

https://dspace.org/

Contributions to the molecular approach of using diatoms in the monitoring of coastal waters

Kalyan, Brishan

2022
https://hdl.handle.net/10539/34490

Downloaded from a DSpace repository. https://dspace.org/



Contributions to the molecular approach of using diatoms in the monitoring of coastal waters

by

Brishan Kalyan

(1312380)

Submitted in fulfilment of the requirements for the degree

Masters of Science

in

Animal, Plant and Environmental Sciences

at the Faculty of Science, University of the Witwatersrand, Johannesburg, South Africa.

Supervisor: Dr. Gavin Snow

Co-supervisor: Dr. Thomas Bornman

07 June 2022

Declaration

I declare that this dissertation is my own, unaided work. It has been submitted in fulfilment of the requirements of a Master of Science at the University of the Witwatersrand. It has not been submitted before for any degree or examination to another university or similar institution.

Brishan Kalyan



07 June 2022

Acknowledgements

I would like to express my gratitude to my supervisors Dr Gavin Snow and Prof Thomas Bornman, as without their support the project would not have been completed. Laboratory and research support by Prof Stuart Sym was invaluable to the completion of this project. Dr Roksana Majewska for her contributions of diatom monocultures and sound advice during this project. The advice and conversations with Phumlile Cotiyane-Pondo and Dr Gwynneth Matcher aided in the development of morphological and molecular analysis for the building and recommendations of a diatom barcode reference library. Prof Alexander Ziegler and Jacques Gerber at the Microscope and Microanalysis Unit (University of the Witwatersrand), William Goosen for assistance with the scanning electron microscopy at the Centre for High Resolution Electron Microscopy (CHRTEM), Nelson Mandela University and Donald McCallum from the University of the Witwatersrand's Herbarium for the assistance with and use of the Zeiss Axio Imager 2 and the Phenom Pure G6 Desktop SEM. Khanyisile Shabangu at the African Centre of DNA Barcoding for constant support and communication regarding DNA sequencing analysis.

Table of contents

Abbreviations 6

List of figures and tables 7

Chapter 1. Introduction 9

1.1. Monitoring of coastal waters 10

1.2. Planktonic diatoms as bioindicators 12

1.3. Algoa Bay 14

1.4. Project aims and objectives 15

Chapter 2. Methods and material 16

2.1. Study sites 16

2.2. Sample collection 16

2.3. Isolation and monoculture protocols 17

2.4. Morphological identification 18

2.5. Molecular identification 19

2.5.1. DNA extraction and quantification 19

2.5.2. PCR amplification 20

2.5.3. DNA sequencing and analysis 21

2.6. Barcode of Life Data Systems (BOLD) 22

Chapter 3. Results 23

3.1. Sample collection 23

3.2. Morphological identification - light and scanning electron microscopy 26

3.3. Molecular identification 31

3.4. Barcode of Life Data Systems submission and analysis 42

Chapter 4. Discussion 44

4.1. Species significance 44

4.2. Morphology and molecular identification workflow 46

4.3. Environmental monitoring 48

4.4. Recommendations and future steps 49
4.5. Conclusions 49
References 50
Supplementary material 1 63
Supplementary material 2 82

Abbreviations

AB – Algoa Bay

ACDB – African Centre of DNA Barcoding

BLAST – Basic Local Alignment Search Tool

BOLD – Barcode of Life Data Systems

BSA – Bovine Serum Albumin

CHRTEM - Centre for High Resolution Electron Microscopy

CTAB – hexadecyl-trimethyl-ammonium bromide

DNA – Deoxyribonucleic acid

eDNA – environmental deoxyribonucleic acid

LTER – Long-Term Ecological Research

MMU – Microscope and Microanalysis Unit

NCBI – National Centre for Biotechnology Information

NWU – North – West University

PCR – Polymerase Chain Reaction

rbcL – Ribulose-1,5-bisphosphate carboxylase/oxygenase plastid gene

RNA – Ribonucleic acid

rRNA – ribosomal ribonucleic acid

SAM – Strategic Adaptive Management

SAEON – South African Environmental Observation Network

SEM – Scanning Electron Microscope

%HQ – percentage of high-quality bases in an untrimmed generated consensus sequence.

List of figures and tables

Figure 1.1. Workflow for morphological- and molecular-based environmental DNA (eDNA) identification techniques for planktonic marine diatoms. 11

Figure 2.1. Map of Algoa Bay indicating the Pelagic Ecosystem LTER stations (modified from Dorrington *et al*., 2018). Sampling stations utilised in this study were Cape Recife, Bay Central and Bird Island (indicated by circles). 16

Table 2.1. PCR primer sequences and annealing temperatures used for the amplification of monocultured diatom DNA samples (Zimmermann et al., 2015; Vasselon et al., 2017). 20

Table 3.1. Planktonic samples collected from Algoa Bay include sites within the bay (Bay Central (BC) (25.98722 E, -33.881778 S); Cape Recife (CR) (25.72771 E, -34.034115 S); Bird Island (BI) (26.29146 E, -33.868695 S)), using different towing methods (Horizontal (H) and Vertical (V)). Epizoic and epiphytic samples obtained from Dr. Majewska (NWU) were collected as scrapes from different organisms and at different times in South Africa. Identification of sample monocultures occurred upon analysis of morphology and molecular data. * Indicates samples not used for analysis. 24

Table 3.2. SEM images using the Vega TESCAN and JOEL - JSM 7001F SEMs, operated by technicians at MMU and NMU respectively. These were images of July 2021 cultures using the hydrochloric acid and potassium permanganate digestion method (Harding and Taylor, 2011). The phenom pure G6 desktop SEM was used to take images of September 2021 cultures digested by the nitric acid and potassium dichromate digestion method (Elias et al., 2017; Cotiyane-Pondo and Bornman, 2021). 27

Table 3.3. Sequencing results from July 2021, showing the number of forward and reverse sequences that were used to generate consensus sequences (barcode) of each monocultured species. The number of ambiguities, sequence length and the percentage of bases that were high quality in the consensus sequence (%HQ) were recorded. Sample IDs highlighted were used for comparing to the NCBI nucleotide database via the blastn algorithm. Samples marked with * were not used in further analysis. 32

Table 3.4. Sequencing results from September 2021, showing the number of forward and reverse sequences that were used to generate consensus sequences (barcode) of each monocultured species. The number of ambiguities, sequence length and the percentage of bases that were high quality in the consensus sequence (%HQ) were recorded. Sample IDs highlighted were used for comparing to the NCBI nucleotide database via the blastn algorithm and samples marked with * were not used in further analysis 33

Table 3.5. Blast results identifying genera or species of July 2021 and September 2021 sequencing. Images of associated cultures are labelled according to Sample ID in Table 3.2. * Indicates morphological identification was used to identify monoculture as molecular data were inconclusive. # Indicates that the samples morphological identification (Table 3.2) matched the blast algorithms output. $ Indicates samples were only identified to genera. LQ Indicates sequences that were re-analysed at a 10% error per base and then consensus sequences used for blast analysis and molecular species identification. 35

Table 3.6. Samples re-analysed at a 10% error per base and blasted to the nucleotide database. Blast results from October 2021. * Indicates morphological identification was used to identify monoculture as molecular data were inconclusive. # Indicates that the samples morphological identification (Table 3.2) matched the blast algorithms output.. 38

Figure 3.1. Genetic distance matrix created by the Geneious tree builder using a pairwise alignment for all sequence pairs. The alignment type was a global alignment with free end gaps using a 93% similarity to build the distance matrix. Similar sequences between monocultures displayed as dark-colored blocks and dissimilar sequences as light-colored blocks. Values in the table written in white signify close relationships between monocultures. 40

Figure 3.2. Unrooted neighbour-joining tree created in January 2022 using the Geneious tree builder to infer relatedness of the diatom monocultures. Branches are labelled with sample name, species or genus identification, sequence event and if it is low quality (LQ). Colour codes indicate the following: dark blue – sequences mined from GenBank to infer relatedness, green – monoculture specimen identified to genera, orange – monoculture specimen identified morphology and molecular data and turquoise – low quality molecular information that provides good quality results. The values on the branches signify the number of substitutes per site..... 41

Table 3.7 Monocultures species identification by morphology and molecular approaches and conclusive species identification when possible. Monocultures with good quality images and barcodes had data published and accession numbers obtained. 43

Chapter 1. Introduction

In the face of global environmental change, one of the greatest challenges that scientists face is the limited understanding of the complex interactions between biotic diversity and a changing abiotic world (Berry *et al*., 2019). Globally, oceanic environments provide essential ecosystem services for the world's population, with organisms like phytoplankton providing approximately 50% of the oxygen requirement for all living organism on the planet (Vallaeys *et al*., 2017; Busseni, 2019). The ecosystem services that the ocean provides may include global weather patterns, the provision of food through fishing and ocean aquaculture, a source of novel pharmaceutical products, energetic and mining resources (Ingole *et al*., 2005; Kato *et al*., 2011; Mohanty *et al*., 2015; Momber *et al*., 2016; Vallaeys *et al*., 2017). Alterations in the oceanic environment, due to an ever-increasing human population and their diversifying needs for resources, means that the ocean is under tremendous growing pressure. Multiple human activities lead to pollution in the ocean, changes in local and global ocean pH, and global warming. Ocean pollution occurs in the form of the introduction of heavy metals, plastics, crude oils, sewage, nutrients, terrigenous materials, organic chemicals, toxins and pathogenic species (Knap *et al*., 2002; Vallaeys *et al*., 2017; Häder *et al*., 2020; Kalenik, 2020).

Studies have shown that changes in oceanic temperatures due to climate change alter species diversity and abundances of many organisms in the ocean. These changes may result in alterations in carbon sequestration and changes in the functional diversity of food webs within a given area (Laakmann *et al*., 2013; Benedetti *et al*., 2019; Berry *et al*., 2019; Brun *et al*., 2019; Hop *et al*., 2019; Nunes *et al*., 2019; Selmeczy *et al*., 2018; Dalu *et al*., 2020).

In order to manage a coastal zone effectively, monitoring is required to fit within a management model that is similar to the Strategic Adaptive Management (SAM) approach used in terrestrial protected areas such as the Kruger National Park, South Africa, i.e. monitoring is essential to ensure that management interventions are being implemented effectively and environmental thresholds are determined accurately (Gaylard and Ferreira, 2011; Kingsford *et al*., 2011; Swanepoel and Sauka, 2019). Much like SAM, the *National Environmental Management: Integrated Coastal Management Act* (24 of 2008) (RSA, 2009) stipulates that; coastal areas are to be managed in an integrative manner. With ocean environments recognised as a vital part of the global economy; policies and regulations have been implemented on a

global scale (Mohanty *et al*., 2015). In South Africa, the *Marine Spatial Planning Act* (16 of 2018) (RSA, 2019) outlines that the ocean provides a vast number of economic opportunities for the country and are being used increasingly more today than in the past. This means that the ocean is also subjected to environmental change, thus, the need has arisen to manage the ocean environment for it to sustainably provide for the local and global environments, society and economy.

A marine spatial plan is required to ensure the sustainable use of marine resources within a given area, allowing it to provide for various socio-economic needs according to all relevant legislative frameworks (Dorrington *et al*., 2018). The maintenance of a given marine area is vital for the provision of all its resources and ecosystem services, thus, understanding the given area from an environmental perspective (physico-chemical and biological factors) is required. This understanding will provide thresholds of potential concern for the given area, which would need to be monitored to ensure that the area is not drastically changed whilst providing the various environmental and socio-economic resources. These thresholds are physico-chemical and biological variables that are identified, specific to the environment, to provide an indication of when and if the environment is changing and what these changes may have on the ecosystem services the environment provides. The monitoring of these thresholds is vital for management as it will provide an indication of what management interventions are needed, whether management measures are effective or not, and if the thresholds need to be adjusted.

1.1. Monitoring of coastal waters

Traditionally, physico-chemical analyses to determine the water quality of a given environment were used to determine changes; however, the cost and difficulty of carrying out chemical analyses, intricacies of interpreting the data and determining the severity of impacts on the receiving environment make it more beneficial to use biomonitoring methods (De la Rey *et al*., 2004). Biomonitoring is the use of biotic communities and different species within an environment (bioindicators) to provide information on how the environment is changing and what may be causing the change through time and space. Biomonitoring allows for efficient, cost-effective and time-integrated techniques of detecting change to abiotic factors within a given environment such as pH, temperature, salinity, turbidity, nutrients, toxin concentrations and the possible identification of introduced pollutants and heavy metals (Chapman and Wang, 2000; De la Rey *et al*., 2004; Vallaeys *et al*., 2017). Therefore, bioindicators

can provide information of changes to the environment in a manner that is easily interpretable by managers, thus, allowing them to implement ways to mitigate these impacts and determine or evaluate the thresholds of the environment being managed (Johnson *et al*., 2011; Dorrington *et al*., 2018; Swanepoel and Sauka, 2019).

Historically, the identification of biotic communities occurred using morphological-based techniques (Figure 1.1), reliant on microscopy for taxonomic identification (Harding and Taylor, 2011; Mohrbeck *et al*., 2015). This approach is limited by the time it consumes, the expertise needed to identify different species, the inability to distinguish between larval stages of certain species and the occurrence of dead or injured individuals within the environmental sample (Laakmann *et al*., 2013; Mohrbeck *et al*., 2015; Zimmermann *et al*., 2015; Berry *et al*., 2019).

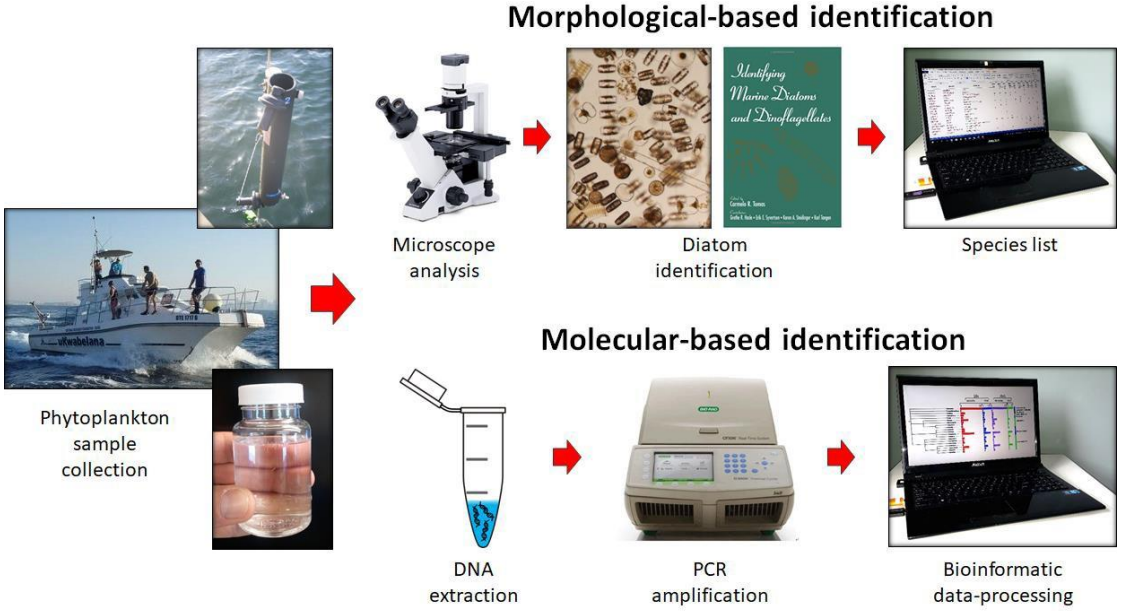


Figure 1.1. Workflow for morphological- and molecular-based environmental DNA (eDNA) identification techniques for planktonic marine diatoms.

Molecular techniques provide an efficient alternative to morphological-based identification of species (Figure 1.1). Environmental samples can be collected and analysed using molecular approaches such as metabarcoding, which is the simultaneous identification of multiple species in an environmental sample (Kermarrec et al., 2013; Zimmermann et al., 2015; Vasselon et al., 2017; Berry et al., 2019; Ruppert *et al*., 2019). In contrast, barcoding identifies just one species from its deoxyribonucleic acid (DNA) or ribonucleic acid (RNA). DNA barcodes are regions of DNA that can be used to identify an unknown specimen in terms of known classification (Hebert *et al*., 2003 and 2003b; Schindel and Miller, 2005; Bucklin *et al*., 2010). In the case of diatoms, individual species can be identified based on specific

gene fragments such as ribulose-1,5-bisphosphate carboxylase/oxygenase plastid gene (*rbcL*) and the 18S V4 region of ribosomal ribonucleic acid (rRNA) (Hebert *et al*., 2003a and 2003b; Zimmermann *et al*., 2011 and 2015; Vasselon *et al*., 2017; Berry *et al*., 2019; Ruppert *et al*., 2019).

The molecular approach to identify species may be more efficient, cost effective and more accurate than the morphological-based approach (Kermarrec *et al*., 2013; Mohrbeck *et al*., 2015; Zimmermann *et al*., 2015; Stat *et al*., 2017; Vasselon *et al*., 2017; Berry *et al*., 2019), however, the use of environmental DNA (eDNA) requires a comprehensive reference database to identify an organism to species level. Thus, a barcode library of possible species within a given environmental sample is needed for accurate processing of eDNA samples. The advancements in sequencing technologies and genetics have led to consolidated databases of processed sequences of multiple organisms, however, processed diatom sequence databases are still in their infancy and are the focus of this study.

1.2. Planktonic diatoms as bioindicators

Phytoplankton are a diverse group of microscopic organisms that occur within a water column (marine, estuarine or fresh) (Cholnoky, 1960; Margalef, 1978; Busseni, 2019; Dutkiewicz *et al*., 2020). Focusing on marine phytoplankton, they move great distances due to oceanic currents and are predominately found within the photic zone of the ocean environment. They are vital to regulating ocean biogeochemistry by means of carbon sequestration through the export of organic matter to the ocean deep thus contributing to alterations in atmospheric CO_2 levels, overall global climate and the provision of oxygen through the process of photosynthesis (Falkowski *et al*., 1998; Fuhrman, 2009; Guidi *et al*., 2009; Cermeño *et al*., 2013; Mbambo, 2014; Busseni, 2019; Dutkiewicz *et al*., 2020). The diversity of phytoplankton species within a given environment promotes the ecosystem's functional and structural stabilities thereby changing the ecosystem services it may provide (McCann, 2000; Ptacnik *et al*., 2008; Cermeño *et al*., 2016; Busseni, 2019; Dutkiewicz *et al*., 2020).

The understanding of phytoplankton community structure and abundance will allow one to infer changes in the environment (Verheye *et al*., 2017; Berry *et al*., 2019; Dalu *et al*., 2020). Changes in the marine environment through global environmental change alter the composition of phytoplanktonic communities resulting in a knock-on effect on higher trophic levels due to their dependence on phytoplankton for food, thus disrupting food webs (McCann, 2000; Ptacnik *et al*., 2008; Cermeño *et al*., 2016;

Verheye *et al*., 2017; Berry *et al*., 2019). This causes larger adverse effects such as limiting the ecological and socio-economic resources that the marine environment can provide.

Diatoms frequently dominate phytoplankton communities and are known to respond rapidly to environmental changes making them ideal bioindicators (De la Rey *et al*., 2004; Morin *et al*., 2008; Verheye *et al*., 2017; Dalu *et al*., 2020). Diatoms react to changes in water quality (De la Rey *et al*., 2004), are sensitive to pollutants (De la Rey *et al*., 2004), have a wide geographical range (Margalef, 1978), are easily sampled in a cost-effective manner (De la Rey *et al*., 2004; Harding and Taylor, 2011; Rimet *et al*., 2018) and can be used to differentiate between anthropogenic and natural stressors (Vallaeys *et al*., 2017). Diatoms have been used to monitor rivers across the world with a variety of morphological and molecular diatom identifications existing in many countries (Harding and Taylor*,* 2011; Zimmermann *et al*., 2015; Vasselon *et al*., 2017; Rimet *et al*., 2018 and 2019). Many diatom indices exist for freshwater purposes to infer environmental conditions; however, no such indices exist for the marine environment. With diatoms being the most diverse microalgal taxa, their use as a bioindicator is vital in monitoring our oceans. Although the morphology of diatoms is well established across the world (Reimann and Lewin 1964; Giffen 1971, 1973,1976; Hasle *et al*., 1996; Sarno *et al*., 2005; Stachura-Suchoples *et al*. 2015; Stock *et al*., 2019; López-Fuerte *et al.,* 2020), there is a lack of molecular information of diatom sequences (Rimet *et al*., 2019; Tapolczai *et al*., 2019). Key diatom species or communities need to be identified as bioindicators thereafter, accurate identification to species-level through morphological features and molecular data will result in their accurate use in eDNA biomonitoring surveys.

With eDNA becoming the forefront of environmental monitoring the use of metabarcoding requires a comprehensive reference library of DNA barcodes within a given environmental sample to provide accurate and unbiased results (Bode et al., 2017; Ruppert et al., 2019; Tapolczai et al., 2019). Studies using metabarcoding as a basis of environmental assessments conclusively show similar results to morphological techniques. However, there are discrepancies indicating a need for an integrated approach of morphological and molecular identification as well as a comprehensive reference library of DNA barcodes. These discrepancies may be reduced in time as the diatom barcode library continues to grow, further expanding the number of known sequences and images of diatoms thereby improving the

accuracy of the output of metabarcoding studies (Kermarrec et al., 2013; Zimmermann et al., 2015; Bode et al., 2017; Vasselon et al., 2017; Rimet et al., 2018; Berry et al., 2019; Ruppert et al., 2019).

1.3. Algoa Bay

Algoa Bay is located on the warm, temperate, south-eastern coast of South Africa (Figure 2.1). It is an open bay with an approximate depth of 70 m and width of 70 km. The log-spiral bay is eastward facing and one of the largest of its of its kind along the coast of South Africa (Klages and Bornman, 2005). The bay is at the receiving end of four rivers; the Swartkops, Sundays, and Coega rivers, that enter the bay between the city of Gqeberha (formerly Port Elizabeth) and Cape Padrone, and the Baakens River that enters the bay at Gqeberha on the western shoreline (Figure 2.1) (Mbambo, 2014; Dorrington *et al*., 2018). The bay itself is highly influenced by seasonal dependent winds, primarily west-southwest winds, and the south-westward flowing Agulhas Current (Klages and Bornman, 2005; Mbambo, 2014; Dorrington *et al*., 2018). These influences guide the chemistry of the bay, altering thermoclines, sea surface temperatures and nutrient concentrations (Klages and Bornman, 2005; Mbambo, 2014; Dorrington *et al*., 2018).

The phytoplankton of Algoa Bay were selected as the focus of the study as the bay was the first case study in creating the first Marine Spatial Plan in accordance with the Marine Spatial Planning Bill of South Africa (Dorrington *et al*., 2018; RSA, 2019). This selection was primarily based on the vast amount of biophysical data available as oceanographic research of the bay began in the 1980s. Furthermore, the establishment of the South African Environmental Observation Network (SAEON) Elwandle Node and the Algoa-Bay Sentinel Site for Long-Term Ecological Research (LTER) in 2007 ensured that the bay is considered the best-monitored coastal area in the southern hemisphere (Dorrington *et al*., 2018). Algoa Bay serves as an ideal area of study where the ocean environment provides a vast number of resources and ecosystem services to the local and global economy and therefore needs to be managed sustainably in accordance with The *National Environmental Management: Integrated Coastal Management Act* (24 of 2008) (RSA, 2009).

1.4. Project aims and objectives

The aim of this project was to accurately obtain barcodes and images of at least 10 planktonic diatoms to identify them and contribute information to their use in biomonitoring surveys aiding in the management of the South African coast. Using phytoplankton samples collected from Algoa Bay, diatoms were isolated, cultured, imaged and barcoded on the *rbcL* plastid gene region. This ensured that the diatom images and barcodes were sourced from accurately identified monocultures. Furthermore, nine epizoic diatoms (diatom found on an animal) and one epiphytic diatom (diatom found on a plant) from Dr. Majewska's culture collection at North-West University were included in this study to increase the number of specimens this study would identify. These diatoms were scraped off other organisms (sea turtles, sea snake and kelp) and isolated by Dr. Majewska.

Objectives:

1) Collect samples to isolate and culture diatom species.

2) Maintain successfully isolated diatoms to ensure that they are monocultures.

3) Obtain light microscope and Scanning Electron Microscope (SEM) images of the successfully monocultured diatoms to identify species by morphology.

4) Generate high quality barcodes of the rbcL region for the successfully monocultured diatom species, and then identify species by comparing the generated barcodes to the nucleotide database of the National Centre for Biotechnology Information (NCBI) using the blastn algorithm. If no barcodes were present, the high-quality barcodes generated in this study would be the first barcodes for the diatom species identified in this study.

5) Compare morphological identification to molecular identification and confirm species identification of each monoculture, associating each monoculture with light and scanning electron microscope images and a barcode sequence of the rbcL region.

6) Upload barcodes and images of successfully monocultured diatoms to the open-source genetic library Barcode of Life Data Systems (BOLD).

Chapter 2. Methods and material

2.1. Study sites

Phytoplankton samples were collected by the SAEON Elwandle Node during monthly biomonitoring trips from sites within Algoa Bay off the coast of Gqeberha, South Africa (33°52'16.6"S, 25°42'55.3"E) (Figure 2.1). Collected environmental samples were couriered overnight to the University of the Witwatersrand, Johannesburg, for isolation, culturing, imaging and sequencing to take place.

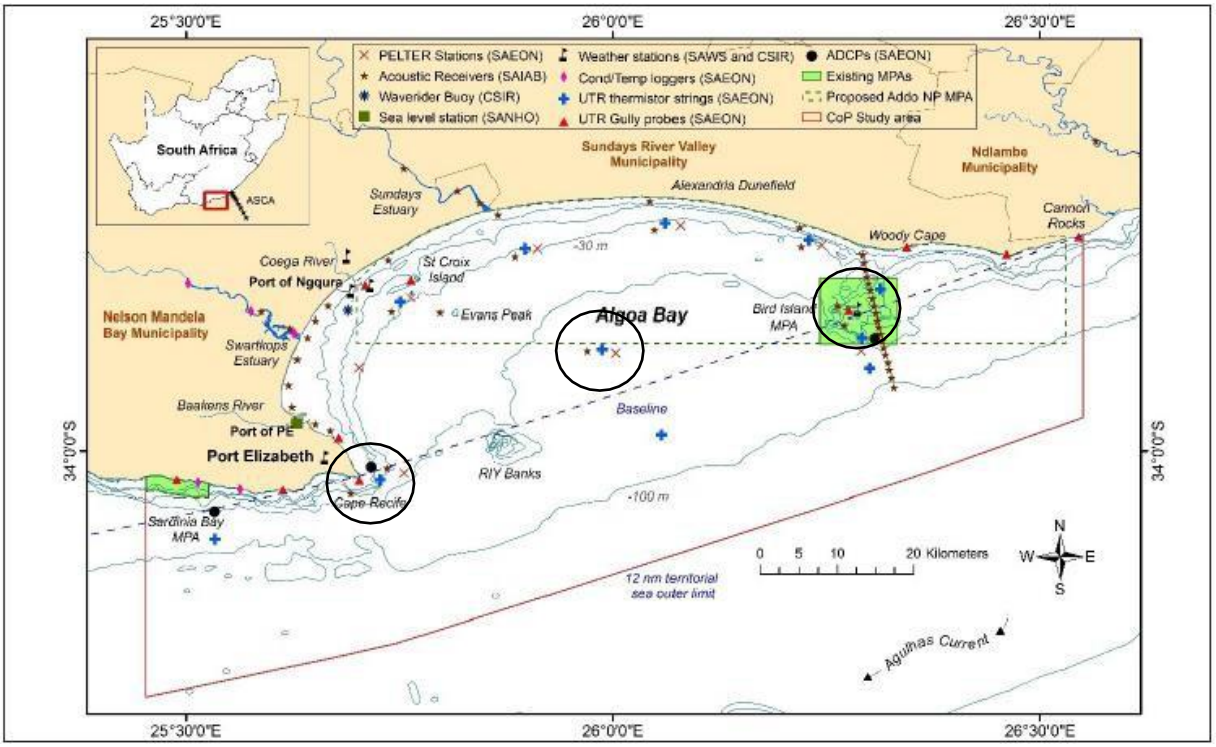


Figure 2.1. Map of Algoa Bay indicating the Pelagic Ecosystem LTER stations (modified from Dorrington *et al*., 2018). Sampling stations utilised in this study were Cape Recife, Bay Central and Bird Island (indicated by circles).

2.2. Sample collection

Physico-chemical variables, nutrients, chlorophyll-a, phytoplankton and zooplankton were collected at Station 6 (Bird Island), 7 (Bay Central) and 8 (Cape Recife) during January and March 2021. (Figure 2.1). Physico-chemical variables measured at each station included, water temperature (ºC), salinity, turbidity (Nephalometric Turbidity Units: NTU), and dissolved oxygen (mg.L^{-1}) were measured using a SeaBird 19 plus V2 CTD mounted on a SBE 55. Water samples for the analysis of chlorophyll-a and nutrients were collected with six Niskin bottles attached to an SBE 55 carousel at predetermined depths. Although these results will not be discussed in this MSc, they will be used to identify the physico-chemical environment of the species in the database.

The phytoplankton samples were collected using a phytoplankton net (Hydro-Bios Apstein net - 438001: 25 cm diameter, 55 µm mesh) towed both vertically and horizontally at each station. Vertical tows were made by hand hauling the net from the seafloor to the surface, while the horizontal tows were done underway at a speed of 2 knots for 3 minutes. The phytoplankton environmental samples were transferred to a 50 ml plastic tube and kept cold (< 10°C) until collection by the courier on the same day for transport to Johannesburg. Monocultures obtained from Dr. Majewska were collected as scrapes from different organism between 2019 and 2021 in South Africa. These monocultures were courier to University of the Witwatersrand under the above-mentioned conditions in May 2021.

2.3. Isolation and monoculture protocols

Monocultures were created using methods outlined by Witkowski *et al.* (2020) and Sym (pers. comm., 2021), where diatoms were isolated using a micropipette and an inverted microscope at 400x magnification (Olympus CK 2 Phase Contrast microscope). Cultures were grown in Daigo's IMK medium made up with filtered sea water (Sym pers. comm., 2021) and stored in an incubator at approximately 22°C with cool white fluorescent lighting on 24 hours. Monocultures were kept alive by subsampling bi-monthly; the old culture was agitated and 5 – 10 ml of this monoculture was transferred to a newly labelled sterile test tube using a clean glass pasture pipette. Thereafter, the old and the new cultures were filled with fresh culture media. Culture maintenance continues today, however, only cultures from July and September 2021 were used in this study.

Five phytoplankton environmental samples were collected on the 28th of January 2021 and isolates were created on the 29th of January 2021 in Johannesburg. Isolates and samples were stored in an incubator under the above-mentioned conditions. An additional attempt of transporting live phytoplankton environmental samples was made in March 2021 following the same protocol as above but was unsuccessful as the samples arrived after 48 hours in Johannesburg, thus, monocultures were only created from the January 2021 sampling session.

Dr Majewska's existing monocultured species that were included in the study were couriered on the 19th of May 2021 and subcultured on the same day at the University of the Witwatersrand. These cultures were kept in the same conditions as those from Algoa Bay and processed in the same manner.

2.4. Morphological identification

Phytoplankton environmental samples from Algoa Bay were inspected, live diatoms isolated and identified, and containers of potential monocultures labelled. Identifying diatom species required using key morphological features of diatoms such as their shape and size, and other unique characteristics such as colour (live chloroplasts) (Reimann and Lewin, 1964; Giffen, 1971, 1973, 1976; Hasle *et al*., 1996; Sarno *et al*., 2005; Harding and Taylor, 2011; López-Fuerte *et al.,* 2020). Isolated diatoms were left for a period of 1 – 2 weeks to grow and cultures were monitored using an Olympus CK 2 inverted microscope to ensure that they were monocultures. When it was clear that the isolation process was successful, representative samples from each monoculture were extracted for the creation of microscope slides and scanning electron microscope stubs. These samples were processed using two different organic matter digestion methods to display the silica frustule of the diatom specimens for accurate species identification. These two methods were evaluated to determine if they can provide comparable results.

The first method used saturated potassium permanganate and concentrated hydrochloric acid to digest the organic matter (Harding and Taylor, 2011) and the second method used nitric acid and potassium dichromate (Elias *et al*., 2017; Cotiyane-Pondo and Bornman, 2021). The hydrochloric acid and potassium permanganate method is more established for processing diatoms, so the nitric acid and potassium dichromate method was only used to compare the efficiency of organic matter digestion. The nitric acid and potassium dichromate method used was as follows, 5 ml of monoculture sample, 5 ml of nitric acid and 5 ml of potassium dichromate (made in a 1:2 weight to volume ratio with distilled water) were combined in a 50 ml glass beaker and left on a shaker at 90 rpm for 24 hours (Elias *et al*., 2017; Cotiyane- Pondo and Bornman, 2021). Samples were then transferred to a centrifuge tube and rinsed with distilled water by centrifugation at 3,000 rpm for 10 minutes, this was repeated until the supernatant was clear in colour. Between repetitions the supernatant was decanted into a sealed labelled disposable container without disturbing the pellet. Once the supernatant was clear, the pellet was resuspended and decanted into Eppendorf tubes for storage. After monoculture samples were digested using either digestion method, they were fixed to microscope slides using Pleurax.

Microscope slides were viewed using a Zeiss Axio Imager 2 and Zen 2012 software (blue edition) was used to annotate and capture images. A Vega TESCAN scanning electron microscope was used to take images by the Microscope and Microanalysis Unit (MMU) (University of the Witwatersrand) and a JOEL JSM–7001F scanning electron microscope was used to take images by the Centre of High Resolution Transmission Electron Microscopy Unit (CHRTEM), (Nelson Mandela University). Both SEMs, operated by their respective technicians, took images of July 2021 monocultures that were digested using the hydrochloric acid and potassium permanganate method (Harding and Taylor, 2011) to remove organic material and coated with ~10 nm of carbon and ~5 nm of gold-palladium (Witkowski et al., 2020). September 2021 monocultures were digested using the nitric acid and potassium dichromate digestion method (Elias *et al.,* 2017; Cotiyane-Pondo and Bornman, 2021) and imaged on a Phenom Pure G6 Desktop SEM (Thermo Fisher Scientific) at the University of the Witwatersrand Herbarium. These images were compared to determine if there were any major differences between the digestion methods, image quality of SEMs and if the monocultures varied morphologically based on the age of the culture.

2.5. Molecular identification

The molecular lab work of successfully monocultured diatoms was conducted by the African Centre of DNA Barcoding (ACDB), University of Johannesburg. They performed DNA extraction, PCR amplification and DNA sequencing providing trace files of the monoculture samples given to them (Powell *et al*., 2018)

2.5.1. DNA extraction and quantification

DNA was extracted from monocultured diatom samples using a standard hexadecyl-trimethyl-ammonium bromide (CTAB) protocol (Doyle and Dickson, 1987). The two buffers used in this protocol are 2x CTAB buffer (100 mM Tris (pH 8.0), 1.4 M NaCl, 20 mM EDTA, 2% CTAB, 0.1% PVPP and add beta-mercaptoethanol to 0.2% (v/v) daily) and TE buffer (10 mM Tris (pH 8.0) and 0.1 mM EDTA).

Prior to DNA extraction 5 ml of the monocultured sample was centrifuged three times at 3,000 rpm for 10 minutes to increase the concentration of diatom individuals in the pellet. Thereafter, the pellet was re-suspended with double-distilled water. DNA was extracted from each monoculture pellet using the following method; 0.01 ml of 2 x CTAB buffer was added per 1 mL of collected cells in a reaction tube. This was then incubated at 60°C for 0.5 -1 h then 0.01 mL of chloroform:isoamyl alcohol (24:1) was

added and samples were mixed by centrifuging for 10 minutes at 5,000 rpm. After centrifugation the aqueous phase (upper layer) was transferred into a new tube and the chloroform:isoamyl alcohol extraction was repeated until the interface was clean. Once clean, 2/3 volume of isopropanol was added and samples were centrifuged for 15 minutes at 10,000 rpm then the supernatant was discarded. The pellet was washed with 70% ethanol, air dried and resuspended in double distilled water. A 1% agarose gel electrophoresis was performed to determine the presence or absence of DNA and band strength was used to provide an estimation of DNA concentration.

2.5.2. PCR amplification

The ribulose-1,5-bisphosphate carboxylase/oxygenase plastid gene (*rbcL*) was selected in this study for identification of diatoms and was amplified based on the method outlined by Vasselon *et al.* (2017) (Table 1). PCR amplification required making up a master mix for each monoculture that had DNA successfully extracted. The master mix per reaction was made to 25 µl, composed of; Dream *Taq* (12.5 µl), Bovine Serium Albumin (BSA) (0.8 µl at 32 mg/mL), 0.3 µl (1 µM) of each forward and reverse primers (Table 1), and 11.1 µl of nuclease free water. After preparation of the master mix, 24 µl of master mix and 1 µl of monoculture template DNA was added to PCR reaction tubes specific to each monoculture. Thereafter, the tubes were cycled a PCR machine and the 25 µl PCR product was used in DNA sequencing.

Table 2.1. PCR primer sequences and annealing temperatures used for the amplification of monocultured diatom DNA samples (Zimmermann et al., 2015; Vasselon et al., 2017).

Gene region	Primer name	Primer sequence 5'-3'	PCR annealing temperature (°C)	Amplicon size (bp)
rbcL	Diat_rbcL_708F_1	AGGTGAAGTAAAAGGTTCWTACTTAAA	45	~315
	R3_1	CCTTCTAATTTACCWACWACTG		

2.5.3. DNA sequencing and analysis

The rbcL region was sequenced in forward and reverse directions using Sanger sequencing. Two master mixes were created, one for the forward primer and one for the reverse primer. The master mixes composed of 1.5 µl of 5x sequencing buffer and 0.3 µl of the appropriate primer that was mixed before the addition of 0.3 µl of Big Dye Terminator v3.1 and 6.9 µl of nuclease free water (Powell *et al*., 2018). These volumes are required for the formation of master mix for a single sequencing reaction. 9 µl of the master mix and 1 µl of the PCR product was used in each sequencing reaction per direction. The ACDB inspected sequencing output and if necessary, re-sequenced the PCR product of an individual culture more than once to generate molecular data for the rbcL region.

Multiple chromatograms were generated in both directions per monoculture by ACDB then visualised on Geneious Prime® 2021.2.2, and regions trimmed with more than a 6% chance of error per base. These regions were excluded from downstream analysis. The chromatograms per monoculture were then *de Novo* assembled using the Geneious Assembler, at the highest sensitivity creating contigs made with at least one forward and one reverse sequence. The contigs per monoculture were inspected and ambiguities noted and not changed. The contigs were used to generate consensus sequences for each monoculture using a 60% threshold for contigs. However, if a contig contained at least one sequence without quality information, then the threshold was adjusted to 95%. A single consensus sequence was created per monocultured species and this was regarded as its barcode. Barcodes were inspected and translated on Geneious; ambiguities, protein coding regions and stop codons were identified and unchanged. For downstream analysis, consensus sequences met the criteria; of being made of at least one forward and one reverse sequence, having a sequence length greater than 250 nucleotide base pairs, and percentage of high-quality bases in an untrimmed generated consensus sequence (%HQ) had to be greater than 65%. These consensus sequences were compared to the National Centre for Biotechnology Information (NCBI) nucleotide database using the blastn algorithm to identify if the monoculture had been previously associated to a diatom species or genus. A batch BLAST was done on Geneious to retrieve matching regions on the nucleotide database, a hit table of 100 hits was created per monoculture consensus sequence.

Trace file sequences that were initially excluded from analysis were re-inspected. Trace files of samples were annotated and trimmed with a 10% chance of error per base, thereafter, the abovementioned method using the Geneious Assembler to generate consensus sequences per monoculture was followed. Consensus sequences were used regardless of a criterion, provided the molecular species identification (blasting) matched morphological identification, the low-quality consensus sequence was used for further analysis, or else excluded.

Sequences of the same region, of similar species or genera that were identified in this study by BLAST results, were mined from GenBank and included in this study as reference sequences of known specimens. These sequences together with the consensus sequences of this study were aligned using the Geneious Alignment, this was a Global Alignment with free end gaps and a cost matrix of 93% similarity. The open gap penalty was set at 12, gap extension penalty at 3 and refinement iterations of 2. The alignment was viewed to determine if monocultures and GenBank specimens aligned over the same rbcL region and if there were any major discrepancies in any of the sequences. Sequences were not altered in any way.

The Geneious Tree builder was used to create an unrooted neighbour-joining tree using consensus sequences of this study and those mined from GenBank. This was done to infer relatedness of the diatom monocultures. The Geneious Tree builder obtained distances from pairwise alignments of all sequence pairs, with the alignment type as a global alignment with free end gaps using a 93% similarity to build the distance matrix. The unrooted tree was built using the Jukes-Cantor Genetic Distance model following the Neighbour-Joining tree building method with no outgroup.

2.6. Barcode of Life Data Systems (BOLD)

A project was created on BOLD systems version 4 and the monoculture details uploaded. Trace files were uploaded and underwent a quality check based on BOLD systems; samples that were successful of the quality check had morphological data uploaded to BOLD (Ratnasingham and Herbert, 2007). Upon meeting the minimum standards set by GenBank for nucleotide records, successful monocultures will have published records through BOLD data systems to obtain accession numbers from GenBank.

Chapter 3. Results

3.1. Sample collection

A total of 14 planktonic diatoms were isolated from Algoa Bay (AB) and 10 monocultures of epizoic and epiphytic diatoms were obtained from Dr. Majewska (NWU) (Table 3.1). A total of 19 of the 24 monocultures were successfully identified and monocultures maintained, 10 from Algoa Bay (AB1, 2, 3, 5, 8, 9, 10, 11, 12, and 13) and nine from NWU (NWU1, 2, 3, 5, 6, 7, 8, 9, and 10). These samples were identified using morphology and molecular data when present. Six cultures were identified by their morphology being analogous to that of published work and their molecular data blasting to published data at greater than 70% (Table 3.5). Nine cultures only had molecular data present at a genus level or were inconclusive, hence, morphology was used to identify the species and associate them with the barcode sequence when possible (Table 3.5). Four cultures were identified by morphology as molecular analyses was unsuccessful with it being a possibility in the future as the cultures are still alive (Table 3.5). The five samples not used in further analysis were due to cultures dying or through failure to maintain as monocultures (Table 3.1).

Table 3.1. Planktonic samples collected from Algoa Bay include sites within the bay (Bay Central (BC) (25.98722 E, -33.881778 S); Cape Recife (CR) (25.72771 E, -34.034115 S); Bird Island (BI) (26.29146 E, -33.868695 S)), using different towing methods (Horizontal (H) and Vertical (V)). Epizoic and epiphytic samples obtained from Dr. Majewska (NWU) were collected as scrapes from different organisms and at different times in South Africa. Identification of sample monocultures occurred upon analysis of morphology and molecular data. * Indicates samples not used for analysis.

Sample ID	Collection site	Sampling protocol/ organism	Collection date	Habitat	Identification and level	Culture status
AB1	BC	V	January 2021	Planktonic	Both	Alive
AB2	BC	V	January 2021	Planktonic	Both	Alive
AB3	BC	V	January 2021	Planktonic	Both	Alive
AB4*	CR	V	January 2021	Planktonic	-	Dead
AB5	CR	V	January 2021	Planktonic	Morpho	Alive
AB6*	CR	V	January 2021	Planktonic	-	Alive
AB7*	CR	V	January 2021	Planktonic	-	Dead
AB8	BI	V	January 2021	Planktonic	Both	Alive
AB9	Unknown	Unknown	January 2021	Planktonic	Morpho	Alive
AB10	CR	V	January 2021	Planktonic	Both	Alive
AB11	CR	H	January 2021	Planktonic	Both	Alive
AB12	BC	H	January 2021	Planktonic	Morpho	Alive
AB13	CR	V	January 2021	Planktonic	Both	Alive
AB14*	CR	V	January 2021	Planktonic	-	Alive

NWU1	uShaka Marine World	Carapace loggerhead sea turtle “Munchkin”	March 2021	Epizoic	Both	Alive
NWU2	uShaka Marine World	Yellow-bellied sea snake	August 2020	Epizoic	Morpho	Alive
NWU3	Hibberdene	Seaweed on tidal plain	March 2019	Epiphytic	Morpho	Alive
NWU4*	uShaka Marine World	Munchkin	March 2021	Epizoic	-	Alive
NWU5	uShaka Marine World	Munchkin	March 2021	Epizoic	Both	Alive
NWU6	uShaka Marine World	Munchkin	March 2021	Epizoic	Both	Alive
NWU7	uShaka Marine World	Yellow-bellied sea snake	August 2020	Epizoic	Both	Alive
NWU8	uShaka Marine World	Yellow-bellied sea snake	August 2020	Epizoic	Both	Alive
NWU9	Bayworld, Gqeberha	Green sea turtle “Vader”	November 2019	Epizoic	Morpho	Alive
NWU10	Bayworld, Gqeberha	Green sea turtle “Jimmy”	December 2020	Epizoic	Both	Alive

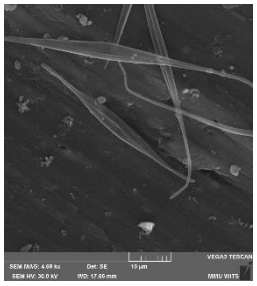
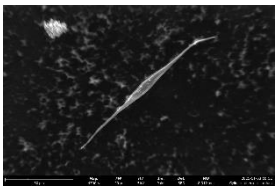
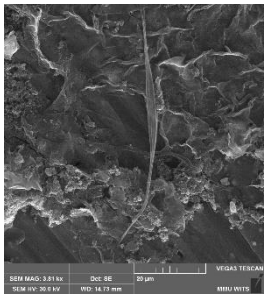
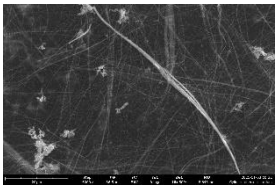
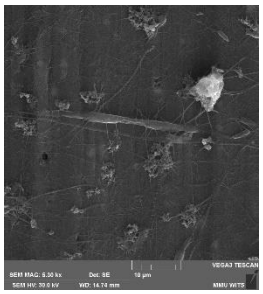
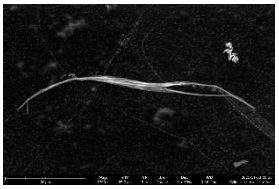
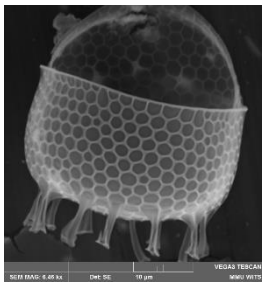
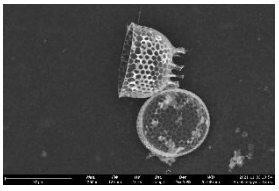
3.2. Morphological identification - light and scanning electron microscopy

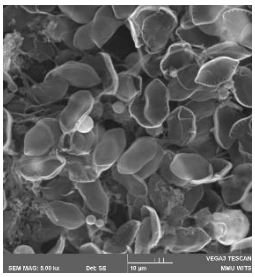
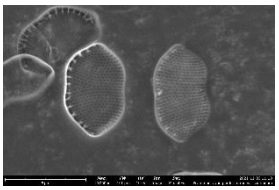
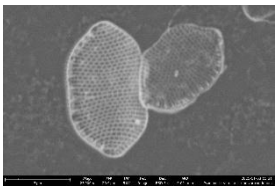
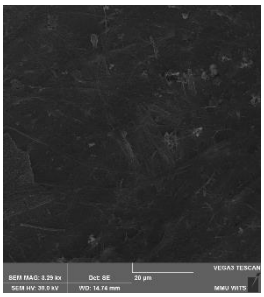
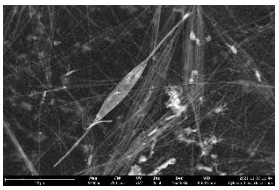
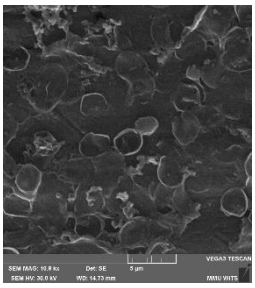
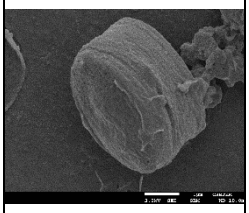
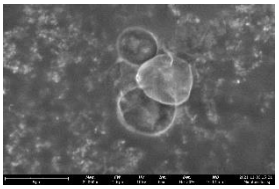
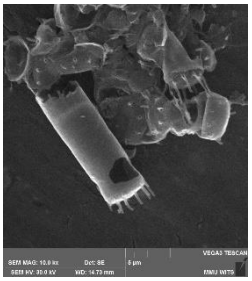
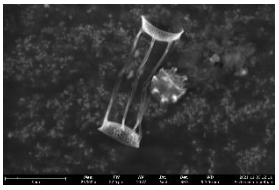
Image plates were created per successfully maintained monocultured sample (AB 1, 2, 3, 5, 8, 9, 10, 11, 12, 13 and NWU 1, 2, 3, 5, 6, 7, 8, 9, 10) displaying light microscope images of the live monocultures and digested (using both methods) monocultures using brightfield, differential interference contrast and phase contrast microscopy. The plates also include images from the Vega TESCAN taken by the MMU, JOEL JSM-7001F taken by the CHRTEM and the Phenom Pure G6 Desktop scanning electron microscopes (Supplementary material 1). These image plates serve as a reference of images of each species identified in this study and represent cultures of July and September 2021.

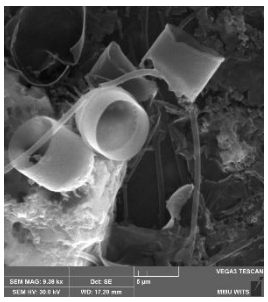
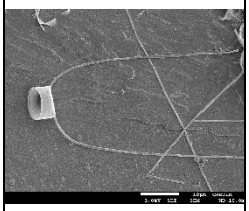
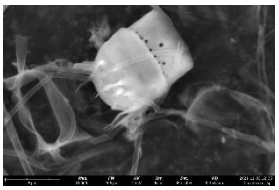
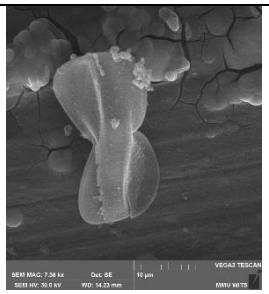
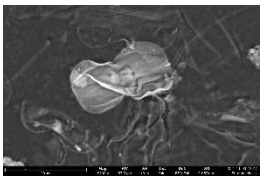
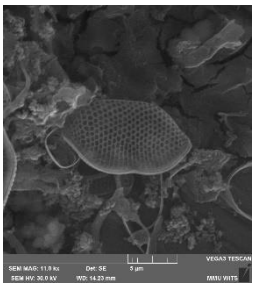
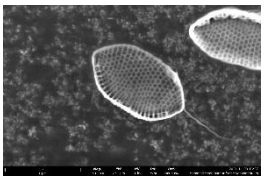
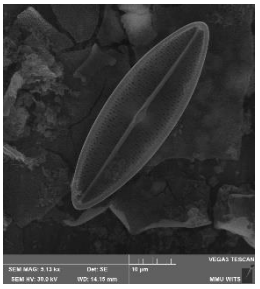
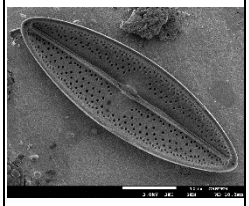
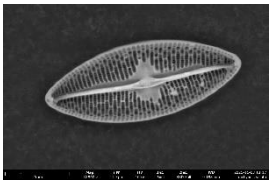
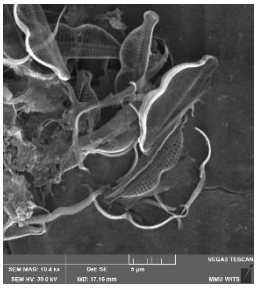
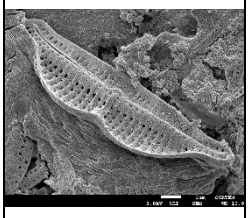
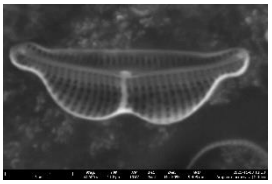
Table 3.2 displays an easy to visualise comparison of July and September 2021 cultures, also showing image comparisons on different SEMs used and the effects of the different organic matter digestion methods implemented in this study. Based on morphology, species or genus identification was provided per monoculture in reference to published literature.

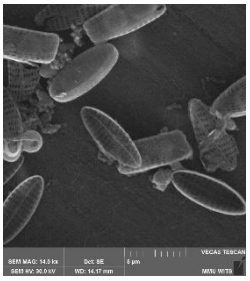
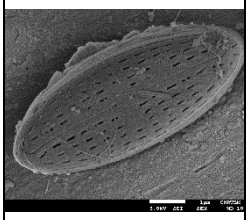
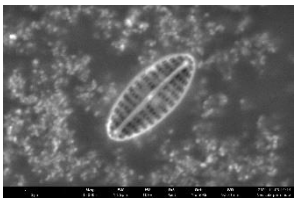
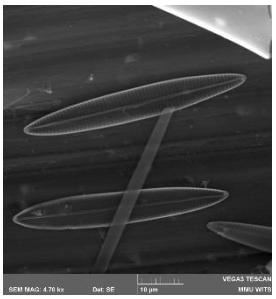
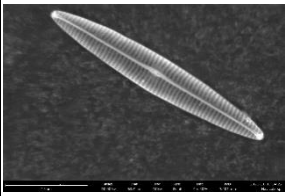
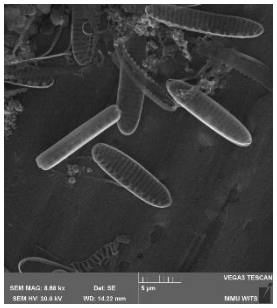
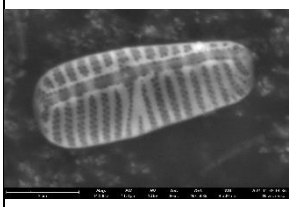
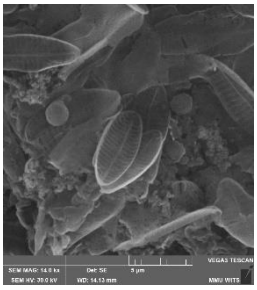
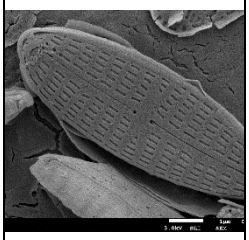
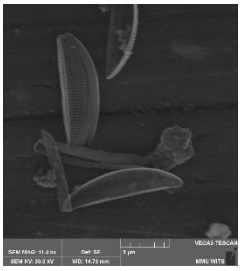
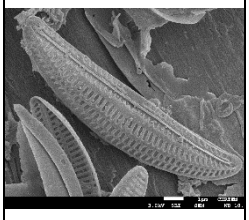
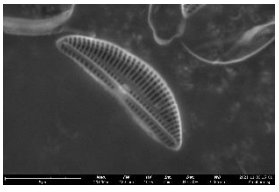
The hydrochloric acid and potassium permanganate digestion method was more aggressive as seen in the images as there were crystalized residue around the frustules. Some frustules appear to be damaged suggesting that the hydrochloric acid may be too strong (Table 3.2, Supplementary material 1). The nitric acid and potassium dichromate digestion method provided comparable images to the hydrochloric acid digestion method with less damage to the diatom frustules and less residue on the samples (Table 3.2, Supplementary material 1). Regardless of the sample digestion method and the SEM used, diatom cultures between July and September 2021 yielded similar looking specimens with minor morphological variations (Table 3.2). Images from different SEMs provide different resolutions however species identification was possible irrespective of the instrument used (Table 3.2, Supplementary material 1).

Table 3.2. SEM images using the Vega TESCAN and JOEL - JSM 7001F SEMs, operated by technicians at MMU and NMU respectively. These were images of July 2021 cultures using the hydrochloric acid and potassium permanganate digestion method (Harding and Taylor, 2011). The phenom pure G6 desktop SEM was used to take images of September 2021 cultures digested by the nitric acid and potassium dichromate digestion method (Elias et al., 2017; Cotiyane-Pondo and Bornman, 2021).

Sample ID	July 2021 cultures digested by hydrochloric acid and potassium permanganate		Sept 2021 cultures digested by nitric acid and potassium dichromate	Morphology specimen ID and reference
	Vega TESCAN SEM	JOEL - JSM 7001F SEM	Phenom Pure G6 Desktop SEM	
AB1		-		*Cylindrotheca closterium* (Ehrenberg) Reimann and Lewin (1964)
AB2		-		
AB3		-		
AB5		-		*Stephanopyxis turris* (Greville) Ralfs (Hasle *et al*., 1996)

AB8		-		*Psammodictyon panduriforme var. continuum* (Grunow) Snoeijis (López-Fuerte *et al*., 2020)
AB9		-		
AB10		-		*C. closterium* (Reimann and Lewin, 1964)
AB11				Minutocellus sp. (Ashworth *et al*., 2013)
AB12		-		*Skeletonema grethae* (Zingone and Sarno) (Sarno *et al*., 2005)

AB13				Chaetoceros sp. (Al-Yamani and Saburova, 2019)
NWU1		-		Entomoneis sp. (Mejdandzic *et al*., 2018)
NWU2		-		*P.parduriforme var. continuum* (López-Fuerte *et al*., 2020)
NWU3				*Trachyneis velata* (Schmidt) (Giffen, 1971)
NWU5				*Amphora incrassata* (Giffen) (Giffen, 1973)

NWU6				*Navicula perminuta* (Grunow) (Stachura-Suchoples *et al*., 2015; Majewska *et al*., 2021)
NWU7		-		Navicula sp. (Hasle *et al*., 1996)
NWU8		-		Nitzchia sp. (Hasle *et al*., 1996)
NWU9				Navicula sp.
NWU10				Amphora sp. (Al-Kandari *et al*., 2009)

3.3. Molecular identification

A total of 24 samples were sent to the African Centre of DNA Barcoding (ACDB) for sequencing in July 2021 after successful confirmation of July 2021 monocultures (Tables 3.2 and 3.3). Furthermore, 12 samples were re-sequenced in September 2021 (Table 3.4) from the September 2021 cultures (Table 3.2). Each sequencing event was treated independently. After the first sequencing event monocultures AB4, 7, 14 and NWU4 were excluded from downstream analysis due to culture death or through failure to be maintained as a monoculture. DNA sequence analysis resulted in a consensus sequence per monoculture, which were generated from multiple forward and reverse sequences of the monocultures DNA. The consensus sequence length, percentage of high-quality bases in an untrimmed generated consensus sequence (%HQ) and number of ambiguities found on the consensus sequence were all noted (Tables 3.3 and 3.4).

Monocultures AB8, 10, 11 and NWU3, 6 and 9 were re-sequenced to confirm species identification and accurately identified specimens (Tables 3.3 and 3.4). Monocultures AB2, 3, 5, 6, 12 and NWU2 were re-sequenced but provided low-quality results as consensus sequences had a high number of ambiguous bases (Tables 3.3 and 3.4). AB6 was excluded from downstream analysis as morphology was not obtained. The first blast result from the July 2021 cultures of AB6 was inconclusive and did not blast to a nucleotide sequence on the NCBI database. The second blast from the September 2021 cultures displayed a match to an *Entomoneis* sp., however, with no morphology data captured the molecular data was not useful.

The consensus sequences per monoculture were blasted to the NCBI nucleotide database and monocultures were identified from their molecular data (Table 3.5). Monocultures AB1, 2, 3, 8, 10 and NWU6 blasted to expected species (Table 3.5) and AB11, 13 and NWU1, 3, 5, 7, 8, 9, 10 blasted to specimens identified to genera or were misidentified when compared to morphology (Table 3.5). Monocultures AB5, 9, 12 and NWU2 were identified by their morphology (Table 3.2 and Supplementary material 1) as molecular data were inconclusive or included as low-quality barcodes (Tables 3.3, 3.4 and 3.6).

Table 3.3. Sequencing results from July 2021, showing the number of forward and reverse sequences that were used to generate consensus sequences (barcode) of each monocultured species. The number of ambiguities, sequence length and the percentage of bases that were high quality in the consensus sequence (%HQ) were recorded. Sample IDs highlighted were used for comparing to the NCBI nucleotide database via the blastn algorithm. Samples marked with * were not used in further analysis.

Sample ID	Number of sequences		Number of ambiguities	%HQ	Sequence length
	Forward	Reverse			
AB1	3	3	7	82.5	383
AB2	3	3	13	85.6	494
AB3	3	3	40	76.3	486
AB4*	3	3	1	93.6	297
AB5	-	-	-	-	-
AB6*	2	3	2	95.2	378
AB7*	3	3	3	96.5	283
AB8	3	3	4	95	281
AB9	1	1	33	57.9	480
AB10	1	1	2	93.2	295
AB11	1	1	2	91.2	295
AB12	1	1	21	63.9	490
AB13	2	2	1	95.8	309
AB14*	-	-	-	-	-
NWU1	2	2	4	86.3	315
NWU2	1	-	-	67.4	273
NWU3	2	1	24	48.8	486
NWU4*	1	1	1	95.2	294
NWU5	1	1	3	93.5	294
NWU6	2	2	1	87.6	386
NWU7	2	2	3	87.3	315
NWU8	2	2	5	76.5	404
NWU9	1	1	20	69.9	489
NWU10	1	1	1	92.2	283

Table 3.4. Sequencing results from September 2021, showing the number of forward and reverse sequences that were used to generate consensus sequences (barcode) of each monocultured species. The number of ambiguities, sequence length and the percentage of bases that were high quality in the consensus sequence (%HQ) were recorded. Sample IDs highlighted were used for comparing to the NCBI nucleotide database via the blastn algorithm and samples marked with * were not used in further analysis

Sample ID	Number of sequences		Number of ambiguities	%HQ	Sequence length	Reason for repeat sequencing
	Forward	Reverse				
AB2	-	-	-	-	-	Confirm first result
AB3	1	1	6	26.9	321	Confirm first result
AB5	1	-	-	27.8	378	Inconclusive first time
AB6*	2	1	4	67.5	305	Confirm first result
AB8	1	1	0	80.3	294	Confirm first result
AB10	1	1	2	74.9	275	Confirm first result
AB11	1	1	13	66.2	287	Confirm first result
AB12	1	1	11	38.7	318	Inconclusive first time
NWU2	1	1	1	53.6	320	Inconclusive first time
NWU3	1	1	2	76	292	Inconclusive first time
NWU6	1	1	2	86.5	310	Confirm first result
NWU9	1	1	4	71.3	314	Confirm first result

Monocultures AB1, 2, 13 and NWU1, 5, 7, 8, 10 were sequenced only during the first sequencing event (Table 3.3). Only monocultures AB1 and AB2 provided molecular data to species level that matched morphological species identification (Tables 3.2 and 3.5). AB13 provided blast results to the wrong species with other hit results to the *Cyclotella* genus and uncultured phytoplankton clones, thus, morphology was used to associate the barcode generated in this study with the *Chaetoceros* genus (Tables 3.2 and 3.5). NWU5 was identified by morphology to a species level and molecular data provided the same genus, other hits from blasting provided similar species to morphological identification, therefore the barcode generated in this study was associated to the species identified by morphology. NWU1, 7, 8 and 10 provided molecular blast results to the same genera predicated by morphological identifications (Tables 3.2 and 3.5). NWU1 blasted to a variety of *Entomoneis* species, namely, *E. umbratical, E. vilicicii, E. gracilis, E. infula and E. tenera*. Due to the cryptic nature of this species morphology was only conclusive to a genus and based on molecular results the genus was confirmed, however, species identification requires better taxonomic resolution of the genus and further morphological and molecular work. NWU7 blasted to a *Navicula sp.* with other matches being *Seminavis* sp. *and Haslea* sp., which were disregarded due to morphological differences to NWU7. NWU8 blasted to various *Nitzschia* sp. and *Fragilariopsis* sp., however, due to morphological differences the culture was not identified to species and associated to the *Nitzschia* genus. NWU10 blasted to *Amphora, Halamphora, Cocconeis, Navicula, Frustulia and Nitzchia* sp., however, was identified to genus *Amphora* by morphology.

Monocultures AB3, 8, 11 and NWU6 were sequenced twice (Tables 3.3 and 3.4), and blast results were a match to the same specimen from the nucleotide database in both sequencing events (Table 3.5). Monoculture AB10 may have been contaminated in the July sequencing as blast results were inconclusive when compared to morphological identification (Tables 3.2 and 3.5). Re-sequencing in September 2021 provided accurate molecular identification as the result matched morphological identification (Table 3.2 and 3.5). This provided further evidence that the July sequencing analysis of this sample may have been contaminated. Monocultures NWU3 and 9 were sequenced twice and blast results provided misidentification between sequencing events and, in comparison, to morphology (Tables 3.2 and 3.5). These two monocultures were only identified by morphology as molecular data were inconclusive and disregarded.

Table 3.5. Blast results identifying genera or species of July 2021 and September 2021 sequencing. Images of associated cultures are labelled according to Sample ID in Table 3.2. * Indicates morphological identification was used to identify monoculture as molecular data were inconclusive. # Indicates that the samples morphological identification (Table 3.2) matched the blast algorithms output. $ Indicates samples were only identified to genera. LQ Indicates sequences that were re-analysed at a 10% error per base and then consensus sequences used for blast analysis and molecular species identification.

Sample ID	Sequencing date	Blast match	% Match	Blast accession number	References
AB1#	July 2021	*Cylindrotheca closterium* (Ehrenberg) Reimann and Lewin	86.3		
AB2#	July 2021	*C. closterium*	76.2	MH807654	
AB3#	July and September LQ 2021	*C. closterium*	72.2	AY866415	Stock *et al*., 2019
AB5*	-	-	-	-	-
AB8#	July and September 2021	*Psammodictyon panduriforme* var. *continuum* (Grunow) Snoeijis	98.4 99.3	MH390344	Romanova, 2018, unpublished
AB9*	-	-	-	-	-
AB10#	July 2021	*Minutocellus. polymorphus* (Hargraves and Guillard) Hasle, Stosch, and Syvertsen	97.9	KC309572	Contamination
	September 2021	*C. closterium*	96.3	MH807639	Stock *et al*., 2019
AB11$	July and September 2021	*Minutocellus Polymorphus* (Hargraves and Guillard) Hasle, Stosch, and Syvertsen	98.3 95.7	KC309572	Ashworth *et al*., 2013
AB12LQ	-	-	-	-	-
AB13$	July 2021	*Thalassiosira pseudonana* (Hasle and Heimdal)	99	XM_002297482	Bowler *et al*., 2008

NWU1[$]	July 2021	*Entomoneis* sp.	97.8	KX120572	Ruck *et al*., 2016
NWU2[LQ]	-	-	-	-	-
NWU3*	July 2021[LQ]	*Psammodictyon* sp.	80.1	MN920680	Mann *et al*., 2020, unpublished
	September 2021	*Entomoneis* sp.	98.3	KX120572	Ruck *et al*., 2016
NWU5[$]	July 2021	*Amphora* sp.	98.6	MH064061	Sabir *et al*.,2018
NWU6[#]	July and September 2021	*Navicula perminuta* (Grunow)	86.4 97.9	MT432484	Majewska *et al*., 2020
NWU7[$]	July 2021	*Navicula* sp.	98.1	MH040271	Lobban *et al*., 2018
NWU8[$]	July 2021	*Nitzchia* sp.	87	MT042784	Contreras *et al*., 2020
NWU9*	July 2021	*Chaetoceros tenuissimus* (Meunier)	79.1	MW848497	Xu and Chen, 2021
	September 2021	*Entomoneis umbratica* sp. nov. (Mejdandžić and Bosak)	98.2	MF000629	Mejdandzic *et al*., 2018
NWU10[$]	July 2021	*Amphora* sp.	96.9	MH064142	Sabir *et al*.,2018

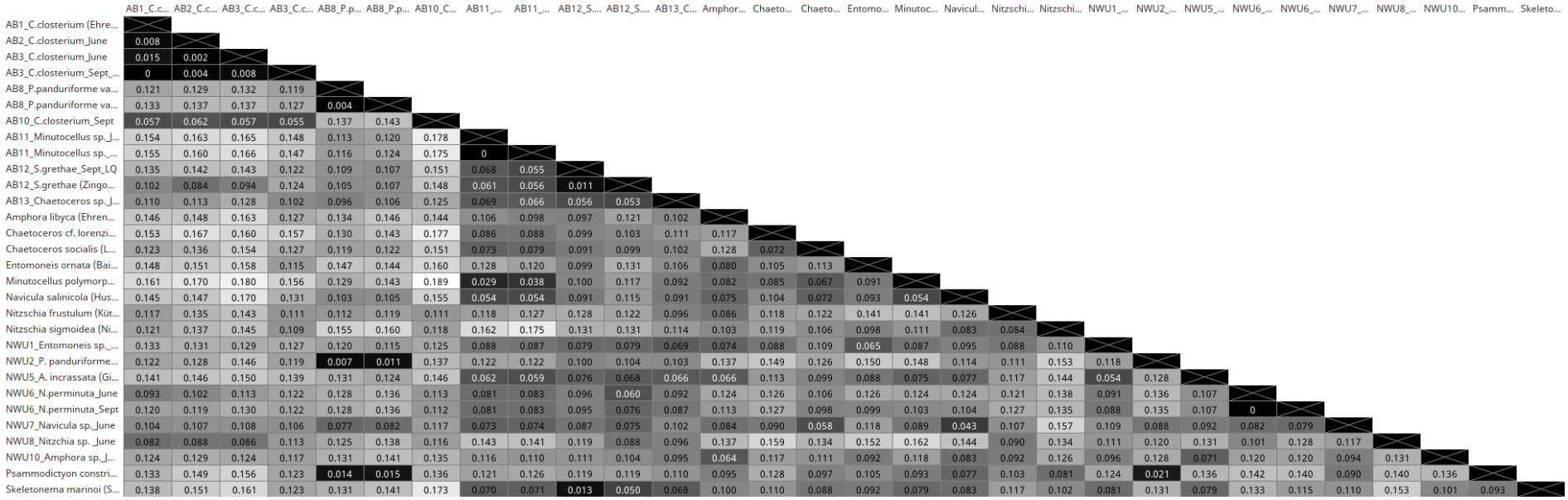
Monoculture sequencing data that were initially excluded from the analysis due to not meeting the 6% threshold of chance of error per base were reanalysed at a 10% chance and assembled to generate consensus sequences of low-quality data that were used in blasting to determine molecular species identification. This provided an evaluation of whether, low quality sequences still provide accurate species identification (Table 3.6), i.e., the blastn results of the low-quality sequences matched morphological identification. Monocultures AB3, 12 and NWU2 provided similar molecular identification to morphological identification (Tables 3.2, 3.5 and 3.6). AB3 blasted to multiple hits of *Cylindrotheca* sp., a few *Nitzschia* sp. and an uncultured phototrophic eukaryote. AB12 blasted *Skeletonema tropicum* (93.5%) in both sequencing events but was matched to the wrong rbcL region however, *Skeletonema gerthae* was a match and across the correct region. NWU2 blasted to the *Psammodictyon* genus and was found to match *P. panduriforme* var. *continuum*. Monoculture AB5 was unable to generate a sequence from amplicons thus no consensus sequence was created (Tables 3.3 and 3.4). AB9 had low quality sequences that blasted to a different species then expected (Table 3.6) and was not re-sequenced in September (Tables 3.3 and 3.4). NWU3 low-quality sequence blasted to a different genus than that of the higher quality sequence from the first sequencing event (Tables 3.5 and 3.6); however, neither of them provided accurate results to morphological identification (Table 3.2). AB9 and NWU3 low-quality sequences were thus excluded from analysis.

Table 3.6. Samples re-analysed at a 10% error per base and blasted to the nucleotide database. Blast results from October 2021. * Indicates morphological identification was used to identify monoculture as molecular data were inconclusive. # Indicates that the samples morphological identification (Table 3.2) matched the blast algorithms output.

Sample ID	Date sequenced	Blasted result	Percentage match	Blast accession number	Reference
AB3#	September 2021	*Cylindrotheca closterium* (Ehrenberg) Reimann and Lewin	98.1	AY866415	Li *et al*., 2004
AB5*		-	-	-	
AB9*	July 2021	*Chaetoceros tenuissimus* (Meunier)	76.6	MK642555	De Luca *et al*., 2004
AB12#	July and September 2021	*Skeletonema grethae* (Zingone and Sarno)	79.5	DQ514818	Alverson *et al*., 2007
			95.3		
NWU2#	July and September 2021	*Psammodictyon panduriforme* var. *continuum* (Grunow) Snoeijis	99.6	MH390348	Romanova, unpublished
			99.2		
NWU3*	July 2021	*Psammodictyon* sp.	80.1	MN920680	Mann *et al*., 2020, unpublished

An unrooted neighbour-joining tree was used to confirm the relatedness of the successfully identified monocultures (Figure 3.2). Sequences were grouped together based on nucleotide differences and were arranged based on sequence similarity, thus, specimens with similar sequences group together. Ten sequences of the same rbcL region of similar species or genera to monoculture diatoms of this study were mined from GenBank to improve accuracy of relatedness inferred from the unrooted tree. The genetic distance matrix that was created when the unrooted tree was made, displays monocultures that were genetically similar to each other in dark-colored blocks with those that were more distantly related in light-colored blocks (Figure 3.1).

The unrooted tree indicated that all *Cylindrotheca* species of this study (AB1,2,3 and 10) grouped together with a *Nitzschia* species, NWU8 forming their own clade. These monocultures were all shown to be similar to *Nitzschia sigmoidea* and *Nitzschia frustulum* mined from GenBank (Figure 3.2). The *Psammodictyon* species of this study all grouped and formed their own clade together with *Psammodictyon constrictum* from published data. The remaining monocultures and mined GenBank sequences formed the last clade of the unrooted tree (Figure 3.2). *Navicula* sp. NWU7 showed close relation to mined data *Navicula salinicola* however, NWU6 *Navicula perminuta* was found to be distant but formed its own clade, thus, further taxonomic classification and molecular analysis is required. AB11, *Minutocellus* sp. showed close relation to *Minutocellus polymorphus* however, morphological features need to be further inspected to confirm species identification (Table 3.2 and Figure 3.2 and Supplementary material 1). *Entomoneis* sp. NWU1 showed close relation to *Entomoneis ornata,* forming their own clade, however, the morphology of NWU1 needs to be further analysed to provide species identification. *Amphora incrassata (*NWU5*) and Amphora* sp. (NWU10) showed similar relations to each other and mined GenBank data *Amphora libyca*. The low-quality sequences of *Skeletonema grethae* (AB12) were shown to be closely related to *Skeletonema marinoi* mined from GenBank. The Chaetoceros sp. of this study (AB13) was found to be closely related to *Skeletonema* sp. rather than *Chaetoceros* sp. mined from GenBank (Figure 3.2). This is of concern as morphology dictates AB13 is different to *Skeletonema* genus and most likely part of the *Chaetoceros* genus, thus, further molecular investigation of this monoculture is required. The consensus sequences of AB13 were of a high-quality, thus, it is unlikely that the sample was contaminated with AB12 as this specimen did not seem to produce good quality sequencing results (Tables 3.3 and 3.4). Based on the genetic distance model *Chaetoceros* sp. mined from GenBank are shown to be closely related to *Skeletonema grethae* (AB12) with values ranges between 0.091 – 0.103 (Figure 3.1), thus it may be that AB13 was a closer relation in the *Chaetoceros* genus to *Skeletonema*.



	AB1_C.c...	AB2_C.c...	AB3_C.c...	AB3_C.c...	AB8_P.p...	AB8_P.p...	AB10_C...	AB11_...	AB11_...	AB12_S....	AB12_S....	AB13_C...	Amphor...	Chaeto...	Chaeto...	Entomo...	Minutoc...	Navicul...	Nitzschi...	Nitzschi...	NWU1_...	NWU2_...	NWU5_...	NWU6_...	NWU6_...	NWU7_...	NWU8_...	NWU10...	Psamm...	Skeleto...
AB1_C.closterium (Ehre...																													
AB2_C.closterium_June	0.008																												
AB3_C.closterium_June	0.015	0.002																											
AB3_C.closterium_Sept_...	0	0.004	0.008																										
AB8_P.panduriforme va...	0.121	0.129	0.132	0.119																									
AB8_P.panduriforme va...	0.133	0.137	0.137	0.127	0.004																								
AB10_C.closterium_Sept	0.057	0.062	0.057	0.055	0.137	0.143																							
AB11_Minutocellus sp._J...	0.154	0.163	0.165	0.148	0.113	0.120	0.178																						
AB11_Minutocellus sp._...	0.155	0.160	0.166	0.147	0.116	0.124	0.175	0																					
AB12_S.grethae_Sept_LQ	0.135	0.142	0.143	0.122	0.109	0.107	0.151	0.068	0.055																				
AB12_S.grethae (Zingo...	0.102	0.084	0.094	0.124	0.105	0.107	0.148	0.061	0.056	0.011																			
AB13_Chaetoceros sp._J...	0.110	0.113	0.128	0.102	0.096	0.106	0.125	0.069	0.066	0.056	0.053																		
Amphora libyca (Ehren...	0.146	0.148	0.163	0.127	0.134	0.146	0.144	0.106	0.098	0.097	0.121	0.102																	
Chaetoceros cf. lorenzi...	0.153	0.167	0.160	0.157	0.130	0.143	0.177	0.086	0.088	0.099	0.103	0.111	0.117																
Chaetoceros socialis (L...	0.123	0.136	0.154	0.127	0.119	0.122	0.151	0.073	0.079	0.091	0.099	0.102	0.128	0.072															
Entomoneis ornata (Bai...	0.148	0.151	0.158	0.115	0.147	0.144	0.160	0.128	0.120	0.099	0.131	0.106	0.080	0.105	0.113														
Minutocellus polymorp...	0.161	0.170	0.180	0.156	0.129	0.143	0.189	0.029	0.038	0.100	0.117	0.092	0.082	0.085	0.067	0.091													
Navicula salinicola (Hus...	0.145	0.147	0.170	0.131	0.103	0.105	0.155	0.054	0.054	0.091	0.115	0.091	0.075	0.104	0.072	0.093	0.054												
Nitzschia frustulum (Küt...	0.117	0.135	0.143	0.111	0.112	0.119	0.111	0.118	0.127	0.128	0.122	0.096	0.086	0.118	0.122	0.141	0.141	0.126											
Nitzschia sigmoidea (Ni...	0.121	0.137	0.145	0.109	0.155	0.160	0.118	0.162	0.175	0.131	0.131	0.114	0.103	0.119	0.106	0.098	0.111	0.083	0.084										
NWU1_Entomoneis sp._...	0.133	0.131	0.129	0.127	0.120	0.115	0.125	0.088	0.087	0.079	0.079	0.069	0.074	0.088	0.109	0.065	0.087	0.095	0.088	0.110									
NWU2_P. panduriforme...	0.122	0.128	0.146	0.119	0.007	0.011	0.137	0.122	0.122	0.100	0.104	0.103	0.137	0.149	0.126	0.150	0.148	0.114	0.111	0.153	0.118								
NWU5_A. incrassata (Gi...	0.141	0.146	0.150	0.139	0.131	0.124	0.146	0.062	0.059	0.076	0.068	0.066	0.066	0.113	0.099	0.088	0.075	0.077	0.117	0.144	0.054	0.128							
NWU6_N.perminuta_June	0.093	0.102	0.113	0.122	0.128	0.136	0.113	0.081	0.083	0.096	0.060	0.092	0.124	0.126	0.106	0.126	0.124	0.124	0.121	0.138	0.091	0.136	0.107						
NWU6_N.perminuta_Sept	0.120	0.119	0.130	0.122	0.128	0.136	0.112	0.081	0.083	0.095	0.076	0.087	0.113	0.127	0.098	0.099	0.103	0.104	0.127	0.135	0.088	0.135	0.107	0					
NWU7_Navicula sp._June	0.104	0.107	0.108	0.106	0.077	0.082	0.117	0.073	0.074	0.087	0.075	0.102	0.084	0.090	0.058	0.118	0.089	0.043	0.107	0.157	0.109	0.088	0.092	0.082	0.079				
NWU8_Nitzchia sp. _June	0.082	0.088	0.086	0.113	0.125	0.138	0.116	0.143	0.141	0.119	0.088	0.096	0.137	0.159	0.134	0.152	0.162	0.144	0.090	0.134	0.111	0.120	0.131	0.101	0.128	0.117			
NWU10_Amphora sp._J...	0.124	0.129	0.124	0.117	0.131	0.141	0.135	0.116	0.110	0.111	0.104	0.095	0.064	0.117	0.111	0.092	0.118	0.083	0.092	0.126	0.096	0.128	0.071	0.120	0.120	0.094	0.131		
Psammodictyon constri...	0.133	0.149	0.156	0.123	0.014	0.015	0.136	0.121	0.126	0.119	0.119	0.110	0.095	0.128	0.097	0.105	0.093	0.077	0.103	0.081	0.124	0.021	0.136	0.142	0.140	0.090	0.140	0.136	
Skeletonema marinoi (S...	0.138	0.151	0.161	0.123	0.131	0.141	0.173	0.070	0.071	0.013	0.050	0.068	0.100	0.110	0.088	0.092	0.079	0.083	0.117	0.102	0.081	0.131	0.079	0.133	0.115	0.110	0.153	0.101	0.093

Figure 3.1. Genetic distance matrix created by the Geneious tree builder using a pairwise alignment for all sequence pairs. The alignment type was a global alignment with free end gaps using a 93% similarity to build the distance matrix. Similar sequences between monocultures displayed as dark-colored blocks and dissimilar sequences as light-colored blocks. Values in the table written in white signify close relationships between monocultures.

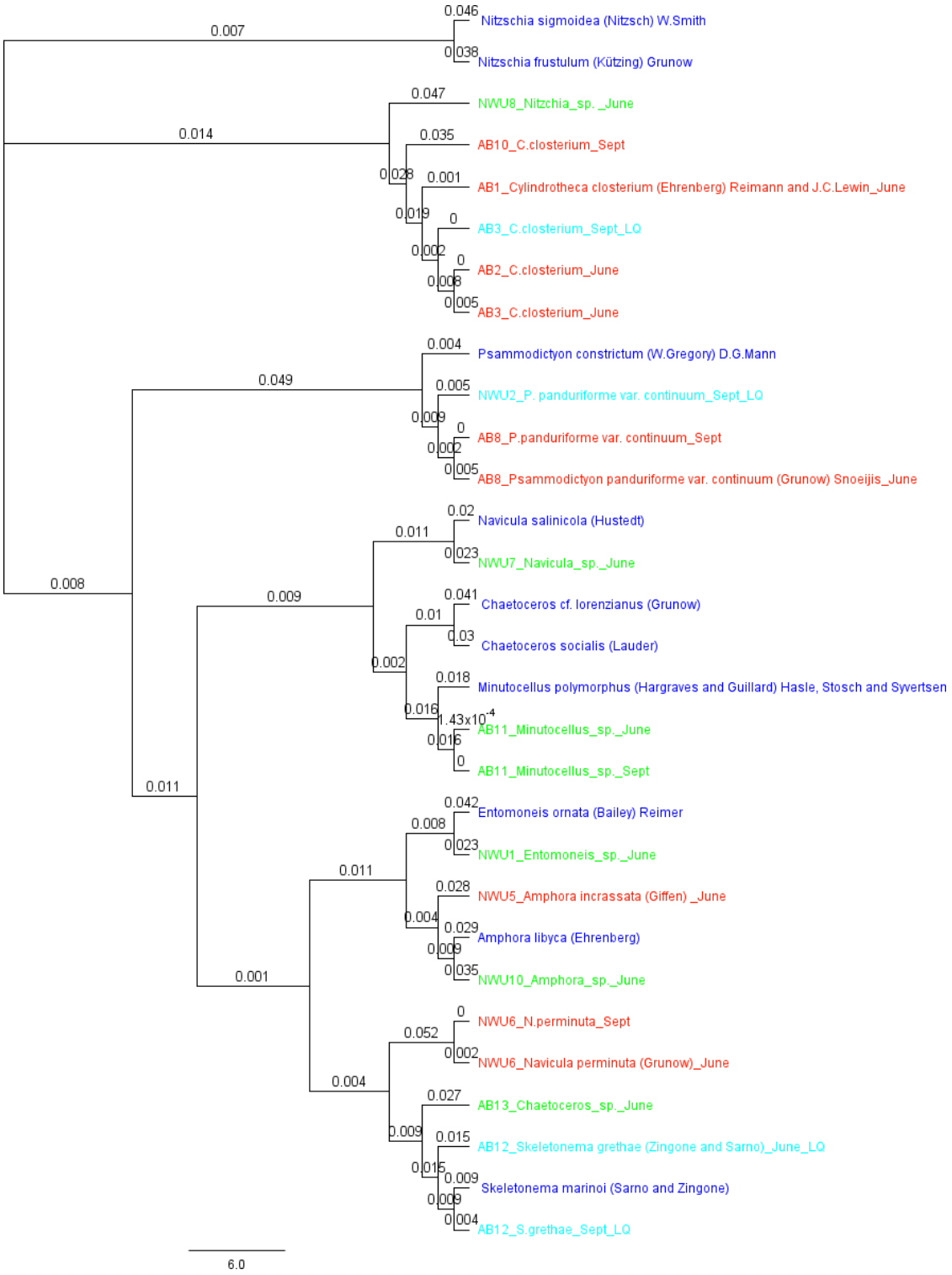


Figure 3.2. Unrooted neighbour-joining tree created in January 2022 using the Geneious tree builder to infer relatedness of the diatom monocultures. Branches are labelled with sample name, species or genus identification, sequence event and if it is low quality (LQ). Colour codes indicate the following: dark blue – sequences mined from GenBank to infer relatedness, green – monoculture specimen identified to genera, orange – monoculture specimen identified morphology and molecular data and turquoise – low quality molecular information that provides good quality results. The values on the branches signify the number of substitutes per site.

Overall, 10 diatom monocultures (AB1, 2, 3, 8, 10, 12 and NWU2, 3, 5, 6) were conclusively identified to species level using morphology and/or molecular data, seven diatoms (AB11, 13 and NWU1, 7, 8, 9, 10) had molecular and/or morphological information but could only be resolved to genus level as better images, molecular data and taxonomic knowledge are required to determine the species (Table 3.7, Supplementary material 2). The remaining two diatoms (AB5, 9) had live cultures available, but we were unable to generate good quality sequences to identify the rbcL barcode region during this study, thus, further molecular analysis is required. Conclusive identification of diatom monocultures (genus or species) using morphology as a basis and matching it to a barcode sequence resulted in the accurate identification of 12 diatom monocultures (AB1, 2, 8, 10, 11, 12, 13 and NWU1, 5, 7, 8, 10) (Table 3.7, Supplementary material 2).

3.4. Barcode of Life Data Systems submission and analysis

Monoculture sample details were uploaded to project DAB (Diatoms of Algoa Bay) on the BOLD interface. Trace files of each sequencing event were uploaded to each specimen file. A total of 17 sequences were accepted by BOLDs quality check with those not accepted due to poor quality trace files. Of the 17 accepted, five were excluded as morphology was not captured due to the monocultures dying or not being able to maintain as a monoculture (AB4, AB7, AB9, AB14 and NWU4). Upon species identification, records were updated, and image plates created per monoculture were uploaded to each specimen file on BOLD (Supplementary material 1). Once the data on BOLD were complete, the data were made public and submitted to GenBank to obtain accession numbers for the 12 accurately identified diatom monocultures (Table 3.7).

Table 3.7 Monocultures species identification by morphology and molecular approaches and conclusive species identification when possible. Monocultures with good quality images and barcodes had data published and accession numbers obtained.

Sample ID	Morphology species ID	Molecular species ID	Species/Genus ID	Monoculture accession number
AB1	*C. closterium*	*C. closterium*	*C. closterium*	OL794286
AB2	*C. closterium*	*C. closterium*	*C. closterium*	OL794283
AB3	*C. closterium*	*C. closterium*	*C. closterium*	-
AB5	*S. turris*	-	*S. turris*	-
AB8	*P. panduriforme* var. *continuum*	*P. panduriforme* var. *continuum*	*P. panduriforme* var. *continuum*	OL794285
AB9	*P. panduriforme* var. *continuum*	-	*P. panduriforme* var. *continuum*	-
AB10	*C. closterium*	*C.closterium*	*C. closterium*	OL794289
AB11	Minutocellus sp.	*M.polymorphus*	Minutocellus sp.	OL794292
AB12	*S. grethae*	*S. grethae*	*S. grethae*	OL794281
AB13	Chaetoceros sp.	*T. pseudonana*	Chaetoceros sp.	OL794282
NWU1	Entomoneis sp.	Entomoneis sp.	Entomoneis sp.	OL794287
NWU2	*P. parduriforme var. continuum*	*P. parduriforme var. continuum*	*P. parduriforme var. continuum*	-
NWU3	*T. velata*	-	*T. velata*	-
NWU5	A. incrassata	Amphora sp.	A. incrassata	L794288
NWU6	*N. perminuta*	*N. perminuta*	*N. perminuta*	-
NWU7	Navicula sp.	Navicula sp.	Navicula sp.	OL794291
NWU8	Nitzchia sp.	Nitzchia sp.	Nitzchia sp.	OL794284
NWU9	Navicula sp.	-	Navicula sp.	-
NWU10	Amphora sp.	Amphora sp.	Amphora sp.	OL794290

Chapter 4. Discussion

The phytoplankton community of an environment is vital for its functioning and stability, thus aiding in the provision of ecosystem services (McCann, 2000; Ptacnik *et al*., 2008; Cermeño *et al*., 2016; Busseni, 2019; Dutkiewicz *et al*., 2020). Diatoms generally dominate phytoplankton communities and respond rapidly to environmental changes making them ideal bioindicators (De la Rey *et al*., 2004; Morin *et al*., 2008; Verheye *et al*., 2017; Dalu *et al*., 2020). This study aimed to contribute to the molecular knowledge of using diatoms in monitoring coastal waters by isolating, culturing, imaging, and barcoding diatoms species collected within Algoa Bay, South Africa. Furthermore, diatom monocultures were included from Dr. Majewska's collection (North-West University) to increase specimens identified in this study. Ultimately this study resulted in an easy-to-follow procedure for the development of barcode reference libraries to aid in species identification of diatoms for management purposes.

4.1. Species significance

Cylindrotheca closterium (Ehrenberg) Reimann and Lewin (AB1, 2, 3, 10)

This diatom species is known to be present within Algoa Bay and is a known planktonic and benthic species (Reimann and Lewin 1964; An *et al*.,2017; Klapper *et al*., 2021; Nunes *et al*., 2021). This is a very robust species of diatom and has already been used as a model in multiple eco-physiological studies showing its ability to adapt to thermal and saline changes (Najdek *et al*. 2005; Apoya-Horton *et al*. 2006; Stock *et al*., 2019). This species of diatom has been extensively studied with its life cycle and reproductive behaviour well established (Vanormelingen *et al*., 2013; Klapper *et al*., 2021), and it is sought after for its high lipid production that may be used for food, feed for aquaculture, nutraceutical products and biofuels. *C. closterium* is also a known harmful species as they produce mucilage and allelopathic substances that may affect organisms around them (Guiry and Guiry, 2021; Zhang *et al*., 2021).

Psammodictyon panduriforme var. *continuum* (Grunow) Snoeijis (AB8, 9 and NWU2)

This species has not been used as a bioindicator because of its cosmopolitan nature and wide tolerance ranges for a variety of environmental variables. According to studies from the North-West Balkan Peninsula and the Red Sea, this species was amongst the most common from sampling sessions. Both studies determined the

relationship between diatoms and environmental variables. Each Canonical Correspondence Analysis conducted found that this species was not influenced by the environmental variables tested (silica, total inorganic nitrogen (TIN), nitrate, nitrite, ammonium, phosphate, oxygen saturation, salinity, chlorophyll *a* concentration and temperature) (Hafner *et al*., 2018; Zalat *et al*., 2021).

Navicula perminuta (Grunow) (NWU6)

Navicula perminuta is a widely distributed species reported across numerous ocean studies (Al-Handal and Wulff, 2008; Stachura-Suchoples *et al*., 2015; Guiry and Guiry, 2021). The morphology and molecular data presented in this study allows this species to be easily identified globally. This species is a known marine and brackish water species that is motile and epiphytic, being found on various macroalgae. It is also known to show directional response with active movement towards light and enhanced lipid utilisation under higher temperatures. This species is also an important source of food for Antarctic krill species *Euphausia superba* (Dana, 1850) (Al-Handal and Wulff, 2008; McLachlan *et al*., 2009; Schaub *et al*., 2017; Guiry and Guiry, 2021), thereby making it an important species to monitor as a bioindicator of the health of the lower trophic Antarctic ecosystem. Its use as a bioindicator in South African coastal waters still need to be established.

Trachyneis velata (Schmidt) (NWU3)

Trachyneis velata is not a well-researched species but has been reported in oceans worldwide (Cleve, 1984; Giffen, 1971; Park *et al*., 2018; Siqueiros-Beltrones, *et al*., 2017). There is no known bioindicator value of this species, more research is required to find out its exact role in marine and coastal ecosystems. The images generated in this study can be used as a reference for morphology; however, further molecular work is required.

Amphora incrassata (Giffen) (NWU5)

This species has been previously reported in Algoa Bay and in other parts of the world (Giffen, 1973; Cotiyane-Pondo *et al*., 2021; Guiry and Guiry, 2021; Risjani *et al*., 2021). Based on work done by Cotiyane-Pondo *et al*. (2021), this species was rarely found in planktonic samples and shown to be slightly influenced by dissolved inorganic nitrogen and temperature. Further research is required to determine if this species has the potential to be a bioindicator.

Stephanopyxis turris (Greville) Ralfs (AB5)

There is no known biomonitoring value of this species, but it has shown importance in various industrial studies (Machill *et al*., 2013; Pytlik *et al*., 2017 and 2019). This species has been shown to take up aluminium and gold, potentially affecting higher trophic species due to bioaccumulation of these substances in the food web in polluted waters. Accurately determining this species' barcode will allow for its easy detection in biomonitoring surveys and accurate identification in other research fields.

Skeletonema grethae (Zingone and Sarno) (AB12)

This species was only recently distinguished from others in its genus by Sarno *et al.* (2005). According to Kooistra *et al*. (2008) this species was exclusively found along the North American coastline, however, Bergesch *et al.* (2009) found this species in the south-western Atlantic Ocean. With this species previously known as *Skeletonema costatum* and part of a species complex, it may have been identified as such in the past. It is important to note that this species has been broadly recognised as a bloom-forming species, additionally this species has been associated with areas of upwelling (a regular occurrence in Algoa Bay), cultivated on commercial scales, and used in molecular N-assimilation experiments. No published records have been found for this species within South African waters, thus, this study would be the first to report it in Algoa Bay. The molecular analysis yielded an average quality barcode for this species that can be used with caution to identify it. It role within the Algoa bay ecosystem needs to be established and thus may be used as bioindicator species in the future.

The abovementioned species may not be conclusive bioindicators, however, the information provided in this study means that future studies will be able to accurately identify these species by morphology and molecular data. Several monocultures of this study have been identified to genus level and will soon be identified to species and future molecular analysis will further improve barcodes.

4.2. *Morphology and molecular identification workflow*

Various laboratory procedures were carried out during this study with different instruments used to image specimens and cultures from different months were imaged and sequenced (Doyle and Dickson, 1987; Harding and Taylor, 2011; Elias *et al.,* 2017; Witkowski *et al.,* 2020; Cotiyane- Pondo and Bornman 2021).

Environmental samples were obtained twice during this study (28 January 2021 and 16 March 2021), but monocultures were only created from January environmental samples. Due to the COVID-19 pandemic, environmental samples were couriered overnight to Johannesburg from Gqeberha. Samples collected in March took two days to get to Johannesburg resulting in the death and decay of some of the diatoms in the environmental sample whilst in transit, which ultimately resulted in no successful monocultures being created. The January samples arrived within a day and were successfully isolated on 29th of January 2021. This is an indication that environmental samples should be processed for isolation immediately or within 24 hours of environmental sampling to increase the percentage success of culture creation (Rosy, 2021).

The culture collection maintenance method, of using Diago's IMK media and bi-monthly subsampling whilst keeping cultures under incubator conditions, is easy to follow and resulted in the continual survival of most monocultures. Thus far, culture maintenance has lasted one year with minimal effort, resulting in the potential for long-term morphology and molecular studies on diatom culture collections.

The laboratory procedures carried out during this study for diatoms resulted in their morphological and molecular identification (Doyle and Dickson, 1987; Harding and Taylor, 2011; Elias *et al.,* 2017; Witkowski *et al.,* 2020; Cotiyane-Pondo and Bornman, 2021). Furthermore, SEM analyses of different cultures (July and September 2021) displayed similar size cells and no major changes in cell shape thus showing minimal changes in monocultures during this study (Table 3.2 and Supplementary material 1). However, as this study did not focus on the morphological variations of monocultures further investigation is required to determine conclusive results for these species.

The circumstances of this study did not allow for direct comparison between digestion methods and SEM instruments, however, the images taken in this study still provided evidence that an affordable desktop SEM can provided images to aid in accurate species identification although a more powerful and expensive SEM can be used to better analyse fine morphology features on different diatom species.

The two methods of organic matter acid digestion used in this study resulted in good quality images, however, the nitric acid method is recommended for monocultures as it produced cleaner frustules. This method was also found to be less aggressive and

allowed for well-kept siliceous frustules of delicate diatoms (Elias *et al.,* 2017; Cotiyane-Pondo and Bornman, 2021). Both acid digestion methods should be used interchangeably and amended if necessary, depending on the type of siliceous structure of the diatom and if the analysis is being carried out on an environmental or monoculture sample (Harding and Taylor, 2011; Elias *et al.,* 2017; Cotiyane-Pondo and Bornman, 2021).

The molecular identification procedure used during this study are recommended for diatoms and primers chosen depending on the gene region of interest (Doyle and Dickson, 1987; Zimmermann *et al*., 2011 and 2015; Vasselon *et al*., 2017). If DNA extraction results are inconclusive it is recommended that the amount of monoculture volume used in centrifugation prior to extraction is increased.

4.3. Environmental monitoring

In the face of global environmental change, the requirement to sustainably manage the environment is an imperative. Effective monitoring is required to determine thresholds of potential concern and manage the environment sustainably (Gaylard and Ferreira, 2011; Kingsford *et al*., 2011; Vallaeys *et al*., 2017; Busseni, 2019; Swanepoel and Sauka, 2019). To monitor the environment, various abiotic and biotic factors need to be recorded. Physical variables are primarily obtained through *in situ* measurements and remote sensing imagery, chemical and biological variables are obtained through sample collection and laboratory work (Chapman and Wang, 2001; De la Rey *et al*., 2004; Johnson *et al*., 2011; Vallaeys *et al*., 2017; Dorrington *et al*., 2018; Swanepoel and Sauka, 2019). With biomonitoring used as an alternative to chemical analysis, species detection of bioindicator species within a given environment are required to allow for accurate detection of changes in abiotic factors (pH, temperature, salinity, turbidity, nutrients and toxin concentrations, as well as identifying introduced pollutants and heavy metals) (Chapman and Wang, 2001; De la Rey *et al*., 2004; Vallaeys *et al*., 2017). The use of environmental DNA for monitoring requires comprehensive reference libraries of DNA barcodes of bioindicator species of a specific environment (Bode *et al*., 2017; Ruppert *et al*., 2019; Tapolczai *et al*., 2019).

Diatom bioindicator species have yet to be defined for most parts of the ocean including Algoa Bay. However, contributing accurately identified specimens to an open-source library (BOLD) increases the percentage of species identified in eDNA

surveys globally. Specimens of this study have a ~250 base pair region and associated images of diatom species to ensure accurate identification. The building of reference libraries of bioindicator species will improve results of eDNA environmental monitoring efforts. The inclusion of eDNA analysis may reduce cost in the future of biomonitoring and provide a holistic view of community composition within the target area.

4.4. Recommendations and future steps

Displaying the silica frustules of diatoms is important for their accurate identification by morphology. The nitric acid digestion method used in this study should be, tested and possibly developed further on environmental samples to see if the protocol can be made more time-efficient, particularly with regards to the use of the shaker (Elias *et al.,* 2017; Cotiyane-Pondo and Bornman, 2021). This method is less time consuming and requires fewer steps for preparation to obtain high quality images compared to the hydrochloric acid digestion method.

Molecular identification of diatoms occurs using the ribulose 1,5–bisphosphate carboxylase/oxygenase plastid gene region (rbcL), and the primers used in the study successfully amplifies the region to greater than 250 nucleotide base pairs (Zimmermann *et al*., 2015; Vasselon *et al*., 2017). However, for an increase in accuracy, the barcode reference library should be built with barcode sequences of more than 500 nucleotide base pairs (Ratnasingham and Hebert, 2007). Alverson *et al*. (2007) and An *et al*. (2017) used rbcL primer pairs that amplified ~1470 and ~1600 nucleotide base pair regions, respectively, that covers the rbcL gene and spacer region between rbcL and rbcS. Either one of these primer pairs would be the recommended pair for the creation of a diatom barcode reference library and should be used when barcoding diatoms.

4.5. Conclusions

Sustainably managing oceanic environments is vital for the planet, with the ocean systems and coastal waters providing vast amounts of ecosystem services locally and globally; understanding the complex interactions within them and managing them is required for the future of the planet (Ingole *et al*., 2005; Kato *et al*., 2011; Mohanty *et al*., 2015; Momber *et al*., 2016; Vallaeys *et al*., 2017; Berry *et al*., 2019; Busseni, 2019). Humans contribute to increasing pressure on the ocean environment through global environmental change, where the increasing population results in other impacts such

as ocean pollution, changes in local and global ocean pH, resource overutilization, and increased greenhouse gas production leading to global warming (Knap *et al*., 2002; Vallaeys *et al*., 2017; Häder *et al*., 2020; Kalenik, 2020). Humans also have the capacity to mitigate the impact or provide solutions to the global change issues by working collaboratively and in an integrative manner, with each coastal or ocean zone management unit managing its area sustainably and contributing to global knowledge of the interactions between abiotic and biotic factors through open-source libraries and publicly available data.

This study contributed diatom species identification by a combination of morphology and molecular data to improve their use as a biomonitoring tool within coastal waters. The procedures and workflows utilised in this study can be adapted and used for many marine organisms in the development of barcode reference libraries. To ensure that data are shared globally, standardised protocols should be used in the generation of barcode libraries for specific organism types. This study made a small, but significant contribution to the global barcode reference libraries for marine species identification. By working together, we can create the ‘Ocean we need for the future we want’ - United Nations Secretary-General, António Guterres (Intergovernmental Oceanographic Commission, 2019).

References

Al-Handal, A. Y. and Wulff, A. 2008. Marine epiphytic diatoms from the shallow sublittoral zone in Potter Cove, King George Island, Antarctica. *Botanica Marina*, 51, 411-435.

Al-Kandari, M., Al-Yamani, F. Y. and Al-Rifaie, K. 2009. Marine Phytoplankton Atlas Kuwait's waters, Kuwait Institute for Scientific Research. 351.

Alverson, A. J., Jansen, R. K. and Theriot, E. C. 2007. Bridging the Rubicon: phylogenetic analysis reveals repeated colonizations of marine and fresh waters by thalassiosiroid diatoms. *Molecular Phylogenetics and Evolution*, 45, 1, 193-210.

Al-Yamani, F. Y. A. S. and Maria A. 2019. Marine Phytoplankton of Kuwait's waters Volume 2 Diatoms, Kuwait Institute for Scientific Research, Kuwait.

An, S. M., Choi, D. H., Lee, J. H., Lee, H. and Noh, J. H. 2017. Identification of benthic diatoms isolated from the eastern tidal flats of the Yellow Sea: Comparison between morphological and molecular approaches. *PLoS One*, 12, 6, e0179422.

Apoya-Horton, M. D., Yin, L., Underwood, G. J. C. and Gretz, M. R. 2006. Movement Modalities and Responses to Environmental Changes of the Mudflat Diatom Cylindrotheca Closterium (Bacillariophyceae). *Journal of Phycology*, 42, 2, 379-390.

Ashworth, M. P., Nakov, T. and Theriot, E. C. 2013. Revisiting Ross and Sims (1971): toward a molecular phylogeny of the *Biddulphiaceae* and *Eupodiscaceae* (Bacillariophyceae). *Journal of Phycology*, 49, 6, 1207-1222.

Benedetti, F., Ayata, S. D., Irisson, J. O., Adloff, F., Guilhaumon, F. and Essl, F. 2019. Climate change may have minor impact on zooplankton functional diversity in the Mediterranean Sea. *Diversity and Distributions*, 25, 568-581.

Bergesch, M., Garcia, M. and Odebrecht, C. 2009. Diversity and Morphology of Skeletonema Species in Southern Brazil, Southwestern Atlantic Ocean. *Journal of Phycology*, 45, 6, 1348-1352.

Berry, T. E., Saunders, B. J., Coghlan, M. L., Stat, M., Jarman, S., Richardson, A. J., Davies, C. H., Berry, O., Harvey, E. S. and Bunce, M. 2019. Marine environmental DNA biomonitoring reveals seasonal patterns in biodiversity and identifies ecosystem responses to anomalous climatic events. *PLoS genetics*, 15, 2, e1007943.

BIOSCAN, 2022. International Barcode of Life: Illuminate Biodiversity. Available at: https://ibol.org/. (Accessed: 20 January 2022).

Bode, M., Laakmann, S., Kaiser, P., Hagen, W., Auel, H. and Cornils, A. 2017. Unravelling diversity of deep-sea copepods using integrated morphological and molecular techniques. *Journal of Plankton Research*, 39, 4, 600-617.

Bowler, C., Allen, A.E., Badger, J.H., Grimwood, J., Jabbari, K., Kuo, A., Maheswari, U., Martens, C., Maumus, F., Otillar, R.P. and Rayko, E., 2008. The Phaeodactylum genome reveals the evolutionary history of diatom genomes. Nature, 456(7219), pp.239-244.

Brun, P., Stamieszkin, K., Visser, A. W., Licandro, P., Payne, M. R. and Kiorboe, T. 2019. Climate change has altered zooplankton-fueled carbon export in the North Atlantic. *Nature Ecology and Evolution*, 3, 3, 416-423.

Bucklin, A., Ortman, B. D., Jennings, R. M., Nigro, L. M., Sweetman, C. J., Copley, N. J., Sutton, T. and Wiebe, P. H. 2010. A “Rosetta Stone” for metazoan zooplankton: DNA barcode analysis of species diversity of the Sargasso Sea (Northwest Atlantic Ocean). Deep Sea Research Part II: *Topical Studies in Oceanography*, 57, 24-26, 2234-2247.

Busseni, G. 2019. Exploring diatoms functional and taxonomic diversity on a global scale. Doctor of Philosophy in Life, Health and Chemical Sciences, Open University.

Centre for Biodiversity Genomics, 2022. Planetary Biodiversity Mission. Available at: https://biodiversitygenomics.net/. (Accessed: 20 January 2022).

Cermeño, P., Chouciño, P., Fernández-Castro, B., Figueiras, F. G., Marañón, E., Marrasé, C., Mouriño-Carballido, B., Pérez-Lorenzo, M., Rodríguez-Ramos, T., Teixeira, I. G. and Vallina, S. M. 2016. Marine Primary Productivity Is Driven by a Selection Effect. *Frontiers in Marine Science*, 3, 173.

Cermeño, P., Rodríguez-Ramos, T., Dornelas, M., Figueiras, F. G., Marañón, E., Teixeira, I. G. and Vallina, S. 2013. Species richness in marine phytoplankton communities is not correlated to ecosystem productivity. *Marine Ecology Progress Series*, 488, 1-9.

Chapman, P. M. and Wang, F. 2000. Issues in Ecological Risk Assessment of Inorganic Metals and Metalloids. Human and Ecological Risk Assessment: An International Journal, 6, 6, 965-988.

Cholnoky, B.J., 1960. The relationship between algae and the chemistry of natural waters. CSIR.

Cleve, P. T. 1894 Synopsis of The *Naviculoid* Diatoms, PA Norstedt and söner. 26, 2-3.

Cotiyane-Pondo, P. and Bornman, T. G. 2021. Environmental Heterogeneity Determines Diatom Colonisation on Artificial Substrata: Implications for Biomonitoring in Coastal Marine Waters. *Frontiers in Ecology and Evolution*, 9, 877.

Cotiyane-Pondo, P., Bornman, T. G. and Dąbek, P. 2021. Insight into nearshore diatom assemblages from an island ecosystem: Composition, spatial variation, and benthic contribution to the pelagic community (Bird Island, South Africa). *Regional*

Studies in Marine Science, 44, 101762.

Dalu, T., Lemley, D. A., Snow, G. C. and Wu, N. 2020. Linking phytoplankton community structure to aquatic ecosystem functioning: A mini-review of the current status and future directions. Handbook of Algal Science, *Technology and Medicine*, 291-302.

De La Rey PA, T. J., Laas A, Van Rensburg L. and Vosloo A 2004. Determining the possible application value of diatoms as indicators of general water quality: A comparison with SASS 5. *Water SA*, 30, 3, 325-332.

De Luca,D., Piredda,R., Sarno,D. and Kooistra,W.H.C.F., 2019. A multigene phylogeny to infer the evolutionary history of Chaetocerotaceae (Bacillariophyta). Unpublished.

Dorrington, R. A., Vrancken, P., Snow, B., Roberts, M., Porri, F., Perissinotto, R., Perissinotto, R., Paterson, A., Parker-Nance, S., Mcquaid, C., Matcher, G. F., Mahler-Coetzee, J., Liu, K., Goschen, W. S., Deyzel, S. H. P., Cawthra, H. C., Adams, J. B., Bornman, T. G. and Lombard, A. T. 2018. Working together for our oceans: A marine spatial plan for Algoa Bay, South Africa. *South African Journal of Science*, 114, 3-4, 1- 6.

Doyle, J. J. and Dickson, E. E. 1987. Preservation of Plant Samples for DNA Restriction Endonuclease Analysis. *Taxon*, 36, 4, 715-722.

Dutkiewicz, S., Cermeno, P., Jahn, O., Follows, M. J., Hickman, A. E., Taniguchi, D. A. A. and Ward, B. A. 2020. Dimensions of marine phytoplankton diversity. *Biogeosciences*, 17, 3, 609-634.

Elias, C. L., Rocha, R. J. M., Feio, M. J., Figueira, E. and Almeida, S. F. P. 2017. Influence of the colonizing substrate on diatom assemblages and implications for bioassessment: a mesocosm experiment. *Aquatic Ecology*, 51, 1, 145-158.

Falkowski, P. G., Barber, R. T. and Smetacek, V. V. 1998. Biogeochemical Controls and Feedbacks on Ocean Primary Production. *Science*, 281, 5374, 200-206.

Fuhrman, J. A. 2009. Microbial community structure and its functional implications. *Nature*, 459, 7244, 193-199.

Gaylard, A. and Ferreira, S. 2011. Advances and challenges in the implementation of strategic adaptive management beyond the Kruger National Park – Making

linkages between science and biodiversity management. Koedoe: African Protected Area Conservation and Science, 53, 2,1-8.

Giffen, M. H. 1971. Marine Littoral Diatoms from the Gordon's Bay, Region of False Bay Cape Province, South Africa. *Botanica Marina*, 14, 1-16.

Giffen, M. H. 1973. Diatoms of the Marine Littoral of Steenberg's Cove in St. Helena Bay, Cape Province, South Africa. *Botanica Marina*, 16, 16.

Giffen, M. H. 1976. A Further Account of the Marine Littoral Diatoms of the Saldanha Bay Lagoon, Cape Province. South Africa. *Botanica Marina*, 19, 15.

Guidi, L., Stemmann, L., Jackson, G. A., Ibanez, F., Claustre, H., Legendre, L., Picheral, M. and Gorskya, G. 2009. Effects of phytoplankton community on production, size, and export of large aggregates: A world-ocean analysis. Limnology and *Oceanography*, 54, 6, 1951-1963.

Guiry, M.D. and Guiry, G.M. 2021. *AlgaeBase*. World-wide electronic publication, National University of Ireland, Galway. https://www.algaebase.org; searched on October 2021.

Häder, D. P., Banaszak, A. T., Villafane, V. E., Narvarte, M. A., Gonzalez, R. A. and Helbling, E. W. 2020. Anthropogenic pollution of aquatic ecosystems: Emerging problems with global implications. *Science of the Total Environment*, 713, 136586.

Hafner, D., Car, A. N. A., Jasprica, N., Kapetanović, T. and Dupčić Radić, I. 2018. Relationship between marine epilithic diatoms and environmental variables in oligotrophic bay, NE Mediterranean. *Mediterranean Marine Science*, 19, 2, 223-239.

Harding, W. R. and Taylor, J. C. 2011. The South African Diatom Index (SADI) A preliminary index for indicating water quailty for rivers and streams of Southern Africa. South Africa: DH Environmental Consulting, Somerset West, South Africa, Department of Biological Sciences, North-West University, Potchefstroom, South Africa.

Hasle, G. R., Syvertsen E.E., Steidinger K.A., Tangen K. and Carmelo R. Tomas 1996. Identifying Marine Diatoms and Dinoflagellates. *Elsevier*.

Hebert, P. D., Cywinska, A., Ball, S. L. and Dewaard, J. R. 2003a. Biological identifications through DNA barcodes. Proceedings of the Royal Society of London. *Series B: Biological Sciences*, 270 (1512), 313-321.

Hebert, P. D., Ratnasingham, S. and Dewaard, J. R. 2003b. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. Proceedings of the Royal Society of London. *Series B: Biological Sciences*, 270, Suppl 1, 96-99.

Hop, H., Wold, A., Vihtakari, M., Daase, M., Kwasniewski, S., Gluchowska, M., Lischka, S., Buchholz, F. and Falk-Petersen, S. 2019. Zooplankton in Kongsfjorden (1996–2016) in Relation to Climate Change. The Ecosystem of Kongsfjorden, Svalbard, *Springer*, Cham, 229-300.

Ingole, B. S., Goltekar, R., Gonsalves, S. and Ansari, Z. A. 2013. Recovery of Deep-sea Meiofauna after Artificial Disturbance in the Central Indian Basin. Marine *Georesources and Geotechnology*, 23, 253-266.

Intergovernmental Oceanographic Commission, 2019. The science we need for the ocean we want: the United Nations decade of ocean science for sustainable development (2021–2030). In Program and meeting document.

Johnson, C. L., Runge, J. A., Curtis, K. A., Durbin, E. G., Hare, J. A., Incze, L. S., Link, J. S., Melvin, G. D., O'brien, T. D. and Van Guelpen, L. 2011. Biodiversity and ecosystem function in the Gulf of Maine: pattern and role of zooplankton and pelagic nekton. *PLoS One*, 6, 1, e16491.

Kalenik, V. 2020. Ocean pollution as a global problem of modern humanity.

Kato, Y., Fujinaga, K., Nakamura, K., Takaya, Y., Kitamura, K., Ohta, J., Toda, R., Nakashima, T. and Iwamori, H. 2011. Deep-sea mud in the Pacific Ocean as a potential resource for rare-earth elements. *Nature Geoscience*, 4, 535-539.

Kermarrec, L., Bouchez A., Rimet F., and Humbert JF. 2013. First Evidence of the Existence of Semi-Cryptic Species and of a Phylogeographic Structure in the *Gomphonema parvulum* (Kützing) Kützing Complex (Bacillariophyta). *Protist*, 164, 20.

Kingsford, R. T., Biggs, H. C. and Pollard, S. R. 2011. Strategic Adaptive Management in freshwater protected areas and their rivers. *Biological Conservation*, 144, 1194-1203.

Klages, N.T.W. and Bornman, T.G., 2005. Pot of Ngqura Marine Biomonitoring Programme. Annual Report 2004 – 2005. Institute for Environmental and Coastal Management Report C, 86.

Klapper, F., Audoor, S., Vyverman, W. and Pohnert, G. 2021. Pheromone Mediated Sexual Reproduction of Pennate Diatom *Cylindrotheca closterium*. *Journal of Chemical Ecology*, 47, 504-512.

Knap, A., Dewailly, E., Furgal, C., Galvin, J., Baden, D., Bowen, R. E., Depledge, M., Duguay, L., Fleming, L. E., Ford, T., Moser, F., Owen, R., Suk, W. A. and Unluata, U. 2002. Indicators of ocean health and human health: developing a research and monitoring framework*. Environment Health Perspective*, 110, 839-45.

Kooistra, W. H., Sarno, D., Balzano, S., Gu, H., Andersen, R. A. and Zingone, A. 2008. Global diversity and biogeography of *Skeletonema* species (bacillariophyta). *Protist*, 159, 177-93.

Laakmann, S., Gerdts, G., Erler, R., Knebelsberger, T., Martinez Arbizu, P. and Raupach, M. J. 2013. Comparison of molecular species identification for North Sea calanoid copepods (Crustacea) using proteome fingerprints and DNA sequences. *Molecular Ecology Resources*, 13, 5, 862-876.

Li,H., Yang,G. and Zhang,X., 2004.Isolation and molecular identification of two strains of Cylindrotheca Closterium. Unpublished.

López-Fuerte, F. O., Siqueiros Beltrones, D. A. and Altamirano-Cerecedo, M. D. C. 2020. Species Composition and New Records of Diatom Taxa on Phyllodictyon pulcherrimum (Chlorophyceae) from the Gulf of California. *Diversity*, 12, 9, 339.

Macgillivary, M. L. and Kaczmarska, I. 2011. Survey of the efficacy of a short fragment of the rbcL gene as a supplemental DNA barcode for diatoms. *Journal of Eukaryotic Microbiology*, 58,6,529-536

Machill, S., Kohler, L., Ueberlein, S., Hedrich, R., Kunaschk, M., Paasch, S., Schulze and Brunner, E. 2013. Analytical studies on the incorporation of aluminium in the cell walls of the marine diatom *Stephanopyxis turris*. *Biometals*, 26, 1, 141-150.

Majewska, R., Ashworth M.P., Bosak S., Goosen W.E., Nolte C., Filek K., Van De Vijver B., Taylor J.C., Manning S.R. and Nel R. 2021. On Sea Turtles-associated *Craspedostauros* (Bacillariophyta), with Description of three novel species. *Journal of Phycology*, 57, 1,119-218.

Mann,D.G., Trobajo,R., Sato,S., Li,C.L., Witkowski,A., Rimet,F., Ashworth,M.P., Hollands,R.M. and Theriot,E.C., 2020. Reassessment of diatom family Bacillariaceae: four-gene dataset reveals extensive cryptic diversity, homoplasy and conficts with previous classifications. Unpublished

Margalef, R. 1978. Life-forms of phytoplankton as survival alternatives in an unstable environment. *Oceanologica acta*, 1, 4, 493-509.

Mbambo, S. W. 2014. Scales of variability of phytoplankton composition and biomass in Algoa Bay, South Africa. Master of Science, University of Cape Town.

McCann, K. S. 2000. The diversity-stability debate. *Nature*, 405, 6783, 228-233.

McLachlan, D. H., Brownlee, C., Taylor, A. R., Geider, R. J. and Underwood, G. J. 2009. Light-Induced Motile Responses of the Estuarine Benthic Diatoms *Navicula Perminuta* and *Cylindrotheca closterium* (Bacillariophyceae). *Journal of Phycology*, 45, 3, 592-599.

Mejdandzic, M., Bosak, S., Nakov, T., Ruck, E., Orlic, S., Gligora Udovic, M., Peharec Stefanic, P., Spoljaric, I., Mrsic, G. and Ljubesic, Z. 2018. Morphological diversity and phylogeny of the diatom genus *Entomoneis* (Bacillariophyta) in marine plankton: six new species from the Adriatic Sea. *Journal of Phycology*, 54, 2, 275-298.

Mohanty, S.K., Dash, P., Gupta, A. and Gaur, P., 2015. Prospects of Blue Economy in the Indian Ocean. *Research and Information System for Developing Countries*, New Delhi.

Mohrbeck, I., Raupach, M. J., Martinez Arbizu, P., Knebelsberger, T. and Laakmann, S. 2015. High-Throughput Sequencing - The Key to Rapid Biodiversity Assessment of Marine Metazoa? *PLoS One*, 10, 10, e0140342.

Momber, A. W., Plagemann, P. and Stenzel, V. 2016. The adhesion of corrosion protection coating systems for offshore wind power constructions after three years under offshore exposure. *International Journal of Adhesion and Adhesives*, 65, 96-101.

Morin, S., Coste, M. and Delmas, F. 2008. A comparison of specific growth rates of periphytic diatoms of varying cell size under laboratory and field conditions. *Hydrobiologia*, 614, 1, 285-297.

Najdek, M., Blažina, M., Djakovac, T. and Kraus, R. 2005. The role of the diatom *Cylindrotheca closterium* in a mucilage event in the northern Adriatic Sea: coupling with high salinity water intrusions. *Journal of Plankton Research*, 27, 9, 851-862.

Nunes, M., Adams, J. B., Van Aswegen, S. and Matcher, G. F. 2019. A comparison between the morphological and molecular approach to identify the benthic diatom community in the St Lucia Estuary, South Africa. *African Journal of Marine Science*, 41, 4, 429-442.

Nunes, M., Lemley, D. A., Matcher, G. F. and Adams, J. B. 2021. The influence of estuary eutrophication on the benthic diatom community: a molecular approach. *African Journal of Marine Science*, 1-16

Park, J. S., Lobban, C. S. and Lee, K.-W. 2018. Diatoms associated with seaweeds from Moen Island in Chuuk Lagoon, Micronesia. *Phytotaxa*, 351, 2, 101-140.

Powell, M., van der Bank, M., Maurin, O. and Savolainen, V. 2018. DNA Barcoding: A practical guide. African Centre of DNA Barcoding, University of Johannesburg.

Ptacnik, R., Solimini, A. G., Andersen, T., Tamminen, T., Brettum, P., Lepisto, L., Willen, E. and Rekolainen, S. 2008. Diversity predicts stability and resource use efficiency in natural phytoplankton communities. *Proceedings of the National Academy of Sciences*, 105, 13, 5134-5138.

Pytlik, N., Kaden, J., Finger, M., Naumann, J., Wanke, S., Machill, S. and Brunner, E. 2017. Biological synthesis of gold nanoparticles by the diatom *Stephanopyxis turris* and in vivo SERS analyses. *Algal Research*, 28, 9-15.

Pytlik, N., Klemmed, B., Machill, S., Eychmüller, A. and Brunner, E. 2019. In vivo uptake of gold nanoparticles by the diatom *Stephanopyxis turris*. *Algal Research*, 39, 101447.

Ratnasingham, S. and Hebert, P.D., 2007. BOLD: The Barcode of Life Data System (http://www. barcodinglife. org). *Molecular Ecology Notes*, *7*, 3, 355-364.

Reimann, B. E. F. and Lewin J.C. 1964. The Diatom Genus *Cylindrotheca* Rabenhorst. *Journal of the Royal Microscopical Society*, 83, 3, 283-296.

Rimet, F., Abarca, N., Bouchez, A., Kusber, W.-H., Jahn, R., Kahlert, M., Keck, F., Kelly, M. G., Mann, D. G., Piuz, A., Trobajo, R., Tapolczai, K., Vasselon, V. and Zimmermann, J. 2018. The potential of High-Throughput Sequencing (HTS) of natural samples as a source of primary taxonomic information for reference libraries of diatom barcodes. *Fottea*, 18, 1, 37-54.

Rimet, F., Gusev, E., Kahlert, M., Kelly, M. G., Kulikovskiy, M., Maltsev, Y., Mann, D. G., Pfannkuchen, M., Trobajo, R., Vasselon, V., Zimmermann, J. and Bouchez, A. 2019. Diat.barcode, an open-access curated barcode library for diatoms. *Scientific Reports*, 9, 1, 1-12.

Risjani, Y., Witkowski, A., Kryk, A., Yunianta, Górecka, E., Krzywda, M., Safitri, I., Sapar, A., Dąbek, P., Arsad, S., Gusev, E., Rudiyansyah, Peszek, Ł. and Wróbel, R.J. 2021. Indonesian coral reef habitats reveal exceptionally high species richness and biodiversity of diatom assemblages. *Estuarine, Coastal and Shelf Science*, 261, 107551.

Romanova, D.Y., 2018. Physiological and morphological features of species Psammodictyon panduriforme var. continua (Grunow) in a culture. Unpublished.

Rosy, A. 2021. The ecological function of intertidal shellfish reefs for benthic algae. PhD, University of Groningen

(RSA), R. O. S. A. 2009. Integrated Coastal Management Act (28 of 2008). In: (RSA), R. O. S. A. (ed.). South Africa: Government Gazette.

(RSA), R. O. S. A. 2019. Marine Spatial Plan Act (16 of 2018). In: (RSA), R. O. S. A. (ed.). South Africa: Government Gazette.

Ruck, E.C., Nakov, T., Alverson, A.J. and Theriot, E.C., 2016. Phylogeny, ecology, morphological evolution, and reclassification of the diatom orders Surirellales and Rhopalodiales. Molecular phylogenetics and evolution, 103, pp.155-171.

Ruppert, K. M., Kline, R. J. and Rahman, M. S. 2019. Past, present, and future perspectives of environmental DNA (eDNA) metabarcoding: A systematic review in methods, monitoring, and applications of global eDNA. *Global Ecology and Conservation*, 17, e00547.

Sabir, J.S., Theriot, E.C., Manning, S.R., Al-Malki, A.L., Khiyami, M.A., Al-Ghamdi, A.K., Sabir, M.J., Romanovicz, D.K., Hajrah, N.H., El Omri, A. and Jansen, R.K., 2018. Phylogenetic analysis and a review of the history of the accidental phytoplankter, Phaeodactylum tricornutum Bohlin (Bacillariophyta). PloS one, 13(6), p.e0196744.

Sarno, D., Kooistra, W. H. C. F., Medlin, L. K., Percopo, I. and Zingone, A. 2005. Diversity in the Genus *Skeletonema* (Bacillariophyceae) II. An Assessment of the Taxonomy of *Costatum*-Like Species with the Description of Four New Species. *Journal of Phycology*, 41, 1, 151-176.

Schaub, I., Wagner, H., Graeve, M. and Karsten, U. 2017. Effects of prolonged darkness and temperature on the lipid metabolism in the benthic diatom *Navicula perminuta* from the Arctic Adventfjorden, Svalbard. *Polar Biology*, 40, 7, 1425-1439.

Schindel, D. E. and Miller, S.E. 2005. DNA barcoding a useful tool for taxonomists. *Nature*, 435.

Selmeczy, G.B., Abonyi, A., Krienitz, L., Kasprzak, P., Casper, P., Telcs, A., Somogyvári, Z. and Padisák, J. 2018. Old sins have long shadows: climate change weakens efficiency of trophic coupling of phyto- and zooplankton in a deep oligo-mesotrophic lowland lake (Stechlin, Germany)—a causality analysis. *Hydrobiologia*, 831, 1, 101-117.

Siqueiros-Beltrones, D. A., Argumedo-Hernández, U. and López-Fuerte, F. O. 2017. Diversity of benthic diatoms in the Guerrero Negro Lagoon (El Vizcaíno Biosfere Reserve), Baja California Peninsula, Mexico. *Revista Mexicana de Biodiversidad*, 88, 1, 21-35.

Stachura-Suchoples, K., Enke, N., Schlie, C., Schaub, I., Karsten, U. and Jahn, R. 2015. Contribution towards a morphological and molecular taxonomic reference library of benthic marine diatoms from two Arctic fjords on Svalbard (Norway). *Polar Biology*, 39, 11, 1933-1956.

Stat, M., Huggett, M. J., Bernasconi, R., Dibattista, J. D., Berry, T. E., Newman, S. J., Harvey, E. S. and Bunce, M. 2017. Ecosystem biomonitoring with eDNA: metabarcoding across the tree of life in a tropical marine environment. *Scientific Reports*, 7, 1, 1-11.

Stock, W., Vanelslander, B., Rudiger, F., Sabbe, K., Vyverman, W. and Karsten, U. 2019. Thermal Niche Differentiation in the Benthic Diatom *Cylindrotheca closterium* (Bacillariophyceae) Complex. *Frontiers in Microbiology*, 10, 1395.

Swanepoel, E. and Sauka S.A. 2019. Ecosystem-based Adaptation in South African Coastal Cities. *JSTOR*.

Tapolczai, K., Keck, F., Bouchez, A., Rimet, F., Kahlert, M. and Vasselon, V. 2019. Diatom DNA Metabarcoding for Biomonitoring: Strategies to Avoid Major Taxonomical and Bioinformatical Biases Limiting Molecular Indices Capacities. *Frontiers in Ecology and Evolution*, 7, 409.

Vallaeys, T., Klink S.P., Fleouter E., Le Moing B., Lignot J.H. and Smith A.J. 2017. Bioindicators of Marine Contaminations at the Frontier of Environmental Monitoring and Environmental Genomics. *Pharmaceuticals*, *57*, 58.

Vanormelingen, P., Vanelslander, B., Sato, S., Gillard, J., Trobajo, R., Sabbe, K. and Vyverman, W. 2013. Heterothallic sexual reproduction in the model diatom *Cylindrotheca*. *European Journal of Phycology*, 48, 1, 93-105.

Vasselon, V., Rimet, F., Tapolczai, K. and Bouchez, A. 2017. Assessing ecological status with diatoms DNA metabarcoding: Scaling-up on a WFD monitoring network (Mayotte island, France). *Ecological Indicators*, 82, 1-12.

Verheye, H. M., Kirkman, S., Crawford, R.J.M. and Huggett, J.A 2017. State of the Oceans and Coasts around South Africa 2016 Report Card (No. 16). Report. South Africa: Department of Environmental Affairs (DEA), Oceans and Coasts.

Witkowski, A., Ashworth, M., Li, C., Sagna, I., Yatte, D., Gorecka, E., Franco, A. O. R., Kusber, W. H., Klein, G., Lange-Bertalot, H., Dabek, P., Theriot, E. C. and Manning, R. 2020. Exploring Diversity, Taxonomy and Phylogeny of Diatoms (Bacillariophyta) from Marine Habitats. Novel Taxa with Internal Costae. *Protist*, 171, 2, 125713.

Xu,Q. and Chen,N., 2021. Direct submission. Key Laboratory of Marine Ecology and Environmental Sciences.

Zalat, A. A., El-Sheekh, M. M., Ghandour, I. M. and Basaham, A. S. 2021. Diatom Assemblages From Surface Sediments of Two Coastal Lagoons, the Central Part of the Red Sea, Saudi Arabia and Their Associated Environmental Variables. Thalassas: *An International Journal of Marine Sciences*, 37, 1, 179-203.

Zhang, C. Z., Peng, R., Wu, Y., Wu, Z., Zhang, Z., Jiang, X. and Xu, Z. 2021. Effects of light qualities on growth, lipid contents and fatty acids profiles of *Cylindrotheca closterium.*

Zimmermann, J., Glockner, G., Jahn, R., Enke, N. and Gemeinholzer, B. 2015. Metabarcoding vs. morphological identification to assess diatom diversity in environmental studies. *Molecular Ecology Resources*, 15, 3, 526-542.

Zimmermann, J., Jahn, R. and Gemeinholzer, B. 2011. Barcoding diatoms: evaluation of the V4 subregion on the 18S rRNA gene, including new primers and protocols. *Organisms Diversity and Evolution*, 11, 3,173-192.

Supplementary material 1

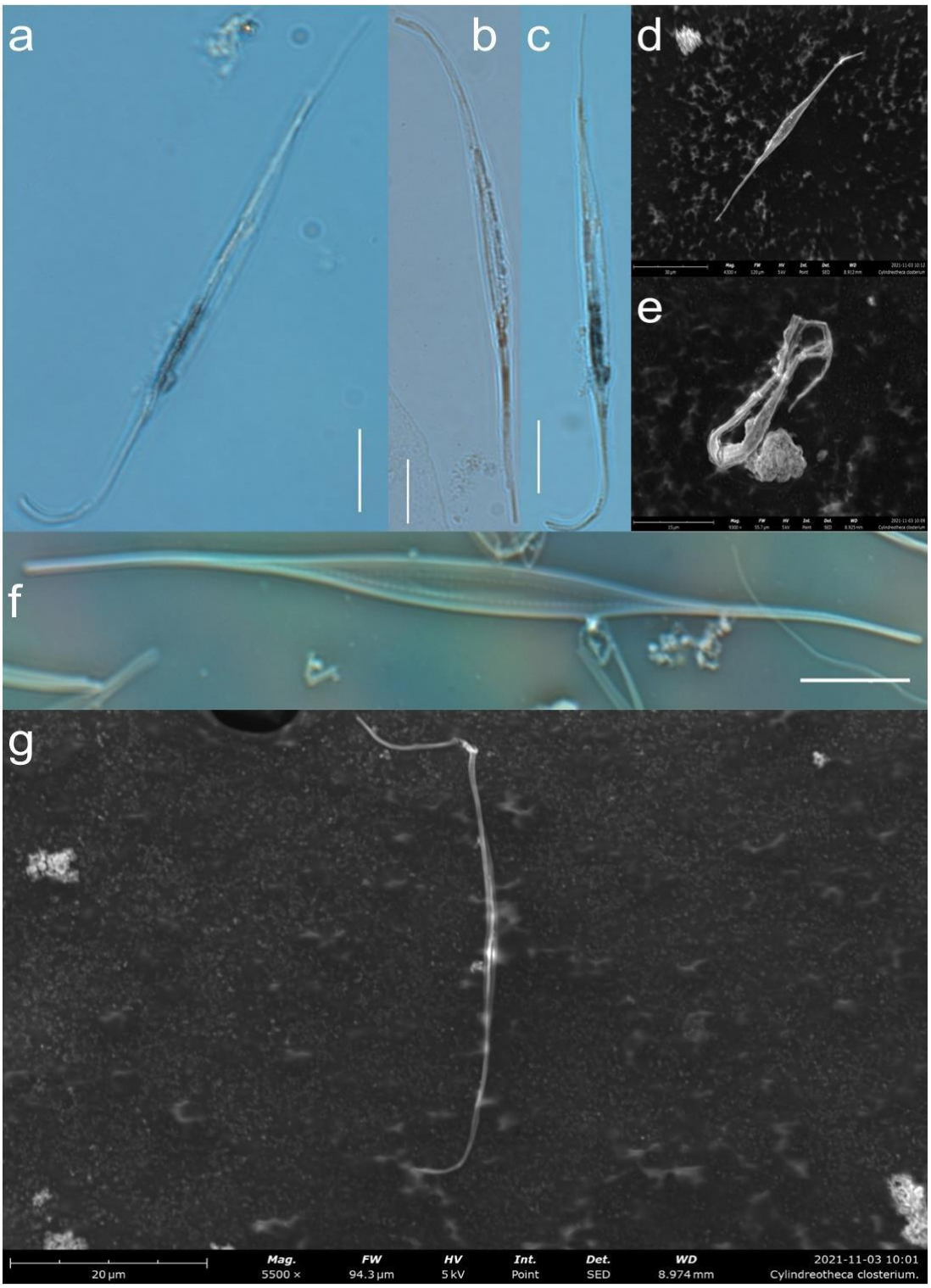


Figure 1. AB1 – *Cylindrotheca closterium* (Ehrenberg) Reimann and Lewin. a – c and f are light microscope images at 1000X magnification (scale bar = 10 µm). A Zeiss Axio Imager 2 using brightfield, differential interference and phase contrast microscopy and Zen 2012 blue edition software was used to capture and annotate images. d, e, and g are Phenom Pure G6 desktop SEM images of samples prepared on aluminium stubs with carbon tape.

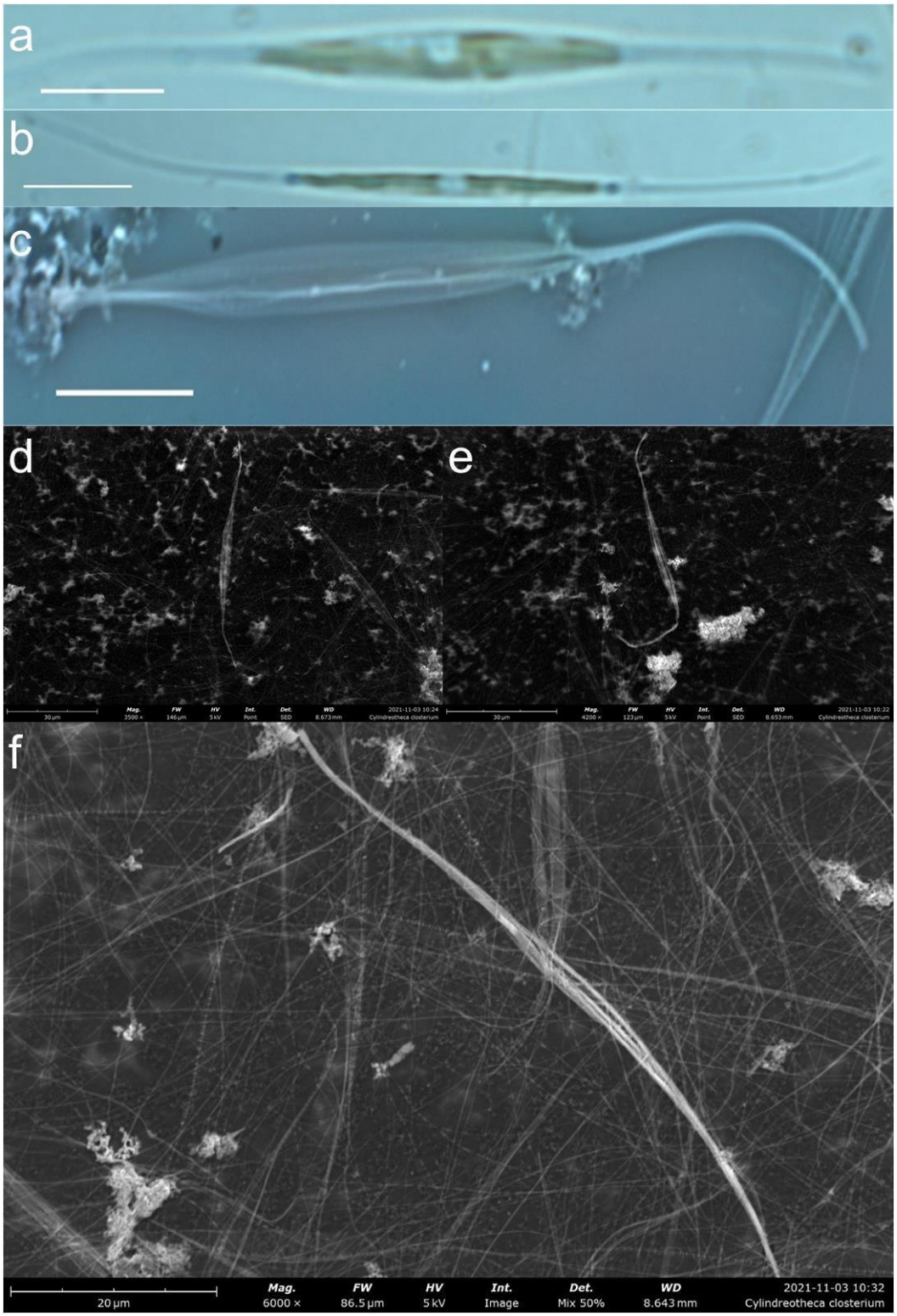


Figure 2. AB2 – *Cylindrotheca closterium* (Ehrenberg) Reimann and Lewin. a – c are light microscope images at 1000X magnification (scale bar = 10 µm), a and b are images of live samples. A Zeiss Axio Imager 2 using brightfield, differential interference and phase contrast microscopy and Zen 2012 blue edition software was used to capture and annotate images. d - f are Phenom Pure G6 desktop SEM images of samples prepared on aluminium stubs with carbon tape.

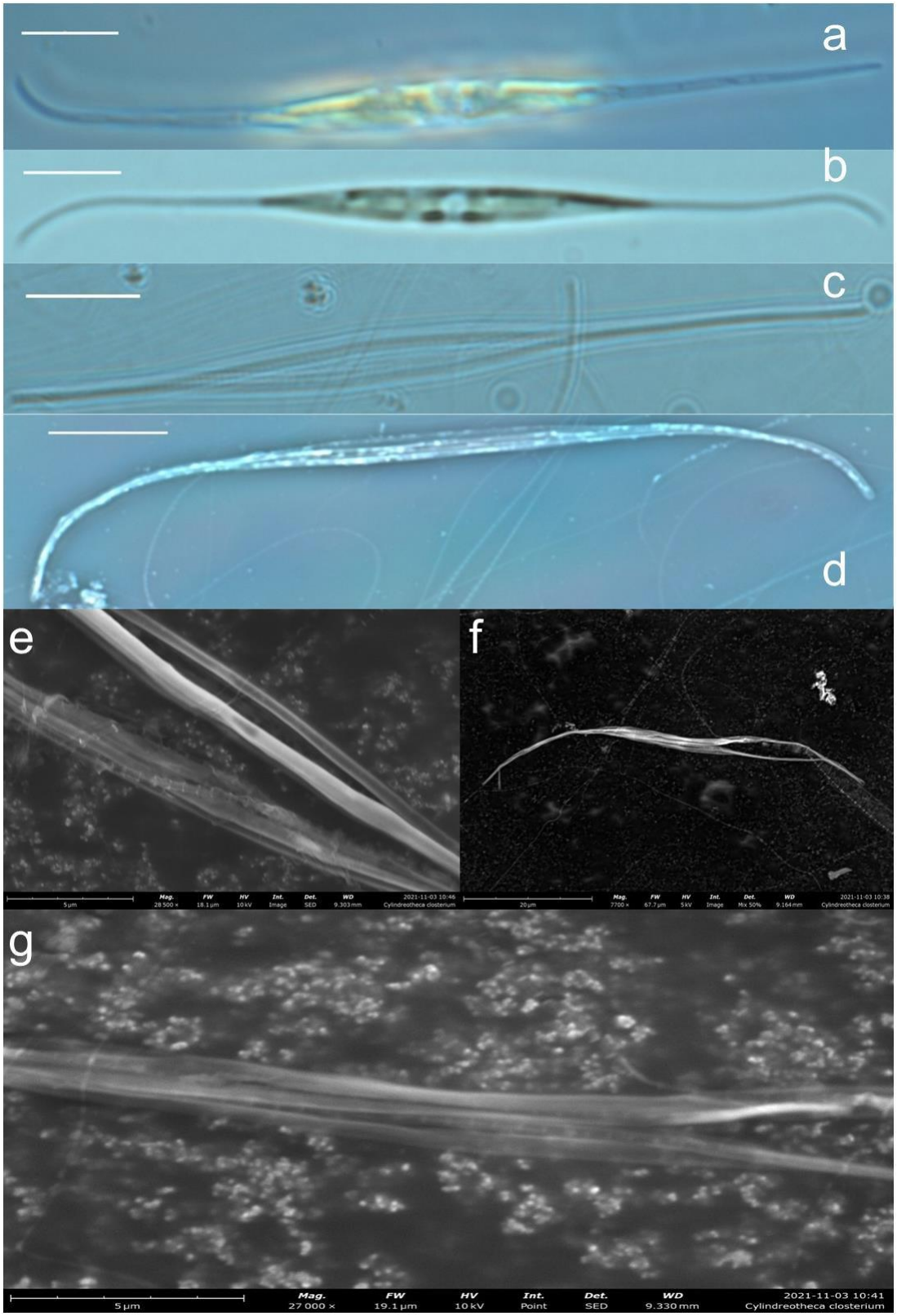


Figure 3. AB3 – *Cylindrotheca closterium* (Ehrenberg) Reimann and Lewin. a – d are light microscope images at 1000X magnification (scale bar = 10 µm), a and b are live samples, and the remaining are digested samples. A Zeiss Axio Imager 2 using brightfield, differential interference and phase contrast microscopy and Zen 2012 blue edition software was used to capture and annotate images. e - g are Phenom Pure G6 desktop SEM images of samples prepared on aluminium stubs with carbon tape.

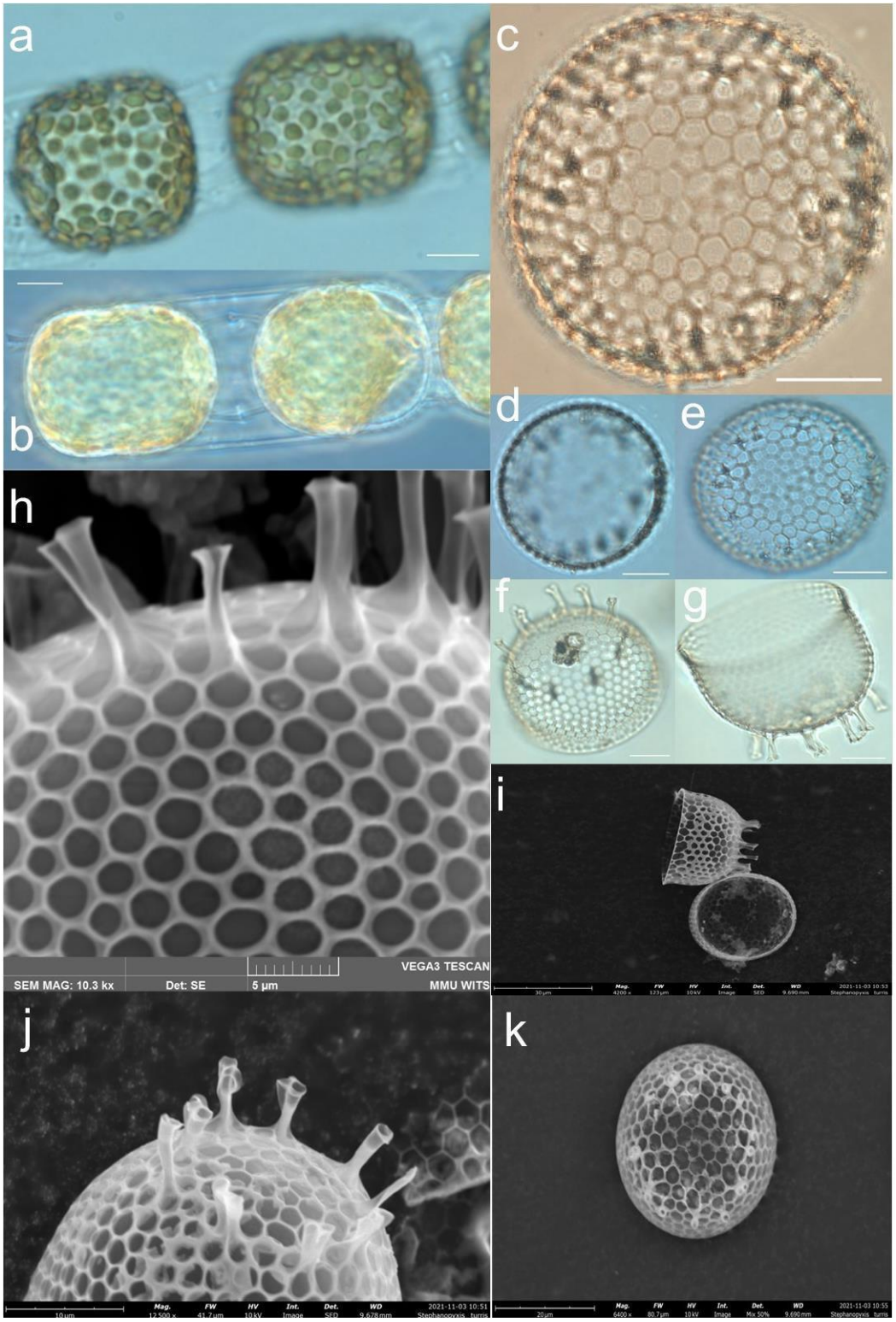


Figure 4. AB5 – ¬*Stephanopyxis turris* (Greville) Ralfs. a – g are light microscope images at 1000X magnification (scale bar = 10 µm), a and b are live samples, and the remaining are digested samples. A Zeiss Axio Imager 2 using brightfield, differential interference and phase contrast microscopy and Zen 2012 blue edition software was used to capture and annotate images. h is an image from the MMU using a Vega TESCAN SEM to image samples coated with carbon and gold palladium and i - k are Phenom Pure G6 desktop SEM images of samples prepared on aluminium stubs with carbon tape.

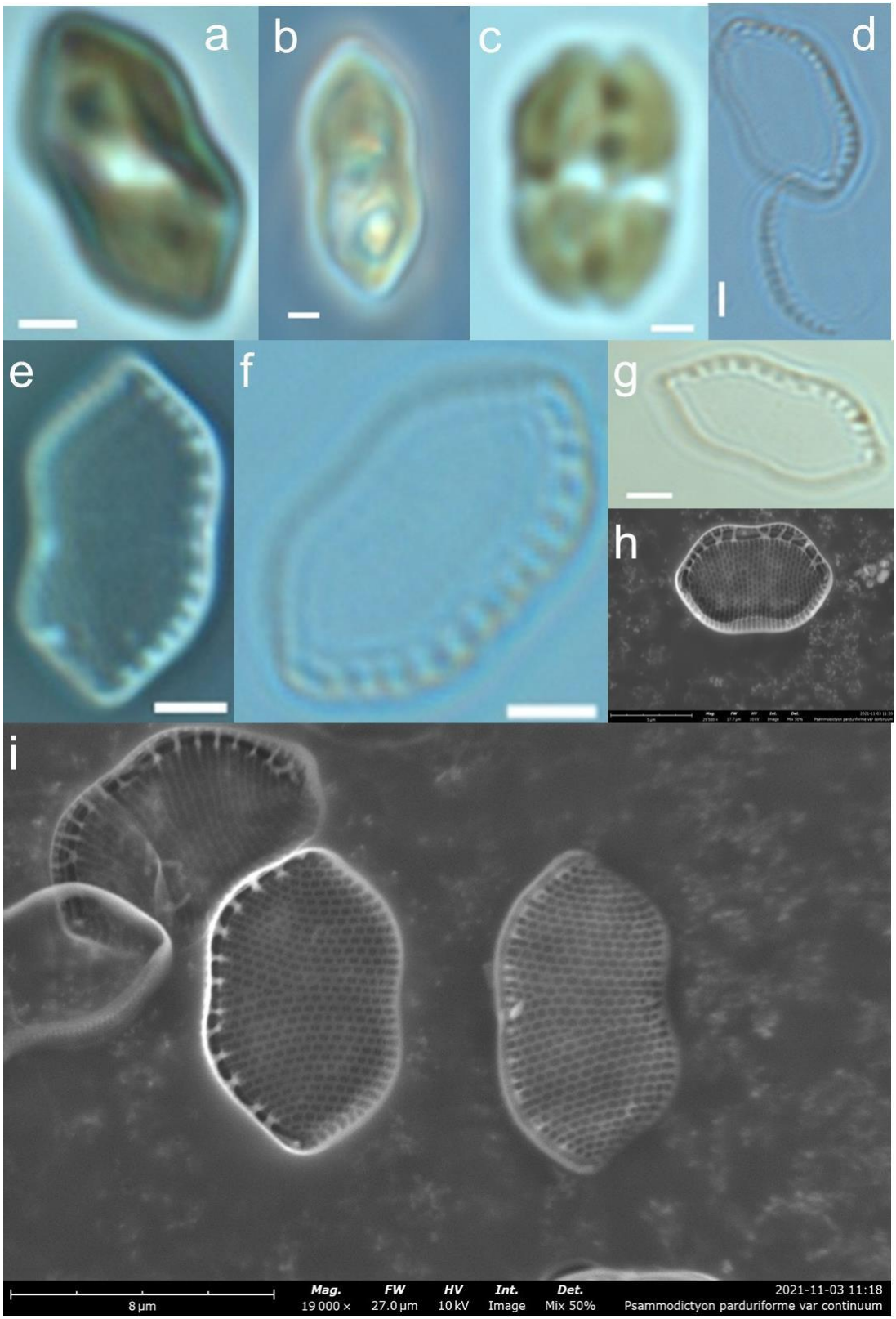


Figure 5. AB8 - *Psammodictyon panduriforme* var. *continuum* (Grunow) Snoeijis. a – g are light microscope images at 1000X magnification (scale bar = 2 µm), a - c are live samples, and the remaining are digested samples. A Zeiss Axio Imager 2 using brightfield, differential interference and phase contrast microscopy and Zen 2012 blue edition software was used to capture and annotate images. h and i are Phenom Pure G6 desktop SEM images of samples prepared on aluminium stubs with carbon tape.

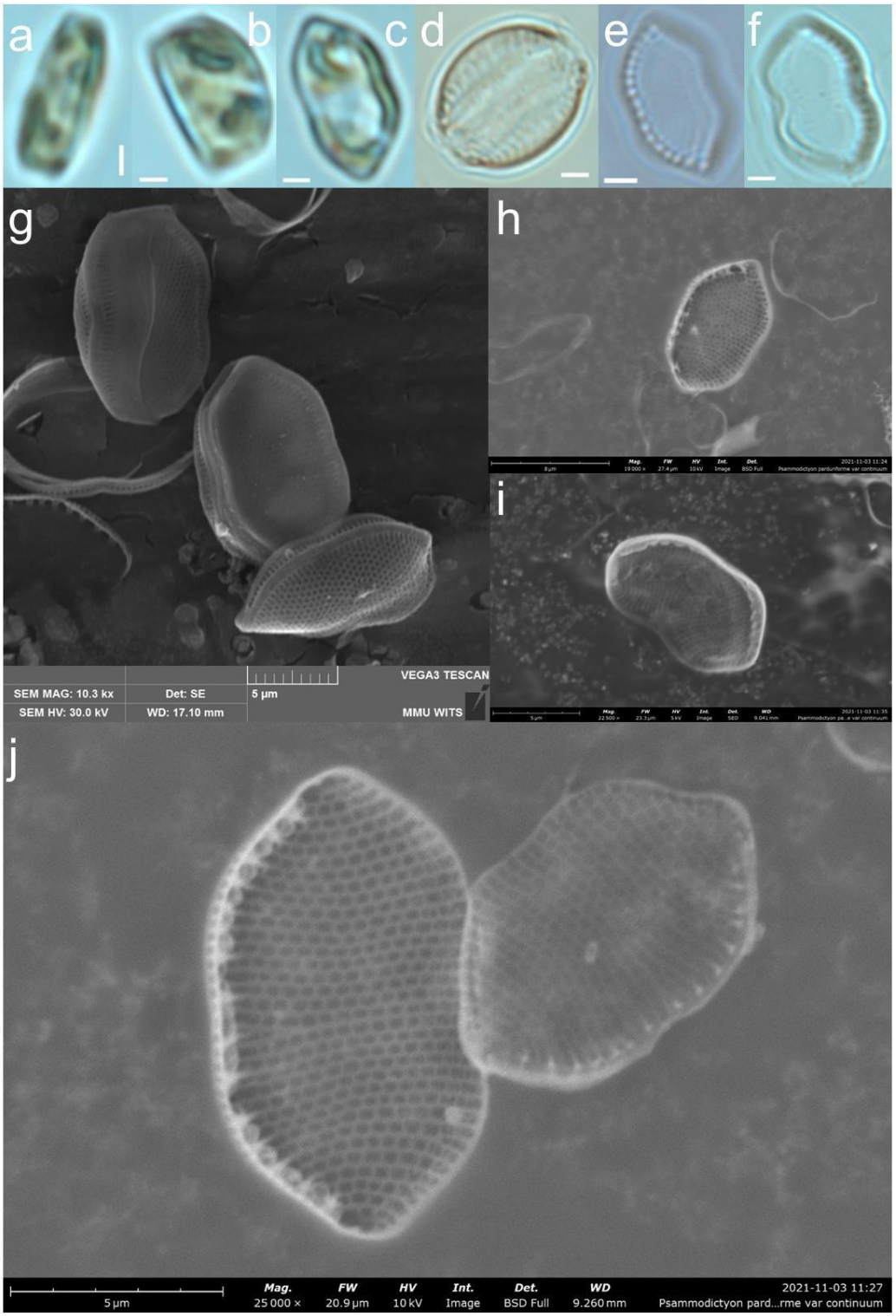


Figure 6. AB9 - *Psammodictyon panduriforme* var. *continuum* (Grunow) Snoeijis. a – f are light microscope images at 1000X magnification (scale bar = 2 µm), a - c are live samples, and the remaining are digested samples. A Zeiss Axio Imager 2 using brightfield, differential interference and phase contrast microscopy and Zen 2012 blue edition software was used to capture and annotate images. g is an image from the MMU using a Vega TESCAN SEM to image samples coated with carbon and gold palladium and h - j are Phenom Pure G6 desktop SEM images of samples prepared on aluminium stubs with carbon tape.

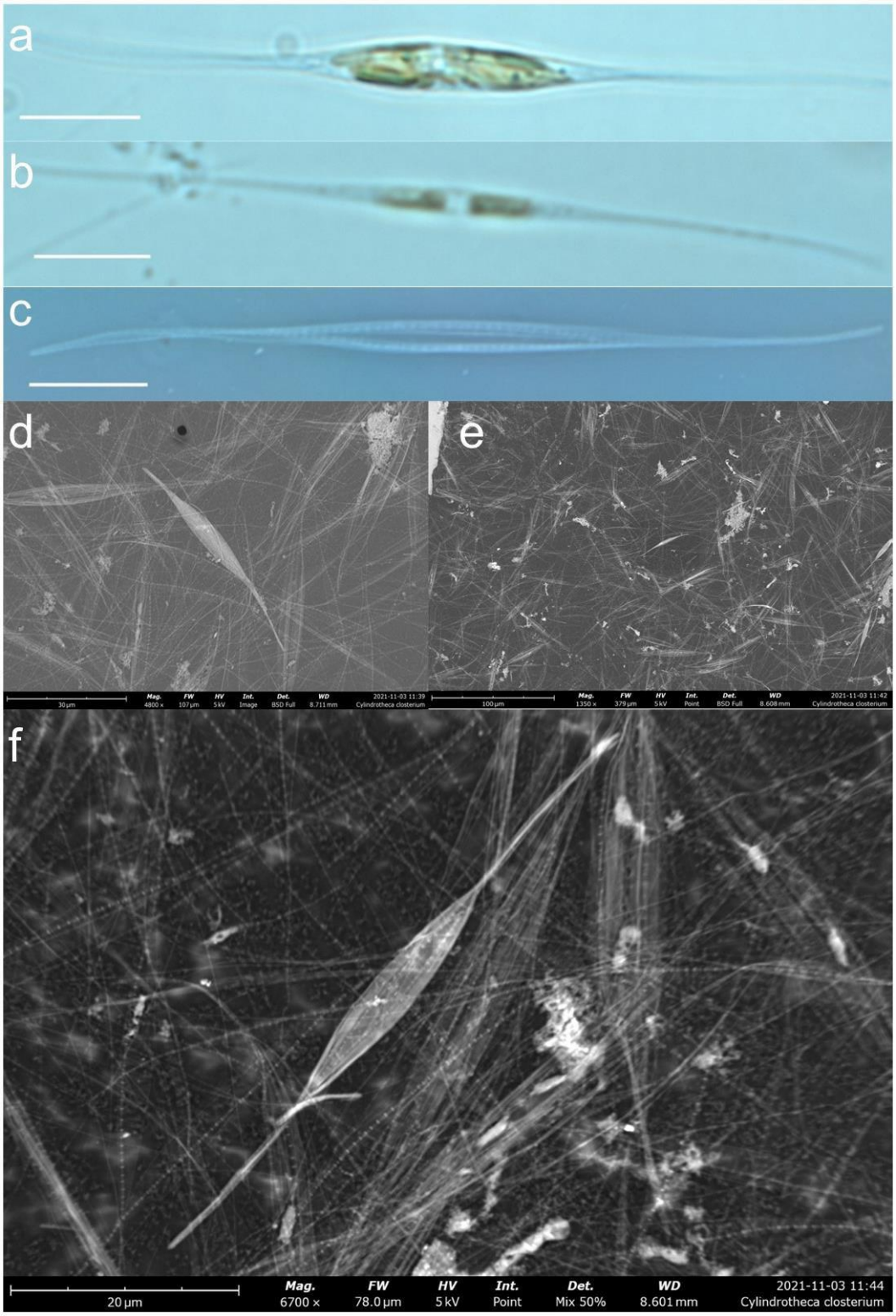


Figure 7. AB10 – *Cylindrotheca closterium* (Ehrenberg) Reimann and Lewin. a – c are light microscope images at 1000X magnification (scale bar = 10 µm), a and b are live samples, and the remaining are digested samples. A Zeiss Axio Imager 2 using brightfield, differential interference and phase contrast microscopy and Zen 2012 blue edition software was used to capture and annotate images. d – f are Phenom Pure G6 desktop SEM images of samples prepared on aluminium stubs with carbon tape.

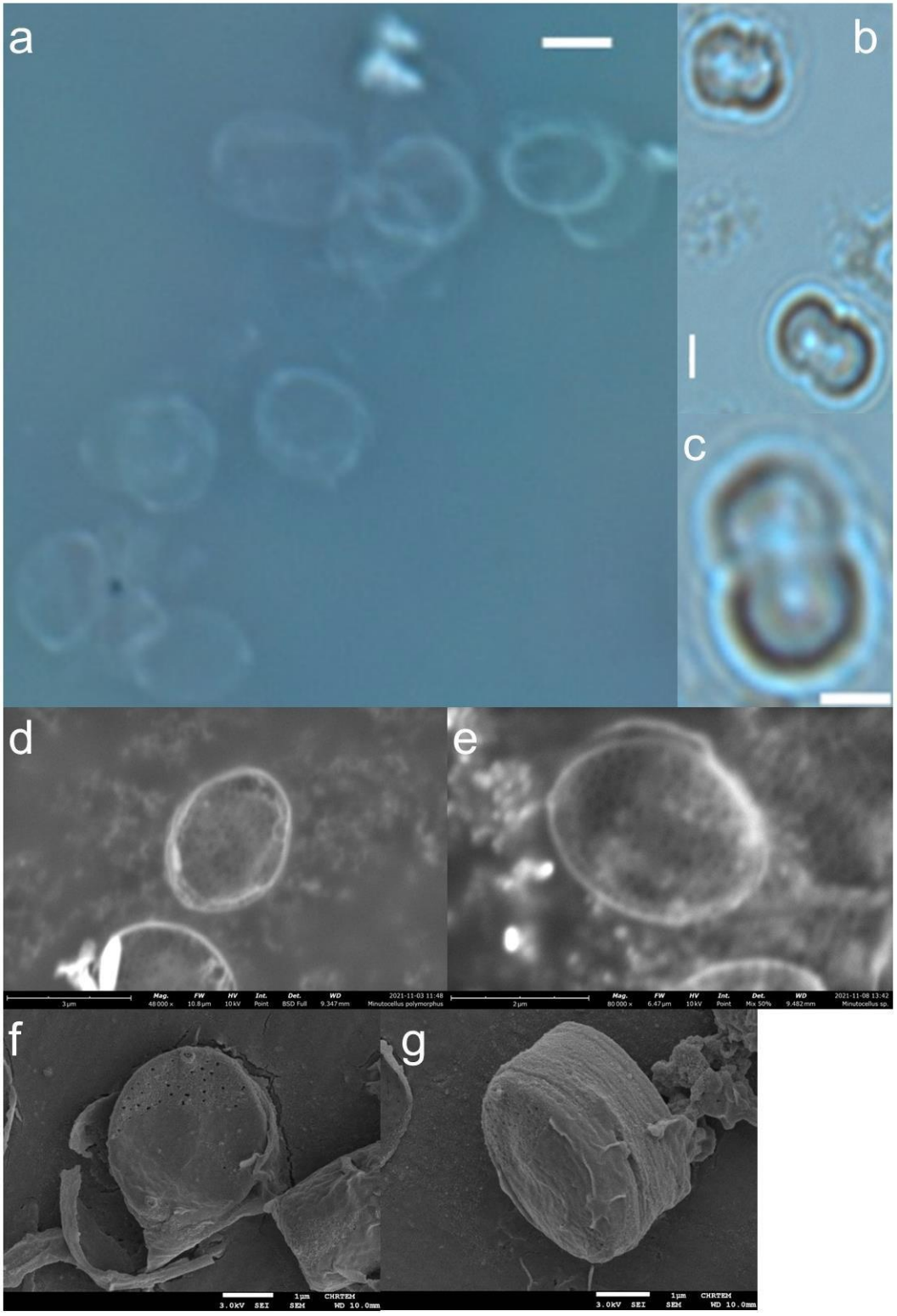


Figure 8. AB11 – *Minutocellus* sp. a – c are Light microscope images at 1000X magnification (Scale bar = 2 µm), no live samples were imaged. A Zeiss Axio Imager 2 using brightfield, differential interference and phase contrast microscopy and Zen 2012 blue edition software was used to capture and annotate images. d, e are Phenom Pure G6 desktop SEM images of samples prepared on aluminium stubs with carbon tape. f, g are images from the HRTEM using a JOEL JSM–7001F SEM, samples were coated with carbon and gold palladium

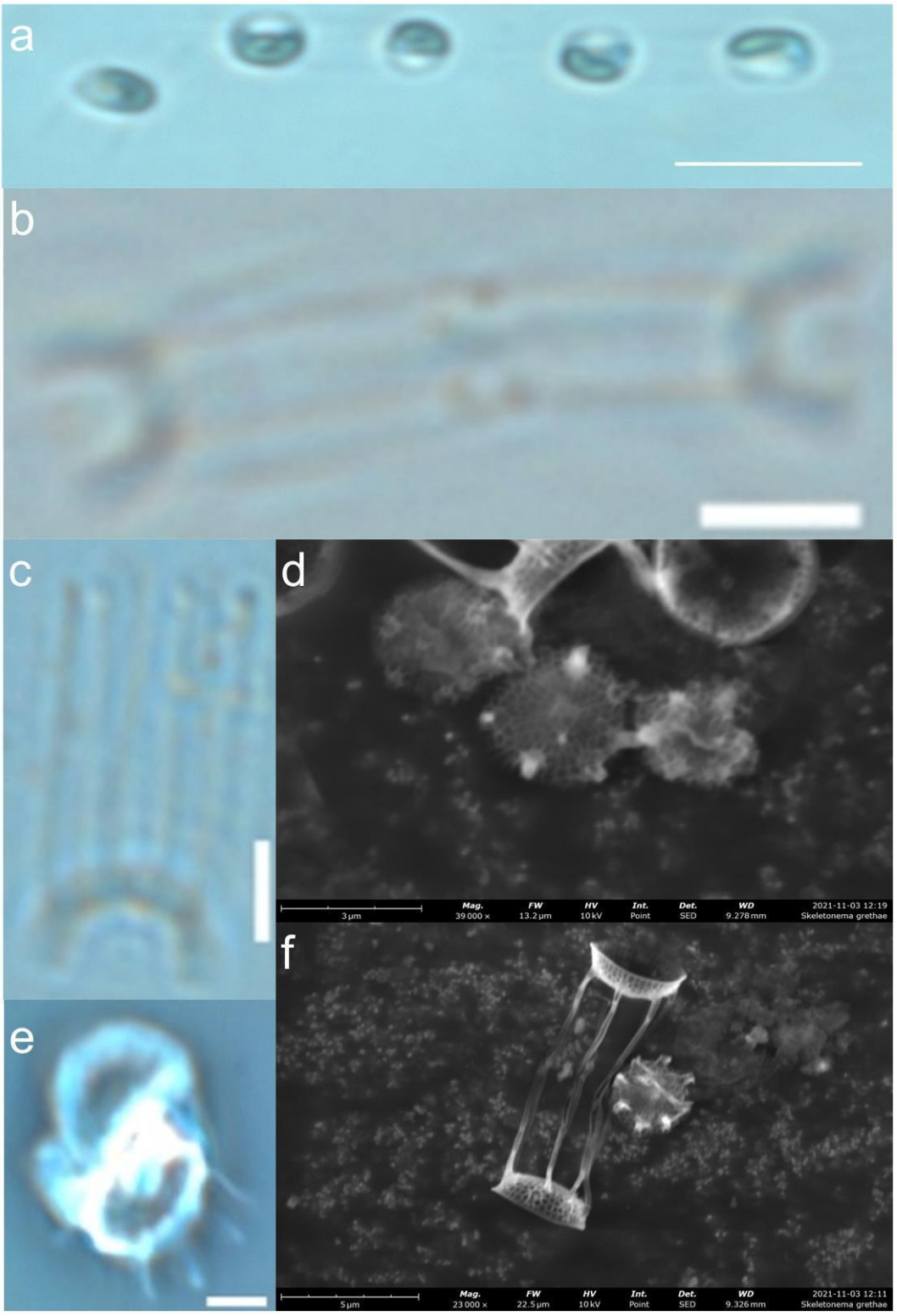


Figure 9. AB12 – *Skeletonema grethae* (Zingone and Sarno). a – c and e are light microscope images at 1000X magnification (a: scale bar = 10 µm b, c, e: scale bar = 2 µm), a is an image of the live sample, and the remaining are digested samples. A Zeiss Axio Imager 2 using brightfield, differential interference and phase contrast microscopy and Zen 2012 blue edition software was used to capture and annotate images. d and f are Phenom Pure G6 desktop SEM images of samples prepared on aluminium stubs with carbon tape.

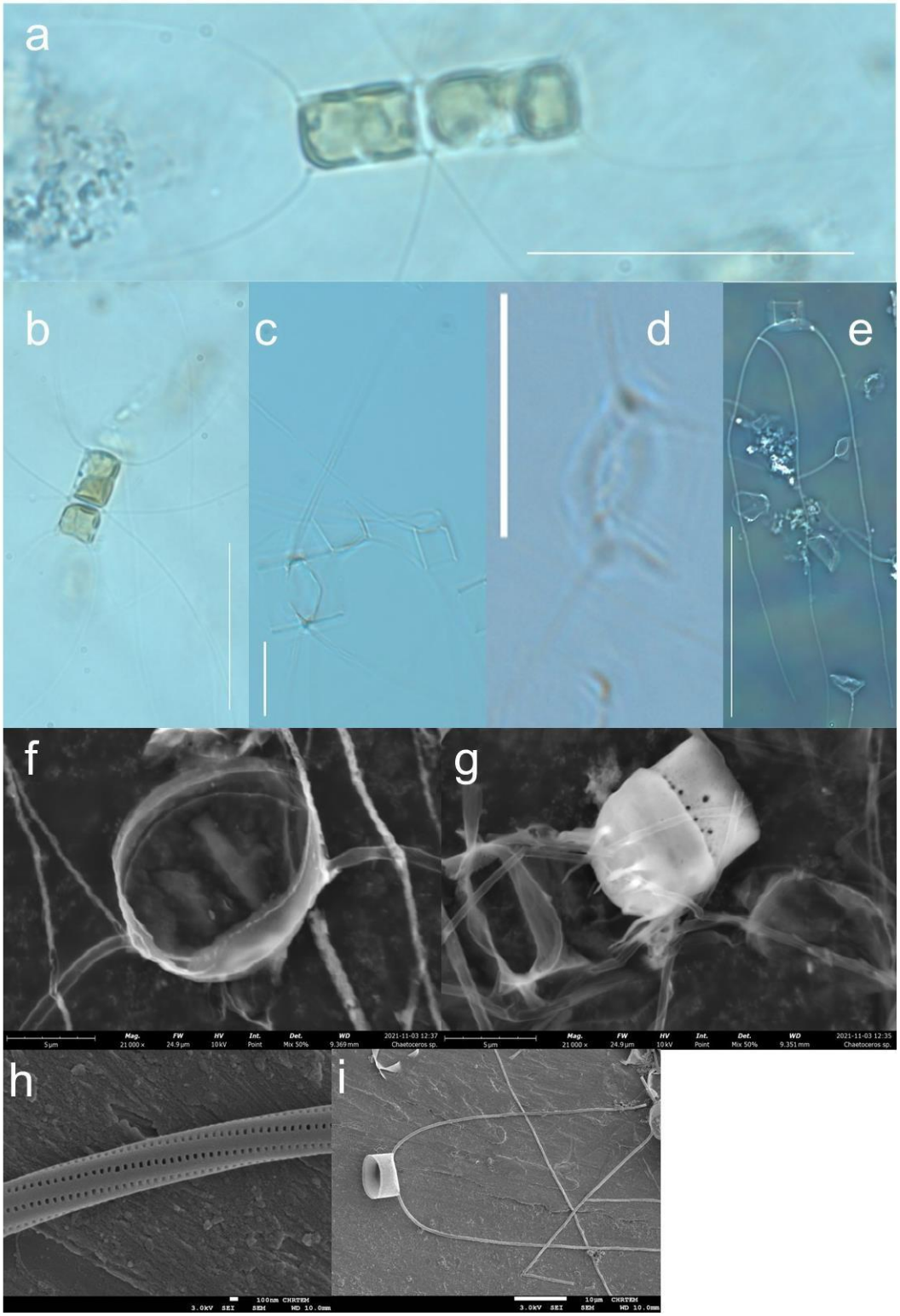


Figure 10. AB13 – *Chaetoceros* sp. a – e are light microscope images at 1000X magnification (a, b, d, e: scale bar = 50 µm c: scale bar = 10 µm), a and b are live samples, and the remaining are digested samples. A Zeiss Axio Imager 2 using brightfield, differential interference and phase contrast microscopy and Zen 2012 blue edition software was used to capture and annotate images. f, g are Phenom Pure G6 desktop SEM images of samples prepared on aluminium stubs with carbon tape. h, i are images from the HRTEM using a JOEL JSM–7001F SEM, samples were coated with carbon and gold palladium

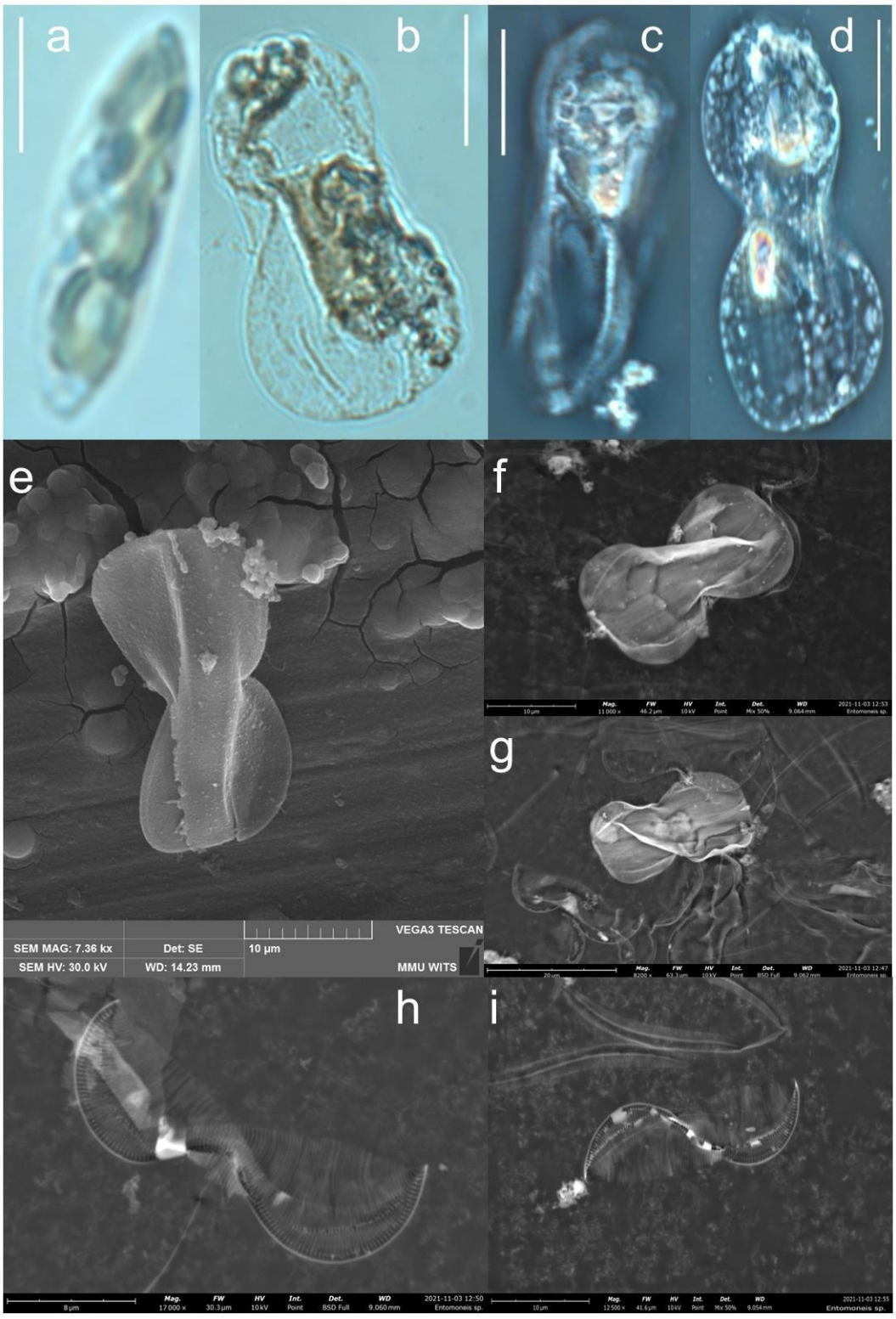


Figure 11. NWU1 – *Entomoneis* sp. a – d are light microscope images at 1000X magnification (scale bar = 10 µm), a and b are live samples, and the remaining are digested samples. A Zeiss Axio Imager 2 using brightfield, differential interference and phase contrast microscopy and Zen 2012 blue edition software was used to capture and annotate images. e is an image from the MMU using a Vega TESCAN SEM, samples were coated with carbon and gold palladium and f - i are Phenom Pure G6 desktop SEM images of samples prepared on aluminium stubs with carbon tape.

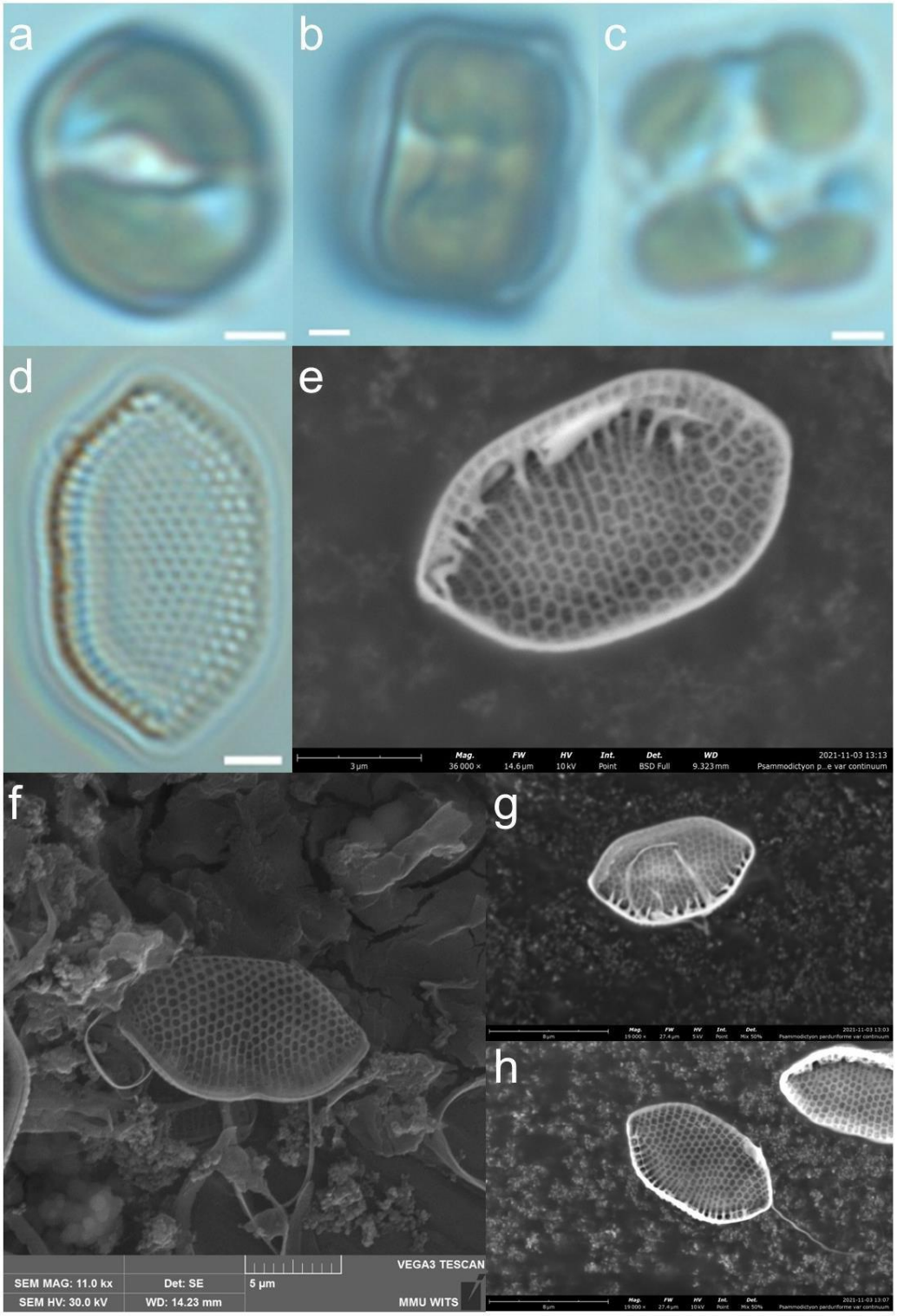


Figure 12. NWU2 - *Psammodictyon panduriforme* var. *continuum* (Grunow) Snoeijis. a – d are light microscope images at 1000X magnification (scale bar = 2 µm), a - c are live samples, and the remaining are digested samples. A Zeiss Axio Imager 2 using brightfield, differential interference and phase contrast microscopy and Zen 2012 blue edition software was used to capture and annotate images. f is an image from the MMU using a Vega TESCAN SEM, samples coated with carbon and gold palladium and e, g and h are Phenom Pure G6 desktop SEM images of samples prepared on aluminium stubs with carbon tape.

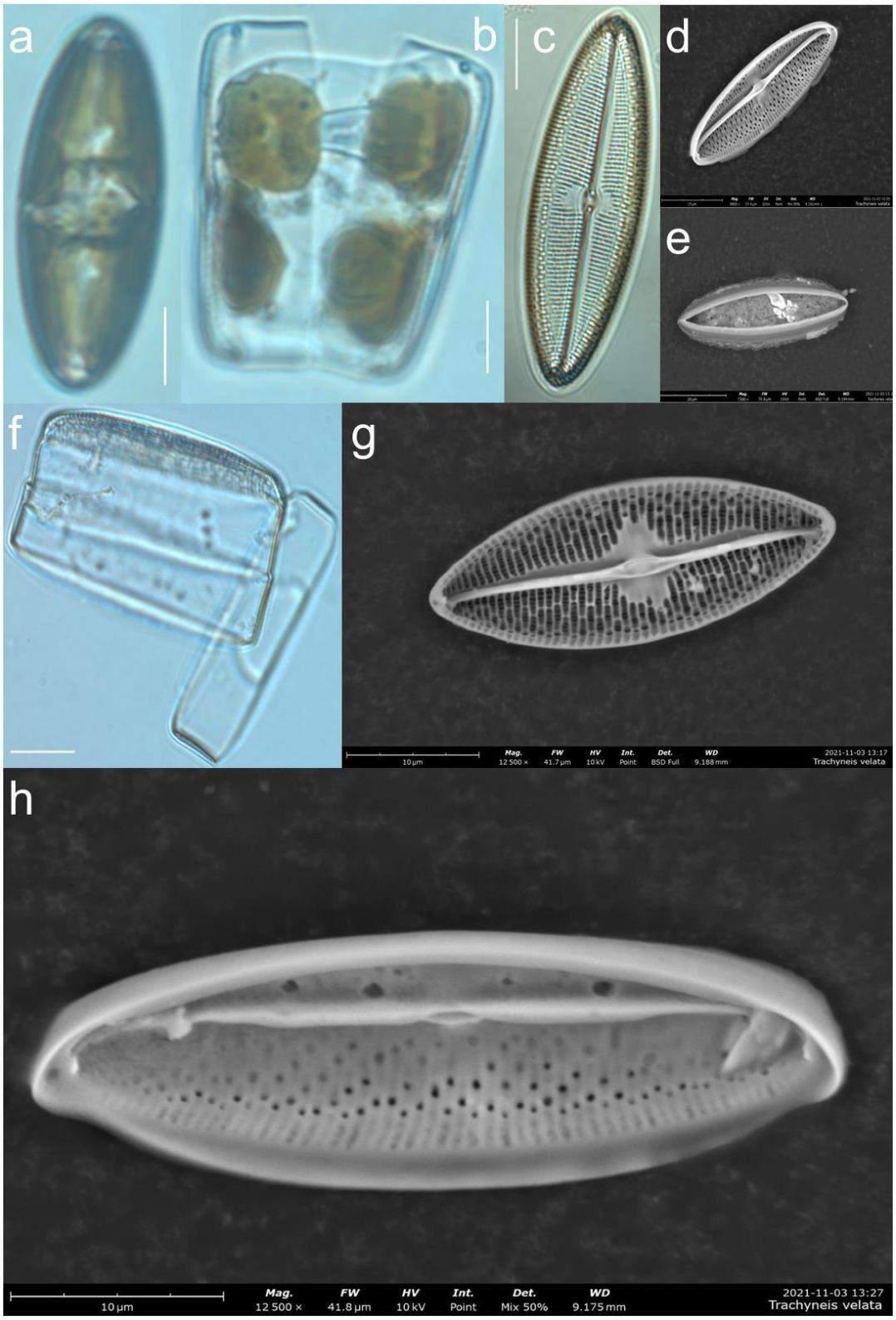


Figure 13. NWU3 - *Trachyneis velata* (Schmidt). a – c and f are light microscope images at 1000X magnification (scale bar = 10 µm), a and b are live samples, and the remaining are digested samples. A Zeiss Axio Imager 2 using brightfield, differential interference and phase contrast microscopy and Zen 2012 blue edition software was used to capture and annotate images. d, e, g, and h are Phenom Pure G6 desktop SEM images of samples prepared on aluminium stubs with carbon tape.

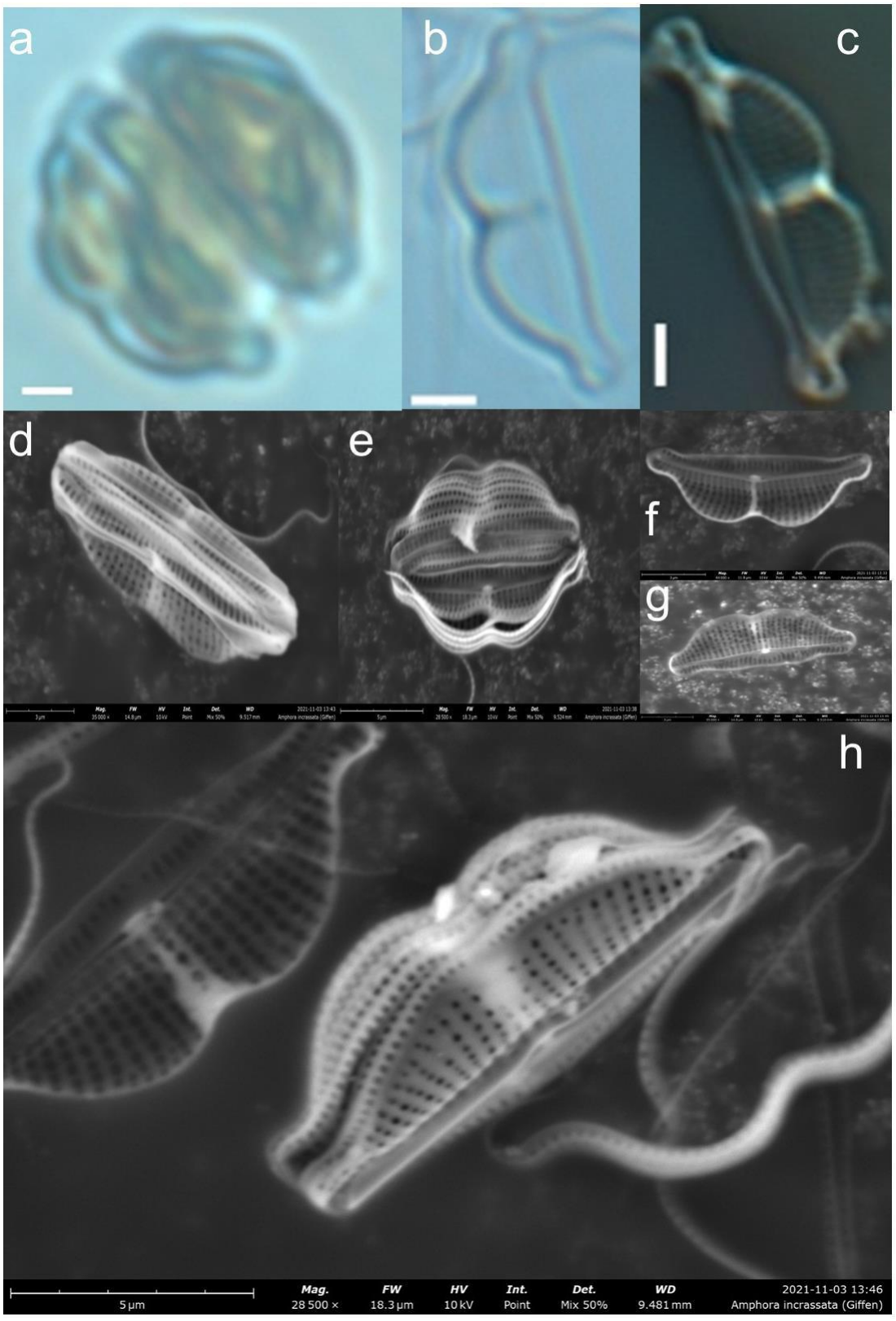


Figure 14. NWU5 – *Amphora incrassata* (Giffen). a - c are light microscope images at 1000X magnification (scale bar = 2 µm), a is of live samples, and the remaining are digested samples. A Zeiss Axio Imager 2 using brightfield, differential interference and phase contrast microscopy and Zen 2012 blue edition software was used to capture and annotate images. d - h are Phenom Pure G6 desktop SEM images of samples prepared on aluminium stubs with carbon tape.

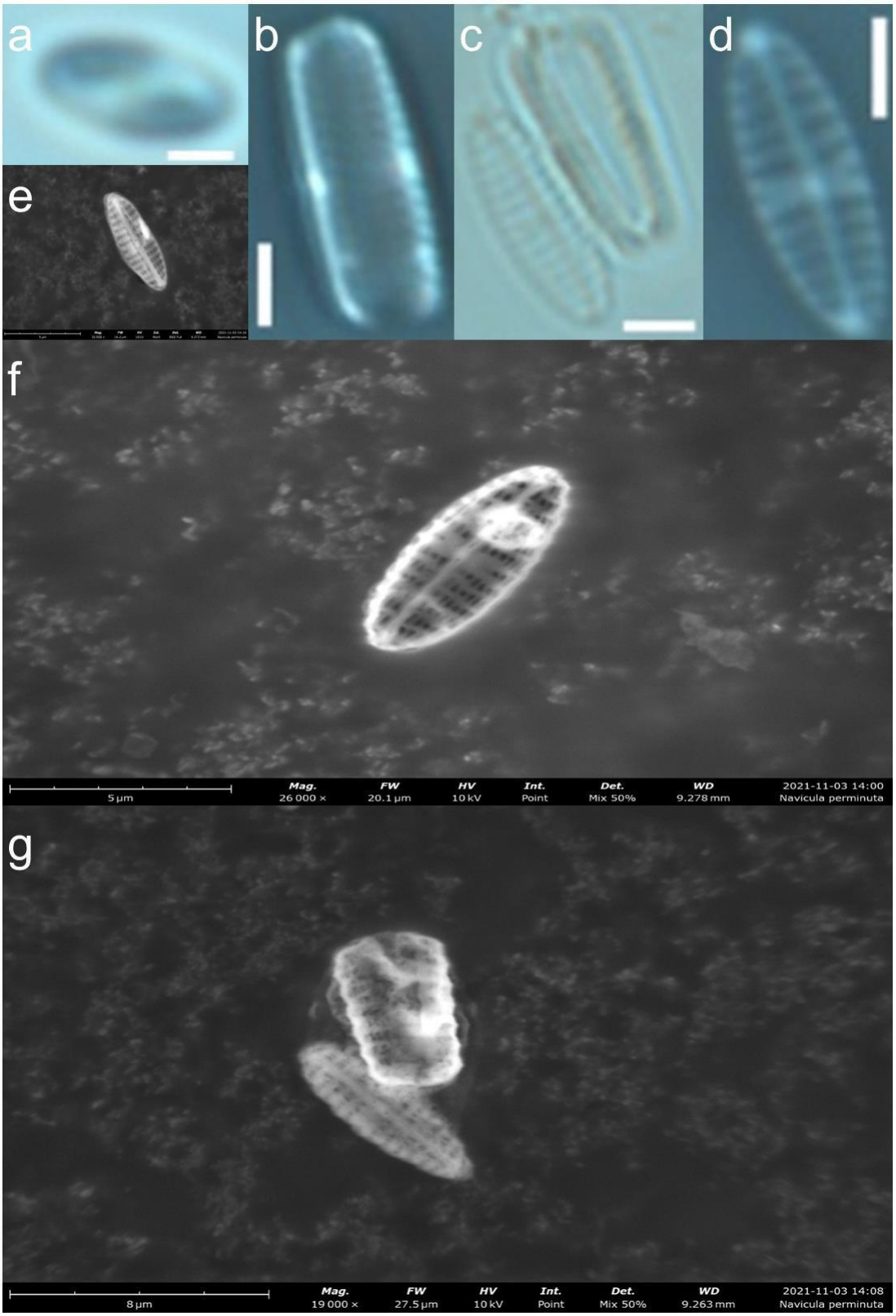


Figure 15. NWU6 – *Navicula perminuta* (Grunow). a – d are light microscope images at 1000X magnification (scale bar = 2 µm), a is of live sample, and the remaining are digested samples. A Zeiss Axio Imager 2 using brightfield, differential interference and phase contrast microscopy and Zen 2012 blue edition software was used to capture and annotate images. e - g are Phenom Pure G6 desktop SEM images of samples prepared on aluminium stubs with carbon tape.

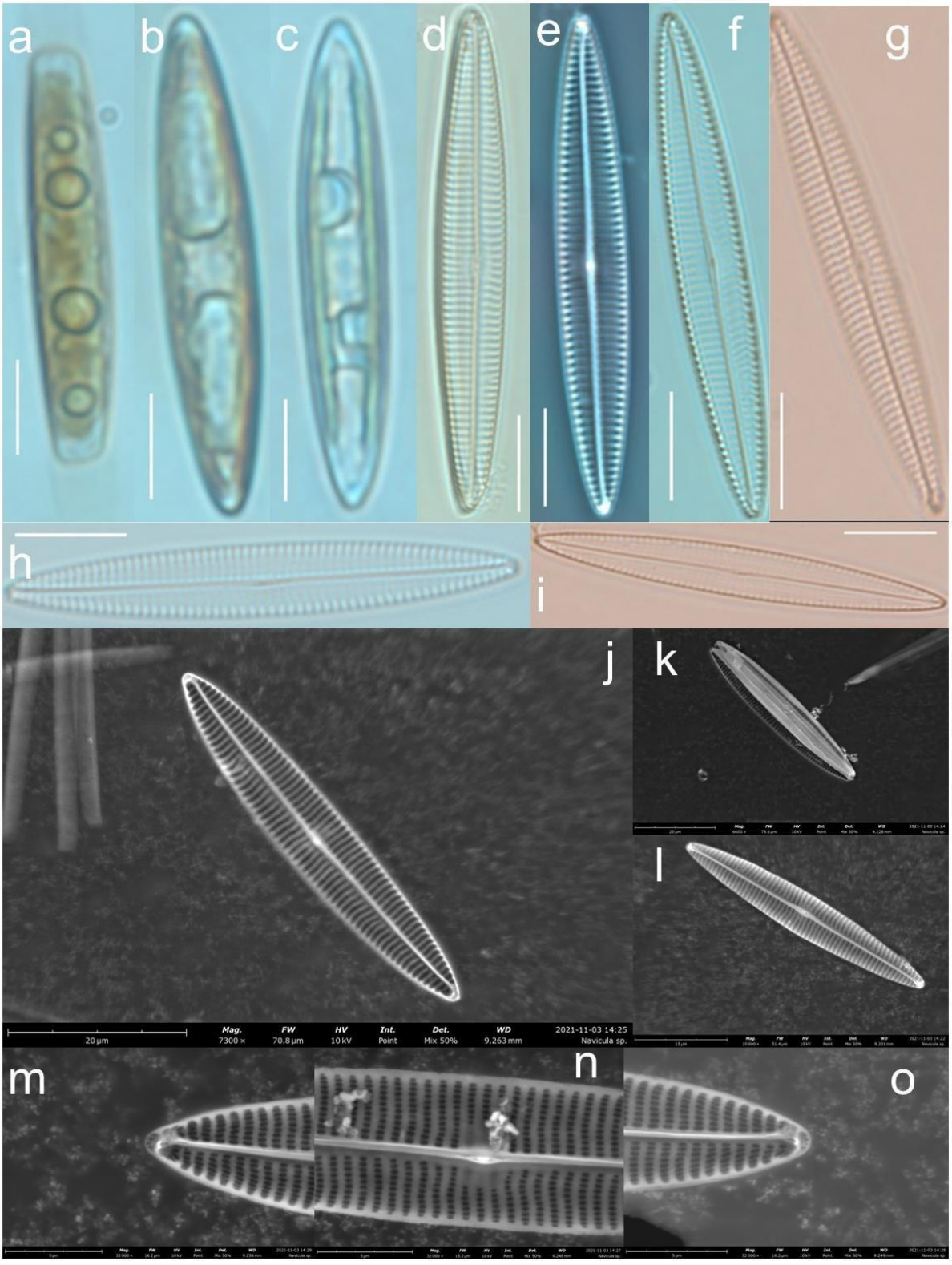


Figure 16. NWU7 – *Navicula* sp. a – i are light microscope images at 1000X magnification (scale bar = 10 µm), a - c are live samples, and the remaining are digested samples. A Zeiss Axio Imager 2 using brightfield, differential interference and phase contrast microscopy and Zen 2012 blue edition software was used to capture and annotate images. j - o are Phenom Pure G6 desktop SEM images of samples prepared on aluminium stubs with carbon tape.

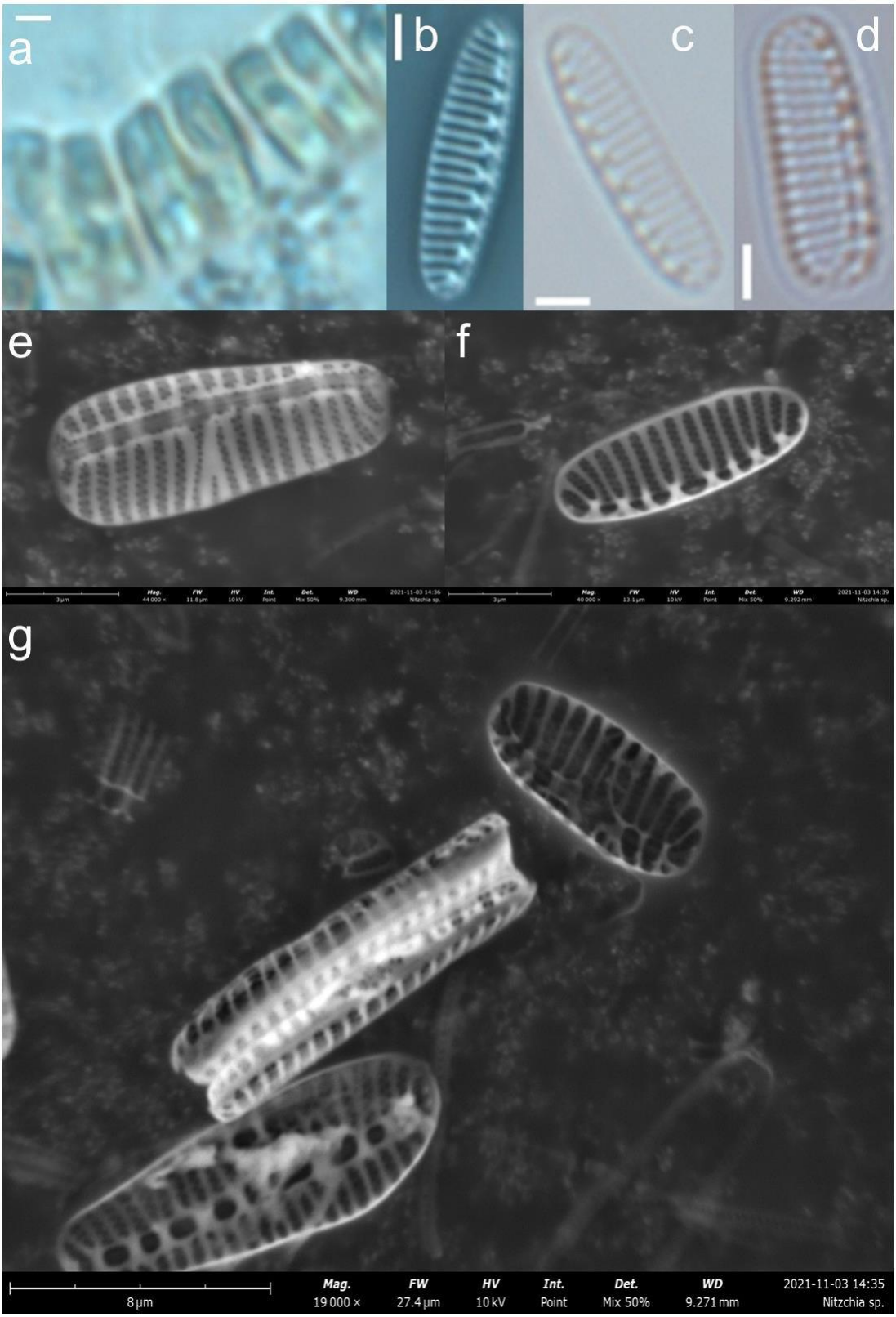


Figure 17. NWU8 – *Nitzschia* sp. a – d are light microscope images at 1000X magnification (scale bar = 2 µm), a is of live samples, and the remaining are digested samples. A Zeiss Axio Imager 2 using brightfield, differential interference and phase contrast microscopy and Zen 2012 blue edition software was used to capture and annotate images. e - f are Phenom Pure G6 desktop SEM images of samples prepared on aluminium stubs with carbon tape.

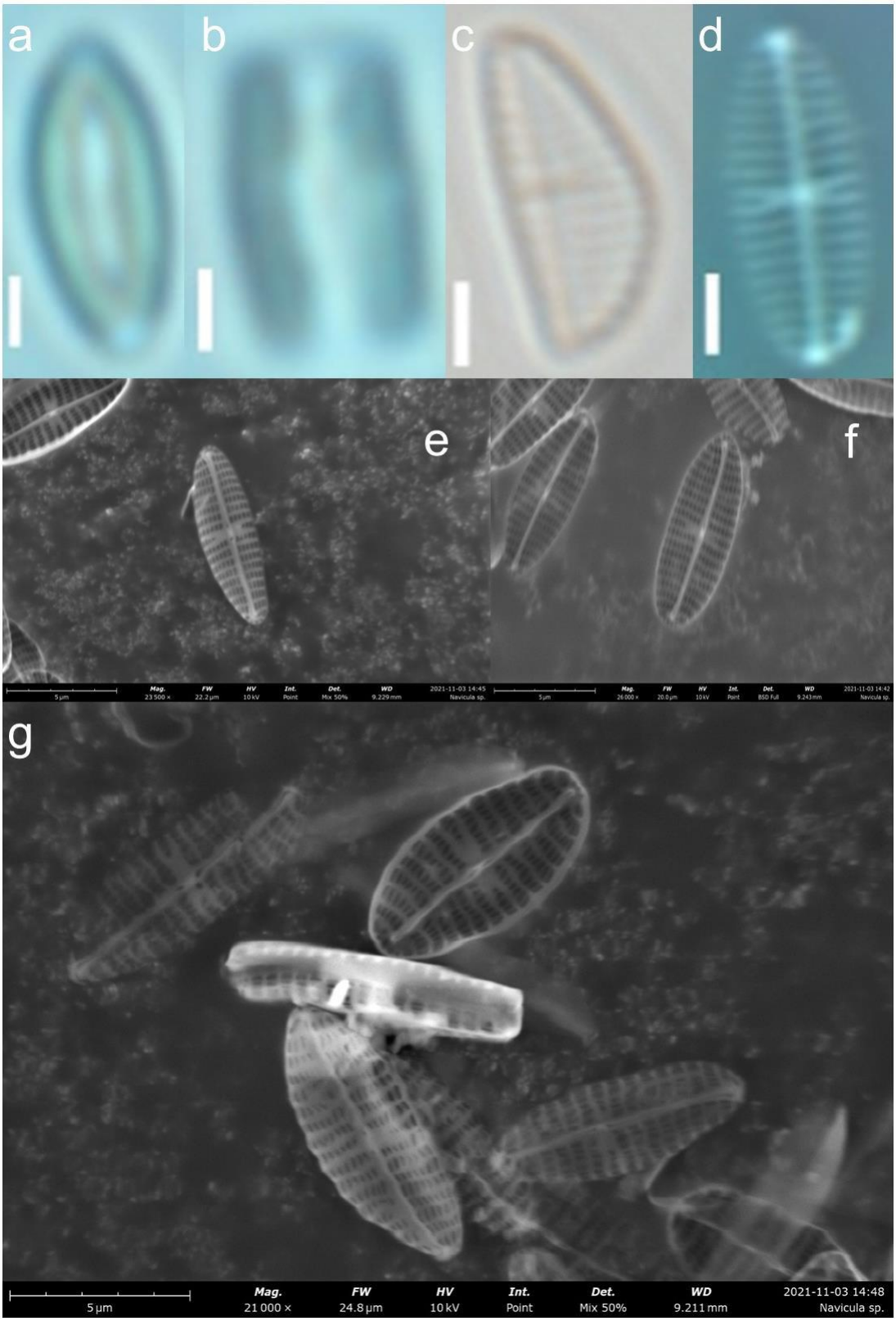


Figure 18. NWU9 – *Navicula* sp. a – d are light microscope images at 1000X magnification (scale bar = 2 µm), a is of a live sample, and the remaining are digested samples. A Zeiss Axio Imager 2 using brightfield, differential interference and phase contrast microscopy and Zen 2012 blue edition software was used to capture and annotate images. e, f and g are Phenom Pure G6 desktop SEM images of samples prepared on aluminium stubs with carbon tape.

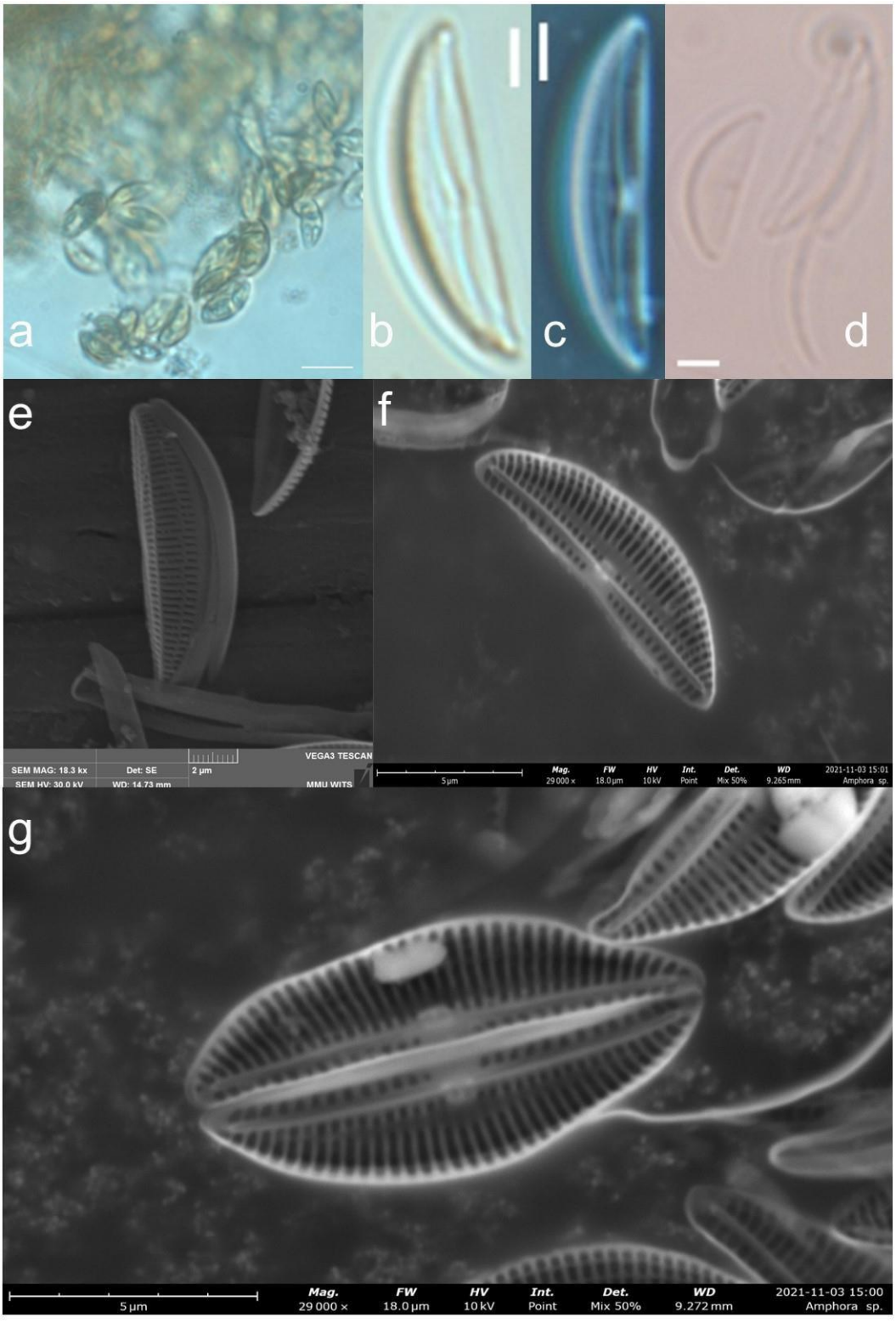
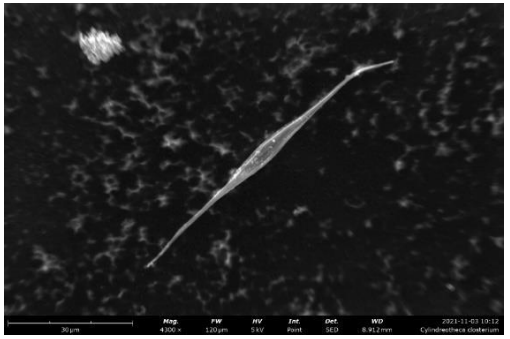
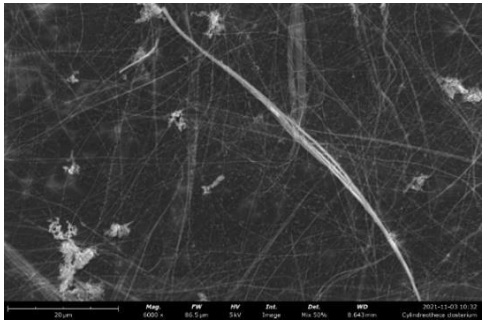
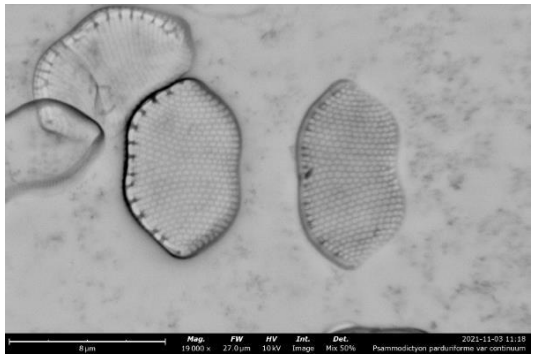
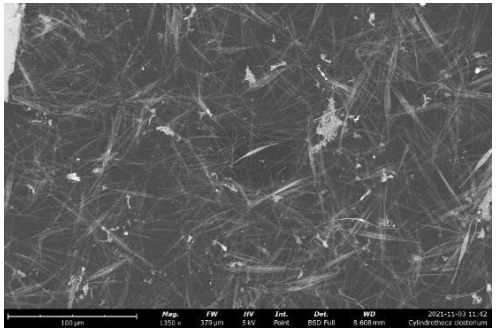


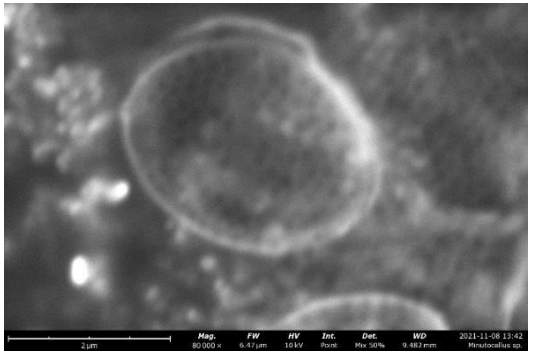
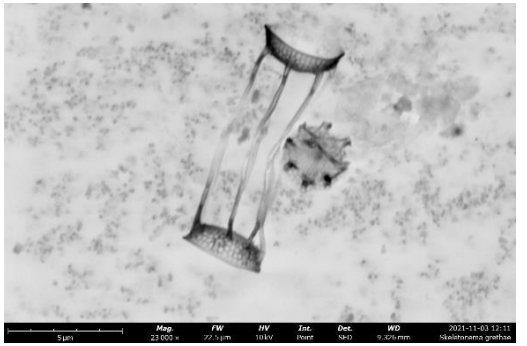
Figure 19. NWU10 – *Amphora* sp. a – d are light microscope images at 1000X magnification (scale bar = 2 µm), a is of a live sample, and the remaining are digested samples. A Zeiss Axio Imager 2 using brightfield, differential interference and phase contrast microscopy and Zen 2012 blue edition software was used to capture and annotate images. e is an image from the MMU using a Vega TESCAN SEM, samples coated with carbon and gold palladium and f - g are Phenom Pure G6 desktop SEM images of samples prepared on aluminium stubs with carbon tape.

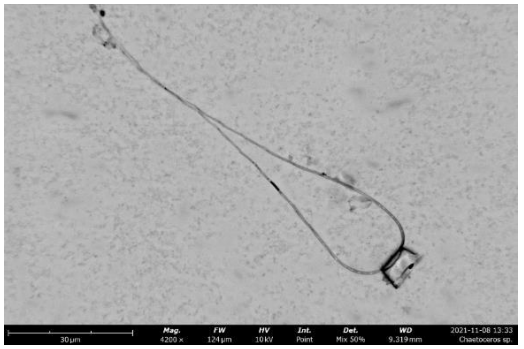
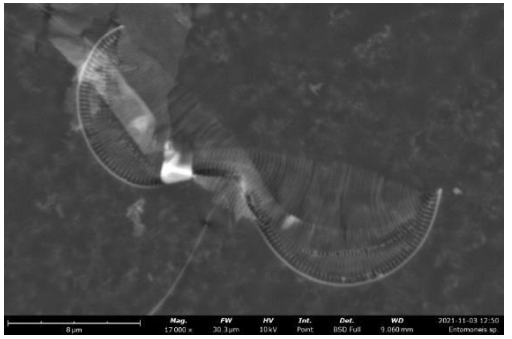
Supplementary material 2

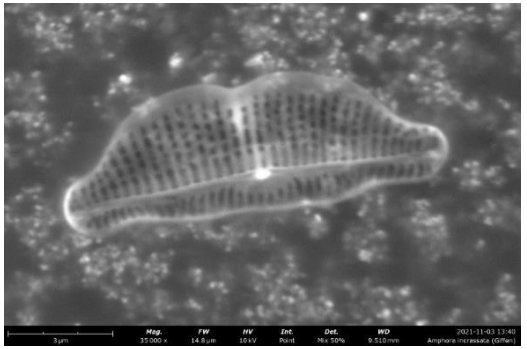
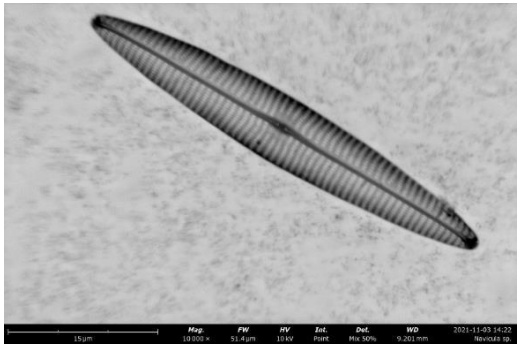
Table 1. Samples with barcode and associated morphological data that has been uploaded on the Barcode of Life Data Systems (BOLD) and relevant accession number for each barcode sequence. The barcode sequences can be used to identify diatom species when doing molecular analysis and associated images used to further confirm identification.

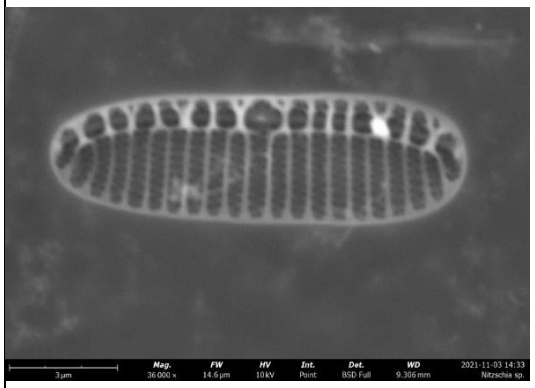
Sample ID	Species/Genus	Accession number	Barcode sequence 5'-3'	Image
AB1	*Cylindrotheca closterium* (Ehrenberg) Reimann and Lewin	OL794286	ASTTAAACGTTACAGGCGCAACAAT GGAAGAAGTATACAAACGTTCAGAG TATGCTAAAGAAGTAGGTTCTATCAT TATCATGATCGATTTAGTTATGGGAT ATACAGCAATCCAAAGTATTGCTTTA TGGGCTCGTGAAAATGATATGCTTTT ACATTTACACCGTGCTGGTAACTCA ACATATGCACGTCAAAAAAATCATG GTATCAATTTCCGTGTAATTTGTAAA TGGATGCGTATGTCTGGTGTAGATC ATATTCATGCTGGTACA	
AB2	*Cylindrotheca closterium* (Ehrenberg) Reimann and Lewin	OL794283	GGCGCWACAATGGAAGAAGTATACAAA CGTTCAGAGTATGCTAAAGAAGTAGGTT STATCATTATCATGATMGATTTAGTTATG GGATATACMGCAATCCAAAGTATTGCTTT ATGGGCTCGTGAAAATGATATGCTTTTACAT TTACACCGTGMTGGTAACTCAWCATATGCA CGTCAAAAAAATCATGGTATCAATTTCCGT GTAATTTGTAAATGGATGCGTATGTCTGG TGTAAATCAYATTCATGCTGGTAC	

AB8	*Psammodictyon panduriforme* var. *continuum* (Grunow) Snoeijis	OL794285	GAAGAAATCTATSAACGTGCTGAGTA CGCTAAAGAAGTAGGTTCTGTTATCA TTATRATCGATTTAGTTATGGGTTATA CTGCTATCCAATCTATCGCTCTTTGGG CTCGTAAMAACGATATGCTTTTACAC TTACACCGTGCTGGTAACTCTACTTAC GCTCGTCAAAAAAACCACGGTATTAAC TTCCGTGTAATCTGTAAATGGATGCGT ATGTCTGGTGTAGATCAYATCCACGCT GGTACA	
AB10	*Cylindrotheca closterium* (Ehrenberg) Reimann and Lewin	OL794289	CTTAAACATCACAGCCGCTACGATG GAAGAAGTTTACAAACGTGCTGACT ACGCTAAGAAAATCGGTTCTATCAT TGTAATGATCGATTTAGTAATGGGTT ACACAGCAATTCAAAGTATTGCTTAC TGGGCTCGTGAAAACGATATGTTATT ACACTTACACCGTGCTGGTAACTCTA CTTACGCTCGTCAAAAGAACCACGGT ATTAACTTCCGTGTTATCTGTAAATG GATGCGTATGTCTGGTGTGGATCAC ATCCATGCTGGTACA	

AB11	*Minutocellus* sp.	OL794292	TTAAACATCACAGCCGCTACGATGG AAGAAGTTTACAAACGTGCTGACTAC GCTAAGAAAATCGGTTCTATCATTG TAATGATCGATTTAGTAATGGGTTAC ACAGCAATTCAAAGTATTGCTTACTG GGCTCGTGAAAACGATATGTTATTAC ACTTACACCGTGCTGGTAACTCTACT TACGCTCGTCAAAAGAACCACGGTAT TAACTTCCGTGTTATCTGTAAATGGAT GCGTATGTCTGGTGTGGATCACATCCA TGCTGGTACA	
AB12	*Skeletonema grethae* (Zingone and Sarno)	OL794281	GCAACAATGGAAGAAGTATACAAACG TGCTGAGTATGCTAAAGCTGTTGGTTC TATCGWTGTTATGATCGATTTAGTAAT GGGTTACACWGCAATTCAATCAATTG CATACTGGGCTCGTGAAAACGATATGT TAWTACACTTACACCGTGCTGGTAACT CTACATACGCTCGTCAAAAGAATCACG GTATTAACTTCCGTGTTATCTGTAAAT GGATGCGTATGTCTGGTGTASATCAC ATCCRCRCTGGTACA	

AB13	*Chaetoceros* sp.	OL794282	TGAAGTAAAAGGTTCATACTTA AACATTACAGCTGCAACGATG GAAGAAGTATACAAACGTGCTG AGTATGCTAAAATGATCGGTTC TGTAATTGTGATGATCGATTTAG TAATGGGTTACACTGCAATTCAA TCAATTGCTTACTGGGCTCGTGA AAACGATATGTTATTACATTTACA CCGTGCTGGTAACTCTACTTACGC TCGTCAAAAAAATCATGGTATTAAC TTCCGTGTAATTTGTAAATGGATGC GTATGTCTGGTGTAGATCATATCCA CGCTGGTACAG	
NWU1	*Entomoneis* sp.	OL794287	TAAACATCACTGCTGGTACAATGGAA GAAGTTTACAAACGTGCTGAATATGC TAAAGCAGTAGGTTCTGTAATTGTTAT GATCGATTTAGTTATGGGTTATACAG CGATTCAATCAATTGCTTACTGGGCTC GTGAAAACGATATGTTATTACACTTAC ACCGTGCTGGTAACTCTACATACGCAC GTCAAAAAAATCATGGTATTAACTTCCG TGTTATCTGTAAGTGGATGCGTATGGC TGGTGTAGATCATATCCACGCTGGTACA	

NWU5	*Amphora incrassata* (Giffen)	L794288	TTAAACATTACTGGTGGTACAATGGAAGA AGTTTACAAACGTGCTGAGTACGCAAAAT CAGTAGGTTCTGTAATTGTGATGATCGAT TTAGTTATGGGTTACACAGCAATTCAAAGT ATTGCTCTTTGGGCTCGTGAAAACGATA TGTTATTACACTTACACCGTGCAGGTAAC TCTACATACGCTCGTCAAAAGAATCATGG TATTAACTTCCGTGTTATCTGTAAATGGA TGCGTATGTCTGGTGTAGATCACATCCA CGCTGGTACA	
NWU7	*Navicula* sp.	OL794291	TAAACGTTACTGCAGCTACTATGGAAGAA GTGTACAAACGTGCAGAGTACGCTAAAAT GGTTGGTTCTATCATTATCATGATCGATTT AGTAATGGGTTATACAGCAATCCAAAGTA TTGCTTTATGGGCTCGTGAAAACGACAT GCTTTTACACTTACACCGTGCTGGTAACT CTACTTCGCTCGTCAAAAGAACCACGGT ATTAACTTCCGTGTTATCTGTAAAGGATG CGTATGTCTGGTGTAGATCACATCCAC GCTGGTACAGT	

NWU8	*Nitzschia* sp.	OL794284	AAAGGTTCATACTTAAATGTTACAGCTGCT ACTATGGAAGAAGTATACAAACGTTGTGA GTATGCAAAAGAAGTAGGCTCTATCATTGT AATGATCGATTTAGTTATGGGCTACACAGC AATTCAAAGTGCTGCAATTTGGGCTCGTGA CAATGATATGCTTTTACATTTACACCGTGCA GGAAACTCTACATATGCACGTCAAAAAAATC ACGGTATTAACTTCCGTGTTATTTGTAAGTG GATGCGTATGTCTGGTGTAGATCATATCCAC GCTGGAACA	
NWU10	*Amphora* sp.	OL794290	TTAAACGTTACTGCCGGTACTATGGAAGAAG TGTACAAACGTGCTGAGTACGCTAAAAATGT GGGATCTGTAATTATCATGATCGATTTAGTTA TGGGTTACACAGCAATTCAAAGTATTGCTAT CTGGTCTCGTGAAAACGATATGCTTTTACAC TTACACCGTGCAGGTAACTCTACATACGCT CGTCAAAAAAATCATGGTATTAACTTCCGT GTTATTTGTAAATGGATGCGTATGTCAGG TGTAGATCATATTCACGCTGGTACA	