Weinhein.

Guillard, R.R.L. (1975) Culture of phytoplankton for feeding marine invertebrates. *Culture of marine invertebrate animals* (ed. by W. Smith and M. Chanley), pp. 29–60. Plenum Press, New York.

HELCOM (1988) Guidelines for the Baltic monitoring programme for the third stage. Part D. Biological determi-

nands. In: *Baltic sea environment proceedings 27D*. Helsinki. *Baltic Marine Environment Protection Commission-Helsinki Commission* 164 pp.

Jansson, B.O. (1978) The Baltic–a system analysis of a semi enclosed sea. *Advances in oceanography* (ed. by H. Charnock and G. Deacon), pp. 131–183. Plenum Press, New York.

Jost, L. (2008) G_{ST} and its relatives do not measure differentiation. *Molecular Ecology*, **17**, 4015–4026.

Kahru, M. & Nömmann, S. (1990) The phytoplankton spring bloom in the Baltic Sea in 1985, 1986-multitude of spatiotemporal scales. *Continental Shelf Research*, **10**, 329–354.

Kashtan, N., Roggensack, S.E., Rodrigue, S., Thompson, J.W., Biller, S.J., Coe, A., Ding, H., Marttinen, P., Malmstrom, R.R., Stocker, R., Follows, M.J., Stepanauskas, R. & Chisholm, S.W. (2014) Single-cell genomics reveals hundreds of coexisting subpopulations in wild *Prochlorococcus*. *Science*, **344**, 416–420.

Keenan, K., McGinnity, P., Cross, T., Crozier, W. & Prodöhl, P. (2013) diveRsity: an R package for the estimation and exploration of population genetics parameters and their associated errors. *Methods in Ecology and Evolution*, **4**, 782–788.

Kenchington, E.L., Patwary, M.U., Zouros, E. & Bird, C.J. (2006) Genetic differentiation in relation to marine landscape in a broadcast-spawning bivalve mollusc (*Placopecten magellanicus*). *Molecular Ecology*, **15**, 1781–1796.

Kooistra, W.H.C.F., DeStefano, M., Mann, D.G., Salma, N. & Medlin, L.K. (2003) Phylogenetic position of *Toxarium*, a pennate-like lineage within centric diatoms (Bacillariophyceae). *Journal of Phycology*, **39**, 185–197.

Kremp, A., Lindholm, T., Dressler, N., Erler, K., Gerdts, G., Eirtovaara, S. & Leskinen, E. (2009) Bloom forming *Alexandrium ostenfeldii* (Dinophyceae) in shallow waters of the Åland archipelago, Northern Baltic Sea. *Harmful Algae*, **8**, 318–328.

Leppäranta, M. & Myrberg, K. (2009) *Physical oceanography of the Baltic Sea*. Springer-Praxis, Berlin.

Lohbeck, K.T., Riebesell, U. & Reusch, T.B.H. (2012) Adaptive evolution of a key phytoplankton species to ocean acidification. *Nature Geoscience*, **5**, 346–351.

Lundholm, N., Ribeiro, S., Andersen, T., Koch, T., Godhe, A., Ekelund, F. & Ellegaard, M. (2011) Buried alive - germination of up to 100 years-old marine dinoflagellate resting stage. *Phycologia*, **50**, 629–640.

Mantel, N. (1967) The detection of disease clustering and a generalized regression approach. *Cancer Research*, **27**, 209–220.

McQuoid, M.R., Godhe, A. & Nordberg, K. (2002) Viability of phytoplankton resting stages in the sediments of a coastal Swedish fjord. *European Journal of Phycology*, **37**, 191–201.

Nosil, P., Funk, D.J. & Ortiz-Barrientos, D. (2009) Divergent selection and heterogenous genomic divergence. *Molecular Ecology*, **18**, 375–402.

Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H. & Wagner, H. (2011) *vegan: community ecology package*. URL: http://cran.r-project.org.

van Oosterhout, C., Weetman, D. & Hutchinson, W.F. (2006) Estimation and adjustment of microsatellite null alleles in nonequilibrium populations. *Molecular Ecology Notes*, **6**, 255–256.

Park, S.D.E. (2001) *Trypanotolerance in West African cattle and the population genetic effects of selection*. PhD Thesis, University of Dublin

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

Rantajärvi, E. (2003) Alg@line in 2003: 10 years of innovative plankton monitoring and research and operational information service in the Baltic Sea. *MERI report series*, p. 55. Finish Institute of Marine Research, Helsinki. *Eija Rantajärvi*, 36 pp.

Raymond, M. & Rousset, F. (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86**, 248–249.

Rice, W.R. (1989) Analyzing tables of statistical tests. *Evolution*, **43**, 223–225.

Richlen, M.L., Erdner, D.L., McCauley, L.A.R., Libera, K. & Anderson, D.M. (2012) Extensive genetic diversity and rapid population differentiation during blooms of *Alexandrium fundyense* (Dinophyceae) in an isolated salt pond on Cape Cod, MA, USA. *Ecology and Evolution*, **2**, 2588–2599.

Rosenberg, M. & Anderson, C. (2011) PASSaGE: pattern Analysis, Spatial Statistics and Geographic Exegesis. Version 2. *Methods in Ecology and Evolution*, **2**, 229–232.

Rynearson, T.A. & Armbrust, E.V. (2005) Maintainance of clonal diversity during a spring bloom of the centric diatom *Ditylum brightwelli*. *Molecular Ecology*, **14**, 1631–1640.

Rynearson, T.A., Newton, J.A. & Armbrust, E.V. (2006) Spring bloom development, genetic variation, and population succession in the planktonic diatom *Ditylum brightwelli*. *Limnology and Oceanography*, **51**, 1249–1261.

Saravanan, V. & Godhe, A. (2010) Physiologic differentiation and genetic heterogeneity among seasonally separated clones of *Skeletonema marinoi* (Bacillariophyceae) in Gullmar Fjord, Sweden. *European Journal of Phycology*, **45**, 177–190.

Serra, I.A., Innocenti, A.M., Maida, G.D., Calvo, S., Migliaccio, M., Zambianchi, E., Pizzigalli, C., Arnaud-Haond, S. & Duarte, C.M. (2010) Genetic structure in the Mediterranean seagrass *Posidonia oceanica*: disentangling past vicariance events from contemporary patterns of gene flow. *Molecular Ecology*, **19**, 557–568.

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

Smayda, T.J. & Reynolds, C.S. (2001) Community assembly in marine phytoplankton: application of recent models to harmful dinoflagellate blooms. *Journal of Plankton Research*, **23**, 447–461.

Smetacek, V. & Passow, U. (1990) Spring bloom initiation and Sverdrup's critical-depth model. *Limnology and Oceanography*, **35**, 228–234.

Stipa, T. (2002) Temperature as a passive isopycnal tracer in salty, spiceless oceans. *Geophysical Research Letters*, **29**, 1953–1957.

Sunday, J.M., Popovic, I., Palen, W.J., Foreman, M.G.G. & Hart, M.W. (2014) Ocean circulation model predicts high genetic structure observed in long-lived pelagic developer. *Molecular Ecology*, **23**, 5036–5047.

Sundqvist, L., Zackrisson, M. & Kleinhans, D. (2013) *Directional genetic differentiation and asymmetric migration*. arXiv, arXiv:1304.0118 [q-bio.PE]

Tahvanainen, P., Figueroa, R.I., Alpermann, T.J., John, U., Hakanen, P., Nagai, S., Blomster, J. & Kremp, A. (2012) Patterns of post-glacial genetic differentiation in marginal populations of a marine microalga. *PLoS ONE*, **7**, e53602.

Taylor, R., Abrahamsson, K., Godhe, A. & Wängberg, S.-Å. (2009) Seasonal variability in polyunsaturated aldehyde production potential between strains of *Skeletonema marinoi* (Bacillariophyceae). *Journal of Phycology*, **45**, 46–53.

Tesson, S.V.M., Montresor, M., Procaccini, G. & Kooistra, W.H.C.F. (2014) Temporal changes in population structure of a marine planktonic diatom. *PLoS ONE*, **9**, e114984.

Thibert-Plante, X. & Hendry, A.P. (2010) When can ecological speciation be detected with neutral markers? *Molecular Ecology*, **19**, 2301–2314.

Utermöhl, H. (1958) Zur Vervollkommnung der quantitativen phytoplankton-methodik. *Mitteilungen Internationale Vereinigung für Theoretische und Angewandte Limnologie*, **9**, 1–38.

Wasmund, N., Nausch, G. & Matthäus, W. (1998) Phytoplankton spring blooms in the southern Baltic Sea–spatio-temporal development and long-term trends. *Journal of Plankton Research*, **20**, 1099–1117.

White, C., Delkoe, K.A., Watson, J., Siegel, D.A., Zacherl, D.C. & Toonen, R.J. (2010) Ocean currents help explain population genetic structure. *Proceedings of the Royal Society B: Biological Sciences*, **277**, 1685–1694.

Wright, S. (1943) Isolation by distance. *Genetics*, **28**, 114–138.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Appendix S1 Supplementary Tables S1–S8.
Appendix S2 Supplementary Text S1–S2.
Appendix S3 Supplementary Figures S1–S3.

DATA ACCESSIBILITY

Microsatellite sequences: GenBank Accession nos EU855763, EU855769-EU855771, EU855775, EU855777, GQ250935, GQ250937.

BIOSKETCH

The authors belong to a Nordic researcher network PRODIVERSA, which foster collaboration within the Nordic countries for research on patterns, mechanisms and the ecological consequences of intra-specific diversity of aquatic planktonic protists (http://prodiversa14.wix.com/prod#!about_us/csgz).

Author contribution: A.G. and A.K. conceived the idea. A.G., C.S., S.Si, J.S., S.H., M.B-F., C.B., S.G., S.K., N.L., I.S., S.Su, A.K. collected the data. A.G., C.S., J.S., E.J., P.R.J., K.E.R., I.S., L.S. analysed the data. A.G., A.K., K.E.R., N.L., C.L., P.R.J. led the writing. All authors commented on the initial version of the manuscript.

Editor: Michael Dawson