Classification: Biological sciences, Population biology

Genetic structure through time - the sediment archive of a long-lived diatom population

(Final title: Hundred years of genetic structure in a sediment revived diatom population)

Karolina Härnström[1], Marianne Ellegaard[2], Thorbjørn J. Andersen[3] and Anna Godhe[1*]

[1] Department of Marine Ecology, University of Gothenburg, Box 461, SE 405 30 Göteborg, Sweden

[2] Department of Phycology, Biological Institute, University of Copenhagen, Øster Farimagsgade 2D, DK-1353 Copenhagen K, Denmark

[3] Institute of Geography, University of Copenhagen, Øster Voldgade 10, DK-1350, Copenhagen K, Denmark

Corresponding author: anna.godhe@marecol.gu.se

Department of Marine Ecology, University of Gothenburg, Box 461, SE 405 30 Göteborg, Sweden.

Telephone: +46 31 7862708

Fax:+46 31 7862727

Author contribution: K.H., A.G. and M.E. designed research; K.H., A.G. and M.E. performed research; K.H., A.G. and T.J.A. analyzed data; and K.H. and A.G. wrote the paper.

Abstract

This paper presents novel research on the genetic structure and diversity of populations of a common marine protist, and their changes over time. The bloom-forming diatom *Skeletonema marinoi* was used as a model organism. Strains were revived from anoxic discrete layers of a ^{210}Pb dated sediment core accumulated over more than 100 years, corresponding to >40,000 diatom mitotic generations. The sediment core was sampled from the highly eutrophic Mariager Fjord in Denmark. The genetic structure of *S. marinoi* was examined using microsatellite markers, enabling the exploration of changes through time and the effect of environmental fluctuations. The results showed a stable population structure among and within the examined sediment layers, and a similar genetic structure has been maintained over thousands of generations. However, established populations from inside the fjord were highly differentiated from open sea populations. Despite constant water exchange and influx of potential colonizers into the fjord, the populations do not mix. One fjord population, accumulated in 1980, was significantly differentiated from the other groups of strains isolated from the fjord. This could be due to the status of Mariager Fjord, which was considered hypereutrophic, around 1980. There was no significant genetic difference between pre- and post-eutrophication groups of strains. Our data shows that dispersal potential and generation time do not have a large impact on the genetic structuring of the populations investigated here. Instead, the environmental conditions, such as the extreme eutrophication of the Mariager Fjord, are deemed more important.

Key words: Population structure, diatom, microevolution, sediment cores, microsatellites

\body

Introduction

Most planktonic protists sink to the bottom of the sea, die, and subsequently degenerate; however some cells have the ability to survive in the sediment as resting stages. These resting stages can act as short or long term survival mechanisms, with cells remaining viable in the sediment for several decades (1). The formation of resting stages is usually initiated by adverse conditions. Resting stages in the sediment are of ecological importance, as they provide genetic material for future years when resuspended in the water column (2). They are well preserved in anoxic sediment, and in the absence of bioturbation and subsequent laminated sediment, the resting stage-forming species are suitable for microevolutionary studies.

Different local or regional factors, such as climate, nutrient concentrations and oceanographic conditions, have relative importance to the sources of sedimentation. Changes in these factors, including anthropogenic disturbances, are sometimes reflected in the sediment. Aquatic organisms in anoxic sediment cores have been extensively used to assess changes in species composition associated with environmental changes, i.e. eutrophication, shifts in salinity or changes in oxygen concentration (e.g. 3, 4). However, very little is known about changes in population genetic structure during periods of natural shift or anthropogenic-caused changes. Diapausing eggs produced by the water flea *Daphnia* provide an exception. Sub-fossil resting eggs, often viable for up to 100 years and containing sufficient quality DNA for genetic analysis (5), have permitted studies of the effects of eutrophication on genetic structure that are important for the ecology and evolution of fresh water bodies (6). However, there are no equivalent studies in the marine environment, and the effects of environmental changes on intra-specific genetic structure and microevolution of eukaryote protists is poorly understood.

Aquatic protists are considered to have high dispersal ability due to their small sizes and high abundances. Consequently, it is thought that they display low genetic diversity and lack biogeographic patterns (7). Due to predominant asexual division, it has also been assumed that populations consist of only a few clones. This has recently been challenged by an increasing number of studies demonstrating high genetic and phenotypic diversity in populations of eukaryotic microorganisms (8, 9, 10).

We use the chain-forming marine diatom *Skeletonema marinoi* as a model organism. This species is abundant during the spring bloom and often dominates the plankton community in temperate waters (11). *S. marinoi* is an important primary producer and constitutes a valuable food source for higher trophic levels. It usually reproduces asexually, but the formation of auxospores and sexual reproduction has been documented (12). Generation time is short, with approximately one division per day under laboratory conditions (13). It has a benthic resting stage, and in Scandinavian sediments up to 50 000 propagules per gram of sediment can be found (1). *S. marinoi* is easy to collect, isolate, and maintain in culture, and the survival of monoclonal cultures after single cell isolation is almost 100% (14).

Laminated sediment cores from the Mariager Fjord in Denmark have previously been analyzed for ecosystem structural changes over time. Anthropogenic loadings of phosphorus (P) and nitrogen (N) have increased in Danish estuarine waters, especially since the 1950s (15). During the 1970s the input of P in the Mariager Fjord reached 80 tons per year, and an annual average phytoplankton biomass of 60 µg chlorophyll *a* L^{-1} was recorded. In the 1980s, several coastal areas were facing oxygen depletion, and mass mortality of lobsters in the Kattegat was associated with low oxygen concentration (16). This scenario led to legislative actions with the goal to reduce nutrient loading - P by 80%, and N by 50%. P-loads have decreased by 75% in the Mariager Fjord, and the chlorophyll *a* concentration has dropped to 4-12 µg L^{-1} (17). However, the Mariager Fjord is still highly eutrophic in spite of efforts to

reduce nutrient levels in the area, and levels of N, P, and pH are considered high (18). The main biological changes in the fjord, such as shifts in abundance of phytoplankton groups and species, occurred from 1915 to the 1940s (19). This suggests that the fjord has been highly eutrophic since the beginning of the 20th century, but the exact consequences are unknown, especially as regular plankton monitoring was initiated only a few decades ago.

Here, we present a novel approach to examine the genetic structure of a marine protist and its changes over time. We germinated *S. marinoi* strains from temporally discrete layers of a sediment core spanning a period of more than 100 years, and investigated how groups of strains from Mariager Fjord related to the open sea population. We hypothesized that changes in environmental conditions, in particular the nutrient loads, would influence the genetic structure and the intra-sample genetic diversity of temporally discrete populations of *S. marinoi* from the fjord.

Results

The analysed groups of strains (1-7) established from the discrete sediment layers of the core represent a time span of 2 years to more than 100 years. The proportion of survival and successful genotyping of the clonal cultures was on average 66%, but with variations between the sediment layers. The highest survival and genotyping proportion (100%) was documented in the group of strains isolated from the sediment originating from 1980, and the lowest (39%) in the oldest sediment layer (Table 1). In total, 158 monoclonal isolates from the sediment core were genotyped. Additionally, 87 clones from the open sea (Kattegat; Anholt and Vinga, Fig. S1) were established and genotyped.

All eight microsatellite loci were polymorphic. The most variable was S.mar5, with 8 to 21 alleles, and the least variable was S.mar3. Significant ($p<0.05$) departure from HWE was observed for all loci, except S.mar2 and S.mar4, at varying numbers of samples. Loci S.mar1, S.mar5 and S.mar7 displayed heterozygote deficiency in all samples (Table S2). No evidence for large allele drop out or stuttering effects was found, but null alleles might be present in some loci in selected samples. Based on the method by Brookfield (20), estimates of null allele frequencies were low or non-existent in S.mar2, S.mar3, S.mar4 and S.mar6, moderate in S.mar7 and S.mar8, and highest in S.mar1 and S.mar5 (Table S2). The indications of null allele coincided with the loci displaying heterozygote deficiencies (Chi-square test, N=64, $p=0.002$). There was no significant correlation between samples and null allele frequencies (Pair sampled T-test, 2-tailed, $p>0.05$). No pairs of microsatellite loci were significantly linked across all samples.

Genetic differentiation was confirmed for the samples from the open sea (Anholt and Vinga), and all of the samples from the Mariager Fjord (Table 2, Fig. 1 A and B). The magnitude of F_{ST} values of the open sea populations and fjord populations indicated great genetic differentiation. The UPMGA (unweighted pair-grouping method with arithmetic

means) dendrogram based on Nei's genetic distance identified two distinct clusters with 100% bootstrap support: 1) the strains established from the open sea, i.e. Anholt and Vinga; and 2) the strains established from germinated resting stages from discrete sediment layers within the Mariager Fjord. From five independent simulations of the Bayesian clustering method implemented in STRUCTURE, the calculated ΔK indicated that two clusters best explained the uppermost hierarchical level of genetic structuring in the *S. marinoi* samples. The genotyped strains were divided into the fjord samples as Cluster 1, and the Kattegat samples as Cluster 2 (Fig. 1B), consistent with the results of the UPMGA (Fig. 1A). All individuals were assigned with high probability (1.0) to one of the clusters. The exception was one outlier from in the oldest Mariager sample, which was assigned to Cluster 1 with a probability of 0.3.

Sequences of the ITS region displayed a maximum of 0.007 difference per site, which is attributed to intra-specific variation and could not be assigned to the open sea or Mariager Fjord individuals. The ITS secondary structure of the 12 individuals examined folded similarly, and the 30 nucleotide motif of the 5'-side of the ITS2 helix III were identical in all clones investigated.

The analyses of genetic differentiation amongst the Mariager Fjord samples showed significant separations of some groups of strains. Strains established from the 28-year-old sediment layer (sample 4) were significantly differentiated ($p<0.05$) from the strains originating from germinated resting stages older than 87 years (sample 1), 19 years, 8 years, and 2 years (samples 5-7). Strains established from the 87-year-old sediment (sample 3) were significantly differentiated from those strains established from the oldest sediment and from the more recent strains (samples 5 and 6, Table 2). The Factorial Correspondence Analysis (FCA) identified six axes with the eigenvalues of 0.104 (axis 1), 0.087 (axis 2), 0.079 (axis 3), 0.069 (axis 4), 0.067 (axis 5) and 0.052 (axis 6), each explaining 22.67%, 19.02%,

17.31%, 15.02%, 14.7% and 11.28% of the variation. The analysis showed gene flow amongst the Mariager Fjord samples, but also separation, especially along axis 1, in accordance with the weak but significant F_{ST} values of samples 3 and 4 versus the other samples (Fig. 2).

All genotyped individuals isolated from the Mariager Fjord and the Kattegat were genetically distinct (100%). Tests for differences in genetic diversity between the Mariager Fjord and the Kattegat populations showed significantly reduced allelic richness ($p<0.05$), F_{IS}, gene diversity, and significantly higher corrected relatedness within the Mariager Fjord population compared to the open sea (Table 3). The observed heterozygosity (H_O) and relatedness were not significantly different between the two groups. Genetic diversity and the degree of clonality amongst the groups of strains sampled in the Mariager Fjord indicated less diversity within sample 4, established from *S. marinoi* resting stages that settled in the sediment in 1980. This population had significantly lower allelic richness ($p<0.05$), and two significantly ($p<0.05$ after Bonferroni adjustment) linked loci were detected (S.mar1 and S.mar6).

Discussion

We have investigated the genetic structure of *S. marinoi* throughout a sediment archive spanning from recent times to more than 100 years, equivalent to approximately 40,000 diatom mitotic generations. By germinating the resting stages from discrete layers of an anoxic sediment core, and subsequently applying powerful molecular markers, we have revealed the population structure within the fjord. The Mariager Fjord maintains an endemic population and a reduced level of gene flow with adjacent populations from the Kattegat, despite continuous water exchange. Throughout the investigated time period, the genetic structure within the fjord was homogenous, although weak genetic differentiation was also recorded within the fjord.

The viability of the diatom resting stages reported here is at least twice as long as that previously documented, of 65 years (1). The age of the upper layers of the sediment core is most certain, whereas the ^{210}Pb dating is uncertain in the oldest two layers (layers 1 and 2, accumulated pre- 1921). Thus, our results suggest we have encountered and successfully cultured viable cells from sediment as old as 150 years, or more (based on extrapolation of the ^{210}Pb-chronology). Survival rate of the strains from the germinated resting stages was high, and most strains were able to resume growth when placed in an illuminated nutrient enriched medium. The highest proportion of strain survival was recorded in the 28 years old sediment accumulation, whereas the germinated strains from the older layers displayed less viability, in agreement with previous studies of other aquatic organisms (21, 22).

S. marinoi reproduces mainly asexually; however high levels of genetic diversity characterizing the populations in this study imply occasional sexual reproduction. Sexual reproduction acts to diversify diatom populations. The frequency of sexual reproduction varies probably amongst different species and populations (23), and therefore the contribution made by the reproductive modes is difficult to estimate. The reproduction strategy has a

significant effect on the allele and genotype frequencies, but the level of heterozygosity will be influenced only when the proportion of asexually reproducing individuals is very high (24). In populations with mainly asexual propagation, large population sizes, high growth rates, and short generation time, genotypic diversity is maintained even if the proportion of sexually derived individuals is very low (25). The populations analyzed here all displayed heterozygote deficiency, possibly due to the mode of reproduction and non-random mating, but also, especially in some loci, to the presence of null alleles.

Our results indicate strong genetic differentiation between the *S. marinoi* populations in the Mariager Fjord and the Kattegat, in spite of weak, or a lack of, physical dispersal barriers. Genetic differentiations of similar magnitudes between populations of seemingly well-connected habitats have previously been confirmed for *S. marinoi* (22), for a different diatom (26), and for a few other marine protists (10, 27). However, the most remarkable outcome is that diatom populations can be extremely long lived, and patterns of genetic structure are well conserved. Groups of strains established from revived resting stages from layers 1, 2, 5, 6 and 7, corresponding to sediment accumulated over a century, could not be distinguished. Thus, in the Mariager Fjord, despite constant water exchange and the influx of new potential colonizers from the Kattegat, the populations do not mix, and most likely have not done so for thousands of generations. Investigations of the secondary structure of the ITS sequence predict meaningful inter-crossing ability between the fjord and the open sea populations, with a constant inflow of Kattegat individuals obtaining close enough proximity to interbreed. The paradox of reduced gene flow despite high dispersal capacities in aquatic organisms has been recorded for multi-cellular animals (cladocerans, rotifers, bryozoans) and macrophytes in aquatic habitats (28, 29). High genetic differentiation between well-connected habitats can be explained by rapid population growth after a historical founder event, enhanced by a large propagule bank buffering against new immigrants, and the rapid adaptation of the resident

population to local conditions. This effectively increases the persistence of the founder event, and initial colonizers will strongly dominate the genetic structure through priority effects that preclude subsequent colonization (30). Because *S. marinoi* reproduces mainly asexually, dividing approximately once per day, selection for individual clonal lineages can be swift, resulting in rapid population differentiation. In addition, a rich seed bank with continuously germinating propagules will enhance the effect of genetic differentiation. Thus, the genetic stability of the Mariager Fjord population over thousands of generations suggests that the original population has adapted to the local environment, and the large population sizes in the plankton and benthos will effectively outcompete later invaders. An additional and complementary explanation for the complete separation between the two neighboring populations could be attributed to the mean of the propagation. During the diatom lifecycle, the cell size is reduced due to the way it divides. Some species can restore their cell size through vegetative cell enlargement, but for the majority, sexual reproduction is a required stage to avoid critical cell sizes and subsequent death (31). Models of population growth and the timing of cell size-dependant sexual reproduction suggests that a small proportion of the parental populations (0.05%) must reproduce at least every four years to avoid local extinction (32). We therefore hypothesize that differential selection pressure coupled to founder effects, infrequent sexual reproduction, and perhaps different periods of sexual induction prevent the open sea cells from reproducing and colonizing the fjord.

Genetic diversity amongst the individuals from the Mariager Fjord was significantly reduced compared to the open sea population. Ecologically marginal environments accommodate populations that experience extreme selection regimes and contain individuals that are genetically impoverished (33). The occurrence of eutrophication processes is known to be associated with reduced diversity (34), and serves as an explanation here, as Mariager Fjord is hypertrophic. Increased frequency of asexual reproduction reduces genetic variation,

and among facultative asexual species, populations living in marginal environments tend to reproduce asexually to a large extent (35, 36). Thus, reduced frequency of sexual reproduction within the Mariager Fjord as a consequence of the perturbated environment, may explain the reduced genetic diversity compared to the open sea population.

Some groups of strains isolated from the discrete sediment layers of the Mariager core were weakly but significantly differentiated (Table 2). Strains originating from the 1980 sediment accumulation (sample 4) were significantly differentiated from all other groups of strains, except those isolated from samples 2 and 3. Four groups of strains (1, 2, 6 and 7) could not be distinguished. The reason for the similarities between strains isolated from the two most recent sediment layers can be attributed to the short time span between them, with the possibility that they overlap in time. The other similarity, between samples 1 and 2, and 6 and 7, lacks the strong coupling in time, but is the more interesting and challenging. All of the revived strains are no less than decades younger than those isolated from sample 1. What the planktonic cells of 1and 2 and 6 and 7 may have in common are the environmental conditions at the time they formed resting stages and sank to the bottom of the fjord. Nutrient loading had not reached an extreme when resting stages of samples 1 and 2 accumulated (19). During the accumulation of resting stages in sample 6 (2000) and 7 (2006) the nutrient loadings, especially phosphorus, had started to drastically decrease due to successful efforts to reduce nutrient concentrations by the end of the 1980s. Because of uncertain dating in pre-1921 sediment layers, no environmental data can be extrapolated from other analyzed sediment cores from the Mariager Fjord. The lower germination and culturing success from the older sediment layers also adds to the speculative nature of any conclusions. However it is possible that similar environmental factors occuring more than 100 years apart in the same geographic area select from a pool of propagules where all clones are represented, and specific genotypes shared by samples 1, 2, 6 and 7 are advantageous in an environment without a nutrient

excess. The distribution of individuals along the FCA axis one (Fig 2) may be an effect of the environmental conditions in the fjord. Individuals originating from sample 3 (1921) and 4 (1980) are positioned on the extreme right of axis one. These are probably a remnant of the corresponding planktonic population that was influenced by eutrophication. Increasingly, during this period the N and P load was reported as extreme (19). The reduced genetic diversity among the strains isolated from sample 4, relative to the other group of strains, further indicates that the *S. marinoi* population was living in an extreme environment during this time. Information in coding loci would have been advantageous for establishing our hypothesis of differential selection in sample 4; however we did not have any coding loci to compare with the non-coding loci. Neutral microsatellite loci will not initially be affected by divergent selection. However, given enough time in divergent environments, especially if more extensive asexual reproduction is present and at larger population sizes, neutral microsatellites will also become differentiated. This is particularly true in markers linked to selected loci, but provided the given parameters (environmental divergence, asexual reproduction, large populations) dependence on the selected marker becomes weaker (37).

We have investigated the genetic structure of *S. marinoi* over a century, and found that the population in a Danish fjord is genetically conserved. Dispersal potential via water exchange had no impact on the genetic structuring of the populations investigated. Instead, environmental conditions, such as the extreme eutrophication of the Mariager Fjord, is perhaps more important. With the exception of a period when the fjord was highly eutrophic and a weakly differentiated population displaying reduced genetic diversity was dominant, the *S. marinoi* population of the Mariager Fjord is genetically homogenous.

Material and Methods

The sampling sites were located in Mariager Fjord on the east coast of Jutland, Denmark, and in the open sea of the Kattegat (Fig. S1). The narrow sill-fjord (42 km long) is connected to the Kattegat, and has high primary production and a permanently anoxic basin (19). Tidal movement controls the exchange of saline water between the fjord and the Kattegat. Salinity is generally stable (20 PSU) in the deeper, inner part of the Mariager Fjord. The eastern part of the fjord is shallow (<6 m), and the maximum depth is 30 m. Thus, the shape of the fjord causes little water exchange, and the deep water in the inner part of the fjord is renewed on average once every 17 months (17). As the deeper part of the inner fjord is anoxic, there is no infauna and therefore no bioturbation. The Mariager Fjord has a sparse mesozooplankton community, and periodically intense blooms of *S. marinoi* (38).

Sediment cores were collected on January 21, 2008, in Mariager Fjord at 30 m depth (56°40.696'N, 10°01.411'E) using a modified HAPS corer. The cores were kept dark at 4°C. On February 2, a single core (MF08/II) was sliced into 1 cm samples from the bottom up in order to avoid smearing from younger to older layers. The outer few mm of each slice were removed. The remains were transferred to zip-lock plastic bags, and stored in the dark at 4 ºC to prevent resting cells from germinating. Sub-samples of every 1 cm layer (1 to 29 cm core depth) were analyzed for ^{210}Pb, ^{226}Ra and ^{137}Cs activity via gamma-spectrometry at the Gamma Dating Centre, Department of Geography, University of Copenhagen. The measurements were carried out on a Canberra ultralow-background Germanium detector. ^{210}Pb was measured by way of its gamma-peak at 46.5 keV, ^{226}Ra by way of the granddaughter ^{214}Pb (peaks at 295 and 352 keV) and ^{137}Cs by way of its peak at 661 keV. Contents of unsupported ^{210}Pb quickly decreased with depth in the sediment core to levels below the detection limit at depths deeper than 10 cm. The calculated flux of unsupported ^{210}Pb was low, 18 Bq m^{-2} y^{-1}. This indicated that the site might have experienced temporary

erosion. The content of ^{137}Cs was high in the top of the core and decreased quickly with depth. The profile showed similarities to that of the unsupported ^{210}Pb, also indicating that the deposition profiles, to some extent, might be influenced by downward transport of surface material, or mobility of ^{137}Cs. The chronology of the core was established using a slightly modified CRS (Constant rate of Supply)-model (39). The activity below 10 cm is calculated on the basis of a regression of unsupported ^{210}Pb versus accumulated dry mass. The chronology is shown in Fig. S2. Monoclonal cultures of *S. marinoi* were established from seven discrete sediment layers from the Mariager Fjord core (Table 1). The layers for germination were selected to represent the periods before, during, and after nutrient excess were discharged to the Mariager Fjord. Recent open sea sediment from the Kattegat (Anholt, depth 53 m, 56°40.00'N, 12°07.00'E) and (Vinga, depth 77 m, 57°33.00'N, 11°31.50'E) was collected once at each location (Fig. S1). A box corer was used to collect undisturbed sediment samples. The top (<1 cm) of the sediment cores was scraped off and kept at 4°C in the dark.

To establish monoclonal strains of *S. marinoi*, approximately 1 g of sediment from each sediment sample was distributed in smaller aliquots and inoculated into 24 wells NUNC plates. The wells were filled with f/2 medium (40), 26 PSU. The sediment slurries were kept at 10°C on a 12:12 h light:dark cycle at an irradiance of 60 µmol photons m^{-2} s^{-1}. Slurries were examined daily for germination and vegetative growth of *S. marinoi* using an inverted microscope (Axiovert 135, Zeiss). One cell chain from each well was isolated by micropipetting, then transferred to a Petri dish (Ø 50mm) with 5 ml of f/2 medium, and incubated under the same conditions as above. When growth was confirmed, the cells were transferred to 50 ml NUNC flasks, with 10 ml of f/2 medium. After further propagation, another 40 ml of medium was added. Monoclonal cultures in the exponential phase were filtered onto 3 µm pore size filters (Ø 25mm, Versapor®-3000, Pall Corporation) and kept at -

80°C.

Genomic DNA was extracted from the monoclonal cultures following a phenol-chloroform based protocol described in (41). DNA was re-suspended in 50 µl of sterile milli-Q water. Eight microsatellite loci (S.mar 1-8) were amplified (42). Polymerase chain reactions (PCR) were run in total volumes of 24 µl consisting of approximately 2 ng template DNA, PE-PCR-buffer, 0.33 µM of each primer, 250 µM of each dNTP, 0.04 U of Taq polymerase (AmpliTaq Gold), and sterile milli-Q water. Amplification conditions and allele size assignment were performed as specified in (22).

The internal transcriber spacer 1 (ITS1), 5.8S rRNA gene, and the ITS2 were sequenced from four isolates from the Mariager Fjord (strains originating from samples 2, 3, 4 and 6); four from Anholt; and four isolates from Vinga. The DNA fragment from the 5' end of the 5.8S ending in the conserved motif of the helix III of the ITS2 has been proposed for barcoding of diatoms (43), and the identity of a 30 nucleotide motif in the highly conserved region of the 5'-side of the ITS2 helix III predicts meaningful intercrossing ability among protists (44). The nuclear encoded ITS region was amplified and sequenced by following the protocol of (14). Consensus sequence for each clone was produced and aligned for comparison using Sequencher 4.1.2 (Gene Codes Corporation, Ann Arbor, MI, USA). The secondary structure of the ITS was analyzed as described in (45). Blast searches (http://www.ncbi.nlm.nih.gov/BLAST) were performed on all obtained DNA sequences to compare the identity with *S. marinoi* sequences in the database. The sequences have been deposited in GenBank (GenBank accession numbers HM582852-63).

Deviations from the Hardy-Weinberg equilibrium (HWE) at each locus in each sample (10000 Markov Chain dememorizations, 20 batches and 5000 iterations per batch), genotypic linkage disequilibrium (LD) between pairs of loci in each sample, and genotypic LD between pairs of loci for all samples (10000 dememorizations, 100 batches and 5000 iterations per

batch) were estimated using Genepop v. 4.0.7 (46). The microsatellite dataset was examined for integer or inconsistent modulus, null alleles and stuttering using MicroChecker 1.0 (47) (95% confidence interval and 1000 randomizations). To investigate if null alleles (20) could explain deviation from the HWE, the relationship between the null allele frequencies and the P-values from the HWE tests was examined using Chi-square tests.

Genetic differentiation between pairs of populations was determined by calculating pair-wise F_{ST} using Genetix version 4.05 (48) with 1000 permutations, and P-values corrected using the Bonferroni test. An UPGMA cluster of the Mariager Fjord and the Kattegat populations was constructed based on Nei's genetic distance (49), using Tools For Population Genetic Analyses TFPGA v. 1.3 (50). Bootstrap values were based on 1000 permutations. Bayesian analysis, as implemented in Structure version 2.2 (51), was used to gain further insight into the gene flow between different populations. We assessed the number of potential clusters (K) among the nine *S. marinoi* samples from five different runs at each K ranging from 1 to 9, and 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 100 000 generations and 100 000 Markov chain Monte Carlo replications were applied. We used the model of no-admixture ancestry, and the assumption of correlated allele frequencies as suggested in (52). 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. ΔK assists in determining the uppermost hierarchical level of structure (53). A Factorial Correspondence Analysis (FCA) was carried out using genotypic data, which displays the structural relationships among the Mariager Fjord populations and individuals using Genetix 4.05 (48).

Microsatellite Tools for Excel (54) was used to search for identical genotypes. Tests for differences in genetic diversity between the Mariager Fjord and the Kattegat populations were examined using Fstat (allelic richness, observed heterozygosity H_O, F_{IS}, gene diversity

H_S, relatedness and corrected relatedness, 1000 permutations). The extent of genetic diversity within the Mariager Fjord samples was assessed by allelic richness corrected for sample size, H_O, H_S and F_{IS} and estimated for each sample using Fstat (55). Significant differences ($p<0.05$) in genetic diversity between the Mariager Fjord groups of strains were estimated with pair sampled T-tests performed in SPSS version 16.0 for Mac (SPSS Inc.).

Acknowledgement

We thank Lovisa Jansson, Jenny Egardt and Dr Adil Yusuf Al-Handal for assistance in the laboratory, Prof Carl André, University of Gothenburg, for help with statistics, Dr Tatiana Rynearson, University of Rhode Island, and two anonymous reviewers for constructive comments on the manuscript. The fragment analysis was performed at RSKC, University Hospital MAS. The ITS sequencing was done at the Genomics Core Facility, Sahlgrenska Academy, University of Gothenburg, by Dr Elham Rekabdar. This project was supported by Sida (SWE-2004-129), the Swedish Research Council Formas (2006-1892), Danish Research Council (2111-04-0011), SYNTHESYS $_{(DK\text{-}TAF)}$, European Community-RI Action ASSEMBLE (227799), Kapten Carl Stenholms Donationsfond, Stiftelsen Lars Hiertas Minne, KVVS, Magnus Bergvalls Stiftelse, Stiftelsen Oscar and Lilli Lamms Minne.

References

1. McQuoid MR, Godhe A, Nordberg K (2002) Viability of phytoplankton resting stages in the sediments of a coastal Swedish fjord. *Eur J Phycol* 37**:** 191-201.
2. Hairston NG, Kearns CM, Ellner SP (1996) Phenotypic variation in a zooplankton egg bank. *Ecology* 77**:**2282-2392.
3. Dale B, Thorsen TA, Fjellså A (1999) Dinoflagellate cysts as indicators of cultural eutrophication in the Oslofjord, Norway. *Estuar Coast Shelf Sci* 48**:**371-382.
4. Karlsen AW, Cronin TM, Ishman SE, Willard DA, Holmes CW, et al. (2000) Historical trends in Chesapeake Bay dissolved oxygen based on benthic foraminifera from sediment cores. *Estuaries* 23**:**488-508.
5. Hairston NG, Vanbrunt RA, Kerrns CM, Engstrom DR (1995) Age and survivorship of diapausing eggs in a sediment egg bank. *Ecology* 76**:**1706-1711.
6. Brede N, Sandrock C, Straile D, Spaak P, Jankowski T, et al. (2009) The impact of human-made ecological changes on the genetic architecture of *Daphnia* species. *Proc Natl Acad Sci USA* 106**:**4758-4763.
7. Finlay BJ (2002) Global dispersal of free-living microbial eucaryote species. *Science* 296**:**1061-1063.
8. Rynearson TA, Armbrust EV (2005) Maintenance of clonal diversity during a spring bloom of the centric diatom *Ditylum brightwellii. Mol Ecol* 14**:**1631-1640.
9. Evans KM, Chepurnov VA, Sluiman HJ, Thomas SJ, Spears BM, et al. (2009) Highly differentiated populations of the freshwater diatom *Sellaphora capitata* suggest limited dispersal and opportunities for allopatric speciation. *Protist* 160**:**386-396.
10. Nagai S, Nishitani G, Sakamoto S, Sugaya T, Lee C, et al. (2009) Genetic structuring and transfer of marine dinoflagellate *Cochlodinium polykrikoides* in Japanese and Korean coastal waters revealed by microsatellites. *Mol Ecol* 18**:**2337-2352.

11. Kooistra WHCF, Sarno D, Balzano S, Gu H, Anderson RA, et al. (2008) Global diversity and biogeography of *Skeletonema* species (Bacillariophyta). *Protist* 159**:**177-193.

12. Migita S (1967) Sexual reproduction of centric diatom *Skeletonema costatum. Bull Jap Soc Sci Fish* 33**:**392-398.

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

14. Godhe A, McQouid MR, Karunasagar I, Karunasagar I, Rehnstam-Holm A-S (2006) Comparison of three common molecular tools for distinguishing among geographically separated clones of the diatom *Skeletonema marinoi* Sarno et Zingone (Bacillariophyceae). *J Phycol* 42**:**280-291.

15. Clarke AL, Weckstrom K, Conley DJ, Anderson NJ, Adser F, et al. (2006) Long-term trends in eutrophication and nutrients in the coastal zone. *Limnol Oceanogr* 51**:**385-397.

16. Diaz RJ, Rosenberg R (2008) Spreading dead zones and consequences for marine ecosystems. *Science* 321**:**926-929.

17. Fallesen G, Andersen F, Larsen B (2000) Life, death and revival of the hypertrophic Mariager Fjord, Denmark. *J Mar Syst* 25**:**313-321.

18. Hansen PJ (2002) Effect of high pH on the growth and survival of marine phytoplankton: implications for species succession. *Aquat Microb Ecol* 28**:**279-288.

19. Ellegaard M, Clarke AL, Reuss N, Drew S, Weckström K, et al. (2006) Multi-proxy evidence of long-term changes in ecosystem structure in a Danish marine estuary, linked to increase nutrient loading. *Estuar Coast Shelf Sci* 68**:**567-578.

20. Brookfield JFY (1996) A simple new method for estimating null allele frequency

from heterozygote deficiency. *Mol Ecol* 5**:**453-455.

21. Brendonck L, De Meester L (2003) Egg banks in freshwater zooplankton: evolutionary and ecological archives in the sediment. *Hydrobiologia* 491**:**65-84.

22. Godhe A, Härnström K (2010) Linking the planktonic and benthic habitat: genetic structure of the marine diatom *Skeletonema marinoi. Mol Ecol* 19**:** 4478-4490.

23. Mann DG (1993) Patterns of sexual reproduction in diatoms. *Hydrobiologia* 269/270**:**11-20.

24. Balloux F, Lehmann L, de Meeus T (2003) The population genetics of clonal and partially clonal diploids. *Genetics* 164**:**1635-1644.

25. Bengtsson BO (2003) Genetic variation in organisms with sexual and asexual reproduction. *J Evol Biol* 16**:**189-199.

26. Rynearson TA, Newton JA, Armbrust EV (2006) Spring bloom development, genetic variation, and population succession in the planktonic diatom *Ditylum brightwellii. Limnol Oceanogr* 51**:**1249-1261.

27. Alpermann T, Beszteri B, John U, Tillmann U, Cembella AD (2009) Implications of life-history transitions on the population genetic structure of the toxigenic marine dinoflagellate *Alexandrium tamarense. Mol Ecol* 18**:**2122-2133.

28. Palsson S (2000) Microsatellite variation in *Daphnia pulex* from both sides of the Baltic Sea. *Mol Ecol* 9**:**1075-1088.

29. Campillo S, Garcia-Roger EM, Carmona MJ, Goméz A, Serra M (2009) Selection on life-history traits and genetic population divergence in rotifers. *J Evol Biol* 22**:**2542-2553.

30. De Meester L, Goméz A, Okamura B & Schwenk K (2002) The monopolization hypothesis and the dispersal–gene flow paradox in aquatic organisms. *Acta Oecologia* 23**:**121-135.

31. Round FE, Crawford R, Mann DG (1990) *The diatoms* (Cambridge University Press, Cambridge).

32. D'Alelio D, d'Alcala M, Dubroca L, Sarno D, Zingone A, et al. (2010) The time for sex: A biennial life cycle in a marine planktonic diatom. *Limnol Oceanogr* 55**:**106-114.

33. Johannesson K, & André C (2006) Life on the margin: genetic isolation and diversity loss in a peripheral marine ecosystem, the Baltic Sea. *Mol Ecol* 15**:**2013-2029.

34. Kerfoot WC, Weider LJ (2004) Experimental paleoecology (resurrection ecology): Chasing Van Valen's Red Queen hypothesis. *Limnol Oceanogr* 49**:**1300-1316.

35. Eckert CG (2002) The loss of sex in clonal plants. *Evol Ecol* 15**:**501-520.

36. Tatarenkov A, Bergström L, Jönsson RB, Serrao EA, Kautsky L, et al. (2005) Intriguing asexual life in marginal populations of the brown seaweed *Fucus vesiculosus. Mol Ecol* 14**:**647-651.

37. Thibert-Plante X, Hendry AP (2010) When can ecological speciation be detected with neutral markers? *Mol Ecol* 19**:**2301-2314.

38. Tiselius P, Borg CMA, Hansen BW, Hansen PJ, Nielsen TG, et al. (2008) High reproduction, but low biomass: mortality estimates of the copepod *Acartia tonsa* in a hyper-eutrophic estuary. *Aquat Biol* 2**:**93-103.

39. Appelby P (2001) in *Tracking environmental change using lake sediments. Volume 1: Basin analysis, coring and chronological techniques*, eds. Last W, Smol JP (Kluwer Academic Publishers), pp 171-203.

40. Guillard RRL (1975) in *Culture of Marine Invertebrate Animals*, eds. Smith W, Chanley, M (Plenum Press, New York), pp 29-60.

41. Godhe A, Asplund ME, Härnström K, Saravanan V, Tyagi A, et al. (2008) Quantifying diatom and dinoflagellate biomass in coastal marine sea water samples

by real-time PCR. *Appl Environ Microbiol* 74**:**7174-7182.

42. Almany GR, De Arruda MP, Arthofer W, Atallah ZK, Beissinger SR, et al. (2009) Permanent Genetic Resources added to Molecular Ecology Resources Database 1 May 2009-31 July. 2009 *Mol Ecol Resour* 9:1460-1466.

43. Moniz MBJ, Kaczmarska I (2010) Barcoding of diatoms: Nuclear encoded ITS revisited. *Protist* 161**:**7-34.

44. Coleman AW (2009) Is there a molecular key to the level of "biological species" in eukaryotes? A DNA guide. *Mol Phylogenet Evol* 50**:**197-203.

45. Mai JC, Coleman AW (1997) The internal transcribed spacer 2 exhibits a common secondary structure in green algae and flowering plants. *J Mol Evol* 44:258-271.

46. Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *J Hered* 86**:**248-249.

47. Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Mol Ecol Notes* 4**:**535-538.

48. 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).

49. Nei M (1972) Genetic distance between populations *The American Naturalist* 106**:**283-292.

50. Miller M (1997) *Tools for Population Genetic Analyses (TFPGA) version 1.3. A Windows® program for the analysis of allozyme and molecular population genetic data* (Department of Biological Sciences, Northern Arizona University, Flagstaff).

51. Pritchard JK, Stephens M, Donnelly P (2000) Interference of population structure using multilocus genotype data. *Genetics* 155**:**945-959.

52. Falush D, Stephens M, Pritchard JK (2003) Inference of population genetic structure: Extensions to linked loci and correlated allele frequencies *Genetics* 164**:**1567-1587.

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

54. Park S (2001) Trypanotolerance in West African cattle and the population genetic effects of selection (University of Dublin).

55. Goudet J (2001) *Fstat, a program to estimate and test gene diversities and fixation indices (version 2.9.3.2)*, http://www2.unil.ch/popgen/softwares/fstat.htm.

Figure legends

Figure 1. Population structure of *S. marinoi* based on 8 microsatellite loci. Samples 1-7 are groups of strains established from resting stages germinated from discrete layers of a sediment core collected in the Mariager Fjord (approximate year when resting stages accumulated within parentheses). Samples Anholt and Vinga are groups of strains established from recent sediment collected in the Kattegat. (A) An unweigthed pair grouping method using an arithmetic average (UPGMA) of clonal cultures established from the 9 different groups of strains based on Nei's genetic distance (49). Bootstrap values are indicated on each node. (B) Genetic clustering estimated by STRUCTURE. Assignment of 245 individuals to K=2 genetically distinguishable groups. Each individual is represented by a vertical bar coloured according to the assigned group, and the average proportions of membership of each sample to the two clusters are shown below.

Figure 2. FCA of the genetic distances based on a plot using F_{ST} values of the individual strains belonging to the seven groups of strains established from resting stages in discrete sediment layers accumulated in Mariager Fjord from prior to 1921 to 2006 (samples 1 to 7 with approximate year when resting stages accumulated within parentheses). Axes 1, 2 and 3 explain 59% of the variance in distribution of the individuals from the discrete sediment layers.

Table legends

Table 1. Chronology of sediment core (MF08/ll) with the depths of the analyzed sediment layers, the approximate ages, and proportions of survival and genotyped *S. marinoi* clonal isolates. All details of the two deepest layers are not shown, due to uncertain dating, but are more than 87 years old.

Table 2. Genetic differentiation between pairs of samples (F_{ST}) from the Mariager Fjord and the open sea (Kattegat; Anholt and Vinga). Bold numbers indicate significant differentiation after Bonferroni correction ($p<0.05$).

Table 3. Test for significant difference of genetic diversity between the Mariager Fjord (N=158) and the Kattegat (N=87) populations.

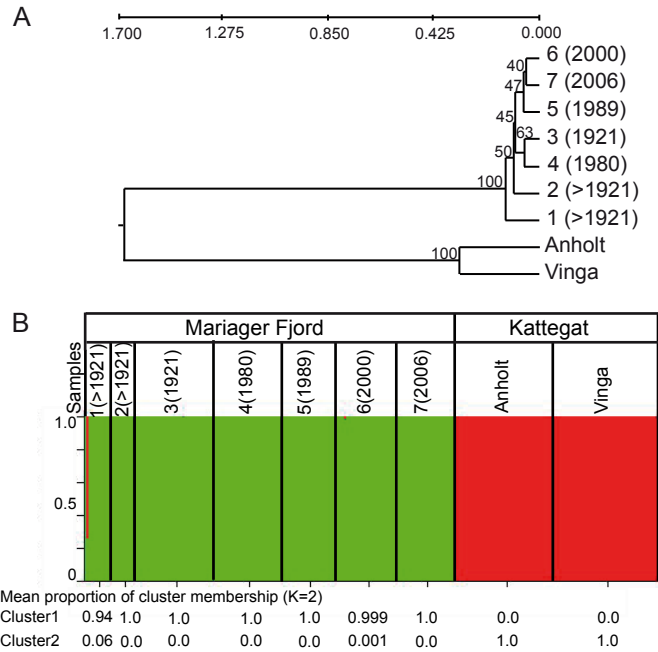
A
1.700
1.275
0.850
0.425
0.000
6 (2000)
7 (2006)
5 (1989)
3 (1921)
4 (1980)
2 (>1921)
1 (>1921)
Anholt
Vinga
40
47
45
63
50
100
100
B
Mariager Fjord
Kattegat
Samples
1(>1921)
2(>1921)
3(1921)
4(1980)
5(1989)
6(2000)
7(2006)
Anholt
Vinga
1.0
0.5
0
Mean proportion of cluster membership (K=2)
Cluster1 0.94 1.0 1.0 1.0 1.0 0.999 1.0 0.0 0.0
Cluster2 0.06 0.0 0.0 0.0 0.0 0.001 0.0 1.0 1.0

Fig. 1 A and B

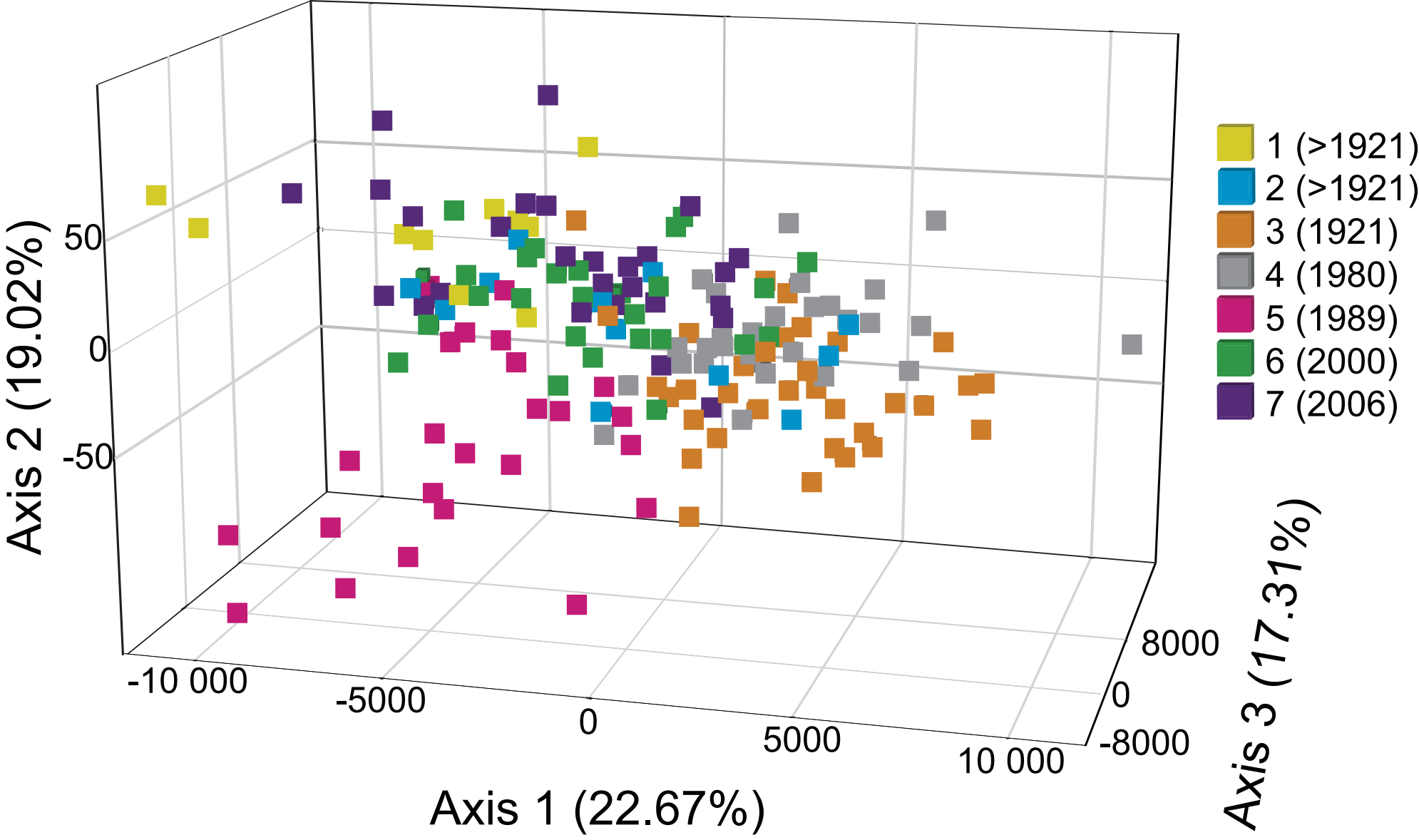
1 (>1921)
2 (>1921)
3 (1921)
4 (1980)
5 (1989)
6 (2000)
7 (2006)
Axis 2 (19.02%)
50
0
-50
-10 000
-5000
0
5000
10 000
Axis 1 (22.67%)
8000
0
-8000
Axis 3 (17.31%)

Table 1. Chronology of the sediment core (MF08/ll) with the depths of the analyzed sediment layers, approximate ages, and proportions of survival and genotyped *Skeletonema marinoi* clonal isolates. Full details of the two deepest layers are not shown due to uncertain dating, but are more than 87 years old.

Depth (cm)	Age (y)	Error age (y)	Date	Sediment accumulation rate (kg m^{-2} y^{-1})	Error rate (kg m^{-2} y^{-1})	Label of strains established	Proportion of genotyped isolates (%)
0.0			2008				
1.0	2	2	2006	0.11	0.02	7	68
2.5	8	2	2000	0.11	0.02	6	79
4.5	19	3	1989	0.09	0.02	5	64
5.5	28	3	1980	0.07	0.01	4	100
6.5	41	4	1967	0.06	0.01		
7.5	57	5	1951	0.05	0.01		
8.5	72	5	1936	0.06	0.01		
9.5	87	7	1921	0.06	0.01	3	62
15-16	>87					2	50
21-22	>87					1	39

Table 2. Genetic differentiation between pairs of samples (F_{ST}) from the Mariager Fjord and the open sea (Kattegat; Anholt and Vinga). Bold numbers indicate significant differentiation after Bonferroni correction ($p<0.05$).

	Sample 2	Sample 3	Sample 4	Sample 5	Sample 6	Sample 7	Anholt	Vinga
Sample 1	-0.018	**0.030**	**0.038**	-0.003	-0.010	0.007	**0.225**	**0.224**
Sample 2	–	0.005	0.007	-0.014	-0.015	0.002	**0.241**	**0.241**
Sample 3		–	0.005	**0.018**	**0.013**	0.007	**0.247**	**0.242**
Sample 4			–	**0.022**	**0.016**	**0.018**	**0.268**	**0.260**
Sample 5				–	-0.013	-0.004	**0.224**	**0.224**
Sample 6					–	-0.011	**0.224**	**0.224**
Sample 7						–	**0.224**	**0.227**
Anholt							–	**0.071**

Table 3. Test for significant difference of genetic diversity between Mariager Fjord (N=158) and Kattegat (N=87) populations.

Diversity index	Differences	P-value
Allelic Richness	Mariager<Kattegat	0.023
F_{IS}	Mariager<Kattegat	0.029
H_S	Mariager<Kattegat	0.004
H_O	Mariager=Kattegat	NS
Relatedness	Mariager=Kattegat	NS
Corrected relatedness	Mariager>Kattegat	0.042