











Adaptive divergence across Southern Ocean gradients in the pelagic diatom *Fragilariopsis kerguelensis*

Ute Postel^{1,2,3}  | Barbara Glemser¹  | Katherine Salazar Alekseyeva^{1,4}  | Sarah Lena Eggers¹  | Marco Groth⁵  | Gernot Glöckner²  | Uwe John^{1,6}  | Thomas Mock⁷  | Kerstin Klemm¹ | Klaus Valentin¹  | Bábk Beszteri^{1,3} 

¹Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Bremerhaven, Germany

²Institute for Biochemistry I, University Cologne, Cologne, Germany

³University of Duisburg-Essen, Essen, Germany

⁴Department of Functional and Evolutionary Ecology, University of Vienna, Vienna, Austria

⁵Leibniz Institute on Aging - Fritz Lipmann Institute (FLI), Jena, Germany

⁶Helmholtz Institute for Functional Marine Biodiversity, Oldenburg, Oldenburg, Germany

⁷University of East Anglia, Norwich Research Park, Norwich, UK

Correspondence

Ute Postel, University of Duisburg-Essen, Universitätsstrasse 5, 45141 Essen, Germany.

Email: ute.postel@uni-due.de

Funding information

Deutsche Forschungsgemeinschaft, Grant/Award Number: BE4316/6-1, BE4316/7-1, GL235/20-1, JO702/10-1 and VA105/24-1; Open access funding enabled and organized by Projekt DEAL

Abstract

The Southern Ocean is characterized by longitudinal water circulations crossed by strong latitudinal gradients. How this oceanographic background shapes planktonic populations is largely unknown, despite the significance of this region for global biogeochemical cycles. Here, we show, based on genomic, morphometric, ecophysiological and mating compatibility data, an example of ecotypic differentiation and speciation within an endemic pelagic inhabitant, the diatom *Fragilariopsis kerguelensis*. We discovered three genotypic variants, one present throughout the latitudinal transect sampled, the others restricted to the north and south, respectively. The latter two showed reciprocal monophyly across all three genomes and significant ecophysiological differences consistent with local adaptation, but produced viable offspring in laboratory crosses. The third group was also reproductively isolated from the latter two. We hypothesize that this pattern originated by an adaptive expansion accompanied by ecotypic divergence, followed by sympatric speciation.

KEYWORDS

adaptive divergence, diatom, pelagic, population structure, Southern Ocean, speciation

1 | INTRODUCTION

The Southern Ocean (SO) surrounds Antarctica in a continuous belt linking the three main ocean basins. With the exception of the Weddell and Ross gyres, circumpolar currents dominate surface water flow: while in the north, the powerful Antarctic Circumpolar Current (ACC) carries water masses eastward, easterlies along the

coast drive southern water masses to the west. These mainly circumpolar currents are crossed by strong meridional gradients in several environmental conditions of importance for planktonic life, including temperature, light and macronutrient concentrations (Pollard, Lucas, & Read, 2002; Post et al., 2014). Environmental heterogeneity is an important driver of diversification (Rainey & Travisano, 1998), and in general, environmental gradients have been proposed to facilitate

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2020 The Authors. Molecular Ecology published by John Wiley & Sons Ltd

adaptive divergence and speciation (Doebeli & Dieckmann, 2003; Endler, 1977). In the particular oceanographic context of the Southern Ocean, however, little is known about how this oceanographic background shapes population structure of phytoplankters.

The pennate diatom *Fragilariopsis kerguelensis* (O'Meara) Hustedt 1952 is one of the best-studied phytoplankton organisms of this region. It has been one of the most important silicate exporting organisms in the pelagic of the SO during the Quaternary (Cortese & Gersonde, 2008) to today, playing a key role in the oceanic silicon cycle (Assmy et al., 2013). Its frustules represent the dominant diatom in terms of valve numbers in the siliceous sediment layers below the ACC (Crosta, Romero, Armand, & Pichon, 2005; Esper, Gersonde, & Kadagies, 2010; Gersonde et al., 2003) and are applied to draw inferences about paleo-environmental conditions (Cortese & Gersonde, 2007; Cortese, Gersonde, Maschner, & Medley, 2012; Kloster, Kauer, Esper, Fuchs, & Beszteri, 2018; Shukla, Crosta, Cortese, & Nayak, 2013; Shukla & Romero, 2018). *F. kerguelensis* is a congener of *Fragilariopsis cylindrus*, the genome of which has recently been sequenced (Mock et al., 2017). Although the biogeographic ranges of both species overlap, *F. cylindrus* is mainly at home in the more southern, colder, temporarily ice-covered zone (Armand, Crosta, Romero, & Pichon, 2005). In contrast, *F. kerguelensis* appears along the entire range of the environmental gradients of the SO, from latitudes displaying subantarctic/temperate climatic conditions (up to near 40°S, a latitude comparable to that of the Italian city Naples in the Northern hemisphere; with summer surface water temperatures near 10°C; ice free all year round), all the way to the year-round subzero water temperatures next to the shelf ice edge (near and beyond 70°S, Crosta et al., 2005; Pinkernell & Beszteri, 2014), thus providing an interesting model to study population differentiation to different temperature and light regimes along these gradients.

Here, we hypothesized that the marked differences in environmental conditions between the northern and southern fringes of its distribution area could lead to an ecophysiological and genomic differentiation between northern (ACC) versus southern, near-ice-edge populations of *F. kerguelensis*. To test this hypothesis, populations were sampled along a meridional transect, ranging from close to the northern boundary of the distribution range of the species near the Subantarctic Front (Pinkernell & Beszteri, 2014), to the shelf ice edge. We applied morphometric, physiological and genome-wide molecular comparisons and tested mating compatibilities for a detailed characterization of patterns of diversity and differentiation.

2 | MATERIAL AND METHODS

2.1 | Cultures

Strains of *Fragilariopsis kerguelensis* were isolated during expedition PS103 on board RV Polarstern between 16.12.2016 and 3.2.2017. Samples were taken with a 20 µm mesh size plankton net from 20 m depth to the surface. Single *F. kerguelensis* chains were picked under an inverted microscope (ZEISS Axiovert 40) into 12-well cell

culture plates filled with filter-sterilized F/2 medium (Guillard & Rytter, 1962) cooled to 4°C (85%–95% of isolated clones per station grew up, indicating a low cultivation bias). In order to minimize contamination, every chain was isolated again in the same manner for at least four further times. The cultures were incubated at 4°C under illumination with cool white fluorescent tubes at an intensity between 15 and 30 µmol/m²s under a 16:8 hr light:dark cycle. Strains were transferred to 15 ml Falcon tubes for cooled transport and transferred into 40 ml culture flasks within a week after arrival at the home laboratory. Strains were held at 4°C at 20–30 µmol/m²s and a 14:10 hr daily light:dark cycle. Table S1 lists the strains used and their sampling metadata.

2.2 | Morphometrics

To test whether strains can be classified in the high versus low rectangularity morphotypes discovered by Kloster et al. (2018), morphometric analyses were performed following that reference. Frustules were prepared using household bleach following Friedrichs (2013) and mounted in Norland Optical Adhesive 61 (Norland Products Inc.). Manually selected valves were imaged in 20 focus levels, and the latter combined to extended depth-of-focus images with a Metafer slide scanning system (Metasystems GmbH). Morphometric measurements were performed by image analysis using SHERPA (Kloster, Kauer, & Beszteri, 2014). Statistical comparison of genotypic groups (see Results) was performed in a mixed model framework in R v. 3.5.1. (R-Core-Team, 2015).

2.3 | DNA extraction

Within a time frame of three to ten months after their arrival in the laboratory, cultures were grown in 200 ml cell culture flasks and harvested using polycarbonate membrane filters (5 µm pore size, Whatman® Nuclepore™, Merck). After rinsing with sterile medium, cells were washed from the filters with SL1 lysis buffer, snap-frozen in liquid nitrogen and stored at –20°C until DNA extraction with the NucleoSpin Soil DNA extraction kit (Macherey-Nagel). After thawing, cells were disrupted by bead beating for 2 × 30 s. The subsequent DNA extractions followed the manufacturer instructions of the kit, and the extracted DNA was stored at –20°C. Quality check was performed using a NanoDrop ND-1000 spectrophotometer (Peqlab) and gel electrophoresis.

2.4 | Ribosomal ITS sequencing

The nuclear ribosomal internal transcribed spacer (ITS) region was amplified using primers ITS1 and ITS4 (White, Bruns, Lee, & Taylor, 1990). PCR amplicons were purified with the High Pure PCR Cleanup Micro Kit (Roche) and directly sequenced on both strands using the BigDye Terminator v3.1 cycle sequencing kit. Products of

sequencing reactions were cleaned up with Agencourt CleanSEQ-Dye Terminator Removal kit (Beckman Coulter) and electrophoresed on an Applied Biosystems 3130xl Genetic Analyzer (Applied Biosystems). Sequences were quality clipped using Pregap4 and assembled and manually checked using Gap4 from the Staden package (Staden, 1996). Positions with secondary peaks were recorded as ambiguities. Multiple alignment and comparative analyses of the ITS sequences were performed using MEGA7 (Kumar, Stecher, & Tamura, 2016).

2.5 | ddRAD analyses

Forty-eight strains from four sampling stations were used for double digest restriction site-associated DNA sequencing, ddRAD (Peterson, Weber, Kay, Fisher, & Hoekstra, 2012), library preparation following the protocols of Hülskötter (2015), Peterson et al. (2012) and Schweyen, Rozenberg, and Leese (2014) (see Figure S1 and Supporting Information for further details).

Preprocessing and demultiplexing of the sequences was performed using the scripts by Schweyen et al. (2014) which also control for biased amplification. The software Stacks Version 2.0 Beta 8c (Catchen, Hohenlohe, Bassham, Amores, & Cresko, 2013) was used to process the data in a de novo approach, combining paired-end reads to single loci. Based on test runs with a subsample of 12 randomly selected strains, a parameter setting of $M=n=5$ was selected for the final analyses following Rochette and Catchen (2017).

Single nucleotide polymorphism (SNP) calls were filtered with the *populations* program of Stacks to include loci shared by at least 80% of individuals, and only keeping one SNP per locus (to avoid linked SNPs), resulting in 1,301 SNPs used for downstream analyses. SNP clustering based on average linkage clustering of an identity by state (IBS) pairwise distance matrix was performed using R 3.2.3 (R-Core-Team, 2015) employing the R-packages *gdsfmt* and *SNPrelate* (Zheng et al., 2012). The software *Structure* 2.3.4. (Falush, Stephens, & Pritchard, 2003, 2007; Hubisz, Falush, Stephens, & Pritchard, 2009; Pritchard, Stephens, & Donnelly, 2000) was used to analyse population structure and to estimate the amount of admixture. The number of populations was determined following Evanno, Regnaut, and Goudet (2005).

2.6 | Low coverage genome sequencing and variant analysis in high copy number regions

To get broader genomic representation, ten strains across all ITS/ddRAD groups were selected for low coverage shotgun sequencing, alongside an older isolate of *F. kerguelensis* (strain L2-C3, isolated on 24.1.2009 at 48.1°S, 24.25°W, also on board RV Polarstern, by Philipp Assmy). Sequencing was performed on one lane of the Illumina HiSeq platform in a Rapid run as 250 bp paired-end reads for the 10 strains from PS103, and as 300 bp PE reads on MiSeq for L2-C3.

The analyses focused on high copy number genomic sequences: the nuclear rDNA operon, and the mitochondrial and plastid genomes. A reference sequence of these three regions was assembled from a single strain (L2-C3 for the rDNA locus, strain 1-09 for both organellar genomes) by running the targeted assembler NOVOPlasty, version 2.6.3 (Dierckxsens, Mardulyn, & Smits, 2017). Publicly available sequences from *Pseudo-nitzschia multiseriis* (KR149143.1) and *Cylindrotheca closterium* (KC509522.1) were used as seeds for the mitochondrial and plastid genomes, respectively, and an ITS sequence determined in this study for the rDNA region.

Read mapping to these reference sequences and variant calling from the 11 sequencing libraries was performed following the protocol "wgs/wes mapping to variant calls," version 1.0 ("WGS/WES Mapping to Variant Calls - Version 1.0," 2018). For a more reliable inference of the phylogenetic history of the variants, a representative "haplotype" for each strain was generated by extracting only the most abundant variant at each SNP position. These sequences were aligned together with an outgroup sequence from *F. cylindrus* using *progressiveMauve* (Darling, Mau, Blattner, & Perna, 2004) and used for phylogenetic analyses with *PhyML* with model selection with *PhyML-SMS* (Guindon et al., 2010; Lefort, Longueville, & Gascuel, 2017) and 1,000 bootstrap replicates.

2.7 | Mating experiments

For testing mating compatibilities between pairs of strains, 250 µl dense late-exponential phase culture of two strains was added to 1.5 ml culture medium and mixed. The crosses were inspected every 2–3 days with an Axiovert 35 inverted microscope over the course of three weeks, and sexual activities were recorded as no sexual activity; detachment (i.e. disintegration of chains or separation of individual cells from them, the first sign of sexualization in *F. kerguelensis*; Fuchs, Scalco, Kooistra, Assmy, & Montresor, 2013); and appearance of auxospores and initial cells. At the end of the experiment, the category corresponding to the furthest progression of sexual reproduction observed was recorded for each cross.

2.8 | Common garden experiment

Twenty-seven strains (nine representatives from each genotypic group collected at the northernmost and southernmost sampling stations) were inoculated simultaneously into two 40 ml culture flasks each. The paired sets of cultures were pre-acclimated to the same daily light cycle (14:10 hr) and illumination intensity (15–30 µmol/m²s, in the following referred to as low light, LL), but different temperatures (6°C and 0°C, respectively) for three weeks.

Subsequently, each of these flasks served to inoculate two fresh cultures, one of which was transferred to an elevated illumination intensity (87–121 µmol/m²s, in the further referred to as high light, HL) at the same temperature, whereas the other was left at low light (growth phase g1). Cultures were kept at exponential growth by

inoculating fresh cultures at the end of the exponential growth phase (growth phase g2). Culture growth at the four conditions (HL and LL, both at 6 and 0°C) was monitored by chlorophyll fluorometry using an Imaging PAM (MAXI Version of M-Series, Heinz Walz GmbH). The dark and maximal fluorescence yield (F_0 and F_m , respectively) were measured after 15 min dark acclimation every 48 to 72 hr, between 6 and 8 hr after “day start” with $9.8 \mu\text{mol}/\text{m}^2\text{s}$ measuring light (ML) intensity applied at 1 Hz. For the measurement of F_m , a saturation pulse of approximately $1,300 \mu\text{mol}/\text{m}^2\text{s}$ (800 ms) was given. The fluorescence images were evaluated using the R package *EImage* (Pau, Fuchs, Sklyar, Boutros, & Huber, 2010). The maximum photosystem II quantum yield (F_v/F_m) of the cultures were calculated as

$$F_v F_m = F_m - F_0 F_m.$$

To determine the growth rates (μ), a logistic model using the R package *growthcurver* (Sprouffske & Wagner, 2016) was fitted to the F_0 values. F_v/F_m values at half maximal F_0 were used as measure for the comparison of photosynthetic fitness among treatments and genotypes.

Differences in the performance between genotypes during the same growth period (g1) at each growth condition were analysed using one-way ANOVA combined with Tukey multiple comparison of means as post hoc test. Where the data failed to be normally distributed, a Kruskal–Wallis rank sum test was used with a Wilcoxon rank sum test as post hoc test. For within-genotype comparison of treatment effects, a paired *t*-test or, in case data were not normally distributed, a paired Wilcoxon signed rank test was used. In a few cases, a strain did not survive at a certain period and condition. These strains were excluded from the respective comparison.

All sequences, submitted with assistance from the German Federation for Biological Data, GFBio (Diepenbroek et. al., 2014), are available under European Nucleotide Archive (ENA) project PRJEB32688 (Beszteri, Glöckner, John, Valentin, & Postel, 2020) (Table S1).

3 | RESULTS

3.1 | Population structure

Nuclear ribosomal internal transcribed regions (ITS) were sequenced from strains collected at the northernmost and southernmost sampling station. Three parsimony-informative positions were detected among the 28 strains sequenced (at positions 481 variants: G/K/T; 571: T/W/A; and 703: C/G; Table S1), all located within the sequence of the ITS2. Additionally, a heterozygous R call was shared by four strains (14-01, 14-02, 14-26 and 14-29) at position 471, as opposed to a G in all others. At eight further positions, singleton variants were detected. Three different nucleotide combinations at the parsimony-informative positions were observed (TTG, TAG and GAC), alongside with two genotypes containing an ambiguity (KAC and TWG). Further variants (GAG, TTC or GTG) which could be expected to occur with random mating and free recombination among strains were not observed.

The mitochondrial genome was assembled as a single circular contig of 37,348 bp in length (Figure S2). Read coverages of the 11 shotgun libraries mapped against this mitochondrial reference genome ranged between 91 ± 37 -fold (strain L2-C3) and 820 ± 83 -fold (strain 3.51b). One thousand seven hundred and thirty seven variant positions were identified, including 100 indels and 1,637 SNPs, 32 of which were multiallelic and excluded from downstream analysis. A cluster analysis of the 1,605 biallelic SNP calls revealed three distinct groups of strains, in accordance with the three ITS groups described above, and clearly grouping the ambiguous ITS genotypes with one nonambiguous group and not in intermediate positions. For instance, strain 14-23 (TWG) was grouped with the TTG ITS group, and strain 1-51 (KAC) grouped with the GAC strains.

The plastid genome was assembled as three contigs with lengths of 131,691 bp (contig 1), 39,794 bp (contig 2) and 39,510 bp (contig 3), respectively. Contig 1 was found to represent the large single copy region (53,998 bp) flanked on each side by one of the two inverted repeat regions (38,795 bp each). Contigs 2 and 3 were reverse complements, corresponding to the single copy region (SSC, 38,792 bp). The pre- and suffixes of each of the three contigs were formed by reverse complement strings of between 246 and 573 bp in length. Since these mutually match either the suffix or prefix of both of the remaining two contigs, the directionality of the SSC within the circular plastid genome cannot be determined unequivocally; the sequence used for further analyses represents one randomly selected ordering. This full plastid reference genome had a length of 170,380 bp (Figure S3), and the mean sequencing coverage ranged between 237 ± 35 (strain 14-27) and $1,809 \pm 266$ -fold (strain 1-51). 1,319 variant positions were identified, comprising 76 indels and 1,243 SNPs, of which 1,155 were biallelic. Clustering of the latter again led to a division into three distinct groups, which were in full congruence with both the ITS and the mitochondrial sequence data. Clustering order within each of these three clusters was, however, different between the mitochondrial and the plastid data sets (Figure 1).

The reference sequence of the nuclear rDNA operon assembled from the same shotgun data sets had a length of 10,223 bp and included the small 18S rRNA subunit, the 5.8S subunit and the 28S rRNA large subunit as well as the nontranscribed spacer regions (Figure S4). Sequencing coverage of the rDNA operon ranged from $1,334 \pm 350$ -fold (L2-C3) to $5,127 \pm 596$ -fold (14-01). 236 variable positions were found, composed of 48 indels, two multiallelic SNPs and 186 biallelic SNPs. Clustering of the latter again revealed a clustering into three groups in full congruence with the ITS, mitochondrial and plastid based groupings (Figure 1).

To extend a broader genomic sampling to a wider selection of strains, ddRAD genotyping was performed. Three of the originally 48 strains, collected at four sampling stations, were excluded from final analyses. Two of these due to a low number of loci obtained (28 and 1,359, <3% of the median count of 46,880, Figure S5), and the third strain was an outlier in terms of its genotype, probably belonging to the morphologically similar species *Fragilariopsis ritscheri* Hust, after verification of its morphology. From the ddRAD sequences, 1,301 unlinked SNPs were extracted. Clustering of these genotypes

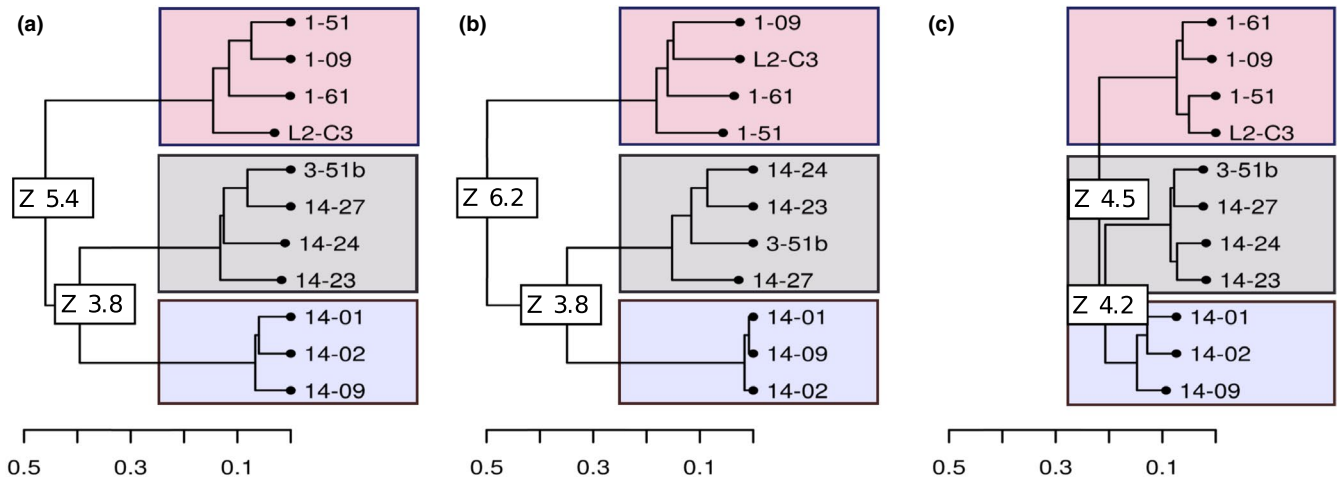


FIGURE 1 Cluster analysis of SNPs demonstrating consistent clustering into an “omnipresent (grey), a northern (red) and a southern (blue) genotypic group, using (a) whole mitochondrial and (b) plastid genomes, and (c) nuclear ribosomal operon sequences. SNPs identified by mapping of shotgun sequencing reads from eleven strains. Cluster analyses are based on 1,605 SNPs of 37,348 reference positions (a); 1,155 SNPs of 170,380 reference positions (b); and 186 SNPs out of 10,223 reference positions of the nuclear ribosomal operon (c), respectively. Tip labels indicate strain IDs; (see Table S1). Permutation Z-scores are given at the nodes

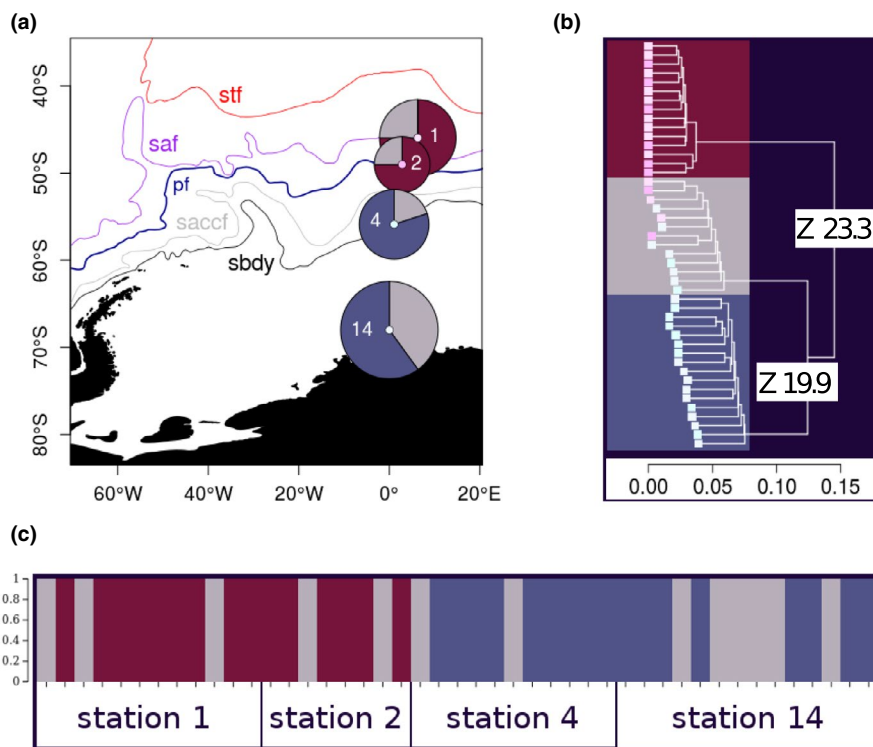


FIGURE 2 ddRAD based (a) genotype composition of sample sites (numbers in pie charts indicate station codes, see Table S1); (b) cluster; and (c) structure analysis. (a) Composition of sampled strains resolved to the three genotypic groups (NG: red, OG: grey and SG: blue) at the four sampling stations studied (size of the pie charts proportional to number of strains analysed); stf: Subtropical Front; saf: Subantarctic Front; pf: Antarctic Polar Front; saccf: Southern ACC Front; and sbdy: Southern Boundary of ACC. (b) cluster analysis of ddRAD genotypes reveals three genotypic clusters in full congruence with mitochondrial and plastid genomes and rDNA operons. Permutation Z-scores are given at the nodes. (c) Structure run with an admixture model indicates a lack of introgression among the three genotypic groups

again revealed three clusters (Figure 2b), in full congruence with those obtained from the ITS, mitochondrial, chloroplast and nuclear rDNA regions.

To obtain rooted phylogenies of the high copy number regions, a consensus sequence was generated for each locus and for each strain, via the mapping of reads to the respective reference sequence, and

used for phylogenetic analysis with outgroup sequences. The three genotypic groups were clearly resolved as separate also in these analyses (Figure 3). These phylogenetic analyses consistently revealed the split between the ancestor of the "blue" (Southern) group versus the ancestor of the other two to precede the separation of the latter.

Analysing the ddRAD genotypes with Structure (Pritchard et al., 2000), the Evanno method (Evanno et al., 2005) also indicated the presence of three populations, with the three groups again corresponding to those resolved by clustering and all other previous analyses. When analysing the genomic contribution of each of these three populations to the multilocus genotypes of each strain with a model allowing admixture, no signal of introgression was found (Figure 2c), indicating that no detectable amount of exchange of genetic information among the three genotypic groups identified took place in the recent past.

The biogeographic distribution of representatives of the three genotypic groups across the four sampled populations showed the following pattern: whereas one group (marked in grey in Figures 1–3, in the further termed OG for "omnipresent" genotype) was found at all four sampling locations, another group (marked in red, from here on termed NG for northern genotype) was only recovered from the two Northern populations, and the third group (blue, further on termed southern genotype or SG) only from the two Southern sampling stations (Figure 2a).

3.2 | Morphometric differentiation

Strains with the ITS pattern TAG belonging to the southern genotype displayed a significantly higher valve outline rectangularity of 0.793 ± 0.014 ($n_{\text{strains}} = 9$, $n_{\text{individuals}} = 348$, "SG") when compared to ITS group GAC ("NG," 0.748 ± 0.014 , $n_{\text{strains}} = 7$, $n_{\text{individuals}} = 280$) or ITS group TTG ("OG," 0.761 ± 0.016 , $n_{\text{strains}} = 8$, $n_{\text{individuals}} = 244$),

whereas both latter groups did not differ significantly from each other (Figure 4).

3.3 | Mating compatibility

Similarly to Fuchs et al. (2013), we observed a heterothallic pattern of sexual compatibility among *Fragilariopsis kerguelensis* strains and inferred mating types from successful crosses. In crosses involving pairs of strains from the same genotypic group belonging to different mating types, auxospore production was mostly observed (Figure 5). To the contrary, we did not observe a single successful mating in crosses involving one partner from the OG and another from either of the other groups SG or NG. However, successful auxospore production was repeatedly observed in crosses including one mating partner from the SG and one from the NG group. The auxospores and initial cells resulting from these latter crosses looked as viable as those from within-group crosses and continued dividing just as vividly. Due to the long time required to reach the sexually inducible size range, we have so far not been able to test whether these presumed "hybrids" are fertile.

3.4 | Common garden experiments

The physiological performance of the three genotypic groups during the first (g1) and second growth periods (g2) under each of the four conditions was monitored by their growth rates (μ) and maximal photosystem II photon yields (F_v/F_m).

The different genotypes showed some common patterns in their responses to the four investigated conditions. During g1, all three genotypes grew fastest at 6°C HL (Figure 6, Figure S6). F_v/F_m was

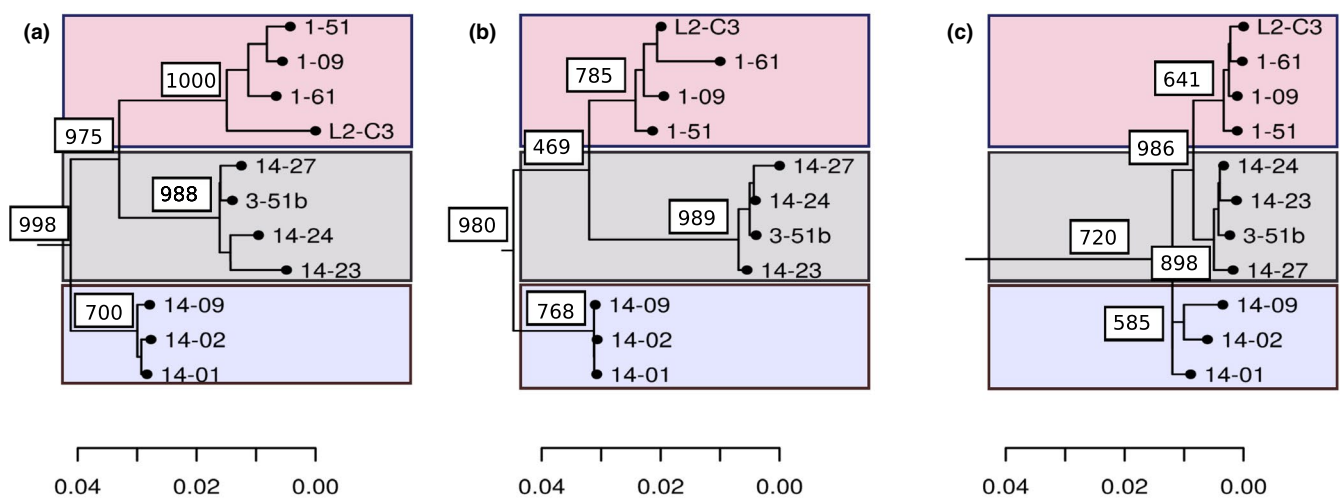


FIGURE 3 Rooted phylogenies of (a) mitochondrial and (b) plastid genome and (c) nuclear rDNA operon nucleotide alignments. For rooting the ingroup phylogeny, homologous sequences from *Fragilariopsis cylindrus* were used. In the figure, the outgroup was removed and only its connection to the ingroup is indicated to allow a better resolution of branch lengths within the ingroup. Branches formed by strains (indicated by their respective strain IDs; see Table S1) belonging to the northern (red), the omnipresent (grey) and the southern (blue) genotypic group are indicated by the colour coding of their background boxes, respectively. Bootstrap values (clade resolved in number of bootstrap replicates out of 1,000) are given at the nodes.

significantly higher at LL when compared with HL conditions at both temperatures in all genotypic groups (Figure S6). Strains from all genotypes seemed to be challenged by growing at 0°C LL, with all genotypes showing a significant reduction in both μ and F_v/F_m in g2 when compared with g1 (Figure 6).

A clear difference observed between the three genotypic groups was a significantly lower μ of the northern genotype ($\mu = 0.19 \pm 0.07 \text{ day}^{-1}$; corresponding to a generation time of $4.0 \pm 1.4 \text{ days}$, $n = 9$) under 0°C HL when compared with the southern genotype ($\mu = 0.34 \pm 0.8 \text{ day}^{-1}$ corresponding to a generation time of $2.2 \pm 0.5 \text{ days}$, $n = 9$, $p = 4.3e^{-4}$). The SG showed a significant increase in their F_v/F_m in g2 versus g1 at 0°C HL ($p = .037$, Figure 6), indicating a pronounced acclimation capacity to these conditions. While the SG performed in general best at 0°C, while struggling under 6°C conditions, the NG performed and acclimated better at 6°C: they showed stable μ and increased F_v/F_m from g1 to g2 under both 6°C conditions, while those of the SG and OG declined significantly (Figure 6). In contrast to both the NG and the SG, none of the four growing conditions seemed to be favourable to the OG, with significantly decreasing μ in g2 versus g1 under all four experimental conditions (Figure 6).

4 | DISCUSSION

Our molecular results using different markers with a broad genomic coverage (including both organellar and nuclear genomes) and an altogether relatively broad sampling (48 strains from four populations) revealed a robust clustering of *Fragilariopsis kerguelensis* strains into three distinct groups in a fully concordant manner. Admixture analyses of 1,301 randomly sampled genomic SNPs from a ddRAD analysis revealed no signal of introgression, indicating that exchange of genetic material among the three genotypic groups was undetectably low in the recent past. To the contrary, phylogenetic incongruence among the three genomes was found within the individual genotypic groups. In light of our knowledge of the biology of the species (Assmy, Henjes, Smetacek, & Montresor, 2006; Fuchs et al., 2013), this probably results from genetic reshuffling by sexual gene exchange in their recent ancestry, that is within each of the populations of the three genotypic groups, which, however, seems to be absent among them.

Using the criterion of absence of gene exchange in sympatry, the OG and SG (found in sympatry at the two southern stations) as well as the OG and NG (sympatric at the two northern sampling stations, Figure 2a) could robustly be diagnosed as pairs of distinct species. This diagnosis is also in line with the lack of mating compatibility observed between these two groups.

On the other hand, the NG and SG were not found in sympatry and produced viable offspring in laboratory crosses, speaking for an interpretation as geographically isolated populations or demes of the same species. Based on the data available thus far, it cannot be excluded that a transitional zone might exist where the geographic distributions of both groups touch or overlap. Admixed genotypes

were nevertheless not encountered in the current survey, indicating that even if such a transitional zone exists, hybrids between both groups must be rather rare apart from a possibly narrow hybrid zone. Another feature underlining the distinctness of NG versus SG is that they can also be distinguished morphometrically. The clear and congruent distinctness of these two groups across all three genomes (the mitochondrial, the plastid and the nuclear genome), and the lack of admixture signals in the ddRAD genotypes thus far indicate a robust genetic divergence and a lack of introgression between the NG and the SG. This appears to be surprising when considering the successful laboratory crosses and the fact that the natural habitats of both genotypes are not separated by any barriers that would limit dispersal. Possible explanations might include rarity or sterility of F1 hybrids, heterozygote fitness disadvantage and/or low dispersal ability of hybrids from a narrow hybrid zone; or the fact that both genotypes indeed rarely occur in sympatry for ecological–oceanographic reasons (blooms and/or sexual events of both groups temporally separated from each other at the same geographic location). The exact nature of the biogeographical and reproductive boundaries between both genotypes in the field remains to be revealed and will be an interesting challenge for future studies.

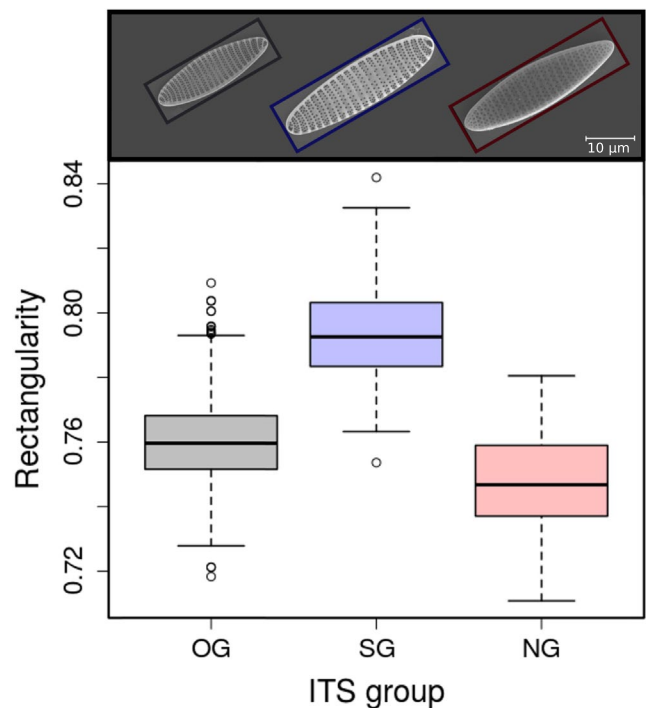


FIGURE 4 Rectangularity of valves across the three ITS groups. OG (grey): omnipresent group, ITS variable positions TTG, 244 valves across eight strains measured (19–46 individual valves per strain); SG (blue): southern group, ITS variable positions TAG, 348 valves across nine strains measured (14–76 individuals per strain); and NG (red): northern group, ITS variable positions GAC, 280 valves representing seven strains were measured (18–89 per strain). The scanning electron microscopy images above the box plots show an example valve for each of the three genotypic groups

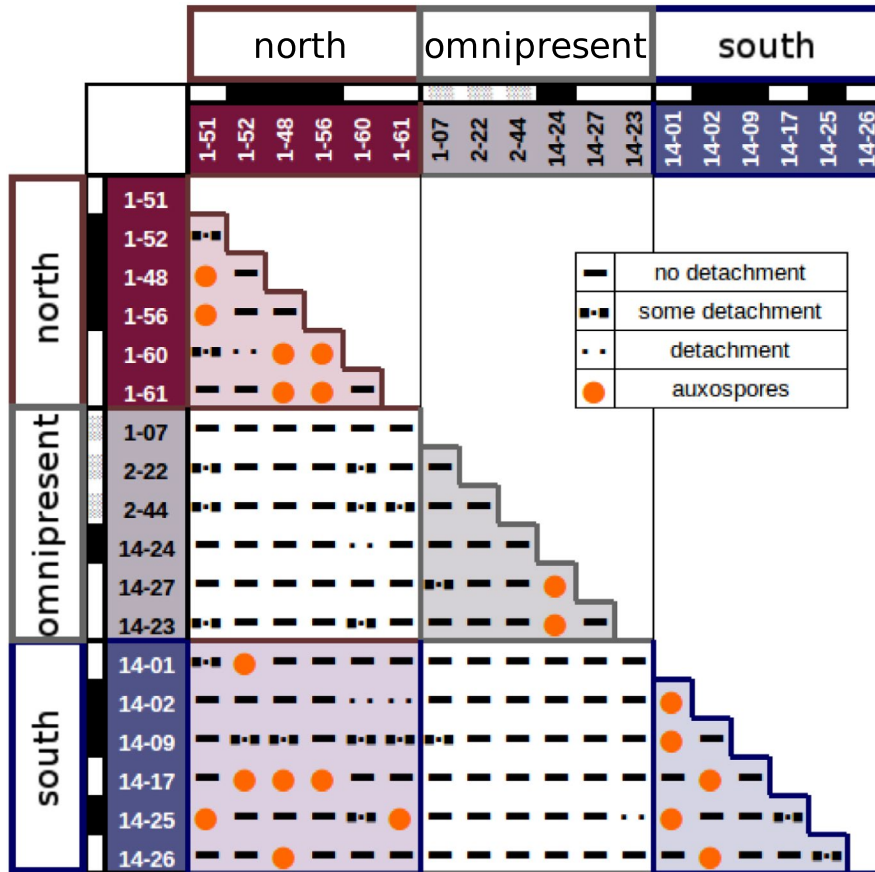


FIGURE 5 Summary of results of crossing experiments. Strains belonging to the three genotypic groups are colour-coded in accord with all other figures, that is northern group – red; omnipresent group – grey; and southern group – blue. No detachment – only vegetative chains observed; some detachment – few detached cells, alongside a majority of chains; detachment – the majority of cells, at least in patches of several dozens, observed as detached single cells; auxospores – elongating auxospores and/or initial cells recorded. Mating types, inferred from mating activities, are indicated at the left and the top by white and black bars, with grey bars indicating strains with unknown mating type

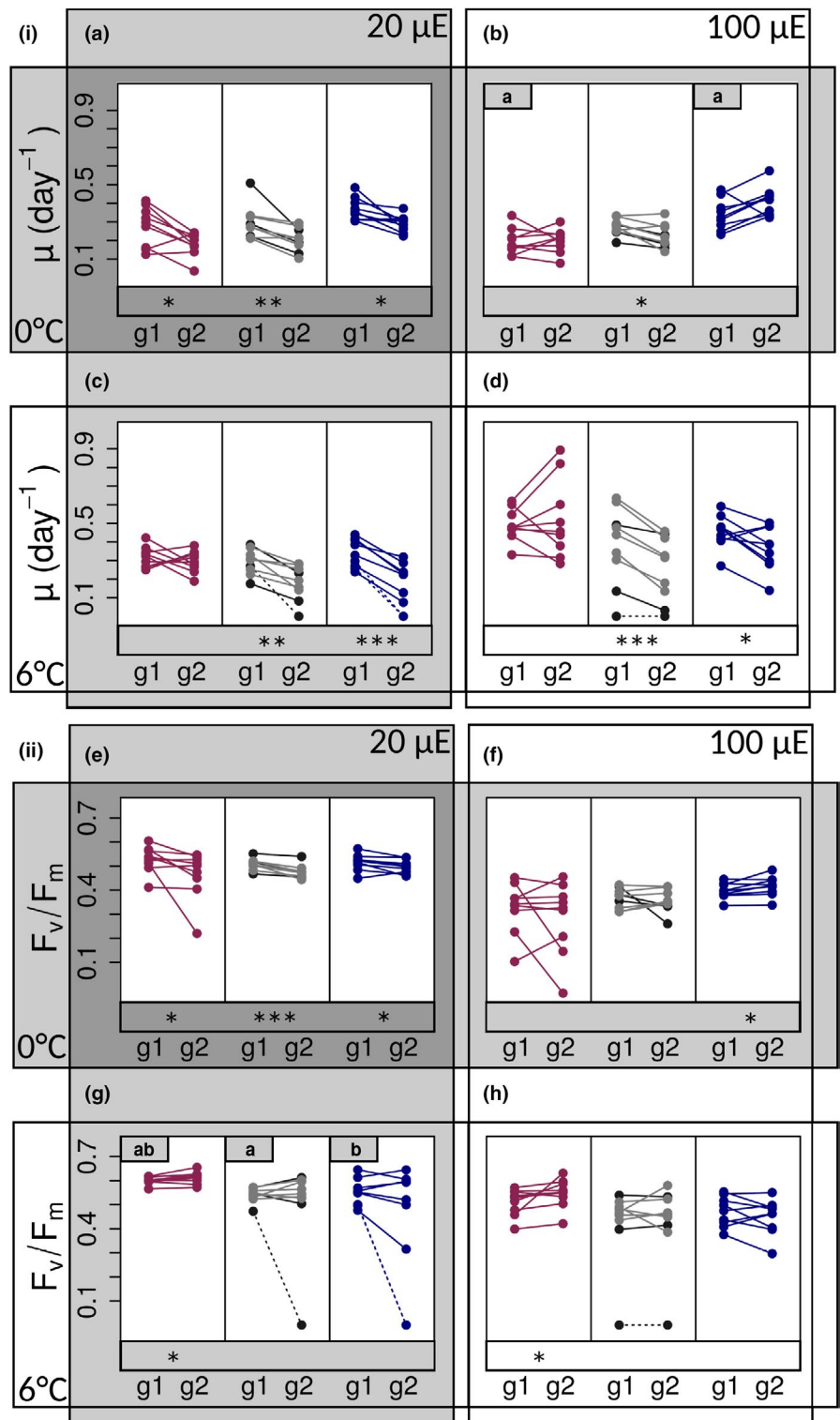
Paradoxically, in spite of the most clear-cut (because sympatric) separation between the OG versus the other two groups, this divergence does not seem to have predated the separation of the SG versus NG groups: in rooted phylogenetic trees, the SG (also morphologically different from the other two) diverged first from the common ancestor of the other two groups (Figure 3). Based on this, one could hypothesize that this SG – common ancestor of NG and OG split – might have taken place as an initial intraspecific differentiation into a high rectangularity population inhabiting southern parts of the Southern Ocean (similar to present-day (SG)) and a low rectangularity genotype/ deme which colonized the more northern, ACC habitat (presumed to roughly correspond to the present-day distribution of the NG). Based on the physiological mating compatibility between the SG and NG, we hypothesize the NG-OG ancestor to have been similar in terms of biogeographic distribution and mating compatibility to the present-day NG. In this case, a subsequent speciation event, presumably in the north (ACC), would have led to a divergence between the NG and the morphometrically highly similar OG groups. The latter would then have recolonized the southern parts of the SO habitat while staying distinct from the there native SG. This might be explained by the possibility that the reproductive barrier, which evolved between OG and NG, also led to a reproductive isolation of OG from the SG deme.

In a common garden experiment, we observed a slight, but significant ecophysiological differentiation among the three genotypic

groups. Strains from the SG showed better performance (growth rates) at and ability to acclimate to cold high light conditions than strains from the NG, along with a significant decline in their growth rate in the warm treatments. To the contrary, strains of the NG showed stable growth rates and an increase in F_v/F_m at 6°C. This ecophysiological differentiation between the two groups, taken together with the tentative biogeographic pattern outlined above (NG – northern and SG – southern), can be interpreted as native advantage indicating local adaptation. The third group (OG, appearing at all latitudes) did not show such clear-cut ecophysiological signal, but seemed to suffer at all experimental conditions (growth rates decreased at every treatment condition) which might indicate a narrower temperature tolerance range enabling a long-term maintenance of good physiological performance when compared to the other two groups (or a lower tolerance than the other two groups against some other aspect of the laboratory growth conditions).

The observed biogeographic–genomic–ecophysiological differentiation between the northern (NG) and southern (SG) demes is somewhat surprising since these locations are not separated by any geographic boundary that would prohibit dispersal, but it is in line with the notion that marked environmental differences along gradients can facilitate divergence and speciation (Doebeli & Dieckmann, 2003; Endler, 1977). Although a continuous transition in environmental conditions, as in the SO pelagic, could also lead to a continuous transition in genetic setup (genetic clines), Endler (1977) showed that it can also lead to abrupt steps in allele frequencies,

FIGURE 6 Comparison of (I) growth rates (μ , expressed as divisions per day) and (II) maximum photosystem II quantum yields (F_v/F_m) during exponential growth of strains from the northern (red), the omnipresent (grey) and the southern genotypic groups (blue), respectively, determined during the first period (g1) and second (g2) period of growth after their transference to one of the four growth conditions 0°C low light (LL; panels a and e) or high light (HL; b and f); and 6°C & LL (c and g) or HL (d and h), respectively ($N = 9$, genetically distinct biological replicates per group; dotted lines with points set to zero indicate cases in which a strain did not grow). Statistically significant differences ($p < .05$) between genotypes at a certain growing condition observed during g1 are indicated at the top (pairs of genotypes marked by identical letters showed a significant difference from each other at the same treatment condition). Significant differences between g1 and g2 of a certain genotype are indicated at the bottom of each plot (* $p < .05$, ** $p < .01$, *** $p < .001$)



for instance with heterozygote disadvantage. Oceanic fronts are usually accompanied by steeper gradients in water properties than water masses lying between them (Sokolov & Rintoul, 2002), and such steepening can facilitate the appearance of sharp genotypic clines and even parapatric speciation (Endler, 1977). In line with this, the Southern Ocean fronts have been described to separate rather distinct plankton communities (Sokolov & Rintoul, 2002). Some fronts might, however, provide more than just steep gradients

to support population differentiation or speciation. A case in point is the Southern Boundary (SB) of the ACC (Orsi, Whitworth, & Nowlin, 1995; Pollard et al., 2002; Talley, 2011), identified previously to follow the biogeographic boundary between the high versus low rectangularity morphotype of *F. kerguelensis* (Glemser et al., 2019). This front or boundary also lies in the area separating the NG and the SG genotypic groups (Figure 2a; see also Figure S7 for oceanographic background), although we cannot pin down exactly the

location of transition between both groups due to our lack of samples in the southernmost band of the ACC. The SB separates the eastward circulation of the ACC from the Weddell Gyre (in our transect; or from the westward flowing Antarctic Coastal Current in other places). Although this boundary is highly permeable in the surface layer, it is also the main place of circumpolar upwelling, driven by opposite directed wind forces, bearing divergent surface waters directed to the north versus south (Post et al., 2014). The upwelling and associated divergent flow pattern can conceivably reduce surface connectivity across this band when compared to that across other fronts of the ACC for instance. In any case, as our study hints at, oceanographic structure can matter for pelagic diversity, refining slightly on the notion that “habitat selection, rather than physical barriers, may be a primary force driving speciation in the pelagic” (Bowen et al., 2016). To resolve in detail how the balance between factors favouring divergence versus homogeneity acts out from the perspective of surface layer eukaryotic microplankton, and whether or how often fronts indeed play special roles in promoting phytoplankton divergence, will need geographically more finely resolved regional surveys of organismal or genotypic distributions, to complement the picture obtained by more global but lower resolution surveys (de Vargas et al., 2015). Although the nature of the transitional zone between the NG and SG groups certainly needs closer study in the future, it is intuitively conceivable that incompatibilities between “warm-adapted” (northern) and “cold-adapted” (southern) genetic backgrounds might indeed play a role in this case, too.

Biogeographic projections suggest that ocean warming will likely lead to poleward shifts in northern distribution boundaries for SO phytoplankton, and also in particular for *F. kerguelensis* (Pinkernell, 2017; Pinkernell & Beszteri, 2014). It makes a difference in this context whether this (and similar other) species are genetically relatively homogeneous but physiologically plastic, ensuring their survival among the wide range of environmental conditions across their distribution range, or genetically structured, with locally adapted sets of genotypes/ demes which might have low fitness in non-native conditions (Villemereuil, Mouterde, Gaggiotti, & Till-Bottraud, 2018). In the former case, the main biogeographic response to ocean warming might be a reversible poleward retraction of the distribution area with little reduction in overall intraspecific genetic variation. If, however, zonal genetic structure and local adaptations are important in these circumpolar plankton, as it seems to be in this particular case, such a biogeographic retraction would result in less reversible losses of genetic variability (in the present case, putting the southern genotypic group into the highest risk of extinction).

ACKNOWLEDGEMENTS

This work was supported by the Deutsche Forschungsgemeinschaft (DFG) in the framework of the priority programme 1158 “Antarctic Research with comparative investigations in Arctic ice areas” by grant nr. BE4316/6-1, GL235/20-1, VA105/24-1, JO702/10-1 and by DFG priority programme 1991 “TaxonOmics” by grant nr. BE4316/7-1. Sample collection was performed in the frame of RV Polarstern

expedition PS103 (grant nr. AWI_PS103_04), thanks for the chief scientist Olaf Boebel and the physical oceanography team for support and CTD handling. Further more we would like to thank Ivaylo Kostadinov for his kind assistance in the process of data submission via GFBio. Open access funding enabled and organized by Projekt DEAL.

AUTHOR CONTRIBUTIONS

The study was planned by B.B., G.G., U.J. and K.V. The samples were collected and the cultures initiated by B.B. and S.L.E. The cultures were maintained by S.L.E. The experiments were designed by B.B. and U.P. The data were collected and the experiments were performed by S.L.E., B.G., K.K., K.S.A., T.M. and U.P. The data were analysed by B.B., B.G., G.G., M.G., K.S.A. and U.P. The paper was drafted by B.B., T.M. and U.P. and written by all authors.


DATA AVAILABILITY STATEMENT

All sequences are available under European Nucleotide Archive (ENA) project PRJEB32688 (Beszteri et al., 2020). For information on sample locations and accession numbers of strain related data sets, see Table S1.

ORCID

Ute Postel  <https://orcid.org/0000-0001-8999-9911>

Barbara Glemser  <https://orcid.org/0000-0002-5747-0244>

Katherine Salazar Alekseyeva  <https://orcid.org/0000-0002-0088-4489>

Sarah Lena Eggers  <https://orcid.org/0000-0001-6094-3201>

Marco Groth  <https://orcid.org/0000-0002-9199-8990>

Gernot Glöckner  <https://orcid.org/0000-0002-9061-1061>

Uwe John  <https://orcid.org/0000-0002-1297-4086>

Thomas Mock  <https://orcid.org/0000-0001-9604-0362>

Klaus Valentin  <https://orcid.org/0000-0001-7401-9423>

Bánk Beszteri  <https://orcid.org/0000-0002-6852-1588>

REFERENCES

- Armand, L. K., Crosta, X., Romero, O., & Pichon, J.-J. (2005). The biogeography of major diatom taxa in Southern Ocean sediments: 1. Sea ice related species. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 223(1), 93–126. <https://doi.org/10.1016/j.palaeo.2005.02.015>
- Assmy, P., Henjes, J., Smetacek, V., & Montresor, M. (2006). Auxospore formation by the silica-sinking, oceanic diatom *Fragilariopsis kerguelensis* (Bacillariophyceae). *Journal of Phycology*, 42(5), 1002–1006. <https://doi.org/10.1111/j.1529-8817.2006.00260.x>
- Assmy, P., Smetacek, V., Montresor, M., Klaas, C., Henjes, J., Strass, V. H., ... Wolf-Gladrow, D. (2013). Thick-shelled, grazer-protected diatoms decouple ocean carbon and silicon cycles in the iron-limited Antarctic Circumpolar Current. *Proceedings of the National Academy of Sciences of the United States of America*, 110(51), 20633–20638. <https://doi.org/10.1073/pnas.1309345110>
- Beszteri, B., Glöckner, G., John, U., Valentin, K., & Postel, U. (2020). *Fragilariopsis kerguelensis* species complex, nucleotide sequences. European Nucleotide Archive (project PRJEB32688). Retrieved from <https://www.ebi.ac.uk/ena/data/view/PRJEB32688>
- Bowen, B. W., Gaither, M. R., DiBattista, J. D., Iacchi, M., Andrews, K. R., Grant, W. S., ... Briggs, J. C. (2016). Comparative phylogeography

- of the ocean planet. *Proceedings of the National Academy of Sciences*, 113(29), 7962–7969. <https://doi.org/10.1073/pnas.1602404113>
- Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). Stacks: An analysis tool set for population genomics. *Molecular Ecology*, 22(11), 3124–3140. <https://doi.org/10.1111/mec.12354>
- Cortese, G., & Gersonde, R. (2007). Morphometric variability in the diatom *Fragilariopsis kerguelensis*: Implications for Southern Ocean paleoceanography. *Earth and Planetary Science Letters*, 257(3–4), 526–544. <https://doi.org/10.1016/j.epsl.2007.03.021>
- Cortese, G., & Gersonde, R. (2008). Plio/Pleistocene changes in the main biogenic silica carrier in the Southern Ocean, Atlantic Sector. *Marine Geology*, 252(3–4), 100–110. <https://doi.org/10.1016/j.margeo.2008.03.015>
- Cortese, G., Gersonde, R., Maschner, K., & Medley, P. (2012). Glacial-interglacial size variability in the diatom *Fragilariopsis kerguelensis*: Possible iron/dust controls? *Paleoceanography*, 27(1), PA1208,14. <https://doi.org/10.1029/2011pa002187>
- Crosta, X., Romero, O., Armand, L. K., & Pichon, J.-J. (2005). The biogeography of major diatom taxa in Southern Ocean sediments: 2. Open ocean related species. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 223(1), 66–92. <https://doi.org/10.1016/j.palaeo.2005.03.028>
- Darling, A. C., Mau, B., Blattner, F. R., & Perna, N. T. (2004). Mauve: Multiple alignment of conserved genomic sequence with rearrangements. *Genome Research*, 14(7), 1394–1403. <https://doi.org/10.1101/gr.2289704>
- de Vargas, C., Audic, S., Henry, N., Decelle, J., Mahe, F., Logares, R., ... Velayoudon, D. (2015). Eukaryotic plankton diversity in the sunlit ocean. *Science*, 348(6237), 1261605. <https://doi.org/10.1126/science.1261605>
- Diepenbroek, M., Glöckner, F., Grobe, P., Güntsch, A., Huber, R., König-Ries, B., & Triebel, D. (2014). Towards an Integrated Biodiversity and Ecological Research Data Management and Archiving Platform: The German Federation for the Curation of Biological Data (GFBio). E. Plödereeder, L. Grunke, E. Schneider & D Ull (Eds.), *Informatik 2014 - Big Data Komplexität meistern*, GI-Edition: Lecture Notes in Informatics (LNI) - Proceedings, GI edn. (Vol. 232, 1711–1724). Bonn: Köllen Verlag.
- Dierckx, N., Mardulyn, P., & Smits, G. (2017). NOVOPlasty: De novo assembly of organelle genomes from whole genome data. *Nucleic Acids Research*, 45(4), e18. <https://doi.org/10.1093/nar/gkw955>
- Doebeli, M., & Dieckmann, U. (2003). Speciation along environmental gradients. *Nature*, 421(6920), 259. <https://doi.org/10.1038/nature01274>
- Endler, J. A. (1977). Geographic variation, speciation, and clines. *Monographs in Population Biology*, 10, 1–246.
- Esper, O., Gersonde, R., & Kadagies, N. (2010). Diatom distribution in southeastern Pacific surface sediments and their relationship to modern environmental variables. *Palaeogeography Palaeoclimatology Palaeoecology*, 287(1–4), 1–27. <https://doi.org/10.1016/j.palaeo.2009.12.006>
- Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software structure: A simulation study. *Molecular Ecology*, 14(8), 2611–2620. <https://doi.org/10.1111/j.1365-294X.2005.02553.x>
- Falush, D., Stephens, M., & Pritchard, J. K. (2003). Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. *Genetics*, 164(4), 1567–1587.
- Falush, D., Stephens, M., & Pritchard, J. K. (2007). Inference of population structure using multilocus genotype data: Dominant markers and null alleles. *Molecular Ecology Notes*, 7(4), 574–578. <https://doi.org/10.1111/j.1471-8286.2007.01758.x>
- Friedrichs, L. (2013). A simple cleaning and fluorescent staining protocol for recent and fossil diatom frustules. *Diatom Research*, 28(3), 317–327. <https://doi.org/10.1080/0269249X.2013.799525>
- Fuchs, N., Scalco, E., Kooistra, W. H. C. F., Assmy, P., & Montresor, M. (2013). Genetic characterization and life cycle of the diatom *Fragilariopsis kerguelensis*. *European Journal of Phycology*, 48(4), 411–426. <https://doi.org/10.1080/09670262.2013.849360>
- Gersonde, R., Abelmann, A., Brathauer, U., Becquey, S., Bianchi, C., Cortese, G., ... Fütterer, D. K. (2003). Last glacial sea surface temperatures and sea-ice extent in the Southern Ocean (Atlantic-Indian sector): A multiproxy approach. *Paleoceanography*, 18(3), 1061. <https://doi.org/10.1029/2002pa000809>
- Glemser, B., Kloster, M., Esper, O., Eggers, S. L., Kauer, G., & Beszteri, B. J. P. B. (2019). Biogeographic differentiation between two morphotypes of the Southern Ocean diatom *Fragilariopsis kerguelensis*. *Polar Biology*, 42(7), 1369–1376. <https://doi.org/10.1007/s00300-019-02525-0>
- Guillard, R. R., & Ryther, J. H. (1962). Studies of marine planktonic diatoms: i. *cytotella nana* hustedt, and *detonula confervacea* (cleve) gran. *Canadian Journal of Microbiology*, 8(2), 229–239. <https://doi.org/10.1139/m62-029>
- Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., & Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology*, 59(3), 307–321. <https://doi.org/10.1093/sysbio/syq010>
- Hubisz, M. J., Falush, D., Stephens, M., & Pritchard, J. K. (2009). Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources*, 9(5), 1322–1332. <https://doi.org/10.1111/j.1755-0998.2009.02591.x>
- Hülskötter, J. (2015). *Detection of genome wide microevolutionary processes in Emiliana huxleyi ((Lohm) Hayard Mohler)*. (Master), Alfred Wegener Institut and University of Bremen, Retrieved from <https://epic.awi.de/id/eprint/39671/>
- Kloster, M., Kauer, G., & Beszteri, B. (2014). SHERPA: An image segmentation and outline feature extraction tool for diatoms and other objects. *BMC Bioinformatics*, 15. <https://doi.org/10.1186/1471-2105-15-218>
- Kloster, M., Kauer, G., Esper, O., Fuchs, N., & Beszteri, B. (2018). Morphometry of the diatom *Fragilariopsis kerguelensis* from Southern Ocean sediment: High-throughput measurements show second morphotype occurring during glacials. *Marine Micropaleontology*, 143, 70–79. <https://doi.org/10.1016/j.marmicro.2018.07.002>
- Kumar, S., Stecher, G., & Tamura, K. (2016) MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*, 33(7), 1870–1874. <https://doi.org/10.1093/molbev/msw054>
- Lefort, V., Longueville, J.-E., & Gascuel, O. (2017). SMS: Smart model selection in PhyML. *Molecular Biology and Evolution*, 34(9), 2422–2424. <https://doi.org/10.1093/molbev/msx149>
- Mock, T., Otilar, R. P., Strauss, J., McMullan, M., Paajanen, P., Schmutz, J., ... Grigoriev, I. V. (2017). Evolutionary genomics of the cold-adapted diatom *Fragilariopsis cylindrus*. *Nature*, 541, 536. <https://doi.org/10.1038/nature20803>
- Orsi, A. H., Whitworth, T., & Nowlin, W. D. (1995). On the meridional extent and fronts of the Antarctic Circumpolar Current. *Deep Sea Research Part I: Oceanographic Research Papers*, 42(5), 641–673. [https://doi.org/10.1016/0967-0637\(95\)00021-W](https://doi.org/10.1016/0967-0637(95)00021-W)
- Pau, G., Fuchs, F., Sklyar, O., Boutros, M., & Huber, W. (2010). EBImage—an R package for image processing with applications to cellular phenotypes. *Bioinformatics (Oxford, England)*, 26(7), 979–981. <https://doi.org/10.1093/bioinformatics/btq046>
- Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., & Hoekstra, H. E. (2012). Double digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS One*, 7(5), e37135. <https://doi.org/10.1371/journal.pone.0037135>
- Pinkernell, S. (2017). *Modeling the biogeography of pelagic diatoms of the Southern Ocean*. (PhD), University of Rostock. Retrieved from <http://>

- hdl.handle.net/10013/epic.38f56694-7abe-4049-8596-1f65e7f1e071
- Pinkernell, S., & Beszteri, B. (2014). Potential effects of climate change on the distribution range of the main silicate sinker of the Southern Ocean. *Ecology and Evolution*, 4(16), 3147–3161. <https://doi.org/10.1002/ece3.1138>
- Pollard, R. T., Lucas, M. I., & Read, J. F. (2002). Physical controls on biogeochemical zonation in the Southern Ocean. *Deep Sea Research Part II: Topical Studies in Oceanography*, 49(16), 3289–3305. [https://doi.org/10.1016/S0967-0645\(02\)00084-X](https://doi.org/10.1016/S0967-0645(02)00084-X)
- Post, A. L., Meijers, A. J. S., Fraser, A. D., Meiners, K. M., Ayers, J., Bindoff, N. L., ... Raymond, B. (2014). Environmental Setting. In C. De Broyer, P. Koubbi, H. J. Griffiths, B. Raymond, C. d'Udekem d'Acoz, A. P. Van de Putte, B. Danis, B. David, S. Grant, J. Gutt, C. Held, G. Hsieh, F. Huettmann, A. Post, & Y. Ropert-Coudert (Eds.), *Biogeographic Atlas of the Southern Ocean* (pp. 46–64). Cambridge, UK: Scientific Committee on Antarctic Research.
- Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, 155(2), 945–959.
- Rainey, P. B., & Travisano, M. (1998). Adaptive radiation in a heterogeneous environment. *Nature*, 394(6688), 69. <https://doi.org/10.1038/27900>
- R-Core-Team (2015). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <http://www.R-project.org/>
- Rochette, N. C., & Catchen, J. M. (2017). Deriving genotypes from RAD-seq short-read data using Stacks. *Nature Protocols*, 12, 2640–2659. <https://doi.org/10.1038/nprot.2017.123>
- Schweyen, H., Rozenberg, A., & Leese, F. (2014). Detection and removal of PCR duplicates in population genomic ddRAD studies by addition of a degenerate base region (DBR) in sequencing adapters. *The Biological Bulletin*, 227(2), 146–160. <https://doi.org/10.1086/BBLv227n2p146>
- Shukla, S. K., Crosta, X., Cortese, G., & Nayak, G. N. (2013). Climate mediated size variability of diatom *Fragilariopsis kerguelensis* in the Southern Ocean. *Quaternary Science Reviews*, 69, 49–58. <https://doi.org/10.1016/j.quascirev.2013.03.005>
- Shukla, S. K., & Romero, O. E. (2018). Glacial valve size variation of the Southern Ocean diatom *Fragilariopsis kerguelensis* preserved in the Benguela Upwelling System, southeastern Atlantic. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 499, 112–122. <https://doi.org/10.1016/j.palaeo.2018.03.023>
- Sokolov, S., & Rintoul, S. R. (2002). Structure of Southern Ocean fronts at 140°E. *Journal of Marine Systems*, 37(1), 151–184. [https://doi.org/10.1016/S0924-7963\(02\)00200-2](https://doi.org/10.1016/S0924-7963(02)00200-2)
- Sprouffske, K., & Wagner, A. (2016). Growthcurver: An R package for obtaining interpretable metrics from microbial growth curves. *BMC Bioinformatics*, 17(1), 172. <https://doi.org/10.1186/s12859-016-1016-7>
- Staden, R. (1996). The Staden sequence analysis package. *Molecular Biotechnology*, 5(3), 233–241. <https://doi.org/10.1007/BF02900361>
- Talley, L. D. (2011). *Descriptive physical oceanography: An introduction*. London, UK: Academic press.
- Villemereuil, P., Mouterde, M., Gaggiotti, O. E., & Till-Bottraud, I. (2018). Patterns of phenotypic plasticity and local adaptation in the wide elevation range of the alpine plant *Arabis alpina*. *Journal of Ecology*, 106(5), 1952–1971. <https://doi.org/10.1111/1365-2745.12955>
- WGS/WES Mapping to Variant Calls - Version 1.0. (2018) Retrieved from <https://www.htslib.org/workflow/>
- White, T. J., Bruns, T., Lee, S., & Taylor, J. (1990). Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. *PCR Protocols: A Guide to Methods and Applications*, 18(1), 315–322.
- Zheng, X., Levine, D., Shen, J., Gogarten, S. M., Laurie, C., & Weir, B. S. (2012). A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics*, 28(24), 3326–3328. <https://doi.org/10.1093/bioinformatics/bts606>

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Postel U, Glemser B, Salazar Alekseyeva K, et al. Adaptive divergence across Southern Ocean gradients in the pelagic diatom *Fragilariopsis kerguelensis*. *Mol Ecol*. 2020;00:1–12. <https://doi.org/10.1111/mec.15554>