THE USE OF METABARCODING APPROACH TO DETERMINE THE DIATOM COMPOSITION OF MUCILAGE IN THE SEA OF MARMARA

by

Anıl Doğan Örün

B.Sc. in Environmental Engineering, Middle East Technical University, 2020

Submitted to the Institute of Environmental Sciences

in partial fulfillment of the requirements for the degree of

Master of Science

in

Environmental Sciences

Boğaziçi University

2023

ACKNOWLEDGEMENTS

First and foremost, I express my heartfelt gratitude to my advisor, Prof. Raşit Bilgin, for his unwavering support. I am forever indebted to him for his encouragement and invaluable insights throughout my academic journey. His mentorship has been exceptional, and I am truly grateful for his endless support throughout the entire thesis process.

I extend my sincerest appreciation to the esteemed committee members, Prof. Dr. Melek İşinibilir Okyar and Prof. Dr. Melek Türker Saçan, for graciously dedicating their time and expertise to provide valuable feedback. Your contribution has been instrumental in shaping the outcome of this project, and I am grateful for your invaluable guidance.

I would like to express my deepest appreciation to Prof. Dr. Melek İşinibilir Okyar and Onur Doğan for their invaluable contributions to this project. Their knowledge and expertise have been indispensable, and I am grateful for their guidance and support. Additionally, I would like to extend my sincere gratitude to Merve Sıvacı and Maha Salameh for their help in the laboratory processes.

I would like to express my heartfelt gratitude to my dear family members and loved ones for their unwavering support throughout my journey. Most notably, my mom Ayten, my dad Önder, my brother Onur, my uncle Tuncay, and my girlfriend Burcu have been my pillars of strength, providing me with emotional support and endless encouragement throughout the years. I am genuinely grateful for their unwavering love and support, which have been instrumental in helping me achieve my goals.

I would like to acknowledge TUBITAK ULAKBIM, High-Performance and Grid Computing Center (TRUBA resources), for their partial support in conducting the numerical calculations reported in this paper. Their cutting-edge infrastructure and resources have provided me with a unique opportunity to analyze a massive amount of data. I am sincerely grateful to TRUBA for their generous support, which has been crucial in the success of this research. This study was supported by a TÜBİTAK Grant, No: 121G096.

ABSTRACT

THE USE OF METABARCODING APPROACH TO DETERMINE THE DIATOM COMPOSITION OF MUCILAGE IN THE SEA OF MARMARA

Marmara Basin has been adversely affected by industrial activities in the recent years. As a result of the rising industrial activities, the occurrence of harmful algal blooms, red-tides and mucilage events have increased in the Sea of Marmara. Mucilage problem has reached peak levels recently due to industrial wastes and the water discharged from wastewater treatment plants without biological treatment. Negative effects of industrial waste and untreated wastewater on the marine ecosystem resulted in a decrease in water quality and became a serious threat to biodiversity. To act against the problems arising in the marine ecosystem, bio-monitoring studies have become tremendously important to ensure the proper management of integrated water resources and to maintain sustainable marine ecosystem services. As a part of bio-monitoring studies, DNA metabarcoding applications based on next-generation DNA sequencing technologies offer reliable and effective approaches in the detection of microorganisms and plankton groups. Such technologies play an important role in the monitoring of the harmful effects of mucilage events and can help as early warning systems for the health of the marine ecosystems. In this thesis, I will study the effects of excessive plankton and mucilage increase, and the impacts of mucilage on water quality by examining diatom genera. According to our results, *Skeletonema* and *Nitzschia* diatom genera are the dominant organism groups for 10 species hits; however, *Guinardia* and *Pseudo-nitzschia* are the dominant ones for 100 species hits. These genera might have a potential effect on the mucilage formation.

ÖZET

MARMARA DENİZİNDEKİ MUSİLAJIN DİATOM BİLEŞİMİNİN BELİRLENMESİ İÇİN METABARCODLAMA YAKLAŞIMININ KULLANILMASI

Marmara Havzası son yıllardaki endüstriyel faaliyetlerden olumsuz etkilenmektedir. Artan endüstriyel faaliyetlerin bir sonucu olarak, Marmara Denizi'nde zararlı alg oluşumları, kırmızı gelgitler ve müsilaj olayları artmıştır. Endüstriyel atıklar ve atık su arıtma tesislerinden biyolojik arıtılmadan deşarj edilen sular nedeniyle müsilaj sorunu son zamanlarda zirve seviyelere ulaşmıştır. Endüstriyel atıkların ve arıtılmamış atık suların deniz ekosistemi üzerindeki olumsuz etkileri, su kalitesinin düşmesine neden olmuş ve biyolojik çeşitlilik için ciddi bir tehdit haline gelmiştir. Deniz ekosisteminde ortaya çıkan sorunlara karşı harekete geçmek için, su kaynaklarının doğru yönetimini sağlamak ve sürdürülebilir deniz ekosistemi hizmetlerini sürdürmek için biyo-izleme çalışmaları büyük önem kazanmıştır. Biyo-izleme çalışmaları kapsamında, yeni nesil DNA dizileme teknolojilerine dayalı DNA metabarkodlama uygulamaları, mikroorganizmaların ve plankton gruplarının tespitinde güvenilir ve etkin yaklaşımlar sunmaktadır. Bu tür teknolojiler, müsilaj olaylarının zararlı etkilerinin izlenmesinde önemli bir rol oynar ve deniz ekosistemlerinin sağlığı için erken uyarı sistemleri olarak yardımcı olabilir. Bu tezde diatom cinslerini inceleyerek aşırı plankton ve müsilaj artışının etkilerini ve müsilajın su kalitesine etkilerini inceleyeceğim. Sonuçlarımıza göre, *Skeletonema* ve *Nitzschia* diatom cinsleri 10 tür isabetinde baskın organizma gruplarıdır; ancak *Guinardia* ve *Pseudo-nitzschia*, 100 tür isabetinde baskın olanlardır. Bu cinslerin müsilaj oluşumu üzerinde potansiyel bir etkisi olabilir.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS/ABBREVIATIONS

Symbol	**Explanation**
L	Liter
µl	Microliter
µM	Micromolar
ºC	Celcius

Abbreviation	**Explanation**
A	Adenine
BOLD	Barcode of Life System
bp	Base pair
C	Cytosine
DNA	Deoxyribonucleic Acid
eDNA	Environmental Deoxyribonucleic Acid
G	Guanine
HPC	High Performance Computing
ml	Milliliter
NC	Negative Control
NCBI	National Center for Biotechnology Information
OTUs	Operational Taxonomic Units
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
qPCR	Quantitative Polymerase Chain Reaction
rbcL	Ribulose-bisphosphate carboxylase
sp.	Species
T	Thymine

1. INTRODUCTION

Besides providing excellent ecosystems for organisms, the seas help supply high quality food sources to humanity. Due to the ecological and economic importance of seas and oceans, it is extremely important to maintain the conscious and sustainable use of marine resources (Marmara, 2021). However, human actions such as overharvesting, land-based sources of pollution, climate change and ocean acidification are causing changes in marine ecosystems (Learnz, 2021). For instance, wastewater that have been discharged to the sea without using biological treatment create significant problems and threat for humans as well as the non-human species (Üçüncü, 2019). Fishing, pollution, and eutrophication are some of the other threats for the marine ecosystems (Learnz, 2021).

In the recent years, mucilage appeared as another phenomenon that has been observed and which comprises an important threat for the marine systems (İşinibilir Okyar et al., 2015). It has been demonstrated that the rising frequency of mucilage outbreaks is closely related to temperature anomalies in the Mediterranean Sea, and as a result, it is possible to expect that this relationship will intensify in response to the accelerated sea surface warming caused by the changing climate (Danovaro et al. 2009; Topcu and Öztürk, 2021).

Mucilage is the physical aggregation of particles and is produced by various marine organisms that can be harmful for the marine food web. Specifically, mucilage is an accumulation of organic matter that are formed by bacteria and unicellular phytoplanktonic organisms in response to changing environmental conditions (Marmara, 2021). It is also stated that exopolymeric chemicals released by stressed marine organisms, mostly filamentous algae in the benthic environment and single-celled microalgae in the pelagic environment, generate mucilaginous aggregates (Topcu and Öztürk, 2021). These molecules' colloidal characteristics transform them into a gelatinous organic substance known as marine snot or mucilage, which can reach considerable distances and cover vast areas (Giani et al. 2005; Aktan et al. 2008; Balkis et al. 2011; Topcu and Öztürk, 2021).

Diatoms are a group of eukaryotic unicellular algae that are known to occur in mucilaginous aggregates, with their extracellular release of polysaccharides implicated in this phenomenon (Aktan et al., 2008). In fact, diatoms are considered to be among the most successful species, along with prokaryotes that can perform photosynthesis (Lopez et al., 2005). They are estimated to generate at least 20% of the world's yearly primary productivity, which is equivalent to the output of tropical rain forests (Mock and Medlin, 2012). Given their ecological significance, diatoms are frequently utilized

in biomonitoring initiatives such as the Water Framework Directive (WFD) of the European Union (Valentin et al., 2019) and similar programs in Canada (Maitland et al., 2020). Furthermore, diatoms play a crucial role in marine ecology by contributing almost 23 percent of the Earth's total primary output (Serôdio and Lavaud, 2020). These photosynthetic activities also have a significant impact on the carbon cycle in benthic sedimentary ecosystems. As complex organic substances, such as carbohydrates in the cell contents of phytoplankton (including diatoms and dinoflagellates), are believed to be the source of mucilage formation in our seas, it is vital to investigate the role of diatoms in the structure of mucilage and help provide corresponding solutions.

Mucilage is a global phenomenon that have been observed in certain parts of the world. For instance, The Adriatic Sea experienced significant accumulations of mucilaginous aggregates in the late 1990s, especially in the northern half of the sea. In another period, mucilage growth was observed in June 2003 in Italy along the rocky cliffs of the Portofino Promontory (Schiaparelli et al., 2007). The 2003 occurrence in the northern Tyrrhenian Sea was distinct from any other in the past. *Acinetospora crinita* from the class Phaeophyceae was responsible for the mucilage in this region, which typically does not predominate in mucilage aggregates for that region (Schiaparelli et al., 2007). In addition to these, in the autumn of 2007, a sudden increase in mucilage production was observed in the central area of Ariake Sound, Japan, which coincided with a bloom of *Coscinodiscus* sp. diatoms. As of yet, the underlying mechanism responsible for the production of marine mucilage in Ariake Sound remains unresolved, warranting further investigation (Fukao and Kimoto, 2009).

Considering Turkey, mucilage was detected for the first time in the Sea of Marmara in 2007 (Aktan et al., 2008), and several diatoms and dinoflagellates were detected as the main contributors (Tüfekçi et al. 2010; Balkis et al. 2011, 2013; Toklu-Alıçlı et al. 2020). Similar incidents had previously been observed near Erdek Bay; however, these events were not documented or examined (Tüfekçi et al., 2010). Most recently, a significant mucilaginous outbreak that affected the entire Sea of Marmara in autumn 2020 was documented, spread far along the shores, and persisted on the water's surface until June 2021 (Özalp, 2021).

“Automated identification of multiple species from a single bulk sample containing entire organisms or from a single environmental sample containing degraded DNA is referred to as DNA metabarcoding” (Taberlet et al., 2012; Aylagas et al., 2014). It is an approach applicable to both modern and old environmental samples. DNA metabarcoding has a tremendous potential to increase the resolution of biodiversity assessments in the foreseeable future (Taberlet et al., 2012). Its applications have been very effective for detecting of microorganisms and plankton groups that can

have detrimental effects on the marine ecosystems (Bohan et al., 2017). Such applications also allow us to monitor planktons associated with mucilage events. For instance, a study conducted by Genitsaris et al. (2020) has indicated that phytoplankton increases and mucilage events were observed in Thessaloniki Bay in Greece. In addition to identifying previously unknown taxonomic taxa in the research region, the metabarcoding analysis undertaken in this study identified the recognized unicellular eukaryotic groups regularly found in the Bay during mucilage formation, which mostly comprised Bacillariophyta and Dinoflagellata. Another advantage of DNA metabarcoding is that it is cost-effective and easy to apply, making it useful for development of biodiversity monitoring programs at various geographic scales (Bohan et al., 2017; Creer et al., 2016; Taberlet, Bonin, Zinger, and Coissac, 2018; Thomsen and Willerslev, 2015).

The purpose of this thesis is to detect the problematic taxa that cannot be identified morphotaxonomically, to detect the species showing an early warning system and to develop emergency action plans for the development of biodiversity composition changes in microorganisms and plankton communities associated with excessive algal growth and the mucilage event occurring accordingly on the Sea of Marmara by investigating the diatoms specifically. Furthermore, we believe that next-generation sequencing based DNA metabarcoding tools will enable the development of more precise and ecologically effective marine management strategies, as well as the assessment of plankton associated with harmful algal blooms, red tides, and mucilage structure.

2. LITERATURE REVIEW

Mucilage is a natural formation that is rich in organic matter, and occurs widely in open seas and oceans. Surface water warming causes the stratification of water column which leads to the formation of the large marine aggregates (mucilage) by the combination of small-sized marine aggregates. These large aggregates can cover the sea surface along hundreds of kilometers of coastline (Danovaro et al., 2009).

Mucilage is dense, gel-like, and highly viscous in its formation. It has hydrogel features, and consists of polymeric substances and extracellular polysaccharides that are produced, released, or leaked by various marine species. It is also rich in dissolved and polymeric organic substances (Öztürk et al., 2021). Due to its gel-like and viscous features, it can contain different sea organisms of various sizes such as viruses, bacteria, phytoplankton and zooplankton. Mucilage is mainly composed of several carbohydrates, such as monosaccharides and complex polysaccharides. In addition to carbohydrates, mucilage contains other organic matters, proteins, nitrogen and phosphorus compounds, and inorganics such as aluminum and silicon. Mucilage also contains ions such as calcium and iron. Calcium and iron are believed to have gelation mechanisms, hence these ions might play a role in the formation of the gel-like structure of mucilage (Giani et al., 2005). Since mucilage is rich in organic matter, it is a substantial source of nutrients for heterotrophic microorganisms that reproduce on organic matter and living creatures, and creates habitat and feeding grounds for these organisms (Öztürk et al., 2021).

There are several types of pollution that can cause the formation of mucilage and other types of marine pollution. Some of the pollution types include household waste, industrial waste, agricultural waste, waste from ships and marine vehicles, and waste carried from other seas. For example, industrial waste is one of the main reasons of the mucilage problem in the Sea of Marmara. There are a lot of industrial facilities around the Sea of Marmara that produce considerable amounts of industrial and chemical waste that can reach the sea after being treated, or without being treated completely. Because industrial waste contains heavy metals and chemicals, these types of waste cannot be sufficiently treated in the treatment plants where domestic waste can be treated. To reduce the adverse effects of industrial waste, it is crucial to prevent them from being mixed into the sea without proper treatment. It is also necessary to develop control systems that will increase the recycling and treatment quality of the treatment systems, and to produce advanced technological solutions (Öztürk et al., 2021).

As well as marine pollution, wastewater treatment plants are a major contributor to the mucilage problem marine pollution. The city of Istanbul, for example, discharges wastewater from both pre-treatment and advanced treatment plants into the Bosporus waters or the Sea of Marmara, which has a complex flow structure impacted by various environmental factors such as temperature, salinity, waves, currents, tides, turbulence mixing, wind, air temperature, atmospheric pressure changes, and solids transport. These factors ultimately determine the water quality in the Bosporus and the Sea of Marmara, including their impact on the formation of mucilage. However, Istanbul is not the only city contributing to this problem. Other highly industrialized cities, such as Tekirdağ, gulf of Gemlik, and Susurluk basin, also transmit their wastewater to the Sea of Marmara, further exacerbating the issue.

As a result of domestic and industrial waste production, organic and inorganic toxic pollutants in the Sea of Marmara are found at alarming levels. Fish and natural life exposed to the pollutants are adversely affected by this situation. Also, large quantities of marine mucilage can result in the obstruction of gillnets and trawl nets, causing significant material harm and presenting a considerable economic challenge for the fishing sector (Fukao and Kimoto, 2009). Moreover, nitrogen and phosphorus in the wastewater transmitted to the Bosporus and the Sea of Marmara constitute a food source for microorganisms in the waters and cause excessive growth of algae. Mucilage formation is also affected by these kinds of conditions (TMMOB, 2015). Figure 2.1 shows a visual example of mucilage in the Sea of Marmara.

Figure 2.1. Mucilage formation in the Sea of Marmara (Hurriyet, 2021).

After the detection of mucilage, bio-monitoring studies have become crucial for the management of the integrated water resources efficiently, and to maintain the sustainability of the marine ecosystems and their services. Biodiversity monitoring is the fundamental basis for assessing the environmental impact of anthropogenic activities. Multiple recent research studies have shown that high-throughput amplicon sequencing of environmental DNA (e-DNA metabarcoding) can overcome many of the constraints of conventional morphotaxonomy-based bioassessment methods for bio-monitoring purposes (Cordier et al., 2018). Also, in compliance with the applicable standards for the environmental water quality determined by nutrient concentrations as well as the phytoplankton quality element, an organized and routine monitoring of coastal waters is also necessary (Water Framework Directive; WFD, 2000/60/EC) (Genitsaris et al., 2019).

Besides the Sea of Marmara, mucilage has been observed in the Thessaloniki Bay (inner part of Thermaikos Gulf) located in northern Greece (Genitsaris et al., 2020), and the metabarcoding approach has been used to evaluate its taxonomic composition. Although wastewater treatment has been applied in that location, the limited water circulation and shallowness of the Thessaloniki Bay has been causing the detrimental algal blooms, red tides, and mucilage aggregates (Genitsaris et al., 2019). As a result of this situation, individuals and officials are becoming increasingly concerned about the Bay's water quality, particularly about the metropolitan front (Genitsaris et al., 2020). In this study, the metabarcoding analysis identified taxonomic taxa that has previously been unreported in the literature. The metabarcoding analysis also identified unicellular eukaryotic groups that are typically seen in the Bay, to be dominated by Bacillariophyta and Dinoflagellata. Genitsaris et al. stated that the most common OTUs were linked to species involved in red tides, hazardous blooms, and mucilage aggregates. In addition to the Thessaloniki Bay study, diatoms were found to be the predominant microalgal species in the benthic mucilage collected from multiple locations in the Tuscan Archipelago between May 1999 and July 2002 (De Philippis et al., 2005). During the entire study period, *Synedra*, *Licmophora*, and *Navicula* were the most commonly observed genera of diatoms in the benthic mucilage, while the presence of the other diatom genera varied based on the month and year of observation (De Philippis et al., 2005).

This thesis will examine the community composition and spatial distribution of diatom groups related to the harmful plankton overgrowth and mucilage structure by the next-generation DNA sequencing technology-based environmental DNA metabarcoding methods. We will conduct our analyses with the samples taken from the Sea of Marmara. Thus, a biological monitoring study will be conducted with DNA-based methods at pre-determined locations in the Sea of Marmara. The diversity of diatom groups associated with the mucilage events, will be analyzed using molecular and

phylogenetic methods. The primary purpose of this thesis will be to identify the diatom genera that contribute to the mucilage problem using DNA metabarcoding tools.

3. MATERIALS AND METHODS

3.1. Sample Collection

The points where field studies were carried out within the framework of the thesis are shown in Figure 3.1 and Table 3.1. In the study, mucilage (three samples) representing each station, were collected from six sampling stations (M-coded, Figure 1 and Table 1) where mucilage was detected in June, 2021 from the surface of the sea. Samples were collected into sterilized plastic sampling bottles with a volume of 1L and DNA sampling was done by passing these samples through Merck Millipore brand Sterivex filters with 0.22 µM pore size. Next, each sample was fixed by adding 2 µl of RNAlater and 2 µl of PBS. Same procedure was also applied for one negative control (1 Sterivex) for each sample point. As we have shown in Figure 3.2 below, the collection of our samples was facilitated by the Yunus-S R/V research ship of Istanbul University.

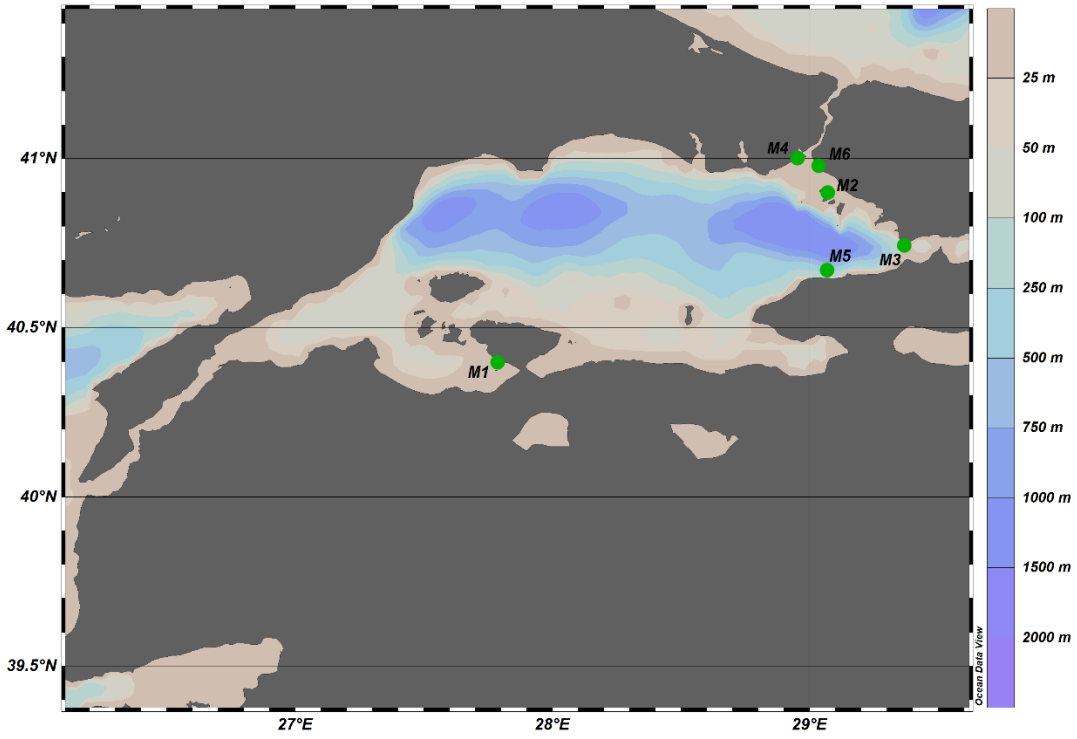

Figure 3.1. Sampling stations. The M code (green) represents the mucilage samples taken from six different locations.

Table 3.1. Sampling point information on mucilage samples.

Code	Location	Coordinates (N, E)	Collection Date
M1	Erdek	40.398; 27.787	16/06/2021
M2	Kınalıada	40.899; 29.070	05/06/2021
M3	Gulf of Izmit	40.743; 29.367	04/062021
M4	Yenikapı	41.001; 28.953	06/06/2021
M5	Çınarcık	40.692; 29.327	05/06/2021
M6	Kalamış	40.978; 29.035	10/10/2021

Figure 3.2. A view from the Yunus-S R/V research ship that helped us collect the samples.

3.2. Laboratory Procedures

The samples obtained within the framework of the thesis were examined in the laboratory, and the following steps have been applied:

- DNA isolation,
- Measuring the concentrations of DNA isolates using Qubit,
- 1st PCR (target rbcL gene) two replicate PCRs for the gene,
- Visual confirmation of 1st PCR products by gel electrophoresis,
- 2nd PCR (addition of Illumina adapters),
- Visual confirmation of 2nd PCR products by gel electrophoresis,
- Assembly of 2nd PCR products and clean-up with Ampure beads
- High-resolution concentration measurement by qPCR.

Details on these steps are given below:

A) DNA isolation

DNA isolation was performed from each mucilage-derived location using three mucilage-containing Sterivex filters, 1 mucilage-free water sample-passed Sterivex filter, and one negative control following the procedures below.

As we can see in Figure 3.3, we took our samples from 15 ml tubes into petri dishes and cut them into small pieces with sterile disposable scalpel tips to make them ready for DNA isolation. Figure 3.4 shows the images of these samples inside the UV hood.

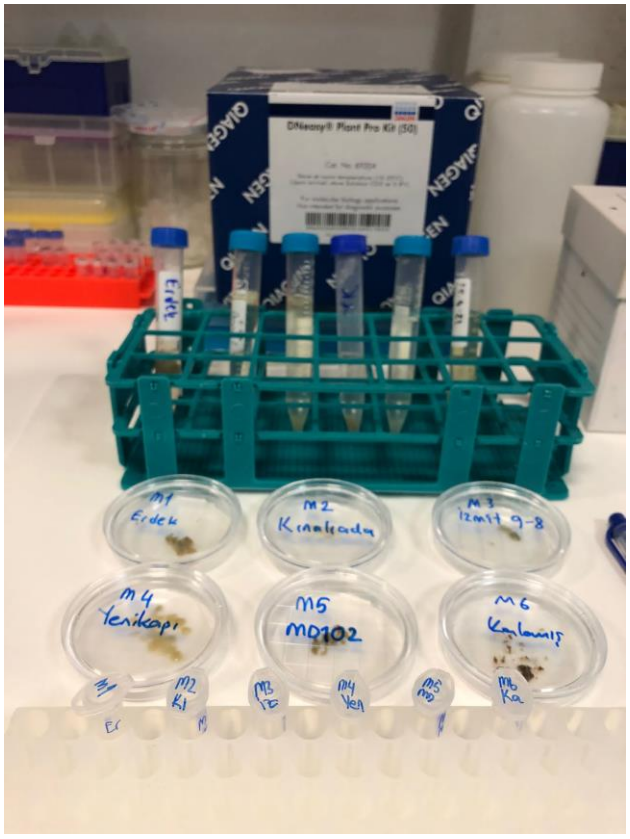

Figure 3.3. Mucilage samples from six different locations during DNA extraction.

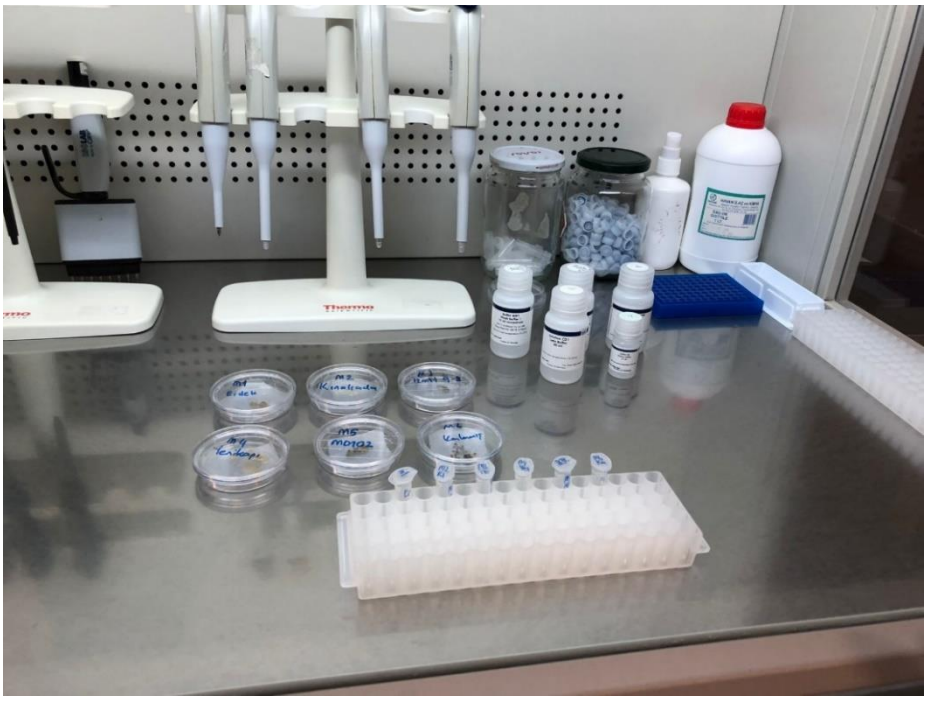

Figure 3.4. UV hood image during DNA extraction with commercial kit (Plant Pro kit, Qiagen).

DNA extractions of mucilage samples were carried out in accordance with the commercial kit DNeasy Plant Pro (Qiagen, Germany). Below are the details of the protocol:

- 5–100 mg of mucilage and 500 µl of Solution CD1 were added to a 2 ml tissue lysis tube. Then the mixture was vortexed briefly.
- Samples were homogenized with the help of the TissueLyser LT device; for this procedure the samples in the tube were placed in the TissueLyser LT device and run at 24 Hz for 2 minutes.
- Samples were centrifuged in tissue disruption tubes at 12,000 x g for 2 minutes.
- The supernatant was transferred to a clean 1.5 ml microcentrifuge tube.
- 200 µl of Solution CD2 was added and vortexed for 5 seconds.
- The samples were centrifuged at 12,000 x g for 1 min at room temperature. Avoiding the pellet, the supernatant was transferred to a clean 1.5 ml microcentrifuge tube.
- 500 µl Buffer APP was added and vortexed for 5 seconds.
- 600 µl of lysate was loaded onto an MB Spin column. Then, the microcentrifuge tube was centrifuged at 12,000 x g for 1 minute.
- After discarding the collection tube, the filter was placed in a new collection tube and step eight was repeated to ensure all lysate had passed through the MB spin column.
- The MB spin column was placed in a clean 2 ml collection tube.
- 650 µl of Buffer AW1 was added to the MB spin column. It was then centrifuged at 12,000 x g for 1 minute. The collection tube was discarded and the MB spin column was placed back in the same 2 ml collection tube.

- 650 µl of Buffer AW2 was added to the MB spin column. It was then centrifuged at 12,000 x g for 1 minute. The collection tube was discarded and the MB spin column was placed in the same 2 ml collection tube.
- Then, the collection tube has been centrifuged at 16,000 x g for 2 minutes. The MB spin column was placed in a new 1.5 ml elution tube.
- 50–100 µl of EB (Elution Buffer) Buffer was added to the center of the white filter membrane and centrifuged at 12,000 x g for 1 minute. The MB spin column was then discarded.
- DNA extract was taken into a new sterile centrifuge tube (1.5 mL) and stored at -20ºC for the subsequent steps.

B) Measuring the densities of DNA isolates with the help of Qubit

C) 1st PCR (target rbcL gene) two replicate PCRs for the gene

After DNA extraction, first PCR was applied to the isolates. In order to detect diatoms in the first PCR step, the optimization process was carried out with primers suitable for amplifying the rbcL gene region. In this context, the appropriate annealing temperature for the primer pair was determined by gradient PCR. Positive and negative controls were included in this process at each step. Different PCR master mixes and high compatibility Taq polymerases were also evaluated during the optimization steps. After optimization, the OnePCR SuperMix (GeneDireX, Inc., Taiwan) commercial kit was seen to provide the best results and was used. A total of 18 PCR products (6 mucilage samples x 3 DNA) isolates / region, were obtained. In addition, the same laboratory procedures were applied to 1 negative control.

Below are these optimal conditions for the rbcL gene:

The PCR regime for the **rbcL gene** was an initial denaturation at 94°C for 2 minutes, followed by 35 cycles of denaturation at 94°C for 45 seconds, annealing at 55°C for 45 seconds, elongation at 72°C for 1 minute, and then a final elongation at 72°C for 10 minutes, followed by cooling at 4°C to complete PCR.

The table below shows the information of the primer pairs we applied to the rbcL gene region for the first PCR.

Table 3.2. The primer pair used for the first PCR appropriate to the rbcL gene region

Primer Name	Sequences
DIV4rev3R (Reverse)	GTCTCGTGGGCTCGGAGATGTGTATAA GAGACAGCTCTGACAATGGAATACGAATA
DIV4F (Forward)	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCTAC GGGNGGCWGCAGGCGGTAATTCCAGCTCCAATAG

Chemical conditions in PCR studies are given in the table below.

Table 1.3. First PCR conditions

Chemical	1 x (25 µl)
Master Mix(GeneDireX)	12.5 µl
Primer F (10 µM)	1.2 µl
Primer R (10 µM)	1.2 µl
DNA	2.5 µl
Water	7.6 µl
Overall	25 µl

D) Visual confirmation of 1st PCR products

After the first PCR, we observed DNA bands in our samples with the help of 1% agarose gel electrophoresis. Figure 3.5 below shows our gel image after PCR. As we can see in the image, we found DNA in our samples, and as we expected, no bands are seen in our last negative sample. After successful PCRs, the obtained PCR products were stored in a -20°C freezer following gel electrophoresis.

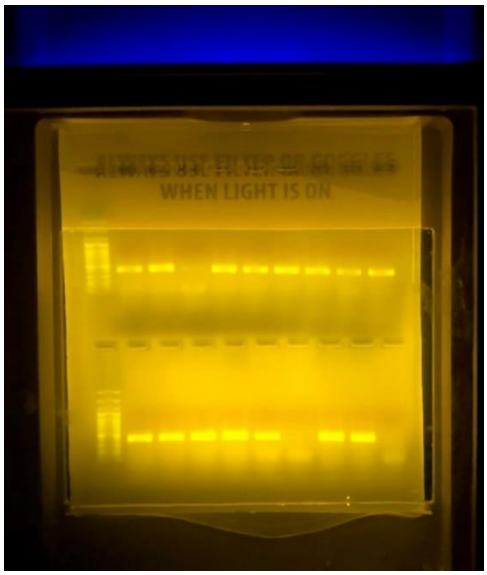

Figure 3.5. Gel images of our mucilage samples after PCR

E) 2nd PCR (addition of Illumina Miseq adapters)

At this stage, adapters for sequencing on the Illumina platform and indices in the i5 and i7 directions were added to each 1st PCR product, which enabled the data to be demultiplexed from the combined final sequencing pool. The chemical conditions of the second PCR are given in Table 3.4. In Table 3.5, the sequences of the added i5 and i7 indexes are provided.

The second PCR regime for the rbcL gene included an initial denaturation at 95°C for 3 minutes, followed by 20 cycles of denaturation at 95°C for 30 seconds, annealing at 60.5°C for 30 seconds, elongation at 72°C for 30 seconds, and then a final elongation at 72°C for 5 minutes, followed by cooling at 4°C to complete PCR.

Table 3.4. Second PCR chemical conditions.

Chemical	1 x (25 µl)
Master Mix(GeneDireX)	12.5 µl
Primer i5	1 µl
Primer i7	1 µl
DNA (PCR-1 product)	2.5 µl
Water	8 µl
Overall	25 µl

Table 3.2. During the second PCR of mucilage samples, barcodes were assigned with i5 and i7 barcodes, considering different location.

Location	**Sample ID**	**I5**	**I7**
Erdek-1	M1.1	ACGCGTGG	TCCTCTAC
Erdek-2	M1.2	ACGCGTGG	ATTACAAT
Erdek-3	M1.3	ACGCGTGG	GAATGATC
Kınalıada-1	M2.1	GGAACTCC	TCCTCTAC
Kınalıada-2	M2.2	GGAACTCC	ATTACAAT
Kınalıada-3	M2.3	GGAACTCC	GAATGATC
Gulf of Izmit-1	M3.1	TGGCCATG	TCCTCTAC
Gulf of Izmit-2	M3.2	TGGCCATG	ATTACAAT
Gulf of Izmit-3	M3.3	TGGCCATG	GAATGATC
Yenikapı-1	M4.1	GAGAGATT	TCCTCTAC
Yenikapı-2	M4.2	GAGAGATT	ATTACAAT
Yenikapı-3	M4.3	GAGAGATT	GAATGATC
Çınarcık-1	M5.1	CGCGGTTA	TCCTCTAC
Çınarcık-2	M5.2	CGCGGTTA	ATTACAAT
Çınarcık-3	M5.3	CGCGGTTA	GAATGATC
Kalamış-1	M6.1	GACCGCCA	TCCTCTAC
Kalamış-2	M6.2	GACCGCCA	ATTACAAT
Kalamış-3	M6.3	GACCGCCA	GAATGATC
NC	DIA_NC	TAAGATGG	GAATGATC

F) Visual confirmation of 2nd PCR products

In the figure below, the gel image after the 2nd PCR is given. As expected, DNA bands are distinct and appear in the expected range of 400-500 bases.

Figure 3.6. Gel images of our mucilage samples after 2nd PCR

G) Cleaning second PCR products with Ampure beads and pooling

In this step, the 2nd PCR products generated in the above-mentioned steps were combined based on the sample points (this way a single pool was created for each of the six mucilage samples separately and for the negative controls) and Ampure XP (Beckman Coulter, USA) was used for clean-up prior to the sequencing step.

H) High-resolution density measurement and assembly by qPCR of sample-based pooled DNA libraries

The concentrations of the pools of each obtained sample point were measured by qPCR for quantification prior to Illumina sequencing. For this, QuantStudio 5 qPCR machine (Applied Biosystems, USA) and qPCR kit designed for the KAPA 4484 (Roche, Germany) Illumina DNA library were used. Accordingly, the amount of DNA used from the samples other than the negative control was adjusted to be equal to the amount of DNA in the sample with the lowest concentration in nM, and all samples were collected in a single pool. A final sequencing pool was created by adding DNA libraries from negative controls, which was sent to Gen-Era (İstanbul, Turkey) for sequencing

on the Illumina NovaSeq platform, with the aim of getting an average of 2 million 150 base paired-end reads per sampling point.

3.3. Data Analyses

For the data analysis, High Performance Computing System (HPC) provided by the Turkish National e-Science e-Infrastructure (TRUBA) was used. After cleaning the data obtained using OBITools v1. (Boyer et al., 2016) in bioinformatics analysis, merging identical sequences, and merging sequences that differed by 5 percent or less from one sequence from the remaining sequences, the closest 10 and 100 species hits were identified for each different sequence available using BLASTn in the GenBank database with the efetch and the esummary commands. PRIMER-e v7 software was utilized to construct dendrograms, which were used to demonstrate the hierarchical clustering patterns between PCR replicates and sampling locations based on the databases. The specific codes used for each analysis is given below.

Ncbi blastn analysis: ./blastn -db nt -query (sample code.fsa) -outfmt 6“ std sscinames pident qcovs staxids" -max_target_seqs (10 and 100) -out (sample code.out) –remote

Efetch analysis: efetch -db taxonomy -id (id) -format xml | \

xtract -pattern Taxon -first TaxId -element Taxon -block "*/Taxon" \

-unless Rank -equals "no rank" -tab "," -sep "_" -element Rank,ScientificName

Esummary analysis: esummary -db taxonomy -id (id) |xtract -pattern DocumentSummary -tab "," -element Id Species

With a home-code written after this, the species with less than 97 percent similarity among these top 100 species were eliminated, and if the characterization in a certain taxonomic category is more than 85 percent identical in the remaining hits, the name was given in that taxonomic category. In order to increase the reliability of the results, the data were evaluated at the genus level and above.

4. RESULTS

4.1. Location-based mucilage characterization

Within the scope of the thesis studies, the characterization of the samples collected from Erdek, Kınalıada, Gulf of Izmit, Yenikapı, Çınarcık and Kalamış (Istanbul) was carried out. The findings of mucilage characterizations at these sampling points are shared below. The closest 10 and 100 species hits were identified for each different sequence available using BLASTn in the GenBank database, with taxonomic assignments undertaken using a custom script.

4.1.1. Erdek (M1) Characterization

The most reads were recorded at the station for the 10 species hits, and the two distinct diatom genera in the data set are *Skeletonema* and *Nitzschia*. On the other hand, the most reads were recorded at the same station for the 100 species hits, and the two distinct diatom genera in the data set are *Nitzschia* and *Achnanthes* (Figure 4.1). It is clear that Nitzschia has been observed the most frequently among both 10 and 100 species hits, based on the available data.

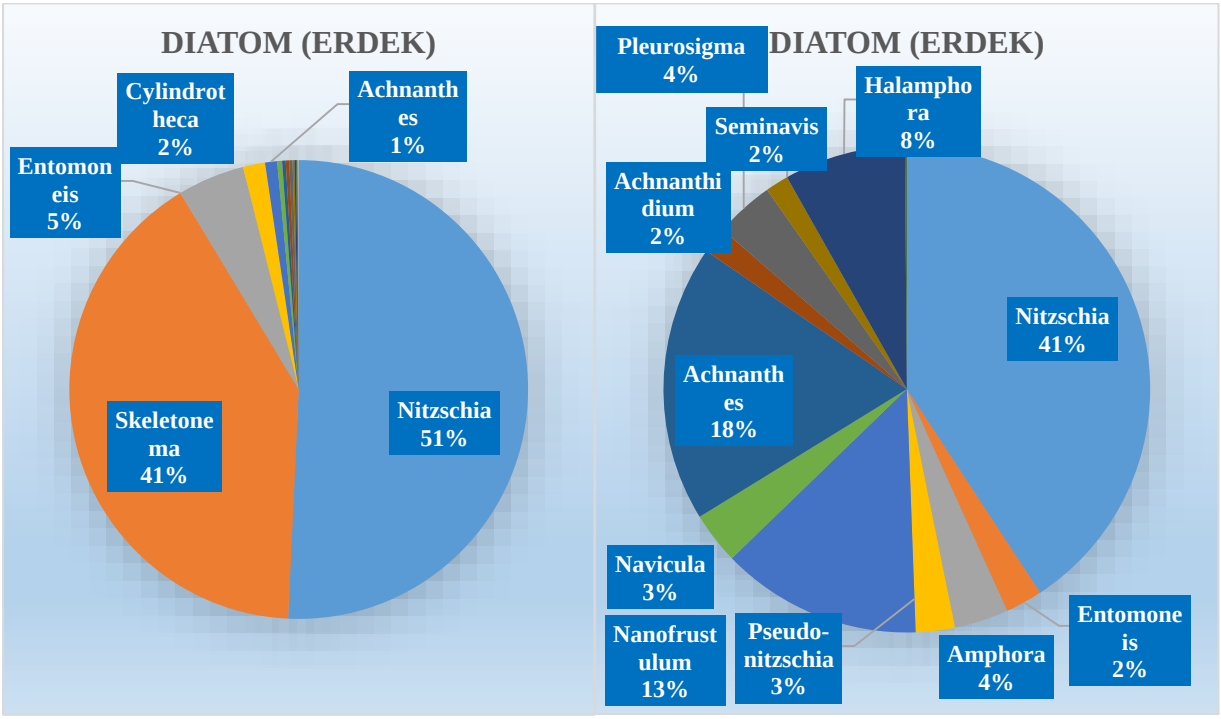

Figure 4.1. Diatom characterization based on genus-based sequence numbers for Erdek (10 hits left, 100 hits right).

4.1.2. Kınalıada (M2) Characterization

Kınalıada location was identified in the 10 and 100 species hits datasets. The former analysis (10 hits) recorded two distinct diatom genera, Cocconeis and Berkeleya. In contrast, for the 100 species hits, the two most distinct diatom genera recorded in terms of highest number of reads were *Guinardia* and *Pseudo-nitzschia* (Figure 4.2).

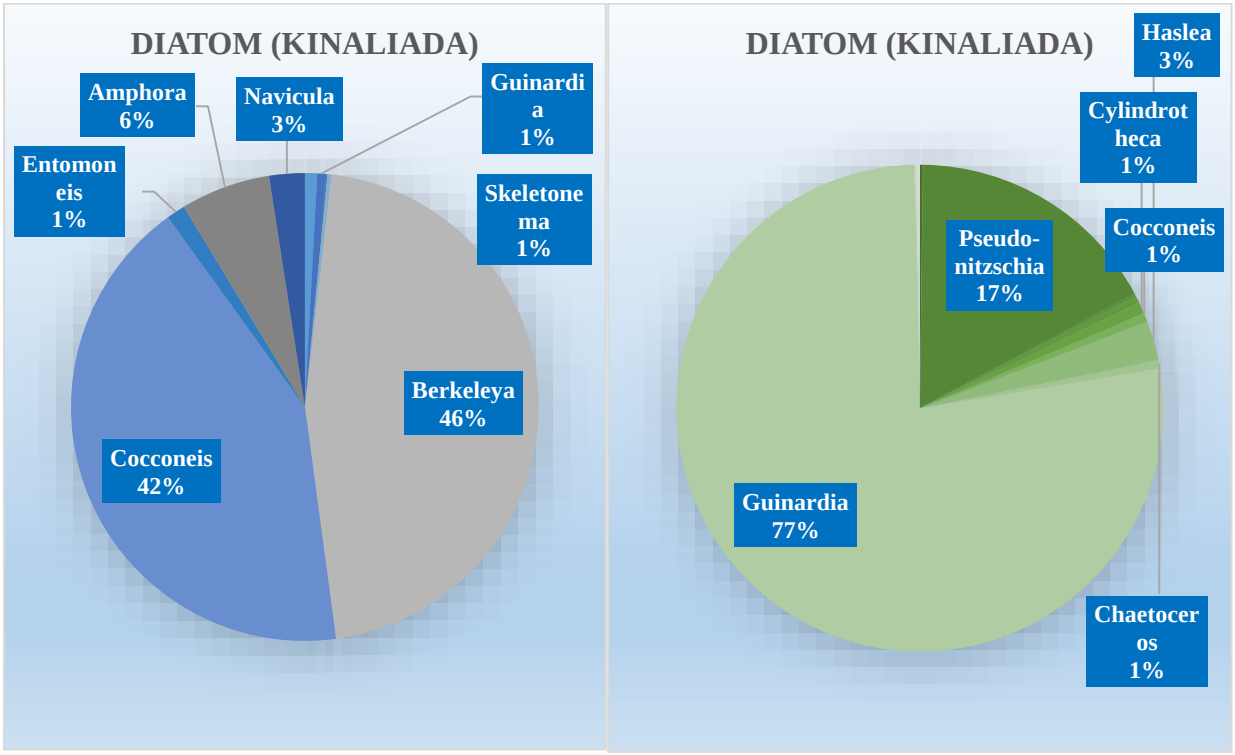

Figure 4.2. Diatom characterization based on genus-based sequence numbers for Kınalıada (10 hits left, 100 hits right).

4.1.3. Gulf of Izmit (M3) Characterization

Izmit location with the highest number of reads was found to harbor two distinct diatom genera, namely *Skeletonema* and *Guinardia*, in the analyzed dataset. However, for the 100 species hits dataset, the same station recorded the most reads, with the dominant diatom genus being *Guinardia* and the additional presence of *Pseudo-nitzschia* (as depicted in Figure 4.3).

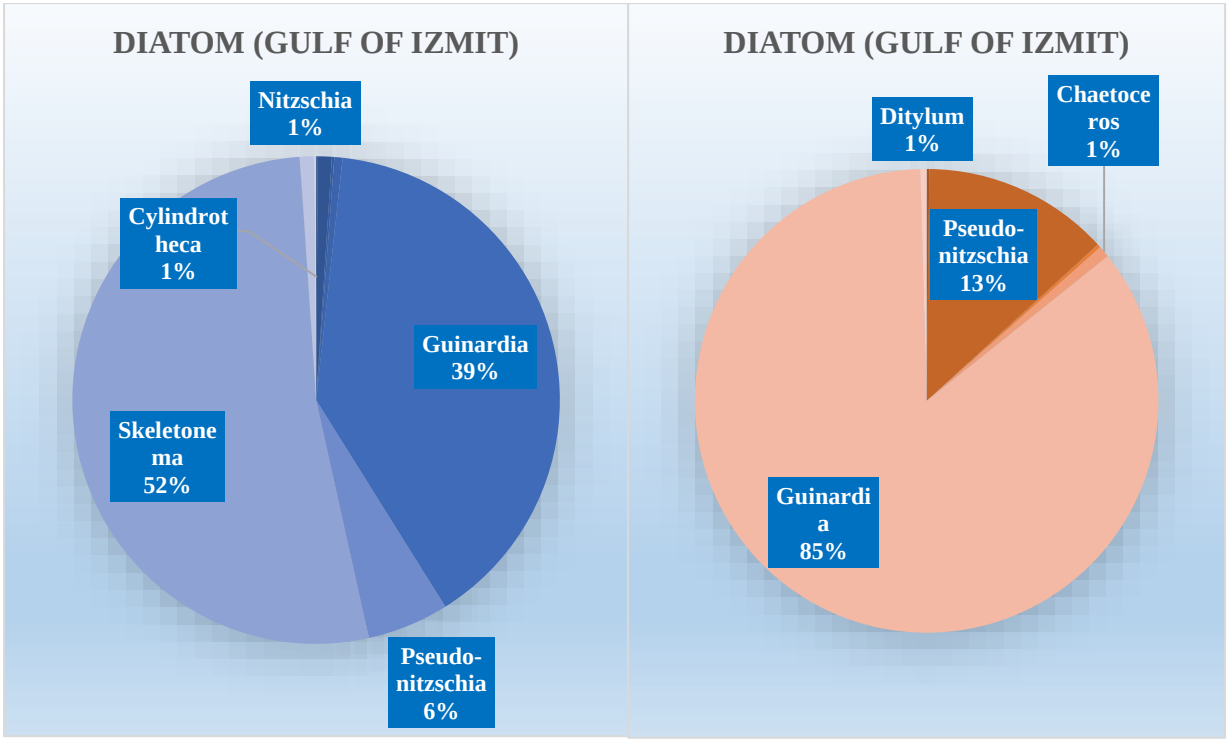

Figure 4.3. Diatom characterization based on genus-based sequence numbers for Gulf of Izmit (10 hits left, 100 hits right).

4.1.4. Yenikapı (M4) Characterization

Among the analyzed dataset for the 10 species hit, the station with the highest number of reads exhibited a notable presence of one dominant diatom genera, namely *Skeletonema* and *Guinardia* is the second genera that has been seen in this data set with a smaller portion. However, for the 100 species hits, the same station recorded the most reads, with *Guinardia* emerging as the dominant diatom genus, accompanied by the additional presence of *Pseudo-nitzschia* (as evidenced in Figure 4.4).

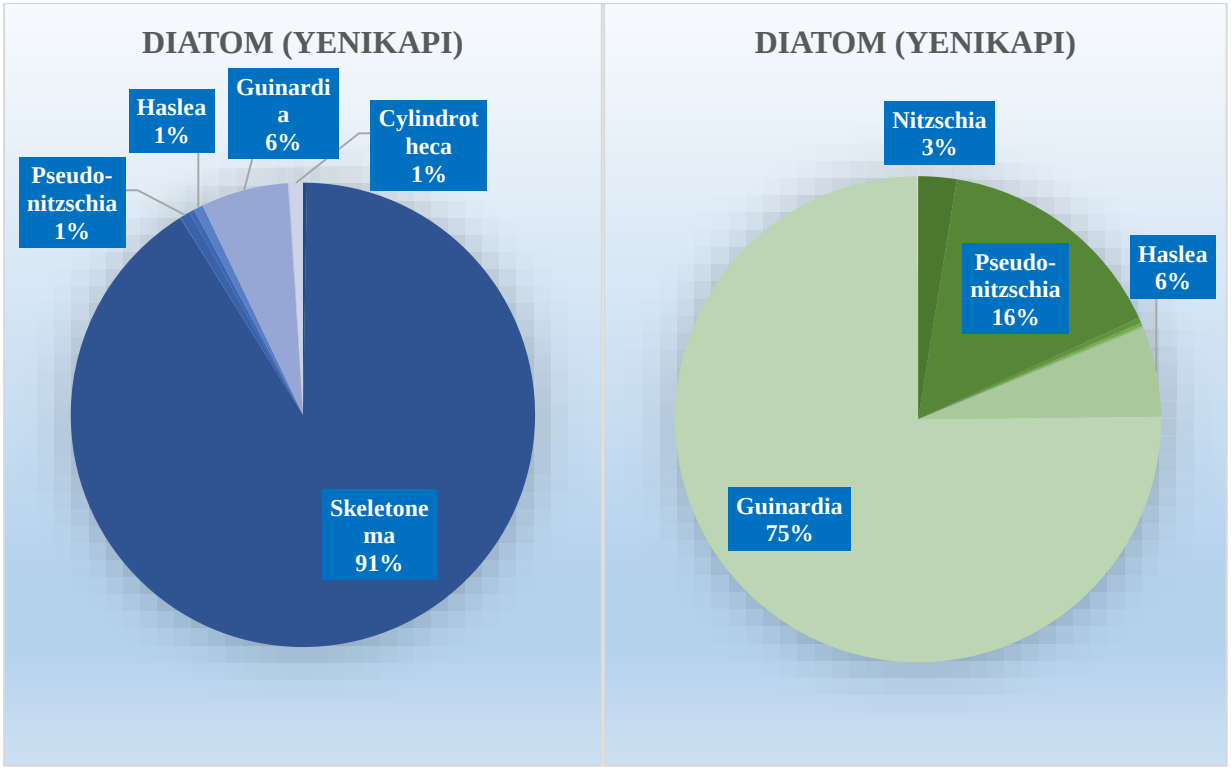

Figure 4.4. Diatom characterization based on genus-based sequence numbers for Yenikapı (10 hits left, 100 hits right).

4.1.5. Çınarcık (M5) Characterization

Looking at the figure below, *Skeletonema* is the predominant genus for the 10 species hits data, and *Pseudo-nitzschia* has been seen as the second dominant. *Guinardia*, *Cylindrotheca*, *Nitzschia*, and *Cocconeis* have also been observed in minor proportions. In contrast, *Pseudo-nitzschia* is a distinct genus for the 100 species hits; in addition, *Guinardia* and *Cocconeis* were accompanied by the additional presence of *Pseudo-nitzschia*.

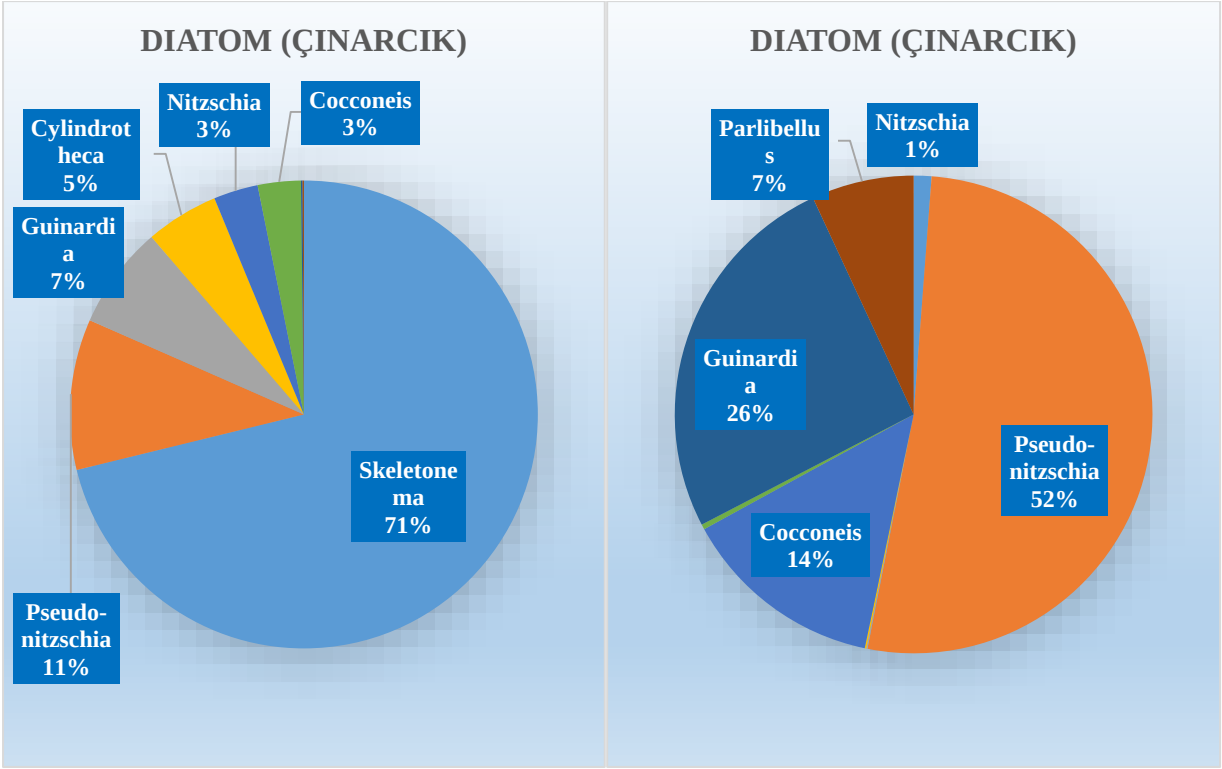

Figure 4.5. Diatom characterization based on genus-based sequence numbers for Çınarcık (10 hits left, 100 hits right).

4.1.6. Kalamış (M6) Characterization

The diatom genus *Skeletonema* appears to be the most prevalent and distinctive in the dataset, as evidenced by 10 species hits. Other genera, such as *Entomoneis*, *Cocconeis*, and *Pleurosigma*, are also observed, albeit at a lower frequency. Furthermore, among the 100 species hits recorded at the same station, the most significant number of reads were attributed to *Skeletonema*, *Cocconeis*, and *Pleurosigma*, which are the three most distinct diatom genera observed in Figure 4.6.

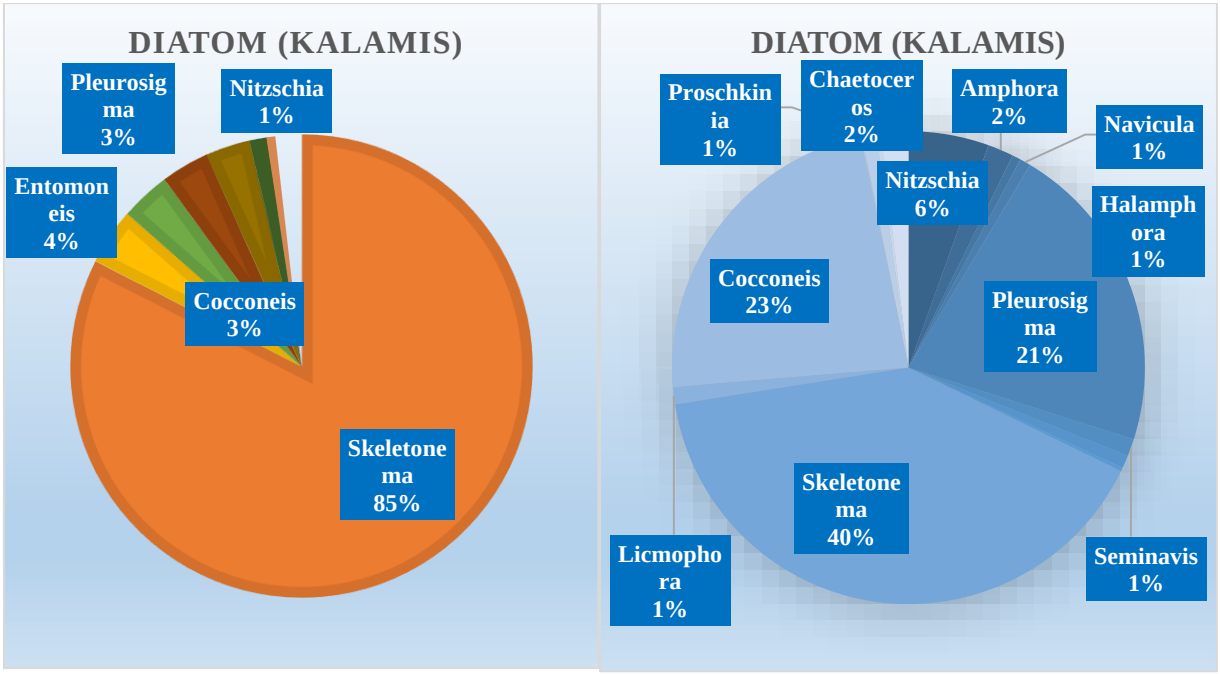

Figure 4.6. Diatom characterization based on genus-based sequence numbers for Kalamış (10 hits left, 100 hits right).

4.2. Diatom Genera Detected: A Graphical Overview of All Locations

In the study for 10 species hit, diatom DNA belonging to 25 genera was observed in the total data set. The three diatom genera with the most reads in the data set are *Skeletonema*, *Guinardia*, and *Nitzschia* (Figure 4.7).

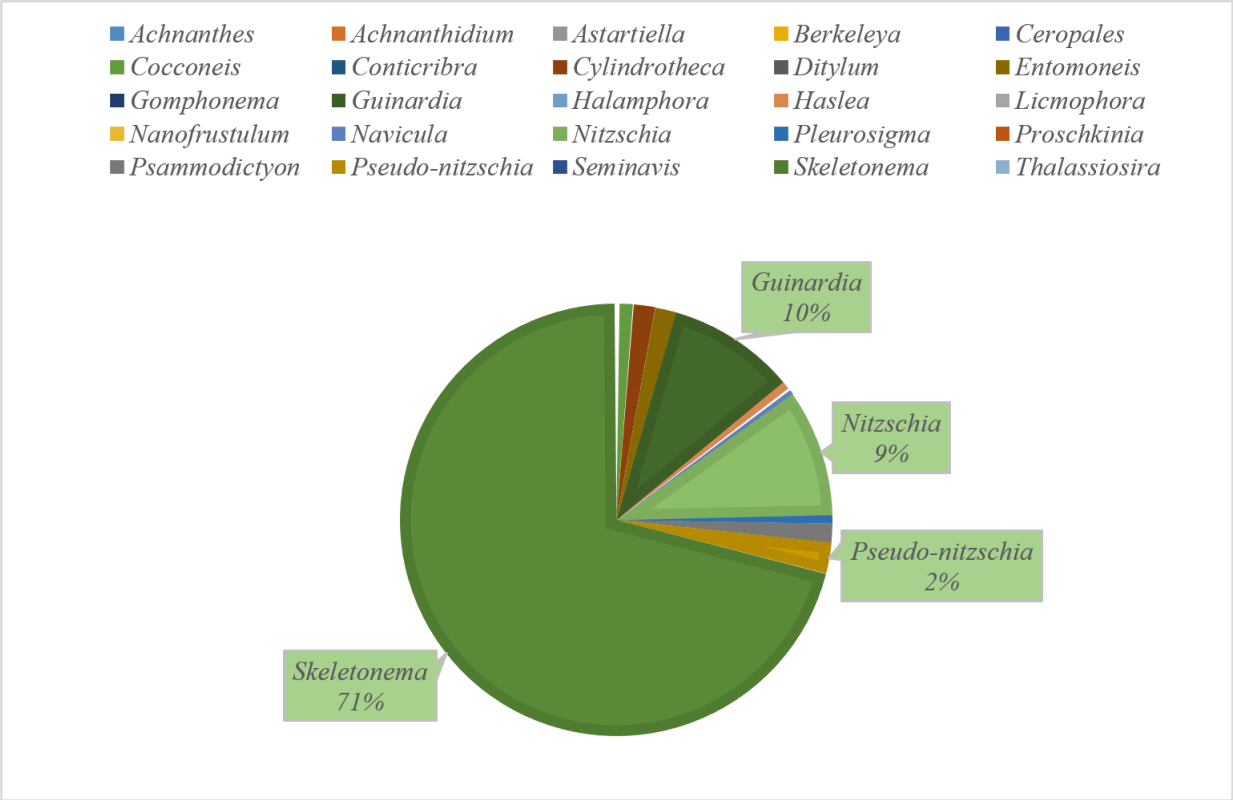

Figure 4.7. Diatom genera recorded in the whole data set.

In the study for 100 species hit, diatom DNA belonging to 24 genera was seen in the total data set. The two diatom genera with the most reads in the data set are *Guinardia* and *Pseudo-nitzschia* (Figure 4.8).

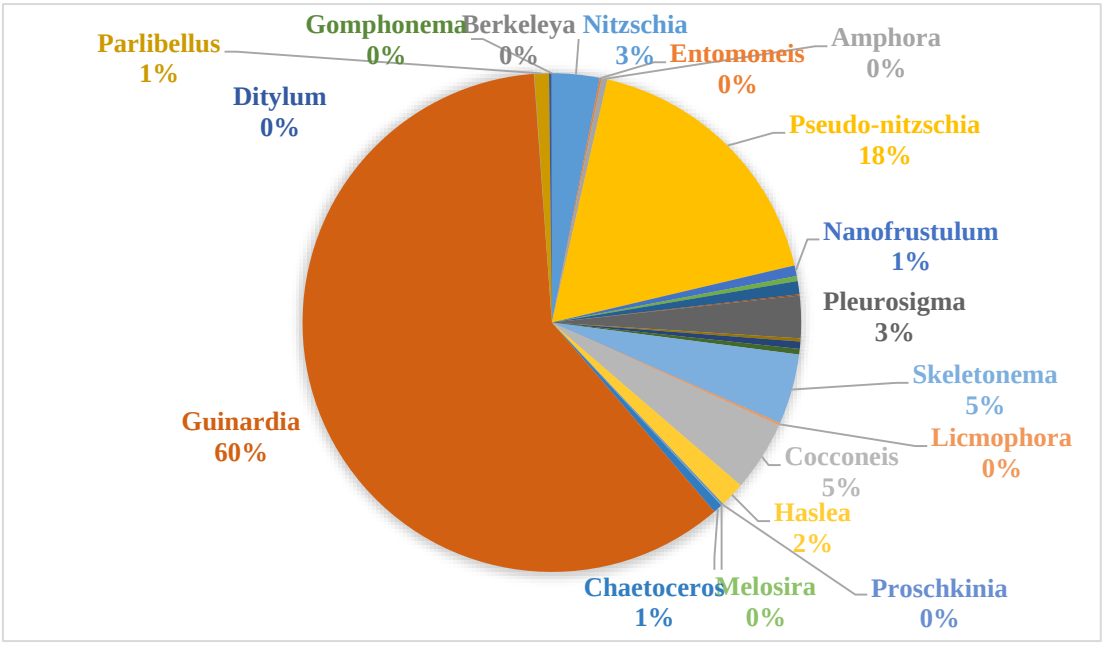

Figure 4.8. Diatom genera recorded in the whole data set for 100 species hit.

In the table below, the observation of diatom genera is evident, as a total of 29 genera have been documented in the analysis of both 10 and 100 species hits. The consistent presence of two genera, namely *Cylindrotheca* and *Nitzschia*, across all six locations and in both analyses, highlights their ubiquity and ecological significance.

Table 4.1. The frequency of occurrence of diatom genera at sampling points (both 10 and 100 species hits).

Genus	**M1(Erdek)**	**M2(Kınalıada)**	**M3(Gulf of Izmit)**	**M4(Yenikapı)**	**M5(Çınarcık)**	**M6(Kalamış)**	**TOTAL**
Achnanthes	(10-100)		10				(10-100)
Achnanthidium	(10-100)						(10-100)
Amphora	100	100				100	100
Astartiella	10					10	10
Berkeleya	10	(10-100)				10	(10-100)
Ceropales	10						10
Chaetoceros		100	100			100	100
Cocconeis		(10-100)			(10-100)	(10-100)	(10-100)
Conticribra	10			10		10	10
Cylindrotheca	(10-100)	(10-100)	(10-100)	(10-100)	(10-100)	(10-100)	(10-100)
Ditylum		100	(10-100)				(10-100)
Entomoneis	(10-100)	10	10	10	10	10	(10-100)
Gomphonema				(10-100)			(10-100)
Guinardia		(10-100)	(10-100)	(10-100)	(10-100)		(10-100)
Halamphora	(10-100)			10		(10-100)	(10-100)
Haslea		(10-100)		(10-100)	(10-100)	(10-100)	(10-100)
Licmophora				(10-100)		(10-100)	(10-100)
Melosira						100	100
Nanofrustulum	(10-100)					(10-100)	(10-100)

Genus	M1(Erdek)	M2(Kınalıada)	M3(Gulf of Izmit)	M4(Yenikapı)	M5(Çınarcık)	M6(Kalamış)	TOTAL
Navicula	(10-100)	(10-100)			100	(10-100)	(10-100)
Nitzschia	(10-100)	(10-100)	(10-100)	(10-100)	(10-100)	(10-100)	(10-100)
Parlibellus		100			100		100
Pleurosigma	(10-100)	100	(10-100)			(10-100)	(10-100)
Proschkinia						(10-100)	(10-100)
Psammodictyon	10					10	10
Pseudo-nitzschia	100	(10-100)	(10-100)	(10-100)	(10-100)		(10-100)
Seminavis	(10-100)			(10-100)		(10-100)	(10-100)
Skeletonema	10	(10-100)	10	(10-100)	10	(10-100)	(10-100)
Thalassiosira	10		10	10			10

The bar plots below provide a clear visual representation of the data in the table above for each location.

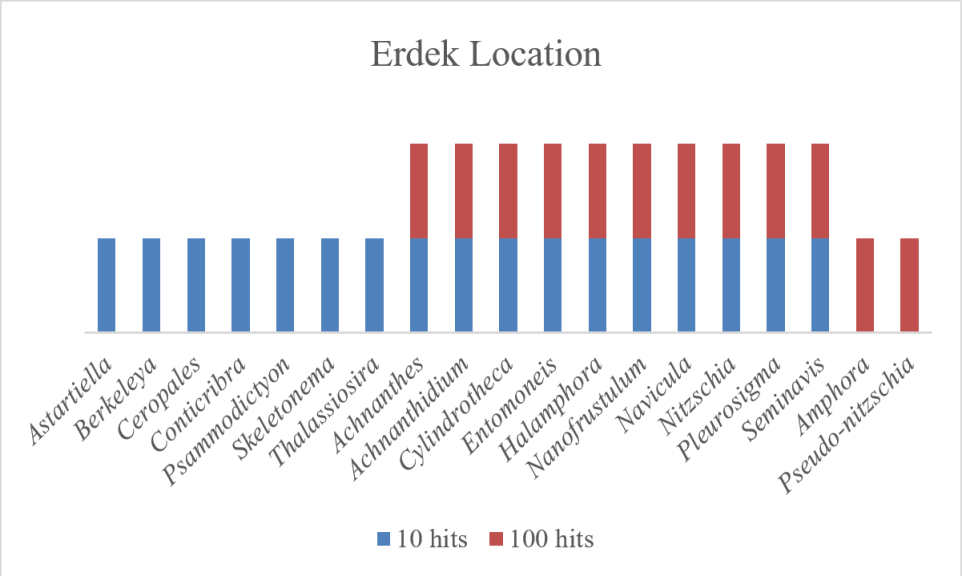

Figure 4.9. A bar plot representation of the location of Erdek.

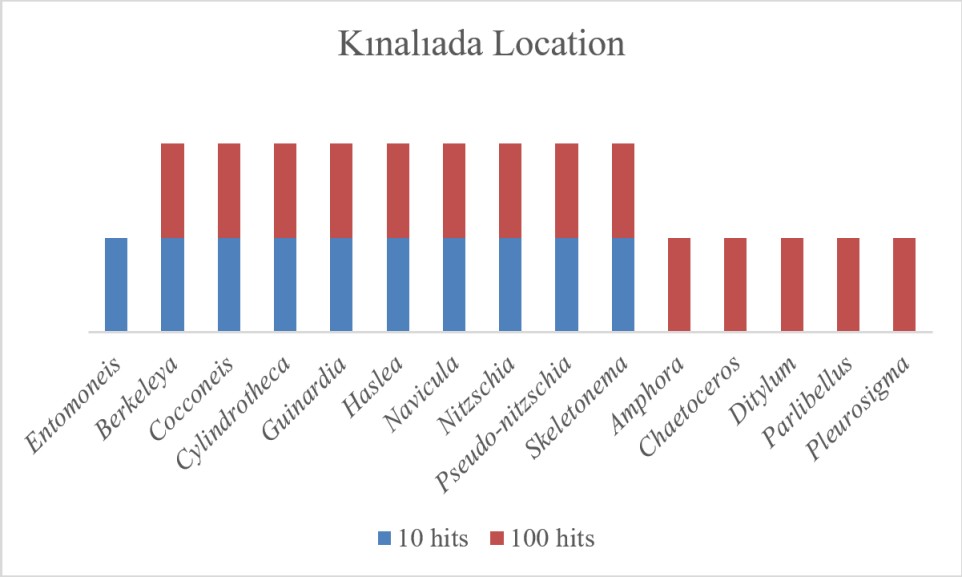

Figure 4.10. A bar plot representation of the location of Kınalıada.

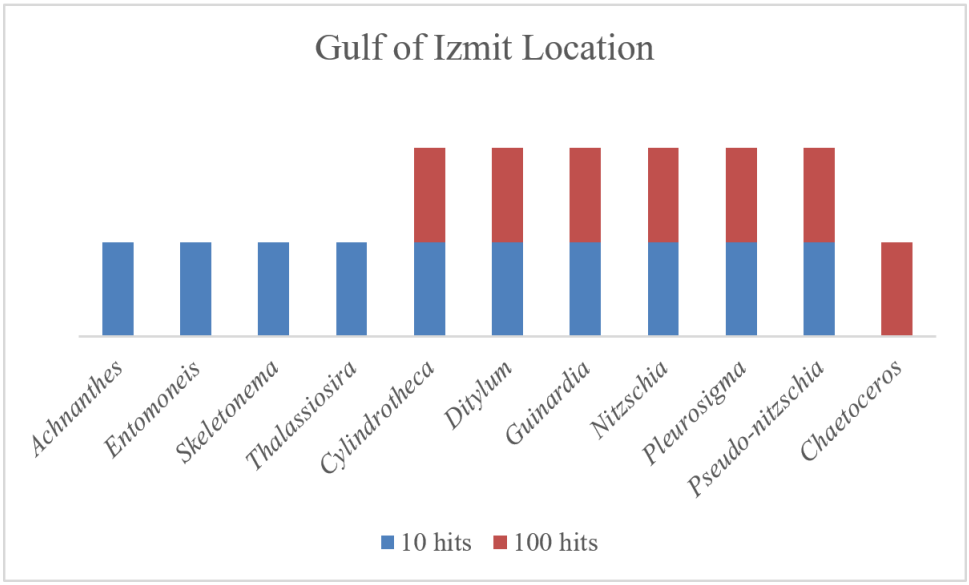

Figure 4.11. A bar plot representation of the location of Gulf of Izmit.

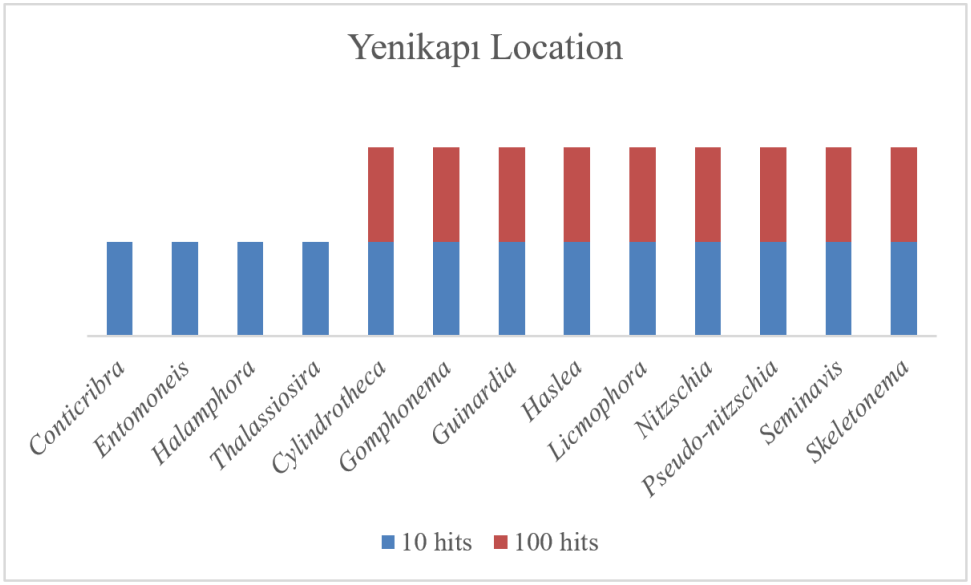

Figure 4.12. A bar plot representation of the location of Yenikapı.

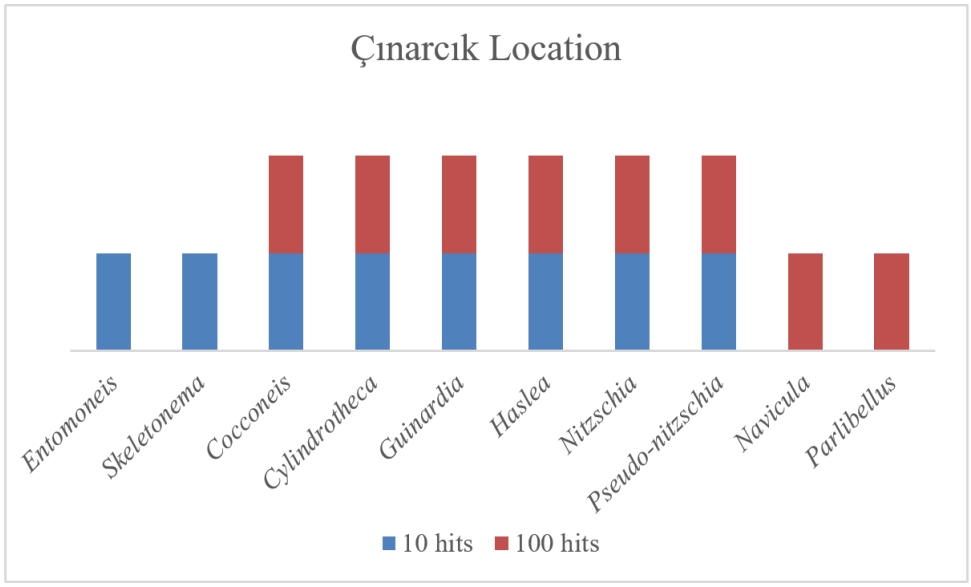

Figure 4.13. A bar plot representation of the location of Çınarcık.

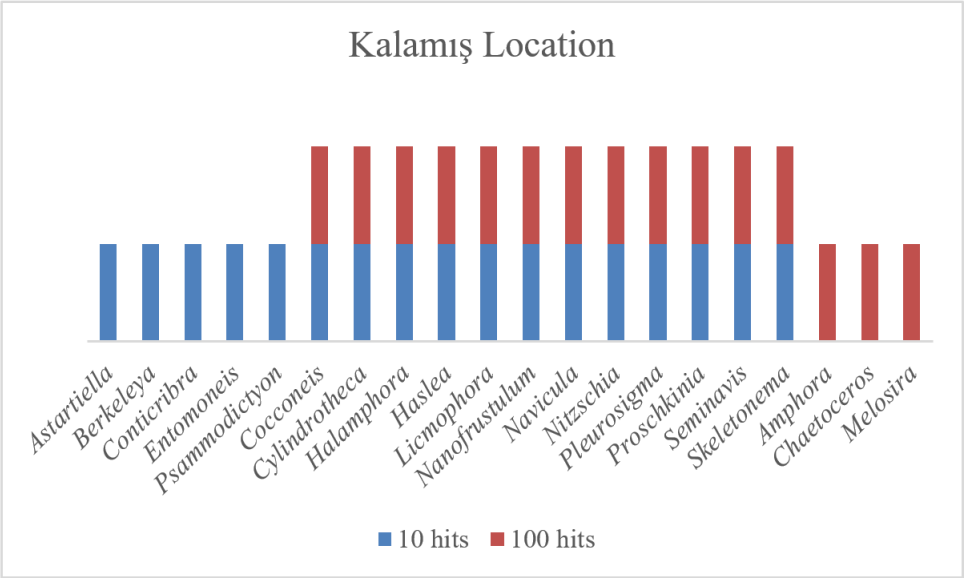

Figure 4.14. A bar plot representation of the location of Kalamış.

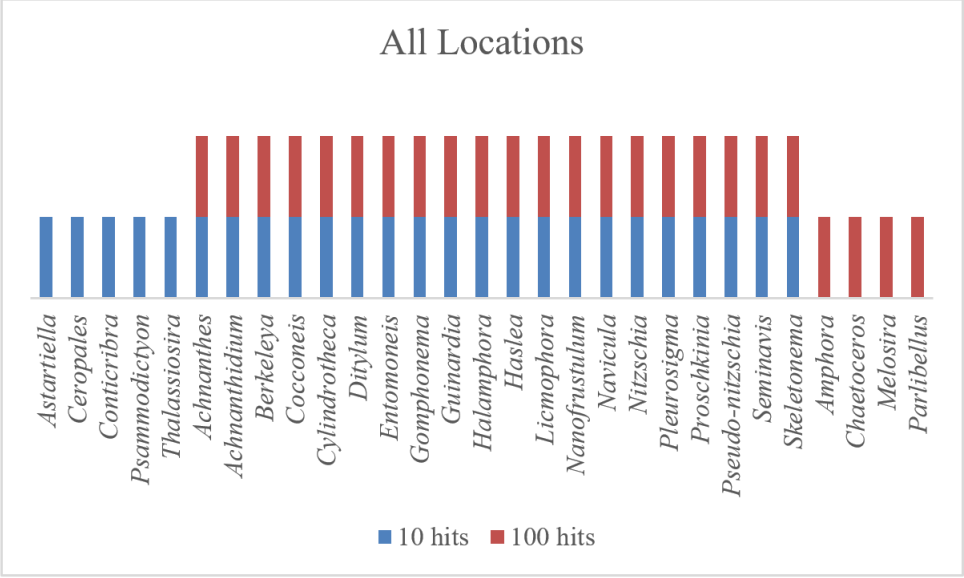

Figure 4.15. A bar plot representation of the location of all locations.

A dendrogram was constructed to visualize the similarities in genus diversity between the PCR replicates (refer to Figure 4.16). Notably, PCR replicates 5.3 and 6.3 exhibited anomalous data and appear to have served as negative controls. On the other hand, PCR replicates 1.1, 1.2, and 1.3 displayed a similarity percentage of nearly 75% in terms of taxa, suggesting that they share similar taxonomic profiles. While PCR replicate 6.1 originated from a different location, it exhibited similar genus diversity to that of M1 (Erdek location). Additionally, PCR replicates 2.1, 2.2, and 2.3, all collected from Kınalıada, were found to be comprising similar taxa as expected. The same pattern was observed with PCR replicates 4.1, 4.2, and 4.3. Notably, samples 3 and 5 showed a close

relationship to each other, likely due to their proximity to M3 (Gulf of Izmit) and M5 (Çınarcık) locations, which are geographically adjacent, as depicted in Figure 3.1, above.

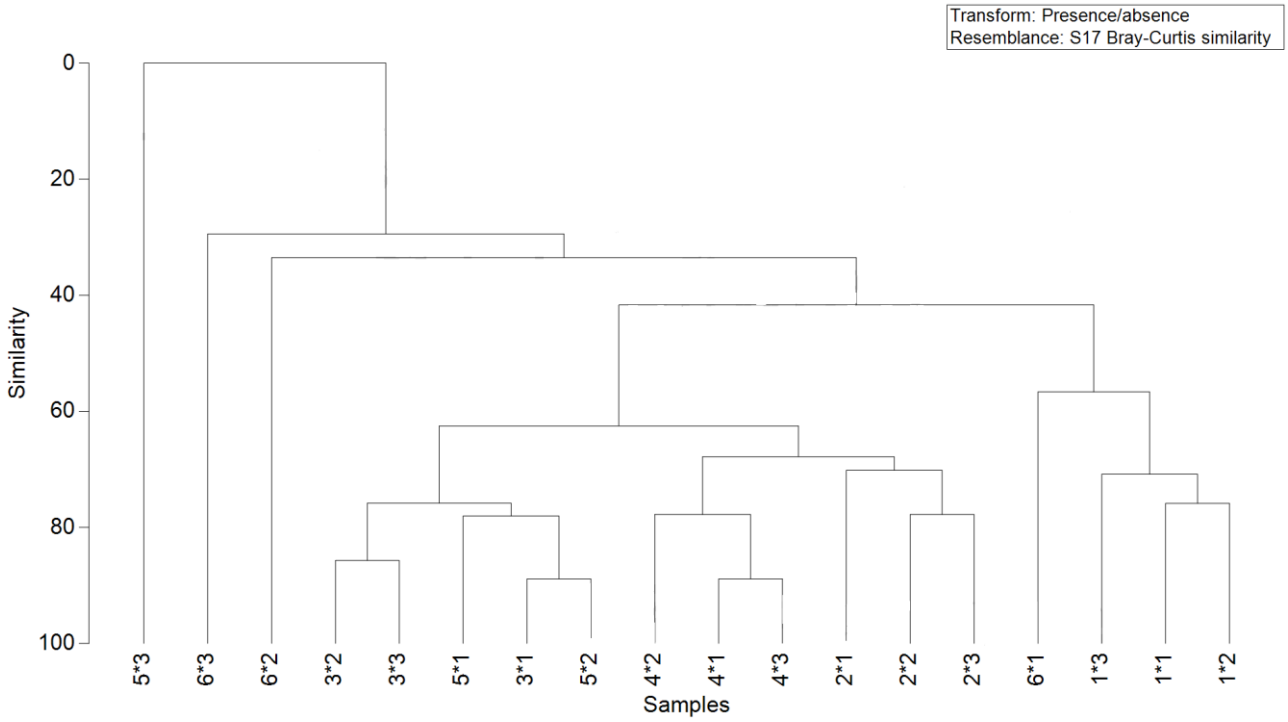

Figure 4.16. The dendrogram illustrates the degree of similarity in genera diversity observed across the various PCR replicates.

Below dendrogram (shown in figure 4.17) was constructed to visualize the hierarchical clustering patterns of genera level diatom diversity from each sampling location. The results indicated that the Gulf of Izmit and Çınarcık locations exhibited similar genera composition, which was expected given their close proximity. Similarly, the Kınalıada and Kalamış locations, which are also geographically adjacent, displayed clustering as anticipated. Interestingly, the Erdek and Yenikapı locations, despite not being geographically adjacent, were found to share a similar diatom genera composition, at a rate of nearly 60%. Overall, these findings suggest that proximity may play a role in shaping the composition of genera diversity observed in different locations.

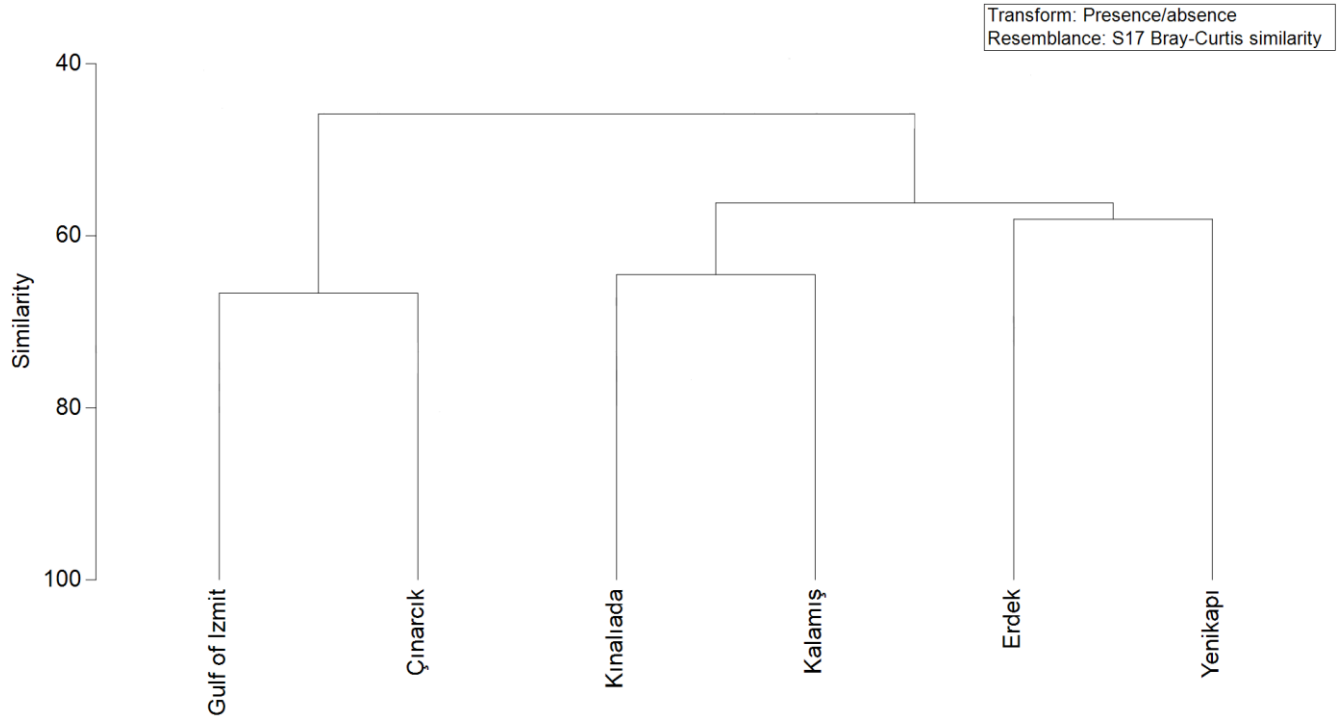

Figure 4.17. The dendrogram shows the degree of similarity in genera diversity observed across our sampling locations.

4.3. Comparison of Genera Identification in GenBank for 10 and 100 hits

By merging data from the GenBank databases, we were able to determine the number of distinct genera that have been identified for 10 and 100 hit results. The analysis revealed 25 distinct genera for the 10 hit results and 24 distinct genera for the 100 hit results, as identified by the GenBank databases. Figure 4.18 shows the genera found for 10 and 100 hit in the form of a Venn diagram. The results indicate a substantial overlap in the genera identified by both analyses, suggesting a considerable degree of similarity between the two datasets.

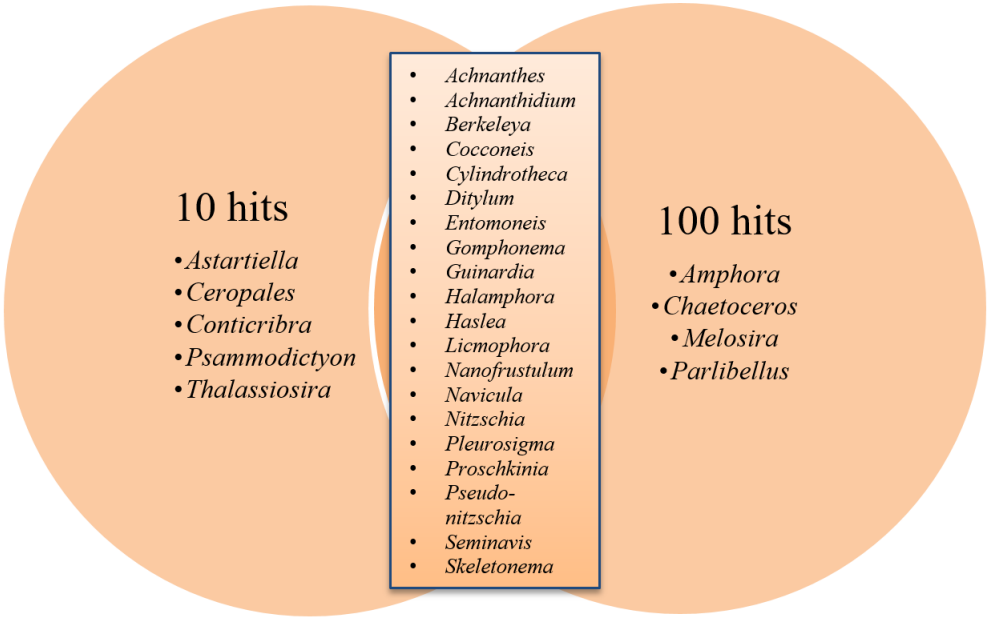

Figure 4.18. The Venn diagram demonstrates the distribution of the diatom genera identified in the GenBank dataset according to two different hit results.

4.3.1. Location-based Comparison of Diatom Genera for 10 and 100 hits

4.3.1.1. Erdek (M1). The Erdek location shows that *Amphora* and *Pseudo-nitzschia* were the diatom genera seen only in 100 hits analysis; however, *Astartiella*, *Berkeleya*, *Ceropales*, *Conticribra*, *Psammodictyon*, *Skeletonema*, and *Thalassiosira* have been observed only in 10 hits. The other 10 genera have been observed both at the 10 and 100 species hit lists (as also shown in Figure 4.19).

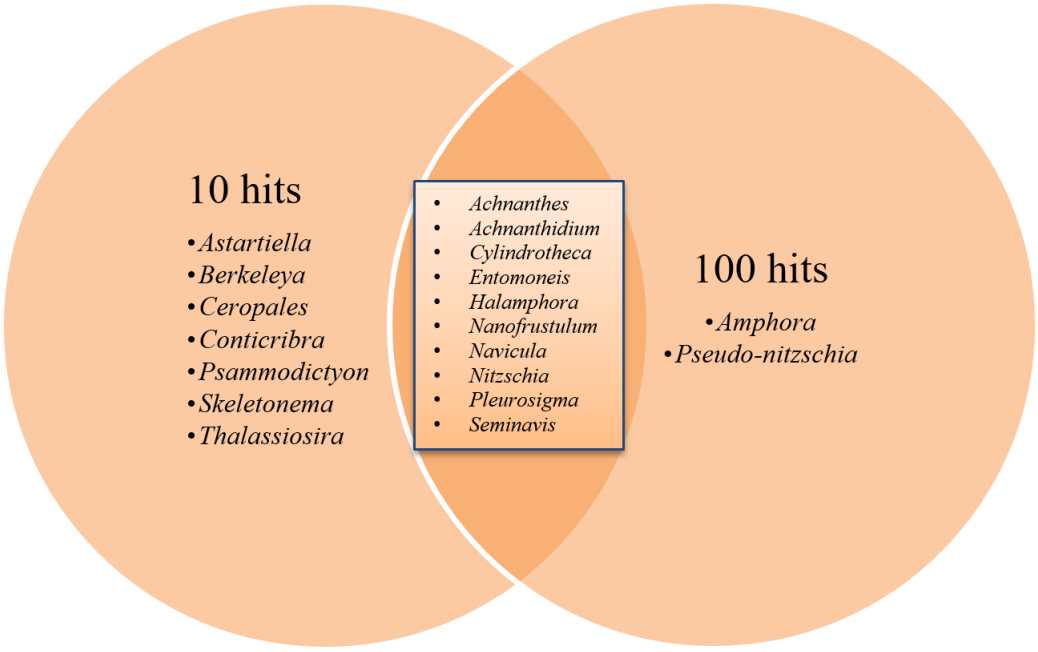

Figure 4.19. The Venn diagram demonstrates the distribution of the diatom genera for M1 station.

4.3.1.2. Kınalıada (M2). A quick glance at the Kınalıada location reveals that only the *Entomoneis* genus was identified in the 10 hits analysis. However, when examining the 100 species hits (as also shown in Figure 4.20), the diatom genera *Amphora*, *Chaetoceros*, *Ditylum*, *Parlibellus*, and *Pleurosigma* were observed.

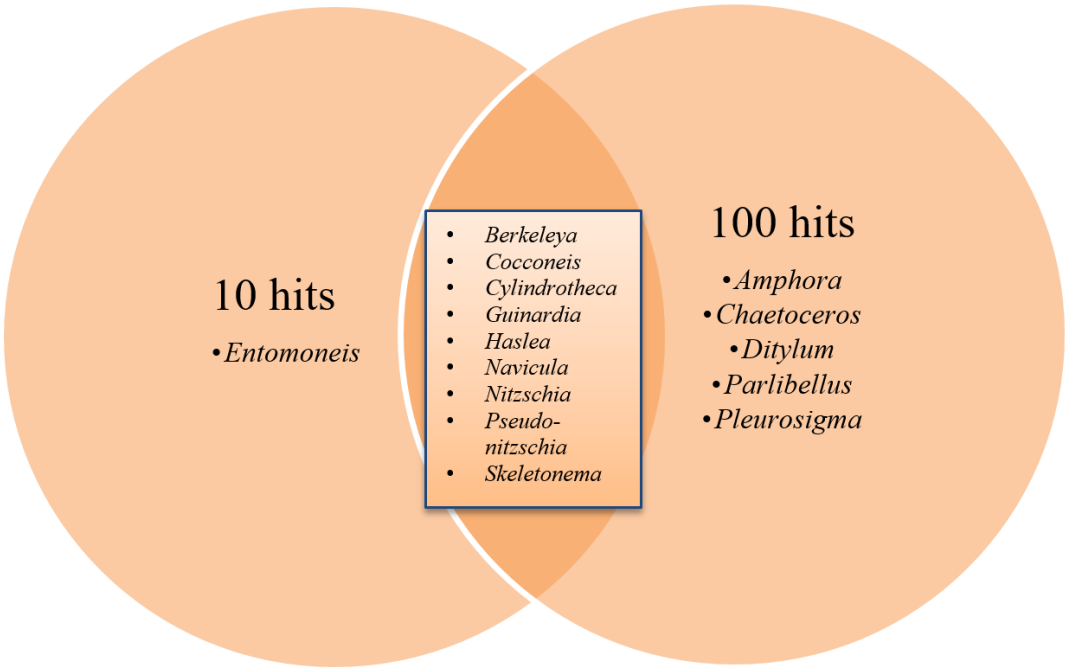

Figure 4.20. The Venn diagram demonstrates the distribution of the diatom genera for M2 station.

4.3.1.3. Gulf of Izmit (M3). *Chaetoceros* was the only diatom genus seen in 100 hits analysis for the Gulf of Izmit location. On the other hand, diatom genera *Achnanthes*, *Entomoneis*, *Skeletonema*, and *Thalassiosira* were identified for the 10 species hits. Other six genera were seen for both analyses, as shown in the figure below.

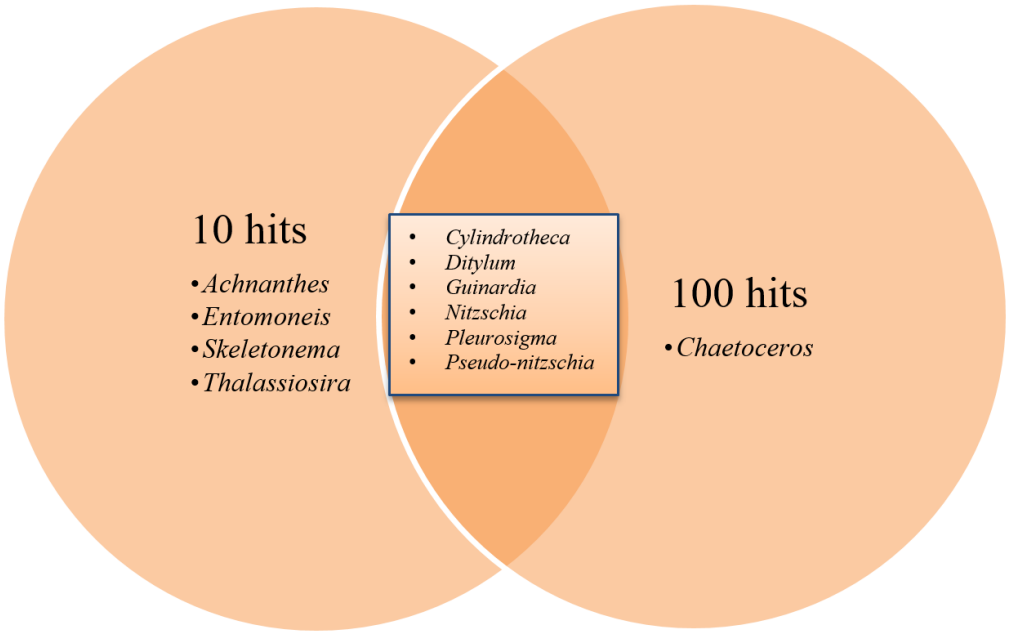

Figure 4.21. The Venn diagram demonstrates the distribution of the diatom genera for M3 station.

4.3.1.4. Yenikapı (M4). In the Yenikapı location, analyses of diatom genera revealed that *Conticribra*, *Entomoneis*, *Halamphora*, and *Thalassiosira* were present in 10 species hits results. However, no diatom genera were observed specifically for the 100 hits. Nine other genera were found in both the 10 and 100 species hits analyses (Figure 4.22).

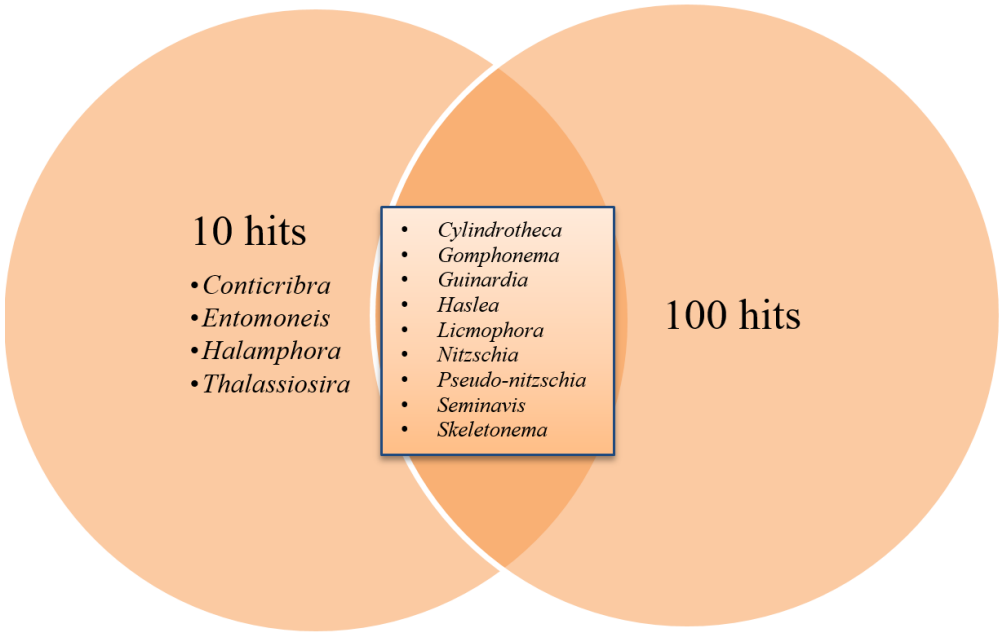

Figure 4.22. The Venn diagram demonstrates the distribution of the diatom genera for M4 station.

4.3.1.5. Çınarcık (M5). *Entomoneis* and *Skeletonema* were only observed in the 10 species hits results, while *Navicula* and *Parlibellus* were exclusively present in the 100 hits. The remaining six diatom genera were observed in both the 10 and 100 species hits analyses, as shown in Figure 4.23.

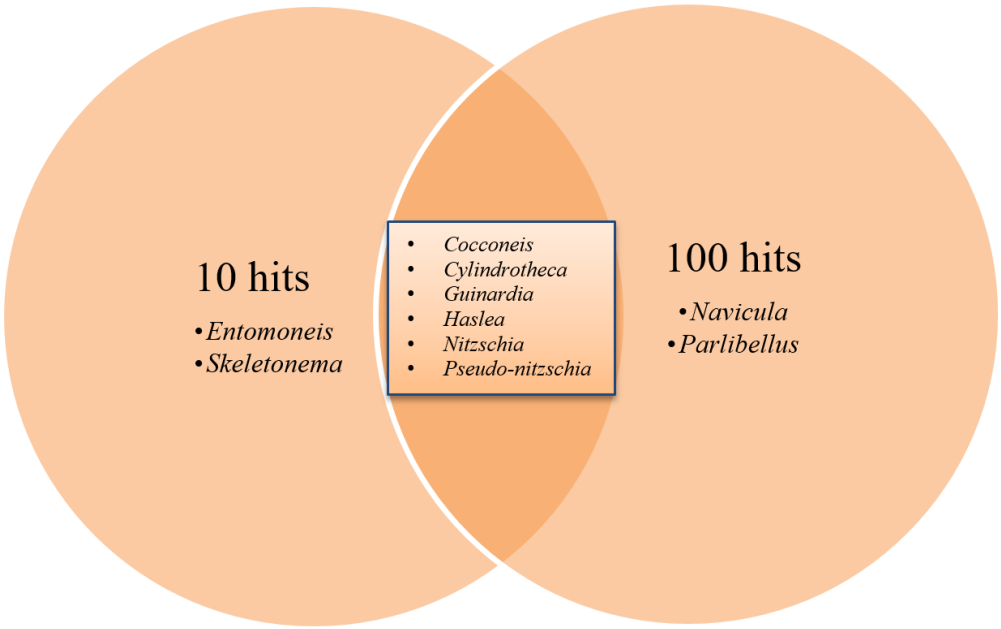

Figure 4.23. The Venn diagram demonstrates the distribution of the diatom genera for M5 station.

4.3.1.6. Kalamış (M6). Looking at the Kalamış results, *Astartiella*, *Berkeleya*, *Conticribra*, *Entomoneis*, and *Psammodictyon* were seen only for the 10 hits; however, *Amphora*, *Chaetoceros*, and *Melosira* were detected explicitly in 100 species hits analyses. The other 12 diatom genera were seen in both 10 and 100 hits, as shown in Figure 4.24.

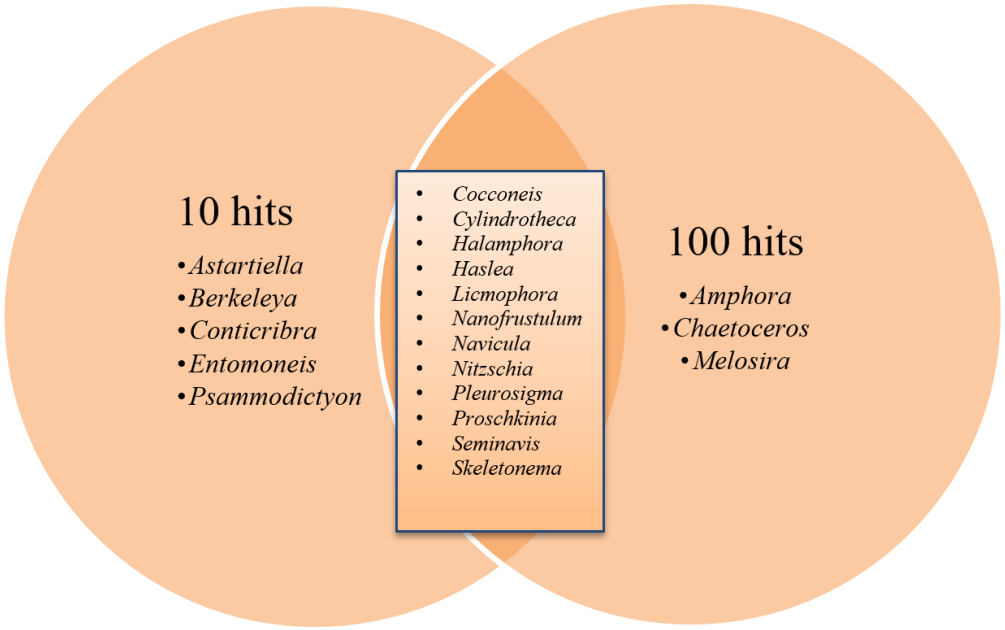

Figure 4.24. The Venn diagram demonstrates the distribution of the diatom genera for M6 station.

5. DISCUSSION

Within the context of the present thesis, beyond the general-level characterizations outlined above, species-level data were examined within our dataset, focusing on species deemed pertinent to mucilage production based on the existing literature. These species are *Phaeocystis pouchetii*, *Skeletonema costatum*, *Thalassiosira rotula*, *Cylindrotheca closterium*, and *Gonyaulax fragilis* in the surface mucilage (Balkis et al., 2009). Also at the genus level *Leptocylindrus spp.* stands out. No species belonging to the *Phaeocystis*, *Gonyaulax* and *Leptocylindrus* were found in the surface mucilage. Although it is not possible to determine at the species level, genus *Thalassiosira* was detected. The genus *Skeletonema* was found at all sampling points mentioned above for the 10 hits analysis, and it was found at Kınalıada, Yenikapı, and Kalamış for 100 hits analysis. The genus *Cylindrotheca* was found at all sampling points for both 10 and 100 species hits. When we go down from the genus level to the species level, the presence of *Skeletonema costatum* and *Cylindrotheca closterium* was confirmed at the species level. Specifically, *Skeletonema costatum* was detected in Kınalıadı, Yenikapı, and Kalamış and *Cylindrotheca closterium* was detected in Yenikapı and Kalamış.

Chrysoreinhardia giraudii and *Nematochrysopsis marina*, which are thought to be effective in the formation of the bottom mucilage, were not detected at the species and genus level in our study. However, species belonging to the Chrysophyceae family, to which these species belong, were found at the Yenikapı sampling point. Also, *Noctiluca scintillans*, *Spatulodinium pseudonoctiluca*, and *Mesodinium rubrum*, which are among the species that make up the “red tide”, or any of the species from these genera, were not found in the data set.

Diatoms have consistently been acknowledged as the predominant group of algae associated with environmental phenomena such as mucilage. This recognition stems from their substantial presence within mucilage aggregates (Rinaldi et al., 1995) and their well-documented ability to release extracellular polysaccharides (Ciglenecki et al., 2003, Pompei et al., 2003). Diatom species are more involved in mucilage formation than dinoflagellates because diatoms can grow and proliferate better within mucilage aggregates (Pompei et al., 2003; Tinti et al., 2007). It should also be emphasized that the phytoplankton communities responsible for mucilage may vary with different dominant species depending on the sampling area and period (Revelante and Gilmartin, 1991). *Cylindrotheca*, *Skeletonema*, *Pseudo-nitzschia* and *Thalassiosira* are diatom genera that have contributed to both the previous and current mucilage events in the Sea of Marmara (Aktan et al.,

2008; Tüfekçi et al., 2010; Balkis-Özdelice et al., 2021). It is known that these phytoplankton species synthesize carbohydrates (organic matter that leads to mucus formation) when they multiply excessively. Organic matter will likely accumulate in the environment, forming mucilage with the co-existence of one or more phytoplankton species.

The formation pattern of mucilage also differs from one phytoplankton to another. For instance, while *Skeletonema* species found in all sampling stations secrete exopolymeric compounds in the absence of bacteria, *Thalassiosira* found in Erdek Bay (M1), Izmit Bay (M3), and Yenikapi (M4) stations produce high amounts of exopolymeric compounds in the presence of heterotrophic bacteria (Grossart et al. 2006). Species of the genera *Cylindrotheca*, *Skeletonema*, and *Thalassiosira* have been reported to form large-scale pelagic mucilage in the Adriatic Sea (Mingazzini and Thake, 1995), while members of *Thalassiosira* have been reported to form filamentous mucilage in nutrient-rich regions (Margalef, 1978; Sournia, 1982). *Cylindrotheca* members are incredibly dominant in the spring-summer phytoplankton community and have been recorded among harmful microalgae species due to their overgrowth (Hasle and Syversten, 1996). *Skeletonema* species also form marine phytoplankton communities with nitrate reductase enzyme characteristics for diatom species, which damages the ecosystem with excessive reproduction (Stearn, 1973).

In addition, it is well known that members of the *Pseudo-nitzschia* genus are abundant in eutrophic coastal regions rich in salts (Parsons et al., 2002; Lundholm et al. 2004). In this study, this genus was detected in environmentally problematic areas such as Kınalıada (M2), Gulf of Izmit (M3), Yenikapı (M4) and Çınarcık (M5). There are also toxin-producing members (Lundholm et al. 2003) of *Pseudo-nitzschia* species, which are commonly observed in the phytoplankton of the Sea of Marmara (Balkis-Özdelice, 2021; Taş et al. 2009; Taş and Okus, 2011) and are involved in harmful algae formations (Türkoğlu and Koray, 2002).

It is known that atmospheric events also play a role in the formation of mucilage. Recently, the Sea of Marmara has been under the influence of Sahara dust, and high chl-a values have been detected (Balkis-Özdelice et al., 2021). It is known that desert dust reaches its maximum level in the Eastern Mediterranean in the spring months (Güllü et al., 1998) and contains aluminum (Al), iron (Fe), phosphorus (P), and lead (Pb) ions (Guieu et al., 2002). Fe ions play an essential role in the nitrogen and carbon cycle and cause an increase in phytoplankton, and this might have also contributed to the detection of this group in the mucilage samples collected during this study.

Several metabarcoding studies have been conducted to examine the impact of mucilage events on diatom communities in various marine environments. In Thessaloniki Bay, a recent study detected high abundances of *Cylindrotheca*, *Chaetoceros, Leptocylindrus*, and *Skeletonema* (Genitsaris et al., 2019). Similarly, during mucilage events in the Adriatic Sea in 1988 and 1999, Revelante and Gilmartin (1991) observed an increase in the abundance of certain diatom genera, including *Nitzschia* and *Cylindrotheca*, at levels higher than those found in the surrounding seawater. In the Büyükada region of the Sea of Marmara, another metabarcoding study by Balkis et al. (2009) detected *Cylindrotheca, Pseudo-nitzschia, Skeletonema*, and *Thalassiosira* as the most dominant diatom genera. A study conducted in the Tyrrhenian Sea during 1999-2002 found *Synedra, Licmophora*, and *Navicula* as the most frequently occurring diatom genera in that region during mucilage events (De Philippis et al., 2005). In situ observations and collections of diatom aggregates in three locations, including the North Sea (Riebesell, 1991, 1992), the Adriatic Sea, and Californian coastal waters, commonly identified *Nitzschia, Chaetoceros, Rhizosolenia, Leptocylindricus, Skeletonema*, and *Thalassionema* as the genera present in diatom aggregates (Thornton, 2002).

We found several diatom species such as *Cylindrotheca, Chaetoceros, Skeletonema, Nitzschia, Pseudo-nitzschia, Thalassiosira, Licmophora*, and *Navicula* were present in our dataset, which is consistent with the literature above. However, we did not observe *Leptocylindrus, Synedra*, and *Rhizosolenia* in our results. It's worth noting that we observed *Guinardia* as the dominant diatom genus with 100 species hits in our results. This genus has not been previously reported in any other study on mucilage events, indicating that it is a novel diatom genus for this most recent phenomenon from the Sea of Marmara that our samples are from.

6. CONCLUSIONS AND RECOMMENDATIONS

As a result of the study, it was determined that the mucilage diversity and composition of diatoms varies by locality in the Sea of Marmara. The Erdek location exhibited the presence of *Nitzschia*, *Skeletonema*, and *Achnanthes*, whereas *Cocconeis*, *Berkeleya*, *Guinardia*, and *Pseudo-nitzschia* were found to be the dominant genera at Kınalıada sampling point. *Guinardia* was identified as the most abundant genus at the Gulf of Izmit location, followed by the occurrence of *Skeletonema* and *Pseudo-nitzschia*. Yenikapı location had a predominance of *Guinardia* and *Skeletonema*. Similarly, Çınarcık showed the dominance of *Skeletonema* and *Pseudo-nitzschia*. Lastly, Kalamış exhibited a high abundance of *Skeletonema*. When the whole sample is examined, *Skeletonema* and *Nitzschia* diatom genera are the dominant organism groups for 10 species hits; on the other hand, *Guinardia* and *Pseudo-nitzschia* are the dominant ones for 100 species hits.

The study constructed dendrograms to visualize the similarities in genus diversity between PCR replicates and diatom diversity from each sampling location. PCR replicates 5.3 and 6.3 showed anomalous data, while PCR replicates 1.1, 1.2, and 1.3 had similar taxonomic profiles. Samples 3 and 5 showed a close relationship with each other due to their proximity to adjacent locations. The results from the diatom dendrogram indicated that proximity might play a role in shaping the composition of genera diversity observed in different locations, such as the Gulf of Izmit and Çınarcık, Kınalıada and Kalamış, and Erdek and Yenikapı locations displayed clustering as anticipated.

Mucilage has incorporated inorganic substances and other organic substances in the environment together with algae. It is understood that the mucilage structure is mostly produced under the dominance of diatoms. The main reasons for the excessive increase of diatoms include the wastewater of the industrial facilities being discharged to the sewer system without pre-treatment or illegally, the wastewater discharge to the aquatic environment without any treatment, and domestic wastewater, and nutrient salts carried to the environment by agricultural activities. It is also crucial to conduct research on the potential impact of toxic substances originating from industries located in the Sea of Marmara, particularly in the Gulf of Izmit. Such substances may alter the biodiversity composition and weaken the defense mechanism of the ecosystem, highlighting the need for further investigation in this area. These toxic substances can also cause mucilage formation due to the stress it creates on microorganisms.

The findings of this study are significant as they contribute to the growing body of research that employs the metabarcoding approach to identify algal diatoms present in mucilage. However, the results should be interpreted cautiously, as different individual studies differ in sampling time, location, and methodology. In addition, the lack of an adequate metabarcoding dataset for the local community on the Turkish coast is an important handicap that prevents comparison with mucilage and community structure of the local community and evaluating community dynamics and successors.

The Sea of Marmara has a dynamic structure, and pelagic ecosystems can change instantaneously depending on current systems and climate change. Likewise, knowledge on the feeding habits of the species that have just entered the system or the behaviors of the species that have replaced the species that have left the system is important for ecosystem health. For this reason, long-term monitoring should be done, preferable in a monthly manner at certain stations, and subsequently local databases containing biodiversity data should be created, and these databases should be kept up-to date with routine monitoring studies.

When the study findings were evaluated, it was seen that the formation of mucilage observed in the Sea of Marmara occurred with the synergistic effect of many physical, chemical, and biological factors. It was concluded that species such as *Cylindirotecha*, *Sketelemona*, *Guinardia*, and *Thalasoirasiras* might have a potential effect on this formation. The formation of mucilage accelerated with the increase in the number of pathogenic bacteria (Danovaro et al., 2009). The recording of these phytoplanktonic diatom species in the environment shows the possibility of mucilage in the Sea of Marmara in the future, as in the past, due to the deterioration of the balance in environmental conditions. For this reason, the life cycles of these species should be studied in detail as well.

REFERENCES

Aktan Y, Dede A, Ciftci PS (2008). Mucilage event associated with diatoms and dinoflagellates in the Sea of Marmara, Turkey. Harmful Algae News 36: 1–3.

Aylagas, E., Borja, Á., & Rodríguez-Ezpeleta, N. (2014). Environmental status assessment using DNA metabarcoding: Towards a genetics based marine biotic index (gAMBI). *PLoS ONE*, *9*(3). https://doi.org/10.1371/journal.pone.0090529

Balkis, N., Atabay, H., Türetgen, I., Albayrak, S., Balkis, H., & Tüfekçi, V. (2011). Role of single-celled organisms in mucilage formation on the shores of Bykada Island (the Marmara Sea). *Journal of the Marine Biological Association of the United Kingdom*, *91*(4). https://doi.org/10.1017/S0025315410000081

Balkis, N., Sivri, N., Fraim, N. L., Balcı, M., Durmuş, T., & Sukatar, A. (2013). Excessive growth of Cladophora laetevirens (Dillwyn) Kutzing and enteric bacteria in mats in the Southwestern Istanbul coast, Sea of Marmara. *IUFS Journal of Biology Research Article IUFS J Biol*, *72*(2).

Balkis-Ozdelice, N., Durmuş, T., Balcı, M. (2021). A Preliminary Study on the Intense Pelagic and Benthic Mucilage Phenomenon Observed in the Sea of Marmara. *International Journal of Environment and Geoinformatics*, *8*(4). https://doi.org/10.30897/ijegeo.954787

Bandelt, H. J., Forster, P., & Röhl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution*, *16*(1). https://doi.org/10.1093/oxfordjournals.molbev.a026036

Bohan, D. A., Vacher, C., Tamaddoni-Nezhad, A., Raybould, A., Dumbrell, A. J., & Woodward, G. (2017). Next-Generation Global Biomonitoring: Large-scale, Automated Reconstruction of Ecological Networks. In *Trends in Ecology and Evolution* (Vol. 32, Issue 7). https://doi.org/10.1016/j.tree.2017.03.001

Boyer, F., Mercier, C., Bonin, A., Le Bras, Y., Taberlet, P., & Coissac, E. (2016). obitools: A unix-inspired software package for DNA metabarcoding. *Molecular Ecology Resources*, *16*(1). https://doi.org/10.1111/1755-0998.12428

Ciglenečki, I., Plavšić, M., Vojvodić, V., Ćosović, B., Pepi, M., & Baldi, F. (2003). Mucopolysaccharide transformation by sulfide in diatom cultures and natural mucilage. *Marine Ecology Progress Series*, *263*. https://doi.org/10.3354/meps263017

Cordier, T., Forster, D., Dufresne, Y., Martins, C. I. M., Stoeck, T., & Pawlowski, J. (2018). Supervised machine learning outperforms taxonomy-based environmental DNA metabarcoding applied to biomonitoring. *Molecular Ecology Resources*, *18*(6). https://doi.org/10.1111/1755-0998.12926

Creer, S., Fonseca, V. G., Porazinska, D. L., Giblin-Davis, R. M., Sung, W., Power, D. M., Packer, M., Carvalho, G. R., Blaxter, M. L., Lambshead, P. J. D., & Thomas, W. K. (2010). Ultrasequencing of the meiofaunal biosphere: Practice, pitfalls and promises. *Molecular Ecology*, *19*(SUPPL. 1). https://doi.org/10.1111/j.1365-294X.2009.04473.x

Creer, S., Deiner, K., Frey, S., Porazinska, D., Taberlet, P., Thomas, W. K., Potter, C., & Bik, H. M. (2016). The ecologist's field guide to sequence-based identification of biodiversity. In *Methods in Ecology and Evolution* (Vol. 7, Issue 9). https://doi.org/10.1111/2041-210X.12574

Cruaud, P., Vigneron, A., Fradette, M. S., Charette, S. J., Rodriguez, M. J., Dorea, C. C., & Culley, A. I. (2017). Open the SterivexTM casing: An easy and effective way to improve DNA extraction yields. *Limnology and Oceanography: Methods*, *15*(12). https://doi.org/10.1002/lom3.10221

Danovaro, R., Umani, S. F., & Pusceddu, A. (2009). Climate change and the potential spreading of marine mucilage and microbial pathogens in the mediterranean sea. *PLoS ONE*, *4*(9). https://doi.org/10.1371/journal.pone.0007006

De Philippis, R., Faraloni, C., Sili, C., & Vincenzini, M. (2005). Populations of exopolysaccharide-producing cyanobacteria and diatoms in the mucilaginous benthic aggregates of the Tyrrhenian Sea (Tuscan Archipelago). *Science of the Total Environment*, *353*(1–3). https://doi.org/10.1016/j.scitotenv.2005.09.078

Deeds, J. R., Kibler, S. R., Tester, P. A., & Place, A. R. (2004). Geographic strain variation intoxin production in *Karlodinium micrum* (Dinophyceae) from Southeastern estuaries of the United States.

In *Harmful Algae 2002. Florida Fish and Wildlife Conservation Commis-sion, Florida Institute of Oceanography, and Intergovernmental OceanographicCommission of UNESCO*.

Fukao, T., Kimoto, K., Yamatogi, T., Yamamoto, K. I., Yoshida, Y., & Kotani, Y. (2009). Marine mucilage in Ariake Sound, Japan, is composed of transparent exopolymer particles produced by the diatom Coscinodiscus granii. *Fisheries Science*, *75*(4). https://doi.org/10.1007/s12562-009-0122-0

Genitsaris, S., Stefanidou, N., Sommer, U., & Moustaka-Gouni, M. (2019). Phytoplankton blooms, red tides and mucilaginous aggregates in the urban Thessaloniki Bay, Eastern Mediterranean. *Diversity*, *11*(8). https://doi.org/10.3390/d11080136

Giani, M., Berto, D., Zangrando, V., Castelli, S., Sist, P., & Urbani, R. (2005). Chemical characterization of different typologies of mucilaginous aggregates in the Northern Adriatic Sea. *Science of the Total Environment*, *353*(1–3). https://doi.org/10.1016/j.scitotenv.2005.09.027

Giani, M., Rinaldi, A., & Degobbis, D. (2005). Mucilages in the Adriatic and Tyrrhenian Sea: An introduction. In *Science of the Total Environment* (Vol. 353, Issues 1–3). https://doi.org/10.1016/j.scitotenv.2005.09.006

Grossart, H. P., Kiørboe, T., Tang, K. W., Allgaier, M., Yam, E. M., & Ploug, H. (2006). Interactions between marine snow and heterotrophic bacteria: Aggregate formation and microbial dynamics. *Aquatic Microbial Ecology*, *42*(1). https://doi.org/10.3354/ame042019

Guieu, C., Loÿe-Pilot, M. D., Ridame, C., & Thomas, C. (2002). Chemical characterization of the Saharan dust end-member: Some biogeochemical implications for the western Mediterranean Sea. *Journal of Geophysical Research Atmospheres*, *107*(15). https://doi.org/10.1029/2001JD000582

Güllü, G. H., Ölmez, I., Aygün, S., & Tuncel, G. (1998). Atmospheric trace element concentrations over the eastern Mediterranean Sea: Factors affecting temporal variability. *Journal of Geophysical Research Atmospheres*, *103*(D17). https://doi.org/10.1029/98JD01358

Hasle, G. R. and Syversten, E. E. (1996). Marine Diatoms. In: Tomas, C. R. (ed.) Identifying Marine Phytoplankton. Academic Press, Inc. San Diego, Calif. pages 292–294.

Hurriyetdailynews.com. (2021, June 5). *Mucilage may kill Black Sea: Expert - Türkiye news*. Hürriyet Daily News. https://www.hurriyetdailynews.com/mucilage-may-kill-black-sea-expert-165273

Innamorati, M., Nuccio, C., Massi, L., Mori, G., & Melley, A. (2001). Mucilages and climatic changes in the Tyrrhenian Sea. *Aquatic Conservation: Marine and Freshwater Ecosystems*, *11*(4). https://doi.org/10.1002/aqc.448

İşinibilir-Okyar, M., Üstün, F., & Orun, D. A. (2015). Changes in abundance and community structure of the zooplankton population during the 2008 mucilage event in the northeastern Marmara Sea. *Turkish Journal of Zoology*, *39*(1). https://doi.org/10.3906/zoo-1308-11

Leray, M., Yang, J. Y., Meyer, C. P., Mills, S. C., Agudelo, N., Ranwez, V., Boehm, J. T., & Machida, R. J. (2013). A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: Application for characterizing coral reef fish gut contents. *Frontiers in Zoology*, *10*(1). https://doi.org/10.1186/1742-9994-10-34

Learnz, (2021). *Threats to Marine Ecosystems*, https://www.learnz.org.nz/sustainableseas181/bg-standard-f/threats-to-marine-ecosystems

Liu, Z., Desantis, T. Z., Andersen, G. L., & Knight, R. (2008). Accurate taxonomy assignments from 16S rRNA sequences produced by highly parallel pyrosequencers. *Nucleic Acids Research*, *36*(18). https://doi.org/10.1093/nar/gkn491

Lundholm, N., Hansen, P. J., & Kotaki, Y. (2004). Effect of pH on growth and domoic acid production by potentially toxic diatoms of the genera Pseudo-nitzschia and Nitzschia. *Marine Ecology Progress Series*, *273*. https://doi.org/10.3354/meps273001

Lundholm, N., Moestrup, Ø., Hasle, G. R., & Hoef-Emden, K. (2003). A study of the Pseudo-nitzschia pseudodelicatissima/cuspidata complex (Bacillariophyceae): What is P. pseudodelicatissima? *Journal of Phycology*, *39*(4). https://doi.org/10.1046/j.1529-8817.2003.02031.x

Margalef, R. (1978). Phytoplankton communities in upwelling areas. The example of NW Africa. *Oecologia Aquatica*, *3*(3).

Marmara Denizi'nde Müsilaj Problemi (2021). *Çevre Koruma ve Kontrol Dairesi Başkanlığı*. İstanbul Büyükşehir Belediyesi | Çevre İBB. https://cevre.ibb.istanbul/deniz-hizmetleri-mudurlugu-sube-mudurlugu/marmara-denizinde-musilaj-problemi/

Mecozzi, M., Acquistucci, R., Di Noto, V., Pietrantonio, E., Amici, M., & Cardarilli, D. (2001). Characterization of mucilage aggregates in Adriatic and Tyrrhenian Sea: Structure similarities between mucilage samples and the insoluble fractions of marine humic substance. *Chemosphere*, *44*(4). https://doi.org/10.1016/S0045-6535(00)00375-1

Mingazzini, M., & Thake, B. (1995). Summary and conclusions of the workshop on marine mucilages in the Adriatic Sea and elsewhere. *Science of the Total Environment*, *165*(1–3). https://doi.org/10.1016/0048-9697(95)04538-C

Özalp, H. B. (2021). First massive mucilage event observed in deep waters of Çanakkale Strait (Dardanelles), Turkey. In *J. Black Sea/Mediterranean Environment* (Vol. 27, Issue 1).

Öztürk, İ., Yanalak, M., Arslan, Ö., Koyuncu, İ., Dilekgürgen, E., Erşahin, M. E., & Türken, T. (2021). Marmara Denizi'nde Deniz Salyası Sorunu ile İlgili Görüş ve Öneriler. In *İstanbul Teknik Üniversitesi*.

Parsons, M. L., & Dortch, Q. (2002). Sedimentological evidence of an increase in Pseudo-nitzschia (Bacillariophyceae) abundance in response to coastal eutrophication. *Limnology and Oceanography*, *47*(2). https://doi.org/10.4319/lo.2002.47.2.0551

Pompei, M., Mazziotti, C., Guerrini, F., Cangini, M., Pigozzi, S., Benzi, M., Palamidesi, S., Boni, L., & Pistocchi, R. (2003). Correlation between the presence of Gonyaulax fragilis (Dinophyceae) and the mucilage phenomena of the Emilia-Romagna coast (northern Adriatic Sea). *Harmful Algae*, *2*(4). https://doi.org/10.1016/S1568-9883(03)00059-3

Prosser, S. W. J., Velarde-Aguilar, M. G., León-Règagnon, V., & Hebert, P. D. N. (2013). Advancing nematode barcoding: A primer cocktail for the cytochrome c oxidase subunit I gene from vertebrate parasitic nematodes. *Molecular Ecology Resources*, *13*(6). https://doi.org/10.1111/1755-0998.12082

Revelante, N., & Gilmartin, M. (1991). The phytoplankton composition and population enrichment in gelatinous “macroaggregates” in the northern Adriatic during the summer of 1989. *Journal of Experimental Marine Biology and Ecology*, *146*(2). https://doi.org/10.1016/0022-0981(91)90027-T

Saydam, A. C. (2014). Desert Dust Cloud Interactions and Natural Iron Enrichment Mechanism. *International Journal of Environment and Geoinformatics*, *1*(1). https://doi.org/10.30897/ijegeo.300712

Sournia, A. (1982). Form and function in marine phytoplankton. *Biological Reviews*, *57*(3). https://doi.org/10.1111/j.1469-185X.1982.tb00702.x

Taberlet, P., Bonin, A., Zinger, L., & Coissac, E. (2018). Chapter 6. DNA amplification and multiplexing. *Environmental DNA: For Biodiversity Research and Monitoring*.

Taberlet, P., Coissac, E., Pompanon, F., Brochmann, C., & Willerslev, E. (2012). Towards next-generation biodiversity assessment using DNA metabarcoding. *Molecular Ecology*, *21*(8). https://doi.org/10.1111/j.1365-294X.2012.05470.x

Tas, S., Yilmaz, I. N., & Okus, E. (2009). Phytoplankton as an indicator of improving water quality in the Golden Horn Estuary. *Estuaries and Coasts*, *32*(6). https://doi.org/10.1007/s12237-009-9207-3

Taş, S., & Okuş, E. (2011). A review on the bloom dynamics of a harmful dinoflagellate Prorocentrum minimum in the Golden Horn Estuary. In *Turkish Journal of Fisheries and Aquatic Sciences* (Vol. 11, Issue 4). https://doi.org/10.4194/1303-2712-v11_4_03

Thomsen, P. F., & Willerslev, E. (2015). Environmental DNA - An emerging tool in conservation for monitoring past and present biodiversity. In *Biological Conservation* (Vol. 183). https://doi.org/10.1016/j.biocon.2014.11.019

Thornton, D. C. O. (2002). Diatom aggregation in the sea: Mechanisms and ecological implications. In *European Journal of Phycology* (Vol. 37, Issue 2). https://doi.org/10.1017/S0967026202003657

TMMOB (2015). Çevre Mühendisleri Odası İstanbul Şubesi, İstanbul Atıksu Arıtma Tesisleri ve Kirleticilerin Değerlendirilmesi.

Toklu-Alicli, B., Polat, S., & Balkis-Ozdelice, N. (2020). Temporal variations in the abundance of picoplanktonic Synechococcus (Cyanobacteria) during a mucilage event in the Gulfs of Bandırma and Erdek. *Estuarine, Coastal and Shelf Science*, *233*. https://doi.org/10.1016/j.ecss.2019.106513

Tüfekçi, V., Balkis, N., Polat Beken, Ç., Ediger, D., & Mantikçi, M. (2010). Phytoplankton composition and environmental conditions of a mucilage event in the Sea of Marmara. *Turkish Journal of Biology*, *34*(2). https://doi.org/10.3906/biy-0812-1

Türkoğlu, M., & Koray, T. (2002). Phytoplankton species' succession and nutrients in the Southern Black Sea (Bay of Sinop). *Turkish Journal of Botany*, *26*(4).