Brandenburgische Technische Universität
Cottbus - Senftenberg

Alfred-Wegener-Institut

Master Thesis

"Population genomic comparison of *Fragilariopsis kerguelensis*, a Southern Ocean diatom''

Katherine Jennifer Salazar Alekseyeva

First examiner: Prof. Dr. rer. nat. habil. Brigitte Nixdorf and Dr. rer. nat. Jacqueline Rücker

Second examiner: Dr. Bank Beszteri

Cottbus, May 31 2018

Declaration

The author declares that he/she has written the thesis at hand independently, without outside help and without the use of any other but the listed sources. Thoughts taken directly or indirectly form external sources (including electronic sources) are marked accordingly without exception. Sources used verbatim and contextual were quoted according to the recognized rules for scientific working. This thesis has not been submitted in the same or similar form, not even partially, within the scope of a different examination. Thus far it also has not been publicized yet.

I herewith agree that the thesis will be examined for plagiarism with the help of a plagiarism detection service.

Location, Date **Author's signature**

Abstract

In the Southern Ocean, diatoms together with other phytoplankton, are the base of the food webs. *Fragilariopsis kerguelensis* as the most abundant and truly Antarctic diatom species, might have important effects on organisms that depend on it. Additionally, this species can influence nutrient cycles and global climate. Therefore, *Fragilariopsis kerguelensis*, can be an important species to predict possible effects of climate change.

Recent distribution studies performed in the Laboratory of Hustedt Diatom Study Centre, Polar Biological Oceanography Section at the Alfred-Wegener-Institut, using public taxon observations alongside observations from material housed there, prospected that ocean warming will affect *Fragilariopsis kerguelensis* around the northern edge of its distribution area. As a result, this species might be developing genotypic adaptations to these changing environmental conditions.

According to the data available and ecological relevance, this thesis reviewed sea surface temperature (SST), as well as other environmental parameters such as nitrate, phosphate, and silicate. Northern and southern populations of this species were characterized using double digest restriction-site associated DNA sequencing (ddRAD). The molecular results obtained suggest that these environmental parameters are correlated with the genetic differentiation of *Fragilariopsis kerguelensis*.

A particular locus of *Fragilariopsis kerguelensis* was identified to be seemingly sensitive to the changing conditions in the Southern Ocean. This was "*Phaeodactylum tricornutum* CCAP 1055/1 predicted protein partial mRNA", also found in another pennate diatom. Nonetheless, the major challenge was to understand the simultaneous effects of these parameters.

Table of contents

List of Figures

List of Tables

Keywords

ACC= Antarctic Circumpolar Current

AZ= Antarctic Zone

CZ= Continental Zone

ddRAD

F_{st}= Fixation index

HNLC= High Nutrient Low Chlorophyll

mRNA= Mitochondrial RNA

rRNA= Ribosomal RNA

PEP= posterior error probability

PF= Polar Front

PFZ= Polar Frontal Zone

SAF= Subantarctic Front

SAZ= Subantarctic Zone

SO= Southern Ocean

SNP= single nucleotide polymorphism

SST= Sea surface temperature

STF= Subtropical Front

1. Introduction

Climate change can influence ocean properties and phytoplankton (Boyd, et al., 2016). In the Southern Ocean, the response of these organisms is important to predict effects on ecosystems, as well as on global climate (Cortese & Gersonde, 2007; Deppeler & Davidson, 2017). Some phytoplankton species have developed genetic differentiation that has allowed them to adapt and evolve to changing environmental conditions (Crosta & Koc, 2007).

1.1. Characteristics of the Southern Ocean

The Southern Ocean (SO), also known as Antarctic Ocean or Austral Ocean, is an ocean located from the coast of Antarctica, north to 30° latitude in the Southern Hemisphere (Moore & Abbott, 2000). It limits with the Atlantic, Indian and Pacific Oceans, and has an area of 20.33 million km² which represents 10 % of the world's oceans area (Tréguer & Jacques, 1992).

The Antarctic Circumpolar Current (ACC) is a predominant current of the SO that flows towards the east entirely around the world (Gille, 2002). The estimated average transport of the ACC is 100-150 Sv (10^6 m^3/s) (Lochte, 2014; Palla, 2014), which represents one of the largest and strongest world's ocean currents (Sandells & Flocco, 2014; Smith, et al., 2013).

An oceanic frontal system or front, is a boundary between two water masses with different hydrographic properties (Shao, et al., 2015). The SO is formed by Polar Front (PF), Subantarctic Front (SAF), and Subtropical Front (STF) (Zielinski & Gersonde, 1997). The zones between these fronts are Continental Zone (CZ), Antarctic Zone (AZ), Polar Frontal Zone (PFZ), and Subantarctic Zone (SAZ) (Figure 1) (Sokolov & Rintoul, 2002).

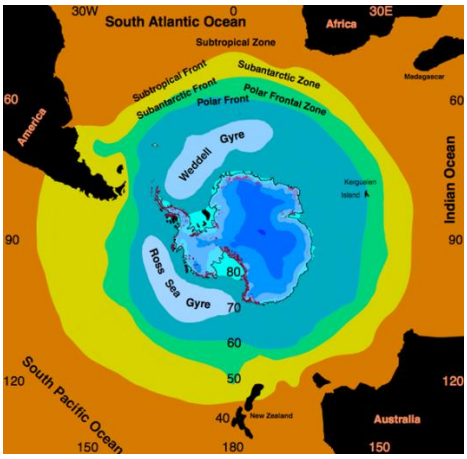

Figure 1. Location of the Antarctic Circumpolar Current (ACC) within the Southern Ocean and its oceanic frontal systems and zones, taken from Grobe (2007).

One important oceanic frontal system is the Polar Front (PF). It is located between 50° and 52° S, approximately 1500 km off the Antarctica coast. It separates warmer and saltier waters of the north from colder and fresher waters of the south (Acha, et al., 2015; Orsi, et al., 1995). The PF is characterized by a high nutrient concentration in surface waters, especially silicate, enabling the growth of thick-shelled diatom species, so it is also the main area where biogenic silica sinks to deep waters (Zielinski & Gersonde, 1997).

The Subantarctic Front (SAF) is located between 48 and 50° S. As the ACC detours northwards, the SAF enters the South Atlantic Ocean through the Drake Passage (Acha, et al., 2015; Talley, 2011). It almost merges with the STF, which creates one of the strongest oceanic frontal systems, together with the PF (Marshall & Speer, 2012), so it oversees most of the transportation in the ACC (Shao, et al., 2015).

The Subtropical Front (STF) is located close to 40° S (Graham, 2013). It is disrupted by South America, so it is considered the northern boundary of the ACC (Orsi, et al., 1995). It separates southern subtropical gyres with a broad current of the ACC further south (Bard & Rickaby, 2009). The STF is characterized by a low nutrient concentration in surface waters (Falkowski, et al., 1998).

A gyre is a system of recirculating waters (Orsi, et al., 1995). The ACC has two important gyres that are the Weddell and Ross gyres. The Weddell gyre location is in the Weddell Ice Shelf, whereas, the Ross gyre is in the Ross Ice Shelf. In the SO, these gyres are created due to significant interactions between the ACC and the Antarctic continental shelf (Palla, 2014).

1.2. Importance of the Southern Ocean

The SO plays an important role in the global climate due to its connection with the Atlantic, Indian, and Pacific Oceans through the ACC (Shao, et al., 2015). The oceans are capable to uptake atmospheric carbon dioxide (CO_2) (Deppeler & Davidson, 2017). The SO just by itself, uptakes approximately 40 % of CO_2 since cold waters are capable to store more CO_2 than warmer waters (Gille, 2002). However, the recent increase of anthropogenic CO_2 can change dramatically this cycle (Arrigo, et al., 1999; Collins, et al., 2013).

Marine phytoplankton plays an important role in the global climate as well (Deppeler & Davidson, 2017). In the SO, one of the largest component of phytoplankton are diatoms (Falkowski, et al., 1998; Ito, et al., 2005), which are also capable to uptake atmospheric CO_2 (Armbrust, et al., 2004; Hülskötter, 2013). Additionally, diatoms are involved in the main

biogeochemical cycles of elements like carbon (C), nitrogen (N), phosphorus (P), and silicon (Si), so they are among the main carriers of organic matter and silicon to deep waters (Assmy, et al., 2006; Brzezinski, et al., 2001).

1.3. Environmental conditions of the Southern Ocean

In surface waters of the SO, except the STF, there are relatively high concentrations of nitrate, phosphate, and silicate, but there are barely phytoplankton blooms, even in the summertime (de Baar, et al., 1997). Nowadays, it is known that limiting factors such as iron, light, and strong grazing, contribute to low phytoplankton standing stocks in the SO (Crosta & Koc, 2007). Other trace metals are also hypothesized to contribute to it, but these are not so clear (Tréguer & Jacques, 1992).

1.3.1. Temperature

Gille (2002) reported that between the 1950s and 1980s, the sea surface temperatures (SST) in the SO have increased by 0.17 °C (Figure 2). Collins, et al. (2013) reported that recently, the SST has risen by 0.70 °C, and an additional 3 °C is expected in a close future. Phytoplankton species are adapted to a defined thermal window, so they are sensitive to an increase in temperatures (Gattuso & Hansson, 2011). Additionally, at higher temperatures, grazers need to consume more phytoplankton which can cause that the rates of grazing exceed the rates of phytoplankton growth (Deppeler & Davidson, 2017).

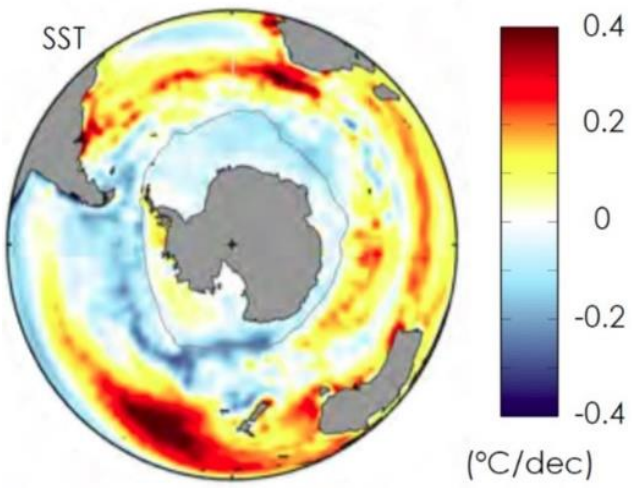

Figure 2. Increase of SST in the SO, taken from Gille (2002).

1.3.2. High Nutrient Low Chlorophyll (HNLC)

The SO is the largest High Nutrient Low Chlorophyll (HNLC) area of the world ocean (Crosta & Koc, 2007; Moore & Abbott, 2000). Compared with other waters, the SO has

macronutrients such as nitrate, phosphate, and silicate that are not significantly depleted, whereas, micronutrients such as iron are in low concentrations (Table 1) (Boyle, 1998). However, phytoplankton biomass remains relatively constant (Pitchford & Brindley, 1999).

Table 1. Comparison of the concentration of macronutrients between the SO and a freshwater body. Source Falkowski, et al. (1998), National Oceanic and Atmospheric Administration (2018), and Walker, et al. (2007).

		Concentration	
		Southern Ocean	Freshwater (natural lakes)
Nitrate	μM	5-50	<1.00
Phosphate	μM	1.00-2.50	0.005-0.05
Silicate	μM	0-90	1-40
Chlorophyll	mg m^{-3}	<0.30	14.00

Phytoplankton as a primary producer, transform light and inorganic nutrients into organic matter (Jones, 2000; Walker, et al., 2007). In the SO, phytoplankton apparently does not use completely all these available nutrients (Cortese & Gersonde, 2007; Ito, et al., 2005), so the chlorophyll concentration does not exceed 0.60 mg m^{-3} (Figure 3). However, the highest chlorophyll concentrations are related to the main fronts and summer (Moore & Abbott, 2000).

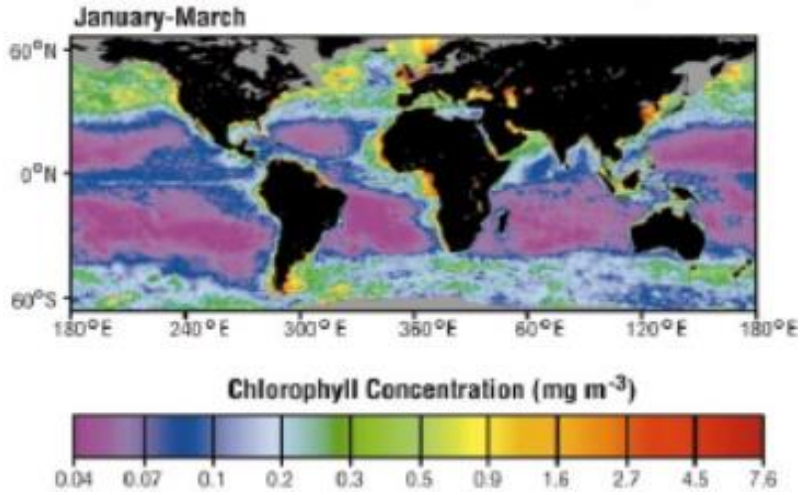

Figure 3. Global chlorophyll concentration (mg m^{-3}), taken from Falkowski, et al. (1998).

In the euphotic zone of the SO, light and iron limits primary production (Falkowski, et al., 1998). In spring and summer, the light limitation is lower compared to autumn and

winter where expansion of heavy sea ice can lead to a strong light limitation (Ito, et al., 2005; Moore & Abbott, 2000). Nevertheless, the light levels are usually low due to cloudiness and deep mixing (Deppeler & Davidson, 2017).

In the study of Moore & Abbott (2000), phytoplankton blooms with a chlorophyll concentration of >1.0 mg m^{-3} were distinguished in coastal/shelf waters, seasonal sea ice retreat waters and close to the fronts. Consequently, phytoplankton blooms are present with optimal light conditions (de Baar, et al., 1997), but low chlorophyll concentration can also be reached due to self-shading (Nelson & Tréguer, 1992).

In deeper waters of the SO, high inorganic nutrient concentrations are transported to the euphotic zone (Figure 4). Grazer and microbial communities are extremely efficient. The grazers consume the low phytoplankton biomass that is produced, while microbes regenerate nutrients. Almost all organic matter is remineralized into frustules of diatoms with an insignificant loss of organic carbon or inorganic nutrients from the euphotic zone (Falkowski, et al., 1998).

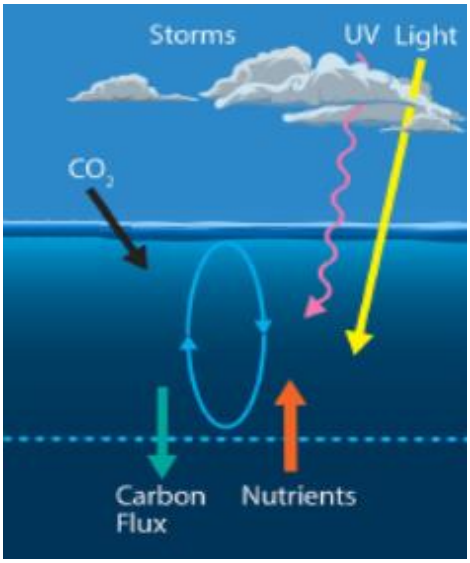

Figure 4. High Nutrient Low Chlorophyll (HNLC) in the SO, modified from Deppeler & Davidson (2017).

1.3.2.1. Silicon

Silicon (Si) is an essential element required for growth of specific biota, such as radiolarians, silicoflagellates (Deppeler & Davidson, 2017), and diatoms (Finkel, 2016; Sokolov & Rintoul, 2002). Si plays an important role in the formation of structures, such as frustules (Brümmer, 2003). In most diatoms species, silicon is taken up before the formation of the cell wall (Hoffmann, et al., 2007).

Silica (SiO_2), also known as opal, is one of the most common minerals produced by living organisms (Brümmer, 2003). Silicification is a process where silicic acid or orthosilicic acid ($Si(OH)_4$) is taken up from water (Brümmer, 2003), and supersaturated and precipitated onto/around an organic matrix (Celan, 2014). The silica deposition vesicles (SDVs) are intracellular compartments where silica is concentrated (Barsanti & Gualtieri, 2014). When the frustule is ready, silica is externalized outside the cell (Raven & Waite, 2004).

Silicification is a process that requires little energy (Timmermans & van der Wagt, 2010). Under unfavorable growth conditions, diatoms can build thicker frustules (Boyle, 1998), which increases the cell wall strength as well as protects against grazers and parasites (Hoffmann, et al., 2007). This in turn, increases the cell density and the sinking rate, thus diatoms can sink to waters with better nutrients concentrations (Raven & Waite, 2004). However, this is a trade-off because diatoms will receive less light for photosynthesis.

Silicon from diatoms frustules is preserved in some ocean sediments (Brzezinski, et al., 2001). In the SO, these sediments are mainly formed by frustules of *Fragilariopsis kerguelensis* and *Thalassiosira lentiginosa*, and spores of *Chaetoceros* spp. (Figure 5) (Hoffmann, et al., 2007; Sachs, et al., 2009). The Opal Belt is the largest deposit of biogenic silica and it is located mainly in the PF and extending southwards (de Baar, et al., 1997; Ito, et al., 2005).

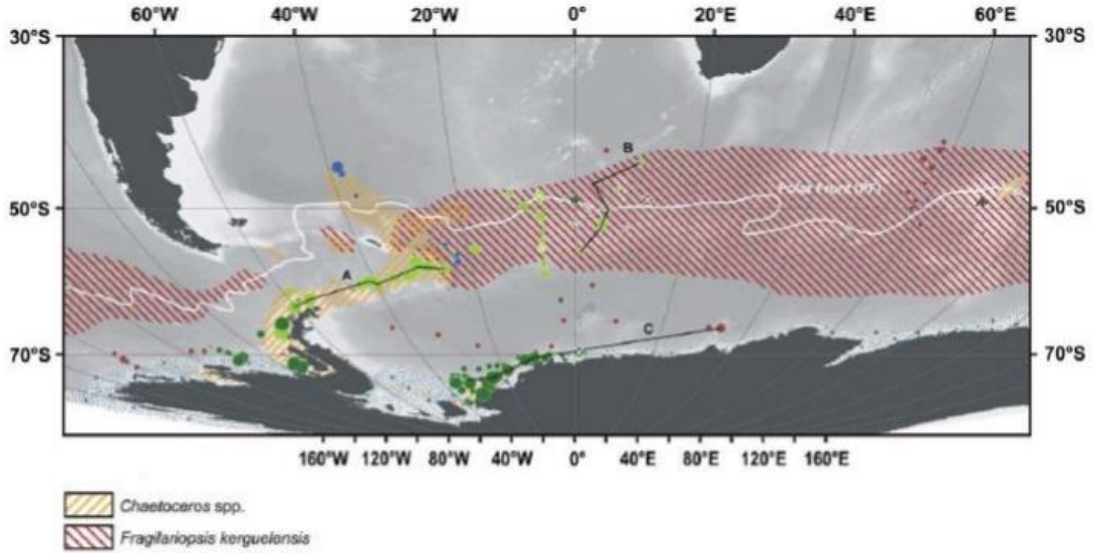

Figure 5. Sediments in the SO, formed mainly by frustules of *Fragilariopsis kerguelensis* and spores of *Chaetoceros* spp., taken from Sachs, et al. (2009).

When a silicified specimen is eaten by other organism or simply die, its siliceous cells sink to deeper waters and form aggregates of biogenic material (Deppeler & Davidson,

2017). During sinking, silicate is dissolved which leads to an enrichment of waters. Further south, these deep water with high silicate concentration are upwelled to the surface. As a result, silicate is reused to build diatoms shells (Gnanadesikan, 1999).

The SO is an area with more than 2/3 of the world's silicate (Gnanadesikan, 1999). As mentioned before, the PF is considered the main area where this nutrient sink to deep waters (de Baar, et al., 1997; Zielinski & Gersonde, 1997). Indeed, the location of this front within the ACC can be recognized due to the silicate concentration in the sediments (Falkowski, et al., 1998).

In this ocean, the distribution of silicate depends mainly on water processes (Gnanadesikan, 1999), which creates a strong latitudinal gradient (Moore & Abbott, 2000). In the southern waters that experience seasonal expansion and retraction of ice, silicate concentration is between 60 to 90 μM. Nonetheless, in northern waters that do not experience the formation of ice, this concentration is <10 μM (Brzezinski, et al., 2001; Ito, et al., 2005; Nelson & Tréguer, 1992). Seemingly, during glacial periods, silicate concentration spread further north than today (Falkowski, et al., 1998).

1.3.2.2. Iron

Iron (Fe) is a trace metal required for an optimal phytoplankton growth. Fe is mainly needed in photosynthetic machinery and also in different enzymes which are essential for nitrogen assimilation (de Baar, et al., 1997; Tréguer & Jacques, 1992). Consequently, iron plays a strong role in phytoplankton distribution, productivity and structure (Strzepek, et al., 2011).

Firstly, in the photosynthetic electron transport chain, iron is necessary for iron-sulfur containing proteins and ferredoxin of the photosystems; and in iron-sulfur containing proteins and heme and of the cytochrome b6f complex (Strzepek, et al., 2011). Therefore, iron is necessary for the photosynthetic electron transfer, photosynthetic energy supply and photosynthetic fixation of carbon (de Baar, et al., 1997; Hoffmann, et al., 2007).

Secondly, in the metabolism of several nitrogen sources, iron is necessary. In the case of nitrate (NO_3^-), its reduction to ammonium requires iron containing cellular enzymes (reductases) and iron within its multiple oxidation states (II and III) (Timmermans & van der Wagt, 2010). These cellular enzymes are nitrate and nitrite reductase, and this last one can also use reduce ferredoxin, an iron-sulfur redox protein, or a non-iron-containing flavodoxin in the reduction (Hoffmann, et al., 2007). As a result, the reductive process of nitrate assimilation requires extra iron (de Baar, et al., 1997).

In the case of nitrogen (N_2), its fixation requires an increase of iron. However, in the cases of ammonia (NH_3) and ammonium (NH_4), its cellular assimilation does not require extra iron (de Baar, et al., 1997). Phytoplanktonic cells can directly assimilate ammonium and integrate it into amino acids without cellular enzymes that contain iron (Hoffmann, et al., 2007).

This means that the iron demand of phytoplankton cells depends on the nitrogen source or the other way around, the range of nitrogen forms that can be utilized by phytoplankton cells depends on iron supply. At sites with low iron concentration (<0.2 nM), the assimilation of ammonium is expected, but at sites with quite low iron concentration (<0.7 nM), the assimilation of nitrate is possible (de Baar, et al., 1997). Therefore, the iron demand is higher when phytoplanktonic cells grow on nitrate compared to ammonium. In the case of Antarctic waters, iron stress is common, so the fixation of nitrogen is unlikely, and the ammonium uptake is preferred (Hoffmann, et al., 2007).

1.3.2.3. Nitrate and phosphate

Nitrate (NO_3^-) and phosphate (PO_4) are required for phytoplankton growth (Sigman, et al., 2000). Nitrogen is an essential component of proteins and chlorophyll, whereas, phosphate is an essential component of nucleic acids, lipids and other metabolites (Walker, et al., 2007). In the SO, approximately half of nitrogen and phosphorus return unused from the surface waters to deeper waters (Falkowski, et al., 1998), probably due to iron and light limitation.

In oceans, nitrate concentration can be affected by processes of a closed or open system. A closed system is dominated by the uptake of nitrate by phytoplankton and oxidation of organic N by bacteria. This influences the cycling and distribution of N within the ocean. On the other hand, an open system is dominated by N fixation and denitrification (Sigman, et al., 2000).

1.4. Diatoms

In the SO, approximately 350 species of phytoplankton have been identified (Deppeler & Davidson, 2017). When growth conditions are optimal, diatoms are the largest component of the SO phytoplankton (Bopp, et al., 2005; Falkowski, et al., 1998). Hence, they are the main primary producers representing 75 % of the SO primary production (Mock, et al., 2017; Zielinski & Gersonde, 1997).

Diatoms are unicellular photosynthetic eukaryotes that grow in both, sea and fresh waters (Huysman, et al., 2010; Raven & Waite, 2004). They are divided into two main groups, *Coscinodiscophyceae* or centric diatoms, and *Bacillariophyceae* or pennate diatoms (Bowler, et al., 2008; Crosta & Koc, 2007). The centric diatoms have cells with a radial symmetry whereas, pennate diatoms have bilateral symmetry (Brümmer, 2003; Spaulding, 2018).

Diatoms can reproduce sexually and asexually. During vegetative reproduction, a frustule with two different sized halves is divided into two different sized daughter cells. Each of them contains an old and new halve, which will result in a progressive decrease of the frustule size (Brümmer, 2003). Nevertheless, at a certain point, the size will be restored through sexual reproduction (Armbrust, et al., 2004).

1.4.1. Characteristics of the genus *Fragilariopsis*

Bacillariaceae is a family belonging to the kingdom Chromista, phylum Ochrophyta, class Bacillariophyceae and order Bacillariales. Currently, this family consists of 25 genera and 1091 species. The biggest genera are *Nitzschia* (830 species), *Hantzschia* (88 species), *Tryblionella* (64 species), *Pseudo-nitzschia* (48 species), *Denticula* (31 species), and *Fragilariopsis* (27 species) (Ehrenberg, 2017).

The genus *Fragilariopsis* Hustedt was described by Friedrich Hustedt in 1913. This genus consists of approximately 27 species of marine diatoms, mainly planktonic species, but also benthic and ice-associated species (Jacques, 1983). The form of living is mainly planktonic, but occasionally also periphytic and epilithic, as well as associated with snow and ice (Wehr, et al., 2015).

The geographic distribution of this genus is mainly polar, especially southwards of the PF, where this genus occurs in great abundance and diversity (Cefarelli, et al., 2010). Ferrario & Licea (2006) also reported *Fragilariopsis* as one of the most dominant genera in the SO during summer and autumn. In the case of this genus, there are only a few exceptions that can occur beyond polar.

1.4.2. Characteristics of the species *Fragilariopsis kerguelensis*

Fragilariopsis kerguelensis (O'Meara) Hustedt is an endemic and pennate diatom of the SO (Shukla, et al., 2013; Zielinski & Gersonde, 1997). *Fragilariopsis kerguelensis* is an open ocean species and lives preferentiality in surface waters with temperatures between 1 and 8 °C (Shukla & Crosta, 2017). The geographic distribution of this species is exclusively polar.

Fragilariopsis kerguelensis is the most abundant species within the diatoms in the ACC between the sea ice edge and the PF, representing 60 to 90 % of the total diatoms (Cefarelli, et al., 2010; Cortese & Gersonde, 2007). The main distribution is between 40° to 70° latitude in the Southern Hemisphere (Zielinski & Gersonde, 1997). Therefore, *Fragilariopsis kerguelensis* is an important contributor to the biomass of the Antarctic diatoms (Hoffmann, et al., 2007).

Fragilariopsis kerguelensis is a heavily silicified species that represents almost 40 % of biogenic silica deposits in the SO (Fuchs, et al., 2013). The heavily silicified frustules formed by valves and girdle bands provides high resistance, and low probability of deformation and breakage, which could be an adaptation against grazing (Fuchs, et al., 2013). In larger specimens, the valve can be lanceolate with curved extremes, but in smaller specimens, it can be elliptic with more curved extremes (Cefarelli, et al., 2010).

Fragilariopsis kerguelensis can be arranged in single cells or in few to hundred cells chains where the valves are united through a mucilaginous material (Figure 6) (Fuchs, et al., 2013; Timmermans & van der Wagt, 2010). The valve mantle is unperforated, and the cingulum is formed by a valvocopula that can have or not have perforations, and a copula unperforated (Cefarelli, et al., 2010).

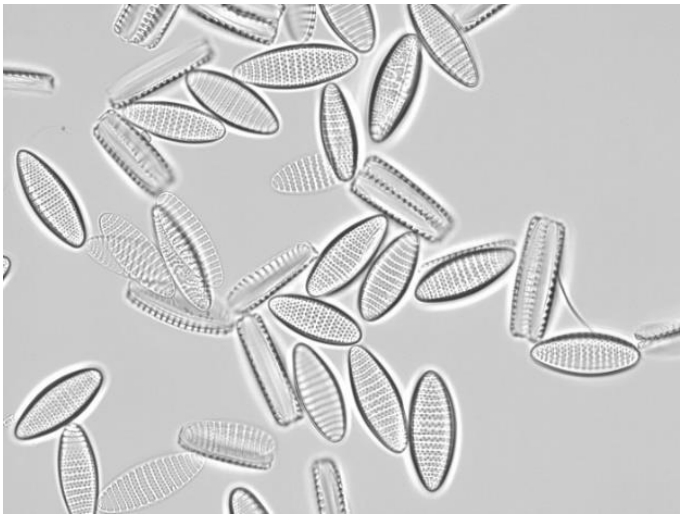

Figure 6. Individuals of *Fragilariopsis kerguelensis*, provided by Hustedt Diatom Study Centre.

The size of *Fragilariopsis kerguelensis* is influenced by physical and biochemical factors (Shukla & Crosta, 2017). The valve area size of *Fragilariopsis kerguelensis* was 30 % larger during glacial than during interglacial (Timmermans & van der Wagt, 2010). During the glacial periods, the concentrations of iron were higher and the sea level lower (by

approximately 120 m) compared with the interglacial periods (Falkowski, et al., 1998). Nowadays, the apical length of this species varies between 10 and 90 μm (Shukla, et al., 2013).

As mentioned earlier, *Fragilariopsis kerguelensis* plays an important role in the cycles of the SO, especially in the cycle of nutrients and silica (Cortese & Gersonde, 2007; Fuchs, et al., 2013). This species exports biogenic silica from the surface to the sediments in the bottom of the ocean (Shukla & Crosta, 2017). Consequently, this species is an important contributor to the global silicon cycle (Cortese & Gersonde, 2007).

1.5. Genotype and phenotype

Genotype is all the genetic information contained in the genes of an organism. This genetic information is transmitted in each generation. Contrary, the phenotype is all the observable characteristics that an organism have, and it can be influenced by genotype as well as the environment. Therefore, different genotypes can produce new phenotypes (de Vries, 2018).

1.6. Population genomics

In the SO, phytoplankton, as the base of the food webs, can influence other ecosystems, nutrient cycles and global climate (Collins, et al., 2013; Deppeler & Davidson, 2017). Compared with other organisms, diatoms have short generation times and high genetic diversity possible due to losses and gains of genes and introns (Bowler, et al., 2008). Therefore, their response to changing environmental conditions will mainly depend on their capacity to adapt and evolve (Boyd, et al., 2016; De Riso, et al., 2009).

Next-generation sequencing methods have been revolutionized by molecular and evolutionary studies (Schweyen, et al., 2014). Single Nucleotide Polymorphism (SNP) is a unique variation in one nucleotide of a specific fixed position in the DNA sequence. This single change can be a beneficial adaptation to the local environment, but also cannot have any effect (de Villemereuil & Gaggiotti, 2015).

Double digestion (dd) is a molecular method that uses two restriction enzymes for simultaneously DNA digestion (Peterson, et al., 2012). Restriction-site associated DNA sequences (RAD) are a useful molecular marker to detect SNPs on a genome-wide scale (Schweyen, et al., 2014). Thus, double digest restriction-site associated DNA sequencing (ddRAD) is a method that uses two restriction enzymes to prepare a library and analyze genomic loci.

In the 1960's, Werner Arber discovered restriction enzymes. These are endonucleases that hydrolyze phosphodiester bonds of double-stranded DNA at restriction sites producing smaller DNA fragments (Hülskötter, 2013; Pray, 2008). The classification of the restriction enzymes is commonly based on the subunit composition, cleavage position, sequence specificity and cofactor requirements.

According to the cleavage position, there are two types of restriction enzymes. Type I are enzymes that cleave DNA at random sites, away from the recognized restriction sites, producing variable restriction fragments (Roberts, et al., 2003). Contrary, Type II enzymes cleave DNA at constant sites, close to or within restriction sites, producing defined restriction fragments. As a result, Type I have low practical value for molecular biological applications, whereas, Type II enzymes are widely used in DNA analysis.

According to sequence specificity, there are two types of enzymes. The frequent cutters are those enzymes that recognize restriction sites with a length of four bases. Contrary, the rare cutters, are those enzymes that recognize restriction sites with a length of six to eight bases (Hülskötter, 2013). For instance, the length of the restriction site of PstI is 6 bp long, so it is usually referred to as a rare cutter, whereas, MspI recognizes a 4 bp long sequence, accordingly it can be called a frequent cutter (Figure 7).

	5'		3'	5'		3'
PstI	CTGCA	↓	G	G	↓	ACGTC
MspI	C	↓	CGG	GGC	↓	C

Figure 7. Sequence of the restriction site of PstI and MspI. Source Thermo Fisher Scientific (2018).

As mentioned before, according to the restriction enzymes, the locus has two types of restriction sites, one is the low frequency cutter (green) and the other is the high frequency cutter (blue) (Figure 8). The DNA digestion is done with two restriction enzymes, for example, PstI and MspI. Later, the overhanging DNA ends (digested fragments) are ligated to two adapters that are barcoded (P5 and P7) (Hülskötter, 2013). During a PCR, the digested and ligated fragments are amplified. However, some fragments are amplified more often than others, according to frequency cut site (Schweyen, et al., 2014).

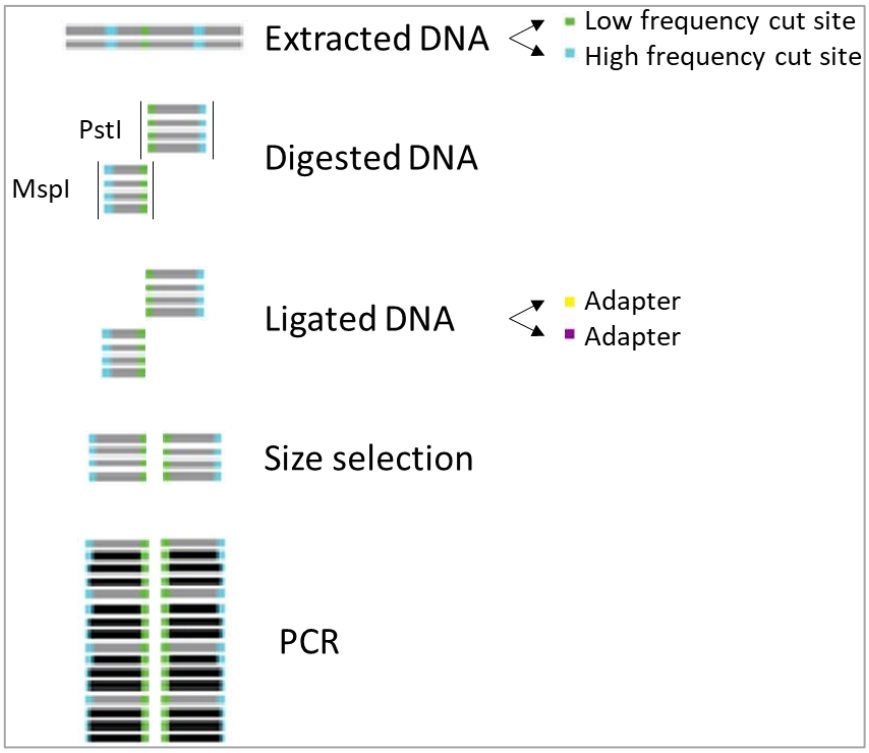

Figure 8. ddRAD process, modified from Schweyen, et al. (2014).

2. Study area

The study area selected was the Southern Ocean. In December 2016, during the cruise PS103 aboard the Polarstern, samples of *Fragilariopsis kerguelensis* were collected from four different stations of a transect from Cape Town – South Africa, southward (Figure 9). Stations 1, 2, 4, and 14 were selected because they represented northern and southern populations, thus different environments and water masses of the SO.

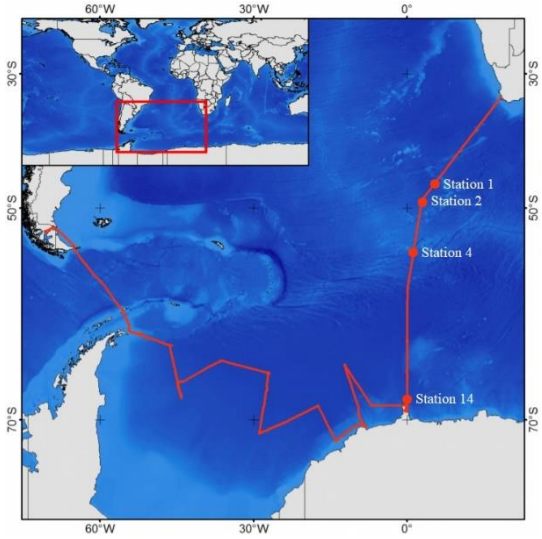

Figure 9. Journey of the cruise PS103 aboard the Polarstern and location of the Stations 1, 2, 4, and 14, modified from Alfred-Wegener-Institut (2017).

As shown in Table 2, Station 1 was located around the -45° latitude and 6° longitude in the Southern Hemisphere, close to the SAF; and Station 2 was located around the -49° latitude and 2° longitude, close to the PF. It is important to mention that the location of the fronts can vary considerably within years (Sokolov & Rintoul, 2002). Station 4 was located around the -55° latitude and 1° longitude, close to the southern boundary of the ACC. Finally, Station 14 was located around the -68° latitude and 0° longitude, in the Weddell Sea, next to sea ice edge and close to the shelf ice (Alfred-Wegener-Institut, 2017).

Table 2. Geographical location of the Stations 1, 2, 4, and 14.

Expedition	Station	Latitude	Longitude
PS103	1	-45	6
	2	-49	2
	4	-55	1
	14	-68	0

3. Objectives

Fragilariopsis kerguelensis (O'Meara) Hustedt is an endemic diatom of the Southern Ocean. Due to its heavily silicified cell wall and its high abundance within diatoms, *Fragilariopsis kerguelensis* represents one of the main constituents of the siliceous sediment belt around Antarctica. Recent distribution studies performed in the Laboratory of Hustedt Diatom Study Centre, Polar Biological Oceanography Section at the Alfred-Wegener-Institut, using public taxon observations alongside observations from material housed there, prospected that ocean warming will affect this species around the northern edge of its distribution area. As a result, *Fragilariopsis kerguelensis* might be developing genotypic adaptations to these changing environmental conditions.

3.1. General objective

The general objective was to determine the genetic diversity of *Fragilariopsis kerguelensis* from four stations in the Southern Ocean using ddRAD (Peterson, et al., 2012).

3.2. Specific objectives

There were three specific objectives of this Master Thesis. First, to analyze the DNA sequences of four populations of *Fragilariopsis kerguelensis*. Second, to correlate these sequences with differences in environmental conditions among the stations. Finally, to identify possible genotypic adaptations of *Fragilariopsis kerguelensis* to these conditions.

4. Materials and Methods

4.1. Sampling and cultivation

Fragilariopsis kerguelensis samples were collected from the Stations 1, 2, 4, and 14 on the Southern Ocean during the Expedition PS103. An Apstein net with a 20 µm mesh size was used to sample *Fragilariopsis kerguelensis* from 20 m of depth to the surface. Approximately 200 strains were sampled, and 20 to 50 ml were kept at 0 to 4 °C (Beszteri, personal communication, May 17, 2018).

In a Utermöhl chamber, single chains of *Fragilariopsis kerguelensis* were obtained and then transferred to 12-well cell culture plates with pre-cooled culture medium F/2 (Table 10) (Guillard & Ryther, 1962). Within 48 hours from the initial isolation, each chain was transferred to fresh sterile cell culture plates to avoid the presence of non-target organisms. When the isolates reached dozens to hundreds of cells, they were transferred to15 ml Falcon tubes with F/2 medium and maintained in a cold container aboard at 4 °C and under a light cycle of 16 hours light and 8 hours dark. Eventually, these isolates were sent from Punta Arenas – Chile to Bremerhaven – Germany in a temperature controlled parcel at 4 °C (Beszteri, personal communication, May 17, 2018).

Once *Fragilariopsis kerguelensis* isolates arrived at Bremerhaven – Germany, they were cultivated in the Laboratory of Hustedt Diatom Study Centre, Polar Biological Oceanography Section at the Alfred-Wegener-Institut. The isolates were grown in 200 ml cell culture flasks with F/2 medium and maintained in a chamber at 4 °C and under illumination with Philips Master TL-D daylight fluorescent tubes of 15-24 µM PAR/m^2s. The light cycle was 14 hours light and 10 hours dark (Beszteri, personal communication, March 2, 2018).

4.2. Filtration and harvest

Strains of *Fragilariopsis kerguelensis* were randomly chosen. A membrane filter with a 5 µm pore size was used to harvest diatom biomass. After filtration, it was washed with the SL1 lysis buffer of the Nucleospin Soil DNA extraction kit (Macherey-Nagel, Düren – Germany) and snap frozen in liquid nitrogen. The lysis buffer containing the harvested cells was stored at -20 °C until further use (Beszteri, personal communication, March 2, 2018).

4.3. DNA extraction

The lysis buffer containing the harvested cells was thawed and transferred into lysis tubes of the Nucleospin Soil DNA extraction kit (Macherey-Nagel, Düren – Germany)

containing homogenization beads. Cells/frustules were opened by bead beating for 2 x 30 s and cooled on ice in between. Subsequent steps were performed according to the protocol provided by the company. DNA from the spin columns was eluted in 30-50 µl elution buffer and stored at -20 °C until further use (Beszteri, personal communication, March 2, 2018).

4.4. ddRAD library construction

The preparation of the ddRAD was a combination of the protocols of Hülskötter (2013); Peterson, et al. (2012) and Schweyen, et al. (2014).

4.4.1. Double digestion

The selected restriction enzymes were MspI and PstI (Thermo Fisher Scientific, Schwerte – Germany). These enzymes are capable to include the available range of recognition sequence length and G/C composition (Peterson, et al., 2012).

For a 45 µL reaction, 4.50 µL of 10X FastDigest Buffer; 2.25 µL of FastDigest PstI; 2.25 µL of FastDigest MspI (Thermo Fisher Scientific, Schwerte – Germany); and a variable volume of extract containing 1.00 µg of *Fragilariopsis kerguelensis* DNA were added, and filled up with nuclease-free water to 45 µL final volume (Table 12). The reaction was incubated for 45 minutes at 37 °C. Finally, the restriction enzymes were removed with the MinElute Reaction Cleanup Kit (QIAGEN, Hiden – Germany) according to the protocol provided by the company and the final product was eluted in 23 µL of nuclease-free water.

4.4.2. Adapter annealing

The adapters P5 and P7 were used (Eurofins, Ebersberg – Germany). For the adapters P5, there were 24 different barcodes and for the adapters P7, there were 4 different barcodes (Eurofins, Ebersberg – Germany). Adapters were synthesized as single-stranded DNA molecules which had to be annealed with their complementary counterparts before use. Each adapter P5 with a different barcode were joined together by heating 10 µL of the upper part; 10 µL of the appropriate lower part; 10 µL of 1X Annealing Buffer and 70 µL of nuclease-free water. The reaction was incubated for 5 minutes at 85 °C. Finally, the adapters were cooled at a rate not greater than 2 °C per minute until they reached room temperature (21 °C). The same procedure was repeated for the adapters P7.

4.4.3. Adapter ligation

For a 40 µL reaction, 22 µL resulted from a previous double digestion of DNA of *Fragilariopsis kerguelensis*; 4.00 µL of 1X T4 DNA Ligase Buffer; 0.80 µL of 0.20 µM P5 adapter; 4.00 µL of 1.00 µM P7 adapter; 0.50 µL of 25 U/µL T4 DNA Ligase (Eurofins,

Ebersberg – Germany and New England Biolabs, Frankfurt – Germany) were added, and filled up with nuclease-free water to 40 µ final volume (Table 13). The reaction was incubated for 2 hours at 16 °C. Finally, the ligase was removed with the MinElute Reaction Cleanup Kit (QIAGEN, Hiden – Germany) according to the protocol provided by the company and the final product was eluted in 52 µL of nuclease-free water.

4.4.4. Dual size selection

The targeted DNA fragment size was 500 to 1000 bp. To achieve this, two ratios of beads Agencourt® AMPure® XP (Beckman Coulter, Krefeld, Germany) were used, first 1:0.60 and afterward 1:0.40. By doing this, in the first size selection, fragments below 400 bp were eliminated, and in the second size selection, fragments above 1000 bp were eliminated.

For the first size selection with a beads ratio of 1:0.60, 30 µL of Agencourt® AMPure® XP (Beckman Coulter, Krefeld, Germany) was added to 50 µL resulted from a previous ligation of DNA fragments of *Fragilariopsis kerguelensis* and incubated at room temperature for 2 minutes. After this time, it was placed into a magnet plate for 5 minutes and the resulting supernatant was carefully removed and discarded. Immediately, a volume of 200 µL of ethanol 70 % was added for 1 minute and afterward removed and discarded. This step was repeated 2 times. After the 2 washes, the sample was taken out from the magnet plate until the formed pellet became matte. A volume of 28 µL of nuclease-free water was added, resuspended with the pellet and incubated at room temperature for 2 minutes. After this time, it was placed again into a magnet plate for 5 more minutes and 25 µL of the resulted supernatant was carefully removed. The beads were finally discarded.

For the second size selection with a beads ratio of 1:0.40, 10 µL of Agencourt® AMPure® XP (Beckman Coulter, Krefeld, Germany) was added to 25 µL resulted from a previous size selection of DNA fragments of *Fragilariopsis kerguelensis*. The same procedure mentioned was repeated, but the first resulted supernatant was conserved, and the beads were discarded.

4.4.5. PCR amplification

The primers P5 and P7 were used (Eurofins, Ebersberg – Germany) (Table 11). The products resulted from a previous double digestion, ligation and size selection of DNA fragments of 48 specimens of *Fragilariopsis kerguelensis* were amplified. For a 50 µL reaction, the following components were used: 1X Q5 DNA Polymerase Buffer; 0.20 mM dNTPs; 1 µM of P5 primer; 1 µM of P7 primer; 2 µL of the obtained sample and 0.50 µL of 0.02 U/µL

Q5 DNA Taq Polymerase ((Eurofins, Ebersberg – Germany and Thermo Fisher Scientific, Schwerte – Germany) (Table 14). The reaction was completed with DNA sample and H_2O.

The amplification conditions were an initial denaturation cycle at 98 °C for 30 seconds; 16 cycles of denaturation at 98 °C for 10 seconds, annealing at 65 °C for 30 seconds, elongation at 72 °C for 30 seconds; and a final extension cycle at 72 °C for 5 minutes. Finally, it was hold at 8 °C (Table 15).

4.4.6. Electrophoresis

A gel of agarose at 1.50% and 0.15 % ethidium bromide was prepared. The electrophoresis chamber was assembled correctly, the mixture was deposited in and, according to the number of samples, the wells were placed. Finally, 5 µL of PCR amplification products were loaded and run for 45 minutes at 120 V. After this time, these products were visualized under UV light.

4.4.7. Pooling and sequencing

The Agilent 2100 Bioanalyzer was used to check the quantity and purity of the DNA of each sample. The individual average size, concentration, and molarity were calculated using Agilent 2100 Expert Software B.02.09 (Agilent, 2017). However, for the concentration of the ddRAD library, the following formula was used:

$$[library] = \frac{\left[individual \frac{ng}{\mu L}\right]}{600 \frac{g}{mol} \times average\ size} \times 10^6$$

The final library concentration selected was 4 nM. Each sample was brought to this concentration, and 5 µl of each equimolar sample were pooled. Finally, the sequencing was done with the Illumina NextSeq 500 according to the protocol provided by the company.

4.5. ddRAD library processing

4.5.1. Demultiplex

The sequences obtained from the Illumina NextSeq 500 were multiplex sequences. The process_radtags program (Catchen, et al., 2013) was used to demultiplex and clean these sequences. The respective combination of DNA barcodes for each individual of *Fragilariopsis kerguelensis* was given and small errors on the final DNA sequence were allowed. Nonetheless, reads with the absence of barcode and restriction site, or low quality, were discarded.

4.5.2. Stacks

The software Stacks Version 2.0 Beta 9 (Catchen, et al., 2013) was used to process the demultiplexed data. Firstly, ustacks was used to construct loci. M is the maximum distance in nucleotides allowed between stacks. Different M values were tested to ensure the soundness of the analysis. The raw results of 12 random DNA samples of *Fragilariopsis kerguelensis* that contained loci shared by at least 80 % of the samples were considered. Then the number of SNPs was plotted against the percentage of loci (Rochette & Catchen, 2017).

Subsequently, cstacks was used to create a loci catalog. The loci obtained were aligned using the de novo approach since there is no a reference genome for *Fragilariopsis kerguelensis*. Later, sstacks was used to match the loci obtained with the loci catalog created. Finally, gstacks was used to assemble and combine paired-end, and to call SNPs (Catchen, et al., 2013).

4.5.3. Structure

The software Structure 2.3.4. (Pritchard, et al., 2000) was used to probabilistically assign specimens of *Fragilariopsis kerguelensis* to respective clusters, and thus to infer population structure. The alleles frequencies were estimated and computed to the likelihood of a genotype of each population. The clustering generated represented genetically different populations.

4.5.4. BayeScEnv

The software BayeScEnv (de Villemereuil & Gaggiotti, 2015) was used to identify possible genetic differences of *Fragilariopsis kerguelensis* and correlate them with differences in environmental conditions among different stations. Four environmental parameters were chosen according to their data availability (National Oceanic and Atmospheric Administration and measurement aboard the Polarstern) as well as ecological relevance. These were monthly average of December 2016 of sea surface temperature (SST), silicate, nitrate, and phosphate.

The software BayeScEnv considers environmental data as a form of 'environmental differentiation' which allows the identification of possible links between genetic differentiation and environmental parameters (de Villemereuil & Gaggiotti, 2015). This software is based on the Bayesian approach, so the fixation index (F_{st}) and posterior error probability (PEP) were calculated using the following formulas:

$$F_{st} = \frac{\sigma_S^2}{\bar{p}(1-\bar{p})}$$

σ^2_S= variance in the frequency of a specific allele among different subpopulations.

p= frequency of a specific allele in all the population.

$$PEP = 1 - P_i(M2\ a_i E)$$

i= locus.

M2= local adaptation model with environmental differentiation.

a_i= alleles count.

5. Results

A total of 48 DNA samples of *Fragilariopsis kerguelensis* were analyzed. The samples analyzed corresponded to 12 samples from Station 1; 9 samples from Station 2; 11 samples from Station 4, and 16 samples from Station 14. However, the sample No. 32 corresponding to ID 204, Strain ID 14-12 and Station 14, was later discarded due to its low mean coverage compared to the number of loci (Figure 13).

5.1. ddRAD library construction

5.1.1. Electrophoresis

The ddRAD library consisted mainly of different DNA fragments derived from the size selection efficiency. As shown in Figure 10, the size obtained was between 500 and 1000 bp. The negative control ensured that there was no external contamination in the reagents used and that the PCR products were derived from a successful amplification of DNA samples.

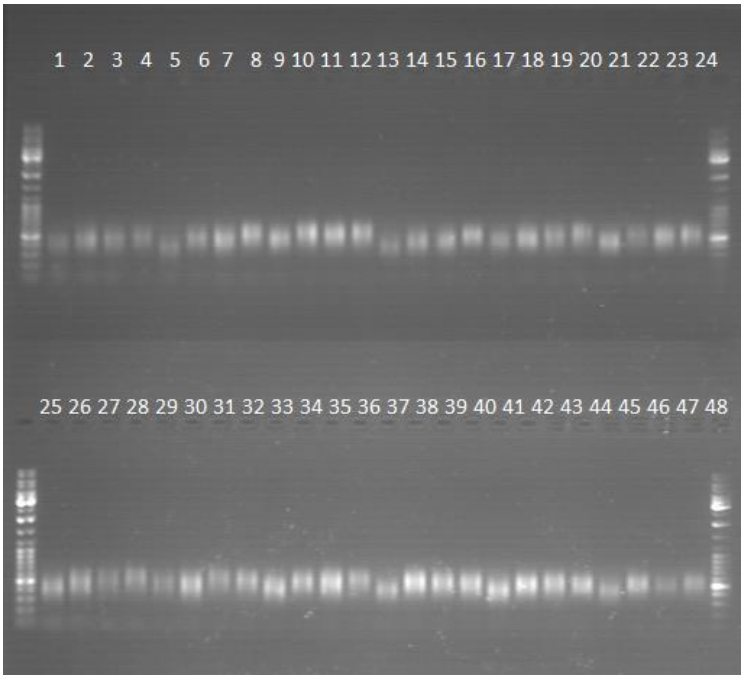

Figure 10. Electrophoresis of a gel of agarose at 1.50 % and 0.15 % ethidium bromide of 48 DNA samples of *Fragilariopsis kerguelensis*.

5.1.2. Pooling

The average size of the pool of 48 DNA samples was 518 bp. The concentration was 1693.42 pg/µl and the molarity 5357.30 pmol/l (Figure 11). Nonetheless, the molarity was again adjusted to the final library concentration selected of 4 nM.

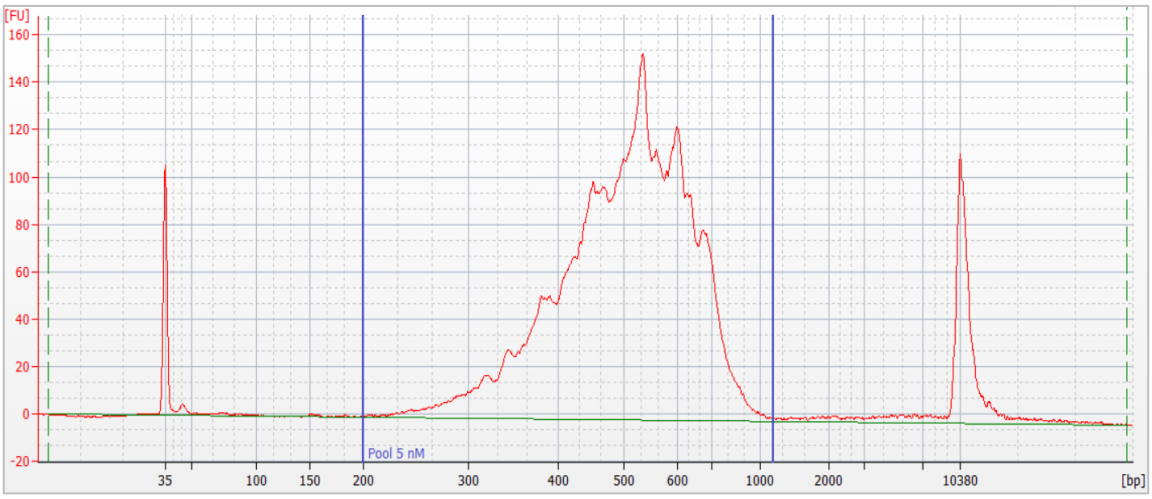

Figure 11. Electrogram of the pool of 48 DNA samples of *Fragilariopsis kerguelensis*.

5.2. ddRAD library processing

5.2.1. Stacks

For the loci construction, the distribution of the number of SNPs per locus was analyzed. The M values tested were 1, 2, 3, 4, 5, 6, 7, 8, and 9. As shown in Figure 12, the final M value selected was 5. This value allowed a stabilization of the proportion of loci, thus a higher coverage (Rochette & Catchen, 2017). Additionally, is important to mention that at least 60 % of the loci did not present any SNP (0).

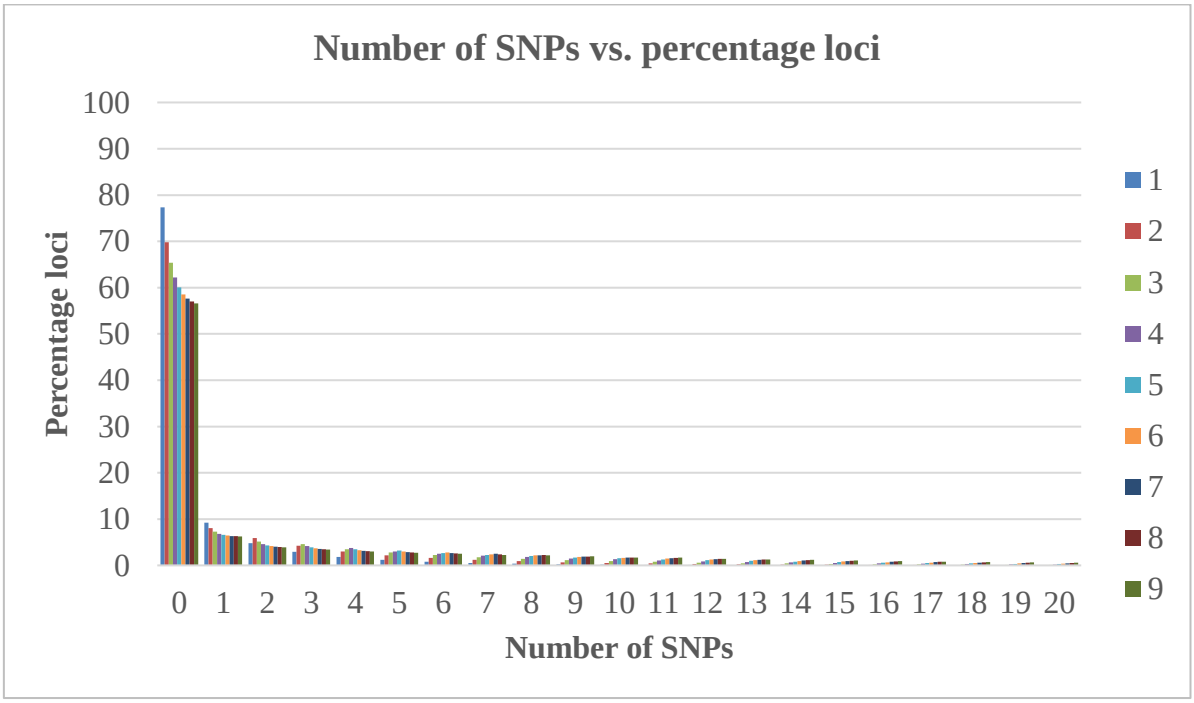

Figure 12. Number of SNPs vs. percentage loci of 12 random DNA samples of *Fragilariopsis kerguelensis*. M= 1, 2, 3, 4, 5, 6, 7, 8, and 9.

The highest number of loci obtained was 68129 which corresponded to individual fk_14_26 (Figure 13). The lowest number of loci permissible was 10000, so the individual fk_14_12 with 28 loci was discarded. As a result, 47 DNA samples of *Fragilariopsis kerguelensis* were further analyzed.

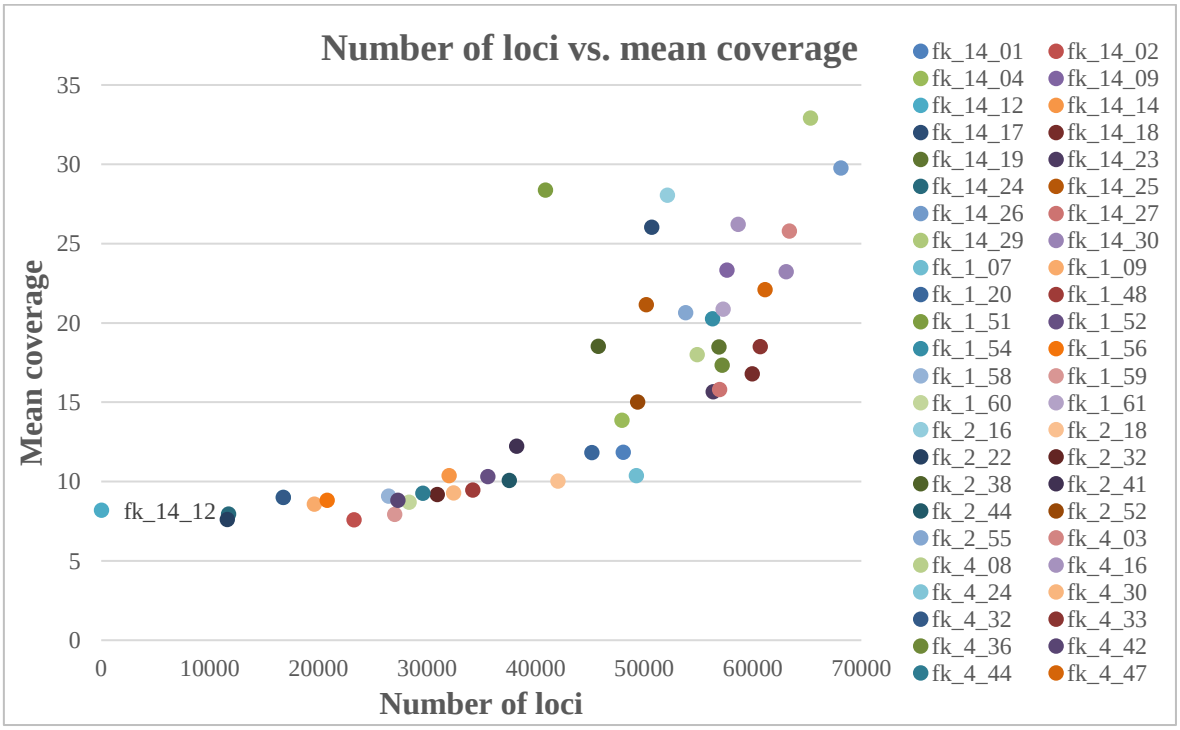

Figure 13. Number of loci vs. mean coverage of 48 DNA samples of *Fragilariopsis kerguelensis*.

5.2.2. Structure

The specimens of *Fragilariopsis kerguelensis* where clustered in three groups (Table 3 and Figure 14). Three Runs were performed, and the results obtained were supported by a probability higher than 0.90. Consequently, there was a clustering of northern populations which contained Station 1 and 2 (Cluster 3), and a clustering of southern populations which contained Station 4 and 14 (Cluster 1). In the case of Cluster 2, it contained specimens from the Stations 1, 2, 4, and 14.

In Figure 15, 16 and 17, there were not apparent morphological differences between representatives of the three different clusters (Cluster 1, 2, and 3). Nonetheless, individuals from fk_1_61 grouped in Cluster 3, seem to have a bigger size compared with the individuals grouped in Cluster 1 and 2 and also elongated shape. Contrary, the individuals from fk_1_20 grouped in Cluster 2, seem to have a smaller size and a more rounded shape.

Table 3. Clustering of the individuals of *Fragilariopsis kerguelensis* according to probability in Run 1, 2 and 3. Orange represents Cluster 1; green Cluster 2; and blue Cluster 3.

Sample ID	Run 1	Run 2	Run 3	Run 1	Run 2	Run 3	Run 1	Run 2	Run 3
	Cluster 1			Cluster 2			Cluster 3		
fk_1_07	0.004	0.002	0.002	0.994	0.996	0.996	0.002	0.002	0.002
fk_1_09	0.003	0.005	0.003	0.006	0.004	0.001	0.991	0.991	0.996
fk_1_20	0.005	0.003	0.004	0.993	0.993	0.992	0.002	0.005	0.003
fk_1_48	0.002	0.002	0.001	0.001	0.001	0.001	0.997	0.997	0.998
fk_1_51	0.004	0.001	0.001	0.002	0.002	0.003	0.994	0.997	0.996
fk_1_52	0.002	0.002	0.001	0.002	0.001	0.002	0.997	0.997	0.997
fk_1_54	0.002	0.001	0.001	0.001	0.002	0.001	0.997	0.997	0.998
fk_1_56	0.002	0.002	0.001	0.001	0.001	0.001	0.997	0.997	0.998
fk_1_58	0.002	0.002	0.001	0.003	0.001	0.001	0.996	0.996	0.998
fk_1_59	0.001	0.001	0.002	0.998	0.996	0.996	0.001	0.003	0.002
fk_1_60	0.001	0.002	0.002	0.002	0.002	0.001	0.996	0.997	0.997
fk_1_61	0.001	0.001	0.001	0.001	0.001	0.001	0.998	0.998	0.998
fk_2_16	0.001	0.001	0.002	0.001	0.001	0.002	0.997	0.997	0.997
fk_2_18	0.002	0.001	0.001	0.001	0.002	0.001	0.998	0.997	0.998
fk_2_22	0.005	0.005	0.002	0.991	0.990	0.991	0.005	0.006	0.007
fk_2_32	0.001	0.001	0.001	0.001	0.001	0.001	0.997	0.998	0.998
fk_2_38	0.001	0.002	0.002	0.002	0.003	0.002	0.997	0.995	0.996
fk_2_41	0.002	0.002	0.002	0.003	0.003	0.003	0.995	0.995	0.994
fk_2_44	0.002	0.002	0.001	0.998	0.997	0.997	0.001	0.001	0.002
fk_2_52	0.002	0.001	0.002	0.002	0.003	0.001	0.996	0.996	0.997
fk_4_03	0.003	0.003	0.003	0.995	0.991	0.994	0.002	0.006	0.003
fk_4_08	0.994	0.995	0.994	0.003	0.002	0.003	0.003	0.003	0.002
fk_4_16	0.990	0.996	0.995	0.007	0.002	0.004	0.003	0.003	0.001
fk_4_24	0.917	0.866	0.911	0.072	0.121	0.079	0.011	0.012	0.010
fk_4_30	0.978	0.977	0.989	0.015	0.018	0.005	0.007	0.005	0.006
fk_4_32	0.996	0.996	0.998	0.002	0.003	0.001	0.002	0.001	0.001
fk_4_33	0.002	0.002	0.002	0.996	0.997	0.996	0.001	0.002	0.003
fk_4_36	0.998	0.997	0.998	0.001	0.001	0.001	0.001	0.002	0.001
fk_4_42	0.997	0.996	0.996	0.002	0.002	0.002	0.001	0.002	0.002
fk_4_44	0.997	0.995	0.994	0.002	0.003	0.003	0.001	0.002	0.003
fk_4_47	0.991	0.992	0.995	0.006	0.004	0.004	0.003	0.004	0.002
fk_14_01	0.996	0.997	0.996	0.002	0.002	0.003	0.002	0.001	0.001
fk_14_02	0.994	0.995	0.994	0.004	0.004	0.003	0.002	0.001	0.003
fk_14_04	0.998	0.996	0.996	0.001	0.002	0.003	0.001	0.002	0.001
fk_14_09	0.998	0.998	0.997	0.001	0.001	0.003	0.001	0.001	0.001
fk_14_14	0.003	0.003	0.002	0.991	0.989	0.992	0.006	0.008	0.007
fk_14_17	0.996	0.996	0.996	0.002	0.002	0.003	0.002	0.002	0.001
fk_14_18	0.011	0.003	0.007	0.986	0.992	0.990	0.004	0.005	0.003
fk_14_19	0.003	0.003	0.002	0.994	0.995	0.996	0.002	0.002	0.002
fk_14_23	0.003	0.002	0.001	0.995	0.996	0.998	0.003	0.002	0.001
fk_14_24	0.003	0.007	0.003	0.994	0.985	0.996	0.002	0.008	0.001
fk_14_25	0.995	0.996	0.998	0.002	0.001	0.001	0.003	0.003	0.001
fk_14_26	0.996	0.996	0.994	0.002	0.002	0.002	0.001	0.002	0.003
fk_14_27	0.004	0.005	0.003	0.995	0.993	0.995	0.001	0.002	0.002
fk_14_29	0.995	0.995	0.995	0.002	0.003	0.002	0.003	0.001	0.003
fk_14_30	0.996	0.997	0.996	0.003	0.002	0.002	0.001	0.001	0.002

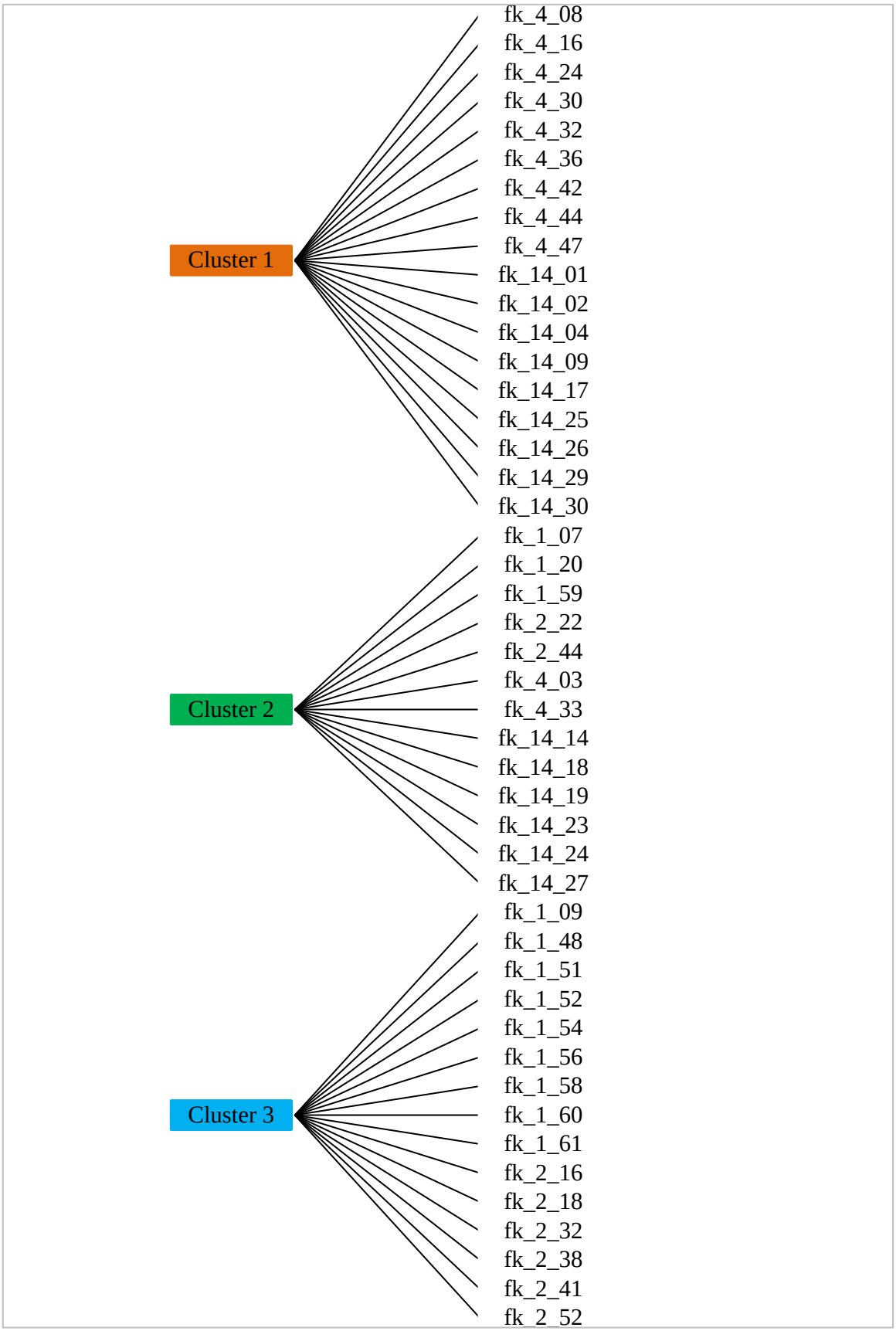

Figure 14. Clustering of the individuals of *Fragilariopsis kerguelensis*. Orange represents Cluster 1; green Cluster 2; and blue Cluster 3.

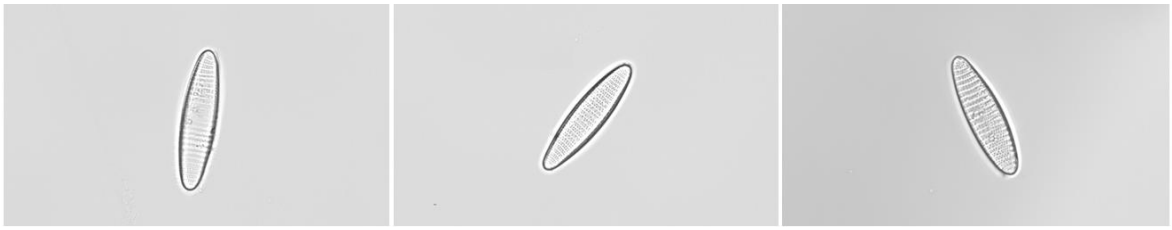

Figure 15.Individuals from fk_14_25 grouped in Cluster 1 taken at 10x magnification, provided by Hustedt Diatom Study Centre.

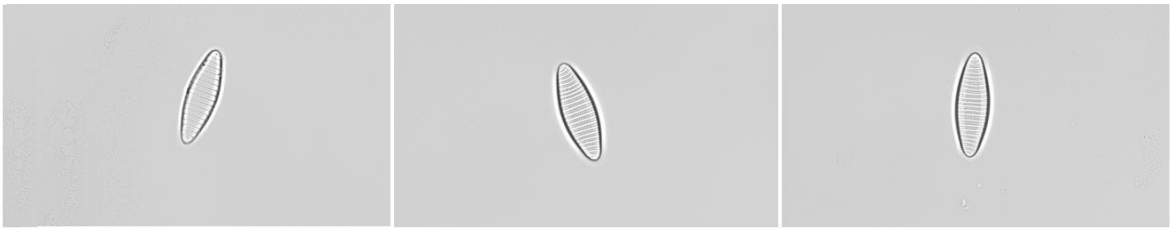

Figure 16.Individuals from fk_1_20 grouped in Cluster 2 taken at 10x magnification, provided by Hustedt Diatom Study Centre.

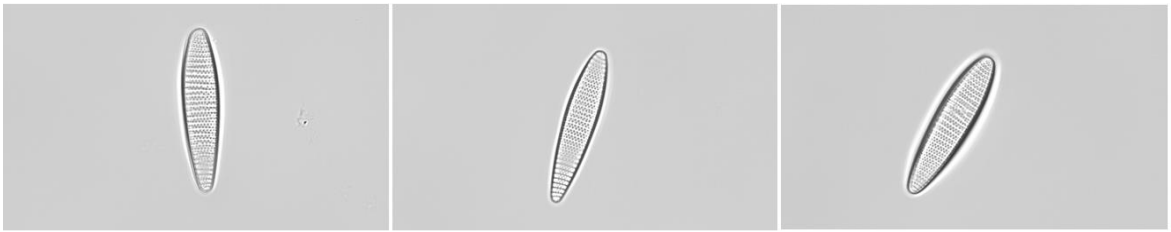

Figure 17. Individuals from fk_1_61 grouped in Cluster 3 taken at 10x magnification, provided by Hustedt Diatom Study Centre.

5.3. Distance analysis between stations

The level of genetic differentiation among populations of *Fragilariopsis kerguelensis* was compared with the geographical distance between stations. The highest fixation index (F_{st}) were found for Station 2 vs. Station 4 and Station 1 vs. Station 4. The F_{st} values obtained were 0.11 and 0.10 respectively (Figure 18). The distance between Station 2 and 4 was 670 km, while between Station 1 and 4 was 1167 km (Table 4).

The lowest F_{st} values were found for Station 4 vs. Station 14 and Station 1 vs. Station 2. The F_{st} values obtained were 0.02 and 0.03 respectively (Figure 18). The distance between Station 4 and 14 was 1446 km, while between Station 1 and 2 was 538 km (Table 4). Finally, the total distance between the first (1) and the last (14) station was 2581 km (Table 4 and Figure 19).

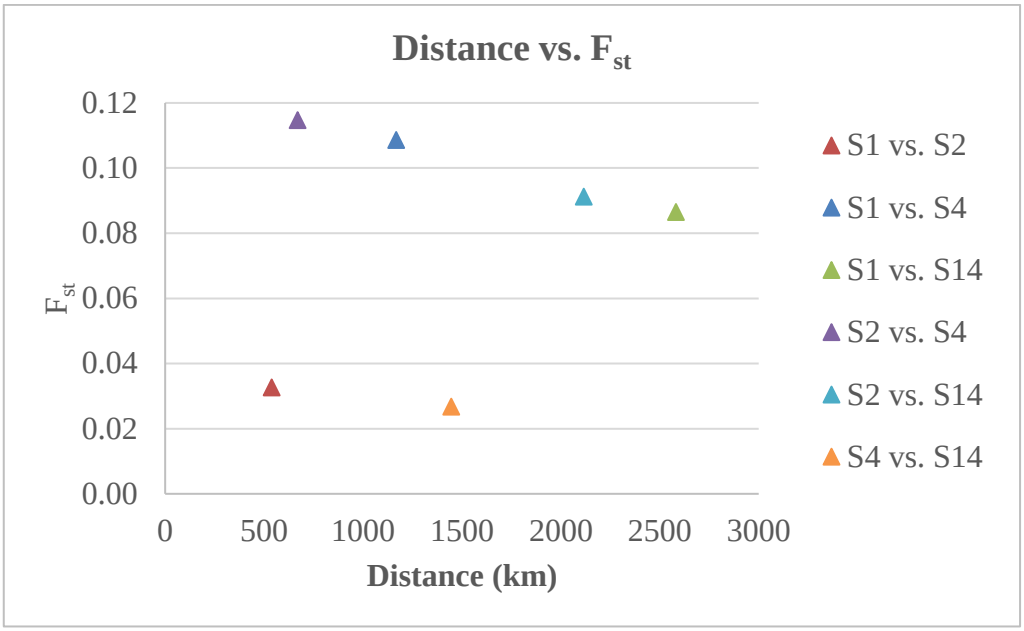

Figure 18. Distance (km) vs. F_{st} value of the individuals of Cluster 2 *Fragilariopsis kerguelensis* of the Stations 1, 2, 4, and 14. "S" represents Station.

Table 4. Distance between stations where specimens of *Fragilariopsis kerguelensis* were collected.

Station	Distance (km)			
	1	2	4	14
1	-	538	1167	2581
2	538	-	670	2116
4	1167	670	-	1446
14	2581	2116	1446	-

Figure 19. Total distance between the first (1) and the last (14) station (2581 km), modified from Google Maps (2017).

5.4. Environmental parameters between stations

According to the data availability as well as ecological relevance, the environmental parameters of sea surface temperature (SST), silicate, nitrate, and phosphate were chosen. Data of SST (°C) was taken from cruise PS103 aboard the Polarstern, while data of silicate, nitrate, and phosphate (μM) were taken from National Oceanic and Atmospheric Administration (NOAA) (2018).

The SST was <6 °C in all four stations (Table 5 and Figure 20). In the case of the macronutrients silicate (Table 5 and Figure 21), nitrate (Table 5 and Figure 22), and phosphate (Table 5 and Figure 23), their concentration increased with higher latitudes. However, for Station 4 that was located close to the southern boundary ACC, this apparent tendency changed significantly.

Table 5. Environmental parameters data from 1 to 31 of December 2016 in the Stations 1, 2, 4, and 14. Source National Oceanic and Atmospheric Administration (2018) and measurement aboard the Polarstern.

Environmental parameters		Station			
		SAF	PF	Southern boundary ACC	Antarctica
		1	2	4	14
SST	°C	5.3	4.4	0.6	1.9
Silicate	μM	6.1381	15.7566	53.6392	66.5772
Nitrate	μM	18.6211	22.0355	26.8446	24.1721
Phosphate	μM	1.3569	1.7157	1.8887	1.7234

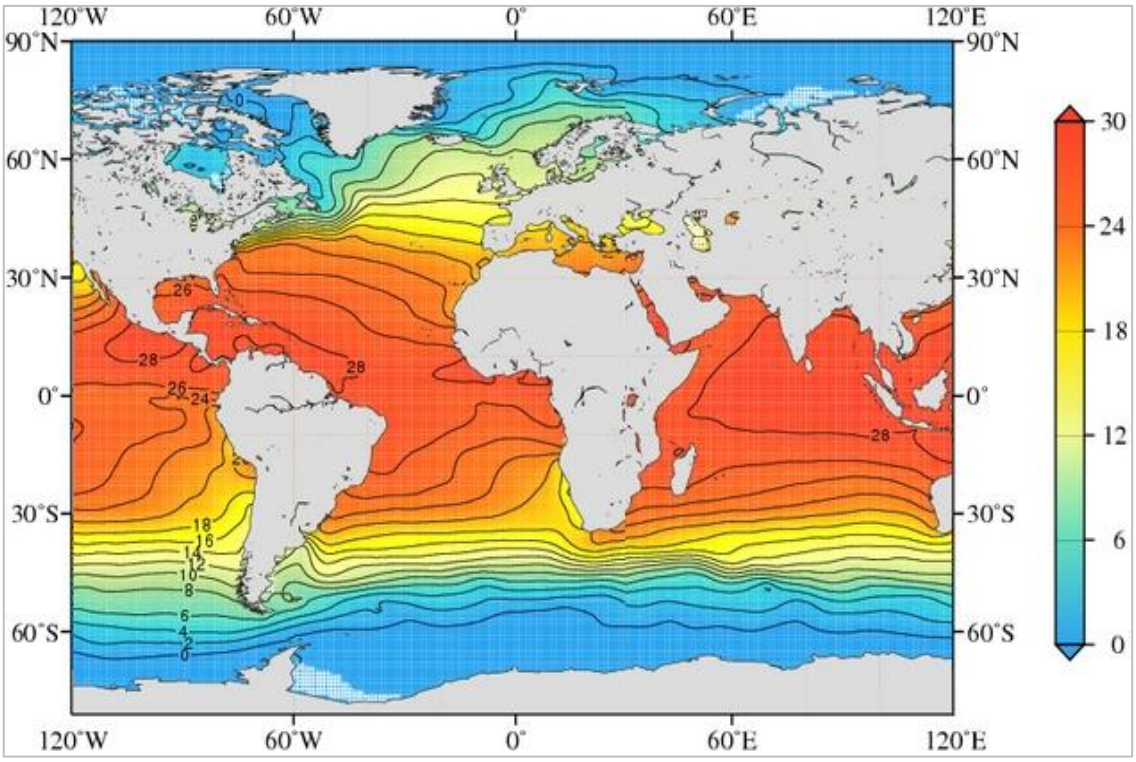

Figure 20. Sea surface temperature (SST) (°C) in the SO during December, modified from National Oceanic and Atmospheric Administration (2018).

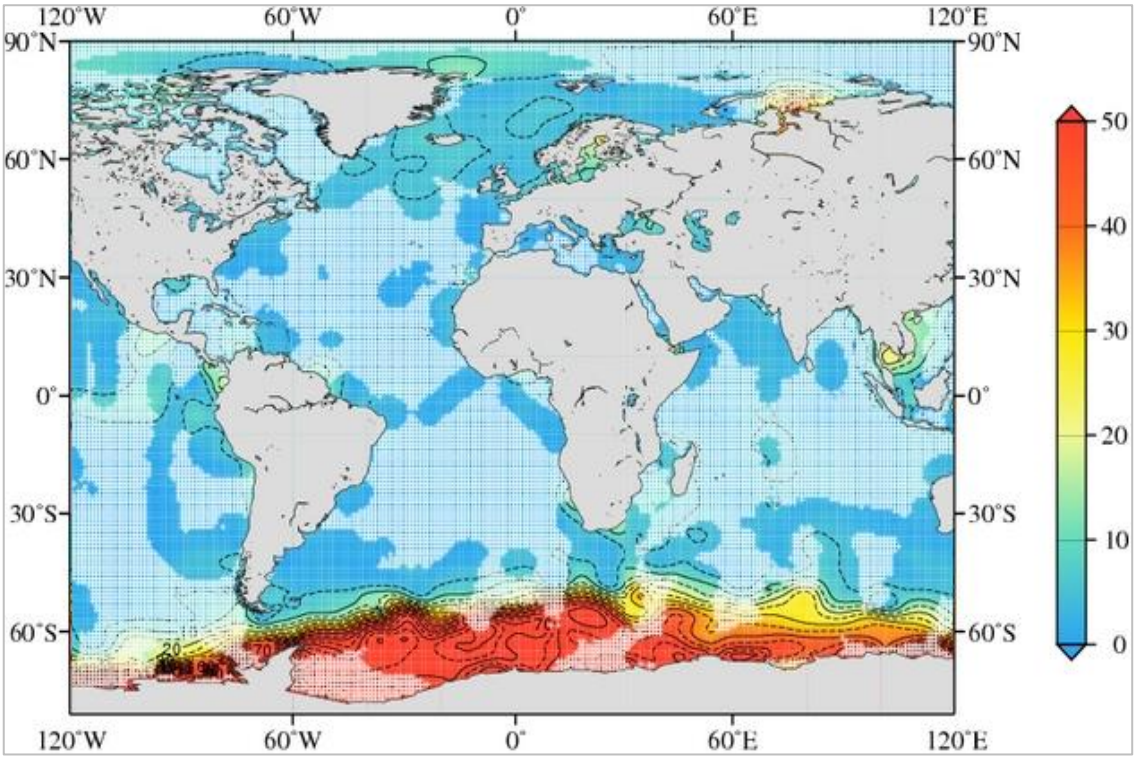

Figure 21. Silicate concentration (μM) in the SO during December, modified from National Oceanic and Atmospheric Administration (2018).

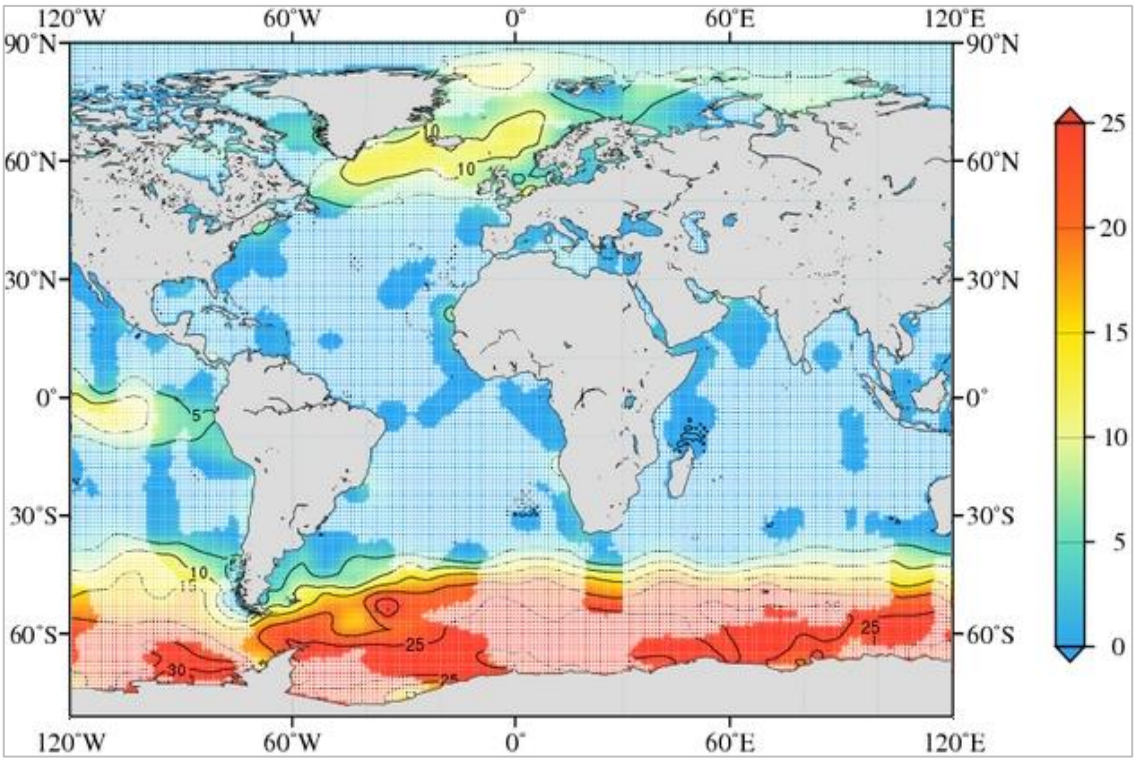

Figure 22. Nitrate concentration (µM) in the SO during December, modified from National Oceanic and Atmospheric Administration (2018).

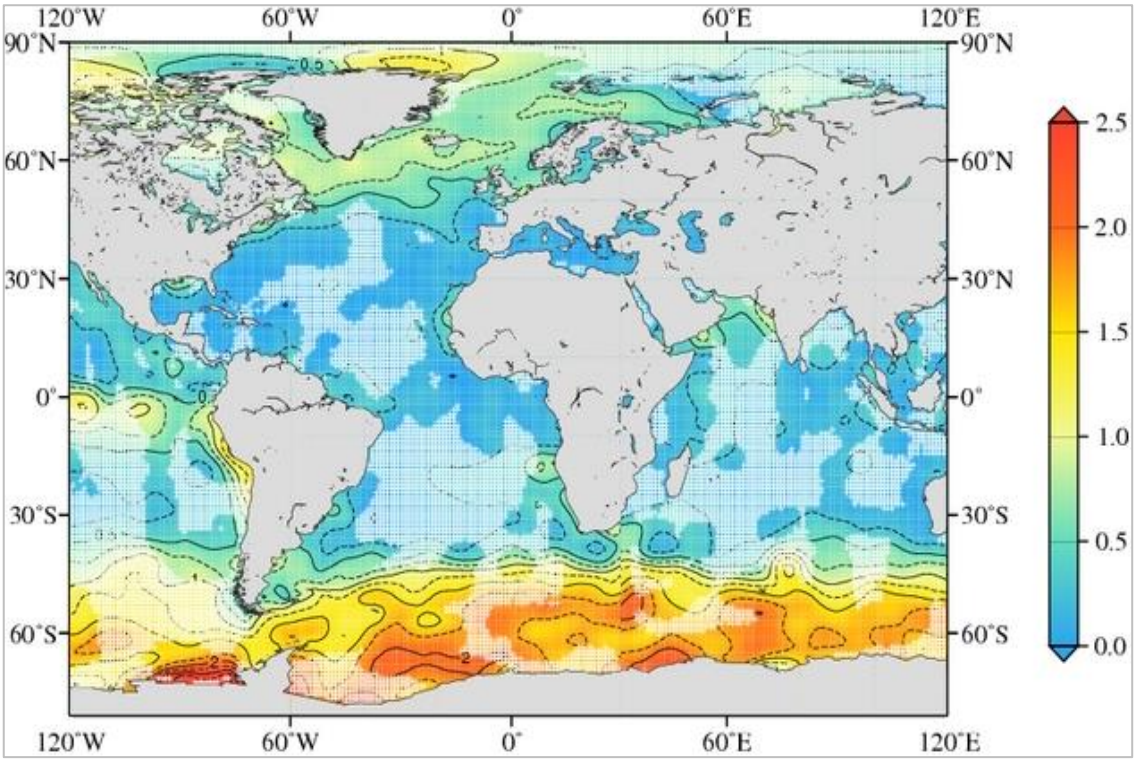

Figure 23. Phosphate concentration (µM) in the SO during December, modified from National Oceanic and Atmospheric Administration (2018).

5.5. Correlation analysis between environmental parameters and loci

The analysis of sea surface temperature (SST) was carried out twice (Run 1 and Run 2) (Figure 24). The results obtained in both runs confirmed a congruence in the methodology used. The highest F_{st} values were 0.69 corresponding to locus 1496 and 0.62 corresponding to locus 765 (Figure 25). Contrary, the lowest F_{st} value obtained was 0.07 and values <0.20 were found on 3757 loci, which corresponded to 99.26 % of the total number of loci.

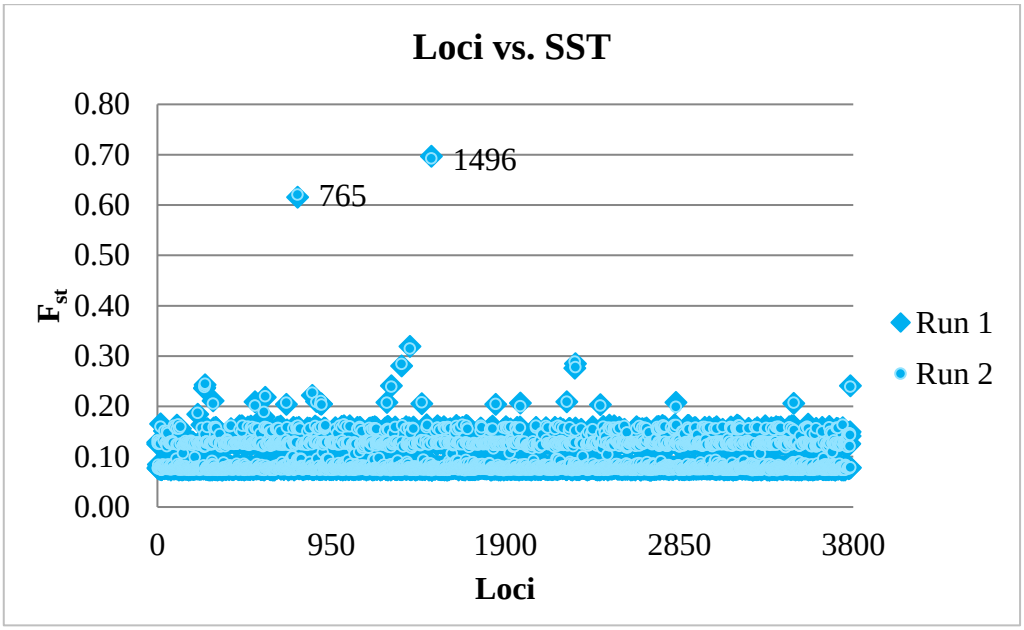

Figure 24. Comparison between Run 1 and 2 of loci and F_{st} in sea surface temperature (SST) of the individuals of *Fragilariopsis kerguelensis*.

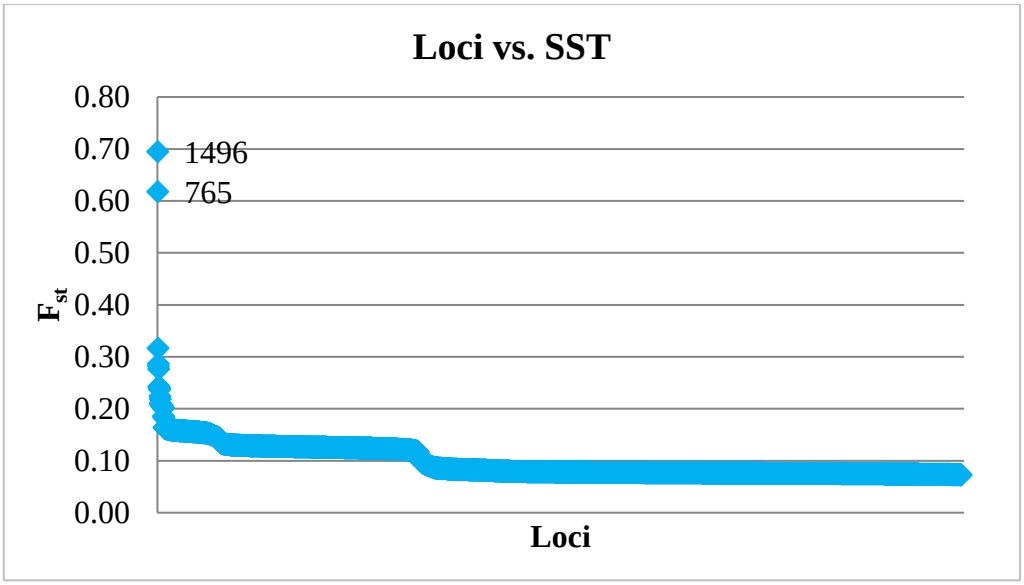

Figure 25. Loci vs. largest to smallest F_{st} in sea surface temperature (SST) of the individuals of *Fragilariopsis kerguelensis*.

Run 1 and Run 2 were carried out for the analysis of silicate (Figure 26). The results obtained in both runs were similar and consistent. For the environmental parameter of silicate, the highest F_{st} value was 0.56 corresponding to locus 1156 (Figure 27). Aversely, the lowest F_{st} value obtained was 0.06 and values <0.20 were found on 3626 loci, which corresponded to 95.80 % of the total number of loci. The remaining loci had F_{st} values between 0.39 and 0.20, which corresponded to 4.17 %.

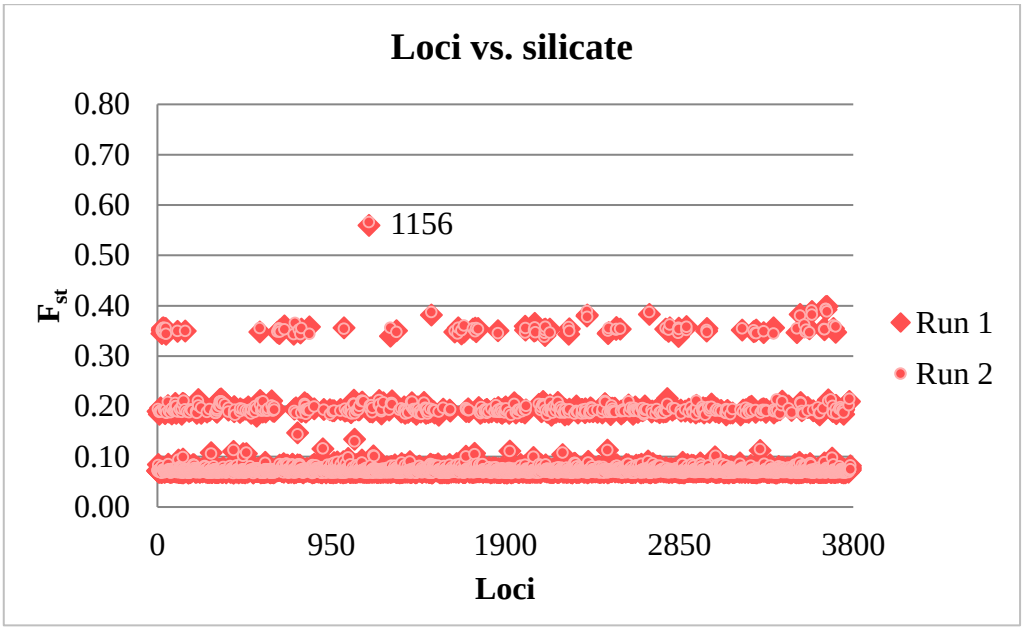

Figure 26. Comparison between Run 1 and 2 of loci and F_{st} in silicate of the individuals of *Fragilariopsis kerguelensis*.

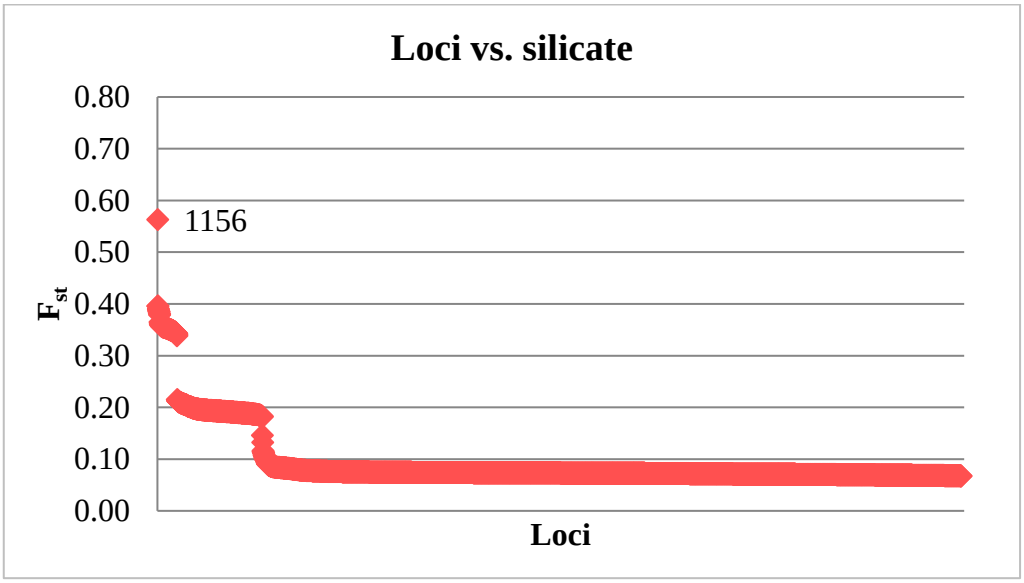

Figure 27. Loci vs. largest to smallest F_{st} in silicate of the individuals of *Fragilariopsis kerguelensis*.

The analysis of nitrate was carried out two times (Run 1 and Run 2) (Figure 28). The results obtained in both runs confirmed once again a congruence in the methodology used. In this case, the highest F_{st} values were 0.61 corresponding to locus 1496 and 0.54 corresponding to locus 765 (Figure 29). Contrary, the lowest F_{st} value obtained was 0.07 and values <0.20 were found on 3771 loci, which corresponded to 99.63 % of the total number of loci.

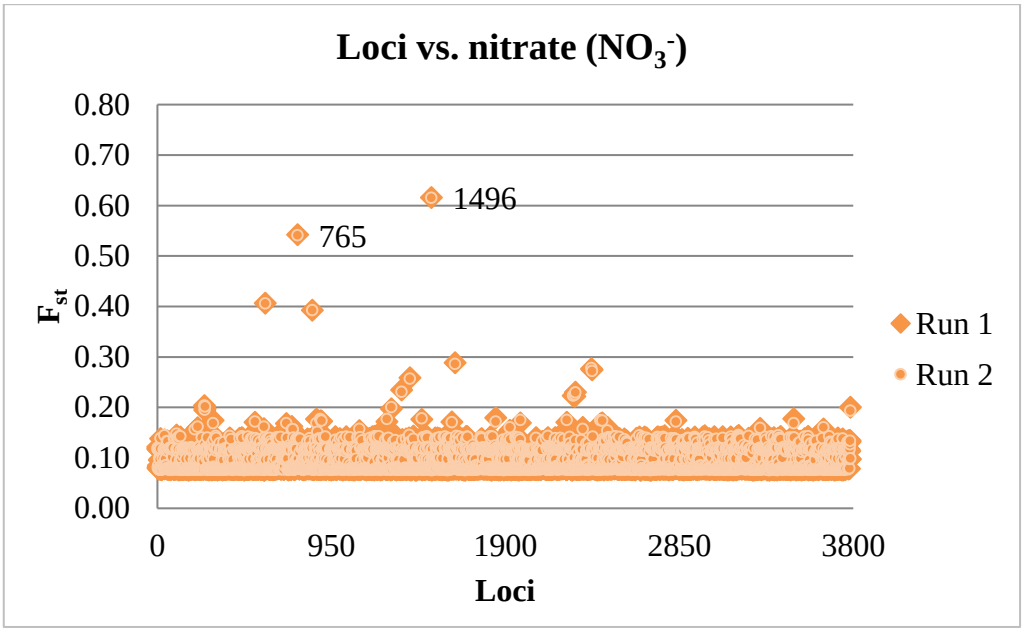

Figure 28. Comparison between Run 1 and 2 of loci and F_{st} in nitrate (NO_3^-) of the individuals of *Fragilariopsis kerguelensis*.

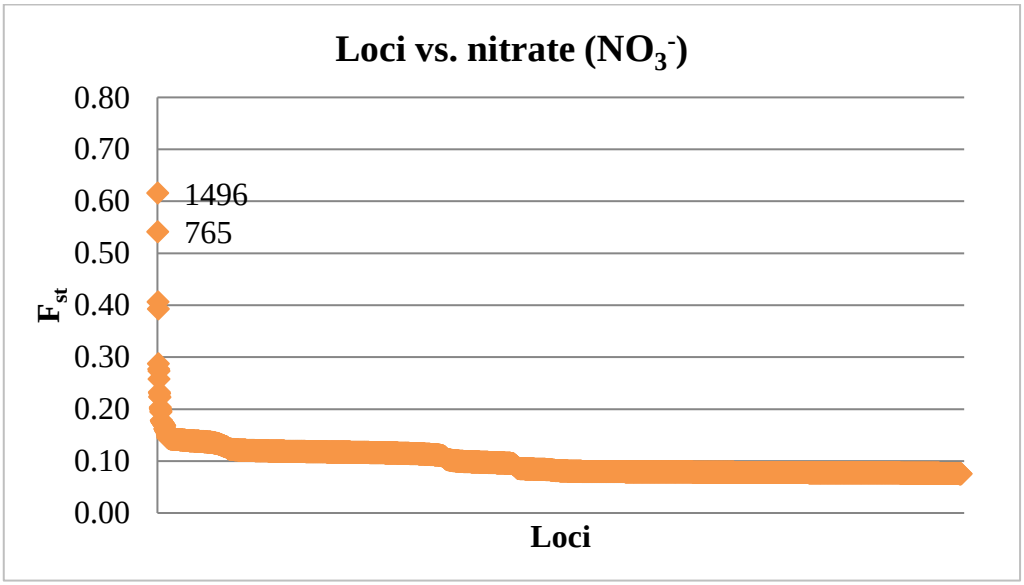

Figure 29. Loci vs. largest to smallest F_{st} in nitrate (NO_3^-) of the individuals of *Fragilariopsis kerguelensis*.

Run 1 and Run 2 were carried out for the analysis of phosphate (Figure 26). The results obtained in both runs were similar and consistent. For phosphate, the highest F_{st} values were 0.58 corresponding to locus 1496 and 0.51 corresponding to locus 765 (Figure 31). Aversely, the lowest F_{st} value obtained was 0.07 and the remaining loci had Fst values between 0.41 and 0.08. Values <0.20 were found on 3773 loci, which corresponded to 99.68 % of the total number of loci.

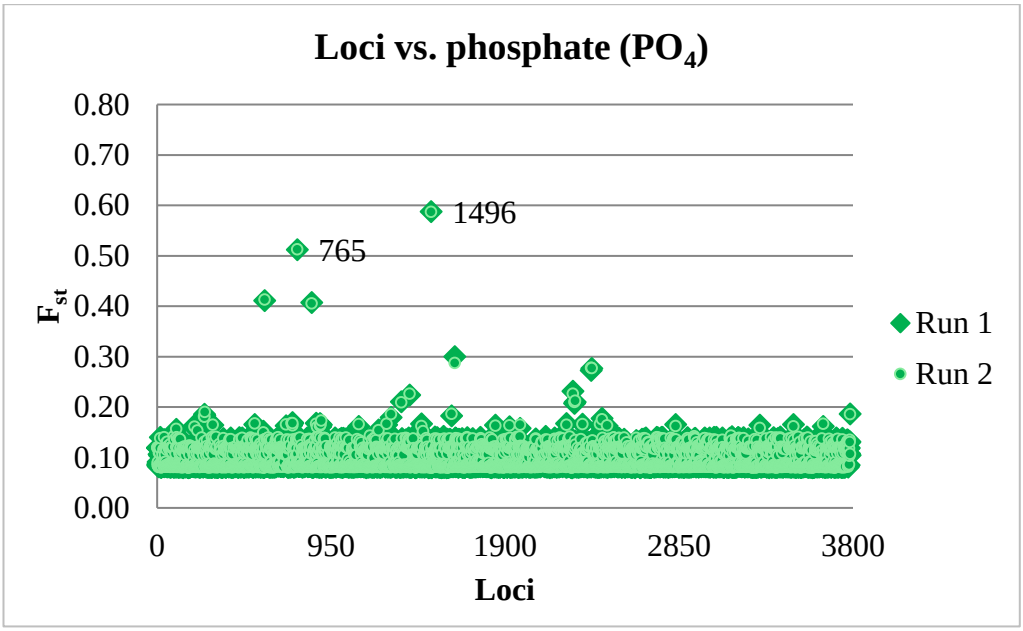

Figure 30. Comparison between Run 1 and 2 of loci and F_{st} in phosphate (PO_4) of the individuals of *Fragilariopsis kerguelensis*.

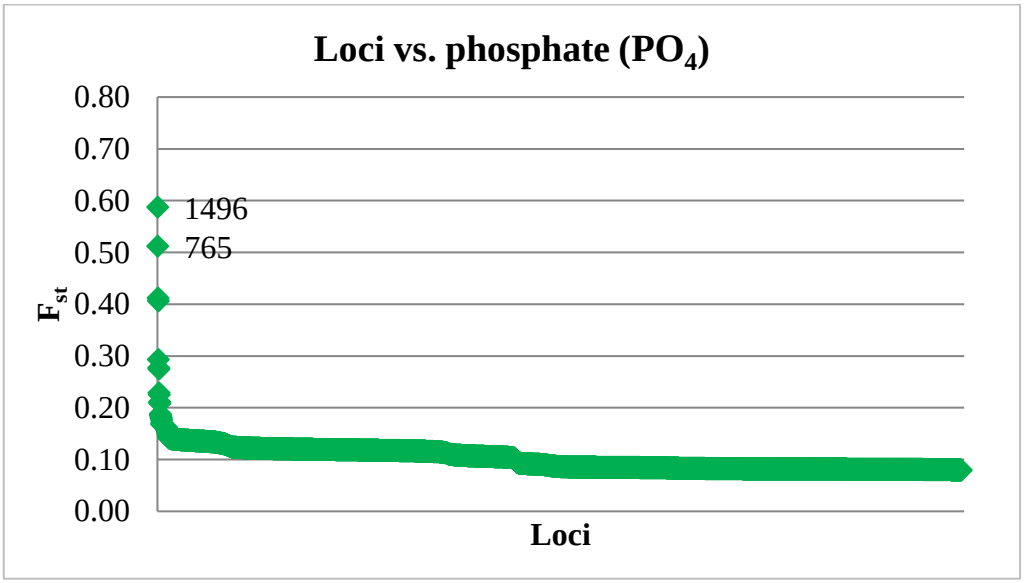

Figure 31. Loci vs. largest to smallest F_{st} in phosphate (PO_4) of the individuals of *Fragilariopsis kerguelensis*.

Table 6. Five of the lowest values of PEP according to Run 1 and 2 of sea surface temperature (SST). The highlighted values are lower than 0.25.

Sea surface temperature (SST)			
Run 1		Run 2	
Loci	PEP	Loci	PEP
1496	0.0286	1496	0.0342
765	0.1346	765	0.1276
1379	0.5979	1379	0.6035
2283	0.6499	2280	0.6429
1334	0.6565	1334	0.6499

Table 7. Five of the lowest values of PEP according to Run 1 and 2 of silicate. The highlighted values are lower than 0.25.

Silicate			
Run 1		Run 2	
Loci	PEP	Loci	PEP
1156	0.2407	1156	0.2291
1496	0.3511	1496	0.3403
3657	0.4715	3650	0.4777
3650	0.4747	3652	0.4803
3652	0.4807	3573	0.4827

Table 8. Five of the lowest values of PEP according to Run 1 and 2 of nitrate (NO_3^-). The highlighted values are lower than 0.25.

Nitrate (NO_3^-)			
Run 1		Run 2	
Loci	PEP	Loci	PEP
1496	0.0458	1496	0.0462
765	0.1560	765	0.1564
588	0.3621	588	0.3639
845	0.3817	845	0.3815
1626	0.5825	1626	0.5829

Table 9. Five of the lowest values of PEP according to Run 1 and 2 of phosphate (PO_4). The highlighted values are lower than 0.25.

Phosphate (PO_4)			
Run 1		Run 2	
Loci	PEP	Loci	PEP
1496	0.0572	1496	0.0538
765	0.1844	765	0.1830
588	0.3415	588	0.3423
845	0.3459	845	0.3465
1626	0.5531	1626	0.5757

The highest PEP accepted was 0.25. According to this, the loci 1496 and 765 have the lowest PEP for sea surface temperature (SST), nitrate, and phosphate (Table 6, Table 7, and Table 8). However, as shown in Table 7, a different locus (1156) had the lowest PEP for silicate. Therefore, the loci 1496 and 765 were selected, as well as locus 1156, for further analysis (Figure 32).

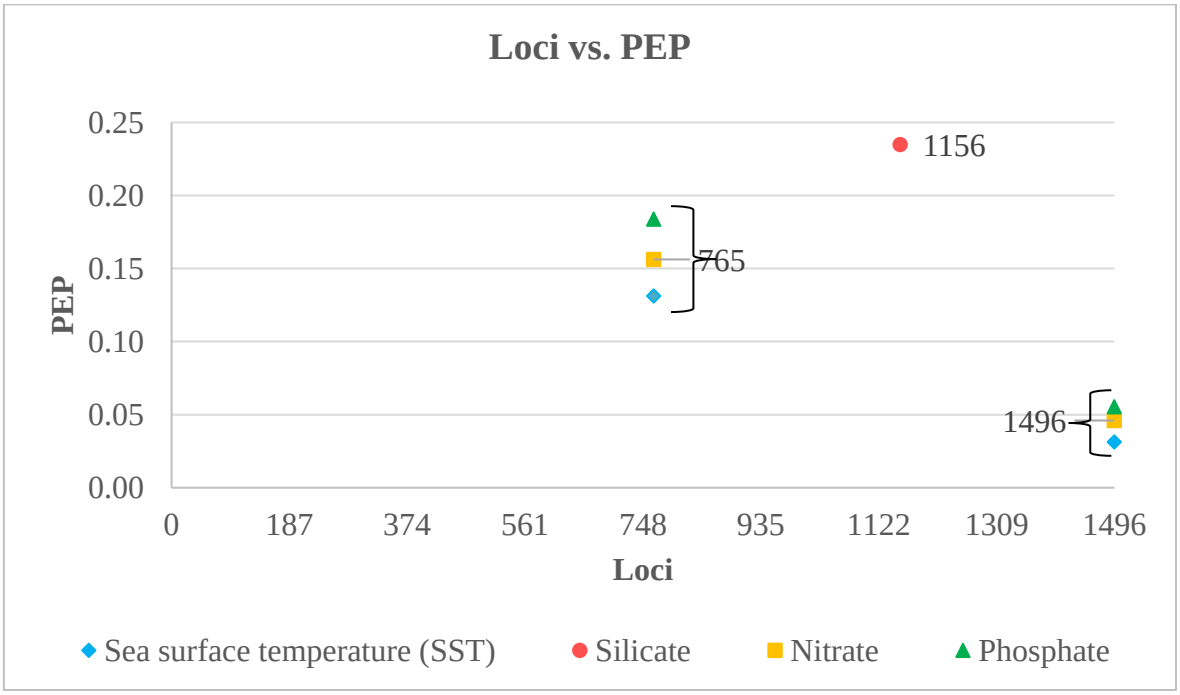

Figure 32. Loci vs. PEP in four environmental parameters of the individuals of *Fragilariopsis kerguelensis*.

The following figures are an overview of the alignment of the nucleotides sequences (query sequences) to the database sequences (subject sequences). The locus 1496 had the highest F_{st} value for the environmental parameters of SST, nitrate, and phosphate. The nucleotides sequences had 241 letters (Table 16) and its Blast analysis created 178 hits (Figure 33). 177 hits had an alignment score between 40-80 so they cannot be considered as specific but more as a random noise. The only exception was one hit that has an alignment score of 80-200 which corresponded to "*Phaeodactylum tricornutum* CCAP 1055/1 predicted protein partial mRNA".

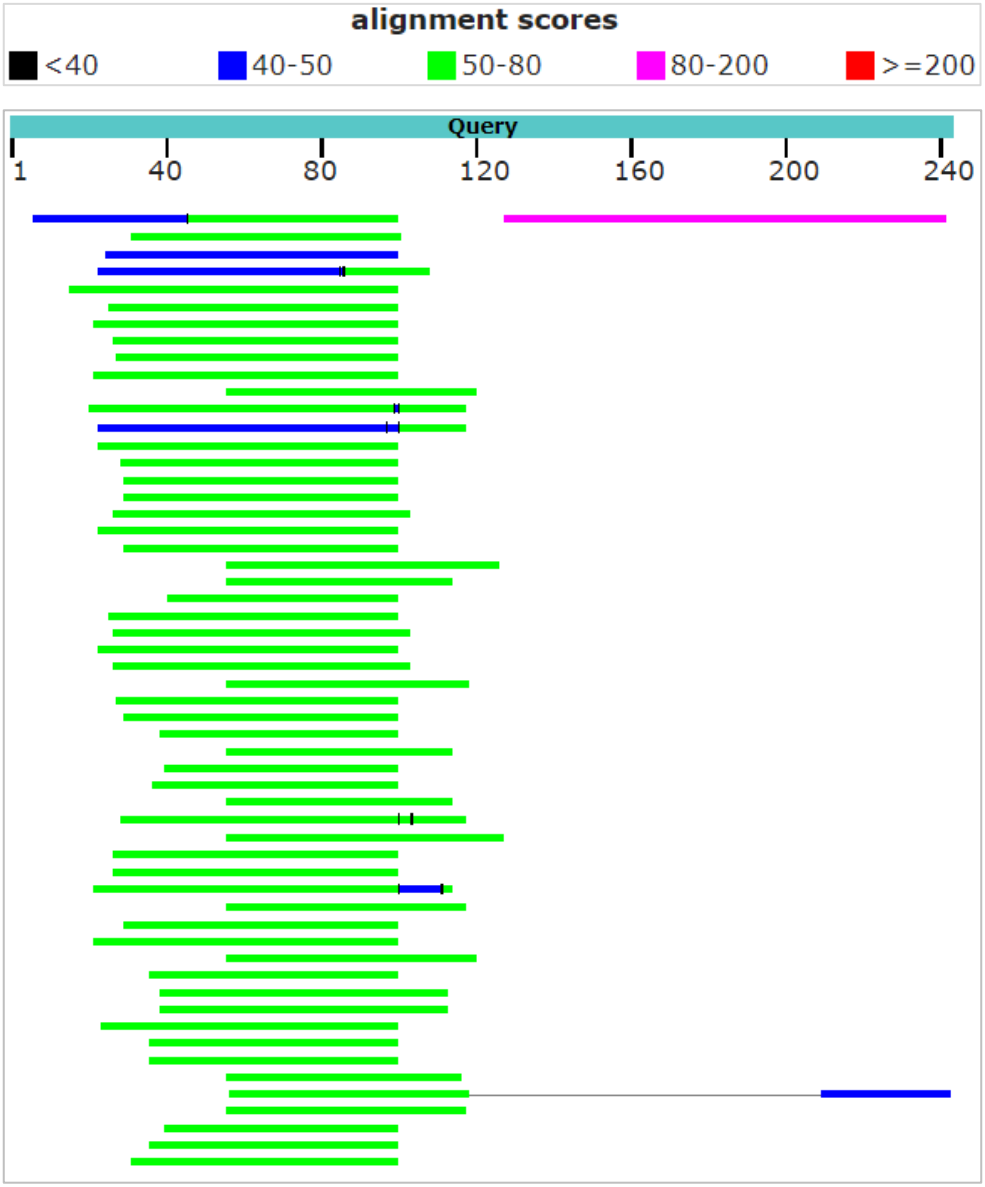

Figure 33. Distribution of 178 blast hits on subject sequences of the locus 1496 (18876_67) related to the environmental parameters of SST, nitrate, and phosphate.

The locus 765 had the second highest F_{st} value for the environmental parameters of SST, nitrate, and phosphate. The nucleotides sequence had 232 letters (Table 17) and its Blast analysis created 17 hits where each alignment had a score or 40-50 (Figure 34).

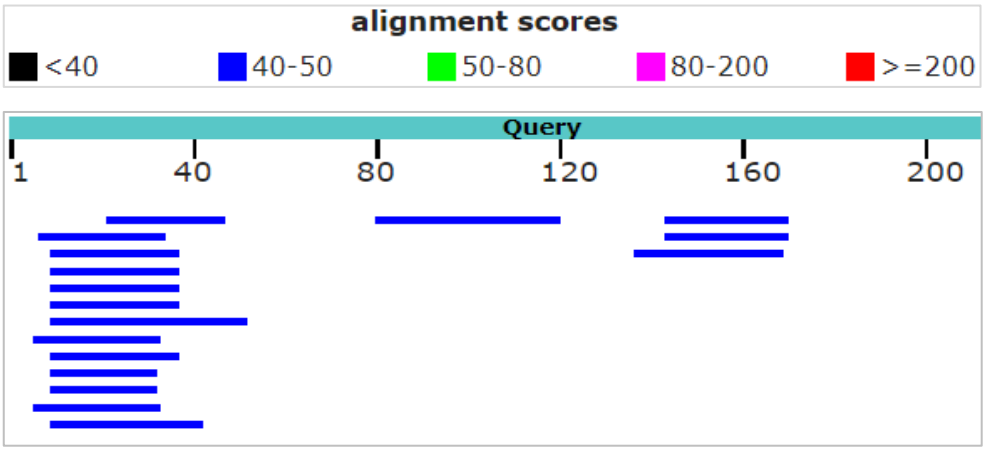

Figure 34. Distribution of 17 blast hits on subject sequences of the loci 765 (9600_47) related to the environmental parameters of SST, nitrate, and phosphate.

The locus 1156 had the highest F_{st} value for the environmental parameter of silicate. The nucleotides sequence had 190 letters (Table 18) and its Blast analysis created 15 hits where each alignment had a score or 40-50 (Figure 35).

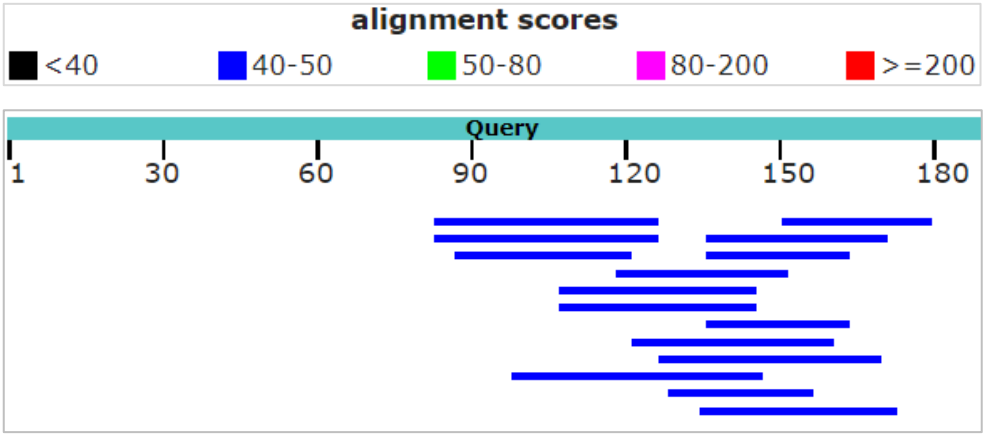

Figure 35. Distribution of 15 blast hits on subject sequences of the loci 1156 (13469_173) related to environmental parameter silicate. 190 letters.

6. Discussion

The potential genetic differentiation of *Fragilariopsis kerguelensis* to four environmental parameters such as sea surface temperature (SST), silicate, nitrate, and phosphate, was analyzed through ddRAD. The restriction enzymes MspI and PstI were selected due to the size and GC content of the genome of *Fragilariopsis kerguelensis* (Schweyen, et al., 2014).

6.1. *Fragilariopsis kerguelensis* clustering

ddRAD allows the analysis of patterns of genetic differentiation between diverse populations (Schweyen, et al., 2014). The 47 individuals under *Fragilariopsis kerguelensis* sampled from four stations in the SO, were grouped into three different clusters. Cluster 1 contained individuals only from the southern stations 4 and 14, whereas, Cluster 3 contained individuals only from the northern stations 1 and 2. In contrast, Cluster 2 contained individuals from the four stations

Fuchs, et al. (2013), with the use of different molecular markers (LSU, ITS rDNA, and rbcL), found a single biological species of *Fragilariopsis kerguelensis*. Contrary, the molecular results of this study suggest that under this species, there are at least three different species grouped in Cluster 1, 2, and 3. Furthermore, according to the morphological results performed in the Laboratory of Hustedt Diatom Study Centre, there are not apparent phenotypical differences between these individuals.

Rynearson & Armbrust (2004), with the use of rDNA, also found related species of the diatom *Ditylum brightwellii*. These species were morphologically indistinguishable, but molecularly identifiable. In diatoms, intraspecific variation is high (Crosta & Koc, 2007), so cryptic species can be expected (Hülskötter, 2013). As *Ditylum brightwellii*, *Fragilariopsis kerguelensis* might be a cryptic species, so the genetic exchange between specimens of different clusters is not possible. Therefore, for further studies, only specimens from Cluster 2 were considered.

The species of *Fragilariopsis kerguelensis* grouped in Cluster 2, contained individuals from the four stations. The SO is an area almost without any geographical barrier, so winds and water can flow relatively unimpeded eastwards around Antarctica (Graham, 2013; Hinkel, 2016). This allows a constant and largescale dispersion of specimens and genetic material (Rynearson & Armbrust, 2004). As a result, gene flow might occur permitting populations to converge (Collins, et al., 2013).

The ACC is one of the largest and strongest world's ocean currents (Sandells & Flocco, 2014; Smith, et al., 2013). Since the 1980s, due to great temperature differences between cold air from Antarctica and warm air from the subtropics, western winds have become stronger (Schneider & Reusch, 2015). The occurrence of these winds can increase the dispersion of specimens (Jacques, 1983), so the ACC is not an insurmountable current (Lochte, 2014).

In the SO, the fronts are boundaries between different water masses (Graham, 2013), that can act as retention or dispersion mechanisms for plankton (Acha, et al., 2015). These mechanisms can be intensified at <100 km (Moore & Abbott, 2000). The individuals of *Fragilariopsis kerguelensis* were sampled from 20 m of depth to the surface as this species lives preferentiality in surface waters (Shukla & Crosta, 2017). In this case, the PF and SAF might act as dispersion mechanisms.

On the other hand, in the case of the species of *Fragilariopsis kerguelensis* grouped in Cluster 1 (only southern stations) and 3 (only northern stations), the PF and SAF might act as retention mechanisms. Genetic differentiation between populations can occur by natural selection or neutral processes such as genetic drift. It can be presumed that the former populations adapted to the conditions offered in those areas and eventually diverged into new species (Safran & Nosil, 2012).

6.2. ddRAD library processing

At least 60 % of loci from individuals of *Fragilariopsis kerguelensis* did not present any SNP (0). However, approximately 10 % of loci presented at least one single nucleotide variation (1). Thus, for further studies, these loci which contained one SNP were considered. Comparable, in the study of Mock, et al. (2017), 24 % of loci from *Fragilariopsis cylindrus* presented highly divergent alleles.

6.3. Distance influence on *Fragilariopsis kerguelensis* genotype

The fixation index (F_{st}) is a statistical value used to measure the level of genetic differentiation among populations (de Villemereuil & Gaggiotti, 2015). A value close to 0 indicates that the populations are similar, whereas, a value close to 1 indicates that they are different. In this context, this value allows the identification of the geographical distance that produces a greater genetic difference between *Fragilariopsis kerguelensis* populations (Rynearson & Armbrust, 2004).

The lowest F_{st} value for distance was found between the southern stations 4 and 14, although the distance was 1446 km. These stations were located close to or in the Weddell Gyre. As mentioned earlier, a gyre is a system of recirculating waters created due to significant interactions between the ACC and the Antarctic continental shelf (Orsi, et al., 1995; Palla, 2014). It can be assumed that even the distance was 1446 km, this gyre provides enough closeness to reproduce (Rynearson & Armbrust, 2004).

The second lowest F_{st} value was found between the northern stations 1 and 2. Station 2 was close to the PF, which lies in an area of maximum Ekman drift (Taylor, et al., 1978). In the Southern Hemisphere, due to the Coriolis effect and wind direction, western winds move constantly surface waters to the north (Crosta & Koc, 2007), and replace them with bottom waters deflected to the left (Lochte, 2014). This also includes plankton populations (Jacques, 1983), so it can be suggested that *Fragilariopsis kerguelensis* is also moved constantly within the ACC.

The relatively highest F_{st} value was found between Station 2 and 4. A comparable value was found between Station 1 and 4. Approximately 88 % of the genotype was similar within the populations, whereas, 12 % was different between populations. As referred before, Rynearson & Armbrust (2004) studied close populations of *Ditylum brightwellii* and stated that if a F_{st} value is higher than 0.20, populations diverge. These populations were genetically distinct, so they concluded that the geographical distance might not be an important driver in diatom genetic differentiation.

The analysis of distance vs. F_{st} value did not demonstrate a significant influence of distance over the genotype even though the total distance was 2581 km. This can be supported by the already mentioned fact that the apparent barriers in the SO might allow the movement of alleles from one population to another within the water. As a result, the geographical distance might not influence significantly the genetic differentiation between individuals of different stations (Collins, et al., 2013; Rynearson & Armbrust, 2004).

6.4. Environmental conditions in the Southern Ocean during December 2016

In the Southern Ocean, the distribution of the main nutrients can vary significantly among the different subsystems. The season can also influence a nutrient to become a limiting factor (Nelson & Tréguer, 1992). Consequently, the SO is a region of great variations that can modify the structure and function of phytoplankton (Crosta & Koc, 2007; Deppeler & Davidson, 2017).

According to the study of Pinkernell & Beszteri (2014), in the SO, SST and nitrate are the main environmental parameters that affect the distribution of *Fragilariopsis kerguelensis*. The potential genetic differentiation of this species to these factors have not yet been investigated. Therefore, in the present thesis, these environmental parameters together with silicate and phosphate, were correlated with possible effects in the genotype of *Fragilariopsis kerguelensis*.

The BayeScEnv software was used to identify possible genetic differences of *Fragilariopsis kerguelensis* and correlate them with differences in environmental conditions among different populations in the SO. Hence, this software allows the identification of potential adaptation and evolve of *Fragilariopsis kerguelensis* under different environmental pressures (de Villemereuil & Gaggiotti, 2015).

As mentioned before, under *Fragilariopsis kerguelensis* there are at least three different species grouped in Cluster 1, 2, and 3. Due to possible variances in genetic changes between species (Collins, et al., 2013), only individuals grouped in Cluster 2 were analyzed. In addition, two runs were performed for each environmental parameter to confirm the congruence of the results obtained.

6.4.1. Sea surface temperature (SST) correlation to *Fragilariopsis kerguelensis* genotype

The SO is characterized by a seasonal formation and decline of sea ice (Sackett, et al., 2013). The northern part is sea ice free all year round, whereas, the south can be seasonally covered by sea ice (Boyd, 2002; Marshall & Speer, 2012). The Weddell Sea and Ross Sea are two major areas where sea ice usually does not thaw even in the summer. Thus, sea ice formation and reduction affects the heat exchange (Altvater, 2014; Schneider & Reusch, 2015), as well as phytoplankton composition (Moore & Abbott, 2000).

Crosta & Koc (2007) found that sea ice cover and SST have an important influence on two species of the genus *Fragilariopsis*. The sea ice diatom *Fragilariopsis curta* had the highest abundance at a heavy sea ice cover and at -1 to 1 °C. Quite the opposite, *Fragilariopsis kerguelensis*, an open ocean diatom, had the highest abundance at a low ice cover and at 1 to 7 °C. Accordingly, for this thesis, specimens were collected from four stations sea ice free on December 2016.

Fragilariopsis kerguelensis has been found in a wide range of temperatures. For instance, Frenguelli & Orlando (1959) reported this species at temperatures close to 9 °C, but

Cefarelli, et al. (2010) at temperatures from -1.33 to 14.06 °C. Similar, Zielinski & Gersonde (1997) reported *Fragilariopsis kerguelensis* at broader temperatures from -1.00 to 18.00 °C, but there was a significant decrease at temperatures higher than 13.50 °C. In the present thesis, this species was found at a range from 0.60 to 5.30 °C.

Phytoplankton species are adapted to a defined thermal window (Gattuso & Hansson, 2011), so temperature plays an important role in diatom growth (Boyd, et al., 2016). In the study of Zielinski & Gersonde (1997), SST was one of the most important factors that influenced the abundance of *Fragilariopsis kerguelensis*. Temperatures below 0 °C, influenced the abundance of this species, but temperatures above 19 °C inhibited the species success. Therefore, at colder SST, the abundance of *Fragilariopsis kerguelensis* decreases abruptly, whereas, at warmer SST, its abundance decreases gently while it is replaced by warmer water species such as *Azpeitia tabularis* (Crosta & Koc, 2007).

A high F_{st} value (0.69) was found for SST. In this context, this value allows the recognition of loci with higher genetic differentiation between populations (de Villemereuil & Gaggiotti, 2015). This result was related to locus 1496 and suggested a 69 % of genetic variation between populations in this fixed position. Similar, a high F_{st} value (0.62) was related to locus 765. A significative percentage of individuals within populations were different, so SST seemingly influences *Fragilariopsis kerguelensis* on loci 1496 and 765.

On the other hand, approximately 99.26 % of the F_{st} values were <0.20, so 3757 out of 3785 loci did not present important genetic differentiation for SST. This high percentage suggests that SST does not have a significative effect on all loci, but only on specific ones such as the loci mentioned above. As a result, SST might have effects on specific fixed positions in the DNA sequence.

6.4.2. Silicate correlation to *Fragilariopsis kerguelensis* genotype

As mentioned earlier, in lower latitudes of the SO, silicate concentration is <10 μM, whereas, in higher latitudes is between 60 and 90 μM (Brzezinski, et al., 2001; Ito, et al., 2005; Nelson & Tréguer, 1992). In this case, the silicate concentration in the northern stations 1 and 2 was between 6 and 16 μM, while in the southern stations 4 and 14 was between 54 and 67 μM.

In the SO, there is a silicate gradient that normally decreases northwards (Jacques, 1983), due to silicate biological consumption (Brzezinski, et al., 2001). Silicate is used for growth of diatoms (Brümmer, 2003), so they play an important role in the cycle of Si (Shukla, et al., 2013). *Fragilariopsis kerguelensis* is a highly silicified species (Timmermans

& van der Wagt, 2010), so it has a high demand for Si (Fuchs, et al., 2013). As a result, *Fragilariopsis kerguelensis* considerably removes silicate from surface waters (Pinkernell & Beszteri, 2014).

The highest F_{st} value for silicate was 0.56. This result was related to locus 1156 and suggested a 56 % of genetic differentiation between populations in this fixed position. A significative percentage of individuals within populations were different, so silicate apparently influences *Fragilariopsis kerguelensis* specifically on locus 1156. Though, no other defined locus was identified.

Compared with the other environmental parameters, silicate had a higher number of locus with relative high F_{st} values. 159 out of 3785 loci had F_{st} values between 0.39 and 0.20, which represented 4.17 % of the total number of loci. According to Rynearson & Armbrust (2004), F_{st} values >0.20 suggests high levels of genetic differentiation. Then, silicate might have a significative effect on a higher number of loci.

In the SO, silicate can act as a limiting factor (Jacques, 1983). Moore & Abbott (2000) reported a silicate limitation in the north of the SO, and possibly, an iron limitation further south. Diatoms distribution and biomass are strongly influenced by silicate, while iron influences uptake of silicate, degree of silicification and growth rate (Bopp, et al., 2005; Boyd, 2002).

Orthosilicic acid ($Si(OH)_4$) is the soluble silicate form that can be biologically assimilable (Brümmer, 2003). The dissolution of silicate depends on temperature (Gnanadesikan, 1999), so it eventually influences the uptake of silicate by diatoms. Consequently, the temperatures not only influence their distribution, but also of an essential element for their growth.

6.4.3. Nitrate and phosphate correlation to *Fragilariopsis kerguelensis* genotype

In most of the surface waters of the world oceans, nitrate is completely utilized, but the SO is an important exception. Due to closed system processes such as nitrate uptake by phytoplankton and organic N oxidation by bacteria, there are different nitrate concentrations within the SO (Sigman, et al., 2000). In this case, the nitrate concentration in Station 1 and 2 was 18.62 and 22.03 μM respectively, while in Station 4 and 14 was 26.84 and 24.17 μM respectively.

The concentration of nitrate affects the productivity of an ocean (Sigman, et al., 2000). Bopp et al. (2005) informed that between 30° and 40° latitude in the Southern Hemisphere, diatoms retreat due to nitrate decrease. Likewise, Sigman et al. (2000) found that nitrate can

limit the production in the large open ocean. Therefore, *Fragilariopsis kerguelensis* is rarely present at low nitrate concentrations (Pinkernell & Beszteri, 2014).

As nitrate, phosphate utilization in the surface waters of the SO is also an exception (Sigman, et al., 2000). The phosphate concentration in Station 1 and 2 was 1.35 and 1.71 μM respectively, whereas, in Station 4 and 14 was 1.88 and 1.72 μM respectively. This gradient observed in the SO is due to a minor biological consumption (de Baar, et al., 1997) and mixing with low nutrients northern waters (Lochte, 2014).

For nitrate and phosphate, the increasing concentration with higher latitudes was stronger in Station 2 and 4. Station 2 was located close to the PF, and Station 4 close to the southern boundary of the ACC. As mentioned before, the PF is one of the strongest oceanic frontal systems (Marshall & Speer, 2012) capable to oversee most of the transportation in the ACC (Shao, et al., 2015). As a result, especially the PF is characterized by a high nutrient concentration in surface waters (Zielinski & Gersonde, 1997).

A high F_{st} value (0.61) was found for nitrate. This result was related to locus 1496 and suggested a 61 % of genetic differentiation between populations in this fixed position. Similar, a high F_{st} value (0.54) was related to locus 765. A significative percentage of individuals within populations were different, so nitrate apparently influences *Fragilariopsis kerguelensis* on loci 1496 and 765.

In contrast, approximately 99.63 % of the F_{st} values were <0.20, so 3771 out of 3785 loci did not present important genetic differentiation for nitrate. This high percentage suggests that this nutrient does not have a significative effect on the majority of loci, but only on specific ones such as the ones mentioned above. As a result, nitrate might influence specific fixed positions in the DNA sequence.

In the case of phosphate, the highest F_{st} values were 0.58 and 051, which were also related to loci 1496 and 765, respectively. More than half of individuals within these populations were not similar in these fixed positions in the DNA sequence. According to these results, phosphate apparently influences *Fragilariopsis kerguelensis* also on loci 1496 and 765.

Compared with the other environmental parameters analyzed, phosphate had a significant number of loci with low F_{st} values (<0.20). These loci had F_{st} values between 0.18 and 0.07, which represented 99.68 % of the total number of loci. This suggests that 3773 out of 3785 loci present low genetic differentiation within populations for phosphate, so there are apparently low levels of genetic differentiation.

As mentioned earlier, the surface waters of the SO are not all depleted as other oceans. There are high concentrations of nitrate, phosphate, and silicate, but there are barely phytoplankton blooms, even in the summertime (de Baar, et al., 1997). Nitrate and phosphate can be limiting factors (Crosta & Koc, 2007), but in the case of the SO, Hart (1934) proposed that these macronutrients might not be limiting factors as they are in high and persistent concentrations.

As a possible explanation, Gran (1931) proposed the Antarctic iron limitation hypothesis. This states that in HNLC regions such as the SO, iron is the main phytoplankton liming factor (Armbrust, et al., 2004; Boyle, 1998; Moore & Abbott, 2000). Iron plays an important role in the nitrogen metabolism of phytoplanktonic cells, so its limitation decreases the N uptake rates and nitrate reductase activity (Hoffmann, et al., 2007).

Iron derives from rivers and continental-shelf sediments (Boyle, 1998; de Baar, et al., 1997). As there are almost no continental sources in the SO, iron comes mainly from dust deposition (Strzepek, et al., 2011). As a result, surface waters of the SO have low iron concentrations (Brzezinski, et al., 2001; Hoffmann, et al., 2007), which probably leads to a limitation on phytoplankton (Tréguer & Jacques, 1992).

In the SO, a variation of the known "Redfield ratio" for nitrate to phosphate from 16:1 to ~14:1 is expected possibly due to iron. The assimilation of these nutrients by phytoplankton leads to an increase in biomass, while their concentration in the environment decreases (de Baar, et al., 1997). In the study of de Baar, et al. (1997), *Fragilariopsis kerguelensis* seemed to vary its nitrate to phosphate uptake in response to iron concentration. Timmermans & van der Wagt (2010) also found that individuals of *Fragilariopsis kerguelensis* growing at low iron concentrations assimilated more phosphate compared with the same species growing at high iron concentration.

On the other hand, de Baar, et al. (1997) proposed that even under optimal iron conditions, *Fragilariopsis kerguelensis* assimilates and uses low concentrations of nitrate and phosphate. This species is apparently adapted to low iron concentration, so is capable to decrease cell size while it reduces photosystems that require iron like cytochrome b6f complex (Strzepek, et al., 2011) or optimizes nutrient-uptake transporters and reservoirs (Huysman, et al., 2010).

6.4.4. ACC

The ACC is one of the most dominant characteristic of the SO (Shao, et al., 2015) where the occurrence of stronger western winds can influence it (Schneider & Reusch,

2015). In turn, this can also influence the fluxes of essential nutrients (Falkowski, et al., 1998) and alter ecosystems functions and nutrient balance (Timmermans & van der Wagt, 2010).

Populations that undergo different environmental conditions, can be exposed to adaptive and selective pressures that could provide an advantage in its environment (de Villemereuil & Gaggiotti, 2015). Therefore, some phytoplankton species, and probably *Fragilariopsis kerguelensis*, might evolve and develop certain mechanisms to ensure their stability in changing conditions (Falkowski, et al., 1998).

6.5. *Fragilariopsis kerguelensis* size

A higher size of *Fragilariopsis kerguelensis* was found in Cluster 3 which corresponded to the northern stations 1 and 2, and a lower size was found in Cluster 1 which corresponded to the southern stations 4 and 14. In phytoplankton, cell size often reflects a response to environmental conditions (Lepistö & Saura, 1998). Nonetheless, in the case of diatoms, size is not an optimal morphological characteristic because the asexual reproduction results in a progressive decrease of the frustule size (Brümmer, 2003).

6.6. Loci

According to de Villemereuil & Gaggiotti (2015), an outlier is a high F_{st} value within the genome that is a suspicion of a local adaptation that can lead to a genetic differentiation between populations. Selection acts on individual locus, so under a selective pressure, the genetic differentiation in a specific locus is expected to be higher compared with the rest of loci in the genome (Safran & Nosil, 2012).

Regarding the results obtained on the analysis of F_{st}, seemingly differences in environmental conditions among different stations leads to a genetic differentiation in a specific loci. The loci identified were 765 and 1496 for SST, nitrate, and phosphate, and 1156 for silicate.

6.6.1. Loci 765 and 1156

The second highest F_{st} value obtained for SST, nitrate, and phosphate, indicated a relation with locus 756. However, the Blast analysis gave a high E-value (1.3-4.6) and a low alignment score (40-50). In the case of E-value, the smaller it is or closer to 0, the match of the alignment is considered more “significant”. Contrary, the smaller the score is, the worse the alignment is (National Center for Biotechnology Information, 2018). As a result, this

locus was not further considered as well as the locus 1156 for silicate with a high E-value (1.0-3.6) and a low alignment score (40-50).

The posterior error probability (PEP) is an error estimation of the approach used (de Villemereuil & Gaggiotti, 2015). As for the E-value, the smaller the PEP is or closer to 0, the result is considered more "significant". In this context, this value allows identifying true and false positives. For the locus 756, a PEP >0.12 was found, and for the locus 1156, a PEP ~0.23. This confirmed the results of the E-value and alignment score.

The combination of double restriction digestion and size selection produces smaller DNA fragments that are within the size-selection window (Peterson, et al., 2012; Pray, 2008). As the locus 756 appeared in many of the correlation analysis, it might be part of a bigger sequence. Further studies are recommended to analyze in detail a larger sequence of this locus.

6.6.2. Locus 1496

The highest F_{st} value obtained for SST, nitrate, and phosphate, indicated a correlation to locus 1496. The Blast analysis gave 178 blast hits on the subject sequences, with low E-values and high alignment scores. The most significant result had an E-value of $2e^{-18}$ and an alignment score of 80-200 which corresponded to "*Phaeodactylum tricornutum* CCAP 1055/1 predicted protein partial mRNA".

As *Fragilariopsis kerguelensis*, *Phaeodactylum tricornutum* is also a pennate diatom belonging to class Bacillariophyceae. According to fossil records, centric diatoms appeared 180 million years ago, whereas, pennate diatoms diverged 90 million years ago. This diverging had led to a high differentiation between these two diatoms groups, so the representatives of them only share ~60 % of genes (Bowler, et al., 2008).

Bowler, et al. (2008) suggested that the greatest force of diatoms evolution was probably horizontal gene transfer from bacteria. For instance, *Phaeodactylum tricornutum* has acquired 5 % of its genome from bacteria. These genes have become specific for diatoms, so they are evolving quicker than other genes in their genomes. Therefore, they might have the potential to encode innovative metabolisms to the extreme conditions in the SO (De Riso, et al., 2009; Mock, et al., 2017).

"*Phaeodactylum tricornutum* CCAP 1055/1 predicted protein partial mRNA" was reported by Bowler, et al. in 2008 as a possible mRNA. However, it was modified as a possible rRNA (Quast, et al., 2013). Ribosomal RNA (rRNA) is the major component of the ribosome

structure. A ribosome oversees mRNA translation into proteins, so it is a link between genotype and phenotype (Bhagavan & Ha, 2011).

In this locus, the highest outlier obtained was for the environmental parameter of SST, and the lowest PEP obtained also supported this result. Genotypic changes to fluctuating SST have also been identified in other diatoms populations such as *Thalassiosira pseudonana* (Brand, et al., 1981). Phytoplankton species exposed to constant and high changing environmental conditions like sea ice can eventually be capable to adapt and evolve to further future changes (Sackett, et al., 2013).

In diatoms, the approximate rate of division is 0.1-.1.8 times per day (Crosta & Koc, 2007), which allows them to have higher biomass, as well as greater populations size, so a higher probability to evolve under changing environmental conditions (Collins, et al., 2013). During asexual reproduction, all genes within a specimen are tightly linked. Under certain environmental conditions, a specific gene can be selected which leads to changes in allele frequencies, and eventually to population differentiation (Rynearson & Armbrust, 2004).

Finally, even though no specific gene could be identified, the results obtained in this study with the software BayeScEnv allows to hypothesize that locus 1496 and probably also 765, “might be part of a specific gene that is under selection” (de Villemereuil & Gaggiotti, 2015). Moreover, a possible correlation between environmental conditions and genotype of *Fragilariopsis kerguelensis* can be suggested.

7. Conclusions

The molecular results obtained in this Master Thesis suggest that *Fragilariopsis kerguelensis* is seemingly a cryptic species, so under this species, there are at least three different species. The barriers in the SO apparently allow the dispersion of specimens of Cluster 2, so the geographical distance between stations do not produce a significative genetic difference between populations.

The environmental data in the SO obtained from the National Oceanic and Atmospheric Administration and measurement aboard the Polarstern, suggest that during December 2016, each station had different SST and nutrient concentration. Generally, there was an increasing trend of macronutrients such as silicate, nitrate, and phosphate from northern to southern stations, but a decreasing trend of SST.

Fragilariopsis kerguelensis grouped in Cluster 2 are probably influenced by environmental parameters such as SST, silicate, nitrate, and phosphate. According to literature, it can also be influenced by iron since it is required for the assimilation of these macronutrients. Moreover, it can be deduced that all environmental parameters correlate with each other quite strongly, so *Fragilariopsis kerguelensis* genotype might be influenced simultaneously by all these parameters.

The loci 1496 was found to be the fixed position with higher genetic differentiation between populations of *Fragilariopsis kerguelensis* no matter the environmental parameter. This was related to "*Phaeodactylum tricornutum* CCAP 1055/1 predicted protein partial mRNA", also from a pennate diatom. Further studies are recommended to analyze in detail this sequence in diatoms.

Finally, *Fragilariopsis kerguelensis* might have an adaptive potential to the changing conditions in the SO. Approximately 10 % of the genome analyzed have SNPs, so it can be suggested that environmental conditions in the SO not only influence the abundance of *Fragilariopsis kerguelensis* but probably also its genotype. As this species is the most abundant in the SO and base of the food web, it might have effects on ecosystems, nutrients cycles and also global climate.

8. Bibliography

Acha, E., Piola, A., Iribarne, O., & Mianzan, H. (2015). *Ecological processes at marine fronts* (Vol. 1). Cham: Springer International Publishing.

Agilent. (2017). *Agilent 2100 Expert Software B.02.09 Powerful software for the analysis of RNA, DNA and proteins.* Santa Clara: Agilent Technologies, Inc.

Alfred-Wegener-Institut. (2017). *Master Track RV Polarstern PS103.* Bremerhaven: AWI.

Altvater, F. (2014). *Analysis of diatom species composition and size distribution in the Southern Ocean.* Bremerhaven: Universität Bremen and AWI.

Armbrust, V., Berges, J., Bowler, C., Green, B., Martinez, D., & Putnam, N. (2004). The genome of the diatom Thalassiosira Pseudonana: Ecology, evolution, and metabolism. *Science, 306*(5693), 79-86.

Arrigo, K., Robinson, D., Worthen, D., Dunbar, R., DiTullio, G., VanWoert, M., & Lizotte, M. (1999). Phytoplankton community structure and the drawdown of nutrients and CO2 in the Southern Ocean. *Science, 283*(5400), 365-367.

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.

Bard, E., & Rickaby, R. (2009). Migration of the subtropical front as a modulator of glacial climate. *Nature, 460*(7253), 380-384.

Barsanti, L., & Gualtieri, P. (2014). *Algae: Anatomy, biochemistry, and biotechnology* (Vol. 1). Boca Ratón: CRC Press.

Beszteri, B. (2018, March and May 2 and 17). Population genomic comparison of Fragilariopsis kerguelensis.

Bhagavan, N., & Ha, C. (2011). *Essentials of medical biochemistry* (Vol. 1). San Diego: Academic Press.

Bopp, L., Aumont, O., Cadule, P., Alvain, S., & Gehlen, M. (2005). Response of diatoms distribution to global warming and potential implications: A global model study. *Geophysical Research Letters, 32*(19), 1-4.

Bowler, C., Allen, A., Badger, J., Grimwood, J., Jabbari, K., & Kuo, A. (2008). The Phaeodactylum genome reveals the evolutionary history of diatom genomes. *Nature, 456*(7219), 239-244.

Boyd, P. (2002). Environmental factors controlling phytoplankton processes in the Southern Ocean. *Journal of Phycology, 38*(5), 844-861.

Boyd, P., Dillingham, P., McGraw, C., Armstrong, E., Cornwall, C., Feng, Y., & Nunn, B. (2016). Physiological responses of a Southern Ocean diatom to complex future ocean conditions. *Nature, 6*(2), 207-2016.

Boyle, E. (1998). Pumping iron makes thinner diatoms. *Nature, 393*(6687), 788-789.

Brand, L., Murphy, L., Guillard, R., & Lee, H. (1981). Genetic variability and differentiation in the temperature niche component of the diatom Thalassiosira pseudonana. *Marine biology, 2-3*, 103-110.

Brümmer, F. (2003). Living inside a glass box - silica in diatoms. In W. Müller, *Silicon biomineralization : Biology - Biochemistry - Molecular Biology - Biotechnology* (Vol. 1, pp. 1-7). Cham: Springer International Publishing.

Brzezinski, M., Nelson, D., Franck, V., & Sigmon, D. (2001). Silicon dynamics within an intense open-ocean diatom bloom in the Pacific sector of the Southern Ocean. *Deep sea research part II: Topical studies in oceanography, 48*(19-20), 3997-4018.

Catchen, J., Hohenlohe, P., Bassham, S., Amores, A., & Cresko, W. (2013). Stacks: an analysis tool set for population genomics. *Molecular ecology, 22*(11), 3124-3140.

Cefarelli, A., Ferrario, M., Almandoz, G., Atencio, A., Akselman, R., & Vernet, M. (2010). Diversity of the diatom genus Fragilariopsis in the Argentine Sea and Antarctic waters: morphology, distribution and abundance. *Polar biology, 33*(11), 1463-1484.

Celan, P. (2014). *Breathturn into timestead* (Vol. 1). New York: Farrar, Straus and Giroux.

Collins, S., Rost, B., & Rynearson, T. (2013). Evolutionary potential of marine phytoplankton under ocean acidification. *Evolutionary applications, 7*(1), 140-155.

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.

Crosta, X., & Koc, N. (2007). Diatoms: From micropaleontology to isotope geochemistry. In C. Hillaire–Marcel, & A. De Vernal, *Developments in marine geology* (Vol. 1, pp. 328-369). Amsterdam: Elsevier.

de Baar, H., Van Leeuwe, M., Scharek, R., Goeyens, L., Bakker, K., & Fritsches, P. (1997). Nutrient anomalies in Fragilariopsis kerguelensis blooms, iron deficiency and the nitrate/phosphate ratio (A. C. Redfield) of the Antarctic Ocean. *Deep sea research part II: Topical studies in oceanography, 44*(1-2), 229-260.

De Riso, V., Raniello, R., Maumus, F., Rogato, A., Bowler, C., & Falciatore, A. (2009). Gene silencing in the marine diatom Phaeodactylum tricornutum. *Nucleic acids research, 37*(14), 1-12.

de Villemereuil, P., & Gaggiotti, O. (2015). AnewFST-based method to uncover local adaptation using environmental variables. *Methods in ecology and evolution, 6*(11), 1248-1258.

de Vries, M. (2018). *Genotype vs. phenotype*. Retrieved from Understanding evolution: https://evolution.berkeley.edu/evolibrary/article/0_0_0/genovspheno_01

Deppeler, S., & Davidson, A. (2017). Southern Ocean phytoplankton in a changing climate. *Frontiers in marine science, 203*(1), 1-28.

Ehrenberg, C. (2017). *Bacillariaceae*. Retrieved from Algae base: http://www.algaebase.org

Falkowski, P., Barber, R., & Smetacek, V. (1998). Biogeochemical controls and feedbacks on ocean primary production. *Science, 281*(5374), 200-206.

Finkel, Z. (2016). Silicification in the microalgae. In M. Borowitzka, J. Beardall, & J. Raven, *The physiology of microalgae* (Vol. 1, pp. 289-300). Cham: Springer International Publishing.

Fuchs, N., Scalco, E., Kooistra, W., Assmy, P., & Montresor, M. (2013). Genetic characterization and life cycle of the diatom Fragilariopsis kerguelensis. *Journal of Phycology, 48*(4), 411-426.

Gal, A. (2018). *Diatom silicification*. Retrieved from Weizmann Institute of Science: http://www.weizmann.ac.il/plants/gal/research-activities/diatom-silicification

Gattuso, J.-P., & Hansson, L. (2011). *Ocean Acidification* (Vol. 1). Oxford: OUP Oxford.

Gille, S. (2002). Warming of the Southern Ocean since the 1950s. *Science, 295*(5558), 1275-1277.

Gnanadesikan, A. (1999). A global model of silicon cycling: Sensitivity to eddy parameterization and dissolution. *Global biochemical cycles, 13*(1), 199-220.

Graham, R. (2013). *The location and variability of Southern Ocean fronts.* Stockholm: Stockholms universitet.

Guillard, R., & Ryther, J. (1962). *f/2 medium.* Retrieved from National Center for Marine Algae and Microbiota (NCMA): https://ncma.bigelow.org/media/wysiwyg/Algal_recipes/NCMA_algal_medium_f_2_1.pdf

Hart, J. (1934). *On the phytoplankton of the South-west Atlantic and the Bellingshausen Sea* (Vol. 1). Cambridge: University Press.

Hinkel, L. (2016). *Southern Ocean cooling in a warming world*. Retrieved from Oceans at MIT: http://oceans.mit.edu/news/featured-stories/southern-ocean-cooling-in-a-warming-world

Hoffmann, L., Peeken, I., & Lochte, K. (2007). Effects of iron on the elemental stoichiometry during EIFEX and in the diatoms Fragilariopsis kerguelensis and Chaetoceros dichaeta. *Biogeosciences, 4*(4), 569-579.

Hülskötter, J. (2013). *Molecular ecological study of the toxic marine dinoflagellate Alexandrium tamarense.* Bremerhaven: Hochschule Bremen and Alfred-Wegener-Institut (AWI).

Huysman, M., Martens, C., Vandepoele, K., Gillard, J., Rayko, E., & Heijde, M. (2010). Genome-wide analysis of the diatom cell cycle unveils a novel type of cyclins involved in environmental signaling. *Genome biology, 11*(2), 1-19.

Ito, T., Parekh, P., Dutkiewicz, S., & Follows, M. (2005). The Antarctic Circumpolar productivity belt. *Geophysical research letters, 32*(13), 1-4.

Jacques, G. (1983). Some ecophysiological aspects of the Antarctic phytoplankton. *Polar biology, 2*(1), 27-33.

Jones, R. (2000). Mixotrophy in planktonic protists: an overview. *Freshwater biology, 45*(2), 219-226.

Lepistö, L., & Saura, M. (1998). Effects of forest fertilization on phytoplankton in a boreal brown-water lake. *Boreal environment research, 3*(1), 33-43.

Lochte, K. (2014). *The consequences of climate change for life in the Southern Ocean.* Bremerhaven: Alfred-Wegener-Institut.

Marshall, J., & Speer, K. (2012). Closure of the meridional overturning circulation through Southern Ocean upwelling. *Nature geoscience, 5*(3), 171-180.

Mock, T., Otillar, R., Strauss, J., Mc Mullan, M., Paajanen, P., & Schmutz, J. (2017). Evolutionary genomics of the cold-adapted diatom Fragilariopsis cylindrus. *Nature, 541*(7638), 536-551.

Moore, J., & Abbott, M. (2000). Phytoplankton chlorophyll distributions and primary production in the Southern Ocean. *Geophysical research, 105*(C12), 709-728.

National Center for Biotechnology Information. (2018). *Blast*. Retrieved from National Center for Biotechnology Information (NCBI): https://blast.ncbi.nlm.nih.gov/Blast.cgi

National Oceanic and Atmospheric Administration. (2018). *Ocean climate laboratory*. Retrieved from National Oceanic and Atmospheric Administration: https://www.nodc.noaa.gov/about/oceanclimate.html

Nelson, D., & Tréguer, P. (1992). Role of silicon as a limiting nutrient to Antarctic diatoms: evidence from kinetic studies in the Ross Sea ice-edge zone. *Marine ecology progress series, 80*(3), 255-264.

Orsi, A., Whitworth, T., & Nowlin, W. (1995). On the meridional extent and fronts of the Antarctic Circumpolar Current. *Deep sea research part I: Oceanographic research papers, 42*(5), 641-673.

Palla, R. (2014). *Krill America* (Vol. 1). Pennsauken Township: Book Baby.

Peterson, B., Weber, J., Kay, E., Fisher, H., & Hoekstra, H. (2012). Double Digest RADseq: An inexpensive method for De Novo SNP discovery and genotyping in model and non-model species. *PLoS ONE, 7*(5), 1-11.

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.

Pitchford, J., & Brindley, J. (1999). Iron limitation, grazing pressure and oceanic high nutrient-low chlorophyll (HNLC) regions. *Journal of Plankton Research, 21*(3), 525-547.

Pray, L. (2008). Restriction enzymes. *Nature education, 1*(1), 1.

Pritchard, J., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics, 155*(1), 945-959.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., & Glöckner, F. (2013). *ABQD01000052*. Retrieved from The SILVA ribosomal RNA gene database project: improved data processing and web-based tools: https://www.arb-silva.de/browser/lsu-132/ABQD01000052

Raven, J., & Waite, A. (2004). The evolution of silicification in diatoms: inescapable sinking and sinking as escape? *New phytologist, 162*(1), 45-61.

Riginos, C., Crandall, E., Liggins, L., Bongaerts, P., & Treml, E. (2016). Navigating the currents of seascape genomics: how spatial analyses can augment population genomic studies. *Current zoology, 62*(6), 581-601.

Roberts, R., Belfort, M., Bestor, T., Bhagwat, A., Bickle, T., & Bitinaite, J. (2003). A nomenclature for restriction enzymes, DNA methyltransferases, homing endonucleases and their genes. *Nucleic acids research, 31*(7), 1805-1812.

Rochette, N., & Catchen, J. (2017). Deriving genotypes from RAD-seq short-read data using Stacks. *Nature Protocols, 12*(12), 2640–2659.

Rynearson, T., & Armbrust, E. (2004). Genetic differentiation among populations of the planktonic marine diatom Ditylum brightwellii (Bacillariophyceae). *Journal of Phycology, 40*(1), 34-43.

Sachs, O., Sauter, E., Schlüter, M., Rutgers van der Loeff, M., Jerosch, K., & Holby, O. (2009). Benthic organic carbon flux and oxygen penetration reflect different plankton provinces in the Southern Ocean. *Deep sea research part I: Oceanographic research papers, 56*(8), 1319-1335.

Sackett, O., Petrou, K., Reedy, B., De Grazia, A., Hill, R., & Doblin, M. (2013). Phenotypic plasticity of Southern Ocean diatoms: Key to success in the sea ice habitat? *Plos one, 8*(11), 1-12.

Safran, R., & Nosil, P. (2012). Speciation: The origin of new species. *Nature, 10*(17), 1-2. Retrieved from Nature.

Sandells, M., & Flocco, D. (2014). *Introduction to the physics of the cryosphere* (Vol. 1). Williston: Morgan & Claypool Publishers.

Schneider, D., & Reusch, D. (2015). Antarctic and Southern Ocean surface temperatures in CMIP5 models in the context of the surface energy budget*. *Journal of Climate, 5*, 1689-1716.

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). *Marine biological laboratory, 227*(2), 146-160.

Shao, A., Gille, S., Mecking, S., & Thompson, L. (2015). Properties of the Subantarctic Front and Polar Front from the skewness of sea level anomaly. *Geophysical Research: Oceans, 120*(7), 5179-5193.

Shukla, S., & Crosta, X. (2017). Fragilariopsis kerguelensis size variability from the Indian subtropical Southern Ocean over the last 42 000 years. *Antarctic science, 29*(2), 137-146.

Shukla, S., Crosta, X., Cortese, G., & Nayak, G. (2013). Climate mediated size variability of diatom Fragilariopsis kerguelensis in the Southern Ocean. *Quaternary science reviews, 69*(1), 49-58.

Sigman, D., Altabet, M., McCorkle, D., Francois, R., & Fischer, G. (2000). The 15N of nitrate in the Southern Ocean: Nitrogen cycling and circulation in the ocean interior. *Geophysical research, 105*(C8), 19599-19614.

Smith, R., Desflots, M., White, S., Mariano, A., & Ryan, E. (2013). *The Antarctic CP Current*. Retrieved from Ocean Surface Currents: http://oceancurrents.rsmas.miami.edu/southern/antarctic-cp.html

Sokolov, S., & Rintoul, S. (2002). Structure of Southern Ocean fronts at 140 E. *Marine systems, 37*(1), 151-184.

Spaulding, S. (2018). *Diatoms*. Retrieved from US Geological Survey / Institute of Arctic and Alpine Research: http://westerndiatoms.colorado.edu/about/what_are_diatoms

Strzepek, R., Maldonado, M., Hunter, K., Frew, R., & Boyd, P. (2011). Adaptive strategies by Southern Ocean phytoplankton to lessen iron limitation: Uptake of organically complexed iron and reduced cellular iron requirements. *Limnology and Oceanography, 56*(6), 1983-2002.

Talley, L. (2011). *Descriptive physical oceanography: An introduction* (Vol. 1). San Diego: Academic Press.

Taylor, H., Gordon, A., & Molinelli, E. (1978). Climatic characteristics of the Antarctic Polar Front Zone. *Geophysical research, 83*(C9), 4572-4578.

Timmermans, K., & van der Wagt, B. (2010). Variability in cell size, nutrient depletion, and growth rates of the Southern Ocean diatom Fragilariopsis kerguelensis (Bacillariophyceae) after prolonged iron limitation. *Journal of Phycology, 29*(2), 497-506.

Tréguer, P., & Jacques, G. (1992). Dynamics of nutrients and phytoplankton, and fluxes of carbon, nitrogen and silicon in the Antarctic Ocean. *Polar biology, 12*(2), 149-162.

Walker, J., Younos, T., & Zipper, C. (2007). *Nutrients in lakes and reservoirs – A literature review for use in nutrient criteria development.* Blacksburg: Virginia Water Resources Research Center and Virginia Tech.

Wehr, J., Sheath, R., & Kociolek, P. (2015). *Freshwater algae of North America: Ecology and classification (aquatic ecology)* (Vol. 2). London: Elsevier.

Zielinski, U., & Gersonde, R. (1997). Diatom distribution in Southern Ocean surface sediments (Atlantic sector): Implications for paleoenvironmental reconstructions. *Palaeogeography, Palaeoclimatology, Palaeoecology, 129*(3-4), 213-250.

9. Annex

Table 10. Reagents for the culture medium f/2. Source Guillard & Ryther (1962).

Reagents	Volume (ml)
$NaNO_3$	1.00
NaH_2PO_4 H_2O	1.00
Na_2SiO_3 $9H_2O$	1.00
Trace metal solution	1.00
Vitamin solution	0.50

Table 11. Sequence of the primers P5 and P7. Source Hülskötter (2013).

Primer	Sequence (5' →3')
P5	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGAC-GCTCTTCCGATCT
P7	CAAGCAGAAGACGGCATACGAGAT

Table 12. Double digestion reaction for a final volume of 45 µL. Source Hülskötter (2013) and Peterson, et al. (2012).

Reagents	Volume (µL)
10X FastDigest Buffer	4.50
PstI	2.25
MspI	2.25
DNA	Variable
Nuclease-free water	Variable
Total volume	45.00

Table 13. Adapter ligation reaction for a final volume of 40 µL. Source Peterson, et al. (2012).

Reagents	Initial concentration		Final Concentration		Volume (µL)
T4 DNA ligase buffer	10	X	1	X	4.00
P5 adapter	10	µM	0.20	Mm	0.80
P7 adapter	10	µM	1	Mm	4.00

T4 DNA ligase	2000	U/µL	25	U/Ml	0.50
Digested sample	-				22.00
Nuclease-free water	-				8.70
Total volume	-				40.00

Table 14. PCR reaction for a final volume of 50 µL. Source Hülskötter (2013) and Peterson, et al. (2012).

Reagents	Initial Concentration		Final Concentration		Volume (µL)
Q5 buffer	5	X	1	X	10.00
dNTPs	10	m M	0.20	m M	1.00
Primer P5	100	µM	1	Mm	0.50
Primer P7	100	µM	1	Mm	0.50
Q5 HF Taq Polymerase	2	U/µL	0.02	U/Ml	0.50
Digested and ligated sample	-				2.50
Nuclease-free water	-				35.00
Total volume	-				50.00

Table 15. PCR amplification conditions. Source Hülskötter (2013).

Cycle	Temperature	Time	Number of cycles
Initial denaturation	98 °C	30 seconds	1
Denaturation	98 °C	10 seconds	14
Annealing	65 °C	30 seconds	
Elongation	72 °C	30 seconds	
Final extension	72 °C	5 minutes	1

Table 16. Sequence of the loci 1496 (18876_67).

Locus	Sequence (5' →3')
1496 (18876_67)	TGCAGCAATTTATGGTAAGCCTAAACAAATATTTTTGTTTACTT TAACTTGAGACTATATAAATATACATATACATATACATATACA TATACATATATACTAACTTTTTCCTTCTTCAATATCATAGCTTG

	GGCACATCTCATACTTTATGGAGAAGAAGCAACACGTCAAGCT CAAAAGGAGGGCGAAGAAGCCGCAGCAGCATACGAGCAAGA GGAAGCGT GATATAGCAGCCGTGAGCC

Table 17. Sequence of the loci 765 (9600_47).

Locus	Sequence (5' →3')
765 (9600_47)	TGCAGTACTAAAAGCAAAGTGCAAAAGAAAATTGAGGGAACT CAAATCGATTCCCTCGCAGAAGCTTCTACTCCTGTCAAACCTA AATTGCGTCCAATTTTAAATTACACAAACTTCTCGTGCATGCGT GGCATACACGGTATCTTATTTCACTTATATTTTATGAGCTCGCG TGGCGCACGAATCACTTATATTTTATGAGCTCGCCGCGGTGCG CGAATCAAGTACCGTC

Table 18. Sequence of the loci 1156 (13469_173).

Locus	Sequence (5' →3')
1156 (13469_173)	TGCAGTGTTAGCATCTACACTAGTATTGGTCCAAGCACTTCCA GCAAATCGACCAGTACCCATAGTAACACGTTTTGCAGAACCAC CTTTTGACTCAGATTGTGAATCAAGATTTTCTTTGATAGCGGCA GCCTTTTCCTTCTTGGTAGCAGCATCTTTACCTCCTTTTGAAAC AATAGTAGTACGGCCG