The Open University, London
Stazione Zoologica Anton Dohrn,
Naples (ARC center)

Integrated study of dinoflagellate diversity in the Gulf of Naples

Solenn MORDRET

Thesis submitted for the degree of Doctor in Philosophy (PhD) in Life and Biomolecular Sciences

December 2017

Abstract

Dinoflagellates are a diverse lineage of protists known as an essential component of marine planktonic communities. To study their seasonal diversity at the LTER-MC station in the Gulf of Naples, I used high throughput sequencing (HTS) metabarcoding of the V4 region in the 18S rDNA. To taxonomically identify the metabarcode sequences, I established a database, called DinoREF, of taxonomically verified and nomenclaturally updated 18S reference barcodes with associated metadata. The reference sequences were organised into Superclades based on phylogenetics and higher taxonomic treatment. DinoREF contains 1,671 sequences for 422 species and covering 22% of the described species. The database revealed that the V4 region alone cannot discriminate between some morphologically and genetically distinct species or genera. Moreover, many species and genera were collapsed together when clustered into 98% similarity OTUs. For the metabarcoding, dinoflagellate HTS V4 reads were gathered from 48 environmental DNA samples taken over three years at LTER-MC. Results of a-taxonomic cluster analysis showed three principal seasonal clusters, one with winter samples (16% of reads), one with mainly spring-summer samples (62%) and one with late summer-autumn samples (22%). Sorting reads into ribotypes and assigning them with DinoREF to taxa showed that reads belonging to the *Gyrodinium,* Gymnodiniales and Gonyaulacales Superclades were the most abundant. Winter samples were characterised by specific taxa thriving only in cold season. Results revealed 85 new records and detected 26 potentially toxic species for the Gulf of Naples. Many dinoflagellate genera such as *Tripos* are underrepresented in DinoREF because many species cannot be cultured. I applied single cell imaging, PCR and sequencing to gather 18S and 28S for 22 *Tripos* species. Using 18S V4 region as barcode I assessed the seasonal abundances of these species. Some were common all year whereas others showed distinct seasonality, mainly occurring in winter.

Acknowledgements

Firstly, I would like to thank to the Stazione Zoologica Anton Dohrn and the Open University for these three years of research. There are a number of people who have guided me and provided me with support to complete this thesis, and without their help this project would not have been possible.

I would like to thank my director of studies, Dr. Diana Sarno, for giving me this opportunity, and for her patience and advice throughout this project. Additional thanks to my internal supervisor, Dr. Wiebe Kooistra, who was particularly helpful during the write-up, whilst also providing essential assistance throughout my three years at Stazione Zoologica. I am also grateful to my external supervisor, Dr. Daniel Vaulot, for his valuable comments, suggestions and corrections. Likewise, to Dr. Marina Montresor for our conversations on the wonderful world of dinoflagellates, and for all her helpful comments on my manuscript.

Particular thanks are extended to Dr. Adriana Zingone for providing the HTS data, which was supported by the Italian project MIUR-FIRB Biodiversitalia (RBAP10A2T4). Thanks also to the MEDA unit for providing me with LTER MareChiara environmental parameters metadata. Also to LifeWatch-Italy (Bari) for sequencing eDNA.

A big thank you to the MareChiara team for providing plankton samples every week, without which this work would not have been possible. In particular I would like to thank the technicians who helped me throughout this project. For the plankton samples and culture, I thank Ferdinando Tramontano, Carmen Minucci and Alessandro Manfredonia. In the Molecular Biology Service, I am grateful to Ellio Biffali, Elvira Mauriello and Raimondo Pannone, and in the electron microscopy group, Annarita Graziano, Franco Iamunno and Giovanna Benvenuto.

I would also like to thank Thomas Mollica for being my extra pair of hands and for helping me conduct the experimental work on single cell extraction, amplification and sequencing.

Enormous thanks to my friend and colleague Dr. Roberta Piredda for teaching me the bioinformatic techniques required to process my data. I cannot express my gratitude for the daily support and patience that she gave me.

Also to my friends and labmates Dr. Chetan Gaonkar and Dr. Katerina Pargana for just being there for me.

Many thanks to all my friends at Stazione Zoologica Arianna, Angela, Daniele, Luigi, Isabella, Greta, Laura V., Laura E., MariPa, Eleonora, Isabella, Ennio, Davide, Gauri, Claudia, Christian, Filomena, Ylenia, Ivan and Filip… and also to my friends outside Stazione Zoologica, Elijah, Hristiana, Mariano, Felix, Celia, Michelle, Dina, Neri, Giulia, Giaccio, Elena, Jovana, Danila, Josh, Lucía, Swaraj, Tara, Mark, Paolo and many more I met along the way.

Finally, thank you to my family and to Chris for all their love and support.

Cette thèse est dédiée à mes parents, Dominique Abguillerm et Guy Mordret, pour leur soutien immense, leur amour et leur patience.

Table of contents

Integrated study of dinoflagellate diversity in the Gulf of Naples:

CHAPTER VI: Bibliography

CHAPTER VII: Appendix

List of Abbreviations

©/®, TM	Copyright/Registered, unregistered TradeMark
°C	Degree Celcius
µL	Microlitre
µm	Micrometer
18S (SSU)	Small Subunit Ribosomal DNA
28S (LSU)	Large Subunit Ribosomal DNA
BioMarKs	Biodiversity of Marine Eukaryotes
BLAST	Basic Local Alignment Search Tool
bp	Base pair
CCA	Canonical-correlation analysis
CEDiT	Centre of Excellence for Dinophyte Taxonomy
Cox gene	Cytochrome c oxydase gene
DCM	Deep Chlorophyll Maximum
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphate
e.g.	*exempli gratia* (for example)
et al.	*et alia* (and other)
etc.	*et cetera* (and so on)
GoN	Gulf of Naples
HCA	Hierarchical Cluster Analysis
HGT	Horizontal Gene Transfer
Hsp	Heat shock protein
HTS	High Throughput Sequencing
ICBN	International Code of Botanical Nomenclature
ICZN	International Code of Zoological Nomenclature
i.e.	id est (that is to say)
ITS	Internal Transcribed Spacer
LM	Light Microscopy
LTER-MC	Long Term Ecological Research station - MareChiara
ML	Maximum Likelihood
NCBI	National Center for Biotechnology Information
OTU	Operational Taxonomic Unit
PCR	Polymerase Chain Reaction
rbcL	Ribulose-Bisphosphate Carboxylase large subunit
RNA	Ribonucleic acid
rRNA	Ribosomal RNA
sensu	In the sens of
SZN	Stazione Zoologica Anton Dohrn
TRIS	tris(hydroxymethyl)aminomethane or tromethamine

List of Tables

Chapter II: DinoREF: a curated dinoflagellate reference database

Chapter III: Dinoflagellate diversity and seasonal changes at the LTER station in the Gulf of Naples as inferred from HTS metabarcode data

Chapter IV: Diversity of the dinoflagellate genus *Tripos* in the Gulf of Naples

List of Figures

Chapter I: General introduction

Chapter II: DinoREF: a curated dinoflagellate reference database

Chapter III: Dinoflagellate diversity and seasonal changes at the LTER station in the Gulf of Naples as inferred from HTS metabarcode data

Chapter IV: Diversity of the dinoflagellate genus Tripos in the Gulf of Naples

CHAPTER I: General introduction

In this chapter, Solenn Mordret made a review of literature for dinoflagellates and wrote the introduction chapter, which was revised by the supervisor team.

1.1. Introduction

Dinoflagellates are a heterogeneous group of protists present in virtually all aquatic ecosystems and occupying various ecological niches (Not *et al.*, 2012). Most dinoflagellates are mobile and have been named accordingly. *Dinos* comes from the Ancient Greek term for "whirling" or "rotation" and the Latin term *flagellum* meaning "whip" or "scourge". Dinoflagellates are unicellular living as individual cells, though some species sometimes form chains or exhibit fusion of many cells. Their cell size ranges from 5 to 2,000 µm (Hoppenrath, 2017). Together with Apicomplexa and Ciliates, Dinoflagellates constitute one of the most diverse phyla of Alveolates in the protist super group SAR (Stramenopiles Alveolates Rhizaria; **Fig.1.1.1**; Burki *et al.*, 2007; Adl *et al.*, 2012). To date, about 2400 living dinoflagellate species have been described (Gómez, 2012a; Guiry and Guiry, 2017; Hoppenrath, 2017) and around 2500 fossils (Taylor, Hoppenrath and Saldarriaga, 2008). They show an incredible diversity in shape and behaviour (**Fig.1.1.2**; Taylor, Hoppenrath and Saldarriaga, 2008; Gómez, 2012b) and new species are described every year (Hoppenrath, 2017). Recent biodiversity surveys have revealed the existence of many undescribed species (Massana *et al.*, 2015; Le Bescot *et al.*, 2016; Piredda *et al.*, 2017).

The vast majority of known dinoflagellate species are marine; only 17 % of them occur in freshwater ecosystems (Gómez, 2012b). Most are planktonic, though many species abound in marine benthic environments (Hoppenrath *et al.*, 2014). Compared to other protistan groups, dinoflagellates are considered poor competitors and rarely dominate the community (Smetacek, 2012; John *et al.*, 2014). Nevertheless, dinoflagellates exhibit complex genetic, morphological and physiological features that have allowed them to adapt to a wide range of environments (Not *et al.*, 2012). They have developed highly specialised life strategies performing many different key ecological functions (host, parasites or symbionts, autotrophs, heterotrophs or mixotrophs) and therefore add complexity to aquatic ecological networks (Taylor, Hoppenrath and Saldarriaga, 2008; Gómez, 2012b; Murray *et al.*, 2016). In general, dinoflagellates are ubiquitous and abundant

in marine planktonic and benthic ecosystems and contribute significantly to marine food webs (Not *et al.*, 2012).

Some dinoflagellate species are emblematic or notorious because of their direct impact on human wellbeing and the economy (Graham *et al.*, 2016). Examples are toxic species (*Alexandrium*, *Gambierdiscus*), symbionts of coral reefs (*Symbiodinium*) and bioluminescent species (*Noctiluca scintillans* and *Lingulodinium polyedra).*

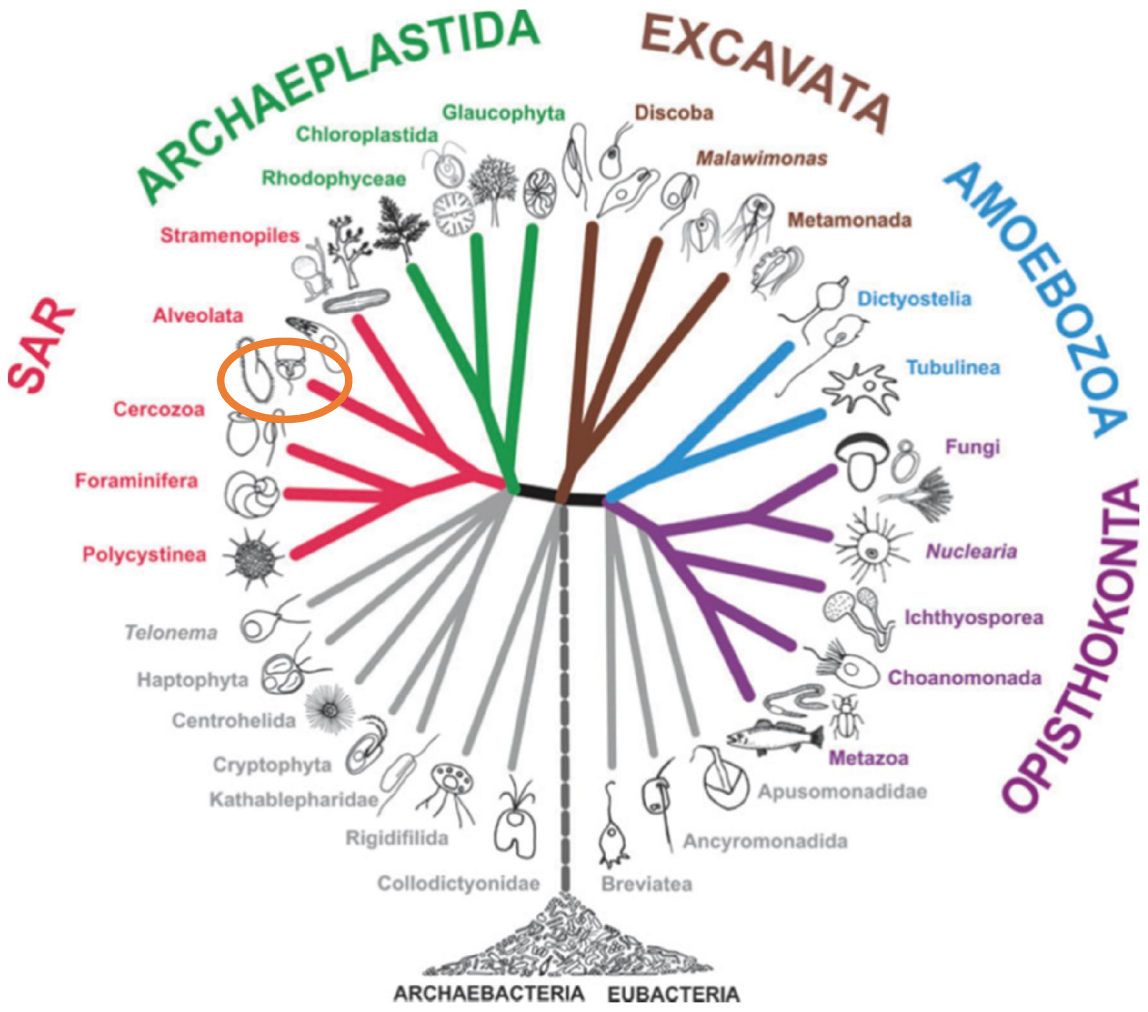


Fig.1.1 1: Schematic phylogeny of the eukaryotes according to Adl *et al.* (2012).

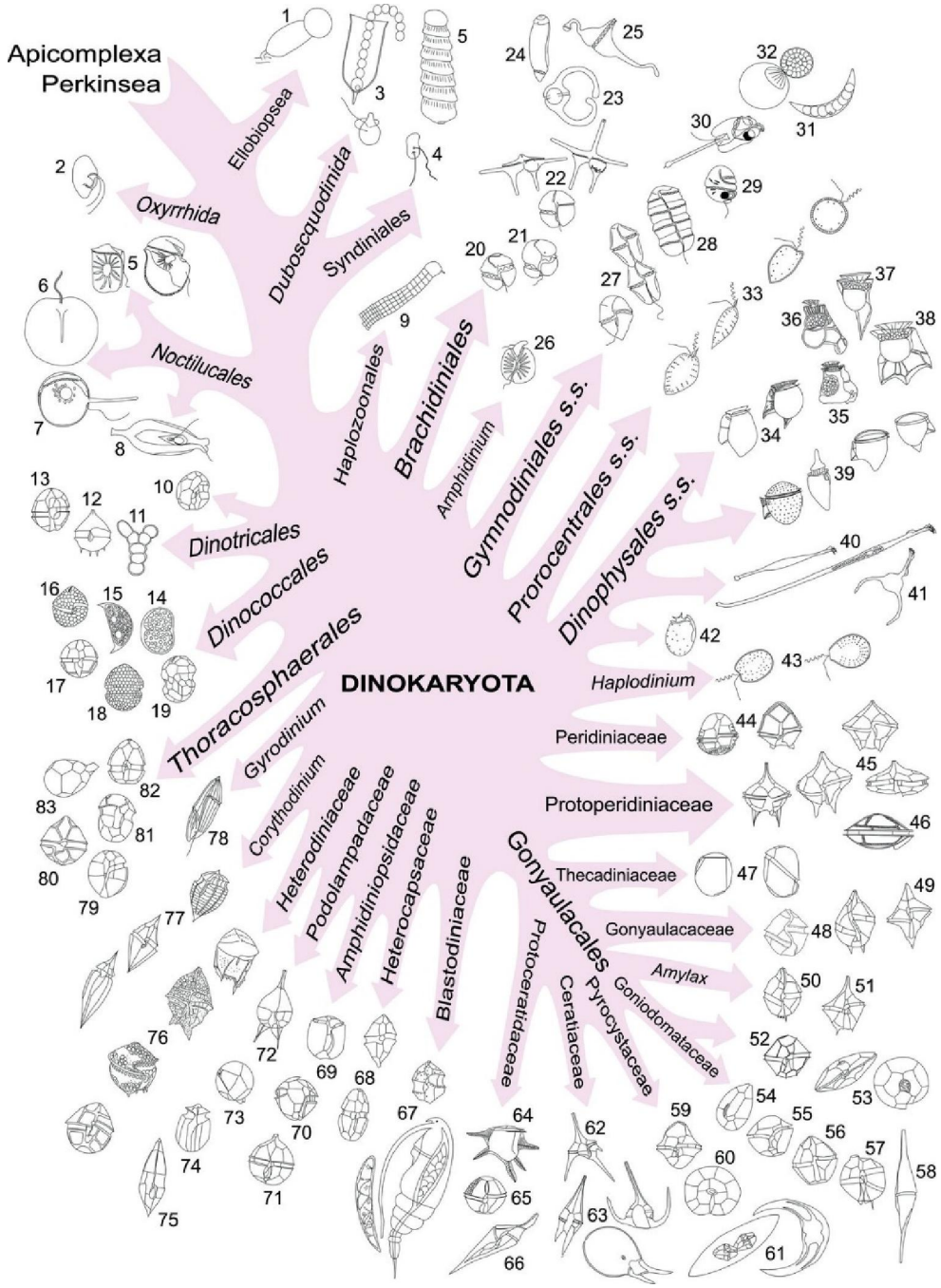


Fig.1.1.2: Diversity of the main lineages of dinoflagellates according to Gómez (2012b).

1.2. Cell Biology

1.2.a. General features of dinoflagellate cells

External morphology and morphological diversity

Morphologically, dinoflagellates are characterised by the presence of two different flagella. In the most widespread morphological type, i.e. dinokont flagellation, the transverse flagellum is located in a furrow encircling the cell called the cingulum (marked light green in **Fig.1.2.3**), and is used by the cell to spin around its central axis; the longitudinal flagellum is shorter and it is located in a ventral furrow, the sulcus (marked darker green in **Fig.1.2.3**), and propels the cell forwards (Gaines and Taylor, 1985; Fensome *et al.*, 1993). The two flagella emerge ventrally, from a pore positioned in the sulcus. In this morphological type, cells are composed of two parts: the epicone (top; marked red in **Fig.1.2.3**) and the hypocone (bottom; marked blue in **Fig.1.2.3**). In the desmokont flagellation (e.g. *Prorocentrum*), the two differentiated flagella do not merge with any furrow, but lie freely out of the anterior part of the cell. In a different way, podolampid dinoflagellates lack cingulum and depressed sulcus (Gómez, Moreira and López-García, 2010a).

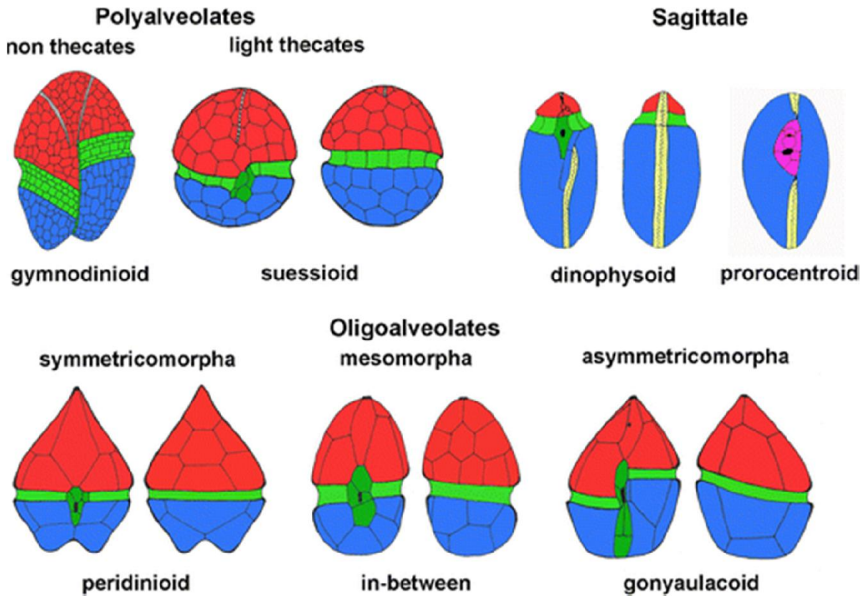


Fig.1.2.3: The six major types of tabulation observed in dinoflagellates, modified from Hoppenrath (2017).

As in other alveolates (ciliates and apicomplexans), dinoflagellate cells are covered with amphiesmal vesicles, called alveoli, organised in an intricate mosaic just inside the plasmalemma. Many dinoflagellate groups possess flat thecal plates inside the alveoli, and are collectively known as the thecate or armoured dinoflagellates. The plates are composed of cellulose or other polysaccharide microfibrils, which protect the cell and give it its shape. The remaining groups of dinoflagellates lack hard structures in their alveoli and are known as athecate or naked dinoflagellates. The number, the size, the shape and the arrangement of alveoli, known as tabulation, constitute the main criteria in dinoflagellate taxonomy for over a century. Six major types of tabulation can be distinguished: Gymnodinoid, Suessoid, Peridinoid, Gonyaulacoid, Dinophysoid, and Prorocentroid tabulation (**Fig.1.2.3**). Dinoflagellates are classified according to these tabulation patterns using the "Kofoid System" (**Fig.1.2.4**; see Hoppenrath, 2017). This conventional nomenclature system consists in naming and counting different rows of successive thecal plates in order to create a formula characterising each thecate species or genus. The formula describes plates starting from the top of the epicone to the bottom of the hypocone. The first plate named in a series is always positioned on the right near to the mid-ventral position and the count continues from the left to the right. In this system, special plates, such as intercalary, cingular or sulcal ones are each designated by a letter.

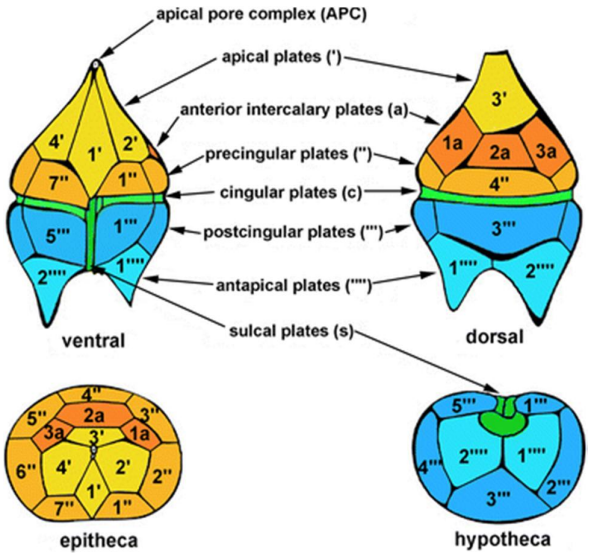


Fig.1.2.4: Kofoidian systemic describing the different plate pattern of thecate dinoflagellates, modified from Hoppenrath (2017).

Thecal plates provide clear characteristics to describe and differentiate dinoflagellates because they can be observed under the microscope. In addition, thecate dinoflagellates can be easily manipulated with micro-tools without squashing the cell. These plates help the cell conserve its shape and remain visible even if the cell is dead or fixed. Instead, in naked dinoflagellates, the alveoli are not easily observed, and lose their shape upon cell death or fixation, rendering the taxonomic treatment and identification of its species challenging. While thecate dinoflagellates possess a robust wall, unarmoured dinoflagellates do not; their cells are prevalently roundish (Gymnodinoid shape) and change shape easily. Parasitic or symbiotic dinoflagellates usually lack defined morphological characteristics.

Recently, apical surface structure arrangements have been shown to be phylogenetically meaningful traits for most of the major groups of athecate dinoflagellates and have been used to delineate and characterise lineages. For instance, the presence and the type of the apical structure complex (ASC or apical groove) has been adopted to distinguish among genera of naked dinoflagellates such as *Gymnodinium, Karenia, Karlodinium, Levanderina, Akashiwo,*and *Polykrikos* (Moestrup *et al.*, 2014; Takano *et al.*, 2014; Hoppenrath, 2017). Other characteristics including chloroplast types (e.g., in de Salas, Bolch and Hallegraeff, 2004; Jorgensen, Murray and Daugbjerg, 2004; Hansen, Daugbjerg and Henriksen, 2007; Hoppenrath and Leander, 2007b) or specific organelles (Moestrup and Daugbjerg, 2007) such as pyrenoids or eyespots, are used also as taxonomic characteristics to distinguish species or groups of them.

Inside a dinoflagellate cell

The dinoflagellate cell is considered among the most complex among eukaryotes, displaying a large variety of components. As in other eukaryotes, cells contain a nucleus, mitochondria, Golgi body, vacuoles and in the case of photosynthetic dinoflagellates, chloroplasts (**Fig.1.2.5**). Yet, several of these organelles exhibit characteristics specific to dinoflagellates.

One of the remarkable features distinguishing dinoflagellates from all other protists is their unique nucleus known as dinokaryon (**Fig.1.2.5**; Fensome *et al.*, 1993). Inside the dinokaryon, DNA is always condensed and organised into many chromosomes, but without nucleosomes or histones structuring them. Instead, core dinoflagellates possess their own unique specific packaging proteins called DVNPs or Dinoflagellate/viral nucleoproteins (Gornik *et al.*, 2012; Janouškovec *et al.*, 2017) and their own histone-like proteins. The number of chromosomes varies among species (from 4 to 200, Bhaud *et al.*, 2000), and the amount of DNA can exceed that of the human genome by several times (Murray *et al.*, 2016). Genes are usually present in multiple copies, from 30 to 5000 arranged in tandem repeats (Hou and Lin, 2009; Lin *et al.*, 2015). Many genes lack introns and the proportion of coding genes in the genomes seems very low (highly redundant genome, Hou and Li, 2009).

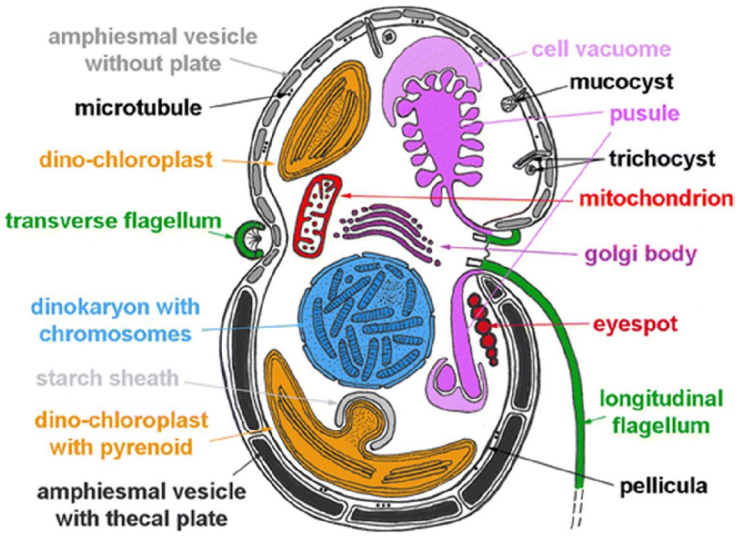


Fig.1.2.5: Principal organisation and organelles of a dinoflagellate cell, modified from Hoppenrath (2017).

Dinoflagellates possess specific mitochondria with tubular cristae. The mitochondrial genome seems to be reduced compared to that of other eukaryotes with only a few genes detected (cox 1, cob and cox3 for proteins and two ribosomal genes), many pseudogenes, non-coding and repetitive DNA or partial gene fragments. In addition, the genetic material has an unusual organisation with extensive transcript editing and a large number of inverted repeats motifs (Nash *et al.*, 2008; Waller and Jackson, 2009).

Materials are stored in cytoplasmic granules and lipid droplets. Food reserves in granules are retained in the form of starch (polyglucan-like polysaccharides), while lipid droplets store long-chained unsaturated fatty acids (Steidinger and Tangen, 1997). All dinoflagellates share a pusule, a distinctive type of vacuole with openings through the cell surface. This organelle can vary in complexity in different types of dinoflagellates but its function remains unknown (Hoppenrath, 2017). It may be involved in water content regulation (Graham *et al.*, 2016).

Photosynthetic dinoflagellates special features

About half of the dinoflagellate species possess chloroplasts (**Fig.1.2.5**) and therefore are able to photosynthesise. While most of the aveolates and early branching (basal) dinoflagellates are heterotrophic, having lost photosynthetic capacity during evolution, part of the dinoflagellates have kept the ancestral peridinin containing chloroplast inherited as a result of the original secondary endosymbiotic event that enabled photosynthesis in the last common ancestor of the entire "brown lineage" (Keeling, 2010). Since then, several dinoflagellate lineages have lost this ancestral chloroplast secondarily, and have regained the ability to photosynthesise through incorporation of various autotrophic protists independently in different lineages. These so-called tertiary endosymbiontic events have resulted in eight different types of chloroplasts, rendering dinoflagellate evolution unique among eukaryotes (Moestrup and Daugbjerg, 2007). For instance, the lineage Kareniaceae acquired its chloroplast as a result of a tertiary endosymbiosis event of a haptophyte cell. Members of the genus *Dinophysis* engulf cryptophytes and retain a still functional chloroplast. Dinotom dinoflagellates incorporated two evolutionary distinct plastids via endosymbiosis of two different diatom cells (Hehenberger *et al.*, 2014). These acquisitions are considered permanent in some cases, or temporary in others, i.e., in kleptoplasty in which functional plastids derived from an ingested algal prey. As a result of all the tertiary endosymbiosis events, photosynthetic dinoflagellates possess a wide variety of photosynthetic pigments able to absorb photons in a wide range of the light spectrum. The majority possesses chlorophyll a/c_2, and peridinin as the main pigments (Zapata *et al.*, 2012; Yamada, Tanaka and Horiguchi, 2015) but a

total of 63 different accessory pigments have been reported and these pigments are encountered in different proportions in different lineages of dinoflagellates. The pigments include betacarotene, dinoflagellate specific xanthophylls, chlorophyll b or c1, fucoxanthins and phycobiliproteins (Zapata *et al.*, 2012).

The chloroplast genome is arranged in many minicircular DNA, each minicircle coding for one, two, or in rare cases, three genes (Howe, Nisbet and Barbrook, 2008). Hence, in peridinin plastids the number of coding genes seems very low (no more than 16) compared to other microalgal lineages such as cryptomonads and diatoms, which retained around 165-185 genes (Green, 2004). Evidence shows that some of these missing genes have been transferred to the dinoflagellate nucleus (Hackett *et al.*, 2004), even if many seem to have been lost.

Recent studies also demonstrated the possibility of endosymbiotic gene transfer in tertiary plastid bearing dinoflagellates. For instance, Kareniaceae with haptophyte endosymbionts integrated nine genes, dinotom dinoflagellates incorporated 90 genes (Burki *et al.*, 2014), and the symbiontic species *Symbiodinium minutum*, a peridinin-containing dinoflagellate, has transferred up to 109 plastid genes to the nuclear genome (Mungpakdee *et al.*, 2014).

Another confirmation of lateral gene transfer in dinoflagellates is the Rubisco protein of the peridinin plastid, which happens to be of bacterial origin (Whitney, Shaw and Yellowlees, 1995), making dinoflagellate evolution one of the most affected by lateral gene transfer among protists and a significant driver of gene innovation (Wisecaver, Brosnahan and Hackett, 2013).

Pyrenoids are cellular micro-compartments found in protists, macroalgae and one lineage of land plants, including many species of dinoflagellates. These protein body structures, present inside the chloroplast, are involved in carbon (CO_2) concentration and fixation; and sometimes in starch formation and storage. Considering the high number of plastid bearing lineages in dinoflagellates, there is a wide diversity of pyrenoid types, even among peridinin bearing dinoflagellates. Five main types have been distinguished by Dodge and Crawford (1971) for dinoflagellates. Yet not all phototrophic dinoflagellate species display pyrenoids.

1.2.b. Special adaptations and complex organelles

From eyespots to complex eye-like structure

Dinoflagellates possess different types of complex photoreceptive organelles. Some species possess relatively simple eyespots (**Fig.1.2.5**) that are capable of detecting light signals and are responsible for phototaxis (Graham *et al.*, 2016). The eyespot, also known as stigma, can be composed of an aggregate of cytoplasmic globules positioned just beneath the cell membrane, or of layers of carotenoids integrating lipid droplets located between the thylakoid membranes or carotene droplets enclosed in multiple membranes (Hoppenrath and Saldarriaga, 2008). A group of phagotrophic predator dinoflagellates, known as the Warnowiaceae, possess a more complex photoreceptor, the ocelloid, an eye-like subcellular structure whose organisation is functionally analogous to that of the lens, cornea, iris and retina in animal eyes. Components of the "ocellus–eye" are believed to be composed of an aggregation of different organellar structures such as mitochondria (cornea) or pigments (retina) originating from different endosymbiosis events (Gavelis *et al.*, 2015). The mechanisms behind the "vision" of Warnowiaceae is poorly understood. One hypothesis states that the ocelloid confers an advantage to predator dinoflagellates because it enables phototaxis and habitat selection, rather than prey detection (Gómez, 2017).

Feeding tools

When it comes to feeding, heterotroph and mixotroph dinoflagellates have developed complex strategies. Three principal feeding mechanisms exist: direct engulfment of the whole cells, feeding through a feeding tube or feeding by means of a feeding veil (Jacobson, 1999). Dinoflagellates can feed on a wide range of prey including many other protist cells, other dinoflagellates and even animals such as nematodes, polychaetes and fish larvae and copepod eggs. Many naked dinoflagellates can stretch their body and possess a mouth-like aperture called a cytosome in order to consume prey directly. A few armoured dinoflagellates like *Fragilidinium* are known to have the same capacity, temporarily detaching their plates from the cell surface and allowing the assimilation of much bigger prey (Skovgaard, 1996). Other dinoflagellates feed thanks to a feeding

tube. This structure known as a peduncle and made of microtubules is used by some dinoflagellates to feed by myzocytosis (suction). The peduncle allows perforation of another membrane and feeding directly on the cell content, like a drinking-straw (Schnepf and Elbrächter, 1992). Dinoflagellates in the genera *Protoperidinium* and *Diplopsalis* have been observed to feed deploying a feeding veil known as pallium; a thin and flexible cytoplasmic extension allowing a dinoflagellate to wrap and capture other protist cells or even entire diatom chains (Jacobson and Anderson, 1986; Naustvoll, 1998). When enveloped in the pallium, prey are digested outside the dinoflagellate body and the pallium is retracted when done.

Extrusomes: defensive and prey capture adaptations

In addition to all common organelles, dinoflagellates present peripheral organelles secreting material at the exterior of the cells known as extrusomes. Extrusomes are budded off from the Golgi apparatus and can be of different nature and function. For example, trichocysts discharge defensive projectiles off the cell wall when disturbed (Livolant, 1982a; b) and mucocysts exude mucilage when stressed by environmental conditions (**Fig.1.2.5**; Hoppenrath and Leander, 2008). Nematocysts create extrusive filaments outside the cell of polykrikoid and warnowiid dinoflagellates (Hoppenrath and Leander, 2007a). It is believed that these complex defensive extrusomes have an important role in defence response against predators.

The warnowiid dinoflagellate *Erythopsidinium* has one of the most complex feeding strategies among dinoflagellate lineages. In addition to an ocelloid eye structure, *Erythropsidinium* displays a piston, a long tube-like contractile apparatus, demonstrating high-speed expansion and retraction. Recent studies (Gómez, 2017) show that this unique organelle could be used as environment-scanning tactile organelle for prey search, capable of attaching prey by a suction structure at its distal end. Pistons are also involved in locomotion of cells and accelerate cell mobility by producing jumps while swimming.

Yet some of these contraptions are used for hunting; some dinoflagellates such as warnowiids and polykrikoids possess harpoon-like nematocysts allowing prey capture (Gavelis *et al.*, 2017). These complex organelles act like weapons using ballistic mechanisms to harm the prey.

Scintillons and Bioluminescence

Dinoflagellates are responsible for a significant part of bioluminescence in the sea. More than 30 species have been reported to show this particularity, including a major part of cosmopolitan photosynthetic dinoflagellates but also some heterotrophic ones like *Noctiluca scintillans* and some *Protoperidinium* species. Bioluminescence is produced in cytoplasmic bodies known as scintillons or microsources. These intracellular structures, located at the cell periphery, contain luciferase enzymes, luciferin pigments and for some dinoflagellates luciferin binding protein. Bioluminescence is triggered by the chemical reaction between luciferase, luciferin and oxygen, creating short blue flashes of light. In dinoflagellates, bioluminescence is usually generated by external mechanical stimuli and is thought to be a defence mechanism against predators. It only occurs at night and is ruled by a circadian cycle. Some species of dinoflagellates include both bioluminescent and non-bioluminescent strains (Marcinko *et al.*, 2013; Valiadi and Iglesias-Rodriguez, 2013).

1.3. Life histories

Dinoflagellates are haplontic species, meaning that they spend the larger part of their life cycle dividing mitotically in a haploid (n) stage (**Fig.1.3.6**; Litaker *et al.*, 2002; Figueroa, Bravo and Garcés, 2006). A mother cell divides into two identical daughter cells by equal binary fission contributing to population growth and sometimes to large bloom formation. They are known to have a zygotic life cycle, in which the zygote is the only diploid (2n) stage. Sexual reproduction has been observed in a few species and is assumed to be widespread in dinoflagellates (Pfiester, 1989). Gametes are formed under specific environmental conditions and usually appear identical to vegetative cells, which complicates their identification as gametes (Graham *et al.*, 2016). Sexual reproduction can be homothallic or heterothallic depending if the fusing gametes derive from the same strain or from complementary mating-type strains (Figueroa *et al.*, 2010). The result of the fusion of two gametes is known as the "planozygote" (vegetative cell 2n). This bi-flagellate cell can divide meiotically to return to the haploid stage (n) or form a quiescent, resistant stage known as resting cyst.

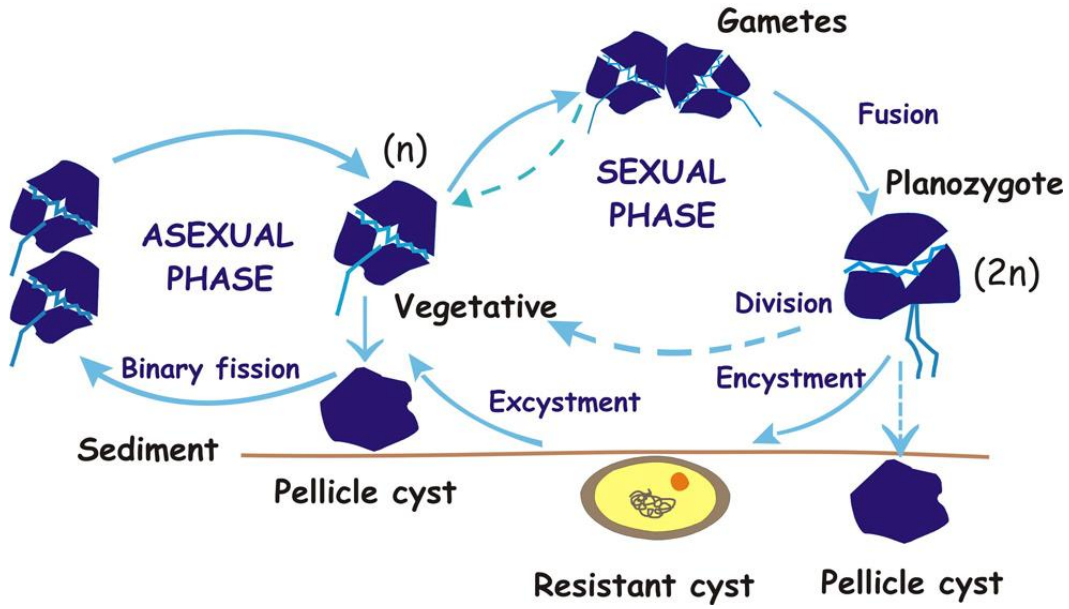


Fig.1.3.6: Dinoflagellate life cycle (ToLweb website dinoflagellate section Hoppenrath and Saldarriaga, 2008; modified after Walker et al., 1984). © Rosa I. Figueroa

Nearly 13-16 % of living dinoflagellate species have been reported to produce resting cysts (Taylor *et al.,* 2008). Encystment is mainly a survival mechanism, allowing dinoflagellates to survive unfavourable environmental conditions such as nutrient depletion (Doucette, Cembella and Boyer, 1989), unfavourable temperature (Grzebyk and Berland, 1996), allelopathic changes (Fistarol *et al.*, 2004) or even interactions with specific bacteria (Adachi *et al.*, 2004). Cyst formation can be enhanced during both sexual and asexual phases of their life cycle. When formed during sexual reproduction, non-motile cysts, called hypnozygots (2n – diploid stage), generally have a mandatory resting period. They can stay in dormancy for over a year in sediments where absence of light and low oxygen concentration inhibit germination. Temporary non-motile cysts can be formed also directly from vegetative cells to avoid unfavourable conditions, but these cysts can germinate easily.

Cell morphology and cell composition are transformed for encystment. From a motile free-living cell, cell walls grow thicker by the formation of a peripheral region grown from the cytoplasm and thecal plates, pulled out from the cell membrane, creating external ornamentations or spines (Graham *et al.,* 2016). Photosynthetic pigments can be reduced, storage products often increase and flagella are lost. In case of toxic species, cysts can also show an increase in toxicity in comparison with the vegetative cells (Persson *et al.*, 2006). Dinoflagellates belonging to the Thoracosphaeracean family are known to produce calcite-coated cysts (Van de Waal *et al.*, 2013; Graham *et al.*, 2016) whereas others produce silicified internal structures (Wetherbee et al., 2012). Dinoflagellates present various cyst morphologies, often differing markedly from that of their motile cells. Due to the robustness of their cell walls, fossil cysts have been found in coastal sediments worldwide (Zonneveld *et al.*, 2013). Cysts have been classified, often independently from their vegetative cells, and their morphology used as taxonomic criteria to identify different species and different layers through geological times. More importantly, cysts are used as tracers of dinoflagellate evolution (Wiggan, Riding and Franz, 2017). However, only some particular lineages of dinoflagellate life history have been studied in detail or have been preserved in the

fossil record. In reality, little is known about the life cycle, and life phases of the vast majority of dinoflagellates have still to be studied and documented.

1.4. Ecology

1.4.a. Diversity of habitats and ecological adaptations

Dinoflagellates thrive in virtually every aquatic habitat. The majority of the species are strictly marine (**Fig.1.4.7**), but several lineages have also colonised rivers, freshwater lakes and continental saline lakes. Some dinoflagellates can adapt to brackish environments and tolerate broad ranges and rapid changes in salinity and oxygen concentration, as encountered in estuaries and mangrove swamps. Marine environments also offer different niches to be exploited by dinoflagellates. These niches can be categorized into planktonic or benthic, coastal or open-ocean pelagic, sunlit or deep-sea ocean waters, or even sea ice (Buck, Bolt and Garrison, 1990; Murray *et al.*, 2016; **Fig.1.4.7**).

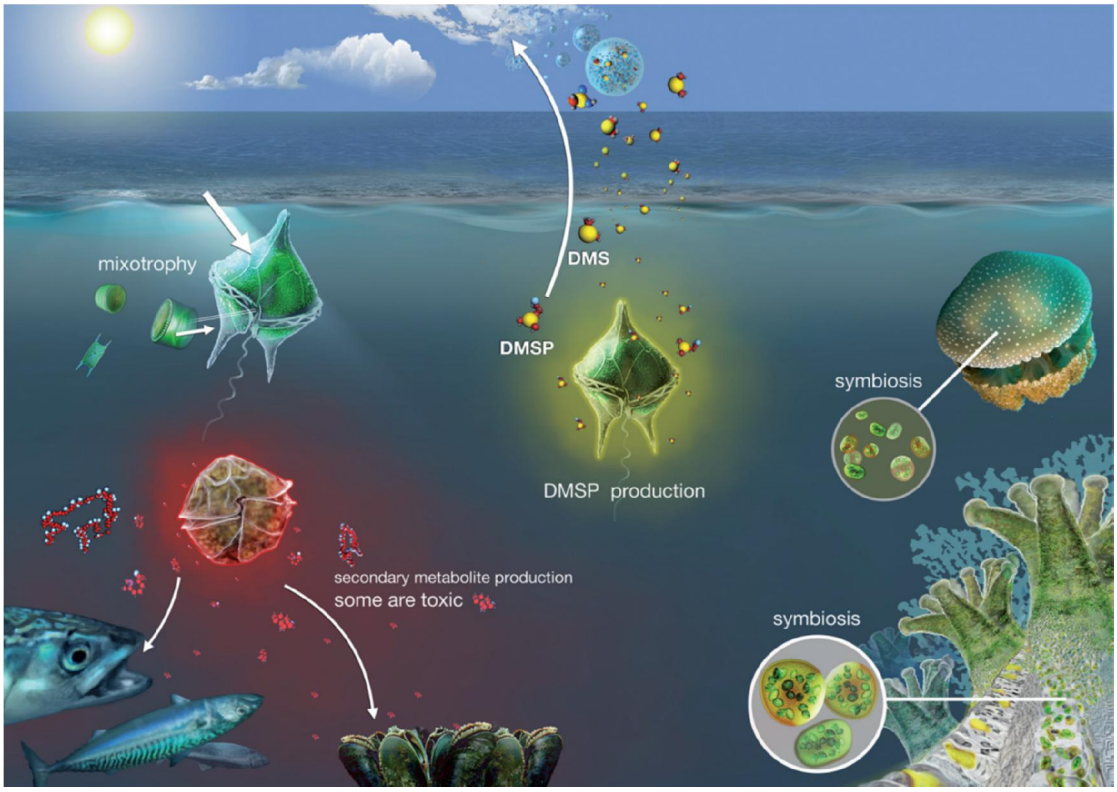


Fig.1.4.7: Illustration of the ecology of dinoflagellate cells, modified from Murray *et al.* (2016).

Among plankton, dinoflagellates rarely dominate planktonic assemblages, growing slower than other protists, such as diatoms or other flagellates (Smayda, 1997). However, some dinoflagellate

species can create massive blooms under specific nutrient, turbulence and light conditions. Photosynthetic dinoflagellates are mainly limited by phosphorus and nitrogen concentrations in the water (Not *et al.,* 2012). Heterotrophic dinoflagellates are found in greatest abundances during mid and late summer, following the temperate diatom spring bloom (Hoppenrath and Saldarriaga, 2008). Growth of bloom-forming dinoflagellates is favoured mainly by non-turbulent conditions. High nutrient concentration may favour growth of toxic dinoflagellates (Not *et al.*, 2012; Smetacek, 2012). The most notable blooms of planktonic dinoflagellates occur in coastal regions (Not *et al.,* 2012) where concentrations during a bloom can reach 10^7-10^8 cells per litre (Taylor, Hoppenrath and Saldarriaga, 2008). In the open ocean, dinoflagellates are less abundant, but usually display a higher diversity (Gómez, 2012b). This is linked to the fact that through various strategies, dinoflagellates have adapted to specific conditions to occupy different niches in the open ocean. For instance, in the sunlit oligotrophic waters such as the tropical zone, some dinoflagellates have adapted their morphology and chemistry, and evolved in symbiosis with other organisms to thrive and dominate in this specific environment (Decelle, Colin and Foster, 2015). Other dinoflagellates are thought to have specialised in deeper layers of the ocean, under the photic zone. However, much less is known about these pelagic dinoflagellates and their dynamics due to the difficult accessibility of open ocean sites for research (Gómez, 2014).

Dinoflagellate diversity is also high in the benthic coastal zone. Many dinoflagellates live associated with a particular substrate in shallow waters and are well adapted to these habitats. Among them, some lineages are known as psammophilic and dwell in the sand of beaches, coral rubble, tidal sand flats or tidal pools. Others are epiphytic or epizoic. Benthic dinoflagellates can be photosynthetic, heterotrophic or mixotrophic and can also form blooms. About 180 benthic species in 38 genera have been described worldwide (Hoppenrath *et al.*, 2014), but most of the benthic dinoflagellates remain unexplored. In addition, many dinoflagellate cysts can be found in the sediment, rendering these habitats important for their life cycle.

1.4.b. Biogeographic distribution

Dinoflagellates are present from the polar regions to the equatorial zone and display the same patterns of distribution as other groups of protists (Dolan, 2005). Their distribution has been qualified as "modified latitudinal cosmopolitanism" by Taylor (1987; 2004) to describe the distribution of the same morphospecies communities at the same specific climatic latitudes in the northern and southern hemispheres. The principal factors influencing the biogeography of dinoflagellates are temperature and currents. Circumtropical communities are similar in different oceans. In the same way, many species present in the Arctic Ocean can also be found in the Antarctic under similar environmental conditions (Montresor *et al.*, 2003).

Usually a clear distinction exists between assemblages from the coast versus the open ocean. In fact, neritic dinoflagellates have a life cycle involving frequent encystment transition periods. Therefore, this relation with the benthos restrict their distribution to shallow waters (Taylor, Hoppenrath and Saldarriaga, 2008). In many neritic dinoflagellates, growth is dependent on intermittent nutrient input from the land while those occurring in the open ocean are adapted to open ocean conditions (e.g. photosymbiosis in tropical areas).

Rare cases of endemism have been reported for dinoflagellates and are mainly due to the extreme singularity of certain environment such as polar habitats, isolated lentic habitats or internal seas (Taylor, 1987; Buck, Bolt and Garrison, 1990; Moestrup *et al.*, 2006). Some morphospecies have only been reported to occur in some specific sea or ocean (Gómez, 2006). However, some of these morphospecies have been reported only once and their status have still to be confirmed (Taylor, Hoppenrath and Saldarriaga, 2008).

Molecular signatures are starting to emerge to define more precisely the biogeographic patterns of dinoflagellates. Metabarcoding has been used to assess global dinoflagellate patterns (Massana *et al.*, 2015; Le Bescot *et al.*, 2016), but this approach does not always allow evaluation of patterns for individual species precisely (Le Bescot *et al.,* 2016). However, globally distributed and abundant

groups of dinoflagellates such as *Alexandrium* or *Symbiodinium* have been studied intensively and their worldwide distribution provides first insights into biogeography and population dynamics.

Symbiodinium diversity is represented in nine lineages (clades A to I; Pochon, Putnam and Gates, 2014). This small dinoflagellate belonging to the Suessiales is known to be symbiont of a vast diversity of benthic hosts in reef ecosystems, including corals, anemones, bivalves, sponges, as well as unicellular protists such as foraminifera and ciliates. Next to reef ecosystems, *Symbiodinium* is also symbiont of bentho-pelagic jellyfish and pelagic ciliates. Recent studies on this genus have demonstrated that the distribution of the genus and its different clades can be host specific (Pochon, LaJeunesse and Pawlowski, 2004; Santos *et al.*, 2004; Pochon and Gates, 2010) but is also linked with environmental and ecological factors such as temperature optimum, irradiance tolerance, water clarity, depth at which the host abounds and resistance to stressful conditions (LaJeunesse *et al.*, 2010; Bongaerts *et al.*, 2015). In some cases, the clade composition of *Symbiodinium* communities was shown to differ between similar reef ecosystems worldwide (Pochon, LaJeunesse and Pawlowski, 2004; Goulet, Simmons and Goulet, 2008; LaJeunesse *et al.*, 2010). Other studies revealed genetic differences between *Symbiodinium* symbionts of the ciliate *Tiarina* originating from the Atlantic, Pacific and Indian Ocean, the Red Sea and the Mediterranean Sea (Mordret *et al.*, 2016).

The genus *Alexandrium* is globally distributed and known for its capacity to grow fast, allowing it to form massive blooms (John *et al.*, 2014). About 33 species are described (Guiry and Guiry, 2017), many of them being toxic. Several of these can co-occur from subarctic to tropical shallow waters of the Northern and Southern Hemisphere (Anderson *et al.*, 2012). Because many strains of the same species are distributed worldwide, genetic analyses showed links between the distribution and evolution of different populations with paleo-geo-oceanographic conditions and some other factors like eutrophication (John, Fensome and Medlin, 2003). Recent studies based on particular species of *Alexandrium* reveal genetic structure of populations correlated with geographic

distances and connectivity (Lilly, Halanych and Anderson, 2007; Kremp *et al.*, 2009; McCauley *et al.*, 2009).

1.4.c. Nutritional strategies

Dinoflagellates display a vast diversity of trophic strategies (**Fig.1.4.7**). Most dinoflagellates are motile and can respond to external ecological stimuli such as light intensity, nutrient availability, chemical signals from prey or symbionts and adapt to their immediate environment. Free-living dinoflagellates can be photosynthetically autotrophic, mixotrophic (i.e. able to both photosynthetise and absorb organic food) or strictly heterotrophic. They can also be osmotrophic, assimilating organic material directly, or phagotrophic, feeding on other organisms or organic particles. They can be considered producers or consumers in marine food chains, or perform both functions at the same time (Smalley and Coats, 2002), and therefore, cannot be represented as an uniform group in ecological modelling (Flynn *et al.*, 2013).

According to Gómez (2012b), 49 % of the dinoflagellates described morphologically do not contain any plastids. These dinoflagellates are considered heterotrophic and their majority qualify as prey-specific predators (See this Chapter, part 2.b. Special adaptations and complex organelles - Feeding tools).

Some plastid bearing dinoflagellates are entirely autotrophic and grow easily in culture, whereas many others can grow only if they also ingest prey. This mixotrophic lifestyle is widespread and the balance between phagotropy and photosynthesis varies between species and the stability of their plastids (in case of kleptoplastidy, the plastids are not permanent and new ones need to be ingested from time to time; (Stoecker, 1999; Stoecker *et al.*, 2009). The types of plastids found in different lineages of dinoflagellates contain a wide diversity of pigments, some being unique to dinoflagellates and capturing light energy across a major part of the spectrum. Remarkably, a recent study show that the pigment composition of plastid bearing dinoflagellates can be directly linked to the habitat, demonstrating again the incredible adaptability of dinoflagellates to their environment (Yamada, Tanaka and Horiguchi, 2015).

Most dinoflagellate are motile allowing them to react and adapt to non-favourable conditions or stimuli, giving them an advantage over non-motile protists such as diatoms (Not *et al.*, 2012). Non-motile dinoflagellates exist among parasitic and symbiotic dinoflagellates but usually conserve motile gametes or a motile stage when free from the host (e.g. in Trench, 1993; Skovgaard, Karpov and Guillou, 2012). These species generally exhibit reduced morphological features. In the same way, other dinoflagellate like *Pyrocystis noctiluca* display an important part of their life cycle in a planktonic non-motile vegetative stage (Seo and Fritz, 2000). Regardless of the different nutritional strategies, dinoflagellates have developed an important capacity to store nutrients by means of food vacuoles that are much more evolved than in other protists, giving them an advantage to survive in temporary adverse conditions (Graham *et al.*, 2016).

1.4.d. Symbiotic associations

Symbiotic associations are common in dinoflagellates and many life strategies include interactions between a wide diversity of organisms (Decelle, Colin and Foster, 2015; Murray *et al.*, 2016). Compared to other unicellular eukaryotes, they show an enhanced propensity to form symbiotic partnerships, as demonstrated by the diversity of acquisitions of plastids and organelles from various partners in the different lineages of dinoflagellates. Associations involving dinoflagellates range from total mutualism to parasitism. Some relationships are permanent, while others are not obligatory and considered unstable.

Mutualistic interactions

When considering mutualist dinoflagellates, the well-known photosymbiosis between corals and the *Symbiodinium* dinoflagellates comes to mind (LaJeunesse, 2001). This particular form of symbiosis sustains one of the most important ecosystems on earth, coral reefs (Stanley, 2006). Many other mutualistic interactions also exist, for instance in pelagic ecosystems. Naked and thecate photosynthetic species belonging to the genera *Symbiodinium, Pelagodinium, Heterocapsa, Azadinium, Scrippsiella* or *Amphidinium* have been found to be symbionts of other protists such as radiolarians and ciliates, and even of invertebrates such as clams, anemones,

jellyfish and flatworms (e.g. in Fitt, 1985; McNally *et al.*, 1994; Trench and Thinh, 1995; Lobban *et al.*, 2002; Mordret *et al.*, 2016).

In addition, dinoflagellates themselves can host other species. For instance, several members of the order Dinophysiales harbour specialised prokaryotic and/or eukaryotic symbionts in their cytoplasm or in a small chamber with openings on the outside. *Amphisolenia* species host both internal eukaryotic photosynthetic symbionts belonging to pelagophytes and external prokaryotic cells (Daugbjerg, Jensen and Hansen, 2013). In *Ornithocercus*, *Histioneis* or *Citharites* cells a mix of bacteria and cyanobacteria can be hosted externally in the girdle list, or in a chamber in the girdle floor (Foster, Carpenter and Bergman, 2006; Decelle, Colin and Foster, 2015). Some *Noctiluca* strains from the Indian Ocean also host photosynthetic prasinophytes (Sweeney, 1976).

Parasitic interactions

Parasitic relationships are also numerous among dinoflagellates and in particular among early branching lineages, but parasitic dinoflagellates can be found in several different orders of dinoflagellates (Horiguchi, 2015). Some parasitic dinoflagellates infect other protists in order to complete their life cycle. For instance, the three genera *Duboscquella, Duboscquodinium* and *Tintinnophagus* are specialist parasites of tinitinnid ciliates, i.e. *Favella ehrenbergi, Eutintinnus fraknoii* and *Tintinnus acuta* respectively, but they are not phylogenetically closely related (Harada, Ohtsuka and Horiguchi, 2007; Coats *et al.*, 2010). Members of the genus *Amoebophrya* parasitise a range of other dinoflagellates (Fritz and Nass, 1992) while *Paulsenella* is known as an ectoparasite of diatoms (Drebes and Schnepf, 1988). However, the majority of parasitic dinoflagellates are specialised in infecting crustaceans: *Blastodinium* and *Syndinium* species infect different species of copepods (Skovgaard, 2005; Skovgaard, Karpov and Guillou, 2012), *Hematodinium* species are parasites of commercially important crab and lobster species (Stentiford and Shields, 2005). Some dinoflagellates are known to cause problems for fisheries and aquaculture. *Ichtyodinium* parasitise fish eggs such as those of tuna and sardines (Gestal *et al.*, 2006). Freshwater *Piscinoodinium* and marine *Amyloodinium* infect gills of common farmed fish (Levy *et al.*, 2007). Rare cases of

dinoflagellate parasitism are reported in marine polychaetes (Rueckert and Leander, 2008) and appendicularians (Gómez and Skovgaard, 2015).

1.4.e. Secondary metabolites

Production of toxins

Dinoflagellates produce ecologically and economically important secondary metabolites, including a number of biotoxins (**Fig.1.4.7**). Indeed, 75-80 % of the toxic phytoplankton species belong to the dinoflagellates (Cembella, 2003) and 95 different species are registered on the Taxonomic Reference List of Harmful MicroAlgae (IOC – UNESCO- 2017 update; Moestrup *et al.*, 2009 onwards). Among these, about 60 species, all of them marine, are recognised to have harmful effects on animals including birds, fish and mammals. Dinoflagellates developing harmful blooms are mainly photosynthetic and thrive in estuarine or coastal environments. The toxins synthesised during the blooms usually accumulate in shellfish or fish. Ingestion of such contaminated seafood causes different poisoning syndromes depending the type of toxin assimilated (Hallegraeff, 2004). Poisoning syndromes induced by dinoflagellates can be separated in five categories (Wang, 2008): Paralytic shellfish poisoning (PSP), neurotoxic shellfish poisoning (NSP), diarrheic shellfish poisoning (DSP), ciguatera fish poisoning (CFP) and azaspiracid shellfish poisoning (ASP). PSP is caused by saxitoxins and gonyautoxins produced by *Alexandrium, Pyrodinium* and *Gymnodinium* species. Brevetoxins synthetised by *Karenia* species result in NSP. CFP can be caused by both ciguatoxins and maitotoxins from *Gambierdiscus* species. Some *Azadinium* species were found responsible for Azaspiracid shelfish poisoning involving serious human incidents in Northern Europe (Tillmann *et al.*, 2009). Finally, some *Dinophysis* and *Prorocentrum* species can produce okadaic acid causing diarrheic poisoning (DSP).

Related to ciguatoxins and brevetoxins, yessotoxins produced by *Protoceratium reticulatum, Lingulodinium polyedra* and *Gonyaulax spinifera* can affect liver, pancreatic and heart function in mice (Tubaro *et al.*, 2010). They have been classified in their own toxin category accumulating in shellfish but no case of deleterious effects in humans have been reported yet.

Palytoxin-like compounds produced by some species of *Ostreopsis* are considered an emerging issue (Biré *et al.*, 2015). This complex fatty alcohol acts as powerful vasoconstrictor leading sometimes to serious illness if ingested or inhaled. In summer 2005 and 2006, along the Ligurian coast (North-Western Mediterranean), hundreds of people were reported ill after a swim in the sea (Brescianini *et al.*, 2006). Recent studies demonstrated toxicity of *Ostreopsis ovata* (Ovatoxin-a), *Ostreopsis mascarenensis* (Mascarenotoxins) and *Ostreopsis siamensis* (Ostreocin-D)(Ciminiello *et al.*, 2010). Though, the exact function of all these toxins remains unclear, their production could be associated with cell osmoregulation or may act as a deterrent against predators and limit grazing during a bloom (Murray *et al.,* 2016).

Production of DMSO

Dinoflagellates are responsible for a significant part of the production of DMSP metabolites in the ocean (**Fig.1.4.7**, Caruana *et al.*, 2012). DMSP is one of principal precursors of DMS metabolites, which in their turn are known to play a crucial role in biogeochemical cycles and the global climate. Most dinoflagellates are known to synthetize DMSP, but the concentrations vary greatly among species (Caruana and Malin, 2014). *Symbiodinium* species, *Prorocentrum* species, *Gyrodinium impudicum*, *Scrippsiella acuminata*, *Dinophysis acuminata* and *Heterocapsa pygmaea* are the biggest known producers of DMSP. The precise role of DMSP for dinoflagellates is still unknown (Murray *et al.,* 2016) but its production seems essential considering that the wide diversity of dinoflagellate (naked as well as thecate) tested showed this capacity to produce DMSP. Studies have suggested that it may act as osmolyte, cryoprotectant, antioxidant, specialised metabolite and/or a defensive element (Caruana and Malin, 2014).

Production of economically important metabolites

Dinoflagellates can produce economically important primary molecules such as lipids with very long chains (Murray *et al.,* 2016). For instance, *Crypthecodinium cohni* is cultured for its production of the polyunsaturated fatty acid docosahexanoic acid (DHA), which is commercially important as nutraceutical (omega-3 dietary supplement) and as aquaculture feed stock (Mendes *et al.*, 2009).

Dinoflagellates are considered potential sources of unique and complex polyketides and sterols. For instance, dinoflagellates are able to synthetise more than 35 different types of sterols (Robinson *et al.*, 1984) and polyketides, being precursor of the majority of dinoflagellate toxins (Rein and Snyder, 2006). These molecules are believed to have high therapeutic value, and understanding their synthesis pathways represents a major future research area. Dinoflagellates are emerging targets of biotechnological applications ever since researchers have used them as potential model organisms for genetic manipulation and mass production of various types of fatty acids, e.g., for biofuels (Radakovits *et al.*, 2010).

1.5. Evolutionary history

1.5.a. Fossil record and origin

The origin and evolution of the dinoflagellate lineages can be traced through geological times through fossil records of thecate dinoflagellates or dinoflagellate resting stages (**Fig.1.5.8**, Finkel *et al.*, 2007; Janouškovec *et al.*, 2017; Wiggan, Riding and Franz, 2017). Organic microfossils known as acritarchs are encountered in strata all the way down into the early Palaeozoic (**Fig.1.5.8**; Brocks and Summons, 2005), but the nature of these fossils is enigmatic. Many acritarchs are hypothesised to be cysts of protistan lineages ancestral to one or more lineages of modern dinoflagellates (Graham *et al.*, 2016). Some acritarchs exhibit a morphology similar to that of modern dinoflagellate cysts, displaying complex projections and ornamentations (Wall and Dale, 1969). Nonetheless, microfossils attributed to dinoflagellates with high confidence first appear in the fossil record at 240 million years before present (BP), i.e., the Early Triassic period (Fensome *et al.*, 1993; Taylor, Hoppenrath and Saldarriaga, 2008). Sediments from the Mid Jurassic to Cretaceous periods show a rapid radiation of dinoflagellates through an increase of the abundance and diversity of fossils recovered (Janouškovec *et al.*, 2017). This radiation probably corresponds to the diversification of thecate dinoflagellates and the explosion of Gonyaulacoid and Peridinoid lineages.

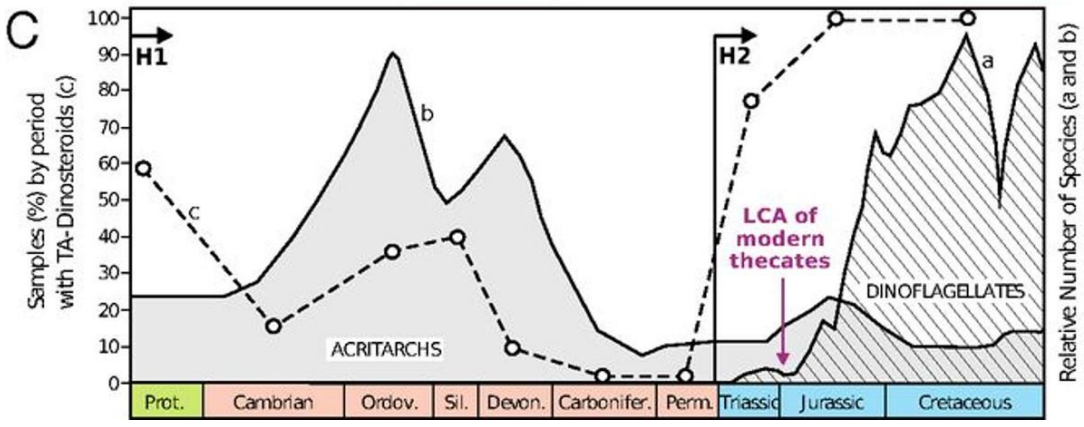


Fig.1.5.8: **a**. Dinoflagellate fossil records (dotted line). **b**. Abundance of acritarchs recovered in sediments in different geological times (line, grey filling). **c**. Dinosterol concentration in the sediments. LCA= Last Common Ancestor. **H1**: First hypothesis of appearance of dinoflagellates. **H2**: Second hypothesis of appearance of dinoflagellates. Figure modified from (Fensome, Saldarriaga and M. F. J. R. Taylor, 1999; Janouškovec *et al.*, 2017).

Examination of steroids produced by dinoflagellates has also helped understanding dinoflagellate evolution because these steroids can be related to geologically stable biomarkers found in petroleum. Dinosterol and dinosterane steroids have been shown to be an indicator of dinoflagellate presence (Fensome *et al.*, 1993). These streroids, often found in sediments, are synthetised by many modern dinoflagellates and rarely produced by other protists making them excellent chemical biomarkers. However, even if these compounds seem to be produced by most dinoflagellate lineages, they have not been found in more basal naked groups such as the Kareniaceae, *Gyrodinium* and *Amphidinium* species. In the same way, all early branching dinoflagellates, like *Noctiluca scintillans* or close relative Apicomplexans do not show the capacity to synthesise dinosterols (Janoušk ovec *et al.*, 2017). These results suggest that thecate dinoflagellates and their close relatives first appeared and diversified in the Early Jurassic (**Fig.1.5.8**) and that the core naked dinoflagellates probably are of Late Paleozoic or Early Mesozoic origin. Estimates based on molecular clock calculations inferred from earliest appearances of dinoflagellate fossils and molecular phylogenies of existing dinoflagellates place the origin of the modern dinoflagellate diversity well into the Paleozoic (Berney and Pawlowski, 2006; Parfrey *et al.*, 2011; Wiggan, Riding and Franz, 2017) supporting the notion that some of these Paleozoic acritarchs are, indeed, dinoflagellates. Molecular clock calculations place the emergence of Peridiniales and Gonyaulacales at 190 MYA and 180 MYA – i.e. Early Jurassic (John, Fensome and Medlin, 2003). In the same way, Shaked and De Vargas (2006) placed the first appearance of Suessiaceae in the Mid Jurassic.

1.5.b. Genetic diversity of modern dinoflagellates

Since the emergence of molecular techniques to study dinoflagellates, data has accumulated, especially for ribosomal markers commonly used in phylogeny and ecology. This data shows a vast diversity in the dinoflagellates, but phylogenies do not always corroborate the traditional taxonomy based on the morphology. One of the main markers historically used to characterise

dinoflagellate is the 18S rRNA encoding gene region in the rRNA operon (**Fig.1.5.9**). The first dinoflagellates 18S rRNA sequence was published in GenBank in 1993 (Rowan and Powers, 1991; McNally *et al.*, 1994). This marker is, together with the nearby 28S rRNA marker, the most commonly used for taxonomic purposes (Gómez, 2014).

The 28S rRNA is known to discriminate among species of dinoflagellates; it is currently used to identify strains and to assess the phylogenetic position of newly discovered species. For highly diverse genera of dinoflagellates like *Symbiodinium* or *Scrippssiella*, the internal transcribed spacer 2 (ITS2) marker between the 5.8S and 28S rRNA coding regions of the nuclear ribosomal operon (**Fig.1.5.9**; (LaJeunesse, 2001; Montresor *et al.*, 2003) is preferred and diversity usually expressed in term of "clades" or species complex. Other markers such as Hsp, ITS1, COI, cox or rbcL genes (photosynthetic dinoflagellates) have been used in phylogenetic studies (Litaker *et al.*, 2007; Hoppenrath and Leander, 2010; Stern *et al.*, 2010, 2012; Pochon, Putnam and Gates, 2014), but none of them show the coverage exhibited by the 18S or 28S markers, and these are markedly biased towards cultivable autotrophic species (Gómez, 2014).

Fig.1.5.9: Visualisation of a ribosomal operon, modified from Mordret *et al.*, (2016).

Nuclear ribosomal markers or any other marker on its own do not exhibit a good phylogenetic resolution of the various dinoflagellate lineages, as the basal ramifications in the tree obtain insufficient support (Fensome, Saldarriaga and Taylor, 1999; Daugbjerg *et al.*, 2000; Saldarriaga *et al.*, 2001, 2004; Zhang, Bhattacharya and Lin, 2007; Bachvaroff *et al.*, 2014). This is probably due to a rapid diversification of dinoflagellates following the last common ancestor of the extant diversity (Murray *et al.*, 2005). Therefore, the phyletic status of many known dinoflagellate higher taxa remains unresolved. Yet, in spite of this issue, the 18S rRNA constitutes the most commonly applied marker for taxonomic identification and phylogenetic placement because of its high

coverage of the dinoflagellate diversity (Gómez, 2014). In fact, most of the species are described using only nuclear ribosomal data. The 18S rRNA is an easy obtainable marker considering its high number of copies in dinoflagellate genomes and the availability of commonly used universal eukaryotic primers. The marker is also the most commonly used in single cell amplification studies targeting heterotrophic dinoflagellates that are still unculturable (Gómez, 2014). Equally important, in meta-barcoding studies performed on protist communities, the main markers used are regions within the 18S, such as the variable V4 or V9 regions. Remarkably, dinoflagellate sequences are both the most abundant and diverse ones in environmental meta-barcoding studies targeting protists (de Vargas *et al.*, 2015; Massana *et al.*, 2015; Piredda *et al.*, 2017). Nonetheless, detailed analyses of environmental data for dinoflagellates are just starting to emerge. This type of studies is mainly limited by the lack of references representing the major diversity of the dinoflagellates.

1.5.c. Molecular phylogenies

The recent phylogenetic literature based on multigene alignments suggests that many dinoflagellate orders and genera are polyphyletic or paraphyletic. A phylogeny based on multiple gene markers (Orr *et al.*, 2012; **Fig.1.5.10**) reveal that the naked dinoflagellates evolved first and the theca was acquired only once, more recently. Before the rise of molecular phylogenies, most taxonomists used to group all thecate dinoflagellates in a single taxonomic group and all naked dinoflagellate in another one. However, from a molecular phylogenetic viewpoint, it appears that both naked and thecate dinoflagellates include different lineages that evolved independently. For instance, it seems clear that groups such as Gymnodiniales (a group of naked dinoflagellates) are paraphyletic, but so are Peridiniales, which belong to thecate dinoflagellates. Yet thecate dinoflagellates probably constitute a clade (Orr *et al.*, 2012; **Fig.1.5.10**). Considering the ever-increasing number of dinoflagellate sequences, this taxonomy is currently undergoing major changes (Hoppenrath, 2017). Therefore, there is a need to update and renew the classification of dinoflagellates, taking into account both molecular and morphological updates. In spite of the

amassed sequence data, many dinoflagellates and entire lineages remain without sufficient coverage of reference sequences, or without any references at all, to be represented in a multi-gene alignment.

Transcriptomics research involving different dinoflagellates lineages is still in its infancy, but initial studies start to shed light on their peculiar genomics, gene transfer and evolution (Janouškovec *et al.*, 2017).

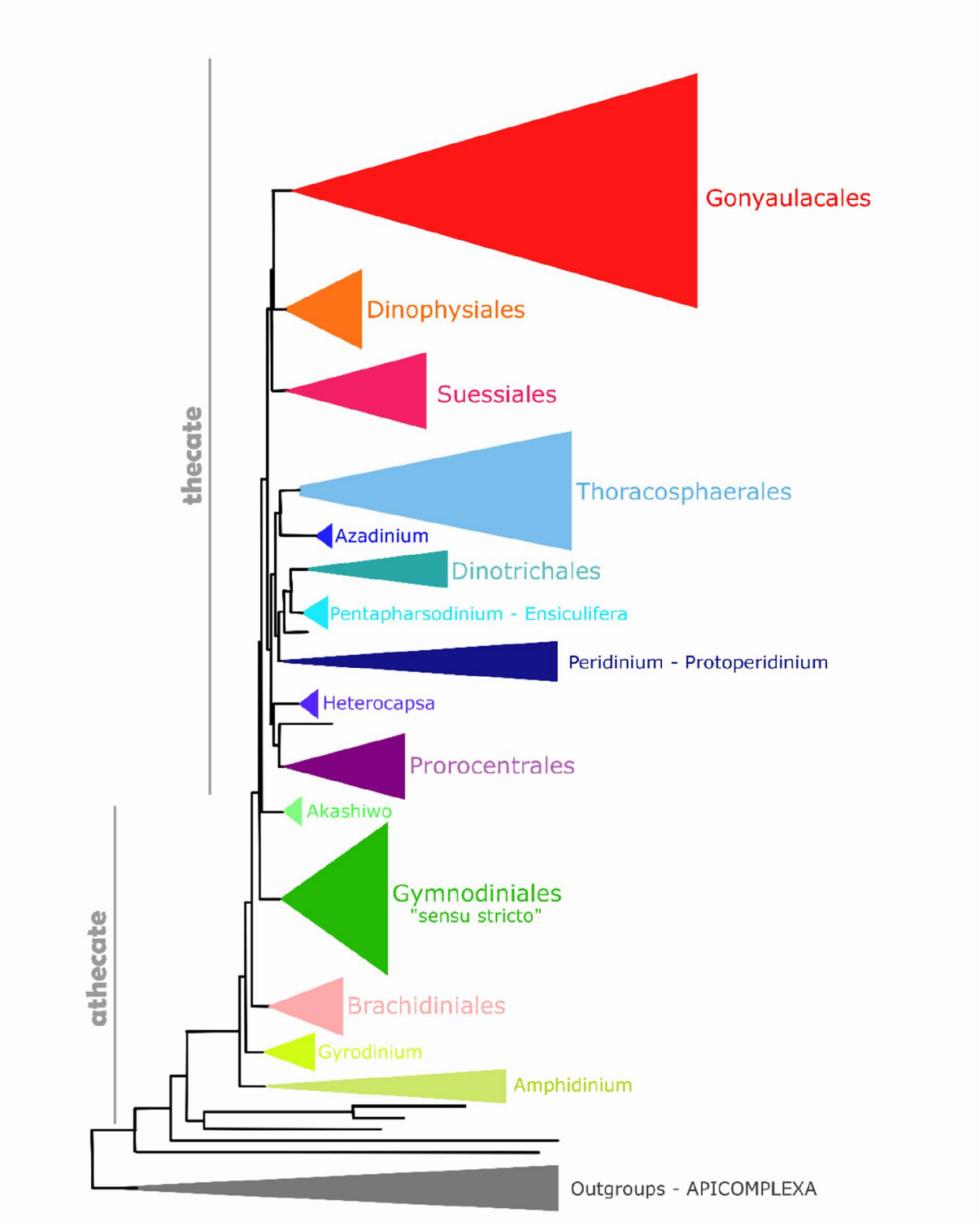


Fig.1.5.10: Multigene phylogeny of dinoflagellate inferred from 18S, 5.8S, 28S, cob, cox1, hsp 90, actin and beta tubulin genes (7138 bp). Figure modified from Orr *et al.* (2012).

1.6. Towards a natural classification

Identification of species is the critical starting point for biodiversity research. Until recently, the description of new species in protists depended on our ability to isolate and cultivate these organisms. In the last two decades, the application of molecular methods to study entire marine ecosystems revolutionized views of microbial community structure and functioning (Caron, 2013; Kim *et al.*, 2014), allowing a complete reassessment of marine microbial diversity in natural environments. Classification satisfies our innate need to distribute organisms into natural groups that share characteristics because of shared evolutionary history. Originally, morphology and pigmentation were the only sources of characteristics available to the classifiers. Since the onset of electron microscopy, ultrastructure was added. More recently, molecular data has added a wealth of characters. Not surprisingly, some conflicts among all these characters abound and the addition of all these new types of data have affected classification schemes (Fensome *et al.*, 1993; Adl *et al.*, 2012; **Fig.1.6.11**).

1.6.a. A brief history of dinoflagellate classification

The first modern dinoflagellate was described in 1753, when the British naturalist Henry Baker depicted what he called "Animalcules which cause the Sparkling Light in Sea Water". Otto Friedrich Müller a Danish naturalist introduced the name "dinoflagellate" in 1773 when arranging the Infusoria group into genera and species for the first time. *Ceratium* is the first genus name still in use (Schrank, 1793). Later in the 1830s, the German microscopist Christian Gottfried Erhenberg made a great contribution to dinoflagellate taxonomy by describing *Peridinium, Prorocentrum* and *Dinophysis* successively. Naked genera *Amphidinium* and *Gymnodinium* were later described by Claparède and Lachmann in 1859 and Stein in 1878, respectively. Dinoflagellate lineages were named in many different ways since their discovery including Cilioflagellata (Claparède and

Lachmann, 1868), Pyrrophyta (Pascher, 1914), Pyrrhophycophyra (Papenfuss, 1946), Arthrodelen Flagellaten (Stein, 1883), Dinophyta (Dillon, 1963) or Dinomastigota (Margulis and Sagan, 1985).

Dinoflagellate nomenclature remains administered under the rules of both the International Code of Botanical Nomenclature (ICBN) and the International Code of Zoological Nomenclature (ICZN).

Over time, many authors have contributed to the publication and updates of dinoflagellate classification. Jakob Schiller gave one of the first one with detailed description of species (1931-1937). Later, Alain Sournia added his contribution by publishing a series of books and checklists, each time updated with the newest taxonomic entities (1973, 1978, 1982, 1990 and 1993). In 1993, Fensome et al. published "A classification of fossil and living dinoflagellates" comparing different classification, illustrating and describing fossil and modern dinoflagellate taxonomy.

More recent reference classification systems of dinoflagellates have been proposed by Gómez (2005; 2012b) and Adl *et al.* (2012) for all protists. Online databases or websites such as www.algaebase.org (Guiry and Guiry, 2017) or www.dinophyta.org (Centre of Excellence for Dinophyte Taxonomy, CEDiT) provide reliable classifications of all algae, being updated constantly.

Nevertheless, no consensus exists between published reference classifications (**Fig.1.6.11**). Different rank names and higher-level classifications can vary greatly. Some names of genera and species are accepted by some authors and not by others. There is currently no harmonisation of the classification of dinoflagellates.

Fensome et al., 1993					Algaebase.com Guiry and guiry, 2017		Adl et al., 2012			
Dinokaryota	Dinophyceae				Dinoflagellata	Dinophyceae	Dinoflagellata	Dinophyceae		
	↳	Gymnodiniphycidea				↳ Actiniscales		↳	Gymnodiniphycidea	
		↳	Gymnodiniales			↳ Amphilothales			↳	Gymnodinium
		↳	Ptychodiscales			↳ Blastodiniales			↳	Amphidinium
		↳	Suessiales			↳ Brachidiniales			↳	Gyrodinium
	↳	Peridiniphycideae				↳ Desmomastigales			↳	Kareniacea
		↳	Gonyaulacales			↳ Dinamoebales			↳	Ptychodiscales
		↳	Peridiniales			↳ Dinophysiales			↳	Borghiellaceae
	↳	Dinophysiphycidae				↳ Dinotrichales			↳	Tovelliaceae
		↳	Nannoceratopsiales			↳ Gonyaulacales			↳	Suessiaceae
		↳	Dinophysiales			↳ Gymnodiniales		↳	Peridiniphycidae	
	↳	Prorocentrophycidae				↳ Haplozoonales			↳	Gonyaulacales
			Prorocentrales			↳ Lophodiniales			↳	Peridiniales
	↳	Subclass uncertain				↳ Peridiniales			↳	Thoracosphaeraceae
		↳	Desmocapsales			↳ Phytodiniales			↳	Podolampadaceae
		↳	Phytodiniales			↳ Prorocentrales		↳	Dinophysiales	
		↳	Thoracosphaerales			↳ Ptychodiscales		↳	Prorocentrales	
	Blastodiniphycea					↳ Pyrocystakes		Noctilucales		
		↳	Blastodiniales			↳ Suessiales	Protoalveolata	Chromerida		
	Noctiluciphyceae					↳ Torodiniales		Colpodellida		
		↳	Noctilucales		Noctilucea			Perkinsidae		
Syndinea	Syndiniophyceae					↳ Noctilucales		Oxyrrhis		
		↳	Syndiniales		Oxyrrhida			Syndiniales		
						↳ Oxyrrhinales	Insertae sedis			Ellobiopsidae
					Syndinea					
						↳ Coccidiniales				
					Ellobiopsea					
						↳ Thalassomycetales				

Fig.1.6.11: Comparison of different reference classification frameworks published for dinoflagellates. Data taken from Fensome *et al.*, 1993; www.algaebase.com, Guiry and Guiry (2017) and Adl *et al.* (2012).

1.6.b. A phylogenetic backbone for classification

To date, molecular data has given us insight into phylogeny and life history of dinoflagellates, but have not provided a precise view of evolution of this group especially due to poorly resolved backbone of the 18S rRNA tree (see above point 1.5. b. Genetic diversity of modern dinoflagellates). There is still not consensus on the definition of higher taxonomic groups which should result from integration of molecular, morphological and ecological data. Multigene

phylogenies (Orr *et al.*, 2012; **Fig.1.5.10**) and transcriptomics (Janouskovec *et al.*, 2017) are contributing to better elucidation of dinoflagellate life history. However, the lack of references covering the entire diversity of dinoflagellates still inhibit robust phylogenies and classification for some groups. A limited number of references, for a limited number of markers (main genes are 18S rRNA and 28S rRNA) and from a limited number of locations in the world are currently available (Gómez, 2014).

In parallel, classification revision is needed for many genera and species still lacking molecular data which could provide information on their taxonomic placement (Gómez, 2014). Recently, the re-investigation of some genera has allowed the solution of polyphyly and paraphyly inside orders, families and genera; and to raise some new genera and species as a consequence (Hoppenrath and Saldarriga, 2008; Hoppenrath, 2017). However, many cases need to be re-investigated.

Traditional taxonomy is undergoing a major transformation with the ever-increasing number of sequence data of dinoflagellates collected and available in online databases. The use of molecular data as a criterion of characterisation and description of species along with morphological and ecological data has helped the understanding of dinoflagellate evolution and modification of traditional classification. Hence, recent knowledge gained about dinoflagellates is still not reflected in current classifications and needs to be integrated into an official re-organisation of higher taxonomic groups.

1.7. Study site: The Gulf of Napoli

1.7.a. Description of the study site LTER-MareChiara

MareChiara is a long-term marine monitoring station (LTER-MC) located two miles offshore in the Gulf of Naples in the Tyrrhenian Sea (40°48.5'N, 14°15'E; on the line of the 75 m isobath; **Fig.1.7.12**). Phytoplankton, zooplankton and physico-chemical data was collected fortnightly between 1984 and 1991 and continue to be collected weekly since 1995. The aim of this series is to analyse the functioning of a coastal pelagic ecosystem (Ribera d'Alcalà *et al.*, 2004). The station is a strategic point to study Mediterranean populations of pelagic phytoplankton since it is located at the boundary between two hydrographic subsystems: coastal eutrophic waters and offshore oligotrophic waters, making the plankton assemblages diverse throughout the year (Ribera d'Alcalà *et al.*, 2004).

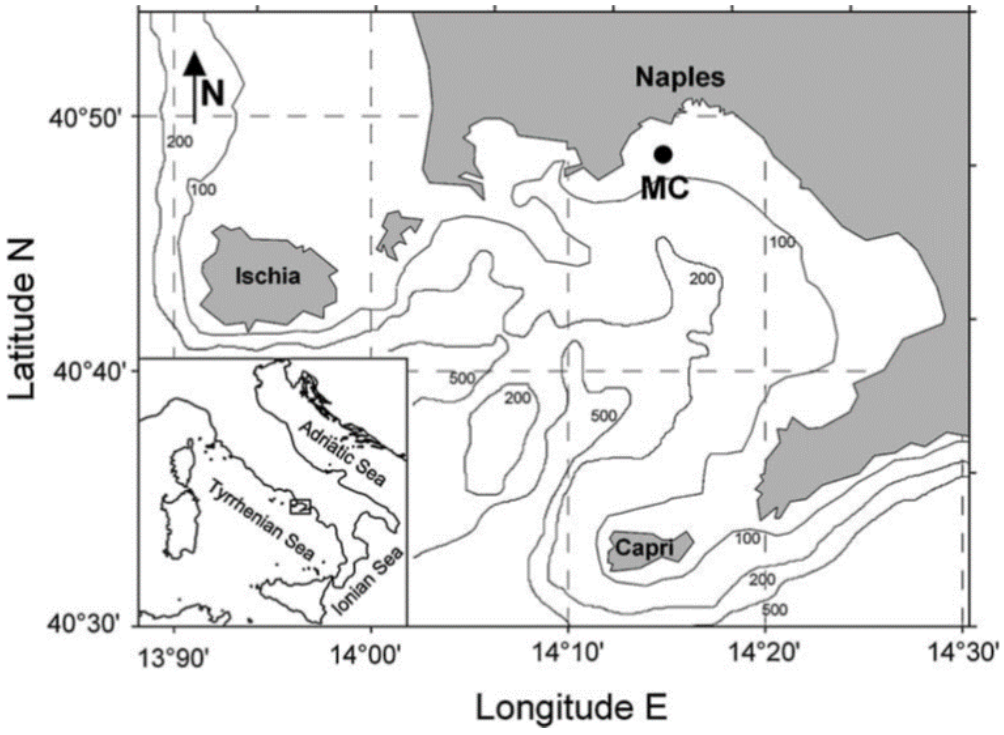


Fig.1.7.12: Map of the Gulf of Naples (Italy) and localisation of the LTER long-term station MareChiara, modified from Zingone *et al.*, (2010).

1.7.b. Diversity and temporal pattern of dinoflagellates at LTER-MC

During the last 40 years about 750 taxa of protists including around 325 dinoflagellates, have been identified morphologically based on observations of net- and Niskin bottle samples (courtesy Diana Sarno). However, 28% of these morphotypes have been observed only once and 66.4% less than 10 times since 1999. The ten most frequent dinoflagellate species recorded at the LTER-MC are thecate dinoflagellate with a highly distinctive morphology, which rarely reach abundances higher than 10-20 cells/ml, i.e., *Tripos furca, Tripos fusus, Protoperidinium diabolus, Prorocentrum triestinum, Prorocentrum compressum, Dinophysis sacculus, Podolampas palmipes, Tripos declinatus, Dinophysis caudata* and *Oxytoxum scolopax*. Naked species are less represented and constitute only 12.9% of dinoflagellate taxa recorded at the LTER-MC. This is mainly due to the difficulties of identification of these fragile organisms which are often classified at the genus or family level, or, more commonly as undetermined.

At LTER-MC, the highest abundances of dinoflagellates, mainly constituted by undetermined naked cells less than 15 µm, usually occur in summer, but different species of dinoflagellates successively appear and overlap throughout the year. Compared to small diatoms and flagellates which frequently bloom at LTER-MC, the abundance of dinoflagellates can be considered low. However, they often contribute significantly to the biomass due to their larger size. Indeed, dinoflagellates can be the major group in summer or even in winter when total phytoplankton abundance is lower (blooms of *Gyrodinium* species in June and many species of *Tripos* in December-January).

Taxonomic knowledge of the dinoflagellates in general is less extensive than that of other protists, such as diatoms, and needs to be explored and referenced with different tools and new methodologies. Due to their lower abundance and the difficulties of identification at species level, there is still no comprehensive description of temporal patterns of dinoflagellate species at the LTER-MC station. Most papers published stating microbial diversity of the Gulf of Naples consider

phytoplankton communities and their global dynamics through the year (Ribera d'Alcalà *et al.*, 2004).

1.7.c. NGS barcoding results

Recent analysis of environmental DNA at LTER-MC, including clone libraries (McDonald *et al.*, 2007; Ruggiero *et al.*, 2015) and metabarcoding data analyses (ERA Biodiversa-BIOMARKS and FIRB-Biodiversitalia projects) have demonstrated a vast protistan diversity (Dunthorn *et al.*, 2014; Logares *et al.*, 2014; Massana *et al.*, 2015; Piredda *et al.*, 2017). However, as with other protists, the majority of the dinoflagellate metabarcodes obtained could not be assigned to any reference with high confidence because they are from entities still unknown to science or lacking reference sequences. Only a minority of the protistan species have been yet formally described (Duff, Ball and Lavrentyev, 2008).

In all High Throughput Sequencing (HTS) metabarcoding studies performed in various marine sites all over the globe, including the LTER-MC, dinoflagellates show a great diversity and represent the majority of the reads obtained. According to Piredda *et al.* (2017) the distribution of dinoflagellate metabarcodes at the LTER-MC indicates clear density shifts and changes in species composition over the seasons, (Ribera d'Alcalà *et al.*, 2004; Cerino *et al.*, 2005). Piredda *et al.* (2017) found that diatoms were dominant throughout the year, except for a few dates in late spring, during the summer and in winter when dinoflagellates and flagellates supercede diatoms in terms of biomass and cell numbers, respectively. In this study, most of the metabarcode reads recovered for dinoflagellates have been assigned to naked dinoflagellate references, but the majority of these had a weak similarity to reference sequences, making their classification at the genus or family-level impossible. When compared with phytoplankton cell abundances estimated from analyses of samples using light microscopy (LM), the number of reads does not match the patterns and protist group composition. For instance, in June 2011, one ribotype peaked (*Gyrodinium* cf. *spirale*) representing 47% of the reads while corresponding to only about 14% of the biomass in LM-based

counts. This difference and the high diversity retrieved for dinoflagellates could be explained by the fact that dinoflagellates possess a high quantity of rDNA copy number due to their to large genome and probably display intragenomic variation (Hou and Lin, 2009).

1.8. Outline of the thesis

In this thesis, I used the HTS data available for 48 dates (3 years, 2011, 2012 and 2013) for the 18S rRNA V4 barcode to assess the global patterns of diversity for dinoflagellates. This chapter has provided an introduction to dinoflagellates, which is built upon and contextualised in the following three chapters. The first main chapter (**Chapter II**) is devoted to DinoREF, a new updated and curated dinoflagellate reference database for the 18S rRNA gene, that I created for this study. Given the increased use of environmental metabarcoding to assess diversity of planktonic organisms, better references are needed to accurately analyse data. This database was generated because the other sources for dinoflagellates were insufficiently curated and updated, often containing assignation mistakes. To address these issues, DinoREF uses a five-step process by: **gathering** all dinoflagellate 18S rRNA sequences present in public databases; **filtering** these sequences based on quality criteria; **validating** the names and verifying their taxonomic assignation; **annotating** the remaining sequences with additional notes and metadata; and then **organising** all the data into a comprehensible framework. Additionally, I used DinoREF to analyse resolution power of the V4 region of the 18S rRNA gene, i.e. the barcode used to produce the HTS dataset at LTER-MC station. This chapter has been submitted as an individual article so as to make this resource as accessible as possible (Mordret *et al.*, 2018). This chapter provides the groundwork for the analysis of the HTS data in Chapters III and IV.

The metabarcode dataset for MareChiara station for dinoflagellates is analysed in **Chapter III** using DinoREF as a reference to assign and classify recovered barcode sequences. I chose to follow a double-pronged approach, analysing the data ataxomically and taxomically in order to observe dinoflagellates as a community, whilst also investigating dinoflagellate patterns at a finer level (such as order, genus and species). Given that MareChiara is a long-term monitoring station, it is an ideal site to study seasonal variation of plankton communities. Here I tried to see if I could detect

any seasonal patterns for dinoflagellate taxa and to test the resolution of the v4 to analyse an environmental metabarcode dataset.

When the reference sequences from DinoREF were cross-referenced with specialised literature, a large number of morphologically described taxa were not characterised molecularly. A good example of this is the genus *Tripos*, for which many species are described morphologically, but only a small number are described molecularly, mainly because they do not grow well in culture. *Tripos* is a very well-known dinoflagellate genus that is present in planktonic samples worldwide, but whose genetic diversity is insufficiently characterised. For this reason, **Chapter IV** is devoted to the *Tripos* genus. Here I used a single cell approach: i.e. I imaged a single cell by LM, extracted, amplified and sequenced the 18S and 28S rRNA for each cell. The 18S and 28S rRNA were used to build *Tripos* phylogenies while the V4 region of the 18S rRNA was extracted in order to assess the diversity and the variation of different *Tripos* species in the LTER-MC HTS dataset.

The final chapter (**Chapter V**) presents my conclusions and outlooks of the thesis.

CHAPTER II: DinoREF: a curated dinoflagellate (Dinophyceae) reference database for 18S rRNA gene

(Mordret et al, 2018, Molecular Ecology Resources)

In Chapter II, Solenn Mordret downloaded sequences from NCBI and designed the framework for molecular and computational analyses under the supervision of Roberta Piredda. S.M. reviewed the taxonomic literature and assembled the database.

2.1. Introduction

Dinoflagellates are an important group of unicellular eukaryotes. Most dinoflagellates are marine and associated with plankton communities whereas some are benthic or live in aquatic terrestrial ecosystems. Dinoflagellates display a remarkable diversity of size and shape as a result of their adaptation to extremely diverse habitats and the adoption of various trophic strategies. Cells are characterised by the presence of flattened vesicles, called alveoli, packed in a continuous layer located beneath the plasma membrane, and the majority possess two asymmetric flagella providing motility capacity to the cell.

Since the discovery of dinoflagellates, their traditional classification has been based on the observed morphology. Nine major lineages have been distinguished following these observations: Gonyaulacales, Dinophysiales, Suessiales, Peridiniales, Prorocentrales, Gymnodiniales, Blastodiniales, Phytodiniales and Noctilucales (Not *et al.,* 2012; Gómez, 2012a, Adl *et al.*, 2012). These orders are still considered valid (Orr *et al.,* 2012; Le Bescot *et al.,* 2016; Guiry and Guiry, 2017). Many dinoflagellates possess a visible, rigid cell wall inside the alveoli, called a theca, which is composed of plates (Gonyaulacales, Dinophysiales, Peridiniales, and Prorocentrales). The organisation and shape of these plates provide taxonomical characteristics to differentiate taxa. Other groups lack such a rigid theca and are called naked dinoflagellates (Gymnodiniales, Apstein 1909.) Several species are large and rather easy to identify in light microscopy (LM) as they possess conspicuous morphological features, whereas others are minute (e.g., Suessiales) and difficult, if not impossible, to identify in LM. A number of species are parasites or symbionts of various hosts and, for those reasons, have lost most or all of their morphological characteristics. Many species are difficult to grow under laboratory conditions (heterotrophic and mixotrophic species mainly).

There is a clear need for morphology-independent methodologies to identify these organisms down to the species level reliably, rapidly and cost-effectively. Sequence barcoding can constitute an alternative to more classical approaches (Pawlowski *et al.*, 2012). Since the emergence of

molecular techniques, the two markers historically used to characterise dinoflagellate species are the 18S rRNA-encoding region and the first 700 bp of the 28S rRNA-encoding region of the nuclear ribosomal cistrons (Murray *et al.,* 2005; Pawlowski *et al.,* 2012; Gómez, 2014). Both regions can be PCR-amplified and sequenced easily because of the high number of copies present in dinoflagellate genomes and the availability of universal eukaryotic primers. However, the 18S rRNA gene seems to be the best choice for a number of reasons.

Of these two regions, I have chosen to establish a Dinoflagellate reference of 18S rRNA sequences for the following four reasons:

1. First, the 18S rRNA gene provides the widest coverage of dinoflagellate diversity. 18S has been sequenced for a higher number of dinoflagellate species than 28S and it shows the best coverage across the phylum (see Gómez, 2014 or Le Bescot *et al.* 2016). This coverage also reaches into groups not amenable to cell culture, such as many heterotrophic species, because the 18S has been applied in single-cell PCR amplification. In the case of single cell amplification for heterotrophic dinoflagellates, the 18S rRNA is the major choice made by the authors (e.g. Ruiz Sebastián and O'Ryan, 2001; Ki, Jang and Han, 2005; Gómez, Moreira and López-García, 2010a; Hoppenrath, Chomérat and Leander, 2013). Other markers such as ITS, COI or HsP 90 (Litaker *et al.,* 2007; Pochon *et al.,* 2012, Hoppenrath and Leander, 2010; Stern *et al.,* 2012) have been proposed, but the numbers of sequences available for each of these markers for dinoflagellates is far smaller and biased toward cultivable dinoflagellates (Gómez, 2014).
2. Despite its rather low variability (Murray et al., 2005), the 18S rRNA genes has a good discrimating power down to the species level (Ki, 2012). Intraspecific variation has also been observed among geographic isolates of what are considered single species but was usually quite low (99% similarity and above; Ki, 2012). (Gómez, 2014; Ki, 2012).

3. The 18S rRNA gene is conserved enough to allow meaningful alignment and phylogenetic analysis, which permits grouping of sequences into meaningful taxa at different levels. A large number of species have been described based on an 18S phylogeny (Gómez, 2014). 18S allows a good clustering of lower levels of classification and offers good resolution mainly at species or genus levels (Ki, 2012; Hoppenrath, 2017). Phylogenetic resolution is not ideal however, and trees built from 18S tend to recover numerous clades showing well resolved internal ramifications, but all emerging from a large polytomy which could be due to rapid basal diversification (Saldarriaga *et al.,* 2001; Murray *et al.,* 2005).
4. Most recent high throughput sequencing (HTS) metabarcoding studies of planktonic protistean diversity uses two regions of the 18S, the variable V4- or V9-regions, rather than the D1-D3 region in the 28S. The two 18S regions can be sequenced entirely using current HTS technology whereas the latter is currently still too long for that. In most of these studies, dinoflagellates represent the largest group in terms of number of sequences detected (Pawlowski *et al.,* 2012; de Vargas *et al.,* 2015; Massana *et al.,* 2015; Piredda *et al.,* in 2017). Here I chose to devote the database to the V4 region because it one of the longest (around 380 bp) and most variable region of the 18S for dinoflagellates and supposedly shows the same phylogenetic resolution as the 18S (Ki, 2012). On the other hand, the V9 has been reported to be inadequate to discriminate between species and even genera (Lie *et al.*, 2013).

Analysis and interpretation of metabarcoding datasets requires very carefully annotated reference databases (Guillou *et al.*, 2013; Decelle *et al.*, 2015). Sequences submitted to Genbank are not checked for quality and not validated for identification. Metadata, i.e., information linked to each sequence, such as location, habitat and methods used are often lacking. Moreover, some groups of dinoflagellates have undergone taxonomic revision, but GenBank entries have not been updated to reflect these changes.

In the same way, recent studies characterising or revising dinoflagellate taxonomy with phylogenetic support have not been translated yet into changing and reforming higher taxonomical levels such as order level (Hoppenrath and Saldarriaga, 2008). For this reason and to better reflect these changes, I developed our own operational taxonomical framework for dinoflagellate higher level of classification.

Even if several reference taxonomic checklists can be listed for dinoflagellates (Gómez, 2012a; Centre of Excellence for Dinophyte Taxonomy ***CEDiT***, Adl *et al.*, 2012) and reference 18S sequence databases exist for protists (Guillou *et al.*, 2013; Quast *et al.*, 2013), none of them supply users with curated sequences, ecological metadata and an up to date taxonomic framework.

The aim of the present study is to provide a taxonomically curated database of 18S rRNA data for dinoflagellates (Dinophyceae) complemented with reliable ecological information called DinoREF. I excluded from the curation process sequences related to taxa clustering at the base of the dinoflagellate lineages, i.e. Noctilucales, Syndiniales, Haplozoonales, Duboscquellales, Oxyrrhinales, Psammosa and Thalassomycetales, and concentrated on "core dinoflagellates" (=Dinokaryots) (Graham *et al.,* 2016). This choice was made because the curation process is particularly problematic for early branching dinoflagellates for two main reasons: poor morphological characterisation of organisms that are principally known from environmental sequences (Okamoto, Horák and Keeling, 2012), and difficulty of aligning sequences, which produce long branches in phylogenies.

In order to build DinoREF, I gathered all 18S rRNA GenBank sequences of dinoflagellates, selected each sequence based on molecular quality and verified their taxonomical assignation with phylogenetic positioning. I chose to include in the database only sequences containing the V4 region of the 18S rRNA because this region is used for many protist metabarcoding studies. Using the database, I summarised information for the number of sequences of each dinoflagellate group and how many sequences were available compared to the number of described species. Therefore,

DinoREF also contains a review of how well the known diversity is covered in terms of 18S rRNA molecular data. This collected molecular information has been used to evaluate the power of the V4 marker to assess dinoflagellate diversity considering described morphospecies and recent multigene phylogenies. This resource will be very useful in the analysis and interpretation of environmental V4 datasets and for annotation of clone library sequences available from public databases.

2.2. Material and Methods

*Curation of sequences includes a **gathering** step in which all possible sequences belonging to the group are included; a **filtering** step in which all sequences that do not fit a set of predetermined quality criteria are removed; a **validation** step in which all remaining sequences are checked that the sequences are identified correctly (the name makes sense from a taxonomic and phylogenetic viewpoint within the context of neighbour sequences); an **annotation** step in which the names are taxonomically updated and the same kind of metadata is associated to it; and finally an Organisation step, in which the dataset is **organised** in such a way that metadata is formatted the same way for all, that the taxonomic levels of organisation are harmonised across the dataset and that the dataset is searchable. These steps are explained below in detail.*

Sequence retrieval

Dinoflagellate 18S rRNA entries available on the 29th August 2016 were downloaded from NCBI GenBank (https://www.ncbi.nlm.nih.gov/) using the following text query: Dinophyceae [Organism] AND (small subunit ribosomal[titl] OR 18S[titl] OR SSU[titl]). The taxonomic information and metadata associated with these sequences were also extracted. Sequences of early branching dinoflagellates were recovered genus by genus when not classified as Dinophyceae in GenBank taxonomy (see below).

Sequence verification

Sequences were inspected to remove those not meeting a list of defined criteria of quality (**Fig.2.2.1** – Phase 1). First, sequences not classified down to the genus level and environmental sequences were excluded as well as entries that did not correspond to 18S rRNA, like plastid 16S rRNA or protein coding genes (**Fig.2.2.1** – Step 1). Sequences belonging to "early-branching" dinoflagellates were set apart from "core" dinoflagellates ("dinokaryon") and did not go through curation process. Regions outside the 18S rRNA gene (ITS, 5.8S and 28S mainly) were removed and sequences without the V4 region were eliminated by aligning all sequences in Seaview (Gouy,

Guindon and Gascuel, 2010) using MUSCLE (Edgar, 2004). In a second step (**Fig.2.2.1** – Step 2) only sequences fulfilling the following criteria were retained: i. sequence length >450 bp; ii. <50 ambiguous nucleotides; iii. <100 bp consecutive deletions; iv. <20 consecutive bp insertions; and v. <20 consecutive ambiguous nucleotides. Sequences poorly aligned or not aligning with any others were removed if BLAST results suggested placement outside dinoflagellates (i.e. BLAST assignation < 90% similarity). This set of sequences constituted the "Quality checked database" (**Fig.2.2.1 – Phase 1**).

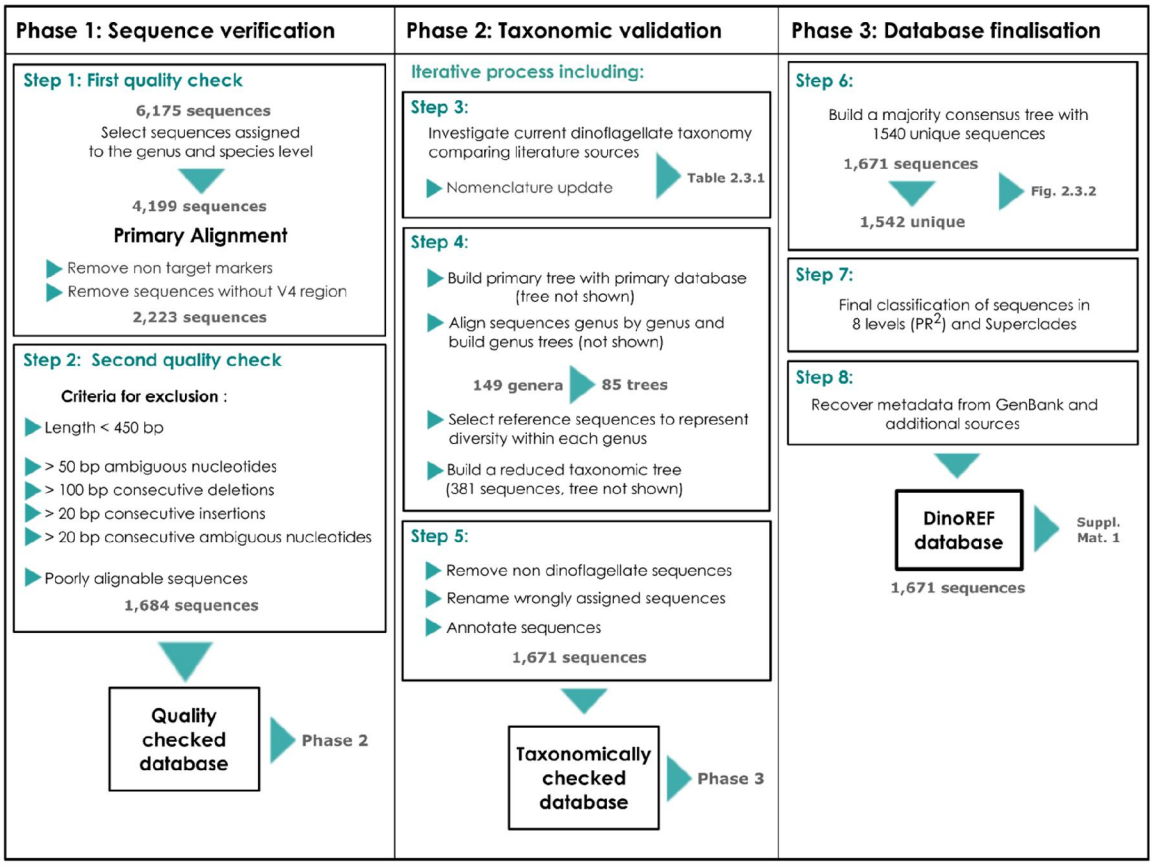


Fig.2.2.1 Workflow describing the steps needed to generate the curated and annotated dinoflagellate 18S rRNA Reference Database from the sequences downloaded from NCBI.

Taxonomy validation

The taxonomic validation was an iterative process including:

i) control on the validity of the nomenclature, based on Fensome *et al.*, (1993), Gómez, (2012b), AlgaeBase (Guiry and Guiry, 2017; http://www.algaebase.org), CEDIT (http://www.dinophyta.org/) and information from the literature of specific groups (**Fig.2.2.1 -** Step 3);

ii) phylogenetic evidences based on the primary tree, the genus-level trees, and the taxonomic reference tree (**Fig.2.2.1 -** Step 4).

Species names were validated following taxonomically accepted names in AlgaeBase (Guiry and Guiry, 2017); names marked as "C"). In order to categorise all sequences of the database in phylogenetically coherent groups, I developed our own classification scheme, which can be used in parallel or as alternative with the classical taxonomic system described above (**Fig.2.2.1** – Step 3). Sequences were classified and grouped in Superclades (described in **Table 2.3.1**) which were supported by molecular and morphological data provided in previous studies (references listed for each Superclade in **Table 2.3.1**). Other dinoflagellates with an uncertain classification were not assigned to any Superclade, but instead, artificially grouped in two categories: "Uncertain naked dinophyceae (UND)" and "Uncertain thecate dinophyceae (UTD)".

Sequences having passed the steps above were aligned with MAFFT v7 (Katoh and Standley, 2013) and a phylogenetic tree was built using FastTree (Price, Dehal and Arkin, 2010) as implemented in the Geneious software (Kearse *et al.*, 2012). This primary tree provided information on the number, statistical support and position of the different terminal clades (**Fig.2.2.1** –Step 4).

To detect sequences assigned to the wrong genus in GenBank, I performed specific alignments for sequences labelled with the same genus name (**Fig.2.2.1** – Step 4). Sequences were split automatically into different fasta files according to their genus name (GenBank assignation) using command "split.groups" in Mothur (Schloss *et al.*, 2009). Then, sequences were aligned genus by

genus using MAFFT and visualised in Seaview (Gouy *et al.*, 2010). When possible (3 or more sequences for a genus), a maximum likelihood tree using PhyML v3.0 (100 bootstraps) was built. Single sequences were grouped with their closest phylogenetic group after verifying their positioning in a global phylogenetic tree.

Based on these trees, I selected a number of sequences representing the diversity within each genus and generated a taxonomic reference dataset containing 381 sequences with length of at least 1700 bp. (**Fig.2.2.1** – Step 4). If possible, sequences with length ≥1,700 bp were selected. Sequences representing the only reference for a given genus and sequences placed at the basal position in their genus clade were included in the dataset even if they were shorter than 1,700 bp. The dataset was aligned using MAFFT (Katoh and Standley, 2013) and the tree built with RAxML (Stamatakis, 2014) (raxmlHPC -f a -m GTRGAMMA -o outgroups -p 12345 -x 12345 -# 100 -s input.phy). Branch support was established using 100 bootstrap replicates.

The comparison between all trees (**Fig.2.2.1** – Step 4) enabled checking of the phylogenetic relationships among all terminal clades, besides allowing the removal of any remaining non-dinoflagellate sequences (**Fig.2.2.1** – Step 5). After these steps, sequences were annotated, renamed or taxonomically updated based on the trees produced and the literature collected. This set of sequences was called the "Taxonomically checked database" (**Fig.2.2.1** – Phase 2).

Database Assembly

From the "Taxonomically checked database" unique sequences were extracted (unique.seqs command in Mothur; Schloss *et al.*, 2009), aligned in MAFFT (Katoh and Standley, 2013) and analysed phylogenetically (RAxML, Majority Rule Consensus tree, 100 bootstraps). In the obtained Majority Rule Consensus tree of 1000 equally most likely trees, nodes with weak support (boostrap >50) were collapsed (**Fig.2.2.1** – Step 6; **Fig.2.3.2**).

The Majority Rule Consensus tree (**Fig.2.2.1** – Step 6; **Fig.2.3.2**) representing all reference sequences was annotated and organised with different colours indicating different Superclades.

All Superclades and clades are presented in **Table 2.3.1** alongside current classical taxonomy. This annotated consensus tree (**Fig.2.3.2**) can be visualised online on iTOL v3 - Interactive Tree of Life (https://itol.embl.de/tree/1932052318357911479398328 ; Letunic and Bork, 2016).

DinoREF database is displayed in a form of an Excel file (**Appendix 1, Supplementary Material 1**) containing information about the traditional taxonomy (Guiry and Guiry, 2017) organised in eight levels (KINGDOM, SUPERGROUP, DIVISION, CLASS, ORDER, FAMILY, GENUS and SPECIES), on the same model as PR² (Guillou *et al.*, 2013)(**Fig.2.2.1** – Step 7). In the case of missing information for a level, the genus name was used together with the name of the missing rank (e.g., *Akashiwo*_order and *Akashiwo*_family). Superclade groups (**Table 2.3.1**) and specific annotations based on the results of this study (phylogenetic position of the sequence or from literature) were added to the DinoREF database for each sequence (**Fig.2.2.1** – Step 8, **Appendix 1, Supplementary Material 1**).

The standard metadata (e.g., strain name, location, see the file for a complete list) extracted from GenBank for each sequence has been supplemented by information on: i) species habitat (Guiry and Guiry, 2017), ii) symbiotic or parasitic lifestyle derived from Genbank or from original papers, iii) potential toxicity obtained from the IOC-UNESCO Taxonomic Reference List of Harmful Micro Algae (Moestrup *et al.*, 2009 - http://www.marinespecies.org/hab/dinoflag.php) accessed on the 13th March 2017), iv) benthic lifestyle according to Hoppenrath *et al.*, (2014) and v) material from which sequence was obtained (culture or single cell), extracted from the original publication (**Fig.2.2.1** – Step 8, **Supplementary Material 1**).

V4 analysis

Reference sequences were used to assess the variation of the V4 region for dinoflagellates. From the final alignment, sequences were cut using V4 primers (primers used by Piredda *et al.*, 2017). Then, using Mothur (Schloss *et al.*, 2009), sequences were de-replicated (unique.seqs) and split by genus and Superclades (split.groups). Each group of sequences was re-aligned using MAFFT

(Katoh and Standley, 2013) and checked manually. Distance matrices calculating the number of pairwise differences between 2 sequences over the length of the V4 (p-distance) were generated through the software MEGA v7 (Kumar, Stecher and Tamura, 2016) for each Superclade and each genus. Two files were created by merging the calculations made by Superclade automatically into a single file and the calculations made by genus in a second file using linux command cat (cat *.* > All_together.csv). The distributions of distances by Superclades and by genus were visualised in the form of boxplots (**Fig.2.3.4** and **Fig.2.3.5**). Graphs for **Fig.2.2.3**, **Fig.2.3.4** and **Fig.2.3.5** were produced via R (R Development Core Team, 2016) using the "ggplot2" library (Wickham, 2009). V4 OTUs at 98% of similarity were generated from alignment. V4 references were clustered together using VSEARCH algorithm with distance-based greedy clustering within Mothur (DGC, Rognes *et al.*, 2016). The OTUs produced were organised by Superclade, the taxonomy of the first sequence in each OTU prevailing over the others. When OTUs collapsed different species, genera or Superclades, lines in the documents were coloured in "blue", "light green" or "purple" respectively. The OTUs at 98% similarity can be found in **Supplementary Material 7**.

DinoREF is available as flat files on Figshare https://figshare.com/s/ebdc8df3cbfaa5690d97 and has been incorporated in Pr2 version 4.7 which is available from https://figshare.com/articles/PR2_rRNA_gene_database/3803709 .

In addition to the DinoREF database (**Appendix 1**, **Supplementary Material 1**), I provided a file containing all DinoREF 18S sequences in fasta format (**Supplementary Material 2**) and three taxonomy files: i) complete GenBank taxonomy (**Supplementary Material 3**), ii) the curated taxonomy (**Supplementary Material 4**), iii) the Superclade classification used in this study (**Supplementary Material 5**). The format of the files is compatible with Mothur (Schloss *et al.*, 2009) and Qiime (Caporaso *et al.*, 2010). I also provide two excel files with V4 sequences and V4 OTUs at 98% similarity (**Supplementary Material 6 and 7**). **Supplementary Material 8** includes a fasta file containing all sequences of early branching dinoflagellates recovered from GenBank but not included in DinoREF.

2.3. Results

The DinoREF database

Phase 1: Sequence verification

A total of 6,175 dinoflagellate sequences were downloaded from GenBank. After selection of the sequences assigned at the genus or species level, 4,199 remained. Of these sequences, 1,976 consisted of other markers than 18S or did not contain the V4 region, and/or aligned poorly (34) or not at all with the 18S. About 60% of the sequences removed had been named «18S rRNA gene partial sequence » in the title, but contained just a very short fragment of 18S gene. After removal of these sequences, 2,223 were kept (**Fig.2.2.1** – Step 1). 539 sequences did not pass the second quality check screening and were discarded from the database (**Fig.2.2.1** – Step 2). Of these discarded sequences, only four sequences contained insertions of ≥20 bp (two in *Phalacroma,* one in *Symbiodinium* and one in *Prorocentrum)*, four *Gambierdiscus* sequences showed an >100 bp deletion, and two sequences exhibited ≥50 ambiguous nucleotides and/or ≥20 consecutive ambiguous nucleotides (as a consequence of poor quality chromatogram). Following removal of these sequences, the Quality checked database contained 1,684 sequences (**Fig.2.2.1** – Step 2).

Phase 2: Taxonomic validation

The primary tree inferred from the aligned Quality checked database revealed that the genus level was the best supported taxonomic level (**Fig.2.2.1** – Step 3 and 4). Indeed, most of the time, the phylogeny grouped together sequences belonging to the same genus. Therefore, the validity of each genus was checked with the literature relevant to that genus. 85 genus trees and a reduced taxonomic tree were built in order to help with curation process (**Fig.2.2.1** – Step 4).

At this stage, 13 of the 1,684 sequences had to be removed because they did not belong to dinoflagellates (**Fig.2.2.1** – Step 5). Then, the names of 440 sequences (26%) had to be curated or updated (**Supplementary Material 1**) because names originally assigned to them were invalid on GenBank or the phylogenetic analyses revealed that they were attributed to the wrong taxon (20 sequences). Annotations based on the phylogenetic position or useful details from literature were

also provided for 300 sequences (18%)(**Supplementary Material 1**). At the end of the second phase (**Fig.2.2.1** – Phase 2), the "Taxonomically checked database" contained 1,671 sequences corresponding to 1,540 unique sequences. Overall, 18S data was available for a total of 149 validated dinoflagellate genera (**Table 2.3.1**) representing 422 species (**Table 2.3.2**).

Three genera included more than 150 sequences each: *Alexandrium* (Superclade **#1**) with 210 sequences, *Gambierdiscus* (**#1**) with 169 sequences and *Symbiodinium* (**#3**) with 173 sequences (**Table 2.3.2**, representing 36% of the sequences in the database). Some species within these three genera were represented by a large number of different 18S sequences. For instance, 54 slightly different sequences were attributed to *Gambierdiscus scabrosus*. A total of 27 genera included between 68 and 10 sequences in the database (40%), 55 genera contained between nine and three sequences (18 %), whereas the remaining 63 were represented by one or two sequences only (6% of the total number of sequences) (**Fig.2.3.5**).

The length of the sequences in the reference database vary from 579 to 1764 bp (**Fig.2.3.3**), with the majority 1200 sequences (72 %) between 1600 and 1764 bp covering almost the full-length sequence of the 18S rRNA gene. 165 (10%) of sequences have a size between 1100 and 1400 bp (**Fig.2.3.3**) and correspond in general to sequences amplified from a single cell.

Phase 3: Database finalisation

The 1,540 unique sequences from the "Taxonomically checked database" were used to build a majority-rule consensus tree (**Fig.2.3.2**) in order to provide a final representation of the database (**Fig.2.2.1** – Step 6). The curated sequences were then organised hierarchically, in the same way as the PR^2 database (8 levels taxonomic classification, **Supplementary Material 1**). For the assignation of the 'order' level, I followed a conservative approach accepting the following six orders: Gonyaulacales, Peridiniales, Dinophysiales, Prorocentrales, Suessiales, and Gymnodiniales. Some sequences could not be placed within any of these orders and were listed as Dinophyceae *ordo incertae sedis*. Classification of the sequences at "family" level was problematic

for many dinoflagellates (see discussion). Algaebase (Guiry and Guiry, 2017) family names were mainly used in this database.

Superclades: an attempt to depict the recent changes in dinoflagellate phylogeny

The Majority-rule Consensus tree (**Fig.2.3.2**) built with the 1,540 unique sequences from the Taxonomically checked database (**Fig.2.2.1** – Step 6) showed a large number of clades with ≥50 bootstrap support onto a polytomy. The grouping of the monophyletic clades (149 genera) included in this study over the Superclades, according to various authors specified in **Table 2.3.1,** did not conflict with the various clades in the tree in **Fig.2.3.2**. However, most of the Superclades were separated in multiple clades (**Table 2.3.1**). The largest of these sub-clades included Gonyaulacales, but additional members of this order were recovered into four smaller clades (1B, 1C, 1D and 1E; **Table 2.3.1**; **Fig.2.3.2**). The two next largest clades, clade 3 and clade 13 included all Suessiales (Superclade 3) and all Gymnodiniales *sensu stricto* (Superclade 13) sequences respectivelly. Superclade 2 included Dinophysiales (clade 2A) except for *Sinophysis* (clade 2B) and *Pseudophalacroma (*clade 2C) genera, each of which grouped into its own clade. Peridiniales *sensu stricto* were recovered in four clades and two single sequences (Superclades 8A, 8B, 8C and 8D). Species of *Protoperidinium* were recovered in three of the four clades (8A, 8C and 8D), each of which also included members of other genera (**Table 2.3.1**). Many dinoflagellate genera cannot be classified in any Superclade category. These genera usually cluster alone, without strong support to any group and lack decisive morphological characters to provide clues for their evolution (**Table 2.3.1, Fig 2.3.2, Supplementary Material 1**). These genera are grouped in UTD: Uncertain Thecate Dinophyceae (23 genera) and UND: Uncertain Naked Dinophyceae (12 genera) (**Table 2.3.1, Fig.2.3.2**).

The number of species represented is highly variable among the different Superclades, clades and genera of dinoflagellates (**Table 2.3.2**). The Gonyaulacales (Superclade 1) includes 92 molecularly characterised species while other Superclades, such as *Akashiwo* (Superclade 12) or Ptychodiscales

(Superclade 20), are represented by only one species (**Table 2.3.2**). The number of genera represented by an 18S sequence varies depending on the lineages (**Table 2.3.2**, **Fig.2.3.4**).

The number of thecate taxa characterised by 18S sequences (Superclade 1 to 11, plus UTD) is higher than that of naked taxa (Superclade 12 to 20, plus UND), i.e. 105 genera and 339 species *vs* 43 genera and 117 species, respectively (**Table 2.3.2**). This pattern reflects the higher number of described thecate taxa: indeed the 2,342 taxonomically described species include 1,626 thecate species, belonging to 163 genera, and 716 naked species, belonging to 69 genera (**Table 2.3.2**).

Overall, only 22% of the dinoflagellates described in literature have a reference 18S sequence such that a large number of species described morphologically are still to be characterised from the molecular point of view. These species include 1639 species belonging to genera supported by molecular data and 247 species belonging to 84 genera that still lack molecular characterisation (**Table 2.3.2**; Additional genera and Additional species).

A total of 1,485 (89%) sequences originate from marine environments, 137 (8%) from freshwater habitats and 50 (2%) were recovered from other environments (brackish, estuarine, athalassohaline). Less than 1% of sequences were annotated with an "x" because no information was given by the authors. In addition, 397 sequences (24%) were annotated as benthic and 571 sequences (34 %) as toxic species. Specific ecological information (symbiont, parasite, host) were provided for 302 sequences (18%) (**Supplementary Material 1**).

Table 2.3.1: Table showing the different Superclade and clade categories of dinoflagellates represented in the database. Genera or species present in each category correspond to the sequences forming Superclades in the consensus tree (**Fig.2.3.2**). All Superclades are supported by specialised literature and molecular evidences. After literature review, dinoflagellates characterised molecularly but lacking morphological information, requiring deeper investigation or clustering alone without phylogenetic support to any group have been arbitrary classified in "Uncertain Thecate Dinophyceae" and "Uncertain Naked Dinophyceae". All these dinoflagellates have been also found clustering alone in the consensus phylogenetic tree (**Fig.2.3.2**). Taxa in bold were included in Orr *et al.* (2012) phylogenetic analyses.

ORDER (AlgaeBase)	SUPERCLADE	CLADE	GENERA or SPECIES
GONYAULACALES	❶ GONYAULACALES (Adl *et al.*, 2012; Orr *et al.*, 2012)	1A	***Alexandrium*, *Fragilidinium*, *Coolia*,** *Ostreopsis, Fukuyoa,* ***Gambierdiscus*,** *Goniodoma,* ***Pyrocystis*, *Pyrodinium*,** *Pyrophacus*
		1B	***Ceratium*, *Tripos***
		1C	***Lingulodinium*, *Amylax*,** *Gonyaulax verior*
		1D	***Gonyaulax***
		1E	***Ceratocorys*, *Protoceratium***
DINOPHYSIALES	❷ DINOPHYSIALES (Adl *et al.*, 2012; Orr *et al.*, 2012; Hoppenrath, Chomérat and Leander, 2013)	2A	***Amphisolenia*, *Dinophysis*,** *Histioneis* ***Ornithocercus*, *Phalacroma*,** *Triposolenia*
		2B	*Sinophysis*
		2C	*Pseudophalacroma*
SUESSIALES	❸ SUESSIALES (Adl *et al.*, 2012; Orr *et al.*, 2012)	3	*Ansanella, Asulcocephalium,* ***Baldinia*, *Biecheleria*,** *Biecheleriopsis, Borghiella, Cystodinium, Leiocephalium,* ***Pelagodinium*,** *Phytodinium, Piscinoodinium,* ***Polarella*, *Protodinium*, *Symbiodinium***
PERIDINIALES	❹ THORACOSPHAERACEAE (Adl *et al.*, 2012; Gottschling *et al.*, 2012)	4A	*Amyloodinium,* ***Cryptoperidiniopsis*,** *Paulsenella,* ***Pfiesteria***
		4B	<u>*Scrippsiella sensu lato*</u>: *Pernambugia,* ***Dubosquodinium*,** *Naiadinium,* ***Scrippsiella*,** *Theleodinium*
		4C	*Apocalathium*
		4D	***Crypthecodinium***
		4E	***Stoeckeria***
		4F	*Chimonodinium*
		4G	***Thoracosphaera***
			<u>Single sequences</u>: *Aduncodinium glandula,* ***Tintinnophagus acutus***

ORDER (AlgaeBase)	SUPERCLADE	CLADE	GENERA or SPECIES
Dinophyceae *ordo incertae sedis*	❺ AMPHIDOMATACEAE (Tillmann *et al.*, 2014)	5A	***Azadinium***
		5B	*Amphidoma*
		5C	*Azadinium dexteroporum*
		5D	*Azadinium polongum, Azadinium concinnum*
		5E	*Azadinium caudatum*
PERIDINIALES	❻ KRYPTOPERIDINIACEAE (Takano *et al.*, 2008; Gottschling *et al.*, 2017)	6	***Durinskia,*** *Galeidinium,* ***Kryptoperidinium,*** *Unruhdinium, Blixaea*
	❼ genera ENSICULIFERA and PENTAPHARSODINIUM	7	***Ensiculifera, Pentapharsodinium***
	❽ PERIDINIALES *sensu stricto* (Gu, Liu and Mertens, 2015; Mertens *et al.*, 2015)	8A	Clade Monovela: *Amphidiniopsis, Archaeperidinium, Herdmania,* ***Islandinium,*** *Protoperidinium americanum, P. fusiforme, P. fukuyoi, P. monovelum, P. parthenopes*
		8B	*Peridinium* clade: ***Peridinium willei,*** *P. volzii,* ***P. cinctum,*** *P. gatunense, P. bipes, P. limbatum*
		8C	*Protoperidinium* sensu stricto: *Protoperidinium abei, P. bipes, P. conicum, P. crassipes, P. divergens, P. denticulatum, P . elegans, P. excentricum, P. leonis, P. pallidum, P. pellucidum, P. pentagonum, P. punctulatum, P. thorianum, P. thulesense, Kolkwitziella*
		8D	Diplopsalioideae III and Oceanica clade: *Diplopsalopsis, Niea, Qia, Gotoius,* ***Protoperidinium claudicans,*** *Protoperidinium depressum*
			Single sequences: *Diplopsalis caspica, D. lenticula, Preperidinium meunieri*
	❾ HETEROCAPSACEAE (Salas, Tillmann, & Kavanagh, 2014)	9	***Heterocapsa***
	❿ PODOLAMPADACEAE (Adl *et al.*, 2012)	10A	*Blepharocysta, Podolampas, Roscoffia*
			Single sequence: *Lessardia elongata*
PROROCENTRALES	⓫ PROROCENTRALES (Adl *et al.*, 2012; Orr *et al.*, 2012)	11A	*Prorocentrum dentatum,* ***P. donghaiense,*** *P. emarginatum, P. fukuyoi, P. mexicanum,* ***P. micans, P. cordatum,*** *P. rhathymum, P. shikokuense, P. texanum, P. triestinum, P. tsawwassenense*

ORDER (AlgaeBase)	SUPERCLADE	CLADE	GENERA or SPECIES
		11B	*Prorocentrum hoffmannianum, P. bimaculatum,* ***P. concavum, P. consutum,*** *P. foraminosum, P. maculosum, P. leve,* ***P. lima***
		11C	*Prorocentrum glenanicum, P. panamense, P. pseudopanamense*
		11D	*Plagiodinium*
		11E	*Prorocentrum cassubicum*
GYMNODINIALES	⓬ genus *AKASHIWO* (Orr *et al.*, 2012)	12	**Akashiwo**
	⓭ GYMNODINIALES *sensu stricto* (Hoppenrath and Leander, 2007b; Reñé, Camp and Garcés, 2015)	13	*Chytriodinium,* ***Dissodinium,*** *Erythropsidinium,* ***Gymnodinium,*** *Gymnoxanthella,* ***Gyrodiniellum, Lepidodinium, Nematodinium,*** *Nusuttodinium, Paragymnodinium, Pellucidodinium,* ***Pheopolykrikos, Polykrikos,*** *Proterythropsis,* ***Spiniferodinium, Warnowia***
	⓮ KARENIACEAE (Adl *et al.*, 2012)	14A	***Brachidinium, Karenia***
		14B	***Karlodinium, Takayama***
	⓯ genus *GYRODINIUM* (Reñé, Camp and Garcés, 2015)	15	***Gyrodinium***
	⓰ genus *AMPHIDINIUM sensu stricto* (Jørgensen, Murray and Daugbjerg, 2004)	16	***Amphidinium***
			Single sequences: ***Amphidinium mootonorum,*** *A. herdmanii, A. longum*
	⓱ TORODINIALES (Boutrup *et al.*, 2016)	17A	*Torodinium*
		17B	*Kapelodinium*
Dinophyceae *ordo incertae sedis*	⓲ TOVELLIACEAE (Lindberg *et al.*, 2005; Adl *et al.*, 2012)	18A	*Esoptrodinium*
		18B	*Jadwigia (*including #JQ639765 *Woloszynskia* sp.*)*
		18C	*Tovellia (*including #AY443025*Woloszynskia leopoliensis)*
PERIDINIALES	⓳ genus *BLASTODINIUM* (Skovgaard and Salomonsen, 2009)	19A	*Blastodinium navicula, B. mangini, B. galatheanum*
		19B	*Blastodinium spinulosum, B. crassum, B. pruvoti, B. inornatum*
		19C	*Blastodinium contortum*
			Single sequence: *Blastodinium oviforme*

ORDER (AlgaeBase)	**SUPERCLADE**	**CLADE**	**GENERA or SPECIES**
Dinophyceae *ordo incertae sedis*	⓴ PTYCHODISCALES (Adl *et al.*, 2012)	20	Single sequence: *Ptychodiscus noctiluca*
	UTD: «Uncertain Thecate Dinoflagellates»	**UTD**	***Adenoides***, *Ailadinium, Amphidiniella, Bysmatrum, Glenoaulax, Glenodiniopsis,* ***Gloeodinium***, *Hemidinium, Heterodinium, Madanidinium, Oodinium, Palatinus, Parvodinium, Peridinium sociale, Peridiniopsis borgei, Pileidinium, Pseudadenoides,* ***Rufusiella***, *Sabulodinium, Stylodinium,* ***Thecadinium***, *Zooxanthella.*
	UND: «Uncertain Naked Dinoflagellates»	**UND**	*Ankistrodinium, Apicoporus, Balechina, Bispinodinium, Ceratoperidinium,, Cucumeridinium, Levanderina, Margalefidinium, Moestrupia, Testudodinium, Togula*

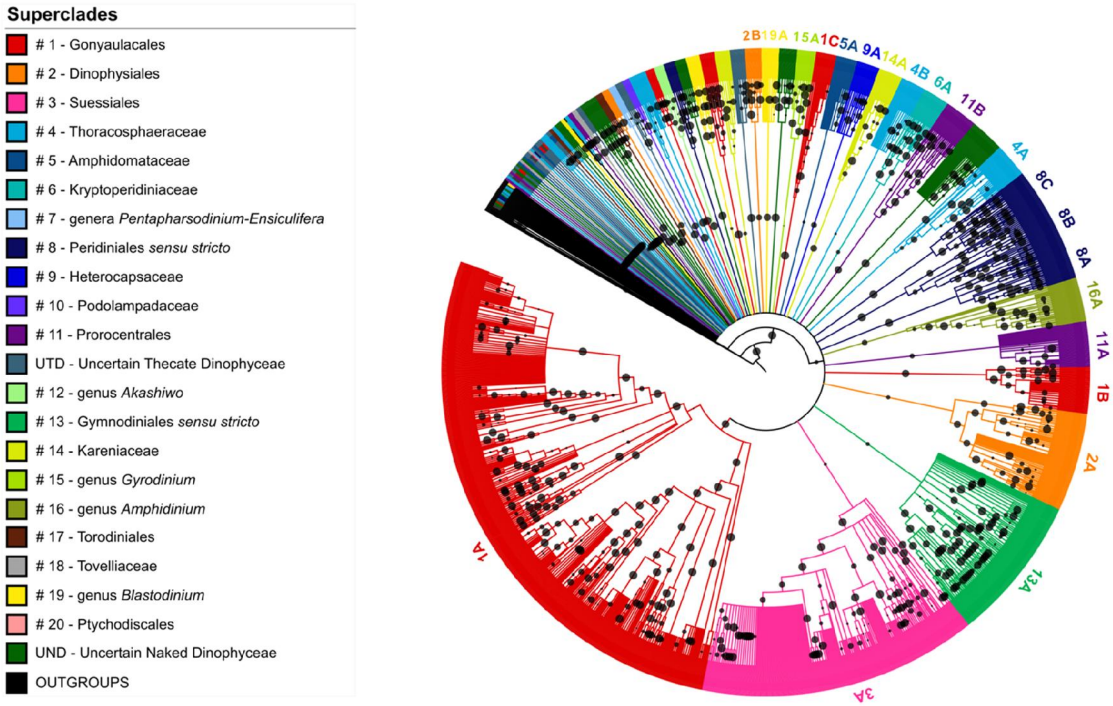


Fig.2.3.2: Consensus phylogenetic tree (RAxML, GTR model) based on 1,540 unique 18S rRNA sequences in the dinoflagellate reference database. Alignment of 2153 bp with three sequences of Ciliates (U97109; X56165 and X03772) and three sequences of Apicomplexa (M97703; AF236097 and AF291427) used as outgroup (OUTGROUPS). Clades are ordered according to their size, they have obtained ≥50 bootstrap support. Bootstrap values are represented by black dots, their size being proportional to their bootstrap value. The colours of the Superclades correspond to those in **Table 2.3.1**. Clades within each Superclade have been marked A, B, C, etc., along the outer rim of the tree, corresponding with their assignment in **Table 2.3.1**. The "Superclades" Uncertain Naked Dinophyceae and Uncertain Thecate Dinophyceae have not been marked and neither have the minute Superclades and sub-clades to the upper left of the tree. The tree can be visualised on i-Tol (https://itol.embl.de/tree/193205231835791147939832 8) in which all Superclades are marked.

The barcode V4 region

Considering only the V4 region, the number of unique sequences from DinoREF shrunk from 1,540 to 946 (**Supplementary Material 6**). The decrease was mainly the result of the collapse of intraspecific diversity in the 18S rRNA gene. However, 11% of the species (48 of 422) represented in the database collapsed together and could not be differentiated with the V4 region. For example, the V4 did not allow the differentiation of some species or all species within potentially toxic genera such as *Dinophysis*, *Karenia*, *Karlodinium* or *Azadinium* (**Table 2.3.3**). In addition, the V4 region did not allow the unambiguous identification of 13 genera represented in DinoREF

(**Table 2.3.3**). As an example, *Karlodinium* shares identical V4 with *Takayama* (**V4 #788**); *Histioneis* shares identical V4 with *Ornithocercus* (**V4 #315**), and some species of *Scrippsiella* share the same V4 with *Duboscquodinium* and *Pernambugia (***V4 #510***)* (**Table 2.3.3**). On the other hand, a single species could be represented by several different V4 sequences. *Alexandrium fundyense* and *Gambierdiscus scabrosus*, for instance, display 103 and 54 18S and 36 and 39 different V4 sequences respectively.

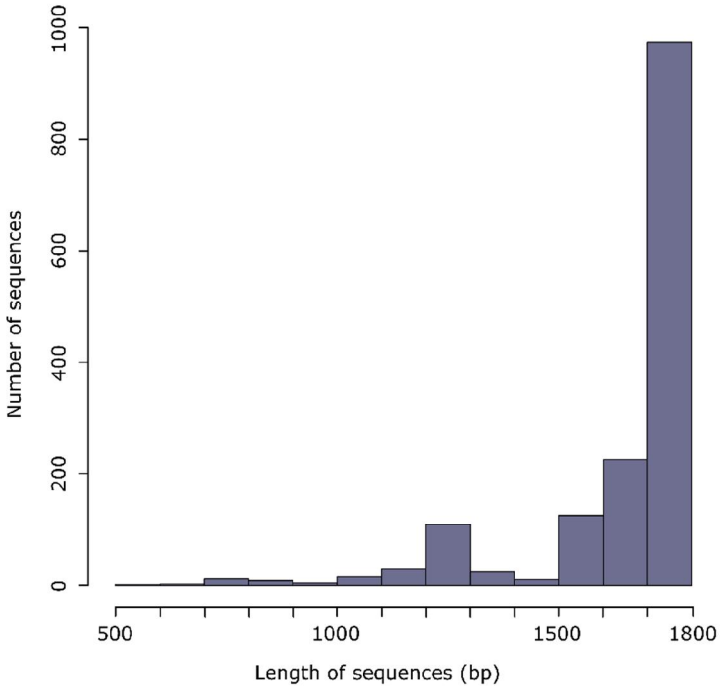


Fig. 2.3.3: Histogram of the length distribution of the 1,540 unique 18S rRNA sequences in the dinoflagellate reference database.

Pairwise p-distance values (number of mismatches divided by the length of the V4 region, approximatively 380 bp) by Superclade and by genus are represented in **Fig.2.3.4** and **Fig.2.3.5**, respectively. Globally, Superclades that include more sequences (and genera) revealed highest p-distance values (**Fig.2.3.4**). However, different patterns were observed. For example, Superclade **#1** (Gonyaulacales) and Superclade **#8** (Peridiniales *sensu stricto*), which showed similar p-distance values were represented by a different number of sequences, genera and species, i.e. 543

sequences for 17 genera and 92 species in Superclade **#1** and 116 sequences, 13 genera and 47 species in Superclade **#8** (**Table 2.3.2, Fig.2.3.4**). Other Superclades, such as Superclade **#16** (*Amphidinium*) or Superclade **#18** (Tovelliaceae) and Superclade **#19** (Blastodiniales) showed a high level of variation for a low number of sequences, genera and species described (**Table 2.3.2, Fig.2.3.4**). Superclade **#2** (Dinophysiales) had similar p-distance patterns to Superclade **#3** (Suessiales), but is represented by fewer than twice the sequences and a lower number of genera and species described.

Table 2.3.2: Number of unique and total dinoflagellate 18S rRNA gene sequences by Superclade included in the database. Number of dinoflagellate genera and species represented in the database by at least one sequence. Sequences not assigned to the species level (annotated as "sp.") were not considered. Total number of genera and species described (based on Gómez, (2012a), AlgaeBase (Guiry & Guiry, 2017), CEDIT (http://www.dinophyta.org/).

Superclades		No. of sequences in DinoREF		No. of taxa in DinoREF		Total No. of described taxa	
		Unique	Total	Genera	Species	Genera	Species
# 1	Gonyaulacales	507	543	17	85	20	296
# 2	Dinophysiales	97	97	8	38	13	358
# 3	Suessiales	223	240	14	29	26	91
# 4	Thoracosphaeraceae	71	82	16	22	19	66
# 5	Amphidomataceae	26	27	2	11	2	20
# 6	Kryptoperidiniaceae	21	23	5	10	6	16
# 7	*Pentapharsodinium - Ensiculifera*	6	6	2	4	2	6
# 8	Peridiniales *sensu stricto*	106	116	13	46	25	475
# 9	Heterocapasaceae	18	20	1	7	1	16
# 10	Podolampadaceae	7	7	4	7	8	42
# 11	Prorocentrales	70	78	2	28	4	68
# 12	*Akashiwo*	8	13	1	1	1	1
# 13	Gymnodiniales *sensu stricto*	115	129	16	41	21	341
# 14	Kareniaceae	31	38	4	9	8	40
# 15	*Gyrodinium*	15	15	1	7	3	112
# 16	*Amphidinium*	36	40	1	10	3	101
# 17	Torodiniales	9	9	2	3	2	3
# 18	Tovelliaceae	6	10	4	4	4	19
# 19	Blastodinium	29	32	1	8	1	13
# 20	Ptychodiscales	1	1	1	1	1	2
UTD	Uncertain thecate Dinophyceae	56	56	23	32	37	172
UND	Uncertain naked Dinophyceae	82	89	11	19	25	84
	Total	1540	1671	149	422	232	2342

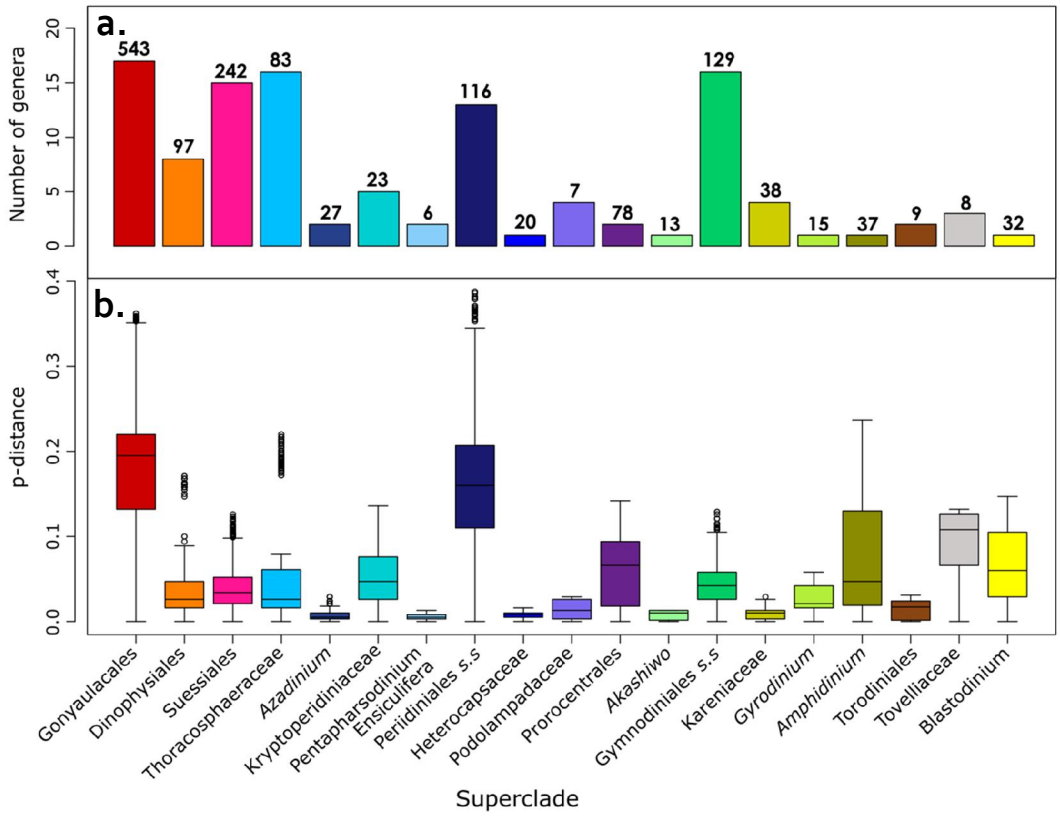


Fig. 2.3.4: a. Barplot showing the number of genera with 18S rRNA information in 19 of the 20 Superclades depicted in **Fig.2.3.2** and described in **Table 2.3.1**. Superclade 20 is not shown as it contains only one sequence. The number of sequences in each Superclade is specified on the top of each bar plot. Numbers indicated on the top of each bar plot represent the number of sequences available in the database for each Superclade. **b.** Boxplot showing the pairwise p-distances of the V4 regions in the 19 dinoflagellate Superclades. Graphs should be interpreted in regard of the numbers of sequences and genera included for the p-distance calculation.

P-distances allowed the pinpointing of specific genera with a high variation in the V4, but also genera with a weak or non-existent variation (**Fig.2.3.5**). The highest divergence rates were found for the genera *Protoperidinium, Gonyaulax, Gambierdicus* and *Amphidinium.* These rates were not linked with the number of sequences or species characterised molecularly for each genus (**Fig.2.3.5**). Some genera such as *Tripos*, *Dinophysis*, *Azadinium*, *Heterocapsa*, *Takayama*, *Karlodinium* or *Karenia* presented low variation or no variation at all in the V4 region, making differentiation of the species problematic or even impossible by barcoding.

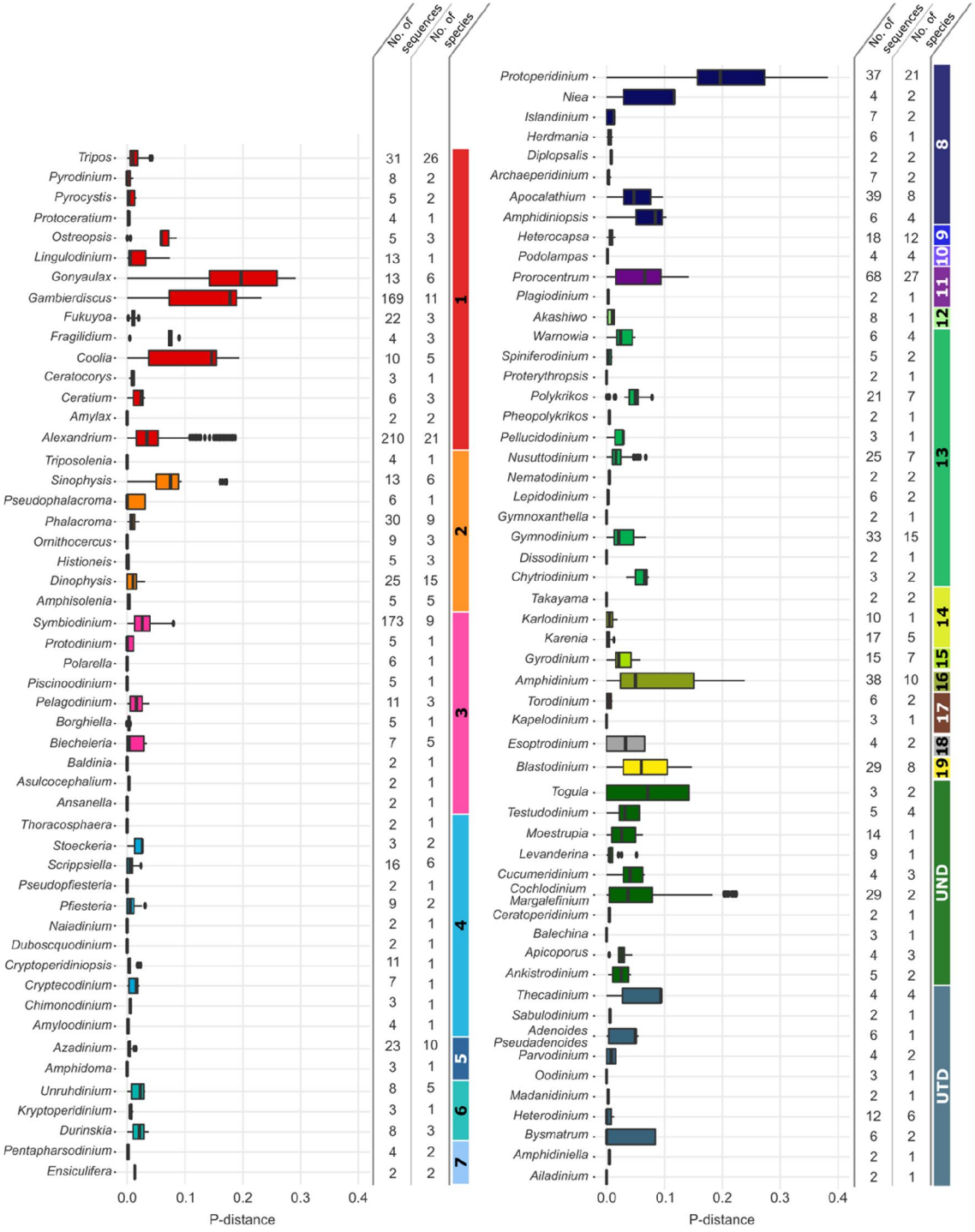


Fig.2.3.5: Boxplots showing the range of the pairwise p-distances over the V4 region of the 18S rRNA sequences within genera for each of the Superclades in **Fig.2.3.2** and described in **Table 2.3.1**. Number of sequences that have been used to calculate pairwise p-distance and number of species represented by those sequences are specified for each genus. Sequences with no species name (annotated "sp.") were not accounted for in the number of species, but still used for the calculation of pairwise p-distance. Graphs should be interpreted in regard of the numbers of sequences and species included for the p-distance calculation.

When clustered into OTUs at 98% of similarity, the unique 946 V4 sequences were reduced to 313 different OTUs (**Supplementary Material 7**, for 946 unique V4). Over 313 OTUs, 33 OTUs collapsed different species from the same genus and 12 OTUs clustered different genera together. In fact, 59 genera (40% of the genera represented in DinoREF) were clustered within the same OTU as other genera and were not discriminated in the analysis. For instance, a single OTU (**OTU #126**) collapsed V4 sequences belonging to *Gyrodinium*, *Scrippsiella*, *Karlodinium*, *Prorocentrum*, *Podolampas*, *Dubosquodinium*, *Azadinium*, *Heterocapsa*, *Kapelodinium* (**Supplementary Material 7**).

Table 2.3.3 List of V4 sequences failing to differentiate between species or orders in the DinoREF database. For each conflictual V4 (40 sequences), species and sequences (GenBank accession number) were specified. This table presents a summary of **Supplementary Material 6** detailing all 946 unique V4 sequences of DinoREF. V4 sequences were classified by Superclade (SC) in the same way as DinoREF (**Table 2.3.1**). Potential toxic taxa were annotated with a **T** and symbiont with a **S** following DinoREF annotations (**Supplementary Material 1**). V4 sequences collapsing different orders together were highlighted in **green**.

V4 No	SC	Sequences sharing the same V4
#26	**#1 : GONYAULACALES**	*Alexandrium hiranoi* **T** (LC056070, LC056068, AY641564), *Alexandrium pseudogonyaulax* **T** (AB088302, JF521638)
#67		*Alexandrium minutum* **T** (JF521631), *Alexandrium insuetum* (EU418967, JF521630, AB088298), *Alexandrium andersonii* **T** (JF521620)
#68		*Alexandrium ostenfeldii* **T** (KJ362003, KJ361992, KJ362001, KJ361998, JF521636, AJ535382, AJ535381, JF521637, AJ535384, KJ361990, AB538439), *Alexandrium andersonii* **T** (JF521621)
#77		*Alexandrium tamiyavanichi* **T** (AB088318, AB088323, AB088316, AB088317, AB088324, AB088325), *Alexandrium cohorticula* (AF113935)
#78		*Amylax buxus* (AB375868), *Amylax triacantha* (JX666361)
#299		*Tripos longipes* (DQ388462), *Tripos arietinus* (FJ402956), *Tripos symmetricus* (FJ402947), *Tripos euarcuatus* (FJ402946)
#300		*Tripos minutus* (FJ402964), *Tripos limulus* (FJ402962, FJ402952), *Tripos paradoxides* (FJ402965), *Tripos kofoidii* (FJ402963)
#301		*Tripos pentagonus* (FJ402948), *Tripos declinatus* (FJ402949)
#302		*Tripos petersii* (FJ402951, FJ402953), *Tripos azoricus* (FJ402954)
#303	**#2 : DINOPHYSIALES**	*Amphisolenia schauinslandii* (HM853766), *Amphisolenia globifera* (HM853765), *Amphisolenia bidentata* (HM853763)
#313		*Dinophysis acuminata* **T** (FJ869120, AJ506972, EU130569, AB073117, KJ508017), *Dinophysis norvegica* **T** (AY260470, AJ506974, AB073119, AF239261), *Dinophysis tripos* **T** (HM853816), *Dinophysis caudata* **T** (EU780644, HM853815), *Dinophysis infundibulum* **T** (AB366002), *Dinophysis fortii* **T** (AB073118), *Dinophysis acuta* **T** (AJ506973)
#315		*Histioneis longicollis* (HM853804), *Histioneis* sp (EU780646), *Histioneis gubernans* (HM853802), *Histioneis cymbalaria* (HM853801), *Ornithocercus quadratus* (EU780647, HM853800, HM853799), *Ornithocercus heteroporus* (HM853795, HM853793, HM853794, HM853796), *Ornithocercus magnificus* (HM853797, EU780651)
#325		*Phalacroma mitra* **T** (HM853775, HM853776, HM853777, HM853778), *Phalacroma rapa* (EU780655, FJ477082, HM853774), *Phalacroma* sp. (AB551248)
#326		*Phalacroma rotundatum* **T** (AJ506975), *Phalacroma oxytoxoides* (JQ996385, HM853782)
#337		*Sinophysis ebriola* (JQ996372, JQ996379), *Sinophysis grandis* (JN587291)

V4 No	SC	Sequences sharing the same V4
#347	#3 : SUESSIALES	*Biecheleria cincta* (JF794059, JN934667), *Biecheleria baltica* (EF058252)
#348		*Biecheleriopsis adriatica* (HG792067), *Protodinium simplex* (EF492493, U41086, DQ388466, EF492491)
#353		*Cystodinium phaseolus* (EF058235), *Phytodinium* sp (EF058251)
#367		*Symbiodinium* sp. S (AB016539, AB055916, AB085912, AB016595, AB126930, AF271291, AJ271761, JN255734, JN255733), *Symbiodinium* sp Clade C S (EF419289, KC816644, KC816643, KC816638, KC816631), *Symbiodinium goreaui* S (EF036539)
#374		*Symbiodinium* sp. S (AB055915, AB055913, AB055912, AB085914, AY051096), *Symbiodinium* sp. Clade E S (AF238261), *Symbiodinium* sp. Clade D (KC848881) S
#392		*Symbiodinium microadriaticum* S (KU900226, EF492514, JN717147), *Symbiodinium* sp. S (AY160124, KT860942, JQ320136), *Symbiodinium californium* S (AF225965)
#510	#4	*Duboscquodinium collinii* (HM483398, HM483399), *Pernambugia tuberosa* (KR362907), *Scrippsiella sweeneyae* (HQ845331), *Scrippsiella* sp. (LC054940, AB183674, JQ246506), *Scrippsiella acuminata* (KF733540, HQ845330, JX661036, AF274277)
#514		*Apocalathium aciculiferum* (EF417313, AY970653, EF417314, KF446621, EF417315), *Apocalathium malmogiense* (EF417316)
#538	#5	*Azadinium caudatum* (JQ247707, JQ247701), *Azadinium concinnum* (KJ481826)
#539		*Azadinium spinosum* T (FJ217814, JX262491, JX559885, JN680857), *Azadinium* sp. (JX661035), *Azadinium trinitatum* (KJ481803, KJ481808, KJ481813, KJ481815, KJ481817)
#651	#9	*Heterocapsa niei* (AF274265, EF492499), *Heterocapsa circularisquama* T (LC054932)
#655	#10	*Podolampas palmipes* (FJ888594), *Podolampas bipes* (FJ888595), *Podolampas spinifera* (FJ888597)
#693	#11	*Prorocentrum hoffmannium* T (JQ638934), *Prorocentrum maculosum* T (Y16236)
#694		*Prorocentrum donghaiense* (AJ841810), *Prorocentrum dentatum* (AY803742), *Prorocentrum shikokuense* (AB781324)
#696		*Prorocentrum mexicanum* T (DQ174089, Y16232, EF492510), *Prorocentrum rhathymum* T (HF565183, JQ616822, HF565181, FJ842096, KF733536, HF565182), *Prorocentrum texanum* (JQ390504), *Prorocentrum micans* (EF492511, EU780638, AY833514), *Prorocentrum cordatum* (DQ028763, JX402086, AJ415520, FJ587221, Y16238)
#722	#13	*Nusuttodinium acidotum* (JQ639760, AB921309), *Nusuttodinium aeruginosum* (AB921315, LC027037, AB921317, LC027038)
#726		*Spiniferodinium palustre* (AB921299), *Spiniferodinium palauense* (AB626150)
#729		*Lepidodinium viride* (DQ499645), *Lepidodinium chlorophorum* (AM184122), *Lepidodinium* sp. (AB686255)
#766		*Proterythropsis* sp. (FJ947036, FJ947037), *Warnowia* sp. (KP790169, KP790168)

V4 No	SC	Sequences sharing the same V4
#784	#14	*Karenia brevis* **T** (AF274259, AF352818, EF492501, FJ587219, EF492504, AJ415518), *Karenia mikimotoi* **T** (AF022195, FR865627), *Karenia* sp. (AJ415517)
#788	#14	*Karlodinium veneficum* **T** (JN986577, AF272045, EF036540, AF272046, AJ415516, HQ832504, AY121855), *Takayama pulchellum* (AY800130), *Takayama acrotrocha* (HM067010)
#830	#17	*Torodinium robustum* (KP790166, KP790167, KR139784), *Torodinium teredo* (KR139783)
#856	#19	*Blastodinium mangini* (JX473664, JX473655), *Blastodinium* sp. (JN257679, JN257677), *Blastodinium navicula* (JX473665)
#911	UND	*Togula britannica* (AY443010), *Togula jolla* (AF274252)
#924	UTD	*Heterodinium scrippsii* (JQ446589, JQ446590, JQ446591), *Heterodinium rigdeniae* (JQ446588), *Heterodinium globosum* (JQ446586, JQ446587), *Heterodinium milneri* (JQ446582, JQ446583, JQ446584, JQ446585)
#939	UTD	*Thecadinium yashimeansi* (AY238477), *Thecadinium inclinatum* (EF492515)

2.4. Discussion

The DinoREF database is a database providing 18S rRNA sequences of dinoflagellates with a curated and updated taxonomical framework. Reference taxonomic checklists exist for dinoflagellates (Gómez, 2012a; Centre of Excellence for Dinophyte Taxonomy ***CEDiT***, Adl *et al.*, 2012), as do databases collecting 18S rRNA gene sequences for protists such as PR² (Guillou *et al.*, 2013) or Silva (Quast *et al.*, 2013). However, none of them combine both an updated taxonomic framework and curated sequence database. Yet I should always consider all genetic, morphological and ecological information in order to study an organism (Keeling, 2013). DinoREF is a ready to use database reconciling and updating both molecular, morphological and ecological data available for dinoflagellates in a single file. Users can use it in many different ways and different types of information can be extracted. For example, it can be used to annotate environmental data, as a reference database for phylogeny, as a check for authors publishing sequences, to extract ecological data, information on toxicity, the place of collection or the technique used to obtain the sequence. Information provided has been validated manually sequence by sequence. No existing database provides such comprehensive information and utility.

Composition and coverage of the DinoREF database

About one third of DinoREF is composed of sequences from three genera: *Alexandrium*, *Symbiodinium* and *Gambierdiscus*. The database is biased toward species that are either toxic, autotrophic or have been the focus of research (e.g., *Symbiodinium* endosymbiont of corals). Noticeably, these species can all be grown in culture, therefore are easier to characterise than uncultured ones and are better referenced in databases (Gómez, 2014). In contrast, heterotrophic and mixotrophic species are under-represented because they are more difficult to isolate and maintain. Data from single cells isolated and sequenced directly from environmental samples is also more difficult to obtain. For 63 dinoflagellate genera, only one or two reference sequences are available. For example, *Ansanella* or *Ailadinium*, include only one species (described recently with

molecular data) whereas others, such as for *Tripos* or *Dinophysis*, included a higher number of morphologically described species and are clearly underrepresented in terms of 18S sequences.

Considering the total number of described dinoflagellate species (more than 2,300 according to Gómez (2012b) and Hoppenrath (2017); 2,342 species validated in this study), only 22% of them are represented by an 18S reference sequence (**Table 2.3.2**). The comparison of my data with the estimation made by Gómez (2014) showed that the number of species and genera represented by at least an 18S reference sequence increased in a five-year period (2012 to 2017) from 345 to 422 (+18%) and from 131 to 149 (+13%) respectively, mainly due to the addition of newly described species and revision of some genera (e.g. *Amphidinium, Katodinium, Gymnodinium, Peridinium*; (Hoppenrath *et al.*, 2012; Takano *et al.*, 2014; Boutrup *et al.*, 2016; Craveiro *et al.*, 2017). This is a trend also reported in the researcher community working on dinoflagellates (Hoppenrath and Saldarriaga, 2008).

Yet 84 genera and 1886 species validly described still need to be characterised molecularly (**Table 2.3.2**). Some genera contain a high number of species that have only ever been observed once suggesting a possible over-estimation of their diversity (Thessen, Patterson and Murray, 2012). This is especially the case for *Gymnodinium* (270 species, 38% of which have not been observed since their original description), but also *Gyrodinium, Amphidinium, Glenodinium, Peridinium or Gonyaulax*. On the other hand, many of described species do exist. Some of them are regularly observed by taxonomists in environmental samples (e.g. *Asterodinium, Gynogonadinium, Lissodinium* or *Centrodinium* genus). Many of these dinoflagellates are heterotrophic and observed by taxonomists on rare occasions. Reports of these dinoflagellates to the scientific community became more difficult because of the necessity of molecular characterisation for description and publication (Gómez, 2014). Therefore, these genera need further investigation to validate their identity and study of their diversity.

Moreover, recent metabarcoding studies suggest that there is still a high biodiversity of unknown dinoflagellates (Massana *et al.,* 2015; LeBescot *et al.,* 2016; Piredda *et al.;* 2017). Hence, an underestimation of dinoflagellates is more likely than overestimation.

DinoREF does not include the sequences related to dinoflagellates clustering at the base of the dinoflagellate lineages (i.e *Amoebophrya, Ichtyodinium, Hematodinium, Syndinium, Psammosa*). Early branching dinoflagellates include mainly heterotrophic parasites with an unstable taxonomical position in the dinoflagellate lineage. Most of them are poorly characterised lacking typical dinoflagellate morphological features and are principally known from environmental sequences (Okamoto, Horák and Keeling, 2012). Two exceptions are *Noctiluca* and *Oxhyrris*, that have been studied intensively and cultivated in laboratory (Fukuda and Endoh, 2008; Ki, 2010; Lowe *et al.*, 2010). Their 18S rRNA sequences have been shown to diverge early from the core dinoflagellates producing long branches in phylogenies and aligning poorly with other dinoflagellates which makes their curation particularly difficult. Sequences of early branching dinoflagellates can be accessed in a separate file (**Supplementary Material 8**).

Most of the sequences selected in the database (1,200 sequences) are full-length or close to full-length (**Fig.2.3.3**). Sequences produced from a single cell are usually between 1100 bp and 1400 bp. This can be explained by the fact that most common single cell protocols involve adding a cell directly to the PCR mix and sequencing with 2 primers (small quantity of DNA).

Curation procedure

The curation procedure allowed us to update the name of 440 sequences and provides notes for 300 sequences in the database (**Supplementary Material 1**). One fifth of the names describing the sequences in GenBank are synonyms, not accepted anymore or inaccurate. Ideally, these names should be updated by authors online according to the evolution of dinoflagellate taxonomy. In reality, a tiny proportion of sequence information is corrected or refined after publication, which generates an accumulation and a propagation of errors. Only specialists of dinoflagellates can understand and keep up with dinoflagellate classification. For each sequence, a minimum set of

metadata should be provided including original location with environmental data (e.g. temperature, salinity etc…), but also isolation methods, association, and toxicity for example.

In addition, the curation process has enabled the detection of sequences wrongly assigned to a genus and/or species of dinoflagellate. Possible reasons for these mistaken assignments include wrong identification due to the morphological convergence of distinct species, sequences belonging to endosymbionts, parasites or contaminations instead of the targeted dinoflagellate, and incorrect assignation of sequences based on "best" Blast results.

In this study, I propose a baseline taxonomic framework classifying the dinoflagellates in "Superclades" (**Table 2.3.1**). These Superclades have been carefully investigated and documented based on the most recent specialised literature (**Table 2.3.1**). It is a working hypothesis erected to create a practical higher-level classification of dinoflagellates for the database. It is meant to evolve with outputs from future research. The need to update higher taxonomical levels have been stated by dinoflagellate specialists (Hoppenrath and Saldarriaga, 2008; Hoppenrath, 2017). Recent investigations of dinoflagellates integrating both molecular, morphological and ecological information have still to be translated into a new classification of dinoflagellates.

The building of an 18S phylogeny was not the goal of the present study, but a means to curate and create the DinoREF database. Unfortunately, the basal polytomy in the 18S rRNA phylogeny, from which the various dinoflagellate sub-clades emerge, leaves the phyletic status of most of the higher taxa unresolved. The recovery of such morphologically and or ultrastructurally defined taxa in multiple clades emerging directly from the polytomy neither supports nor opposes the natural status of these taxa. Yet I have retained these taxa in our classification system because other studies with multiple sequence markers have shown evidence for their monophyly (e.g., Orr *et al.*, 2012; **Table 2.3.1**). Exceptions occur of course but that is mostly a problem of a taxonomy not fully established yet.

Nonetheless, the sequences grouping within the recovered sub-clades in the 18S tree are usually from species that share morphological and ultrastructural characteristics, i.e., belong to related species and genera or even families. Remarkably, the results (**Fig.2.3.2**) show that the phylogenetic patterns – Superclade classification (**Table 2.3.1**) - correspond reasonably well with that higher taxonomy of orders and families. No conflict was detected between the sub-clades recovered in the majority consensus tree and the Superclades because most of the species of dinoflagellates have been described based on 18S phylogenies which mainly show strong support at lower taxonomical level such as for genera or families (Gómez, 2014; Hoppenrath, 2017). However, in some cases, genera found in different sub-clades are suspected to be poly or paraphyletic and need to be amended or further investigated (Hoppenrath, 2017). For instance, *Peridinium, Gonyaulax,* and *Warnowia* sequences have been found to cluster in different clades. Other genera like *Protoperidinium* are clearly paraphyletic. The same patterns have been observed for *Protoperidinium* by other authors (Gu, Liu and Mertens, 2015; Liu *et al.*, 2015; Mertens *et al.*, 2015).

V4 variation

The success of metabarcoding studies mainly rely on the choice of the marker and its capacity to detect and discriminate all taxa of interest (Bendif *et al.*, 2014; Hu *et al.*, 2015). The V4 region of the 18S rRNA gene is considered variable enough to distinguish between different species of dinoflagellates (Ki, 2012) and has been used to target dinoflagellates in metabarcoding studies (Onda *et al.*, 2017; Smith *et al.*, 2017). However, our results show that when sequences are restricted to the V4 region, 11 % of the species collapse and 13 different genera show conflicting ribotypes (i.e. unique sequence of V4) included in DinoREF (**Table 2.3.3**). In some cases, V4 ribotypes are identical among different species while in other cases species belonging to specific genera possess a large number of different V4 (such as *Alexandrium fundyense* or *Gambierdiscus scabrosus* for example). The V4 region then does not often allow discrimination between morphologically and ecologically different species. For instance, in the case of surveys of toxic

species belonging to *Karenia* or *Azadinium*, the V4 cannot distinguish between closely related toxic and non-toxic species (**Table 2.3.3**).

V4 pair-distance distribution varies between dinoflagellate Superclades and genera (**Fig 2.3. 4** and **Fig 2.3.5**) illustrating perfectly the power of resolution of the V4 between different lineages of dinoflagellates. V4 variation is of course related to the number of sequences included in the analysis of a given group. However, there are some clear differences in the V4 region evolutionary rates between some genera (e.g. *Symbiodinium, Gambierdiscus, Alexandrium, Prorocentrum*) which confirm what was previously noted by Ki (2012) when he compared hypervariable regions (V1-V9) of the dinoflagellate 18S. Therefore, the resolution of the marker and the ability to capture genetic diversity varies greatly among different lineages of dinoflagellates. Yet, large variations in the V4 pairwise distance may also highlight artificial grouping of some organisms within a genus (Ki, 2012).

For environmental surveys, it is current practice to reduce the number of sequences obtained from Next Generation Sequencing in OTU (Operational Taxon Unit) generally at 97%, 98% or 99% of similarity (Hu et al., 2015) and is systematically applied for dinoflagellates (Massana *et al.,* 2015; LeBescot *et al.,* 2016; Onda *et al.,* 2017; Smith *et al.,* 2017). This practice, even if very useful to reduce the number of machine errors and collapse intra species variability, has its pitfalls. As previously reported for dinoflagellates in Massana *et al.,* (2015), even at 97% similarity OTUs include more than one species or genera. In addition, other authors, recently reported that such clustering with short marker (<400 bp) resulted in fewer OTUs and produced lower value for common diversity indices than full 18S sequences for protist (Hu *et al.,* 2015).

In our DinoREF dataset the cluster at 100% of similarity (unique V4 sequences or ribotypes) highlighted that differentiation of some species is problematic or even impossible with the V4 region. (**Table 2.3.3, Supplementary Material 6**). In collapsing dinoflagellates at 98% (**Supplementary Material 7**), I found several patterns, however, several OTUs cluster together

sequences belonging to different orders. This OTU approach could potentially be used to study species showing high intraspecific variation such as *Alexandrium*, *Gambierdiscus* or *Protoperidinium* but should not be used to assess dinoflagellate diversity from environmental surveys.

Given these points, I recommend general caution in the use of OTUs of V4 environmental metabarcoding data, especially in fine grain explorations of diversity, and advise a ribotype (OTUs at 100%) approach rather with OTUs at 99, 98 or even 97 % similarity for taxonomic characterisation of dinoflagellates from molecular data.

CHAPTER III: Dinoflagellate diversity and seasonal changes at the LTER station in the Gulf of Naples as inferred from HTS metabarcode data

In Chapter III, Solenn Mordret collected data and the designed framework for molecular and computational analyses under the supervision of Roberta Piredda.

3.1. Introduction

The term "dinoflagellate" is commonly used to describe one of the major lineages of modern unicellular eukaryotes (Adl *et al.*, 2012). These unicellular organisms can range from 5 µm to a few millimetres in size and display a wide spectrum of different shapes and functions (Hoppenrath, 2017). They occur in all aquatic environments from marine to freshwater, pelagic to benthic, neritic to open ocean habitats and can be extremely abundant and diverse (Not *et al.*, 2012). In marine environments, dinoflagellates are known to represent an essential part of the "plankton" community, small organisms drifting passively in the sea along with the current. These planktonic dinoflagellates mostly occupy the sunlit surface layer of the world's ocean but can also thrive in deeper layers. About half of the dinoflagellate species are phototrophic while the remainder are considered to be mixotrophic or completely heterotrophic, having lost their plastids. All of them act as important primary producers, consumers, decomposers or even symbionts and parasites of global marine trophic chains (Gómez, 2012b; Murray *et al.*, 2016; Taylor, Hoppenrath, & Saldarriaga, 2008). Dinoflagellates are also of societal relevance because many produce biochemically complex bio-active compounds with possible blue biotechnological applications (Murray *et al.*, 2016) and several can produce toxic compounds that affect animal and human wellbeing (Anderson, Cembella and Hallegraeff, 2012; Berdalet *et al.*, 2016; Graham *et al.*, 2016). These potentially harmful dinoflagellates are one of the reasons why many coastal administrations have implemented plankton monitoring stations.

At the Long-Term Ecological station MareChiara (LTER-MC), phytoplankton, zooplankton and physico-chemical data was collected every two weeks between 1984 and 1991, and weekly since 1995, in order to study the functioning of a coastal pelagic ecosystem (Ribera d'Alcalà *et al.*, 2004). The LTER-MC is a coastal station located in the Gulf of Naples (Tyrrhenian Sea, Italy) and directly influenced by one of the most densely populated metropolitan area in Europe. The station is positioned inside a semi-enclosed embayment with complex oceanographical dynamic, mainly influenced by two hydrological systems: one oligotrophic coming from offshore of the Gulf and

another one with more eutrophic characteristics coming from the coast (D'Alelio *et al.*, 2015). These two systems undergo regular seasonal shift impacting LTER-MC communities and diversity. In winter LTER-MC is more influenced by coastal inputs and strong water column mixing, while summer is known as stratified and often characterised by oligotrophic offshore water masses entering in the bay (Ciannelli *et al.*, 2015; 2017). Surface temperatures normally vary from around 13°c in late winter (February-March) to 28°C in the middle of summer (July- August) showing a regular sinusoidal pattern along the years (**Appendix 2;** Carada *et al.*, 1980). Surface salinity range between 36,7 and 38,1 psu. Spring samples are characterised by low salinity values while late autumn–winter show higher salinity values. Surface chlorophyll *a* usually reached highest concentrations in March-May and September-October. Nutrient concentrations (i.e., Nitrate, Nitrite, Ammonium, Phosphate and Silicate) vary greatly through and between different years and where interpreted as punctual events. In terms of climatic characteristics, LTER-MC site can be compared with other subtropical long-term stations in the Mediterranean Sea, such as Villefranche (point B) and Blanes Bay Microbial Observatory, or in other areas, such as as San Pedro Station (SPOT, California) in the Pacific Ocean. However, the proximity to the coast of a highly urbanised area and, at the same time, the influence of the oligotrophic water from Tyrrhenian Sea makes the LTER-MC station a unique site.

Over 30 years, protist communities have been intensively studied and more than 750 species of phytoplankton, including 325 dinoflagellates taxa, have been morphologically identified by taxonomists analysing samples from net samples and Niskin bottles (courtesy Diana Sarno). Results of cell counts and taxonomic studies over the decades have revealed marked seasonal and yearly variation of phytoplankton species. Dinoflagellates form an important constituent of the plankton community at the LTER-MC monitoring site, especially in terms of biomass (Ribera d'Alcalà *et al.*, 2004). Knowledge has been collected over many years comprising many species descriptions and information on life cycles, population structure, morphological and genetic diversity but also on seasonal trends of plankton communities throughout the year (Montresor and

Zingone, 1988; Montresor *et al.*, 2003, 2004; Siano and Montresor, 2005; Zingone *et al.*, 2010; Percopo *et al.*, 2013; D'Alelio *et al.*, 2015). However, observation and enumeration using light Microscopy (LM) can only uncover dinoflagellate taxa that are easily identifiable such as a number of thecate species, while others have been neglected and remain unknown, principally because of their lack of particular morphological features allowing identification, their small size and indistinct "ball-like" shapes or because they are symbionts or parasites. Likewise, naked dinoflagellates (lacking cellulose armour plates) can also be damaged or under sampled by traditional sampling methods.

In recent years alternative methods have been developed and applied to assess protist diversity. The application of molecular methods to study marine ecosystems revolutionised our view of microbial community structure and functioning (Caron, 2013; Kim *et al.*, 2014). Overall, DNA sequence data has allowed a better understanding of the phylogenetic relationships of dinoflagellate lineages, as well as a better understanding of the boundaries between species (Caron, 2013; De Clerck *et al.*, 2013; Hu *et al.*, 2015). Genetic diversity was discovered in some well-studied morphotypes that were previously considered to be single species (Nishimura *et al.*, 2013; John *et al.*, 2014), and entire new groups of marine alveolates were also detected (Guillou *et al.*, 2008).

Recent technological advances such as High Throughput Sequencing (HTS), used for a metabarcoding approach, provide a large amount of new sequences from the sequencing of environmental samples. This metabarcoding approach has now been widely applied to assess microbial biodiversity worldwide in both terrestrial and aquatic environments. The methods were adapted successfully to study planktonic diversity and have been used in an increasing number of studies in recent years (Amaral-Zettler *et al.*, 2009; Lallias *et al.*, 2015; Piredda *et al.*, 2017), including many important projects such as the European coast survey (BioMarks)(Massana *et al.*, 2015), the Tara Ocean expedition sampling world oceans (de Vargas *et al.*, 2015; Malviya *et al.*,

2016) and the Ocean Sampling Day program (Koft *et al.*, 2015; Penna *et al.*, 2017; Tragin, Zingone and Vaulot, 2017).

These results offered the opportunity to study a community at the ecological level and catch the diversity of microbial assemblages with a high-resolution. This also facilitates the discovery of large numbers of protists, whilst participating in biodiversity monitoring and detection of small species and rare taxa in natural samples (Caron, 2013). DNA sequencing of common genetic eukaryotic markers revealed an unpredicted and unknown level of protist diversity in the ocean and particularly for dinoflagellates (Stern *et al.,* 2010). Indeed, most studies of protist communities showed that dinoflagellates and alveolates in general represent the most important number of reads obtained and display a high diversity (de Vargas *et al.*, 2015; Genitsaris *et al.*, 2015; Massana *et al.*, 2015; Brannock *et al.*, 2016).

For all these reasons, at LTER-MC station, weekly seawater sampling was implemented including a filtration step, which provided material for DNA extraction, in parallel with collection of fixed and live material cell counts and observations. These samples collected over 3 years were selected for a HTS metabarcoding approach which provided me with a meta-barcode dataset of the V4 region in the 18S for planktonic protists. An initial study performed on the first eight sampling dates (2011) showed the potential of the approach to assess the variation and the diversity of plankton communities at a temporal scale (Piredda *et al.*, 2017). In this study, I used an expanded dataset (48 dates) including the first 8 dates samples to explore dinoflagellate diversity specifically at finer taxonomic resolution.

Only a few studies using metabarcoding technologies have focused directly on dinoflagellates, considering their fine taxonomic and ecological diversity. Initial studies favoured the use of the mitochondrial barcodes such as cox1 and cob genes (Lin *et al.*, 2009; Stern *et al.*, 2010; Kohli *et al.*, 2014) to evaluate dinoflagellate diversity. However, the lack of references for a majority of characterised dinoflagellates hindered the detection and classification of most of the diversity. Other metabarcoding studies demonstrated the potential of the fast-evolving Internal Transcribed

Spacers (ITS) in the nuclear encoded ribosomal RNA cistrons to identify dinoflagellates at the species level (Litaker et al., 2007). However, these loci show a high propensity to chimaera, paralogues and poor primers, binding across dinoflagellates, and therefore should not be considered for HTS environmental studies (Stern *et al.*, 2012). Recently, metabarcoding studies adopted rRNA barcodes for dinoflagellate metabarcoding. New studies selected the 28S rRNA-encoding (D1-D2) region as barcode (Grzebyk *et al.*, 2017; Smith *et al.*, 2017) with success and high-resolution power for dinoflagellates. However, the 18S rRNA encoding region remains the most used marker in environmental studies of dinoflagellates due to i) a high number of reference sequences in public databases, ii) good primer binding and iii) easy PCR amplification given the relatively high numbers of copies of rRNA cistrons in dinoflagellates. The short V9 region (about 150 bp) was used to characterise Tara Ocean planktonic dinoflagellate global patterns by Le Bescot and colleagues (2016) but most of studies chose the somewhat longer V4 rRNA region (about 380 bp) as barcode (Kohli *et al.*, 2014; Onda *et al.*, 2017; Smith *et al.*, 2017). The V4 rRNA gene is considered the most variable region of 18S rRNA for dinoflagellates (Ki, 2012) and protists in general (Hu *et al.*, 2015). Introduced as an universal eukaryotic pre-barcode (Pawlowski *et al.*, 2012), it is also the barcode for which most references exist at a species level (422 for the V4 rRNA compared to 329 for V9 rRNA, Le Bescot, 2014). However, while most studies focus on benthic and harmful dinoflagellate distribution at a spatial level, only one provides information on the variation patterns of dinoflagellate diversity over time (Onda *et al.*, 2017) but exclusively for the Arctic region.

The aim of the present study is to explore the diversity of dinoflagellates at the coastal station Long Term Ecological Research – MareChiara (LTER-MC) in the Gulf of Naples (Tyrrhenian Sea, Mediterranean Sea) by means of a metabarcoding approach. The station represents a strategic point to study the Mediterranean population of pelagic phytoplankton and provides an ideal framework to set a metabarcoding experiment with the necessary research background. Many groups of dinoflagellates remain poorly studied and represent "a black box of diversity" that is still

waiting to be elucidated. In this work, dinoflagellate sequences were extracted from the total planktonic community by blasting reads against the protist database. The newly curated reference database DinoREF was tested on the HTS dataset and used to assign and classify all ribotypes recovered. Then the results were analysed via two different approaches: first ataxonomically at the community level without considering sequence assignations and second, taxonomically, classifying ribotypes according to the DinoREF database.

The objectives of our query of the LTER-MC meta-barcode dataset is to assess i) which part of the species diversity can be uncovered using the DinoREF database, ii) what is the relative importance of the various dinoflagellate taxa at the site, iii) if it is possible to detect specific annual and seasonal patterns, iv) what is the "capacity" of the V4 marker to capture dinoflagellate diversity in an environmental dataset.

3.2. Material and methods

3.2.1. Sampling at LTER MareChiara

The Long-Term Ecological Research station MareChiara (LTER-MC) located in the Gulf of Naples (Tyrrhenian Sea, Italy; 40°48.5' N and 14°15'E) has been sampled for physical and biological parameters every week since 1995 (**Fig.3.2 1a** and **1b**). The station belongs to the International Long Term Ecological Research network and is considered as a reference of marine plankton monitoring in the Mediterranean Sea. To assess marine protist diversity in the Gulf of Naples and as part of the Italian and European projects "EU-BioMarKs" and "FIRB Biodiversitalia", HTS metabarcoding data was gathered for 48 dates over a three-year time window (2011 to 2013) at MareChiara station (**Fig.3.2.1c**; Piredda *et al.*, 2017). These dates were selected by Dr. Adriana Zingone and Dr. Diana Sarno as representative of different trophic and seasonal conditions of the pelagic system of the Gulf of Naples, based on phytoplankton abundance data. Except for two months (August 2012 and November 2013), all months were sampled at least once (**Fig.3.2.1c**). In parallel with plankton sampling, environmental parameters including temperature, salinity, chlorophyll, ammonium, nitrate, nitrite, silicate and phosphate were surveyed at each date (**Appendix 2**).

Three litres of surface seawater were collected weekly and divided in three biological replicates. The seawater was filtered on cellulose ester filters (47 mm Ø, 1.2 µm pore size, EMD Millipore, MA, USA) immediately after sampling. Filters were instantly frozen in liquid nitrogen and preserved at -80 °C.

3.2.2. DNA extraction, amplification and sequencing

Two filters were divided in two and DNA were extracted from each half for the 48 dates (2 x 48 samples) using the DNeasy 96 Plant Kit (Qiagen GmbH, Hilden, Germany) following manufacturer's instructions. DNA concentration and quality were verified with a NanoDrop Spectrophotometer (NanoDrop 1000, Thermo Fisher Scientific Inc., Wilmington, DE, USA).

Filtration was performed by Anna Italiano. DNA extraction and sample preparation procedures were conducted by Dr. Roberta Piredda and Dr. Maria Paola Tomasino.

Amplification and sequencing of the DNA were carried out at the Molecular Biodiversity Lab (MoBiLab) of the ESFRI LifeWatch-Italy (CNR, Bari, Italy) using an Illumina MiSeq platform (Illumina, San Diego, CA, USA; Kozich *et al.,* 2013). A first amplification (PCR 1) was performed using V4 BioMarKs universal primers (Stoeck *et al.*, 2006) slightly modified in order to optimise protist specificity (**Table 3.2.1**; Piredda *et al.,* 2017). The second round of amplification (PCR2) involved using the same primers but with adapters (**Table 3.2.1**).

Table 3.2.1: Primers and adapter used for amplifying and sequencing on the Illumina Platform

	Adapter + **V4 primer sequences (modified by Piredda *et al.*, 2017)**
V4_18SNext forward	5′ -tcg tcg gca gcg tca gat gtg tat aag aga cag **CCA GCA SCY GCG GTA ATT CC**-3'
V4_18SNext reverse	5'-gtc tcg tgg gct cgg aga tgt gta taa gag aca g**AC TTT CGT TCT TGA TYR ATG A**-3'

The first PCR mixture (PCR 1) amplifying the V4 region (25 µL final volume) contained 2.5 ng or 5.0 ng of extracted DNA, 1X Buffer HF, 0.2 mM dNTPs, 0.5 µM of each primer, and 1U of Phusion High-Fidelity DNA polymerase (New England Biolabs Inc., Ipswich, MA, USA). Amplification cycles were conducted with the following PCR program: initial denaturation step at 98 °C for 30 s, followed by 10 cycles of denaturation at 98 °C for 10 s, annealing at 44 °C for 30 s, extension at 72 °C for 15 s, and subsequently 15 cycles of denaturation at 98 °C for 10 s, annealing at 62 °C for 30 s, extension at 72 °C for 15 s, with a final extension step of 7 min at 72 °C. After visualisation on agarose gel (1.2%), PCR products recovered from the amplification of two separate half filters were pooled in one sample per date. The amplicons obtained in the first PCR (PCR 1) were purified using the AMPure XP Beads (Agencourt Bioscience Corporation, Beverly, MA, USA) at concentration of 1.2x vol/vol and re-amplified in a second PCR (PCR2). In PCR 2, templates were amplified using a mixture of the Nextera index primers and Illumina P5 and P7 primers as required by the Nextera dual index procedure. Incorporation of unique molecular codes at both ends of the amplicons

allowed sample identification for bioinformatic multiplexing (Kozich *et al.*, 2013). This second PCR (PCR2) contained: DNA amplicons from PCR1 (40 ng), 1X Buffer HF, dNTPs (0.1 nM), P5 and P7 index primers and 1 u of Phusion High-Fidelity Polymerase for a final PCR volume of 50 µL. PCR 2 was performed following the standard Illumina Nextera cycle instructions. Both PCR1 and PCR2 were executed with a negative control (RNase/DNase-free water).

At the end of the process, DNA amplicons were about 550 bp long including 400 bp of V4 region and 150 bp of Illumina Nextera adapters. All amplicons were purified using an AMPure XP Beads kit and re-suspended at a concentration of 0.6X vol/vol. The quality of the DNA amplicons was checked using a 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA) and quantified by fluorimetry using the Quant-iTTM PicoGreen-dsDNA Assay Kit (Thermo Fisher Scientific, Waltham, MA, USA) on a NanoDrop 3300 (Thermo Fisher Scientific). Equimolar quantities of V4 amplicons were pooled and subjected to 2x250 bp sequencing on a MiSeq platform to obtain a total of about 375,000 pair-end reads per sample.

3.2.3. Bioinformatic pre-processing of the dataset

All first steps of the process were performed by Dr. Roberta Piredda (**Fig.3.2.2** – all steps in purple boxes). The initial quality control check of reads in fastq files was performed using FastQC tool in the Galaxy web-based platform (https://usegalaxy.org/) (Giardine *et al.*, 2005; Blankenberg *et al.*, 2010; Goecks *et al.*, 2010). Illumina paired-end reads (2x250 bp) were processed using the Mothur v.1.33.0 software (Schloss *et al.*, 2009) following the standard operating procedure (http://www.mothur.org/wiki/MiSeq_SOP). Reads were then assembled into contigs. The overlap in V4 fragment was 81 bp on an average (st. dev. 11.3); differences in base calls in the overlapping region were solved using ΔQ parameter as described in Kozich et al. (2013).

Primer sequences were removed and no ambiguous bases were allowed; the maximum homopolymer size was 8 bp. Redundant sequences were removed from the remaining sequences and aligned to a reference alignment (silva.seed v119), and the sequences that did not align to the target region were removed. The pre-clustering algorithm (Huse *et al.*, 2010) was used to further

remove sequence noise, allowing 1 nucleotide difference for every 100 bp of sequence, and the resulting sequences were screened for chimeras using UCHIME in *de novo* mode (Edgar *et al.*, 2011).

3.2.4. Taxonomic assignation and extraction of dinoflagellates sequences

A primarily taxonomic assignment was performed using a naïve Bayesian classifier (Wang *et al.*, 2007) trained with the PR² database (Guillou *et al.*, 2013), with an 80% bootstrap confidence threshold, in order to detect non-protist groups (including Bacteria, Archaea and Metazoa) and exclude the sequences from the analyses (step performed by Dr. Roberta Piredda). Then, taxonomic assignment was performed for all sequences using BLAST (Altschul *et al.*, 1990) against a custom version of PR² database containing the DinoREF database, early-branching dinoflagellates sequences and all new Dinophyceae sequences published in GenBank between August 2016 and April 2017. A few sequences (23) produced in this thesis from culture or single cells of dinoflagellates from the Gulf of Naples were also used as references and listed in **Appendix 3**. Sequences blasting against a dinoflagellate reference (blastn -db custom_PR2_dataRef -perc_identity 90 -query 48_NGSsamples.fasta -outfmt 6 -max_target_seqs 1 -out 48_NGSsamples_besthit.txt). The output file was filtered and I retained only the sequences with a minimum query coverage of 70% (270 bp). All ribotypes assigned to a dinoflagellate reference were extracted to create the dinoflagellate Mare Chiara dataset. At this stage, the taxonomic framework used was the same as in DinoREF (Phylum, Class, Superclade, Order, Family, Genus and Species).

All downstream analyses were performed using dinoflagellate ribotypes with an abundance of more than 3 reads (**90% dataset**).

Ribotypes recovered assigned in a range between 90 – 100% similarity (**90% dataset**) were used to perform ataxonomic analyses on alpha and beta diversity. All taxonomic descriptions were then based on a subset of the dinoflagellate dataset considering ribotypes assigned in a range of 97-100% similarity (**97% dataset**) for which I could reasonably apply a classification at the genus level.

3.2.5. General description of the dataset

The **90 % dataset** was explored in terms of number of reads, number of ribotypes and similarity with the reference. I generated plots of distribution of similarity for ribotypes and distribution of the abundance of ribotypes (number of reads). I defined as most abundant ribotypes, the ribotypes with an abundance higher than 1,000 reads corresponding to ribotypes contributing to at least 0.045% of the total number of reads.

In order to compare dinoflagellate patterns with those of another very well-studied and abundant protist group, the diatoms, I generated a table reporting the relative abundance of dinoflagellates (Dinophyceae and early branching dinoflagellates) and diatoms (Bacillariophyta) as compared with the total protists (**Table 3.3.1**).

3.2.6. Ataxonomic explorative analyses

For alpha diversity, the Shannon-Wiener index was calculated for each date. All analyses were performed in R environment (R Core Team, 2013) using vegan package (Oksanen *et al.*, 2016).

For temporal diversity analyses, normalisation among samples was performed randomly subsampling the ribotype table (**90% dataset**) to the number of sequences present in the sample with the lowest amount with exclusion of two samples (10,333 reads, rrarefy function in Vegan R package). Using this normalised table, community dissimilarity data matrices were computed (vegdist function in vegan) using the Bray–Curtis index (abundance data) and then used in subsequent analyses (clustering and correlation with environmental parameters). Significance of clustering was performed using an ANOSIM test. The ANOSIM algorithm assesses if the similarities within clusters are smaller or equal to the similarities between clusters. The same analysis was also performed on the **97% dataset** (4,558 reads, rrarefy function in Vegan R package).

To further explore the relationships between environmental and sequence data, the BIO-ENV analysis (Clarke and Ainsworth, 1993) was performed using the Bray-Curtis dissimilarity matrix.

BIO-ENV allows the identification of a subset of variables that shows the highest explanatory values. The identified variables were used in a Canonical Correspondence Analysis (CCA).

3.2.7. Taxonomic patterns

A first overview of the **90% dataset** was performed at the genus-level generating boxplots showing the range of similarity of ribotypes with the reference sequences recovered for each genus. In the following stages the **97% dataset** was used for all analyses. Taxonomic patterns were explored at Superclade and genus level. First, I generated treemaps (Tennekes and Ellis, 2017) showing the taxonomic composition for all 48 samples and then clustering samples according to the hierarchical seasonal clustering results (Cluster 1, Cluster 2 and Cluster 3).

Linear Discriminant analysis (LDA) and Effect Size (LEfSE) (Segata *et al.*, 2011) was used to identify differentially abundant taxa that are consistent with biologically meaningful categories (seasonal clusters). In particular, I performed the two tests: i) to detect taxa with significant differential abundance in three categories (Cluster 1: "winter", Cluster 2: "spring – mixed seasons" and Cluster 3: "late summer-autumn"); ii) to detect taxa with significant differential abundance in two categories (Cluster 1: "winter against Cluster 2 and Cluster 3 together). The outlier (16^{th} of July 2013 for the **97% dataset**) was excluded from the analysis and results were considered significant if the LDA score >2 with p-value<0.05 according to default parameters.

Moreover, I first built a heatmap at Superclade-level, then, at the genus-level, using the normalised abundance of reads by sample over the three years. All data was $\log_2$ transformed. Finally, I calculated the percentage distribution at Superclade- and genus-level based on the average abundance for each month over the three years.

For the exploration at species level, I used ribotypes with 100% similarity to the reference. Each ribotype at 100% similarity was checked for shared V_4 sequence with other species or genera. Ribotypes were annotated for toxicity and symbiontic or parasitic behaviour following DinoREF.

c

	January	February	March	April	May	June	July	August	September	October	November	December
2011	1	15	3	27	11	7 21	19 26	16 30	6 27	25	15	20
2012	17	14	7	3	4	5 19	10 31		7 18	2 23	13	23
2013	28	19	28	16 30	21	4 18	4 16	6 20	10	2 28		4 30

Fig.3.2.1: **a**. Map showing the sampling site – LTER-MareChiara in red (maps downloaded from www.google.com/maps/). **b**. Location of the station on a world map. **c**. HTS sampling dates over three years (2011, 2012 and 2013) and organised by month. Each month is divided in four weeks. Numbers written within red boxes represent the precise sampling dates.

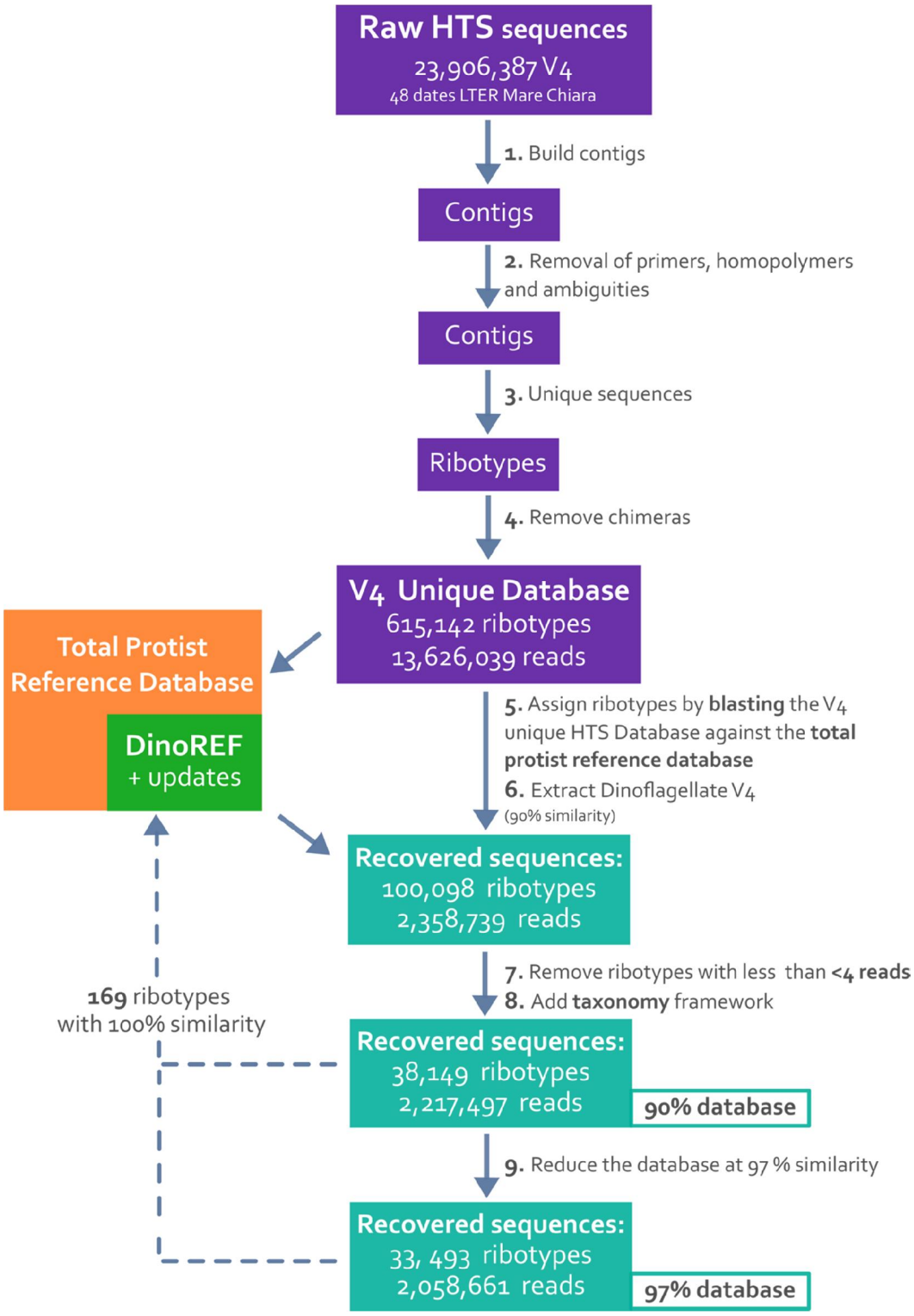


Fig.3.2.2: Flowchart detailing all pre-processing steps of data analysis, from raw HTS data to dinoflagellate V4 database at 90% and 97% similarity. Different steps performed during the analysis are marked from 1 to 9. The first steps of the data analysis (in purple in the flowchart) were performed by Dr Roberta Piredda.

3.3. Results

Dataset description

The 48 samples collected over the three years generated a total of 23 906,387 V4 reads. The raw HTS data was 1: assembled in contigs (**Fig.3.2.2 – step 1**) and filtered following a standard procedure including: 2: removing of primers, homopolymers and contigs showing sequence ambiguities, 3: collapsing unique sequences together (called ribotypes) and 4: removing chimeras (**Fig.3.2.2 – step 2, 3 and 4**). At the end of the filtering procedure performed by Dr. Roberta Piredda, the V4 curated dataset totalised 13,894,744 reads, with 13,040,961 (94%) reads assigned displaying at least 90% similarity to Eukaryota. This dataset was blasted against the protist databases including the new updated database for dinoflagellate DinoREF (See Chapter II of the thesis). 1,756,744 V4 ribotypes (7,201,686 reads) were assigned to a protist reference. (**Fig.3.2.2 – step 5**). 100,098 V4 ribotypes (5.70%) and 2,358,739 reads (30.79%) were assigned to dinoflagellates (Dinophyceae and early branching dinoflagellates) with at least 90% similarity and 70% query coverage (**Fig.3.2.2 – step 6**). Comparatively, 293,044 V4 ribotypes (16.68%) and 1,759,587 reads (24.43%) were assigned to diatom references (Bacillariophyta) with the same criteria (**Table 3.3.1**). In terms of reads, dinoflagellates represented always at least 10% of the total protist community with the exception of 10th of September 2013. The percentage of dinoflagellate ribotypes compared to total protists in each of the samples ranged from 3.79 to 17.33%. The correlation between the number of ribotypes and the number of reads for dinoflagellates (orange) and diatoms (light blue) in the 48 dates showed a different pattern (**Fig.3.3.1**). A linear correlation was detected for both groups, with a high coefficient of determination (R^2=0.78 for dinoflagellate and R^2=0.91 for diatoms). Overall, dinoflagellate ribotypes were represented by a higher number of reads compared to diatoms.

The exclusion of ribotypes represented by less than four reads, reduced the number of dinoflagellate ribotypes from 100,098 to 38,149 (38.11%), while the number of reads were far less effected, from 2,358,739 to 2,217,497 (94% of the dataset, **Fig. 3.2.2 – step 7**). 87.80% (33,493

ribotypes) of the **90% dataset** was assigned to a reference with at least 97% similarity (**Fig. 3.2.2 – step 8**; **Fig. 3.3.2**). On the other hand, 70.67% of the ribotypes had an abundance between 4 and 10 reads, 23.35% had between 10 and 100 reads and only 2.47% had between 100 and 1,000 reads. Very few ribotypes showed an abundance higher than 1,000 reads and only one ribotype more than 100,000 reads (**Fig.3.3.3**). The 196 most abundant ribotypes summed (>0.045%; over 1,000 reads) corresponded to only 0.51% of the total number of ribotypes and represented 1,557,596 reads (70% of the total number of reads – **Table 3.3.2**). Among the 196 ribotypes, 174 (95% of the reads) were assigned within the 97-100% similarity range (**Table 3.3.2**).

This **97% dataset** included 2,058,661 reads. V4 sequences ranged from 272 to 396 bp but 99% of sequences ranged between 375 and 385 bp in length.

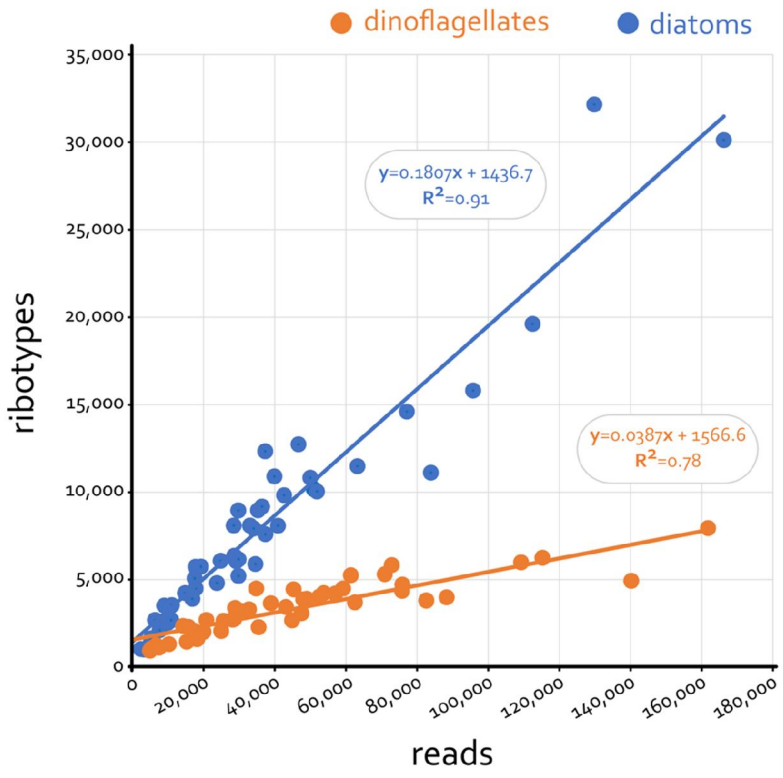


Fig.3.3.1: Scatter plot showing the correlation between the number of ribotypes and the number of reads for dinoflagellates (orange) and diatoms (light blue). Each dot corresponds to a sample (48 dates per 2).

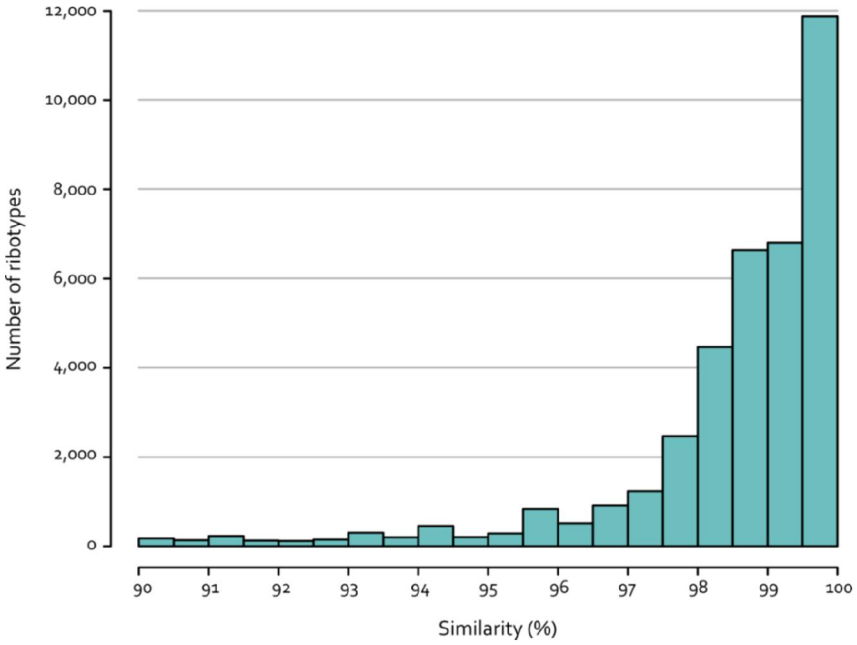


Fig.3.3.2: Histogram showing the distribution of V4 ribotypes obtained according to their similarity with the closest V4 reference sequence. 88% (33,493 ribotypes) of the **90% dataset** was assigned to a reference with at least 97% similarity.

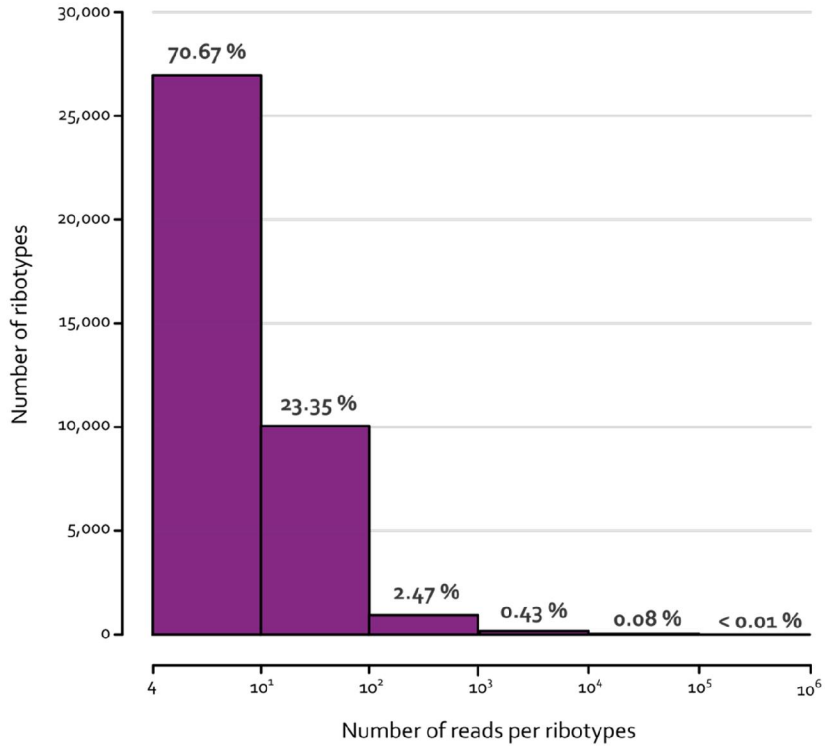


Fig.3.3.3: Histogram showing the distribution of the total number of reads obtained per V4 ribotype. 70,67% of the ribotypes showed an abundance between 4 and 10 reads, 23.35% between 10 and 100 reads, 2.47% between 100 and 1,000 and only 0.43% between 1,000 and 10,000.

Table.3.3.1: Comparison between the number (No.) and percentage (%) of reads and ribotypes between diatoms (Bacillariophyta) and dinoflagellates for the 48 dates organised based on the different clusters obtained in **Fig.3.3.5**. These results were put in perspective with the total number of reads and ribotypes of protists. Shannon-Wiener index was calculated for each sample based on the ribotypes obtained for dinoflagellates. Colours in the table highlight high and low value in each column (from red to blue for the Protist and from red to green for diatoms and dinoflagellates).

Date	cluster	Total Protists		Diatoms				Dinoflagellates				
		No. reads	No. ribotypes	No. reads	% reads	No. ribotypes	% ribotypes	No. reads	% reads	No. ribotypes	% ribotypes	Shannon
20th Dec 2011	Winter	30309	11978	5323	17.56	1648	13.76	5376	17.74	1067	8.91	5.59
8th Mar 2011	Winter	166026	43253	24837	14.96	6074	14.04	28627	17.24	2753	6.36	5.79
14th Feb 2012	Winter	265350	68985	33101	12.47	8089	11.73	57014	21.49	4205	6.09	5.37
17th Jan 2012	Winter	183070	58747	17551	9.59	5073	8.63	49020	26.78	3903	6.64	5.10
23rd Dec 2012	Winter	72408	32975	8987	12.41	3513	10.65	14264	19.70	2347	7.12	5.79
4th Dec 2013	Winter	139593	44225	36468	26.12	9174	20.74	30455	21.82	3202	7.24	5.40
19th Feb 2013	Winter	187402	52740	28506	15.21	8091	15.34	38953	20.79	3650	6.92	5.55
28th Jan 2013	Winter	134016	40204	14838	11.07	4230	10.52	28898	21.56	3377	8.40	5.77
30th Dec 2013	Winter	314473	87317	19291	6.13	5747	6.58	61421	19.53	5246	6.01	5.31
18th Jan 2011	Winter	166814	39719	50019	29.98	10826	27.26	28223	16.92	2708	6.82	5.22
15th Nov 2011	Winter	91406	30865	17781	19.45	5737	18.59	17252	18.87	2033	6.59	5.58
25th Oct 2011	Winter	41407	11814	10078	24.34	2547	21.56	7492	18.09	1113	9.42	5.00
16th Jul 2013		126356	38902	29791	23.58	8952	23.01	52566	41.60	4018	10.33	5.04
28th Mar 2013		274056	59794	42635	15.56	9830	16.44	115261	42.06	6247	10.45	4.98
26th Jul 2011	Spring - mix season	107795	21625	29866	27.71	6160	28.48	44914	41.67	2672	12.36	3.67
19th Jul 2011	Spring - mix season	103486	26488	34112	32.96	7924	29.91	20044	19.37	2009	7.58	4.63
4th May 2012	Spring - mix season	253462	40827	37403	14.76	7598	18.61	140204	55.32	4930	12.07	3.28
3rd Apr 2012	Spring - mix season	283366	54508	16876	5.95	3901	7.16	161750	57.08	7943	14.57	4.44
7th Mar 2012	Spring - mix season	142336	27916	17785	12.49	4496	16.10	82592	58.03	3813	13.66	3.79
13th Nov 2012	Spring - mix season	106593	29296	11140	10.45	3548	12.11	62573	58.70	3698	12.62	4.03
23rd Oct 2012	Spring - mix season	78907	24073	10612	13.45	3429	14.24	32772	41.53	3302	13.72	4.93
6th Aug 2013	Spring - mix season	268680	49514	95767	35.64	15807	31.92	75920	28.26	4712	9.52	4.68
7th Jun 2011	Spring - mix season	118826	21875	4890	4.11	1238	5.66	82633	69.54	3791	17.33	3.51
19th Jun 2012	Spring - mix season	147958	32394	40961	27.68	8077	24.93	59226	40.03	4493	13.87	5.36
5th Jun 2012	Spring - mix season	210220	39861	34597	16.46	5890	14.78	109242	51.97	5991	15.03	5.14
21st Jun 2011	Spring - mix season	155489	28518	29821	19.18	5212	18.28	88310	56.79	3988	13.98	4.57
27th Apr 2011	Spring - mix season	135211	30411	51019	37.73	10169	33.44	48032	35.52	3869	12.72	4.76
15th Feb 2011	Spring - mix season	87408	20412	10955	12.53	2685	13.15	43174	49.39	3437	16.84	4.65
11th May 2011	Spring - mix season	107150	26210	23779	22.19	4800	18.31	25629	23.92	2640	10.07	5.49
30th Apr 2013	Spring - mix season	270391	55713	112470	41.59	19612	35.20	70937	26.23	5310	9.53	5.32
16th Apr 2013	Spring - mix season	188998	43230	77119	40.80	14603	33.78	72911	38.58	5834	13.49	5.45
18th Jun 2013	Spring - mix season	124908	38497	35299	28.26	8965	23.29	45333	36.29	4446	11.55	5.77
4th Jun 2013	Spring - mix season	58103	16244	2405	4.14	1034	6.36	6749	11.62	1246	7.67	5.39
21st May 2013	Spring - mix season	99814	44418	6404	6.42	2686	6.05	34847	34.91	4497	10.12	5.82
2nd Oct 2012	Late summer - autumn	20930	7032	7694	36.76	2108	29.98	4980	23.79	931	13.24	5.21
27th Sep 2011	Late summer - autumn	66158	15943	28667	43.33	6361	39.90	18227	27.55	1596	10.01	4.67
28th Oct 2013	Late summer - autumn	91901	29026	39956	43.48	10901	37.56	20831	22.67	2691	9.27	5.51
2nd Oct 2013	Late summer - autumn	99335	33832	37361	37.61	12339	36.47	15735	15.84	2300	6.80	5.33
7th Sep 2012	Late summer - autumn	62352	14877	10061	16.14	2541	17.08	10333	16.57	1311	8.81	5.01
10th Sep 2013	Late summer - autumn	199413	54914	129818	65.10	32147	58.54	17311	8.68	2082	3.79	5.37
30th Aug 2011	Late summer - autumn	134580	33791	3342	2.48	1001	2.96	18337	13.62	1623	4.80	4.91
18th Sep 2012	Late summer - autumn	121630	26410	29208	24.01	6054	22.92	15312	12.59	1450	5.49	4.61
20th Aug 2013	Late summer - autumn	157748	35714	63256	40.10	11486	32.16	35502	22.50	2281	6.39	4.21
4th Jul 2013	Late summer - autumn	299995	64696	166180	55.39	30117	46.55	53803	17.93	4229	6.54	4.75
6th Sep 2011	Late summer - autumn	132212	34202	46707	35.33	12737	37.24	24998	18.91	2057	6.01	4.20
10th Jul 2012	Late summer - autumn	266670	43987	83897	31.46	11118	25.28	47535	17.82	3075	6.99	4.74
16th Jul 2013	Late summer - autumn	126356	38902	29791	23.58	8952	23.01	52566	41.60	4018	10.33	5.04
31st Jul 2012	Late summer - autumn	228351	52230	51870	22.71	10037	19.22	75819	33.20	4353	8.33	4.85

Table.3.3.2: Distribution of similarity of the most abundant ribotypes, i.e., more than 1,000 reads, in five classes of similarity ranging from 90 to 100%.

Similarity	No. of ribotypes	No. of reads
100 %	53 (27.04%)	868,486 (55.76%)
]100,99] %	54 (27.55%)	279,559 (17.95%)
]99,97] %	67 (34.18%)	337,832 (21.69%)
]97,95] %	11 (5.61%)	39,406 (2.53%)
]95,90] %	11 (5.61%)	32,316 (2.07%)
TOTAL	196 (100%)	1,557,596 (100%)

Ataxonomic patterns

The Shannon-Wiener index varied from 3.28 (4th of May 2012) to 5.82 (21st of May 2013) among the 48 dates (**Fig.3.3.4**; **Table 3.3.1**). No specific seasonal pattern was detected in this analysis, even if most winter samples seemed to show high diversity (>5). The number of ribotypes in a single sample ranged from 931 (2nd of October 2012) to 7,943 (3rd of April 2012). Also, samples from late August, September and October had globally less ribotypes in comparison with other samples. In terms of reads, samples ranged from 4,980 (2nd of October 2012) to 161,750 reads (3rd of April 2012) with an average of 46,197.85 reads (standard error = 35,114.14) and varying regardless of the month, season or year. Only four samples had a low number of reads (<10,000 reads): 25th of October 2011 (7,492 reads), 4th of June 2013 (6,749 reads), 20th of December 2011 (5,376 reads) and 2nd of October 2012 (4,980 reads).

The cluster analysis separated dinoflagellate communities into three seasonal clusters (**Fig.3.3.5**), all of them statistically validated with the ANalysis Of SIMilarities test (ANOSIM, value of R = 0.5464, $p<0.001$). This analysis highlighted two main groups: Group 1 corresponding to Cluster 1 versus Group 2, itself sub-divided in two clusters resulting in Cluster 2 and 3. The first cluster (Cluster 1) gathering samples collected mainly in December, January, February, was called "winter". Cluster 2 contained samples collected in spring (April, May, June), early-summer and a few other samples (15th of February 2011, 23rd of October 2012, 13th November 2012 and 6th of August 2013) while Cluster 3 grouped samples collected in late-summer and autumn (August, September, October). These two last clusters were called "spring-mixed seasons" and "late

summer-autumn" respectively. Two samples 16th of July 2011 and 28th March 2013, were connected with the winter cluster but with high dissimilarity levels (>0.80) (**Fig.3.3.5**). The spring-mixed seasons cluster was the less defined one, not strictly containing spring samples.

In the same way, the clustering analysis performed on the **97% dataset** gave almost the same results (**Fig.3.3.6**) displaying the three clusters "winter", "spring-mix-season" and "late summer-autumn", significantly validated (ANOSIM: R=0.5258, p < 0.001).

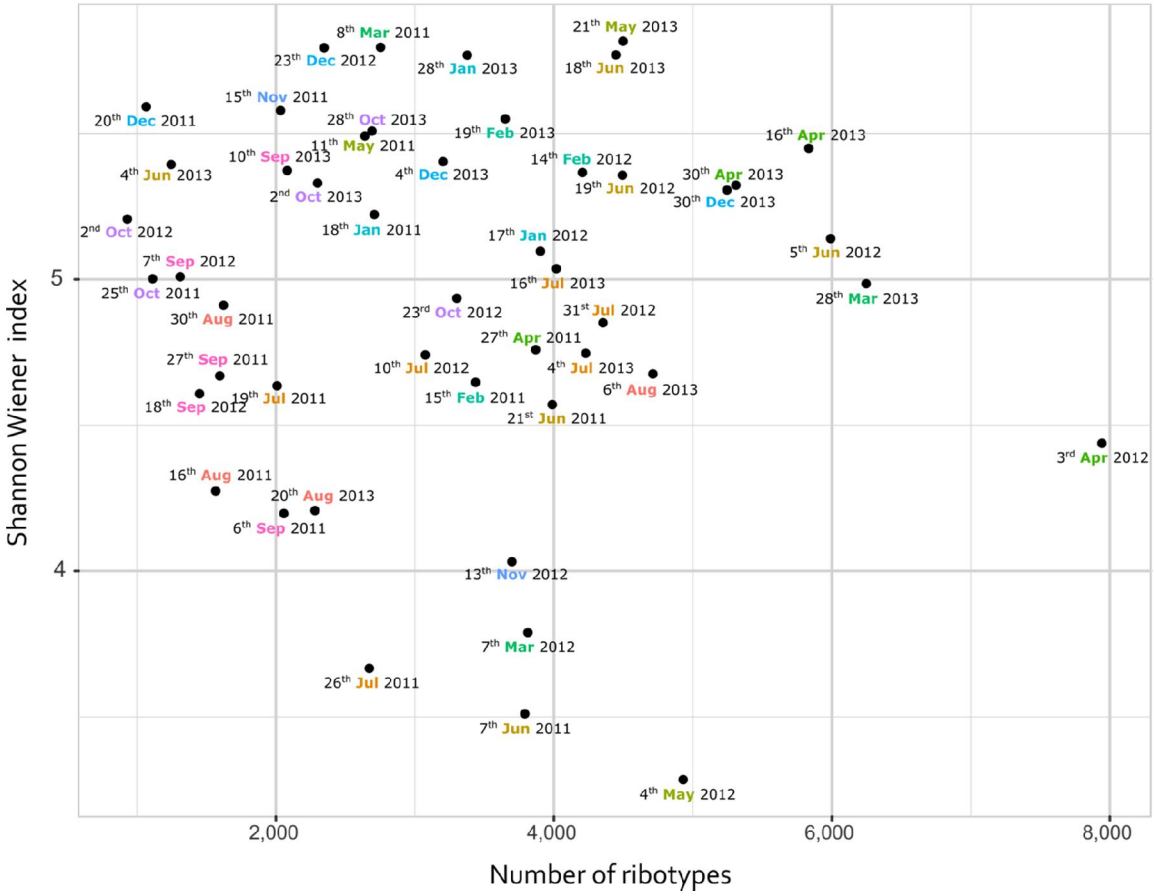


Fig.3.3.4: Scatterplot representing the relation between the number of ribotypes reported and the Shannon-Wiener index calculated for each sample.

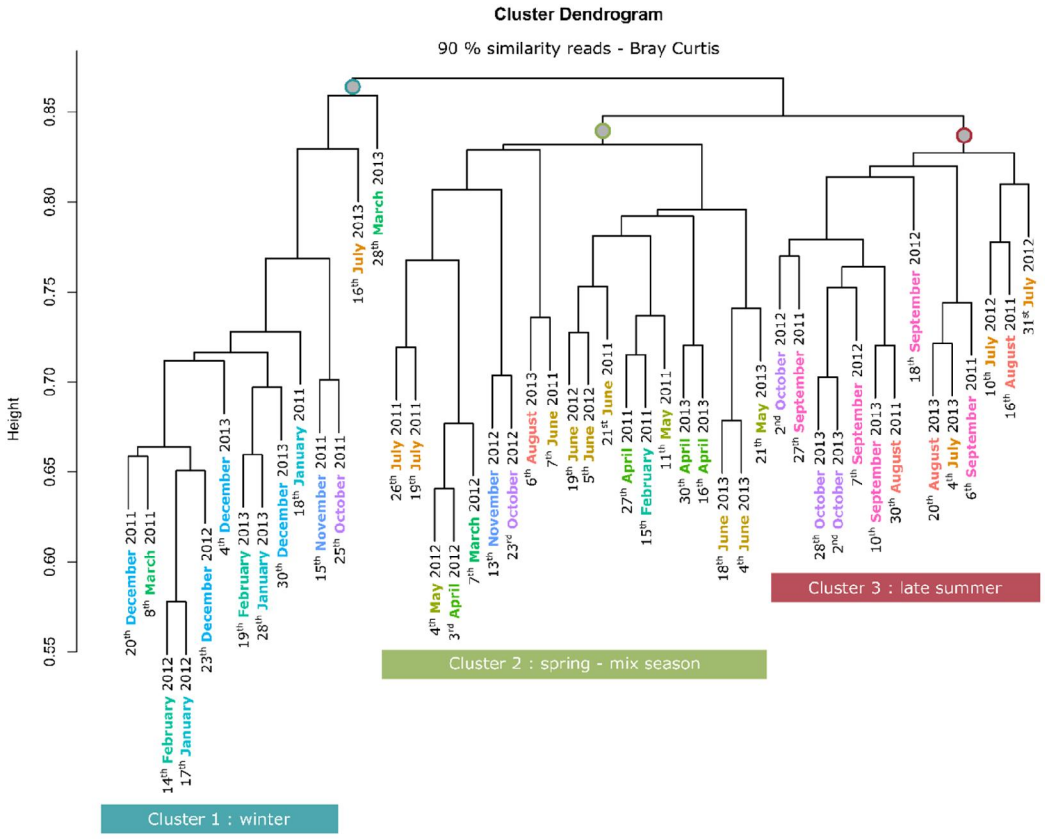


Fig.3.3.5: Hierarchical cluster dendrogram based on Bray-Curtis dissimilarity matrices (ribotypes at 90% similarity). The analysis identified three main clusters ("winter" in blue, "spring-mixed seasons" in green and "late summer-autumn" in red). The grey dots represent the cluster statistically tested with the ANOSIM package on R.

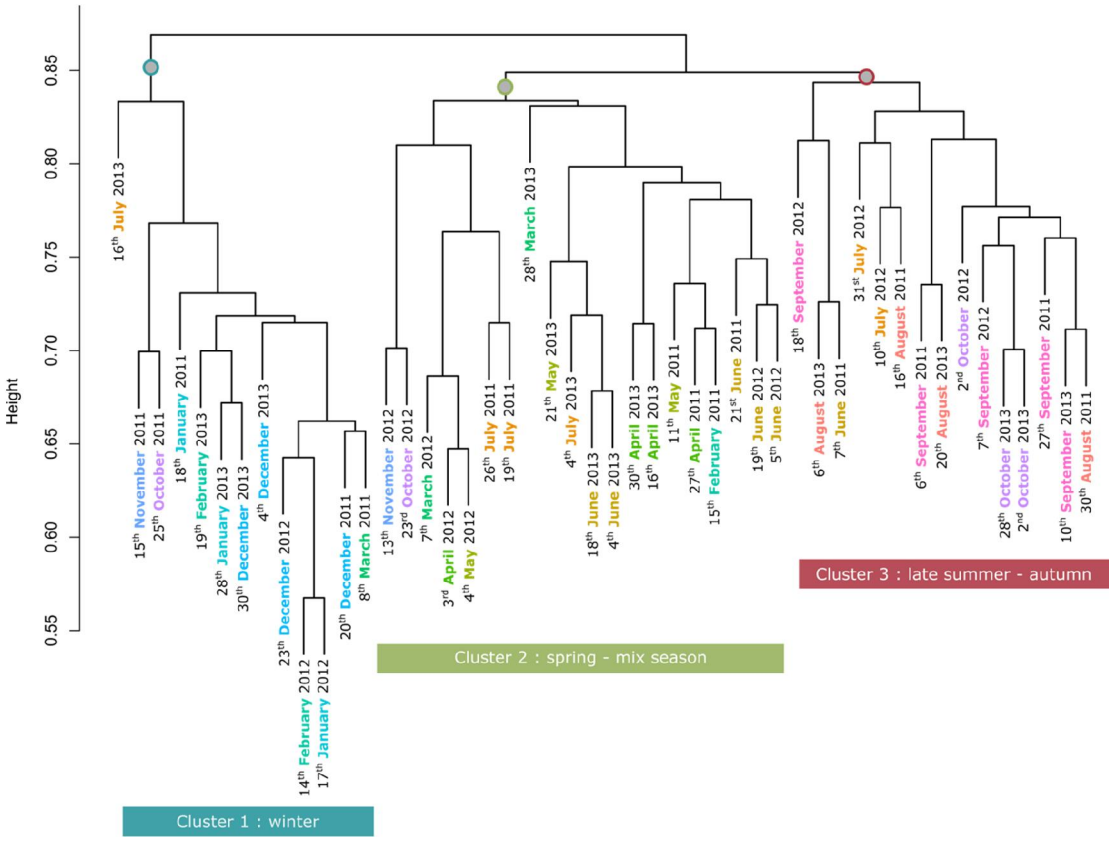


Fig.3.3.6: Hierarchical cluster dendrogram based on Bray-Curtis dissimilarity matrices (ribotypes at 97 % similarity). The grey dots represent the cluster statistically tested with the ANOSIM package on R. The pattern of the 90% dataset was confirmed.

The BIO-ENV analysis revealed that a subset of three environmental parameters (temperature, salinity and chlorophyll *a*) displayed a correlation with dinoflagellate data (**90% dataset**). The CCA analysis confirmed the three clusters observed with the hierarchical analysis (**Fig.3.3.7**). Relationship with environmental parameters suggested that the "winter" samples were linked with lower temperature and lower chlorophyll *a* conditions. The dates sampled in April, May, and June were associated with higher chlorophyll *a* and lower salinity, while dates sampled in "late summer – autumn" seem to be related with higher surface temperature conditions. However, the two main axes explained only 4.98% and 3.89% of the total variance.

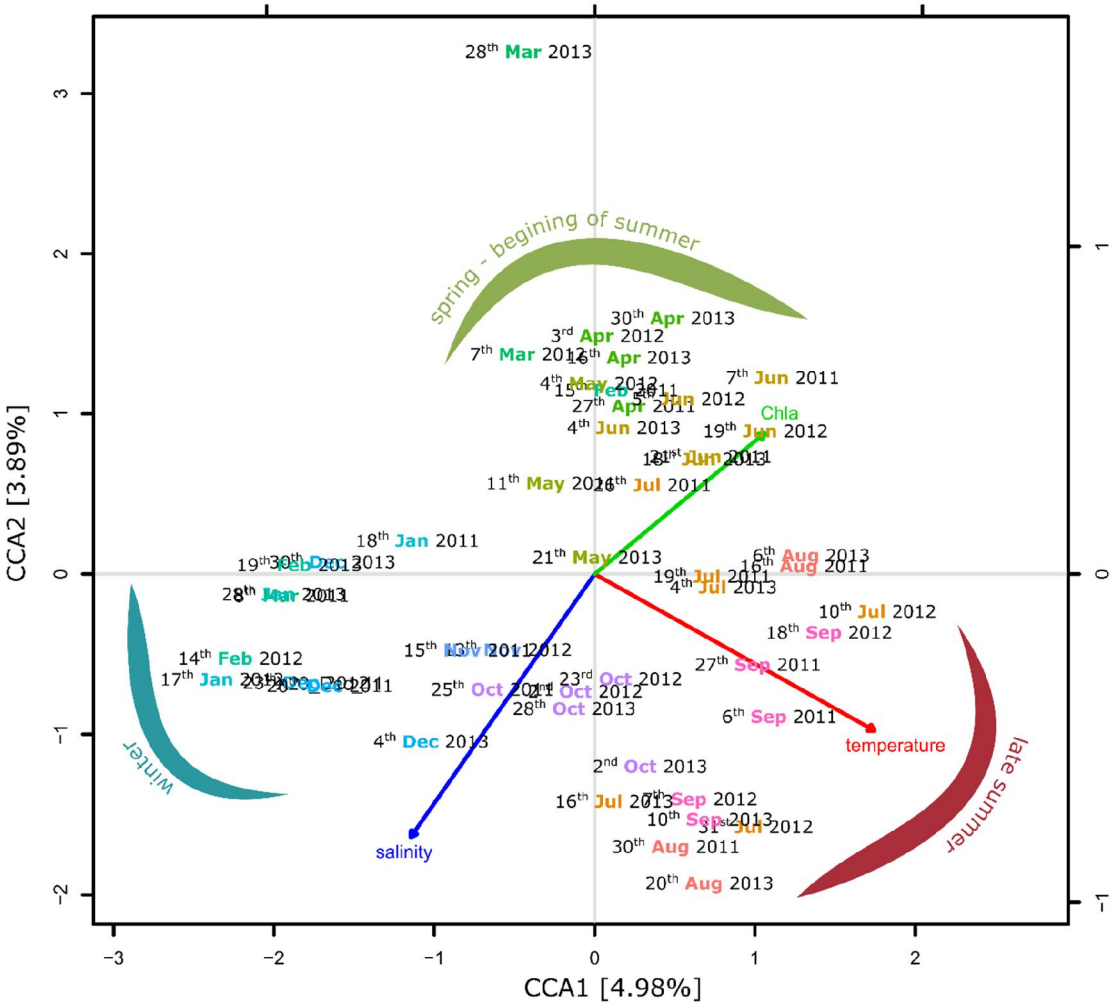


Fig.3.3.7: CCA (Canonical-Correlation Analysis) performed using dinoflagellate ribotypes at 90% similarity and specific environmental parameters selected through the BIO-ENV function (see 3.2. Material and methods section).

Taxonomic patterns

Of the 149 genera in the DinoREF database, I retrieved in the dataset (**90% dataset**) 96 genera belonging to 20 Superclades, Undetermined Thecate Dinophycea (UND), Undetermined Naked Dinophyceae (UTD) or Early branching dinoflagellates (**Fig.3.3.8**).

Boxplots showing the range of similarity values between ribotypes and reference sequences revealed different patterns for different genera (**Fig.3.3.8**). In some cases, as for the genera *Tripos*, *Heterocapsa* or *Akashiwo*, the similarity values were quite high and ranged within a short range (between 100% and 98%) meaning that reads obtained for these genera are similar to reference

sequences. In some cases, the values also varied on a short range but at lower levels of similarity (<95%), like for sequences assigned to *Amylax, Woloszynskia* or *Pseudopfiesteria*, signifying that reads obtained for these genera are quite similar among them but very dissimilar to references and probably belonged to a different genus, still to be discovered or referenced molecularly. In other cases, the range of similarity values was wide (e.g. *Pfiesteria*, *Qia* or *Nussuttodinium*) suggesting that some reads clustered close to a reference but other reads belonged to unknown or unreferenced dinoflagellate taxa. Outlier distribution showed different patterns; as for example, the genera *Gyrodinium, Gymnodinium, Karenia* or *Ptychodiscus* showed regular tail of outlier points ranging for the top of boxplot to 90% similarity, but other genera did not show any outlier.

The treemap of the **97% dataset** at Superclade level (**Fig.3.3.9**) highlighted *Gyrodinium* Superclade (42.51%) as the most abundant Superclade in terms of number reads, followed by Gymnodiniales (20.25%), Gonyaulacales (7.83%), Ptychodiscales (5.20%), Kareniaceae (4.47%) and Suessiales (4.44%) Superclades. The remaining groups, i.e., 13 Superclades + UND and UTD + Early-branching dinoflagellates, shared the last 15% of the reads.

Exploration of the abundance of reads in three seasonal clusters revealed that most of the reads (62% of the total reads) were obtained from the "spring-mixed seasons" cluster containing 20 samples, followed by the "late summer – autumn" cluster (22%, 12 samples) and "winter" cluster (16%, 15 samples) (**Fig.3.3.10**). The treemaps built according to these seasonal clusters showed differences in taxonomic composition at the Superclade level (**Fig.3.3.11**). Gymnodiniales *sensu stricto* dominate in winter with 38,59% of the reads, followed by *Gyrodinium* Superclade (26.88%) and Ptycodiscales (12.77%) (**Fig.3.3.11a**). Cluster 2 (**Fig. 3.3.11b**) and Cluster 3 (**Fig.3.3.11c**) were both dominated by *Gyrodinium*, representing 47.21% and 42.08% of the total reads respectively, which was followed by Gymnodiniales (17.85%) in the "spring-mixed seasons" cluster and Gonyaulacales (15.30%) in the "late summer –autumn" cluster. Other minor groups showed differential patterns among "winter", "spring-mix-season" and "late summer – autumn" clusters. For instance, Suessiales are 257 times more abundant in "spring-mixed seasons" and 97 times in

"late summer – autumn" compared to "winter". Thoracospaheraceae abundance is 335 times more important in "spring-mixed seasons" than in "winter", while Dinophysiales and Torodiniales seem to be characteristic of "late summer – autumn" cluster. In addition, Syndinea and Noctiluciphyceae were present with higher abundance in winter.

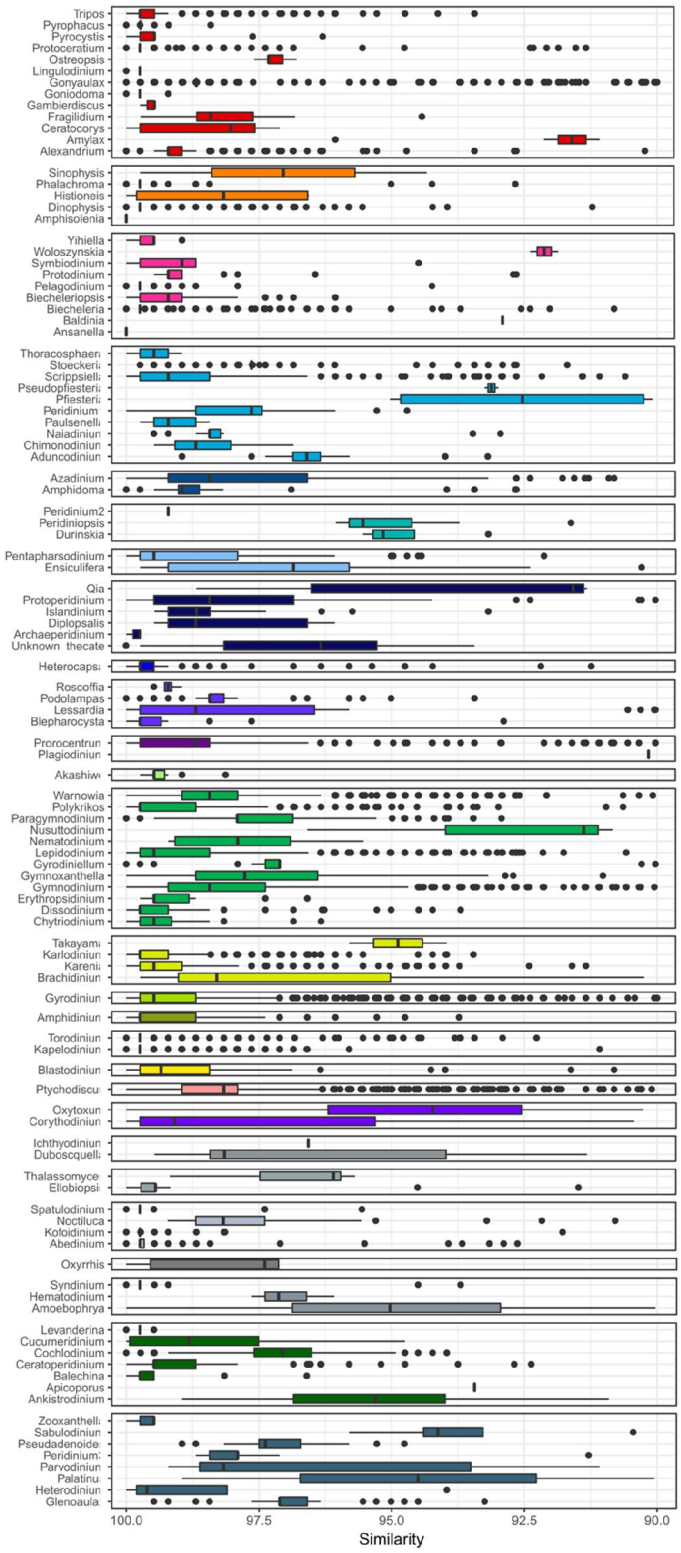


Fig.3.3.8: Similarity of dinoflagellate ribotypes to reference sequences. All ribotypes were clustered based on their taxonomic affiliation and organised under 96 genera. Boxplots show the distribution of the pairwise similarities.

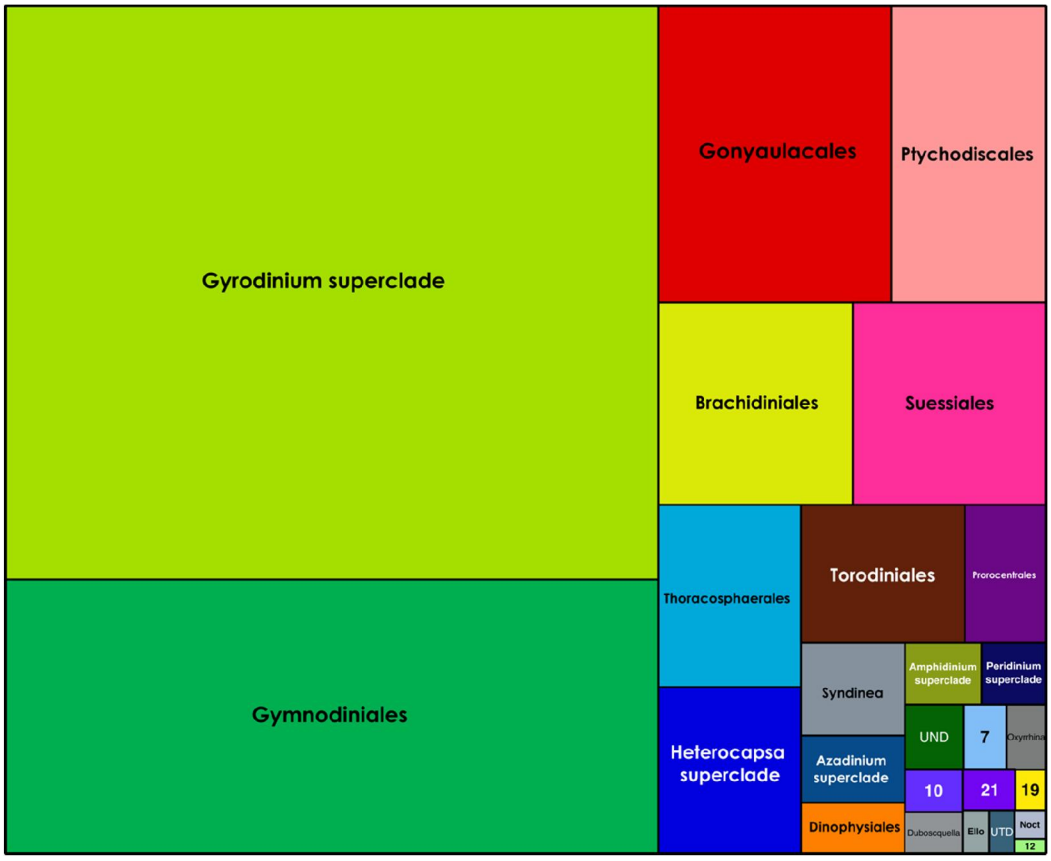


Fig.3.3.9: Treemap representing the proportion of reads assigned to different Superclades inferred from the **97% dataset**. The colours of the Superclades correspond to those used in the DinoREF described in Chapter II of this thesis (submitted article). Smallest boxes have been annotated with the number of corresponding Superclade: **#1**-Gonyaulacales, **#2**-Dinophysiales, **#3**-Suessiales, **#4**-Thoracosphaeraceae, **#5**-Amphidimotaceae, **#6**-Dinotrichales, **#7**-Ensiculifera-Pentapharsodinium Superclade, **#8**-Peridiniales *sensu stricto*, **#9**-Heterocapsaceae, **#10**-Podolampadaceae, **#11**-Prorocentrales, **#12**-genus Akashiwo, **#13**-Gymnodiniales *sensu stricto*, **#14**-Kareniaceae, **#15**-Gyrodinium, **#16**-Amphidinium *sensu stricto*, **#17**-Torodiniales, **#19**-genus Blastodinium, **#20**-Ptychodiscales, **#21**-Oxytoxaceae, **UTD** – "Uncertain Thecate Dinoflagellates", **UND** – "Uncertain Naked Dinoflagellates", **Dub** – "Duboscquella", **Ello** – "Ellobiophyceae", **Noct**-"Noctilucaphyceae", **Oxy** –"Oxyrrhina", **Syn**-"Syndinea".

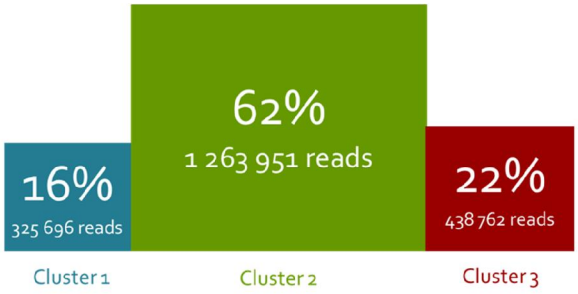


Fig.3.3.10: Distribution of reads over the three clusters identified using the dendrogram analysis and ANOSIM function (Cluster 1: "winter", Cluster 2: "spring-mixed seasons", Cluster 3: "late summer - autumn"). Cluster 1, 2 and 3 were defined in **Fig.3.3.6**.

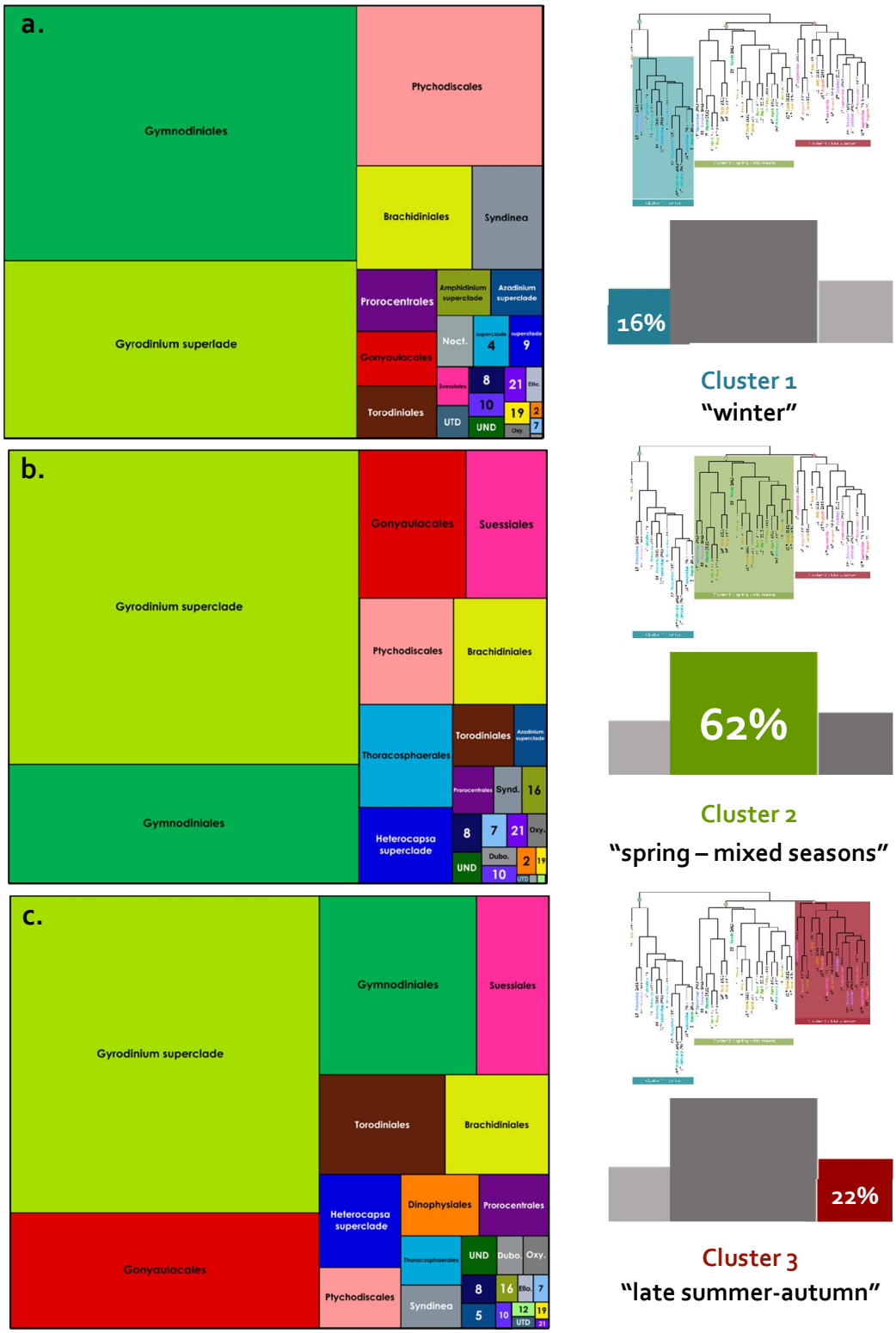


Fig.3.3.11: Treemaps representing the proportion of reads assigned to different Superclades at 97% similarity to three different clusters tested: **a.** Cluster 1: "winter", **b.** Cluster 2: "spring-mixed seasons" and **c.** Cluster 3: late summer-autumn". For each cluster, the proportion of reads on the total is visualised on the right hand side using the same sketches and colour-code as in **Fig.3.3.6** and **Fig.3.3.10**.

The LEFsE approach did not find any significant differential abundance between three categories (Cluster 1: "winter", Cluster 2:" spring – mixed seasons" and Cluster 3: "late summer-autumn"). However, I detected taxa with significant differential abundance between two categories (Cluster 1: "winter" against Cluster 2 and Cluster 3 together). Calculated LDA score showed that 61 taxa with different taxonomic level were detected with significantly higher abundance in winter and 8 taxa for the mixed season group incorporating both Cluster 2 and 3 (**Fig.3.3.12**). As observed with the treemap, Gymnodiniales (**#13**) and a few Gymnodiniales genera such as *Gymnodinium, Chytriodinium* and *Warnowia* were found preferentially in winter. Many parasites like *Amoebophrya, Blastodinium, Hematodinium, Ellobiopsis*, *Chytriodinium*, Thalassomycetales were also detected in significant proportions in winter. Prorocentrales (**#11**), *Ensiculifera-Pentapharsodinium (#7)*, Thoracosphaeraceae (**#4**) and Kareniaceae (**#14**) presented high LDA scores. Mix season group results confirmed the treemap visualisation with *Gyrodinium* Superclade **#15**, Suessiales **#3** and Calciodinellaceae (Thoracosphaeraceae **#4**) statistically reported with higher abundance in spring, summer and autumn.

The variation of abundance of dinoflagellate reads at the Superclade level over three years highlighted very different profiles among different Superclades and among different years (**Fig.3.3.13**). For example, Superclades such as Gonyaulacales (**#1**) or Suessiales (**#3**) seemed to be more represented in spring and summer. Smaller Superclades only containing one or two genera had more defined patterns. For instance, the monospecific genus *Akashiwo* (**#12**) occurred mainly in July, August, September and October; Heterocapsaceae (**#9**) sequences were recovered principally from spring and summer months; while Oxytaceae (**#21**) seemed to be less abundant in summer months. Other Superclades did not show detectable seasonal patterns and/or seemed to occur all year (e.g. Prorocentrales **#11**, Gymnodiniales *sensu stricto* **#13** or Kareniaceae **#14**). *Gyrodinium* genus (**#15**) was the most abundant genus over three years thriving mainly in spring and summer.

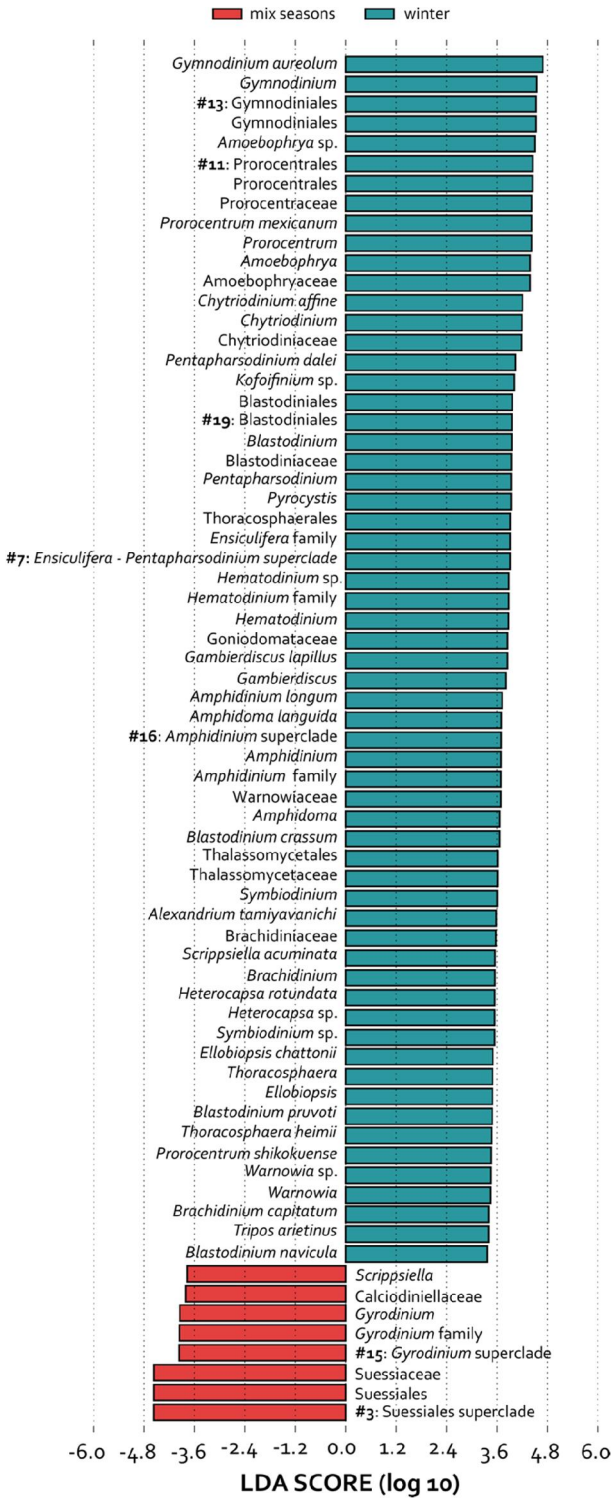


Fig.3.3.12: Superclade, order, family and species of dinoflagellates differentially detected in winter compared to other seasons inferred by linear discriminant analysis coupled with effect size (LEfSe) (LDA>2 and P<0.05). The analysis considers as "winter", dates grouping in the first cluster in the dendrogram analysis (Fig.2). The outlier 16th of July 2013 was excluded from the analysis.

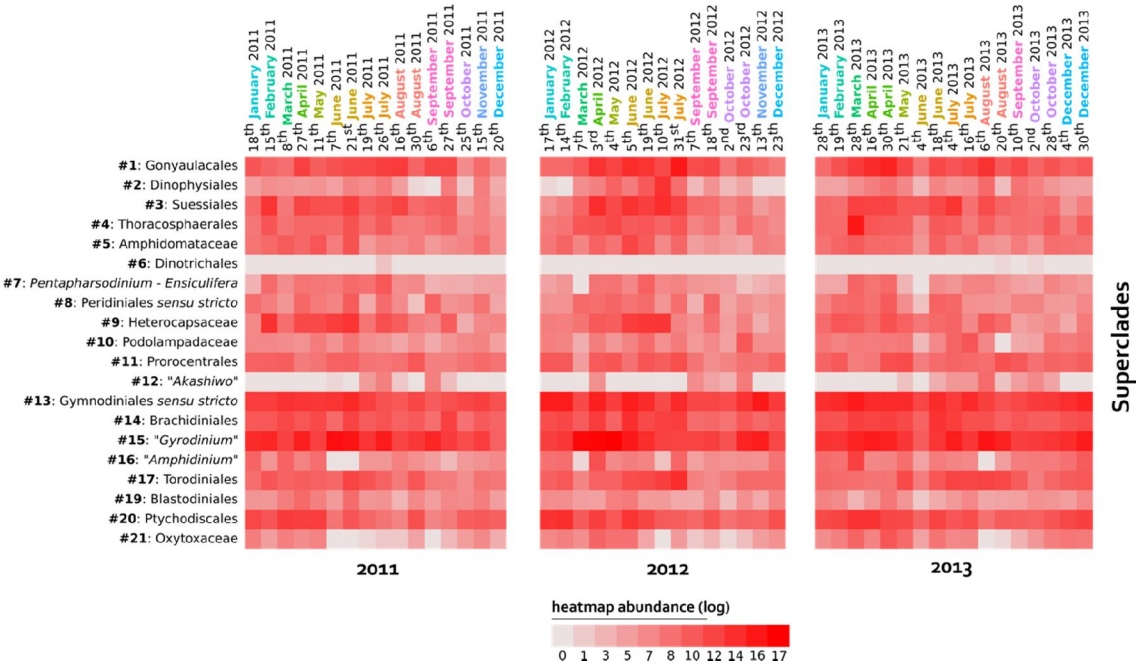


Fig.3.3.13: Heatmap showing the abundance of reads (log_2) for the 48 dates per Superclade.

Variation of abundance at the genus level logically resulted in a more detailed picture of dinoflagellate seasonal variation (**Fig.3.3.14** and **Fig.3.3.15**) and similar trends were observed for Superclades and genera when data was analysed with a percentage perspective regardless of the abundance of reads (**Fig.3.3.16**; **Fig.3.3.17**; **Fig.3.3.18**). Clear seasonal patterns were detected for many genera including *Alexandrium* (spring-summer), *Fragilidium* (summer), *Ostreopsis* (summer), *Pyrophacus* (late summer), *Biecheleria* (spring-summer), *Symbiodinium* (winter), *Akashiwo* (late summer), *Nematodinium* (winter), *Brachidinium* (winter) or *Hematodinium* (winter). Other genera presented a multimodal distribution (e.g. *Protoceratium*, *Heterocapsa, Erythropsidinium*) or seemed to occur all year (e.g. *Tripos*, *Prorocentrum, Gymnodinium, Gyrodinium, Ptychodiscus*). For some genera higher abundance detected for a single month were the result of an isolated bloom event (e.g. *Dinophysis* spp. on the 10th of July 2012, *Gonyaulax fragilis* on the 31st of July 2012, *Stoeckeria* sp. on the 28th of March 2013).

Overall, taxonomic groups with the most abundant ribotypes were **#13** Gymnodiniales *sensu stricto* (64), **#15** *Gyrodinium* genus (23), Early branching dinoflagellate (16, in grey colour in figures), **#14** Kareniaceae (15), **#1** Gonyaulacales (14) and **#3** Suessiales (13).

Fig.3.3.14: Heatmap showing the seasonal distribution of different genera. Relative abundance of reads (log_2) was calculated by normalising values by samples using the **97% dataset**. When reads where absent in samples the heatmap pixel was left white.

Fig.3.3.15: Heatmap showing the seasonal distribution of different genera. Relative abundance of reads (log_2) was calculated by normalising values by samples using the **97% dataset**. When reads where absent in samples the heatmap pixel was left white.

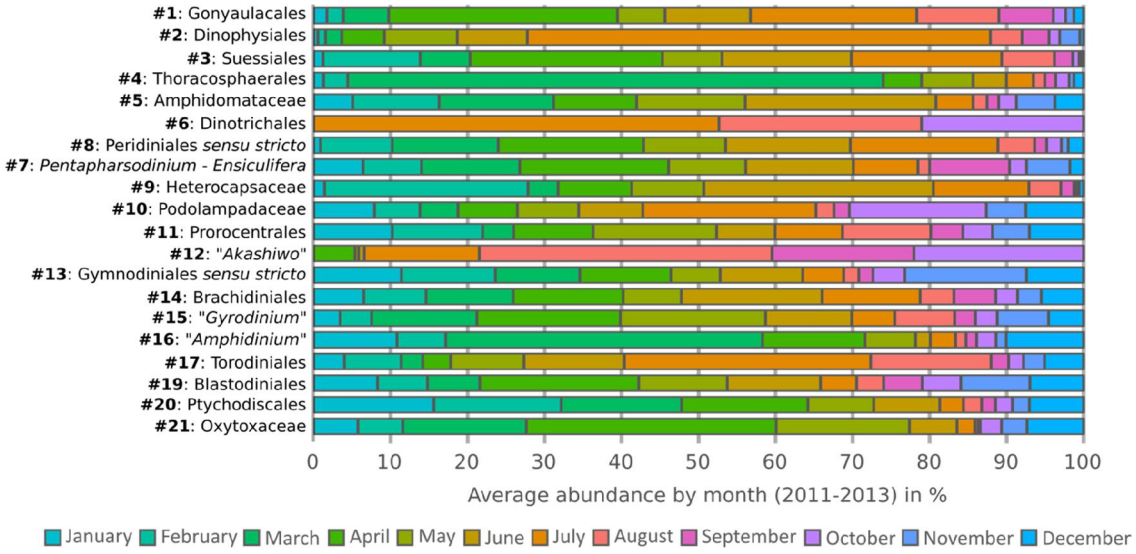


Fig.3.3.16: Percentage distribution of Superclades across the annual cycle. Since different months have different number of samples, the value for each month is the average of the contribution of each sample in that month.

Exploration at species level found 169 ribotypes at 100% similarity to the reference (**Table 3.3.3**). According to DinoREF database (Chapter II of this thesis, **Table 2.3.3**), 24 ribotypes could not be assigned unambiguously because they blasted against a V4 reference sequence with multiple possible assignation. For instance, seven different *Dinophysis* species shared the same V4 ribotype.

In addition, one abundant ribotype was assigned to a V4 reference sequence shared by the two genera *Karlodinium* and *Takayama*. For some genera (*Pelagodinium, Symbiodinium, Scrippsiella, Heterocapsa, Warnowia*, etc…) references were not assigned at species level (i.e. sp.)

Remarkably, among the 169 ribotypes, at least **85** (50.30%) were new dinoflagellate taxonomic records for the LTER MareChiara. **26** ribotypes corresponded to "toxic" species in accordance with the IOC-UNESCO Taxonomic Reference List of Harmful Micro Algae (Moestrup *et al.*, 2009), **10** ribotypes to symbiont species, and **20** ribotypes to parasite species.

In quantitative point of view, of the ten most abundant ribotypes, four were assigned to *Gyrodinium* species (*Gyrodinium spirale* (AB120002) - 365,850 reads; *Gyrodinium spirale* (KP790157)

– 88,819 reads; *Gyrodinium heterogrammum* – 42,836 reads and *Gyrodinium gutrula* – 21,565 reads). Other abundant ribotypes included *Lepidodinium viride* (or *chlorophorum* – 40,614 reads - **#13**), *Kapelodinium vestificii* (26,884 reads - **#17**), *Protoceratium reticulatum* (23,746 reads - **#1** - potentially toxic), *Heterocapsa* sp. (JX661031 – 20,850 reads – **#9** - symbiont), *Biecheleria brevisulcata* (19,260 reads - **#3**) and *Biecheleria tirezensis* (18,989 reads - **#3**).

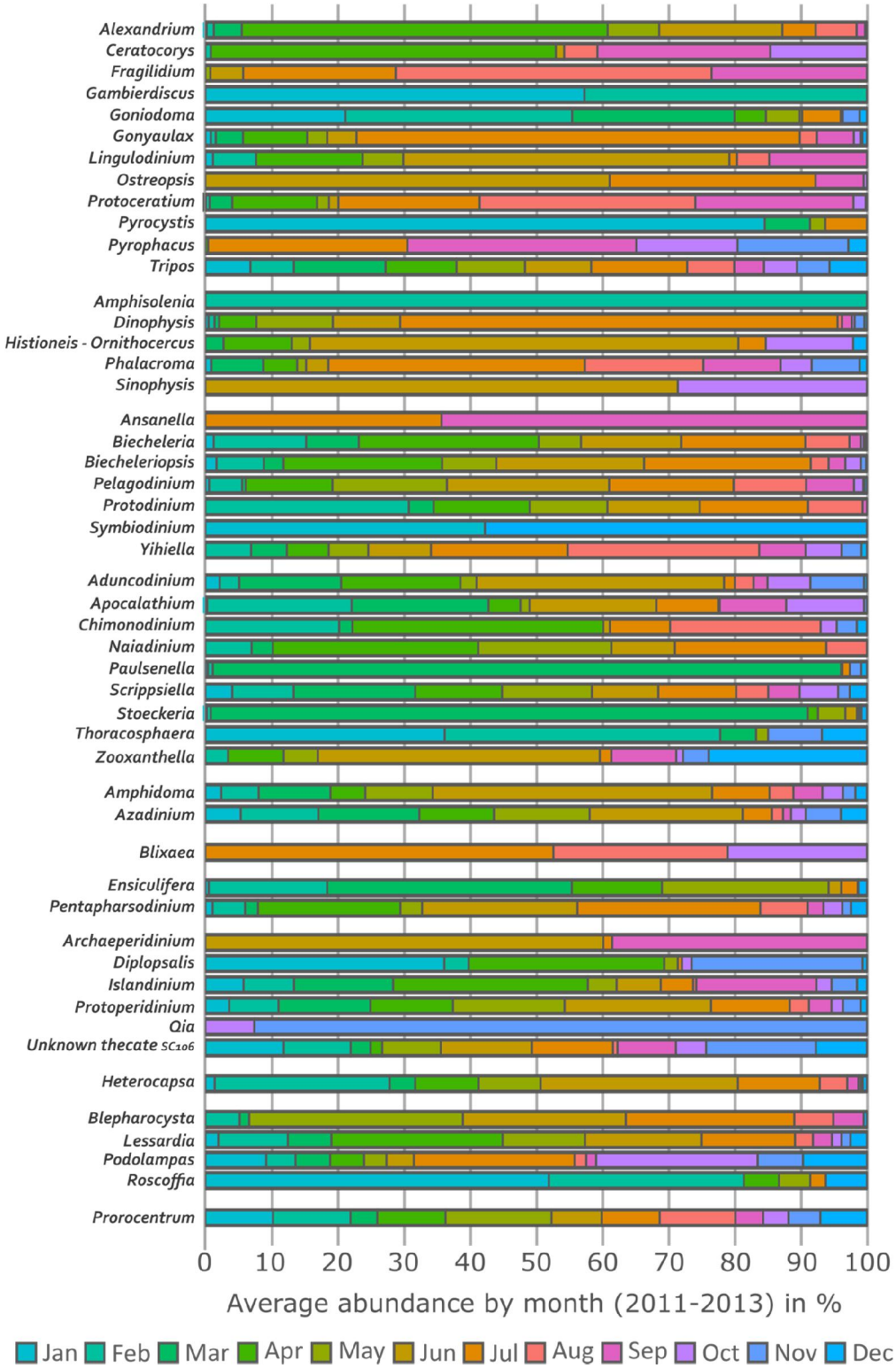


Fig.3.3.17: Percentage distribution of genera across the annual cycle. Since different months have different number of samples, the value for each month is the average of the contribution of each sample in that month. First panel.

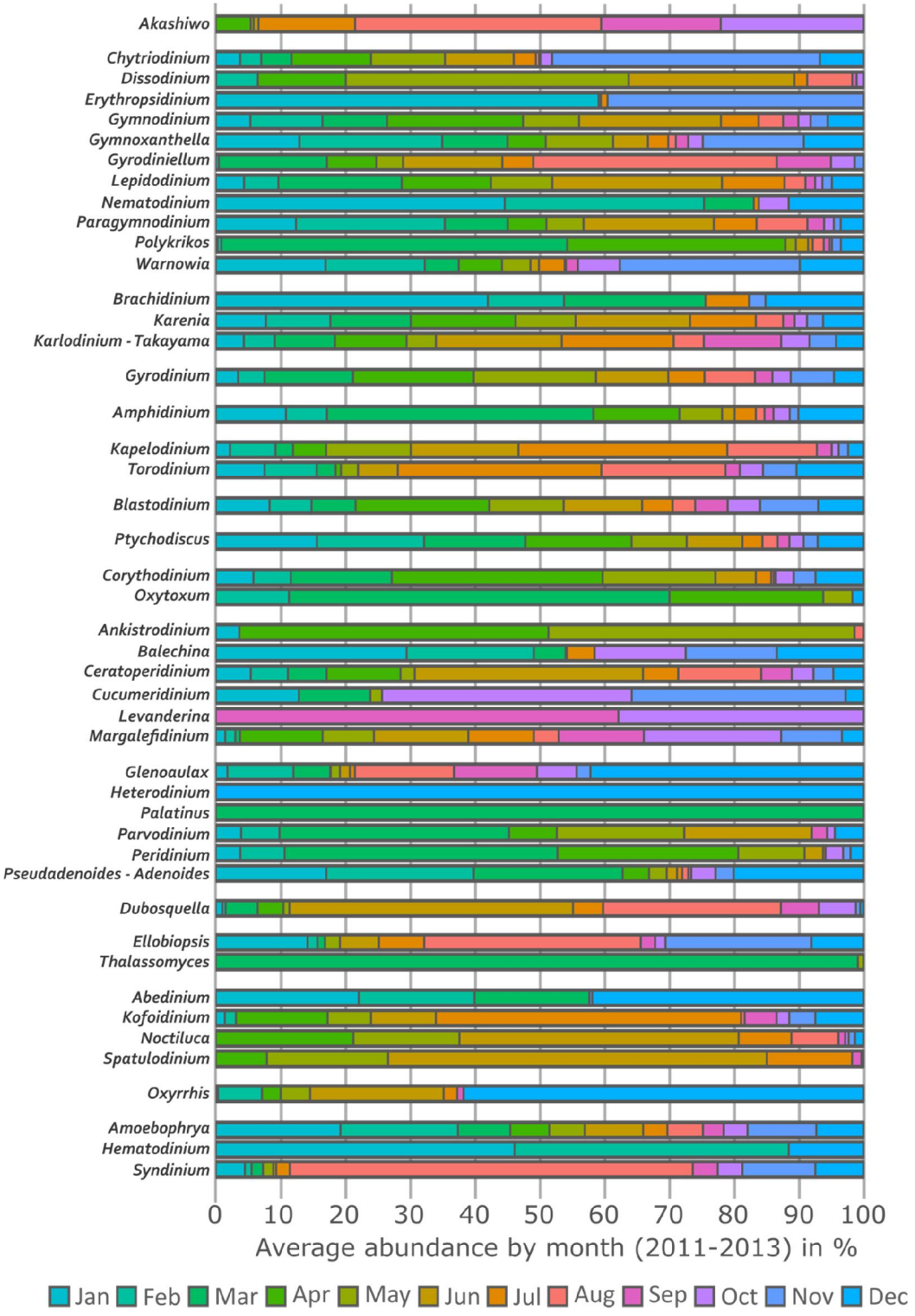


Fig.3.3.18: Percentage distribution of genera across the annual cycle. Since different months have different number of samples, the value for each month is the average of the contribution of each sample in that month. Second panel.

Table.3.3.3: List of species detected with 100% similarity to a reference in LTER-MareChiara over three years sampling (2011, 2012 and 2013). References and Genbank accession numbers are listed in the DinoREF database (Chapter II). Reference sequences produced in this thesis from single cell or culture extraction, amplification and Sanger sequencing are listed in **Appendix 4** (in bold in the table). When multiple species or genera shared the same V4 sequence, this information was specified in the column "same V4 sequence". Read abundance for each ribotype was given in the last column. Ribotypes were classified by Superclade together in the same way than the DinoREF database. New dinoflagellate taxonomic records for LTER MareChiara monitoring stationwere marked with an asterisk (*). Potentially toxic taxa were annotated with a **T**, symbiont and parasites references with a **S** and **P** respectively; following DinoREF annotations.

	Ref. code	Reference species name	same V4 sequence	reads
# 1: GONYAULACALES	AB693196	*Lingulodinium polyedrum*		415
	FJ402956	*Tripos arietinus*	*Tripos longipes*, T. euarcuatus* and *T. symmetricus*	212
	FJ402954	*Tripos azoricus*	*Tripos petersii**	151
	FJ402955	*Tripos candelabrus*		334
	FJ402950	*Tripos concilians*		26
	FJ402959	*Tripos contrarius*		247
	FJ402949	*Tripos declinatus*	*Tripos pentagonus*	1,691
	FJ402957	*Tripos extensus*		241
	FJ402966	*Tripos furca*		10,714
	FJ402958	*Tripos fusus*		2,726
	FJ402943	*Tripos hexacanthus*		280
	FJ402942	*Tripos massiliensis*		353
	FJ402965	*Tripos paradoxides*	*Tripos limulus, T. minutus* and *T. kofoidii*	308
	FJ824911	*Tripos platycornis**	*Tripos horridus*	100
	JF521616	*Alexandrium affine*		310
	KF925334	*Alexandrium andersonii* **T**		363
	LC056070	*Alexandrium pseudogonyaulax* **T**		1,089
	KF908797	*Alexandrium mediterraneum Group II*		58
	JF521631	*Alexandrium minutum* **T**	*Alexandrium andersonii* **T** and *A. insuetum*	84
	JF906998	*Alexandrium minutum* **T**		705
	KJ362003	*Alexandrium ostenfeldii** **T**	*Alexandrium andersonii* **T**	45
	KF908799	*Alexandrium tamarense Group III** **T**		31
	AB088325	*Alexandrium tamiyavanichi** **T**	*Alexandrium cohorticula**	5
	KM886380	*Goniodoma polyedricum*		348
	DQ388465	*Gonyaulax cochlea**		42
	LC036590	*Gonyaulax ellegaardiae**	*Gonyaulax spinifera* **T**	1,247
	FR865625	*Gonyaulax spinifera* **T**		1,785
	DQ388456	*Ceratocorys horrida*		126
	AB727654	*Protoceratium reticulatum* **T**		23,746
	FR865629	*Pyrocystis lunula*		6
	SC80	*Pyrocystis pseudonoctiluca**		38
	AY443024	*Pyrophacus steinii*		460

	Ref. code	Reference species name	same V4 sequence	reads
# 2: DINOPHYSIALES	GU196149	*Amphisolenia bidentata*	*Amphisolenia schauinslandii*, Amphisolenia globifera*	18
	KJ508017	*Dinophysis acuminata* T	*Dinophys norvegica** T, *D. tripos* T, *D. caudata* T, *D. acuta* T, *D. infundibulum* T and *D. fortii* T	8,105
	HM853810	*Dinophysis monacantha**		8
	HM853804	*Histioneis longicollis*	*Histioneis gubernans, H. cymbalaria*, Ornithocercus quadratus, O. magnificus* and *O. heteroporus*	61
	HM853781	*Phalacroma doryphorum*		269
	AB551248	*Phalacroma mitra* T		4
	JQ996385	*Phalacroma oxytoxoides*	*Phalacroma rotundatum* T	1,416
	HM853792	*Phalacroma porodictyum*		69
	HM853784	*Phalacroma rotundatum* T		29
# 3: SUESSIALES	HG529978	*Ansanella granifera**		10
	LC068842	*Biecheleria brevisulcata**		19,260
	FR690459	*Biecheleria cincta*		3,562
	NaD22	***Biecheleria* sp.***		8,785
	KF463288	*Biecheleria tirezensis**		18,989
	HG792067	*Biecheleriopsis adriatica**		3,384
	NaD26	*Biecheleriopsis adriatica**		1,968
	KF422623	*Pelagodinium beii**		299
	U41087	*Pelagodinium beii** **S**		3,085
	JX661026	*Pelagodinium* sp.* **S**		377
	JX661027	*Pelagodinium* sp.* **S**		2,628
	JX661028	*Pelagodinium* sp.* **S**		920
	AB863030	*Symbiodinium* sp.* **S**		41
	LN898222	*Yihella yeosuense**		92
#4: THORACOSPHAERACEA	KF446621	*Apocalathium aciculiferum**	*Scrippsiella hangoei**	145
	KR535602	*Scrippsiella acuminata*		4,357
	AB183671	*Scrippsiella acuminata*		676
	DQ847435	*Scrippsiella precaria*		694
	AM494499	*Scrippsiella* sp.		35
	HM483396	*Scrippsiella* sp.		153
	NaD25	***Scripsiella* sp. *NaD25****		2,146
	LC054944	*Thoracosphaera heimii*		31
#5: genus AZADINIUM	KR362881	*Amphidoma languida** T		29
	KJ481826	*Azadinium concinnum**	*Azadinium caudatum*	2,729
	KJ481822	*Azadinium cuneatum**		44
	KR362890	*Azadinium dexteroporum* T		788
	GQ914935	*Azadinium obesum**		208
	JX559886	*Azadinium polongum**		65
	HQ324899	*Azadinium poporum** T		9
	KJ481817	*Azadinium trinitatum**	*Azadinium spinosum** T	2,211
#7	JX262492	*Pentapharsodinium dalei*		5
	AF274270	*Pentapharsodinium* sp.*		2,596

	Ref. code	Reference species name	same V4 sequence	reads
#8: PERIDINIALES	AB564309	*Archaeperidinium minutum*		12
	SC106	*SC106 Unknown thecate 18S**		36
	AB716911	*Protoperidinium americanum**		86
	AB716913	*Protoperidinium monovelum*		6
	AY443022	*Protoperidinium pellucidum*		427
	AB181904	*Protoperidinium punctulatum**		173
#9: HETEROCAPSACE	EF492492	*Heterocapsa niei*		2,873
	LC189145	*Heterocapsa* sp.*		8
	LC054932	*Heterocapsa circularisquama** T	*Heterocapsa niei*	99
	AF274266	*Heterocapsa pygmaea**		1,605
	DQ388464	*Heterocapsa rotundata*		2,292
	FJ549370	*Heterocapsa* sp.* **S**		5
	JX661031	*Heterocapsa* sp.* **S**		20,850
	JX661033	*Heterocapsa* sp.* **S**		5
#10	FJ888593	*Blepharocysta* sp		654
	AF521100	*Lessardia elongata*		794
	FJ888597	*Podolampas spinifera*	*Podolampas palmipes* and *P. bipes*	476
#11	AB781324	*Prorocentrum shikokuense*	*Prorocentrum donghaiense** and *P. dentatum*	202
	EF492512	*Prorocentrum triestinum*		6,342
	KY426837	*Prorocentrum mexicanum** T	*Prorocentrum rhathymum* T, *P. texanum** and *P. micans*	691
#13: GYMNODINIALES *sensu stricto*	FJ473380	*Chytriodinium affine** **P**		679
	FJ663049	*Chytriodinium roseum** **P**		7
	FJ473378	*Dissodinium pseudolunula** **P**		2,319
	KR362891	*Gymnodinium aureolum*		2,252
	DQ779992	*Gymnodinium impudicum*		22
	KP790152	*Gymnodinium litoralis**		1,371
	AB265963	*Gymnodinium nolleri*		9
	HG005135	*Gymnodinium smaydae**		1,820
	JQ639761	*Gymnodinium* sp.		11
	FR720082	*Gyrodiniellum shiwhaense**		135
	AB686254	*Lepidodinium chlorophorum**	*Lepidodinium viride*	8,019
	AB686255	*Lepidodinium* sp.		40,614
	AM408889	*Paragymnodinium shiwhaense**		23
	AY421789	*Polykrikos hartmannii* T		58
	AB466294	*Polykrikos kofoidii*		12,319
	DQ371292	*Polykrikos kofoidii*		11
	AB466288	*Polykrikos schwartzii*		94
	AB920349	*Gymnoxanthella radiolariae** **S**		15
	FJ467492	*Warnowia* sp.		6,680
	KP790169	*Warnowia* sp.	*Proterythropsis* sp.*	10,627
	KP790170	*Warnowia* sp.		1,738
#14: KARENIACEAE	HM067002	*Karenia bicuneiformis**		79
	HM067005	*Karenia papilionacea* T		7,965
	HM067007	*Karenia selliformis** T		89
	FN357291	*Karlodinium* sp.		656
	AF274262	*Karlodinium veneficum* T		270
	KU314866	*Karenia mikimotoi* T		10,581
	NAD46	***Karlodinium decipiens****		5,502
	KU314867	*Karlodinium veneficum* T	*Takayama pulchellum** and *T. acrotrocha*	12,113

	Ref. code	Reference species name	same V4 sequence	reads
#15: genus GYRODINIUM	FN669510	*Gyrodinium dominans**		352
	FN669511	*Gyrodinium gutrula**		21,565
	FJ024299	*Gyrodinium helveticum**		13
	KP790159	*Gyrodinium heterogrammum**		42,836
	AB120003	*Gyrodinium rubrum*		340
	AB120001	*Gyrodinium spirale*		169
	AB120002	*Gyrodinium spirale*		365,850
	KP790157	*Gyrodinium spirale*		88,819
#16	AF274254	*Amphidinium longum**		6,056
	EU046336	*Amphidinium* sp.		9
#17	KP790162	*Kapelodinium vestifici**		26,884
	KR139784	*Torodinium robustum*	*Torodinium teredo**	12,940
#19: BLASTODINIACEAE	JX473667	*Blastodinium contortum** **P**		340
	FJ541188	*Blastodinium galatheanum** **P**		23
	JN257674	*Blastodinium mangini** **P**		394
	JX473656	*Blastodinium mangini** **P**		40
	JX473659	*Blastodinium mangini** **P**		43
	JX473662	*Blastodinium navicula** **P**		94
	JX473665	*Blastodinium navicula** **P**		572
	HQ226071	*Blastodinium spinulosum** **P**		66
	JN257671	*Blastodinium spinulosum** **P**		5
	JX473663	*Blastodinium spinulosum** **P**		16
#20	KU640194	*Ptychodiscus noctiluca**		8,617
#21	KY421383	*Corythodinium cristatum**		334
	KY421378	*Corythodinium tessellatum*		2,770
	KY421375	*Oxytoxum scolopax*		5
UND	KR139792	*Balechina pachydermata**		287
	KR139786	*Cucumeridinium coeruleum**		42
	KR139787	*Cucumeridinium lira**		8
	KP790150	*Ceratoperidinium falcatum*		7
	KP790151	*Ceratoperidinium falcatum*		221
	KJ561350	*Cochlodinium polykrikoides* T		40
	DQ388457	*Levanderina fissa**		212
UTD	JQ446581	*Heterodinium doma*		11
	U52357	*Zooxanthella nutricula** **S**		13
	GU355680	*Kofoidinium pavillardii**		2,175
	GU355681	*Kofoidinium* sp.		23
	GU355682	*Spatulodinium* sp.*		596
	GU355679	*Abedinium dasypus**		1,864
	AY566418	*Oxyrrhis marina*		246
	AY775284	*Amoebophrya ceratii** **P**		618
	AY775285	*Amoebophrya ceratii** **P**		464
	KF791347	*Amoebophrya ceratii** **P**		620
	AY208893	*Amoebophrya* sp.* **P**		128
	JX173253	*Amoebophrya* sp.* **P**		394
	DQ146406	*Syndinium* sp.* **P**		2,009
	FJ593708	*Ellobiopsis chattonii** **P**		837

3.4. Discussion

This study focuses on dinoflagellate diversity at different taxonomic levels, their abundances and seasonal changes in the Mediterranean Sea using a HTS metabarcoding approach. For this purpose, I used a 48-dates dataset covering three years of sampling (2011, 2012 and 2013) of planktonic communities at the Long-Term station MareChiara (Gulf of Naples, Tyrrhenian Sea, **Fig.3.2.1**). I used the DinoREF dinoflagellate database (**Chapter II**) to extract and annotate dinoflagellate reads from the marine protist dataset.

Overall, the 100,098 unique V4 ribotypes assigned at ≥90% similarity to dinoflagellate reference sequences represent 30.79% of the total number of protist reads. Yet they represent only 5.70% of the number of ribotypes (**Table 3.3.1**). In comparison, Bacillariophyceae, which constitute a dominant component of planktonic communities at LTER-MC, represent 24.43% of the total number of reads (protists), but a much higher proportion in terms of ribotypes (16.68%). This result is in agreement with other studies where a large number of reads could be assigned to dinoflagellates (de Vargas *et al.*, 2015; Massana *et al.*, 2015; Piredda *et al.*, 2017). The result could be explained in several ways:

1. Dinoflagellates are less diverse but each individual species is more abundant than diatoms or than the total protistan community in general;
2. A few dinoflagellate species dominate samples at abundances much higher than diatoms do;
3. Dinoflagellates may dominate especially in periods when plankton densities are low. Therefore, even though their cell number is small in absolute terms they dominate plankton samples.
4. On average, dinoflagellate cells contain a far larger number of 18S rRNA copies than diatoms;

5. Dinoflagellates show less intraspecific and intra-individual ribotype variation than diatoms;

Quantitative data based on light microscopy observations shows that dinoflagellates rarely dominate plankton assemblages at the LTER-MareChiara (data not shown – courtesy Diana Sarno) but also at various marine monitoring site in the world suggesting a possible overestimation with HTS data of dinoflagellate abundance and invalidating hypotheses 1 and 2. Moreover, counts from LTER-MC and HTS results in this study also clearly show that dinoflagellate have a lower abundance in winter season following the trend of phytoplankton abundance suggesting that hypothesis 3 should not be considered.

The disproportionally high number of reads for dinoflagellates is probably related to the high copy number in ribosomal genes reported for dinoflagellates (hypothesis 4), which is usually higher than other protists (Godhe *et al.*, 2008). This tendency is also confirmed by the higher values of reads corresponding to low number of ribotypes compared with diatoms which support the multi-copy hypothesis (**Fig.3.3.1)**.

Several authors try to explain this high copy number with different hypothesis linking this bias with cell size, length, bio-volume or genome size (Zhang, Bhattacharya and Lin, 2005; Godhe *et al.*, 2008; Hou and Lin, 2009; de Vargas *et al.*, 2015; Murray *et al.*, 2016).

The multi-copy hypothesis is also supported by the fact that when ribotypes with less than four reads were removed the number of ribotypes dropped by 62 % and reads by 6 % (**Fig.3.2.2 –Step 7**). The dataset (**90% dataset**) is mainly composed (70.62%) of ribotypes with low abundance (between 4 and 10 reads; **Fig.3.3.2**) and most of these ribotypes (88% of the ribotypes) in this dataset were assigned to a reference with at least 97% similarity **(Table.3.3.2**). These ribotypes with high similarity to a reference but represented by only few reads could mainly be the result of intragenomic 18S rRNA polymorphism of dinoflagellate genes reported for other protists (Alverson and Kolnick, 2005; Decelle *et al.*, 2014) and some dinoflagellates (Thornhill, Lajeunesse

and Santos, 2007; Miranda *et al.*, 2012). Intragenomic variation has also been reported for diatoms (Alverson and Kolnick, 2005) but dinoflagellate usually possess a much bigger and highly redundant genome with a large gene copy number (Hou and Lin, 2009; Murray *et al.*, 2016) making hypothesis 5 unlikely.

Part of this variation is probably also due to sequencing errors but this effect was limited by removing ribotypes less than four reads. If real, this could mean that there is a confusion between intragenomic diversity, intergenomic diversity and sequencing errors. This confusion could make it hard to detect rare taxa.

Yet the number of ribotypes assigned to dinoflagellates show an important diversity exceeding by far the number of described dinoflagellate species. The majority of them (88%) are assigned with high level of similarity (>97%), which allow an annotation with a reasonable confidence at least to order, family or genus level. Among the 196 most abundant ribotypes (**Table.3.3.2**), 53 were identified at 100% similarity and 57 between 99 and 100% similarity showing that the most abundant ribotypes thriving in the Gulf of Naples can be identified at the species or genus level. However, 67 of the most abundant ribotypes are assigned with 97-99% similarity suggesting that dinoflagellate diversity remains to be characterised. Only a minority (5.61%) of the most abundant ribotypes are assigned to a reference with less than 97% similarity signifying that few abundant ribotypes belong to unknown groups of dinoflagellates or to known taxa missing molecular characterisation.

Seasonal pattern

A strong seasonal signal was found for dinoflagellates in the surface of the water column confirming the seasonal pattern reported for the total protist communities at LTER-MareChiara using a single year (Piredda *et al.*, 2017), but also in other subtropical and temperate regions such as the Gulf of Mexico (Brannock *et al.*, 2016) or the English Channel (Genitsaris *et al.*, 2015).

Among the three seasonal clusters, the "winter" (Cluster 1), contains the lower number of reads but also the lower number of cells counted (Diana Sarno, personal communication; Ribera d'Alcalà *et al.*, 2004). Cluster 1 mainly groups samples characterised by low temperature, high salinity (**Appendix 2 and 3**), shorter photoperiod and high mixing in a complete homogeneous water column as described for winter season by Carrada and colleagues (1980) for the Gulf of Naples. The taxonomic characterisation of the clusters with LEfSe analysis highlighted a strong presence of particular taxa under winter conditions. Some groups of parasites such as *Amoebophrya, Chytriodinium, Blastodinium, Hematodinium, Ellobiopsis* or the ocelloid bearing dinoflagellate groups Warnowiaceae seem to be prevalent in the winter season. Copepod parasites like *Blastodinium* and *Ellobiopsis* were previously reported mainly in winter e.g. off Barcelona (Skovgaard and Saiz, 2006). Analysis of alpha diversity highlights high Shannon-Wiener values in all "winter" samples (Cluster 1) suggesting a seasonal cycle also for alpha diversity. Shannon-Wiener diversity can also be high in other seasons indicating that community can demonstrate similar levels of complexity under different conditions.

The "spring-mixed seasons" (Cluster 2) contains a large number of the reads (1,263,951 reads, **Fig.3.3.10**). The cluster contains heterogeneous seasonal samples, which may result from strong and rapid changes in surface layer conditions as a result of variable weather and hydrographic conditions (as also described for bacteria in San Pedro station, SPOT, Southern California, subtropical region, (Kim *et al.*, 2014). This cluster is often related to a higher relative abundance of dinoflagellates with respect to the total number of protists and very different Shannon-Wiener values (ranging from 3.28 to 5.82), coinciding with the typical seasonal succession observed in spring when the dinoflagellate bloom follows the diatom bloom (Sherr and Sherr, 2007; D'Alelio *et al.*, 2015; Bunse and Pinhassi, 2017).

The absence of any significant differential abundance between three categories in the LEfSe approach (Cluster 1:"winter", Cluster 2:"spring – mixed seasons" and Cluster 3:"late summer-autumn") confirms the variable conditions within this cluster (Cluster 2). Few taxa, such as

Gyrodinium, *Scrippsiella* and Suessiales, are characteristic of the mixed season conditions (Cluster 2 and 3) that are mainly defined by the absence or strong reduction of winter taxa.

The "late summer-autumn" (Cluster 3) represents stable and warm conditions with stratification of the water column occurring in the Gulf of Naples usually in summer (July-September) (Carrada *et al.*, 1980). In this period the abundance of dinoflagellate reads is not very high compared to total protists, even if specific taxa are probably adapted to these particular conditions as for example for the genera *Protoceratium, Akashiwo* or *Margalefidinium* (**Fig.3.3.14** and **Fig.3.3.15**). Shannon-Wiener values ranging between 4.20 and 5.51 do not show any particular trend.

Patterns observed in the treemaps and relative abundances of different taxa seem to highlight two different ecological states. The data suggests a shift from spring-summer autotrophic toward a heterotrophic-mixotrophic late-summer state. Small autotrophic thecate taxa (e.g. *Alexandrium, Biecheleria, Heterocapsa*, *Scrippsiella*) are mainly dominant in spring and beginning of summer when they co-occur with different diatom species blooming in that period. Then, in summer and late summer, large heterotrophic taxa (e.g. *Kapelodinium, Torodinium, Levanderina*) are more abundant, as previously reported at MareChiara station (D'Alelio *et al.*, 2015).

Some Superclades or genera are present all year and do not show any seasonal pattern. This is due to the fact that different species belonging to the same genus (e.g. *Tripos* see **Chapter IV**, *Prorocentrum, Karenia* or *Protoperidinium*) succeed or overlap throughout the year, thriving under different conditions.

Temperature, salinity and chlorophyll *a* were identified as the main parameters by correspondence analysis associated with dinoflagellate succession as previously detected for total protist community in the Gulf of Naples (Piredda *et al.*, 2017). These three parameters were identified as the main factors shaping seasons (**Appendix 2**) in a subtropical "Mediterranean climate".

The CCA plot confirms the clustering obtained in the hierarchical analysis. Winter samples are related to lower temperature and lower chlorophyll *a.* Samples from "late summer –autumn" were

linked with higher surface temperature values. CCA2 axis clearly separates samples associated with high autotrophic biomass peaks from samples with lower chlorophyll a. This result indicates that seasonal trends including the cycling variation of several physico-chemical parameters shape plankton communities over the year. However, these three proxies explained only a small percentage (4.98 for CCA1 and 3.89 for CCA2) of the total variance. In comparison, when considering the total protist community the same parameters explained 45% of the variance for year 2011 (Piredda *et al.,* 2017). These results suggest that dinoflagellate communities are mainly driven by biotic factors, such as life-cycle traits and interspecific relationships, rather than by abiotic parameters (at least not the ones measured).

Interpretation of the results (Superclade, order or genus level) is difficult since very few studies report seasonal patterns for dinoflagellate communities at finer taxonomical level. Most studies on phytoplankton temporal patterns focused on inter-groups dynamic (i.e. diatoms versus dinoflagellates) rather than functional groups based on taxa characteristics, life strategies and habitat preferences (Barton *et al.*, 2013).

Dinoflagellates present a large range of adaptive and survival strategies and different taxa seem to dominate in specific conditions (Smayda and Reynolds, 2003). Due to their diversity of life strategies, bigger size and complex shape, most dinoflagellates are less efficient in obtaining nutrients, grow more slowly and therefore supposedly less competitive than other protists such as diatoms or other small flagellates (Smayda, 1997). Unlike diatoms, which present a much more regular seasonal species succession (Ribera d'Alcalà *et al.*, 2004; Aubry *et al.*, 2012), dinoflagellates are seen as favoured by specific conditions, often stressful such as nutrient or physical perturbations (i.e turbulence, salinity, reduced light)(Smayda and Reynolds, 2003). Most dinoflagellates are motile and have the capacity to move in the water column to avoid disturbance or predation. Other dinoflagellates are known to be tolerant of specific conditions such as heavy rain (*Prorocentrum cordatum* syn. *P. minimum* – Garrido *et al.*, 2016), abnormal high nutrients (Heisler *et al.*, 2008), or low turbulence (Margalef, Estrada and Blasco, 1979; Smayda, 2002). This

tendency can be observed in our results (**Fig.3.3.14** and **Fig.3.3.15**): many dinoflagellate taxa do not show a regular cyclic pattern over the three years analysed and some ribotypes corresponding to a single species are detected in few events over the 48 dates sampled (e.g. *Ansanella* (**#3**), *Qia (**#8**), Erythropsidinium* **(#13**)).

In contrast, an increasing number of papers have described specific seasonal patterns for some dinoflagellates species easily recognisable with light microscopy (Ribera d'Alcalà *et al.*, 2004; Aubry *et al.*, 2012). Nonetheless, specific growth conditions or temporal niches are unknown for the majority of dinoflagellate species and many authors are currently suggesting a reexaminaton of the classical Margalef mandala for dinoflagellates (Smayda and Reynolds, 2003; Klais *et al.*, 2011) for which only some specific blooming "red tide" species were considered. These "red tide" and often harmful blooms of dinoflagellates rarely coincide in space or in time with the spring-bloom where diatoms and dinoflagellates co-occur successfully (Kremp, Tamminen and Spilling, 2008; Klais *et al.*, 2011). Yet in our results, most reads assigned to dinoflagellates are recovered predominantly in spring-early summer (Cluster 2; **Fig.3.3.10**) and a strong overall seasonal pattern is obtained (**Fig.3.3.5** and **Fig.3.3.6**) showing an important contribution of dinoflagellates to the planktonic community during this period. Despite these facts, the ecological literature rarely addresses dinoflagellate succession and seasonal variation, especially when dinoflagellates are abundant. This is mainly due to methodological limits of counts by light microscopy where different species cannot be distinguished accurately. Therefore, acquisition of molecular data is crucial to address this issue.

Species detected in MareChiara

Of the 169 ribotypes assigned at 100% similarity to a reference, 85 (50.30%) were new taxonomic records for the Gulf of Naples. There are a number of reasons why these taxa were not previously detected at LTER-MC. Routine screening of the diversity is done in light microscopy (LM). Thus, parasites and symbionts living inside other organisms are hard to detect using this method (e.g. *Blastodinium* spp., *Amoebophrya ceratii*, *Chrytriodinium* spp., *Pelagodinium beii*, *Symbiodinium* sp.,

Gymnoxanthella radiolariae, Heterocapsa sp.). The same is true for small species (<15 µm) (e.g. *Biecheleria brevisculcata, B. tirezensis, Azadinium obesum*, *A. spinosum, Heterocapsa pygmaea*) and species that have only small morphological differences (*Dinophysis monacantha, Protoperidinium americanum, P. punctulatum, Gymnodinium smaydae, G. litoralis*, *Corythodinium cristatum).* As an example, I detected the presence of *Ptychodiscus noctiluca*, which has never been reported at the LTER-MC despite its large size and possibly high abundance (8,617 reads). Naked cells of *P. noctiluca* usually are destroyed by sampling and fixation procedures (Remsen, Hopkins and Samson, 2004), or if observed can be confused easily with species of the family Kareniaceae, with which they share many characteristics (Gómez *et al.*, 2016). Finally, genera and species recently described such as *Ansanella granifera, Yihiella yeoense, Gyrodiniellum shiwhense, Balechina pachydermata, Cucumeridinium* spp., were attributed previously to other species or undetermined groups before their formal description.

Some taxonomic entries were assigned at 100% similarity thanks to new reference sequences produced during this thesis from cultures or single cell isolations (e.g. *Pyrocystis noctiluca S80, Biecheleriopsis adriatica NaD26, Scrippsiella sp. NaD25*, see **Appendix 3**).

Of the 26 potentially toxic dinoflagellate species observed in the metabarcode data of the LTER MC at least three are new taxonomic entries, i.e. *Alexandrium tamarense, Amphidoma languida* and *A. poporum*. The presence of *Karenia selliformis*, hypothesised by Zingone *et al.* (2006, as *K.* cf. *selliformis*) has been confirmed by molecular data. Similar biodiversity studies in different geographic regions have demonstrated the effectiveness of the metabarcode approach for detecting and monitoring toxic species (Kohli *et al.*, 2014; Grzebyk *et al.*, 2017; Smith *et al.*, 2017).

Potential biases and limitations of the approach

When using a HTS metabarcoding approach to study protist diversity, one needs to be aware of this method's limitations (Collins and Cruickshank, 2013). Interpretation of the data hinges on: i) availability of comprehensive reference databases, ii) the potential biases such as amplification

and sequencing errors, iii) taxonomy-related differences in the multi-copy numbers of the target sequence, iv) the capacity of the marker to capture the diversity and v) the capacity of the primers to amplify the target groups.

Reference database

This analysis relied of the DinoREF database. At the LTER-MC, the majority of ribotypes (88%) were assigned with high level of similarity (>97%) and more specifically than 95% of the most abundant ribotypes (**Table 3.3.2**; 70% of the reads) are assigned with a similarity higher than 97% suggesting few unknown or unreferenced lineages of dinoflagellates. However, our results also reveal that many species and genera still need to be characterised by their 18S sequences inside families or order with already described members (**Fig. 3.3.8**), since only 169 ribotypes out of the 33,493 ribotypes (**97% dataset**) present a perfect match (100 % similarity) with a reference from DinoREF (**Fig. 3.2.2**). Moreover, reference sequences obtained during this thesis from the Gulf of Naples (**Appendix 4**) are highly valuable to annotate the HTS dataset at a specific geographical site.

Copy number of target genes

In our dinoflagellate dataset, the principal bias is probably the high copy number of fundamental functional genes as previously discussed. This is most likely related to their body size and by extension to the size of dinoflagellate genomes. This characteristic leads the overestimation of dinoflagellate importance in most protist metabarcoding studies. From a taxonomic perspective, the phenomenon is well illustrated in our dataset by some dinoflagellates taxa such as *Gyrodinium*, Gymnodininiales, Gonyaulacales, Ptychodiscales and Kareniaceae which always present a high number of reads, even if these groups include species which are not always particularly big for dinoflagellates (>40 microns). In particular, one ribotype attributed to *Gyrodinium spirale* has an abundance of 365,850 reads which suggest a large number of ribosomal rRNA gene copies for this big (> 40µm) heterotrophic dinoflagellate. The same bias has been reported before for the same genus by other authors (Massana *et al.*, 2015; Piredda *et al.*, 2017).

Primers capacity to capture diversity and different lineages of dinoflagellates

Comparatively, other groups of dinoflagellates seem to be less represented in the Gulf of Naples and could be underestimated if their copy number of ribosomal genes is low or if V4 primers do not amplify these groups effciently. For example, *Protoperidinium* species are among the most abundant dinoflagellates counted at the MareChiara station but, the number of reads recovered with the HTS metabarcoding approach is low compared to other groups.

It is known that number of gene copies and genome size can vary greatly between different orders, genera and species of dinoflagellates (Lajeunesse, Andersen and Galbraith, 2005; Hou and Lin, 2009). However, this information is missing for the majority of dinoflagellates and particularly for heterotrophic ones, which usually do not grow in culture. In the same way, primer design is mainly based on references, which are also largely biased toward autotrophic species. In addition, a recent study performed on dinoflagellate mock communities showed differential detection of dinoflagellate species for the V4 marker as well as between different markers (Smith *et al.,* 2017). In this study, authors found a much better detection and higher number of reads of Gonyaulacales species (*Fukuyoa paulensis*, *Alexandrium ostenfeldii, Gonyaulax* sp., *Ostreopsis* cf. *siamensis*, *Coolia malayensis*) compared with species belonging to other Superclade such as *Symbiodinium* sp., *Prorocentrum micans, Amphidinium thermaeum, Vulcanodinium rugosum and Gymnodinium catenatum.* Comparatively, I observed the same patterns in our natural samples for Gonyaulacales species detected in MareChiara samples, genera such as *Symbiodinium* and Amphidinium being much less represented.

Barcode resolution power

Finally, the resolution of the V4 marker is known not to be optimal for dinoflagellates as this region does not always discriminate between the different species or genera (DinoREF database; **Chapter II**; Mordret *et al.,* 2018). This lack of resolution hampers the discrimination between different species, including toxic *versus* non-toxic taxa, or even genera. Our dataset perfectly illustrates this problem. For 169 ribotypes assigned at 100% of similarity (**Table 3.3.3**), 24 ribotypes were assigned

to a reference sharing a V4 between different genera or species. The V4 regions does not allow discrimination between some species of *Karlodinium* and *Takayama*, although these dinoflagellates are observed in samples in light microscopy and represent an important number of reads (12,113 reads for KU314867). Additionally, in eight cases, I could not discriminate between a non-toxic and a toxic species respectively (*Alexandrium tamiyvanichi vs A. cohorticula; Gonyaulax ellegaardiae vs G. spinifera; Phalacroma oxytoxoides vs. P. rotundatum; Azadinium trinitatum vs. A. spinosum, Heterocapsa circularisquama vs H. niei*). More variable barcode regions are necessary for metabarcoding studies focused on dinoflagellate diversity. Possible marker includes 28S rRNA or longer barcodes following innovation in metagenomics technologies which would allow comparison of phylogenies with barcodes reads and make easier the detection of new dinoflagellate lineages (Grzebyk *et al.*, 2017).

Sample choice

Another limitation of the metabarcoding approach can be biological. Indeed, samples selected for this study each represent the biological state of a single time and were not chosen randomly but based on specific events such as maximum-minimum of plankton abundance, chlorophyll a or particular bloom conditions (**Fig. 3.2.1c**). It is then normal to find biological outliers in the analysis. For the dinoflagellate community, two outliers were detected in the cluster and in the CCA analysis: the 28th of March 2013 and the 16th of July 2013. The 28th of March 2013 a bloom of an unknown dinoflagellate assigned to the Thoracosphaeraceae blasting against the best reference (*Stoeckeria algicida* at 97,9 % similarity) was detected. At this date, 30% of the reads were assigned to the Superclade Thoracosphaeraceae, which is quite unusual. This bloom was associated with high concentration of nutrients (Nitrates, Nitrites, Phosphates and Silicates – **Appendix 2**) suggesting a punctual abnormal nutrient input from land. Bloom conditions can create a bias toward blooming species overestimating dominant taxa compared to the rarer ones (Piredda *et al.*, 2017). In contrast, 16th of July 2013 does not represent a dinoflagellate bloom condition but is characterised by high number *Warnowia* sp. (3 bp mismatch) which are usually more abundant in winter. However, these cases could be related to the particular geographical position of

MareChiara station where the opposite influences of coastal and off-shore waters are able to generate rapid change of microbial communities. These conditions were previously described for summer season in D'Alelio *et al.* (2015) and defined as "blue" and "green" conditions. Thus, natural perturbations can also generate "outlier communities" and are an important source driving high diversity in the site.

Concluding remarks

In conclusion, the temporal signals of dinoflagellate communities corroborated the trend previously reported for protist communities at LTER-MareChiara over the samples taken in a single year (Piredda *et al.*, 2017; Piredda unpublished). This is not surprising because dinoflagellate taxa cover all the functional traits (from autotrophs to parasites) with characters and complexity similar to the protists as a whole but at a smaller scale. Winter is characterised by parasites and very specific genera of dinoflagellates. The other seasons highlighted a shift from photosynthetic conditions with dinoflagellate displaying the same trend than Bacillariophyta to late summer when heterotrophic conditions seem to be prevalent. The metabarcoding approach allowed us to identify some dinoflagellates occurring at LTER-MC when references exist and proved the importance of quality databases such as DinoREF to analyse the results of these studies. Nonetheless, this study demonstrated that a large part of dinoflagellate diversity remains to be described or characterised molecularly even at a well-studied site such as MareChiara. Our results also highlight a certain number of pitfalls unique to dinoflagellates such as high rRNA copy-number, possible intra-genomic variation or shared V4 barcode between different species or genera. These stumbling blocks need to be taken in consideration while analysing HTS data for dinoflagellates in future studies. Yet, the barcoding approach shows a great potential to explore dinoflagellate diversity and monitor different taxa in time and space. At LTER-MC it offers an important perspective if this study is expanded and replicated over a longer time period in order to assess seasonal trends of dinoflagellate diversity in the Gulf of Naples.

CHAPTER IV: Diversity of the dinoflagellate genus *Tripos* in the Gulf of Naples

In Chapter IV, Solenn Mordret and Thomas Mollica isolated the Tripos cells in natural plankton samples. Solenn Mordret, Thomas Mollica performed molecular analyses (Extraction, Amplification, Purification and Sequencing). S.M assembled raw reads and performed all phylogenetic analyses.

4.1. Introduction

The marine dinoflagellate genus *Tripos* Bory can be considered as one of the most iconic genera of dinoflagellates. The highly recognisable "anchor" shape, large cell size (from around 50 microns to 1 mm) and presence of thick thecal plates make sampling and recognition of *Tripos* cells relatively easy. Members of the genus *Tripos* belong to family Ceratiaceae and the order Gonyaulacales with the Kofoidian plate formula of 4', 6'', 5c, 6''', 2'''' (Steidinger and Tangen, 1997). Generally, *Tripos* cells exhibit three elongated horns, one anterior and two posteriors. The length, the orientation, the shape and the ornamentations of these horns vary greatly within the genus and underlie species classification. *Tripos* species are widespread, thriving from polar to equatorial seas and in open oceans as well as coastal waters. Together with *Protoperidinium* species, representatives of this cosmopolitan genus are among the most commonly observed unicellular species in the marine plankton.

History of the genus

Due to its characteristic morphological features and ubiquitous presence, *Tripos* was among the earliest phytoplanktonic taxa described. Documented for the first time by Müller in 1786, the marine species *Cercaria tripos* was grouped at that time with the freshwater species *Bursaria hirundiniella* and non-dinoflagellate organisms based on morphological similarities. In 1793, Schrank erected the genus *Ceratium* for the three freshwater species *C. pleuroceras*, *C. tetraceras* and *C. macroceras*; later he transferred *Cercaria tripos* and *Bursaria hirundiniella* to the genus *Ceratium* based on their morphological similarity. Since then, numerous species have been described. An abundant literature exists, including a few monographs classifying the genus into subgenera and sections. Four subgenera are distinguished based on the morphology by Vanhöffen (1986), Kofoid and Swezy (1921), Ostenfeld (1903) and Jørgensen (1911): *Amphiceratium*, *Biceratium*, *Archaeceratium* and *Tripoceratium*. Each subgenus itself being divided in sections based on the cell silhouettes, horn orientation and ornamentation.

Revision of the genus with molecular data

The genus *Ceratium* was revised by Gómez and co-authors in 2010 based on the first 18S phylogenies (Gómez, Moreira and López-García, 2010). They demonstrated that the phylogenies did not validate the division of species in subgenera based on cell shape and thecal ornamentation. Instead, none of the subgenera were monophyletic and the variation of the 18S sequences was quite low. Yet, the study supported a clear separation of the freshwater and marine species into two distinct clades, which was supported by morphological differences in a number of the plates. The marine species were then transferred into the new genus *Neoceratium*, reserving the genus *Ceratium* for the type species and the other freshwater taxa. However, for priority of older synonyms (Calado and Huisman, 2010), the name *Neoceratium* had to be considered illegitimate and the genus *Tripos* was reinstated to refer to marine taxa (Gómez, 2013).

The ecological role of *Tripos*

Most *Tripos* species are considered mixotrophic (Jacobson, 1999), that is, they can photosynthesise but in addition need to ingest organic material. A few species such as *Tripos furca* or *Tripos fusus* can be grown in culture and have been used as models, shedding light on many aspects of dinoflagellate ecophysiological processes, mobility and life cycle (Hasle and Nordli, 1951; Jacobson, 1999; Smalley, Coats and Adam, 1999; Smalley, Coats and Stoecker, 2003; Baek, Shimode and Kikuchi, 2007; Baek *et al.*, 2009). Yet, the culture of many species remains difficult, if not impossible. So, despite the fact that the morphological diversity of this conspicuous genus is well characterised, the taxonomy of many species has never been evaluated by molecular means.

Due to its frequent occurrence in net samples and relatively easy identification, the genus has been reported extensively and is one of the rare lineages of dinoflagellates for which the biogeographical distribution has been studied (Taylor, Hoppenrath and Saldarriaga, 2008). Biogeographic data of *Tripos* suggests that temperature affects distribution (Semina and Levashova, 1993). Dodge and Marshall (1994) presented a biogeography based on thermal affinities of different species of *Tripos* including six thermal categories (**Table 4.1.1**). In addition,

many studies showed consistent and regular occurrence in of several *Tripos* species at specific times of the year (Dodge and Marshall, 1994; Tunin-Ley *et al.*, 2007; Aubry *et al.*, 2012). Other studies reported an increase in abundance of the genus with increased surface layer temperatures (Li *et al.*, 2011; Vázquez-Domínguez, Vaqué and Gasol, 2012). This sensitivity to temperature, rapid detection and worldwide distribution renders the genus a promising candidate as ecological indicator for monitoring global warming (Tunin-Ley and Lemée, 2013).

Table 4.1.1: Biogeographic categorization of *Tripos* species proposed by Dodge and Marshall (1994).

		Temperature affinities/characteristics
1	**Arctic-temperate**	Temperature < 15°C
2	**Cosmopolitan**	Ubiquitous and frequently bloom forming species
3	**Intermediate**	Species absent from coldest and warmest waters
4	**Temperate-tropical**	Lower temperature limit: 5-12°C
5	**Warm-temperate-tropical**	Lower thermic boundary: 14-15°C
6	**Tropical**	Species rarely found when temperature < 20°C

Limits of morphological classification

The genus displays a remarkable diversity of shapes that led to the description of more than 150 morphological taxa including infraspecific diversity (i.e varieties) (Guiry and Guiry, 2017). However, several studies showed that the high morphological diversity is also the expression of morphological plasticity in response to environmental factors. Comparing different taxa occurring at different seasons, Sournia (1967) suggested that some morphological changes could be induced by temperature, which affects the viscosity of the water. He noticed that slender morphotypes with long, thin or wide horns are better adapted to warm waters (lower viscosity) while, robust silhouettes with short horns and thick theca are favoured in colder water (higher viscosity). Therefore, he hypothesized that the morphological variability was a result of phenotypical adaption to environmental conditions (in this case an adaptation to the floatability of the cell). In the same way, Lyakh and Bryantseva (2014) observed a seasonal polymorphism in three species of *Tripos*, displaying distinct winter and summer forms. Other studies showed high phenotypic

plasticity even in a single strain. In *Tripos ranipes* cell shape can change over the day: the finger-shaped appendages are formed during the day and reabsorbed during the night (Pizay *et al.*, 2009).

To date, 77 *Tripos* species are considered valid (Gómez, 2013) compared to the 150 morphological taxa but uncertainties exist for many species and varieties (Guiry and Guiry, 2017). A state of the art of the generic diversity is provided in **Annex 1** of this chapter including all species names and their taxonomic references. The classification based on morphological characters seems to lead to an overestimation of the number of species (Gómez, 2013) and the identification of phylogenetically significant characters useful to discriminate different species is difficult. There is a need to assess the diversity of the genus *Tripos* integrating the morphological and molecular characterisation of different strains or specimens. The rise of HTS metabarcoding approaches and other novel molecular techniques to study planktonic community at LTER-MC provided opportunities to investigate the genus in a more detailed way.

The aim of the study

The aim of this thesis chapter was to study morphological and genetic diversity of *Tripos* species in the Gulf of Naples. To achieve this goal, I and Thomas Mollica (master student), isolated single cells, and for each of them took a picture, isolated its DNA, and amplified and sequenced its 18S and 28S rRNA markers. These sequences were used to delineate the genetically distinct taxa, to infer the relationships among these taxa and to assess with the obtained phylogenies the acquisition of morphological character states. These sequences, together with their linked morphological information, were added to the DinoREF reference database to improve the interpretation of metabarcode datasets.

The V4 region was extracted from the 18S sequences of *Tripos* produced in this study as well as from those retrieved from GenBank to assess the diversity of the genus and seasonality of its various species in the HTS metabarcode dataset generated from plankton samples taken at the Long Term Ecological Research station MareChiara (LTER-MC).

4.2. Material and methods

4.2.1. Single cell isolation of *Tripos* cells

All cells isolated for this study were obtained at the LTER MareChiara (40°48.5′ N; 14°15′ E) in the Gulf of Naples (Tyrrhenian Sea, Italy) between February 2016 and May 2017 (**Table 4.2.1**). Plankton samples were collected with a plankton-net (20 µm mesh size) in the surface of the water column. Additional samples were collected further offshore at the L20 station (40°42′ N; 14°10′ E, about four nautical miles from the coast and on the 300 m isobath line) using a 40 µm mesh size plankton-net during winter and spring 2017 (**Table 4.2.1**).

Immediately after collection, samples were transported to the laboratory and diluted in sterile seawater. *Tripos* cells were isolated individually using a P100 micropipette under an inverted light microscope (Leica DMIL LED). Several pictures of each isolated cell were taken with a Leica MC170 HD photocamera at 20x magnification (10x magnification for larger cells) to document its morphology. All cells were identified tentatively based on their gross morphology. The assignations followed Sournia (1967), Rampi and Bernard (1980) and Steidinger and Tangen (1997). *Tripos* cells were then washed carefully in three successive baths of 0.22µm-filtered, sterile seawater and transferred each in a 1.5 mL Eppendorf tube and preserved in 25 µL of Lysis buffer (Tissue and Cell Lysis Solution from MasterPure™ DNA and RNA Purification Kit, Epicentre, Lucigen, Middleton WI, USA). The tubes were placed on ice for 15 min and then centrifuged for a few seconds to ensure that the cell was on the bottom of the tube. The tubes were stored at -20°C or processed immediately for DNA extraction.

4.2.2. Total DNA extraction

Cells were lysed by adding an additional 300 µL of Lysis buffer and 1 µL of Proteinase K, and incubated at 65 °C for 15 min under constant agitation (1000 rpm). Subsequently, 1 µL of RNAse was added to the solution, tubes were incubated at 37°C for 30 min under gentle agitation (550 rpm) to destroy RNA, and then put on ice for 5 min. Then, 150 µL of MPC solution (MasterPure Kit)

was added to the solution to precipitate proteins and organic membranes. Upon centrifugation (11 000 g) for 10 min, the supernatant was transferred in new 1.5 mL Eppendorf tubes and 500 µL of 100% isopropanol was added to precipitate the DNA. The solution was centrifuged at maximum speed for 10 min and the liquid discarded. The DNA pellet was rinsed in 500 µL of 70% ethanol, and dried 10 min under a chemical hood at room temperature. The resulting DNA pellet was dissolved in 25 µL TRIS buffer and the resulting solution stored at -20°C.

This DNA extraction protocol has been tested several times and optimised for *Tripos* cells by Thomas Mollica (master student) in collaboration with Raimondo Pannone and Elio Biffali. Modifications in the protocol involved lyse the freshly isolated cells directly in 300 µL of lysis buffer and 30 µL of Proteinase K overnight at 65°C. In addition, isopropanol precipitation was extended to overnight at -20°C. These modifications of the original protocol lead to a higher probability of PCR success. All cells isolated by Thomas Mollica (labelled **TM**) were extracted with this optimised protocol.

4.2.3. Amplification and purification

To genetically identify *Tripos* cells, I targeted the PCR amplification of the 18S and partial 28S rRNA coding regions using eukaryote-generalist primers or slightly modified primers adapted to dinoflagellates (**Table 4.2.2**; **Fig.4.2.1** and **Table 4.2.3**). Amplifications of both markers were conducted with the Phusion® High-Fidelity DNA Polymerase (Finnzymes). The PCR mixture (25 µL final volume) contained ca. 0.5 ng (2.5 µL) of single cell DNA, 0.5 µM (final concentration) of each primer, 3% of DMSO and 5X of GC Phusion Reaction buffer (Finnzymes), 200 µM dNTPs, and 0.02 u/ µL of Phusion DNA Polymerase. The PCR reactions were performed in the following conditions: one initial cycle of denaturation at 98 °C for 30 s, followed by 40 cycles of denaturation at 98 °C for 10 s, annealing at 60 °C (28S) or at 62 °C (18S) for 30 s, and extension at 72 °C for 30 s and final extension at 72 °C for 10 min.

When the Phusion DNA Polymerase failed to amplify the 18S rRNA, an attempt was made with the Expand ™ Long Template PCR system and the Expand™ High Fidelity PCR system (Roche,

Mannheim, Germany). The 25 µL PCR mixture was composed of the following reagents: 2.5 µL DNA extracts, 2.5 µL of 10x Expand™ Long Template Buffer 1, 2.5 µL (10 mM concentration) of each primer, 4 µL of dNTPs (10 mM concentration) and 0.5 µL of Expand Long™ Polymerase. The PCR cycle conditions were as followed: one initial cycle of denaturation at 94 °C for 2 min, followed by 40 cycles of denaturation at 94 °C for 15 s, annealing at 54 °C for 30 s, and extension at 72 °C for 2 min and final extension at 72 °C for 7 min.

Negative and positive controls were used to verify putative sample contamination from exogenous sources. To determine the presence of DNA amplicons and to visualise their length, PCR products were examined by means of agarose gel electrophoresis with TBE buffer 0.5x and the DNA-dye ethidium bromide.

DNA amplicons for both 18S and 28S rRNA were purified using the microCLEAN kit (Microzone, Haywards Heath, UK) or an ExoSAP-IT™ PCR Product Cleanup Reagent (ThermoFisher Scientific, Waltham, MA, USA).

For the microCLEAN kit, the protocol provided by the manufacturer was optimised in order to maximize DNA recovery from purification. An equal volume of microCLEAN were added directly to the PCR products. All tubes were mixed 5 min at room temperature (1000 rpm) on MixMate® Eppendorf machine and centrifuged at high speed (12,500 g) for 7 min. The supernatant was pipetted out letting up to 2-3 µL in the tubes and centrifuged again for 3 min in the same conditions. Finally, all the supernatant was removed and the samples were eluted in 25 µL of lukewarm TRIS buffer. To maximise the elution of DNA, samples were mixed on the MIxMate® machine at room temperature for 15 min.

With ExoSAP-IT Reagent, samples were processed as follows. Samples were placed on ice, 8 µL of reagent was added per 20 µL of PCR template, and mixed on a MixMate® machine. Tubes were incubated 15 min at 37°C to degrade primers and nucleotides, followed by incubation at 80°C for 15 min to inactivate the ExoSAP-IT reagent. When the amplification presented some supernumerary bands, a PCR gel extraction kit was used following the manufacturer's instructions

in order to purify only the band of interest. Either Gen-Elute gel extraction kit (Sigma-Aldrich, St-Louis, MO, USA) or High Pure PCR Product purification kit (Roche, Mannheim, Germany) was used.

4.2.4. Cloning

A cloning approach was adopted by T. Mollica for both 18S and 28S when the quantity of DNA detected on the gel was too low to perform direct sequencing. To obtain more DNA, amplicons were cloned using The Zero Blunt® TOPO® PCR Cloning Kit for Sequencing (Invitrogen ™ by Life Technologies, Carlsbad, CA, USA) according to manufacturer's instructions. The TOPO reaction was performed over 15 min instead of 5 min in order to maximise the transformation of the bacteria with the highest number of plasmid. After cloning and amplification of obtained colonies, DNA amplicons were purified with the GenElute™ Plasmid Miniprep Kit (Sigma-Aldrich, St-Louis, MO, USA) kit according to company's instructions.

4.2.5. Sequencing

PCR amplicons were sequenced using the ABI-PRISM Big Dye Terminator Sequencing kit (Applied Biosystems) according to the Sanger method and following the recommendations of the manufacturer. Sequencing was performed by the Molecular Biology and Sequencing Service (Stazione Zoologica Anton Dohrn), using fragment analyzer machines (3730 DNA Genetic Analyzer, Applied Biosystems Inc., Foster City, CA, USA)).

Sequencing the 18S rRNA required four primers (two external and two internal) to obtain a full-length sequence of around 1750 bp (for dinoflagellates). For the 28S rRNA, I PCR-amplified and sequenced the first part of the gene (D1 – D3 regions) corresponding to about 900 bp. Details on the primers used are present in **Table 4.2.2, Fig.4.2.1** and **Table 4.2.3**.

4.2.7. Phylogenetic analyses

Amplicon sequences obtained by PCR were cleaned and assembled using Chromas Pro (version 1.7.5, Technelysium, South Brisbane, Australia). For both 18S and 28S markers, a matrix of

sequences was built with all sequences obtained in the present study and sequences from GenBank published by other authors (**Table 4.2.4**). Outgroup sequences of close relatives were also included in each matrix for subsequent phylogenetic analyses. The sequence matrices were aligned with MAFFT (Katoh and Standley, 2013), implemented in Seaview software (Gouy *et al.*, 2010) and visualised. Phylogenetic analyses were conducted with the TOPALI software (Topali V2, Milne *et al.*, 2009) and Geneious software (Kearse *et al.*, 2012). According to Modeltest v0.1.1 (Posada, 2008), a general time-reversible (GTR) model of nucleotide substitution was selected for the 28S rRNA (939 bp nucleotide positions) and the 18S rDNA (1,137 and 1,577 nucleotide positions), respectively. The phylogenetic inference by Bayesian methods was performed with MrBayes (Huelsenbeck and Ronquist, 2001)(http://mrbayes.sourceforge.net/index.php), and robustness of inferred topologies was assessed by posterior probabilities.

4.2.8. Analysis of the HTS data from LTER Mare Chiara

To study the diversity of *Tripos* species in the Gulf of Naples,I had access to the V4 dataset built from plankton samples collected at the LTER-MC on 48 dates from 2011 to 2013 (**97 % dataset**, see Chapter III). The most abundant ribotypes recovered (>50 reads) and ribotypes matching references with 100 % similarity were then extracted from the total HTS metabarcode dataset. The heatmap of these selected ribotypes was prepared using the normalised abundance of reads by sample over the 48 sampling dates. All data was $\log_2$ transformed. Each ribotype was annotated following DinoREF (See **Chapter III**) and the cases of V4 sequences shared among multiple species were marked (See **Chapter III**). To explore the relationships between *Tripos* ribotypes and environmental parameters (**Appendix 2**), the BIO-ENV analysis (Clarke and Ainsworth, 1993) was performed using the Bray-Curtis dissimilarity matrix. BIO-ENV allows identification of a subset of variables that shows the highest explanatory values. The identified variables were used in Canonical Correspondence Analysis (CCA) plotting both samples and *Tripos* different ribotypes.

Table 4.2.1: Summary table of all single cells for which we obtained 18S rRNA or 28S rRNA sequences were obtained and used in phylogenies. **a**. Codes of the cells. Cells isolated by Solenn Mordret are annotated as "SC" and cells isolated by Thomas Mollica as "TM". **b**. Tentative identification made based on the morphology while isolating the cell. **c**. Specific date of isolation. **d**. Collection site. **e**. Sequences obtained for 18S and/or 28S rRNA. When part of the sequence obtained sequences was annotated with a p. **f**. Taxonomic assignation after phylogenetic analyses. Names that were corrected are in bold. **g**. Picture code to link sequence individual cell morphology in **Fig.4.3.5, Fig.4.3.6** and **Fig.4.3.7.**

a. Single cell code	**b.** Tentative identification	**c.** Date of isolation	**d.** Collection site	**e.**Sequence		**f**. Taxonomic assignation	**g.** Picture code
				18S	28S		
SC81	*T.* cf. *fusus*	18/02/2016	LTER-MC	**x**	**x**	***T. extensus***	Fig.4.3.5.**i**
SC82	*T. pentagonus*	18/02/2016		**x**	**x**	*T. pentagonus*	Fig.4.3.5.**a**
SC86	*T.* cf. *pavillardii*	02/03/2016			**x**	*T. pavillardii*	Fig.4.3.5.**o**
SC92	*T.* cf. *pavillardii*	02/03/2016		**x**		*T. pavillardii*	Fig.4.3.5.**m**
SC93	*T.* cf. *horridus/ massiliensis*	02/03/2016		**x**	**x**	*T. horridus/ massiliensis*	Fig.4.3.6.**k**
SC94	*T.* cf. *pavillardii*	02/03/2016		**x**	**x**	*T. pavillardii*	Fig.4.3.5.**p**
SC96	*T.* cf. *massiliensis*	02/03/2016		**x**	**x**	***Tripos* sp.**	Fig.4.3.6.**b**
SC98	*T.* cf. *massiliensis*	02/03/2016		**x**	**x**	*T. massiliensis*	Fig.4.3.6.**q**
SC99	*T.* cf. *macroceros*	02/03/2016		**x**		***T. deflexus***	Fig.4.3.7.**q**
SC100	*T.*cf. *fusus*	02/03/2017		**x**		***T. extensus***	Fig.4.3.5.**h**
SC101	*T.* cf. *contortus*	02/03/2016		**p**		*T. contortus*	Fig.4.3.7.**n**
SC102	*T.* cf. *horridus/ massiliensis*	02/03/2016		**x**	**x**	*T. horridus/ massiliensis*	Fig.4.3.6.**m**
SC105	*T. macroceros*	02/03/2016		**x**		*T. macroceros*	Fig.4.3.7.**o**
SC110	*T. concilians*	02/03/2016		**x**	**x**	*T. concilians*	Fig.4.3.7.**h**
SC113	*T.* cf. *horridus/ massiliensis*	15/03/2016			**x**	*T. horridus/ massiliensis*	Fig.4.3.6.**i**
SC114	*T.* cf. *horridus/ massiliensis*	15/03/2016		**x**	**x**	*T. horridus/ massiliensis*	Fig.4.3.6.**j**
SC116	*T.* cf. *trichoceros*	15/03/2016		**x**	**x**	*T. trichoceros*	Fig.4.3.6.**c**
SC117	*T.* cf. *trichoceros*	15/03/2016		**x**	**x**	*T. trichoceros*	Fig.4.3.6.**e**
SC118	*T.* cf. *horridus/ massiliensis*	15/03/2016		**x**	**x**	*T. horridus/ massiliensis*	Fig.4.3.6.**h**
SC120	*T.* cf. *trichoceros*	15/03/2016		**x**	**x**	*T. trichoceros*	Fig.4.3.6.**g**
SC121	*T.* cf. *trichoceros*	15/03/2016		**x**	**x**	*T. trichoceros*	Fig.4.3.6.**d**
SC122	*T.* cf. *trichoceros*	15/03/2016		**x**	**x**	*T. trichoceros*	Fig.4.3.6.**f**
SC123	*T. furca*	15/03/2016		**x**	**x**	*T. furca*	Fig.4.3.5.**e**
SC130	*T. carriensis*	21/03/2016		**x**	**x**	*T. carriensis*	Fig.4.3.5.**l**
SC134	*T.* cf. *horridus/ massiliensis*	21/03/2016		**x**	**x**	*T. horridus/ massiliensis*	Fig.4.3.6.**l**
SC138	*T. ranipes*	21/03/2016			**x**	*T. ranipes*	Fig.4.3.6.**a**

a. Single cell code	**b.** Tentative identification	**c.** Date of isolation	**d.** Collection site	**e.**Sequence		**f**. Taxonomic assignation	**g.** Picture code
				18S	28S		
SC139	*T.* cf. *pavillardii*	21/03/2016		**x**	**x**	*T. pavillardii*	Fig.4.3.5.**n**
TM2	*T. contortus*	14/02/2017			**x**	*T. contortus*	Fig.4.3.7.**k**
TM6	*T. euarcatus*	21/02/2017			**x**	***T. muelleri***	Fig.4.3.7.**d**
TM8	*T. azoricus*	21/02/2017		**p**	**x**	*T. azoricus*	Fig.4.3.5.**f**
TM10	*T. fusus*	21/02/2017			**x**	*T. fusus*	Fig.4.3.5.**g**
TM15	*T. pentagonus*	28/02/2017	Offshore	**p**	**x**	*T. pentagonus*	Fig.4.3.5.**b**
TM18	*T. candelabrus*	28/02/2017			**x**	*T. candelabrus*	Fig.4.3.7.**s**
TM20	*Tripos* sp.	28/02/2017			**x**	***T. declinatus***	Fig.4.3.7.**c**
TM24	*T. paradoxides*	28/02/2017			**x**	*T. paradoxides*	Fig.4.3.5.**k**
TM45	*T. pentagonus*	09/03/2017			**x**	*T. pentagonus*	Fig.4.3.5.**c**
TM46	*Tripos* sp.	09/03/2017			**x**	***T. contortus***	Fig.4.3.7.**m**
TM52	*T. gibberus*	14/03/2017		**x**	**x**	***T. concilians***	Fig.4.3.7.**i**
TM53	*T. candelabrus*	14/03/2017		**x**	**x**	*T. candelabrus*	Fig.4.3.7.**r**
TM54	*T. declinatus*	14/03/2017		**x**	**x**	***T. contortus***	Fig.4.3.7.**j**
TM55	*T. macroceros*	14/03/2017		**x**	**x**	***T. massiliensis***	Fig.4.3.6.**p**
TM58	*Tripos* sp.	14/03/2017		**x**	**x**	***T. declinatus***	Fig.4.3.7.**b**
TM59	*Tripos* sp.	14/03/2017		**x**	**x**	***T. contortus***	Fig.4.3.7.**l**
TM63	*Tripos* sp.	21/03/2017		**x**		***T. horridus/ massilensis***	Fig.4.3.6.**o**
TM65	*T. paradoxides*	21/03/2017		**x**		***T. hexacanthus***	Fig.4.3.7.**p**
TM68	*T. pulchellus*	28/03/2017		**x**	**x**	*T. pulchellus*	Fig.4.3.7.**g**
TM70	*Tripos* sp.	28/03/2017			**x**	***T. pulchellus***	Fig.4.3.7.**f**
TM71	*T. massiliensis*	05/04/2017		**x**	**x**	*T. horridus/ massiliensis*	Fig.4.3.6.**n**
TM72	*T. pulchellus*	05/04/2017		**x**	**x**	*T. pulchellus*	Fig.4.3.7.**e**
TM73	*T. declinatus*	05/04/2017			**x**	*T. declinatus*	Fig.4.3.7.**a**
TM74	*T. furca*	05/04/2017		**x**	**x**	*T. furca*	Fig.4.3.5.**d**
TM98	*T. extensus*	16/05/2017			**x**	*T. extensus*	Fig.4.3.5.**j**
52 cells	***20 different taxa***			**38**	**46**	***22 different taxa***	

Table 4.2.2: Primer sets and annealing temperatures used in this study for PCR amplification and sequencing of 18S rRNA and 28S rRNA of *Tripos* cells. M13 Primers are primers used for cloning.

	MARKER	FRAGMENT SIZE	Forward PRIMER	Reverse PRIMER	PCR Annealing T°C
PCR with Phusion	18S rRNA gene	1750 bp	18S-FV	18S-RV	62°C
	28S rRNA gene	1500 bp	Dino-D1R-C	28S-1483R	60°C
Sequencing	18S rRNA gene	1750 bp	18S-FV/ SR4-F	1055R/18S-RV	57°C
	28S rRNA gene	900 bp	Dino-D1R-C	Dino-D3Ca-R	57°C
Cloning primers	Cloning Topo 4 vector	Cloned fragment	M13-F	M13-R	56°C

Fig.4.2.1: Relative position of the primers used in this study for amplification and sequencing on the ribosomal operon.

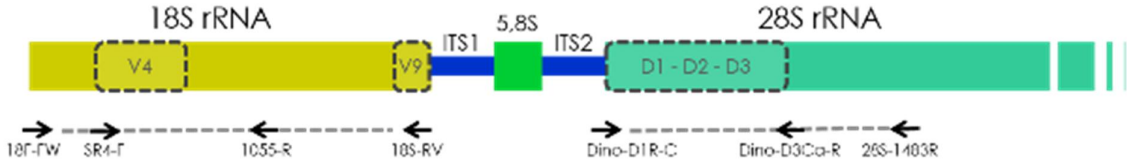


Table 4.2.3: List of primers used in this study, corresponding sequence and reference.

	Sequence (5′ – 3′)	Reference
18S-FW	TCC TGC CAG TAG TCA TAT GC	Chomérat *et al.*, 2010
SR4-F	AGG GCA AGT CTG GTG CCA G	Yamaguchi and Horiguchi, 2005
Dino-D1R-C	ACC YGC TGA ATT TAA GCA	This study
18S-RV	TGA TCC TTC GGC AGG TTC AC	Chomérat *et al.*, 2010
1055R	GGT GGT GGT GCA TGG CCG TTC TTA G	Elwood, Olsen and Sogin, 1985
28S-1483R	GCT ACT ACC ACC AAG ATC TGC AC	Daugbjerg *et al.*, (2000)
Dino-D3Ca-R	GAC GAA CGA TTT GCA CGT CAG	This study

Table 4.2.4: List of sequences of 18S rRNA or 28S rRNA and GenBank Accession Numbers used to build the phylogenies. Sequences produced by Gómez *et al.* (2010) were annotated with an asterisk.

Taxon	Marker	GenBank Accession Number
*Tripos candelabrus**	**18S**	FJ402955
*Tripos candelabrus**	**18S**	FJ402945
*Tripos concilians**	**18S**	FJ402944
*Tripos concilians**	**18S**	FJ402950
*Tripos pentagonus**	**18S**	FJ402948
*Tripos declinatus**	**18S**	FJ402949
*Tripos hexacanthus**	**18S**	FJ402943
*Tripos horridus**	**18S**	FJ402960
Tripos platycornis	**18S**	FJ824911
*Tripos massiliensis**	**18S**	FJ402942
*Tripos contrarius**	**18S**	FJ402959
*Tripos minutum**	**18S**	FJ402964
*Tripos kofoidii**	**18S**	FJ402963
Tripos longipes	**18S**	DQ288462
*Tripos symmetricus**	**18S**	FJ402947
*Tripos arietinus**	**18S**	FJ402956
*Tripos limulus**	**18S**	FJ402962
*Tripos limulus**	**18S**	FJ402952
*Tripos paradoxides**	**18S**	FJ402965
*Tripos furca**	**18S**	FJ402966
Tripos furca	**18S**	AJ276699
*Tripos extensus**	**18S**	FJ402957
*Tripos fusus**	**18S**	FJ402958
Tripos fusus	**18S**	AF022153
*Tripos gravidus**	**18S**	FJ402961
Tripos digitatus	**18S**	FJ824940
*Tripos petersii**	**18S**	FJ402953
*Tripos petersii**	**18S**	FJ402951
*Tripos euarcatus**	**18S**	FJ402946
*Tripos azoricus**	**18S**	FJ402954
Ceratium hirundinella	**18S**	JQ636759
Ceratium furcoides	**18S**	JQ639758
Ceratium furcoides	**18S**	JQ639757
Tripos balechii	**28S**	JQ638944
Tripos muelleri	**28S**	AF260389
Tripos sp.	**28S**	KT389993
Tripos fusus	**28S**	AF260390
Tripos fusus	**28S**	EF517276
Tripos lineatus	**28S**	AF260391
Alexandrium margalefii	**28S**	AY154957
Alexandriumpseudogonyaulax	**28S**	AY549558

4.3. Results

Morphological characterisation

During live observations of plankton samples from the Gulf of Naples, over 150 *Tripos* cells were isolated and photographed between February 2016 and March 2017. Based on the pictures, various morphological characters of the newly isolated cells were observed and compared with specialised literature to assign specimens at species level. This tentative identification based on morphology can be challenging due to the presence of cells in which horns are broken due to the isolation procedure, hiding one of the most important characteristic for species differentiation. The cell code, the tentative isolation made based on the morphology and all isolation details were summarised in **Table 4.2.1**. A total of 21 morpho-species were recognised. When the identification remained uncertain, the cell was annotated as *Tripos* "sp." or identified by two possible names. All cells isolated by myself were named "SC" and cells isolated by Thomas Mollica as "TM".

Molecular characterisation

Out of all *Tripos* cells isolated in this study, molecular information was obtained for 52 cells (listed in **Table 4.2.1**), including 37 partial sequences of the 18S rRNA-encoding region and 46 of the partial 28S rRNA encoding region. All 28S sequences were of more or less the same length (around 900 bp – D1-D3 domains) while the length of the 18S marker varied greatly depending on how large a part of the sequence could be read depending on the sequencing success of some or all of the four sequence primers (**Fig.4.2.1**). In order to include the maximum number of references, an initial phylogenetic tree was built using 18S partial sequences (**1,137 bp** alignment covering the V4 to V7 regions; **Fig.4.3.2**). This tree comprised 28 single cell sequences produced in this study and 30 sequences from literature, which corresponded to this length. A second 18S phylogenetic tree was inferred from a longer alignment (**1,577 bp**, 23 sequences from single cells produced in this study and 7 published reference sequences; **Fig.4.3.3**). The 28S phylogenetic tree (939 bp - **Fig.4.3.4**) included products of 45 single cells obtained in this study and only 6 published reference sequences.

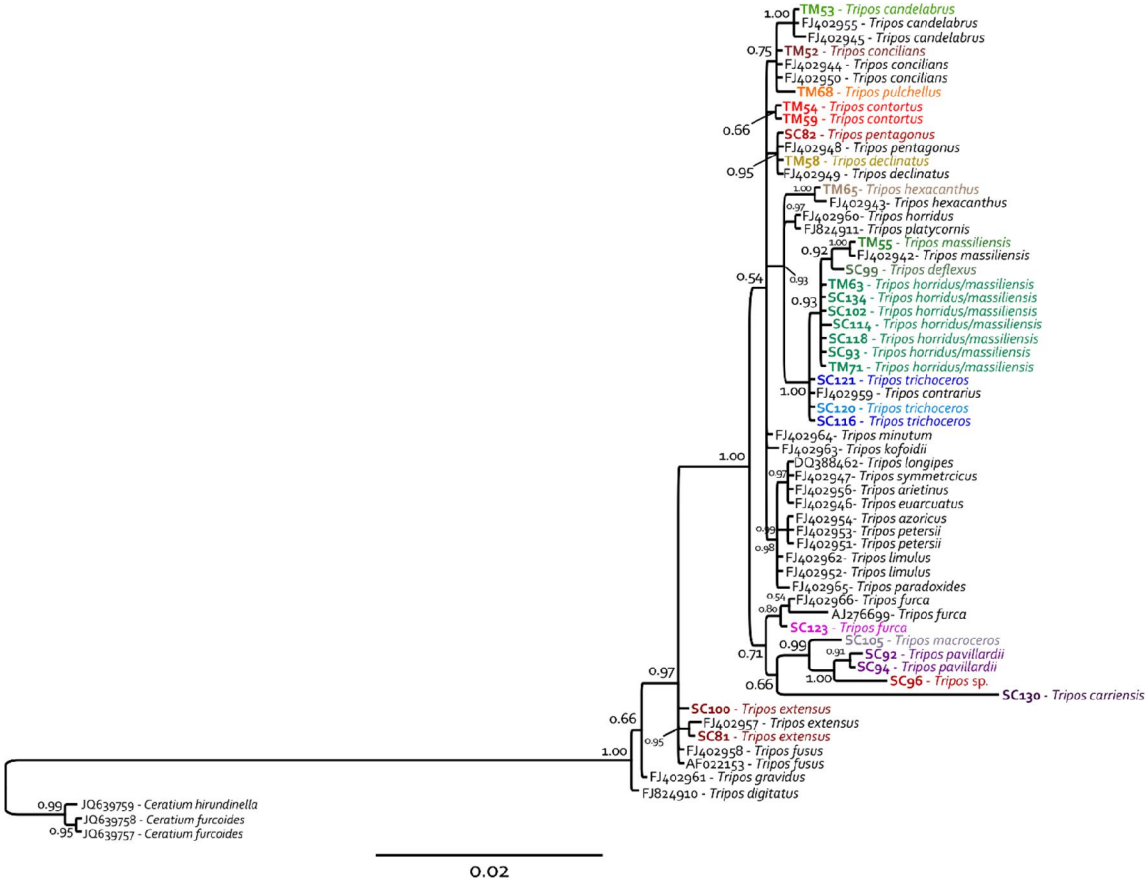


Fig.4.3.2: Bayesian phylogeny inferred from partial nuclear 18S rRNA sequences (**1,137 bp**; Evolution model: GTR) of *Tripos* obtained in this study (listed in **Table 4.2.4**) and from literature. Freshwater *Ceratium* reference sequences were used as outgroup. Numbers at nodes represent posterior probability values. The sequences obtained in this study are in boldface and coloured according to the clades identified in this study.

In both 18S and 28S phylogenies, all *Tripos* sequences grouped with high support values (1.00 posterior values) in a monophyletic group different from other close relative genera. In these phylogenies, different well-supported terminal clades could be detected. However, both phylogenies were largely polytomic, the backbone of the trees was not supported (low posterior values) and the phylogenetic relationships amongst these clades were not resolved. Nonetheless, the 28S phylogeny (**Fig.4.3.4**) seemed to have a much better resolution in comparison with the 18S rRNA (**Fig.4.3.2** and **Fig.4.3.3**), which did not provide a clear distinction between the different genotypes.

In the 28S phylogenetic tree (**Fig.4.3.4**), 17 well-supported terminal clades were distinguished (posterior values >0.9). Within the *Tripos* genus, several of these clades clustered together

sequences with the same morpho-species assignation such as for *T. pentagonus*, *T. furca*, *T. pavillardii*, *T. candelabrus* and *T. pulchellus*. Yet, cryptic diversity was also detected. This was the case for sequences assigned to *T. massiliensis*, which grouped in different clades and were genetically different even if no morphological difference could be observed. Similarly, in a close sister clade, two clades of morphologically similar *T. trichoceros* were distinguished genetically with maximal statistical support (1.00 posterior values). In other cases, the Bayesian 28S phylogenetic tree allowed genetical discrimination among species for which morphological characteristics setting species boundaries are poorly defined. For instance, in the *T. muellerii* group, different cells belonging to the highly similar species *T. muellerii*, *T. declinatus*, *T. contortus* were re-assigned to the "right" species based on the phylogenetic analysis. Overall, in this study we produced 28S rRNA sequences for 19 new *Tripos* lineages (i.e only *Tripos fusus* was characterised) that had never been characterised molecularly before.

In the 18S Bayesian tree (**Fig.4.3.2**), the terminal clades were consistent with the ones of the 28S phylogeny (**Fig.4.3.4**). Nonetheless there were less sequences in the 18S tree than in the 28S tree, and some "morpho-species" were not represented in one of the 18S phylogenies. In the same way, a few cells were represented only by an 18S sequence (SC92, SC99, SC100 & TM63; **Table 4.2.1**; **Fig.4.3.2**). For these reasons, some clades/species were not represented in both 18S and 28S tree. Moreover, unlike in the 28S phylogeny, a few *Tripos* taxa such as *T. digitatus, T. gravidus, T. fusus* and *T. extensus* occupied a basal position excluded from the principal *Tripos* clade (1.00 posterior probabilities). The same pattern was observed in the long alignment 18S tree (**Fig.4.3.3**). In this tree, the terminal clades were similar but generally with a better support than with a shorter alignment. For example, the clade grouping *T. pavillardii*, *T. macroceros* and *T. carriensis* was statistically significant (0.98) in the tree built with the longer sequences, while the support was lower (0.66) in the tree built with shorter sequences. Yet, in some cases the phylogenetic resolution obtained with the sequences of different length was different. This was the case of *T. declinatus* (TM58) and *T. pentagonus* (SC82). In 1,137 bp tree (**Fig.4.3.2**), the two species clustered

in the same well-supported clade (0.95) whereas they have a polytomic position in tree built with 1,577 bp-long sequences (**Fig.4.3.3**).

Overall, 22 phylogenetic morpho-species corresponding to different clades and comparing both 18S and 28S rRNA phylogenies were identified. One clade (*T. "horridus/massiliensis"*) kept a dual assignation because we were not able to determine a unique name based on the photographs. In the same way, the cell **SC96** was annotated "*Tripos* sp." because the sequence did not cluster with other sequences and the picture did not allow a precise species identification. Moreover, due to a lower resolution capability of the 18S, some morphospecies/genotypes could not be discriminated in this phylogeny (**Fig.4.3.2**). For example, *T. declinatus* (**TM58**) and *T. pentagonus* (**SC82**) cluster into two different parts of the 28S phylogeny, while in the 18S tree they grouped together with a good support (>0.95). At the end of the analysis, out of the 52 isolated cells, 15 were renamed based on an iterative process in which we compared phylogenies and morphological information.

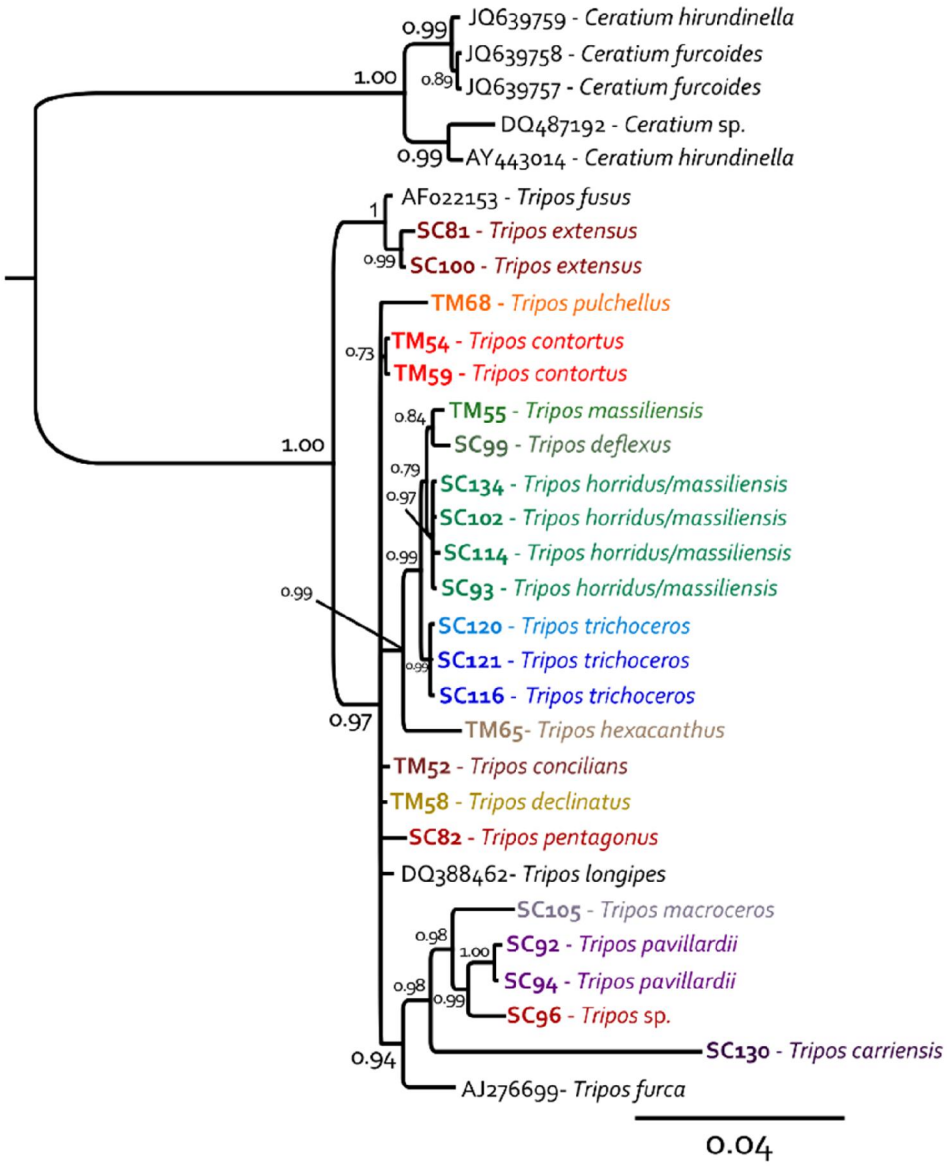


Fig.4.3.3: Bayesian phylogeny inferred from partial 18S rRNA sequences (**1,577 bp**; Evolution model: GTR) of *Tripos* (listed in **Table 4.2.4**). Freshwater *Ceratium* reference sequences were used as outgroup. Numbers at nodes represent posterior probability values. The sequences obtained in this study are highlighted in bold and coloured according to the clades identified in this study.

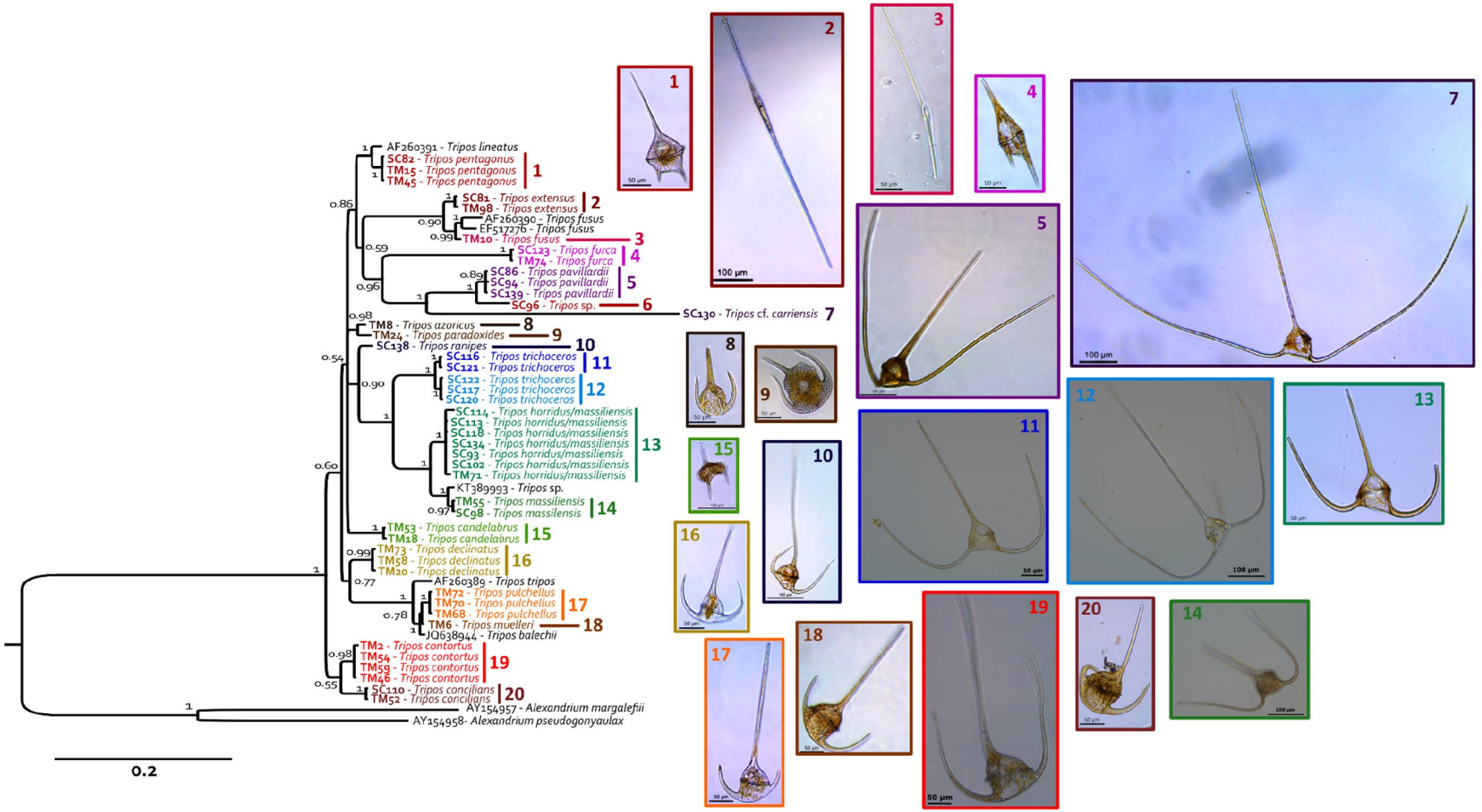


Fig.4.3.4: Bayesian phylogeny inferred from partial nuclear rRNA 28S sequence (939 bp; Evolution model: GTR) with reference sequences of *Tripos* (listed in **Table 4.2.4**) as well as environmental sequences from single cell obtained in this study. *Alexandrium* reference sequences were used as outgroup. Numbers at nodes represent posterior probabilities values. The sequences obtained in this study were highlighted in bold and coloured according to their morpho-phylogenetic clade.

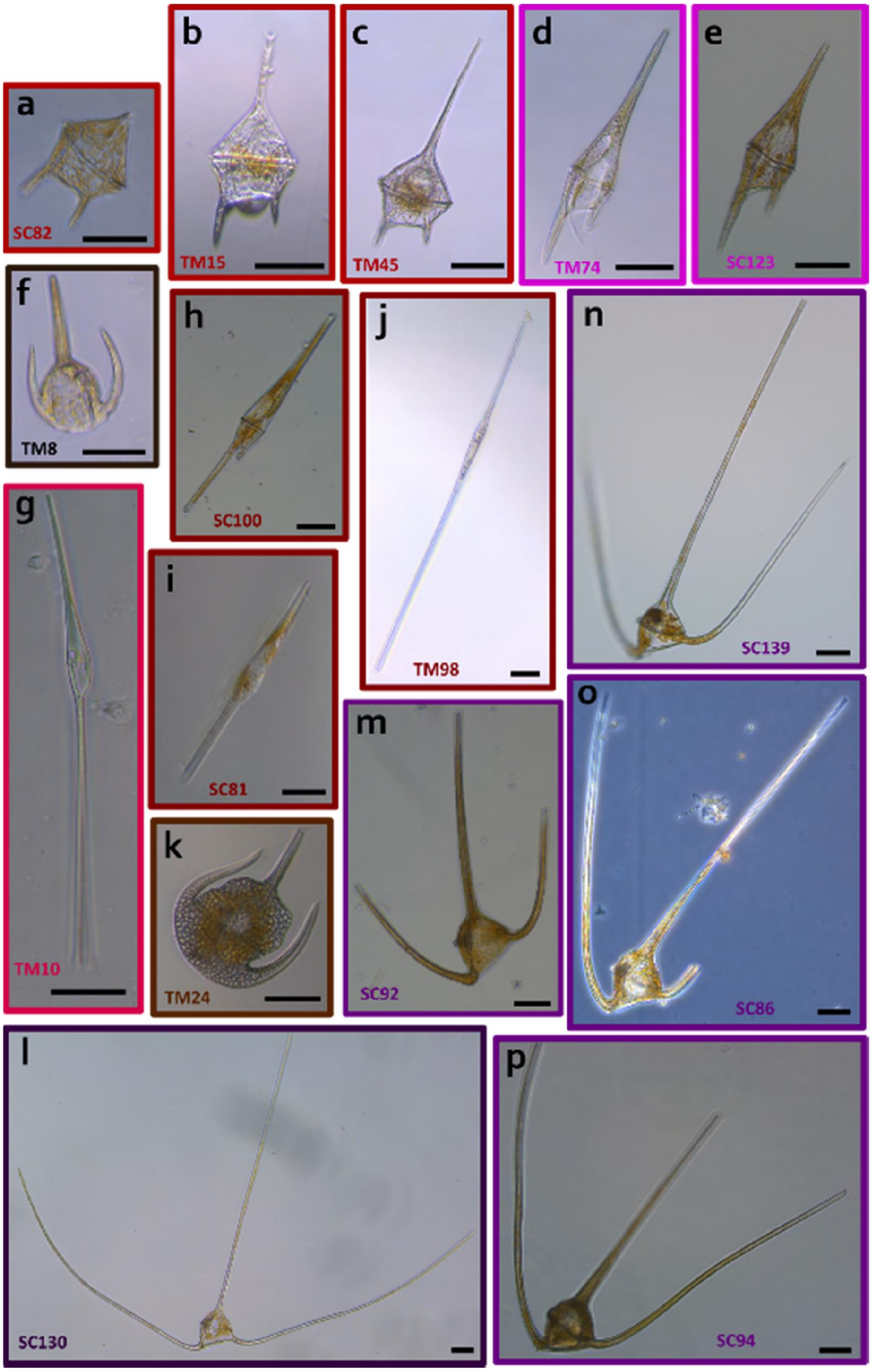


Fig.4.3.5: Light micrographs of cells isolated and sequenced in this study. The scale bars correspond to 50 µm. Information about each cell is found in Table 4.2.1. Colours match different clades of the 28S rRNA phylogeny (**Fig.4.3.4**). **a-c.** *Tripos pentagonus*. **d-e.** *T. furca*. **f.** *T. azoricus*. **g.** *T. fusus*. **h-j.** *T. extensus*. **k.** *T. paradoxides*. **l.** *T. carriensis*. **m-p**. *T. pavillardii*.

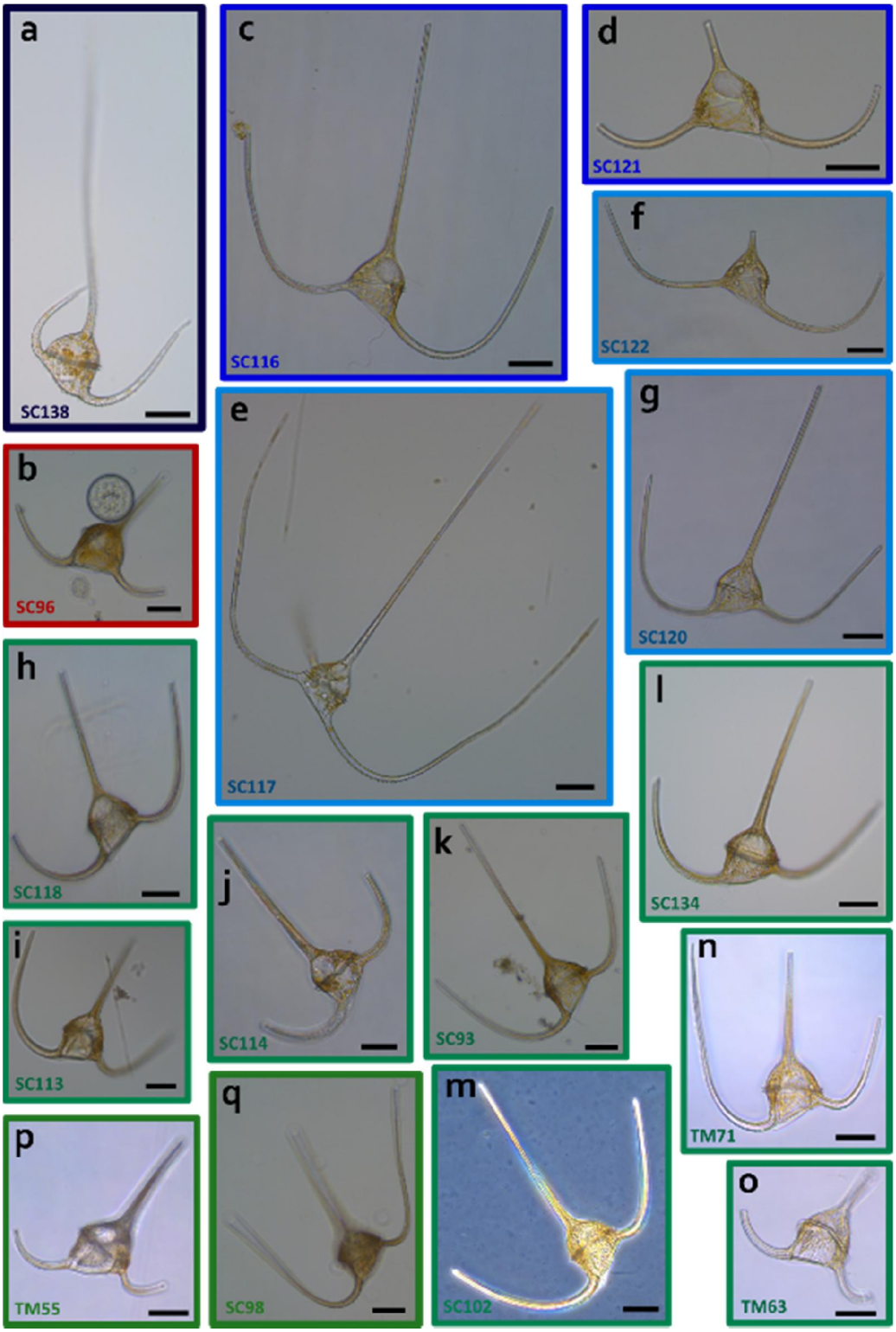


Fig.4.3.6: Light Microscopy pictures of cells isolated and sequenced in this study. The scale bars correspond to 50 µm. Information about each cell is found in Table 4.2.1. Colours match different clades of the 28S rRNA phylogeny (**Fig.4.3.4**). **a.** *Tripos ranipes*. **b.** *Tripos* sp. **c-g.** *T. trichoceros*. **h-o.** *T. horridus-massiliensis*. **p-q.** *T. massiliensis*.

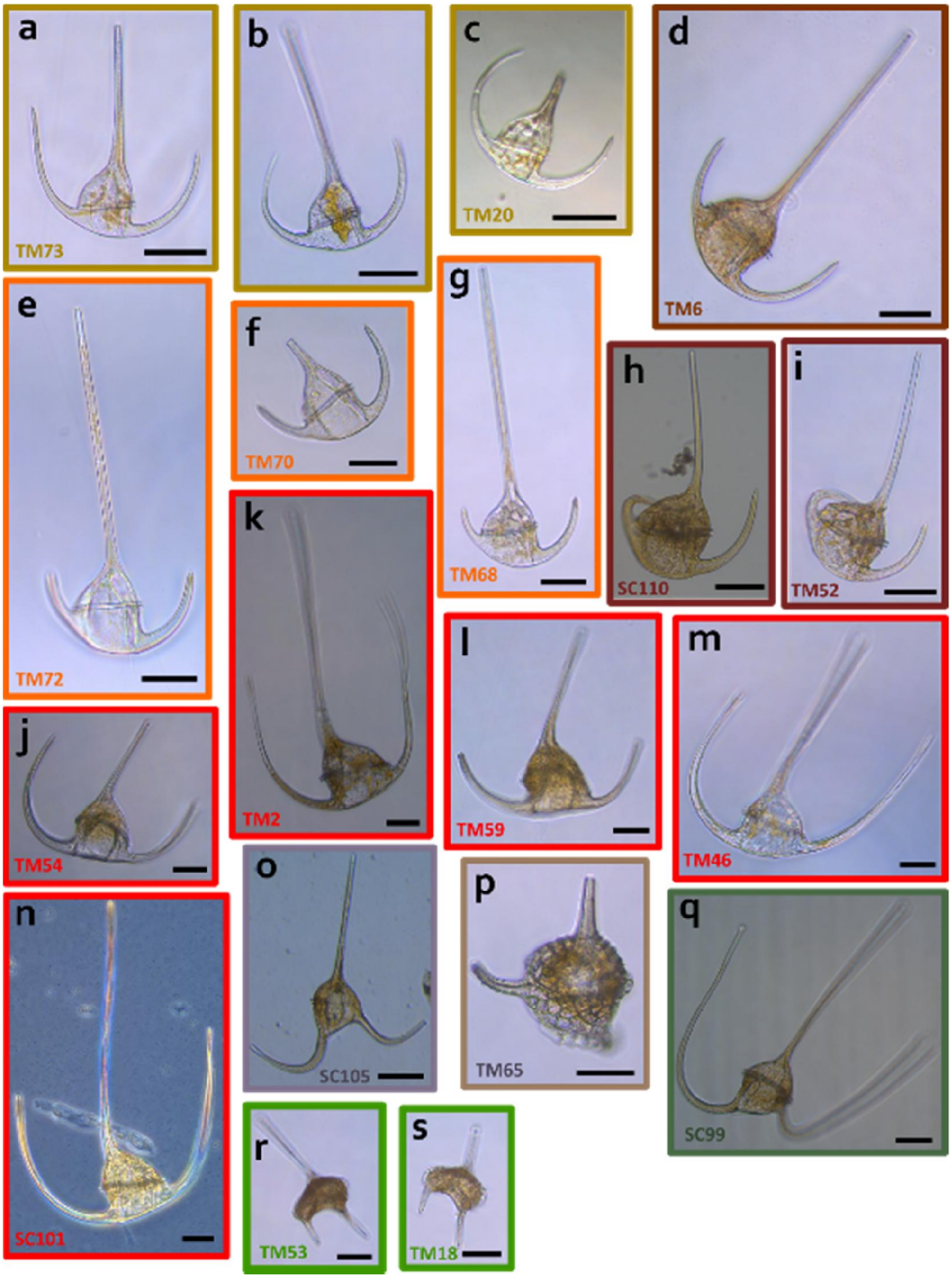


Fig.4.3.7: Light Microscopy pictures of cells isolated and sequences in this study. The scale bars correspond to 50 µm. Information about each cell is found in Table 4.2.1. Colours match different clades of 28S rRNA phylogeny (**Fig.4.3.4**). **a-c.** *Tripos declinatus*. **d.** *T. muelleri*. **e-g.** *T. pulchellus*. **h-i.** *T. concilians.* **j-n.** *T. contortus.* **r-s.** *T. candelabrus*. A few cells (**o, p, q**) were not present in the 28S but only in 18S rRNA phylogeny (**Fig.4.3.2**) were coloured in other colours.

Exploration of HTS metabarcodes for *Tripos*

Out of the 37 18S rRNA sequences produced in this study, 33 included the full V4 region (TM8, TM15 and TM54 having only partial V4). Moreover, 30 sequences of 18S rRNA were available from GenBank, representing a total number of 24 *Tripos* species (**Table 4.2.4**).

Out of the 63 unique 18S rRNA sequences (i.e produced in this study and references) corresponding to 32 morpho-species, 28 different V4 sequences were obtained for *Tripos* (**Table 4.3.5**). The V4 sequences were used in the LTER-MC dataset to detect the signal of different species during the tree year sampling. In some cases, the ribotypes matched perfectly with reference sequences already published, such as for *Tripos extensus* (V4 **#1**), *T. furca* **(**V4**#5**) or *T. candelabrus* (V4 **#22**). In one case, sequences produced in this study and already published sequences shared the same V4 sequences, but did not possess the same assignation (*T. trichoceros* vs. *T. contrarius* (V4 **#15**)). In addition, some sequences with the same taxonomic assignation exhibited a different V4 sequence (i.e. *T. furca* (V4 **#6**) or *T. hexacanthus* **(**V4 **#21**)). Moreover, in the cases in which morphologically identical specimens exhibited different V4 sequences, the differences did not exceed 1 to a few base changes (max 3). In contrast, a number of cases occurred in which specimens with distinct morphology shared identical V4 sequences. Examples include *T. kofoidii/ T. minutus/ T. limulus/ T. paradoxides* (V4 **#13**); *Tripos arientinus/ T. symmetricus/ T. euarcatus/ T. longipes* (V4 **#14**) and *Tripos declinatus/ T. pentagonus* (V4 **#23**).

A total of 11 new genetically different V4 sequences were produced in this study with seven detected in the LTER-MC dataset. Four V4 sequences, corresponding to *Tripos furca* (V4 **#6**), *T. pavillardii* (V4 **#8**), *T. carriensis* (V4 **#11**) and *T. hexacanthus* (V4 **#21**), were not retrieved in the LTER-MC metabarcode dataset. Overall, 20 different V4 sequences were retrieved with 100% similarity within the LTER-MC dataset. Moreover, 11 ribotype sequences with an abundance higher than 50 reads were detected, matching with one of the 28 different V4 sequences with 1 to 11 mismatches to a reference (99.74 to 97.11% similarity).

Table 4.2.5: List of reference V4 sequences available and used to explore the LTER-MC dataset. Some sequences were available in GenBank and some sequences were produced in this study by Thomas Mollica (TM) and Solenn Mordret (SC). When a reference shares the same V4 sequence with other references, the code (GenBank number or Single Cell code) was detailed in each line. Each different V4 sequence was assigned one name; or more than one if different species collapsed in the same V4. When retrieved with 100% similarity in the LTER-MC dataset, an abbreviation (Abbr.) code was given to the V4 sequence. The same abbreviations were used in the heatmap (**Fig.4.3.8**) and the CCA analysis (**Fig.4.3.9**). V4 reference sequences not recovered in the LTER-MC dataset were annotated with NF, i.e., not found.

N°	Abbr.	Assigned Name(s)	Same V4 sequence between *Tripos*	
			Single cell from this study	Published references
#1	**ext1**	*Tripos extensus*	SC81	FJ402957 *Tripos extensus*
#2	**ext2**	*Tripos extensus*	SC100	
#3	**fus1**	*Tripos fusus*		FJ402958 *Tripos fusus*
#4	NF	*Tripos fusus*		AF022153 *Tripos fusus*
#5	**furc1**	*Tripos furca*	SC123	FJ402966 *Tripos furca*
#6	NF	*Tripos furca*	TM74	
#7	NF	*Tripos furca*		AJ276699 *Tripos furca*
#8	NF	*Tripos pavillardii*	SC92, SC94, SC139	
#9	**Sc96**	*Tripos* sp.	SC96	
#10	**macr**	*Tripos macroceros*	SC105	
#11	NF	*Tripos carriensis*	SC130	
#12	**azo_pet**	*Tripos azoricus/ T. petersii*		FJ402951 *Tripos petersii* FJ402953 *Tripos petersii* FJ402954 *Tripos azoricus*
#13	**palikomi**	*Tripos kofoidii/ T. minutus/ T. limulus/ T. paradoxides*		FJ402963 *Tripos kofoidii* FJ402964 *Tripos minutus* FJ402952 *Tripos limulus* FJ402962 *Tripos limulus* FJ402965 *Tripos paradoxides*
#14	**arie4**	*Tripos arientinus/ T. symmetricus/ T. euarcuatus/ T. longipes*		FJ402956 *Tripos arietinus* FJ402947 *Tripos symmetricus* FJ402946 *Tripos euarcuatus* DQ388462 *Tripos longipes*
#15	**tric**	*Tripos trichoceros*	SC116, SC117, SC120, SC121, SC122.	FJ402959 *Tripos contrarius*
#16	**hor_mas**	*Tripos horridus-massiliensis*	SC93, SC102, SC114, SC118, SC134, TM63, TM71.	
#17	**mass**	*Tripos massiliensis*	S98, TM55.	FJ402942 *Tripos massiliensis*
#18	**def**	*Tripos deflexus*	SC99	
#19	**plat**	*Tripos platycornis/ "T. horridus"*		FJ824911 *Tripos platycornis* FJ402960 "*Tripos horridus*"
#20	**hex1**	*Tripos hexacanthus*		FJ402943 *Tripos hexacanthus*
#21	NF	*Tripos hexacanthus*	TM65	
#22	**cand**	*Tripos candelabrus*	TM53	FJ402955 *Tripos candelabrus*
#23	**dec_pent**	*Tripos declinatus/ T. pentagonus*	SC82, TM58	FJ402949 *Tripos declinatus* FJ402948 *Tripos pentagonus*

N°	Abbr.	Assigned Name(s)	Same V4 sequence between *Tripos*	
			Single cell from this study	Published references
#24	**pulc1**	*Tripos pulchellus*	TM68, TM72	
#25	**cont**	*Tripos contortus*	SC101, TM54, TM59	
#26	**conc**	*Tripos concilians*	SC110, TM52	FJ402944 *Tripos concilians* FJ402950 *Tripos concilians*
#27	NF	*Tripos gravidus*		FJ402961 *Tripos gravidus*
#28	NF	*Tripos digitatus*		FJ824910 *Tripos digitatus*

In comparison with other dinoflagellates, *Tripos* reads were not among the most abundant ribotypes representing only 1,36% of the reads in the 48 dates LTER-MC dataset. Few species were present and abundant all year such as *Tripos furca* (10,714 reads – 0.52 %) and *Tripos fusus* (2,726 reads – 0.13%) (**Fig.4.3.8**). Comparatively, other taxa were recovered with a low abundance during the three years like *Tripos concilians* (**conc**) or *Tripos* cf. *extensus* (**ext2**)(**Fig.4.3.8**). Overall, results showed seasonal patterns for some ribotypes (**Fig.4.3.8**). For instance, *Tripos azoricus/ T. petersii* (**azo_pet**), *T. paradoxides/ T. limulus/ T. kofoidii/ T. minutus* (**palikomi**), *Tripos* sp. (**Sc96**) or *T. arietinus, T. longipes, T. symmetricus, T. euarcuatus* (**arie4**) ribotypes displayed a winter seasonal pattern even if the V4 was shared between different species (**Fig.4.3.8**). *Tripos massiliensis* (**mass**) was predominantly detected in late summer (**Fig.4.3.8**), and *T. furca* (**furc1**), even if present almost all year, seemed most abundant in spring, summer and autumn. The same patterns were observed when analysed using CCA (**Fig.4.3.9**). Temperature and Chlorophyll *a* were the only environmental parameters related with *Tripos* abundance and distribution in time. The first axis explained 13% and the second 3% of the observed variation. However, for many ribotypes a signal was absent and no specific patterns were detected. In general, the relative abundance of *Tripos* seemed to increase over the three years.

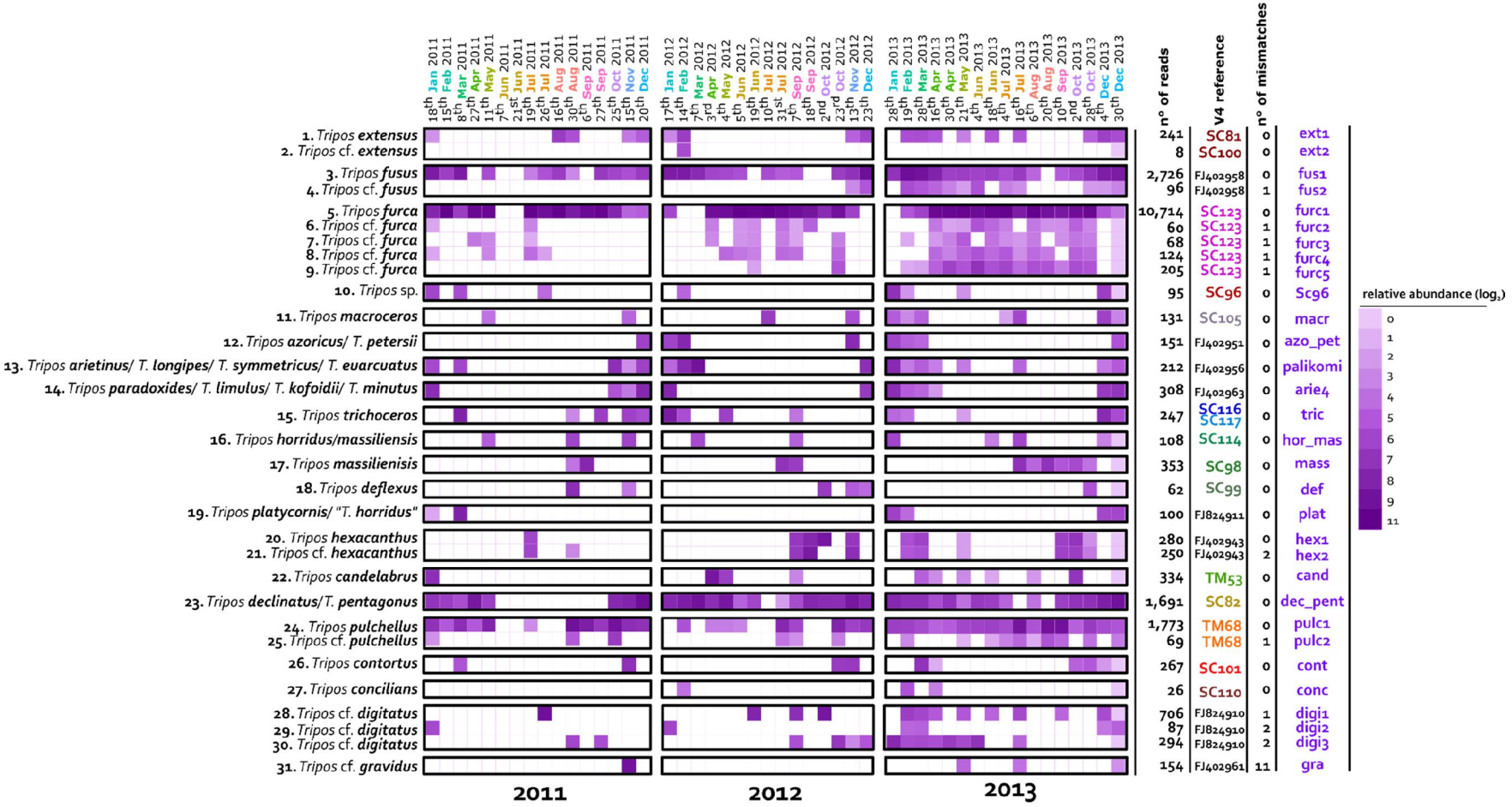


Fig.4.3.8: Heatmap showing the relative abundance of reads (log_2) for the 48 dates for the 31 *Tripos* ribotypes with an abundance of at least 50 reads or matching perfectly with a reference (100% similarity). Numbers of raw reads are specified in the first column on the right of the heatmap; V4 reference sequences matching with each ribotype in a second column, and the number of mismatches with the reference is detailed in a third column. When no *Tripos* reads were detected the cell was left white.

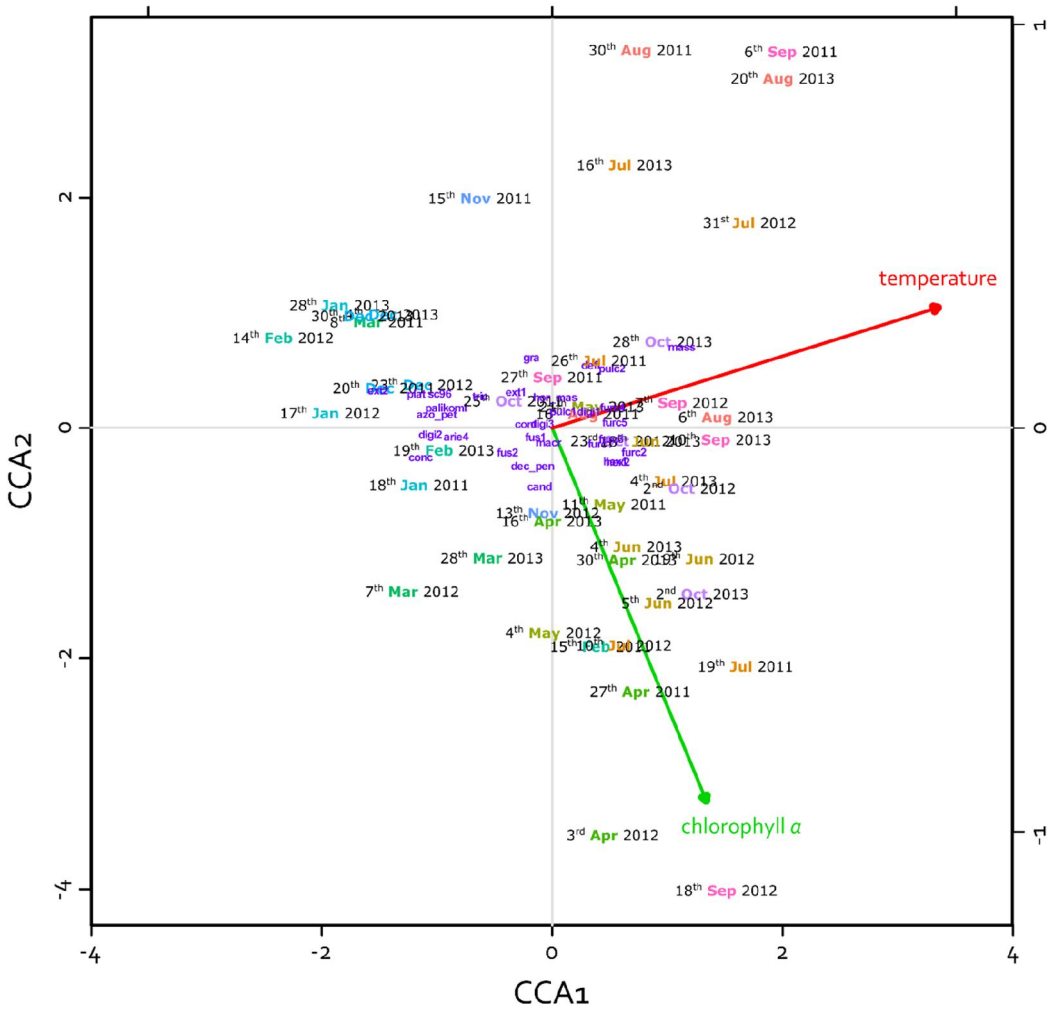


Fig.4.3.9: CCA (Canonical-Correlation Analysis) performed using *Tripos* ribotypes (>50 reads or assigned at 100% similarity, **Fig.4.3.8**) and specific environmental parameters selected through the BIO-ENV function (see 4.2. Material and methods section). Each ribotype was abbreviated in this way: 1. **ext1**: *Tripos extensus* (SC81), 2. **ext2**: *T. extensus* (SC100), 3. **fus1**: *T. fusus*, 4. **fus2**: *T. fusus*, 5. **furc1**: *T. furca*, 6. **furc2**: *T. furca*, 7. **furc3**: *T. furca*, 8. **furc4**: *T. furca*, 9. **furc5**: *T. furca*, 10. **sc96**: *Tripos* sp. (SC96), 11. **macr**: *T. macroceros* (SC105), 12. **azo_pet**: *T. azoricus/ T. petersii*, 13. **arie4**: *T. longipes/ T. arietinus/ T. symmetricus/ T. euarcuatus*, 14. **palikomi**: *T. paradoxides/ T. limulus/ T. kofoidii/ T. minutus*, 15. **tric**: *T. trichoceros* (SC116,SC117), 16. **hor_mas**: *T. horridus-massiliensis* (SC114), 17. **mass**: *T. massiliensis* (SC98), 18. **def**: *T. deflexus* (SC99), 19. **plat**: *T. platycornis/ "T. horridus"*, 20. **hex1**: *T. hexacanthus*, 21. **hex2**: *T. hexacanthus*, 22. **cand**: *T. candelabrus*, 23. **dec_pent**: *T. declinatus/ T. pentagonus* (SC82), 24. **pulc1**: *T. pulchellus* (TM68), 25. **pulc2**: *T.* cf. *pulchellus* (TM68), 26. **cont**: *T. contortus* (SC101), 27. **conc**: *T. concilians*, 28. **digi1**: *T.* cf. *digitatus*, 29. **digi2**: *T.* cf. *digitatus*, 30. **digi3**: *T.* cf. *digitatus*, 31. **gra**: *T.* cf. *gravidus*.

4.4. Discussion

The genus *Tripos* has been known for more than two centuries (Schrank, 1793) and more than 77 morpho-species including many varieties are currently described in literature (Gómez, 2013). These species are usually characterised using subtle morphological variability, but species delimitation is not always clear and no consensus is available for several species (Sournia, 1967; Jørgensen, 1911; Gómez, 2012a). In addition, morphological characterisation is often difficult when *Tripos* cells present broken arms due to isolation methods. The rise of molecular phylogeny represents a new tool to test, and possibly further refine, species identification based on informative morphological characters. Moreover, the increasing use of eDNA and metabarcoding offer an opportunity to assess diversity and monitoring of protists, including the dinoflagellate genus *Tripos*. However, in order to obtain reliable phylogenies and assessment of *Tripos* diversity, reference sequences covering the maximum diversity are needed. However, only a limited number of sequences are currently published for the genus.

As a result of this study, 11 new species of *Tripos* are now characterised genetically adding to those have been characterised in earlier studies (mainly Gómez *et al.*, 2010 and a few unpublished sequences present on GenBank), by obtaining 18S and partial 28S rRNA sequences from single cells. All these sequences were used to build phylogenies, improving the knowledge of the *Tripos* genus and confirming that molecular data largely corresponds to morphology. I also used the V4 regions of all the *Tripos* 18S sequences as reference barcodes to investigate *Tripos* diversity and variation in time in the LTER-MC dataset and the results of that exercise revealed that *Tripos* species show distinct seasonality, with most of them occurring in winter.

Morphology vs. Phylogeny

The first phylogeny investigating the diversity and evolutionary relationship of the *Tripos* genus was produced by Gómez *et al.* (2010). This study based on an 18S phylogeny (**1,137 bp**) showed a clear separation between freshwater (*Ceratium*) and the marine species. All marine *Ceratium* species, including the 27 species molecularly characterised by Gómez *et al.* (2010) were transferred

to the genus *Tripos*. Our study is the first to provide a phylogenetical analysis for the 28S rRNA gene, which is known to provide a better resolution for dinoflagellates in comparison with the 18S for which we also obtained sequences for the same cells. In both 18S and 28S phylogenies, all *Tripos* sequences clustered in a well-supported clade, confirming the separation among marine and freshwater species found by Gómez *et al.* (2010). Within the *Tripos* genus, both phylogenies produce a polytomic backbone and many internal nodes are not well-supported. The same results were obtained by Gómez *et al.* (2010) and this may result from a rapid diversification of *Tripos* species (Taylor, Hoppenrath and Saldarriaga, 2008; Wiggan, Riding and Franz, 2017). Despite the basal polytomy, both 18S and 28S show well- supported terminal clades. In general, the same terminal clades are found in both the 18S and 28S trees even if the 18S phylogeny shows a lower resolution of some clades. In addition, several of these clades group together sequences with the same morpho-species assignation, confirming that morphological differences corroborate molecular differences. In the same way, some of the sequences produced in this study cluster with sequences obtained by Gómez forming well-supported clades of sequences of the same morpho-species (i.e. *T. candelabrus*, *T. hexacanthus*, *T. furca* or *T. massiliensis*). In contrast, in the 18S tree, our sequences assigned to *T. trichoceros* (SC116, SC120 and SC121) group the sequence *T. contrarius* from Gómez. This finding supports the hypothesis that the two species are synonyms (Steidinger and Tangen, 1997).

The main difference between the 18S and 28S rRNA phylogenies is the basal position in the 18S phylogeny of *T. fusus*, *T. extensus, T. gravidus* (not present in 28S) and *T. digitatus* (not present in 28S), which form a supported clade (0.90 posterior probabilities) in the 28S phylogeny. All these species, clustering in a clade, possess a "modified" antapical horn and a cell shape very different from the "classical" anchor shape. I suggest that the basal position detected with the slower evolving 18S rRNA, better reflects the ancestral origin of the clade. This result is supported by Gómez *et al.,* (2010), who also reported the same tree topology.

In addition, the 18S tree also showed low resolution for some sequences assigned morphologically to distinct species and produced by Gómez were placed in a polytomy. A possible explanation for

this low resolution may result from the fact an alignment of 1,137 bp is not sufficient to discriminate between some species. Our observations show that when the alignment is longer including V2 and V3 regions such as in **Fig.4.3.3**, the resolution of the trees increases.

In contrast, the 28S phylogeny does not show good resolution for the basal position of *T. fusus* and *T. extensus*, but almost all terminal groups can be discriminated clearly. Remarkably, and as predicted by Smetacek (2012) cryptic diversity was discovered inside some clades for which all cells had the same assignation and could not be distinguished based on the morphology (i.e. *Tripos horridus-massiliensis*, *T. massiliensis* or *T. trichoceros*). In these particular cases, the genetic diversity may be intraspecific (i.e different varieties of *Tripos horridus* such as *Tripos horridus* var. *horridus* or *Tripos var. buceros*). We are unable to give a specific name to the *T. horridus-massiliensis* due to confusing morphology. In order to further characterise the identity of this taxa a range of morphological measurements should be performed on a greater number of cells.

The 28S rRNA region allowed the distinction between *T. pentagonus* and *T. declinatus* in contrast with the 18S, which clusters together these two morphologically distinct species. Genetic characterisation also allowed discrimination between closely related species with similar morphologies such as *T. declinatus* vs. *T. contortus* or *T. pulchellus* vs. *T. muelleri*. The tentative species identification of 15 cells out 52 was amended based on the phylogeny, underlining the fact that the morphological characterisation of *Tripos* species is challenging and boundaries among species are not well defined.

Globally, we noticed a morphological coherency for some groups of clades at species but also at higher levels such as groups of species in both phylogenies. For example, *T. extensus* and *T. fusus* group together and both display long apical and antapical horns sequences. In the same way, *T. pavillardii*, *Tripos* sp. (**Sc96**), *T. macroceros* (only in 18S phylogeny), and *T. carriensis* all present a similar left anterior horn with different degrees of inclination and always cluster in the same clade.

Limitations of the study

Tripos is a well-known and common genus of dinoflagellates. Its anchor-shape and its size make the taxa recognisable but most of them are not easy, if not impossible to cultivate. For this reason, most molecular investigations on the genus have been carried out on single cells. Yet, single cell isolation, amplification and sequencing involves a long, meticulous and risky procedure, prone to contamination and rarely producing high success rates. Starting from a low quantity of DNA obtained from single isolated *Tripos* cells, we produced sequences for 52 cells for two different markers (37 sequences for 18S and 46 sequences for 28S rRNA). This number is low considering the total number of cells isolated (around 150). The extraction protocol was improved throughout the experiments by Thomas Mollica (see his Master thesis, Mollica, 2017). However, these 52 molecularly characterised cells, covering 22 morpho-species, represent a net increase of the molecular information available for *Tripos* and contribute to building better phylogenies, improving the knowledge of this genus. Overall, this study provided precious reference sequences for 11 new and different *Tripos* lineages in 18S rRNA, as well as 19 different new lineages in 28S rRNA, for which molecular information was not available.

However, a part of the *Tripos* diversity occurring in the Gulf of Naples still remains to be characterised. Since 1984 a total of about 50 *Tripos* taxa have been reported at LTER-MC (D. Sarno, personal communication). One of the limits of this study is that *Tripos* cells occur at low densities and authors (Tunin-Ley *et al.*, 2007) recommend to sampling on at least a volume of 70 L to assess species diversity for *Tripos* because biodiversity estimates depend mainly on sampling method and sampling effort. A net sample is taken from hundreds of litres of seawater but only a small part of a dense net sample ends up in the sample container to be examined. And from that sample container only a Pasteur pipette-volume is examined in LM for cell isolation. So, chances are that species occurring at low densities are missed when searching for specimens in such a small sample. Therefore, only the most abundant species present in samples have a good chance to be observed and isolated.

Another issue centers on damage caused by net sampling on *Tripos* cells. Often, cell horns are broken during collection rendering morphological characterisation challenging. This is the case for SC96 for which we obtained a molecular signature, but we were unable to make a precise species assignation. More sampling of identical cells is needed to overcome this problem and finally provide species names.

Detection of *Tripos* species at LTER-MareChiara

All 18S rRNA references for *Tripos* species were used to investigate the genetic diversity in the Gulf of Naples using a metabarcoding approach. The LTER-MC presents a unique opportunity to study the variation of *Tripos* species over a 3-year dataset (2011-2013) based on the V4 marker. A total of 28 different V4 reference barcode were available to explore the dataset including 11 *Tripos* species newly isolated at the same site (LTER-MC). As expected based on the 18S rRNA phylogeny, the V4 region alone shows low resolution within several *Tripos* clades and some genotypes cannot be discriminated (i.e **palikomi**: *T. paradoxides, T. limulus, T. kofoidii* and *T. minutus* or **arie4**: *T. arietinus, T. longipes, T. symmetricus* and *T. euarcuatus*). The addition of new *Tripos* sequences produced in this study confirmed the limited capacity of the V4 region to discriminate species, already observed based on published references (see **Chapter II**: variation of V4 at genus level).

Despite this limitation, 31 different V4 ribotypes were recovered from the LTER-MC dataset showing large diversity of *Tripos* species in the Gulf of Naples as previously reported by taxonomists using LM (counts courtesy Diana Sarno,Tunin-Ley *et al.*, 2007, 2009). Notably, the V4 signatures of four single cells produced in this study were not detected in LTER-MC dataset. These results can probably be explained by the fact that some isolated cells are large (>100 µm) such as *T. pavillardii* or *T. carriensis* and cannot be retrieved with the sampling methods used to generate the environmental V4 dataset. Only three litres of sea water were collected and filtered weekly (see Chapter III) to produce the V4 dataset while 70 liters would be needed to cover *Tripos* diversity (Tunin-Ley *et al.*, 2007). Therefore, it is possible that the four V4 sequences not recovered in LTER-MC dataset are part of the rare diversity occurring in some years in the Gulf of Naples and these

were not detected. Finally, we cannot totally exclude the possibility of sequencing mistake in the V4 sequences produced in this study.

The exploration of the V4 gives us an insight into the diversity that is still awaiting to be characterised with molecular information. Indeed, 52 Tripos taxa are reported to occur in the Gulf of Naples (counts data checklist for LTER-MC station, courtesy Diana Sarno). Interestingly, out of 28 different V4 retrieved from the LTER-MC, 11 were not assigned at 100% similarity with a reference but still cluster close to a reference (i.e. generally 1 or 2 nucleotides difference) and display an abundance of more than 50 reads. These species represent unknown or uncharacterised *Tripos* diversity, which remains to be described genetically. Single cell isolation, amplification and sequencing are still required to produce more reference sequences to fully characterise this *Tripos* diversity.

Species diversity and temporal patterns based on morphological observations in light microscopy were published by several authors especially for the western part of the Mediterranean Sea for which long-term records are available (Tunin-Ley *et al.,* 2007; Tunin-Ley *et al.,* 2009, Aubry *et al.,* 2012). These studies discussed and compared the categorisation of *Tripos* species (and varieties) based on biogeographical distributions and thermal affinities observed by many scientists such as Sournia (1967), Dodge and Marshall, (1994) and Semina and Levashova (1993). Our study is the first to test metabarcoding to explore temporal variability of *Tripos* species on a relatively long-time scale. The results obtained from this analysis show interesting seasonal patterns for part of the 31 ribotypes detected at LTER-MC suggesting a preference for winter conditions (i.e. *Tripos* sp. (**SC96**), *T. contortus* (**cont**) or *T. concilians* (**conc**)). Also, some V4 ribotypes regroup different species, i.e. *T. azoricus/ T. petersii* (**azo_pet**), *T. paradoxides/ T. limulus/ T. kofoidii/ T. minutus* (**palikomi**), or *T. arietinus, T. longipes, T. symmetricus, T. euarcuatus* (**arie4**) present a winter seasonal pattern. In these cases, it is impossible to discriminate if all or some of the species sharing the same V4 region co-occur in the samples or if the seasonal pattern is due to a single species. Overall, the relative abundance of the different ribotypes shows that several species occur

predominantly in winter confirming a relationship between species presence and temperature. These results are in agreement with Tunin-Ley and colleagues (2007, 2009) who also reported the presence of *T. contortus*, *T. concilians* or *T. paradoxides* in winter in Villefranche Bay, although Dodge and Marshall (1994) classified these species as tropical taxa (i.e. species rarely found when the temperature goes below 20 °C). In the same way, species such as *T. declinatus*, *T. pavillardii* or *T. ranipes* detected in February and March in Villefranche, were also isolated in the same months in Naples even if at times no V4 signature was detected in the HTS dataset, probably due to their low abundance or big size. This is consistent with Dodge and Marshall's (1994) categories, which define *T. declinatus*, *T. pavillardii* or *T. ranipes* as "Warm-temperate-tropical" taxa with temperature boundaries of 14-15°C, which corresponds to the winter temperature in the Gulf of Naples.

As reported by several authors before (Halim, 1960; Gómez and Gorsky, 2003; Tunin-Ley *et al.*, 2007; Aubry *et al.*, 2012), *T. furca* was the dominant species in the LTER-MC dataset with abundance of 10,714 reads. Other *T. furca* ribotypes were also detected (**furc2**, **furc3**, **furc4** and **furc5**) in the LTER-MC dataset, only diverging one nucleotide difference from the dominant ribotype. As discussed in **Chapter III**, these secondary ribotypes are probably an expression of infra-genomic diversity since the temporal variation of these secondary ribotypes almost perfectly follows the temporal variation of the principal ribotype.

The second most abundant ribotype at LTER-MC is *T. fusus* with 2,726 reads. These two abundant species co-occur during the year but while *T. furca* seems to be most abundant in spring and summer, *T. fusus* seem to dominate in colder season. This result matches the temporal trend found for this two species in Villefranche by Tunin-Ley *et al.*, (2007).

Other *Tripos* species such as *T. massiliensis* (**mass**) seem to thrive mainly in late summer – autumn. The same trend is reported for *T. trichoceros* in surface water by Aubry and collaborators (2012) in the Adriatic Sea. Yet in Naples, *T. trichoceros* shows a winter-distribution. Since *T. massiliensis* and

T. trichoceros are morphologically similar, I can hypothesise that these species can be easily confused.

Besides trends described above, no significant seasonal pattern could be detected for some *Tripos* ribotypes. While the presence of some *Tripos* species seems to be linked to water temperature, others did not show any thermal preference (Tunin-Ley *et al.*, 2007, 2009). These observations could be explained in two different ways and probably both cases occur in our case. i) Some *Tripos* species are independent from temperature. Perennial and almost perennial species, as for example *Tripos furca*, *T. fusus* and *T. pentagonus* were already identified at LTER-MC in previous studies based on morphological analyses (Tunin-Ley *et al*., 2009). ii) The V4 region does not permit discrimination of different species occurring at different time during the year and for this reason the pattern of sinlge species can not be distinguished. This could be, for example, the case of *T. horridus*, which was reported as absent from surface waters during the warm period (Tunin-Ley *et al*., 2009).

The presence of different *Tripos* species could also be related to peculiar hydrographic conditions and different water masses at LTER-MC. As described by D'Alelio and colleagues (2015), seasonal shifts occur regularly at LTER-MC, the station swinging between coastal eutrophic or offshore oligotrophic influence. In winter LTER-MC is characterised by a sharp alternation between the currents that favour retention of surface waters and winds that generate a rapid renewal of coastal waters with off-shore waters (Cianelli *et al.*, 2015; 2017). This type of circulation, coupled with the enahnced vertical mixing typical of winter season, probably favours the presence in surface waters of species that are preferentially encountered in deeper waters (Tunin-Ley *et al*., 2009). However, our data supports the idea that *Tripos* species thriving in winter are mainly influenced by temperature.

Given the sensitivity of different *Tripos* species to temperature, our results support the idea that *Tripos* could be tested as a world-wide ecological indicator to monitor global warming (Tunin-Ley and Lemée, 2013). Other studies have already showed that even a low increase in water

temperature seems to have a positive effect on *Tripos* growth (Li *et al.*, 2011; Vázquez-Domínguez, Vaqué and Gasol, 2012). It has been shown that the distribution of *T. furca* extended northward as a possible consequence of climate change (Edwards *et al.*, 2006) suggesting monitoring of *Tripos* species in order to anticipate their response to climate change. In our case, metabarcoding analysis seems to show an increase of *Tripos* species abundance between 2011 and 2013. This trend should be confirmed by the analysis of more data over a longer time period in the coming years.

I think that metabarcoding is a powerful tool to assess and monitor *Tripos* diversity despite the limited resolution power of the V4 region for some of the *Tripos* species. As previously discussed in **Chapter III**, HTS metabarcode data for dinoflagellates should be implemented with other markers such as the 28S in order to obtain more resolution (Grzebyk *et al.*, 2017). By producing single cell sequences we increase the knowledge of the diversity of *Tripos* genus and provide a reliable tool – a genetic signature to harmonise and compare data in space and time.

Annex 1: List of species and their taxonomic authorities as obtained from Guiry and Guiry, (2017) accessed 21 Dec. 2017.

Tripos aequatorialis (Schröder) F. Gómez

Tripos aestuarius (Schröder) F. Gómez

Tripos allieri (Gourret) F. Gómez

Tripos angustocornus (N. Peters) F. Gómez (Uncertain taxonomic status)

Tripos angustus (A.S. Campbell) F. Gómez

Tripos arcticus (Vanhöffen) F. Gómez

Tripos arcticus var. *ventricosus* (Ostenfeld) F. Gómez (Uncertain taxonomic status)

Tripos arietinus (Cleve) F. Gómez

Tripos aultii (H.W. Graham & Bronikovsky) F. Gómez C

Tripos axialis (Kofoid) F. Gómez

Tripos azoricus (Cleve) F. Gómez

Tripos balechii (Meave del Castillo, Okolodkov & M.E. Zamudio) F. Gómez

Tripos balticus (F.Schütt) F. Gómez

Tripos batavus (Paulsen) F. Gómez

Tripos belone (Cleve) F. Gómez

Tripos berghii (Gourret) F. Gómez

Tripos biceps (Claparède & Lachmann) F. Gómez

Tripos bicornis (Gourret) F. Gómez

Tripos bigelowii (Kofoid) F. Gómez

Tripos boehmii (H.W. Graham & Bronikovsky) F. Gómez

Tripos brevis (Ostenfeld & Johannes Schmidt) F. Gómez

Tripos brunellii (Rampi) F. Gómez

Tripos bucephalus (Cleve) F. Gómez

Tripos buceros (Zacharias) F. Gómez

Tripos californiensis (Kofoid) F. Gómez

Tripos candelabrum (Ehrenberg) F. Gómez

Tripos carnegiei (H.W.Graham & Bronikovsky) F. Gómez

Tripos carriensis (Gourret) F. Gómez

Tripos cephalotus (Lemmermann) F. Gómez

Tripos ceylanicus (Schröder) F. Gómez

Tripos claviger (Kofoid) F. Gómez

Tripos coarctus (Pavillard) F. Gómez

Tripos compressus (Gran) F. Gómez

Tripos concilians (Jørgenen) F. Gómez

Tripos contortus (Gourret) F. Gómez

Tripos contrarius (Gourret) F. Gómez

Tripos curvicornis (Daday) F. Gómez

Tripos dalmaticus (Schröder) F. Gómez

Tripos declinatus (G. Karsten) F. Gómez

Tripos deflexus (Kofoid) F. Gómez

Tripos dens (Ostenfeld & Johannes Schmidt) F. Gómez

Tripos denticulatus (Jörgenen) F. Gómez

Tripos depressus (Gourret) F. Gómez

Tripos digitatus (F. Schütt) F. Gómez

Tripos dilatatus (Gourret) F. Gómez

Tripos divaricatus (Lemmermann) F. Gómez

Tripos egyptiacus (Halim) F. Gómez

Tripos ehrenbergii (Kofoid) F. Gómez

Tripos elegans (Schröder) F. Gómez

Tripos euarcuatus (Jörgenen) F. Gómez

Tripos eugrammus (Ehrenberg) F. Gómez

Tripos extensus (Gourret) F. Gómez

Tripos falcatiformis (Jörgenen) F. Gómez

Tripos falcatus (Kofoid) F. Gómez

Tripos filicornis (Steemann Nielsen) F. Gómez

Tripos flagelliferus (Cleve) F. Gómez

Tripos furca (Ehrenberg) F. Gómez

Tripos fusus (Ehrenberg) F. Gómez

Tripos gallicus (Kofoid) F. Gómez

Tripos geniculatus (Lemmermann) F. Gómez

Tripos gibberus (Gourret) F. Gómez

Tripos globatus (Gourret) F. Gómez

Tripos globosus (Gourret) F. Gómez

Tripos gracilis (Pavillard) F. Gómez

Tripos gravidus (Gourret) F. Gómez

Tripos heterocamptus (Jörgenen) F. Gómez

Tripos hexacanthus (Gourret) F. Gómez

Tripos hircus (Schröder) F. Gómez

Tripos horridus (Cleve) F. Gómez

Tripos humilis (Jörgenen) F. Gómez

Tripos hundhausenii (Schröder) F. Gómez

Tripos hyperboreus (Cleve) F. Gómez

Tripos incisus (Karsten) F. Gómez

Tripos inclinatus (Karsten) F. Gómez

Tripos inflatus (Kofoid) F. Gómez

Tripos inflexus (Gourret) Gómez U

Tripos intermedius (Jörgenen) F. Gómez

Tripos inversus (Karsten) F. Gómez

Tripos japonicus (Schröder) F. Gómez

Tripos karstenii (Pavillard) F. Gómez

Tripos kofoidii (Jörgenen) F. Gómez

Tripos lamellicornis (Kofoid) F. Gómez

Tripos lanceolatus (Kofoid) F. Gómez

Tripos leptosomus (Jörgensen) F. Gómez

Tripos limulus (Pouchet) F. Gómez

Tripos lineatus (Ehrenberg) F. Gómez

Tripos longinus (Karsten) F. Gómez

Tripos longipes (J.W.Bailey) F. Gómez

Tripos longirostrus (Gourret) F. Gómez

Tripos longissimus (Schröder) F. Gómez

Tripos lunula (Schimper ex Karsten) F. Gómez

Tripos macroceros (Ehrenberg) F. Gómez

Tripos massiliensis (Gourret) F. Gómez

Tripos minor (Gourret) Gómez

Tripos minutus (Jörgensen) F. Gómez

Tripos mollis (Kofoid) F. Gómez

Tripos muelleri Bory C - type

Tripos neglectus (Ostenfeld) F. Gómez

Tripos obesus (Pavillard) F. Gómez

Tripos obliquus (Gourret) F. Gómez

Tripos obtusus (Gourret) F. Gómez

Tripos okamurae (Schröder) F. Gómez

Tripos orthoceras (Jörgensen) F. Gómez

Tripos ostenfeldii (Kofoid) F. Gómez

Tripos oviformis (Daday) F. Gómez

Tripos pacificus (Schröder) F. Gómez

Tripos palmatus (Schröder) F. Gómez

Tripos paradoxides (Cleve) F. Gómez

Tripos parvus (Gourret) F. Gómez

Tripos patentissimus (Ostenfeld & Johannes Schmidt) F. Gómez

Tripos pavillardii (Jørgensen) F. Gómez

Tripos pellucidus (Gourret) F. Gómez

Tripos pennatus (Kofoid) F. Gómez

Tripos pentagonus (Gourret) F. Gómez

Tripos petersenii (Steemann Nielsen) F. Gómez

Tripos petersii (Steemann Nielsen) F. Gómez

Tripos platycornis (Daday) F. Gómez

Tripos porrectus (Karsten) F. Gómez (Uncertain taxonomic status)

Tripos praelongus (Lemmermann) Gómez

Tripos procerus (Gourret) F. Gómez

Tripos protuberans (G. Karsten) F. Gómez

Tripos pulchellus (Schröder) F. Gómez

Tripos ramakrishnae (Subrahmanyan) F. Gómez

Tripos ranipes (Cleve) F. Gómez

Tripos recurvatus (Schröder) F. Gómez

Tripos recurvus (Jørgesen) F. Gómez

Tripos reflexus (Cleve) F. Gómez

Tripos reticulatus (Pouchet) F. Gómez

Tripos robustus (Ostenfeld & Johannes Schmidt) F. Gómez

Tripos rostellus (Gourret) F. Gómez

Tripos saltans (Schröder) F. Gómez

Tripos scapiformis (Kofoid) F. Gómez

Tripos schmidtii (Jørgesen) F. Gómez

Tripos schrankii (Kofoid) F. Gómez

Tripos schroederi (Nie) F. Gómez

Tripos schroeteri (Schröder) F. Gómez

Tripos semipulchellus (Jørgesen) F. Gómez

Tripos seta (Ehrenberg) F. Gómez

Tripos setaceus (Jørgesen) F. Gómez

Tripos strictus (Okamura & Nishikawa) F. Gómez

Tripos subcontortus (Schröder) F. Gómez

Tripos subrobustus (Jørgesen) F. Gómez

Tripos subsalsus (Ostenfeld) F. Gómez

Tripos sumatranus (Karsten) F. Gómez

Tripos symmetricus (Pavillard) F. Gómez

Tripos tasmaniae (E.J.F.Wood) F. Gómez

Tripos tenuis (Ostenfeld & Schmidt) F. Gómez (Uncertain taxonomic status)

Tripos tenuissimus (Kofoid) F. Gómez

Tripos teres (Kofoid) F. Gómez

Tripos tricarinatus (Kofoid) F. Gómez

Tripos trichoceros (Ehrenberg) Gómez

Tripos tripodioides (Jørgesen) F. Gómez

Tripos truncatus (Lohmann) F. Gómez

Tripos uncinus (Sournia) F. Gómez

Tripos uteri (A.S. Campbell) F. Gómez

Tripos varians (Mangin) F. Gómez

Tripos volans (Cleve) F. Gómez

Tripos vultur (Cleve) F. Gómez

CHAPTER V: Conclusions and outlook

5. Conclusions and outlook

In my PhD studies, I set out to explore different aspects of dinoflagellate diversity. I was interested in this group of organisms because of their diversity in overall cell morphology, their range of ecological strategies (autotrophy, mixotrophy, heterotrophy), interactions with other organisms (symbiosis, parasitism, herbivory), and the capacity of many of their species to produce toxins and form harmful algal blooms. I focused my study on the Long Term Ecological Research station MareChiara (LTER-MC) in the Gulf of Naples, which is one of the few Mediterranean sites that regularly monitors marine plankton. The knowledge generated at this site is comprised of: i) weekly records of phytoplankton species diversity monitored in LM, ii) detailed taxonomic and population genetic studies mainly focused on key diatom genera, and iii) ecological studies investigating the structure and functioning of planktonic communities in relation to the environmental variability and climate change. Over the last decade, HTS metabarcode data has been added to this body of data, and it is this type of data that I explored further to study dinoflagellate diversity season to season.

Chapter II

To accurately identify the metabarcodes, a curated dataset of 18S reference barcodes was required. Since reference datasets available at the onset of my study were incomplete and contained many nomenclatural errors, I set out to compile such a dataset myself; the result is DinoREF (Chapter II). DinoREF represents an updated and validated repository of 18S rRNA sequences made available for the scientific community. The database included 1,671 sequences of dinoflagellates representing 149 genera and 422 species. DinoREF now covers 22% of the total dinoflagellate described species. In addition, DinoREF allowed for the checking of how comprehensive the V4 primers amplify and HTS sequence the entire dinoflagellate diversity into metabarcodes, i.e., how well they fit their intended target regions in the 18S sequences across the dinoflagellate diversity and how thoroughly the V4 regions can discriminate all the known dinoflagellate species. Out of 1,671 sequences in DinoREF, 946 unique V4 sequences were

obtained. The V4 region could unequivocally discriminate 374 of the 422 species. Among the species and genera sharing the same V4 marker, several toxic ones could not be distinguished. Moreover, a significant number of morphologically and genetically distinct taxa (species, and even genera) exhibit V4 sequences that cannot be discriminated at the 98% similarity level; a level usually applied to cluster sequences OTUs. This implies that with the further development of HTS technology, enabling the sequencing of longer markers, longer regions in the 18S or even faster evolving markers such as the 28S rRNA should be considered as reference barcode for metabarcoding studies.

Chapter III

In Chapter III, I assessed dinoflagellate diversity in 48 samples taken at the LTER over the seasonal cycles in three consecutive years (2011 – 2013) using a metabarcoding approach. The V4 variable region in the 18S rDNA was used as a metabarcode because this region has been the one of choice in many such studies and because the ca. 380 bp sequences can be obtained with current HTS technology (Illumina). First, I performed an ataxonomic cluster analysis in which I clustered samples in a hierarchical fashion based on their sequence composition. Result showed that samples clustered into two principal groups in which the winter samples (16% of the sequences) grouped in one cluster, and the remainder of the samples clustered into two other clusters. One of these contained the spring and early summer samples with a few autumn and winter samples added (62% of the sequences) and the other cluster the late summer-autumn samples (22% of the sequences). The HTS sequences sorted into ribotypes were taxonomically assigned to species and higher rank taxa with the help of the DinoREF reference sequences as queries. Overall, the dataset was dominated by the *Gyrodinium* Superclade, the Gymnodiniales Superclade and the Gonyaulacales Superclade. The winter samples showed a high dinoflagellate diversity and were characterised by the presence of very specific taxa occurring only in winter (e.g. some parasites, symbionts or specific Gymnodiniales dinoflagellates such as Warnowiceae). Within each of the Superclades some of the genera were very common and occurred year-round whereas others were

seasonal. The genera that occurred year-round were represented usually by different species in different periods of the seasonal cycle.

Chapter IV

A comparison between the species represented in DinoREF and all the morphologically described dinoflagellate species revealed that many taxonomic groups are still underrepresented in the 18S rDNA or even are missing altogether. Examples of such taxa are parasites because they have to be maintained together with their hosts and are cumbersome to isolate and maintain in steady culture. Likewise, heterotrophs and facultative autotrophs are difficult to maintain because they need their source(s) of food, which are often unknown. In addition, large and conspicuous dinoflagellates often grow only very slowly, if at all, in culture, and finding out what is needed to grow them is a topic of a PhD study of its own. One of these taxa constitutes *Tripos,* the target genus of **Chapter IV.** The importance of the genus lies in the fact that it is considered as a possible indicator of global warming (Tunin-Ley and Lemée, 2013). *Tripos* species display very variable morphology and the classification of the genus is still based on these morphological characteristics which has been shown to vary infra and inter specifically. Therefore, the identification of phylogenetically significant characters useful to set species delimitation is needed. Yet, the amount of species characterised molecularly is still low.

I gathered specimens belonging to this genus from LTER-MC samples and deployed a culture-free method to gather reference barcodes from these specimens, i.e., I took an image in LM from each individual and then applied a single-cell DNA extraction – PCR – sequencing protocol to gather a partial 18S sequence including the V4 region and a partial 28S sequence. Results revealed 22 genetically distinct species of which 11 were not yet characterised molecularly (28S and 18S). All sequences obtained in this study and already published *Tripos* references were used to build phylogenies for both 18S and 28S markers. Remarkably, phylogenies confirmed that morphological variation is reflected in the phylogenetic relationship with similar morpho-species clustering in the same clades. However, some cryptic diversity was also detected for some taxa

such as for the *Tripos massilensis* clade. 28S phylogeny mainly offered a better resolution than the 18S even if both phylogenies had a polytomic backbone. Using the obtained V4 sequences I assessed the seasonal abundances of these species; some were common all year round whereas others showed distinct seasonality, mainly occurring in winter.

Issues that can be raised with all data taken together

DinoREF useful elsewhere?

The DinoREF is composed of 18S reference sequences of dinoflagellate strains gathered from various sites worldwide. Using this dataset, most of the dinoflagellate LTER-MC metabarcodes could be identified down to the species or generic level. However, many morphologically characterised species are not present in DinoREF. Understandably, there is a need to characterise more species and improve diversity coverage from the Gulf of Naples as well as from other geographical areas, especially tropical regions. The tropics include over half of the coastal regions on earth and are still seriously under-sampled for phytoplankton, including dinoflagellates. Results of phytoplankton diversity studies carried out in these regions show high diversity and many species and genera appear to be typical for the tropics. Therefore, taxonomic efforts need to be focused on different geographical regions to make DinoREF, and other such reference datasets globally applicable.

Metabarcode results in accordance with common principles on dinoflagellate occurrence?

The results of my study confirm LM observations over the years that dinoflagellates are highly diverse in the Gulf of Naples. However, the fact that this lineage is particularly diverse in the winter season and that it is the most abundant in spring-beginning of summer seems to contrast the model of phytoplankton succession that Ramon Margalef proposed that dinoflagellates are typical for nutrient-depleted, stratified summer conditions. A possible explanation for this discrepancy between my result and the Margalef model is that only a restricted number of generally large, but conspicuous, species are common for such summer conditions. The knowledge of dinoflagellate diversity and biology in general has greatly increased and newly discovered dinoflagellates were

shown to display very different temporal occurrences. This knowledge still remaines to be integrated in a new model considering all dinoflagellate diversity and their specific niches.

Winter conditions characterised by turbulence and thorough mixing of the water column, show high dinoflagellate diversity but low abundances of each of the species. Moreover, many winter species are smallish parasites or symbionts, morphologically inconspicuous ones that are generally pooled into unspecified categories in routine phytoplankton counting. Our results show that some specific parasitic dinoflagellates seem to be typical for the winter, infecting a range of organisms occurring in that season. In addition, temporal variation of *Tripos* species obtained from HTS data also reflects LM observation identifying higher diversity in winter.

In contrast, the seasonal patterns observed at LTER-MC also confirmed some temporal phytoplankton trends. For example, small autotrophic thecate dinoflagellates dominate in spring – beginning of the summer co-occurring with diatom blooms whereas the following season, i.e. summer and late summer, is mainly dominated by large heterotrophic dinoflagellates.

Societal relevance, toxic dinoflagellates at the LTER

Reference sequences of potentially toxic dinoflagellate species are well represented in DinoREF because such species draw societal and scientific attention. So, this reference dataset can aid signalling of toxic species when deploying environmental metabarcoding. Among the potentially toxic species detected this way in the metabarcodes generated from environmental samples taken at the LTER-MC, several were not previously known to occur in the Gulf of Naples. This is important knowledge relevant for the regional shellfish farming. Those who monitor the plankton in light microscopy can be made aware of their morphology and be sensitised to their possible occurrence. Moreover, the results illustrate that metabarcoding enables detection of all the toxic species present in a sample in one and the same experiment. It is superior to screening methods using probes on microchips because the metabarcode sequences to be identified are much longer than the probe sequences, allowing more precise identification, there are no issues with hybridisation conditions, and toxic species are detectable, not only those for which a probe is present on the

microchip. However, as stated by the V4 analysis (**Chapter II and III**), a few toxic species share the same V4 or could not be distinguished unambiguously from non-toxic species. These ambiguities are specified in Chapter II and III and should be taken in account for future matabarcoding studies using the V4 18S rRNA region.

Yet in order to become a standard tool for the rapid detection of toxic species, the detection procedure needs to be automatised and become more speedy, meaning that sampling, DNA extraction, HTS, and downstream screening of the raw data for the presence of metabarcodes belonging to toxic species needs to be cast in standard operational procedures performed in rapid succession.

Main contributions and future perspectives of this PhD thesis

In this thesis, I attempted to characterise dinoflagellate diversity in the Gulf of Naples through different techniques involving microscopy, isolation of single cells, biomolecular labwork and bioinformatic treatment of metabarcoding data. One of my main achievements was to gather, filter, validate, annotate and organise all the dinoflagellate 18S rRNA sequences available in GenBank to create the DinoREF reference database (**Chapter** II). DinoREF has been the benchmark for the analyses of the other chapters in this thesis and allowed me to review dinoflagellate literature and learn about dinoflagellate diversity. DinoREF is easy to use and provides all metadata necessary for ecological analyses.

I believe that DinoREF can be an extremely useful tool for many researchers worldwide. With the popularisation of sequencing technologies, a growing number of research centres began investing in metagenomics analyses to characterise protist diversity. However, until DinoREF, no good quality reference database existed to analyse this data for dinoflagellates.

The creation of DinoREF allowed me to detect a large number of mistakes and non-updated names for dinoflagellate sequences on GenBank (more than 25% of the database). Mistakes and curation of names of sequences (in the title) published on Genbank can only currently be corrected by the

authors of each sequences themselves, even if researchers can signal mistakes to GenBank administrators. This will probably be an issue in the future with the exponential accumulation of molecular data (and mistakes) in public depository. In my case, I and DinoREF coauthors decided to send feedback to GenBank and highlighted the curations made in DinoREF. Other initiatives such as UniEuk and EukRef (Berney *et al.*, 2017; del Campo *et al.*, 2018) are also currently developing other tools to update protist classification. I hope that my work and those of other researchers on protists (e.g. Decelle *et al.*, 2015, Morard et *al.*, 2015,) will lead to stronger collaboration between taxonomic expert groups and public databases – in order to update and curate regularly public molecular data.

DinoREF is currently accessible online on Figshare, but all data was also integrated to the latest version (v.4.9) of the PR^2 database in order to be spread as much as possible. To follow updates and to integrate new reference 18S rRNA sequences into the database, I will try to upload new versions of DinoREF periodically on Figshare. Another future development could be the creation of a website for DinoREF which would be an interactive reference platform for dinoflagellates. In the same way, I would like to use the same pipeline developed for the 18S rRNA gene to create a database dedicated to the 28S rRNA. The V4 18S rRNA is one of the main markers used for protist metabarcoding but the 28S rRNA (D1-D2) is also often used as a comparison. For dinoflagellates, the 28S rRNA is known to be more resolutive (Murray *et al.*, 2005) and showed promising result as a metabarcode (Grzebyk, *et al.*, 2017; Smith et *al.*, 2017). The use of both markers simultaneously would help to better interpret metabarcoding in characterising dinoflagellate diversity.

Through the creation of DinoREF, I was also able to analyse the variability of the V4 region for an important number of dinoflagellate sequences and species. These analyses revealed that some dinoflagellates species (and in rare cases genera) shared the same V4 sequence and that the V4 region would often differ only by a few base pairs between species belonging to the same genus (**Chapter II – Fig.2.3.5**). I demonstrated that, as a consequence of this, most of the clustering methods used to create OTUs for the analysis of HTS data for protists, collapsed dinoflagellate

diversity and in some cases grouped species from various lineages in the same OTUs. These findings convinced me to use the "ribotypes" to analyse HTS data obtained at the LTER-MC station and without performing OTU clustering. This practice goes against all similar studies on protists which usually cluster diversity with 97% or 98% similarity OTUs in order to reduce HTS data main biases (e. g. sequencing errors, multi-copy and intragenomic diversity). Nonetheless, the results of the HTS analyses at the LTER-MC (**Chapter III**) supported the use of the ribotypes in the study of dinoflagellate diversity. The use of ribotypes allowed to detect a strong seasonal pattern (3 clusters –winter – spring-mixed and late summer – autumn) and identify different species of dinoflagellate which would have been grouped together with a "classical" clustering approach. I would strongly recommend this approach for any future study using the V4 barcode to characterise dinoflagellate diversity.

On another hand, the LTER-MC dataset seemed to be extremely biased toward large and naked dinoflagellates, *Gyrodinium* species, Gymnodiniales, Ptychodiscales and Kareniaceae representing the majority of the reads obtained over the three-year sampled. Dinoflagellates vary greatly in size, genome size and possess a highly repetitive genome (Hou and Lin, 2009) which could artificially overestimate the importance of some lineages. Nonetheless, little is known yet about dinoflagellate genomics including the number of multi-copies and intragenomic diversity for the major part of known dinoflagellates. As for bacteria (Stoddard *et al.*, 2015), it would be extremely interesting to gain knowledge on the number of ribosomal genes copies per cell for different dinoflagellate species. Even at genus level, this information would allow a better weighting/calibration of HTS dinoflagellate analyses and therefore a realistic interpretation.

A major problem in the interpretation of HTS data for dinoflagellates is the very limited availability of validated references squences. This is why the study performed in **Chapter IV** for the ***Tripos* genus** was absolutely fundamental. Obtaining references from single-cells is time-consuming and an extremely meticulous work but essential to provide reference to interpret HTS data worldwide. Though arduous, continuous effort in this regard will be incredibly fruitful in the future, allowing

for easier interpretations of metabarcoding data, particularly for heterotrophic and mixotrophic species which are largely uncharacterised but play a significant role in planktonic ecosystems.

One of the main recommendations from this work is to continue over a longer period the monitoring of the whole protist community, including dinoflagellates, at the LTER-MC station, coupling HTS and counting techniques with the isolation/characterisation of species, in order to expand our knowledge in the functioning and long-term trends of the coastal planktonic ecosystem.

CHAPTER VI: Bibliography

6. Bibliography

Adachi, M., Kanno, T., Okamoto, R., Shinozaki, A., Fujikawa-Adachi, K. and Nishijima, T. (2004) '*Jannaschia cystaugens* sp. nov., an *Alexandrium* (Dinophyceae) cyst formation-promoting bacterium from Hiroshima Bay, Japan', *International Journal of Systematic and Evolutionary Microbiology*, 54(5), pp. 1687–1692. doi: 10.1099/ijs.0.03029-0.

Adl, S. M., Simpson, A. G., Lane, C. E., Lukeš, J., Bass, D., Bowser, S. S., Brown, M., Burki, F., Dunthorn, M., Hampl, V., Heiss, A., Hoppenrath, M., Lara, E., leGall, L., Lynn, D. H., McManus, H., Mitchell, E. A. D., Mozley-Stanridge, S. E., Wegener Parfrey, L., Pawlowski, J., Rueckert, S., Shadwick, L., Schoch, C., Smirnov, A. and Spiegel, F. W. (2012) 'The revised classification of eukaryotes', *Journal of Eukaryotic Microbiology*, 59(5), pp. 429–493. doi: 10.1111/j.1550-7408.2012.00644.x.

D'Alelio, D., Mazzochi, M. G., Montresor, M., Sarno, D., Zingone, A., Di Capua, I., Franzè, G., Margiotta, F., Saggiomo, V. and Ribera d'Alcalà, M. (2015) 'The green-blue swing: Plasticity of plankton food-webs in response to coastal oceanographic dynamics', *Marine Ecology*, 36(4), pp. 1155–1170. doi: 10.1111/maec.12211.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J. (1990) 'Basic Local Alignment Search Tool', *Journal of Molecular Biology*, 215(3), pp. 403–410. doi: 10.1016/S0022-2836(05)80360-2.

Alverson, A. J. and Kolnick, L. (2005) 'Intragenomic nucleotide polymorphism among small subunit (18S) rDNA paralogs in the diatom genus *Skeletonema* (Bacillariophyta)', *Journal of Phycology*, 41(6), pp. 1248–1257. doi: 10.1111/j.1529-8817.2005.00136.x.

Amaral-Zettler, L. A., McCliment, E. A., Ducklow, H. W. and Huse, S. M. (2009) 'A method for studying protistan diversity using massively parallel sequencing of V9 hypervariable regions of small-subunit ribosomal RNA Genes', *PLoS ONE*, 4(7). doi: 10.1371/journal.pone.0006372.

Anderson, D. M., Cembella, A. D. and Hallegraeff, G. M. (2012) 'Progress in Understanding Harmful Algal Blooms: Paradigm Shifts and New Technologies for Research, Monitoring, and Management', *Annual Review of Marine Science*, 4(1), pp. 143–176. doi: 10.1146/annurev-marine-120308-081121.

Anderson, D. M., Alpermann, T. J., Cembella, A. D., Collos, Y., Masseret, E. and Montresor, M. (2012) 'The globally distributed genus *Alexandrium*: Multifaceted roles in marine ecosystems and impacts on human health', *Harmful Algae*. Elsevier B.V., 14, pp. 10–35. doi: 10.1016/j.hal.2011.10.012.

Aubry, F. B., Cossarini, G., Acri, F., Bastianini, M., Bianchi, F., Camatti, E., De Lazzari, A., Pugnetti, A., Solidoro, C. and Socal, G. (2012) 'Plankton communities in the northern Adriatic Sea: Patterns and changes over the last 30 years', *Estuarine, Coastal and Shelf Science*. Elsevier Ltd, 115, pp. 125–137. doi: 10.1016/j.ecss.2012.03.011.

Bachvaroff, T. R., Gornik, S. G., Concepcion, G. T., Waller, R. F., Mendez, G. S., Lippmeier, J. C. and Delwiche, C. F. (2014) 'Dinoflagellate phylogeny revisited: Using ribosomal proteins to resolve deep branching dinoflagellate clades', *Molecular Phylogenetics and Evolution*, 70(1), pp. 314–322. doi: 10.1016/j.ympev.2013.10.007.

Baek, S. H., Shimode, S. and Kikuchi, T. (2007) 'Reproductive ecology of the dominant dinoflagellate, *Ceratium fusus*, in coastal area of Sagami Bay, Japan', *Journal of Oceanography*, 63(1), pp. 35–45. doi: 10.1007/s10872-007-0004-y.

Baek, S. H., Shimode, S., Shin, K., Han, M. S. and Kikuchi, T. (2009) 'Growth of dinoflagellates,

Ceratium furca and *Ceratium fusus* in Sagami Bay, Japan: The role of vertical migration and cell division', *Harmful Algae*, 8(6), pp. 843–856. doi: 10.1016/j.hal.2009.04.001.

Baker, M. (1753) *Employment for the microscope*. Dodsley, London, 403 pp.

Barton, A. D., Finkel, Z. V., Ward, B. A., Johns, D. G. and Follows, M. J. (2013) 'On the roles of cell size and trophic strategy in North Atlantic diatom and dinoflagellate communities', *Limnology and Oceanography*, 58(1), pp. 254–266. doi: 10.4319/lo.2013.58.1.0254.

Bendif, E. M., Probert, I., Carmichael, M., Romac, S., Hagino, K. and Vargas, C. (2014) 'Genetic delineation between and within the widespread coccolithophore morpho-species *Emiliania huxleyi* and *Gephyrocapsa oceanica* (Haptophyta)', *Journal of Phycology*, 50(1), pp. 140–148. doi: 10.1111/jpy.12147.

Berdalet, E., Fleming, L. E., Gowen, R., Davidson, K., Hess, P., Backer, L. C., Moore, S. K., Hoagland, P. and Enevoldsen, H. (2016) 'Marine harmful algal blooms, human health and wellbeing: challenges and opportunities in the 21st century', *Journal of the Marine Biological Association of the United Kingdom*, 96(1), pp. 61–91. doi: 10.1017/S0025315415001733.

Berney, C. and Pawlowski, J. (2006) 'A molecular time-scale for eukaryote evolution recalibrated with the continuous microfossil record', *Proceedings of the Royal Society B: Biological Sciences*, 273(1596), pp. 1867–1872. doi: 10.1098/rspb.2006.3537.

Berney, C., Ciuprina, A., Bender, S., Brodie, J., Edgcomb, V., Kim, E., Rajan, J., Wegener Parfrey, L., Adl, S., Audic, S., Bass, D., Caron, D. A., Cochrane, G., Czech, L., Dunthorn, M., Geisen, S., Glöckner, F. O., Mahé, F., Quast, C., Kaye, J. Z., Simpson, A. G. B., Stamatakis, A., del Campo, J., Yilmaz, P. and de Vargas, C. (2017). UniEuk: Time to speak a common language in Protistology! *The Journal of Eukaryotic Microbiology*. 64, pp. 407-411. doi:10.1111/jeu.12414.

Bhaud, Y., Guillebault, D., Lennon, J., Defacque, H., Soyer-Gobillard, M. O. and Moreau, H (2000) 'Morphology and behaviour of dinoflagellate chromosomes during the cell cycle and mitosis.', *Journal of Cell Science*, 113 (Pt 7(7), pp. 1231–1239.

Biré, R., Trotereau, S., Lemée, R., oregioni, D., Delpont, C., Krys, S. and Guérin, T. (2015) 'Hunt for palytoxins in a wide variety of marine organisms harvested in 2010 on the French Mediterranean coast', *Marine Drugs*, 13(8), pp. 5425–5446. doi: 10.3390/md13085425.

Blankenberg, D., Von Kuster, G., Coraor, N., Ananda, G., Lazarus, R., Mangan, M., Nekrutenko, A. and Taylor, J. (2010) 'Galaxy: A web-based genome analysis tool for experimentalists', *Current Protocols in Molecular Biology*. doi: 10.1002/0471142727.mb1910s89.

Bongaerts, P., Carmichael, M., Hay, K. B., Tonk, L., Frade, P. R. and Hoegh-Guldberg, O. (2015) 'Prevalent endosymbiont zonation shapes the depth distributions of scleractinian coral species', *Proceedings of the Royal Society Open Science*, 2(140297), pp. 1–11. doi: 10.1098/rsos.140297.

Boutrup, P. V., Moestrup, Ø., Tillmann, U. and Daugbjerg, N. (2016) '*Katodinium glaucum* (Dinophyceae) revisited: proposal of new genus, family and order based on ultrastructure and phylogeny', *Phycologia*, 55(2), pp. 147–164. doi: 10.2216/15-138.1.

Brannock, P. M., Ortmann, A. C., Moss, A. G. and Halanych, K. M. (2016) 'Metabarcoding reveals environmental factors influencing spatio-temporal variation in pelagic micro-eukaryotes', *Molecular Ecology*, 25(15), pp. 3593–3604. doi: 10.1111/mec.13709.

Brescianini, C., Grillo, C., Melchiorre, N., Bertolotto, R., Ferrari, A., Vivaldi, B., Icardi, G., Gramaccioni, L., Funari, E. and Scardala, S. (2006) '*Ostreopsis ovata* algal blooms affecting human health in Genova, Italy, 2005 and 2006.', *Euro surveillance : bulletin européen sur les maladies transmissibles = European communicable disease bulletin*, 11(9).

Brocks, J. and Summons, R. (2005) 'Biomarkers for early life', in *Biogeochemistry*, p. 970.

Buck, K. R., Bolt, P. A. and Garrison, D. L. (1990) 'Phagotrophy and fecal pellet production by an athecate dinoflagellate in Antarctic sea ice', *Marine Ecology Progress Series*, 60(1–2), pp. 75–84. doi: 10.3354/meps060075.

Bunse, C. and Pinhassi, J. (2017) 'Marine Bacterioplankton Seasonal Succession Dynamics', *Trends in Microbiology*, pp. 494–505. doi: 10.1016/j.tim.2016.12.013.

Burki, F., Shalchian-Tabrizi, K., Minge, M., Skjæveland, Å., Nikolaev, S. I., Jacobsen, K. S. and Pawlowski, J (2007) 'Phylogenomics reshuffles the eukaryotic supergroups', *PLoS ONE*, 2(8). doi: 10.1371/journal.pone.0000790.

Burki, F. Imanian, B., Hehenberger, E., Hirakawa, Y., Maruyama, S. and Keeling, P. J. (2014) 'Endosymbiotic gene transfer in tertiary plastid-containing dinoflagellates', *Eukaryotic Cell*, 13(2), pp. 246–255. doi: 10.1128/EC.00299-13.

Bütschli O. (1885) *3. Unterabtheilung (Ordnung) Dinoflagellata*. – In: Dr. H.G. Bronn's Klassen und Ordnungen des Thier-Reichs, wissenschaftlich dargestellt in Wort und Bild. Erster Band Protozoa. – C.F. Winter'sche Verlagshandlung, Leipzig und Heidelberg. Pp. 906-1029; Pl.

Calado, A. J. and Huisman, J. M. (2010) 'Commentary: Gómez, F., Moreira, D., and López-García, P. (2010). *Neoceratium* gen. nov., a new genus for all marine species currently assigned to *Ceratium* (Dinophyceae). Protist 161: 35-54', *Protist*, 161(4), pp. 517–519. doi: 10.1016/j.protis.2010.04.003.

Del Campo, J., Kolisko, M., Boscaro, V., Santoferrara, L. F., Massana, R., Guillou, L., Simpson, A., Berney, C., de Vargas, C., Brown, M., Keeling, P. and Wegener Parfrey, L. (2018). EukRef: phylogenetic curation of ribosomal RNA to enhance understanding of eukaryotic diversity and distribution. *bioRxiv*, 278085, doi.10.1101/278085.

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., Fierer, N., Gonzalez Peña, A., Goodrich, J. K., Gordon, J. I., Huttley, G. A., Kelley, S. T., Koenig, J. E., Ley, R. E., Lozupone, C. A., McDonald, D., Muegge, B. D., Pirrung, M., Reeder, J., Sevinsky, J. R., Turnbaugh, P. J., Walters, W. A., Wildmann, J., Yatsunenko, T., Zaneveld, J. and Knight, R. (2010) 'QIIME allows analysis of high-throughput community sequencing data', *Nature methods*, 7(5), pp. 335–336. doi: 10.1038/nmeth.f.303.QIIME.

Carada, G. C, Hopkins, T. S., Bonaduce, G., Ianora, A., Marino, D., Modigh, M., Ribera D'Alcalà, M. and Scotto di Carlo, B. (1980). 'Variability in the hydrographic and biological features of the Gulf of Naples'. *Marine Ecology*, 1(2), pp. 105-120.

Caron, D. A. (2013) 'Towards a molecular taxonomy for protists: Benefits, risks, and applications in plankton ecology', *Journal of Eukaryotic Microbiology*, 60(4), pp. 407–413. doi: 10.1111/jeu.12044.

Caruana, A. M. N. and Malin, G. (2014) 'The variability in DMSP content and DMSP lyase activity in marine dinoflagellates', *Progress in Oceanography*, pp. 410–424. doi: 10.1016/j.pocean.2013.10.014.

Caruana, A. M. N., Steinke, M., Turner, S. M. and Malin, G. (2012) 'Concentrations of dimethylsulphoniopropionate and activities of dimethylsulphide-producing enzymes in batch cultures of nine dinoflagellate species', *Biogeochemistry*, 110(1–3), pp. 87–107. doi: 10.1007/s10533-012-9705-4.

Cembella, A. D. (2003) 'Chemical ecology of eukaryotic microalgae in marine ecosystems', *Phycologia*, 42(4), pp. 420–447. doi: 10.2216/i0031-8884-42-4-420.1.

Cerino, F., Orsini, L., Sarno, D., Dell'Aversano, C., Tartaglione, L. and Zingone, A. (2005) 'The alternation of different morphotypes in the seasonal cycle of the toxic diatom *Pseudo-nitzschia galaxiae*', *Harmful Algae*, 4(1), pp. 33–48. doi: 10.1016/j.hal.2003.10.005.

Chomérat, N., Sellos, D. Y., Zentz, F. and Nézan, E. (2010) 'Morphology and molecular phylogeny of *Prorocentrum consutum* sp. nov. (Dinophyceae), a new benthic dinoflagellate from South Brittany (northwestern France)', *Journal of Phycology*, 46(1), pp. 183–194. doi: 10.1111/j.1529-8817.2009.00774.x.

Cianielli, D., Falco, P., Iermano, I., Mozzillo, P., Uttieri, M., Buonocore, B., Zambardino, G. and Zambianchi, E. (2015). Inshore/offshore water exchange in the Gulf of Naples. *Journal of Marine Systems*, 145, 37-52, doi:10.1016/j.jmarsys.2015.01.002.

Cianelli, D., D'Alelio, D., Uttieri, M., Sarno, D., Zingone, A., Zambianchi, E. and Ribero D'Alcalà, M. R. (2017). Disentagling physical and biological drivers of phytoplankton dynamics in a coastal system. *Scientific reports*, 7, 15868. doi:10.1038/s41598-017-15880-x.

Ciminiello, P., Dell'Aversano, C., Fattorusso, E. and Forino, M. (2010) 'Palytoxins: A still haunting Hawaiian curse', *Phytochemistry Reviews*, pp. 491–500. doi: 10.1007/s11101-010-9185-x.

Clarke, K. R. and Ainsworth, M. (1993) 'A method of linking multivariate community structure to environmental variables', *Marine Ecology Progress Series*, pp. 205–219. doi: 10.3354/meps092205.

De Clerck, O., Guiry, M. D., Leliaert, F., Samyn, Y. and Verbruggen, H. (2013) 'Algal Taxonomy: A Road to Nowhere?', *Journal of Phycology*, 49(2), pp. 215–225. doi: 10.1111/jpy.12020.

Coats, D. W., Kim, S., Bachvaroff, T. R., Handy, S. M. and Delwiche, C. F. (2010) '*Tintinnophagus acutus* n. g., n. sp. (Phylum Dinoflagellata), an ectoparasite of the ciliate *Tintinnopsis cylindrica* Daday 1887, and its relationship to *Duboscquodinium collini* Grassé 1952', *Journal of Eukaryotic Microbiology*, 57(6), pp. 468–482. doi: 10.1111/j.1550-7408.2010.00504.x.

Collins, R. A. and Cruickshank, R. H. (2013) 'The seven deadly sins of DNA barcoding', *Molecular Ecology Resources*, 13(6), pp. 969–975. doi: 10.1111/1755-0998.12046.

Craveiro, S. C., Daugbjerg, N., Moestrup, Ø. and Calado, A. J. (2017) 'Studies on *Peridinium aciculiferum* and *Peridinium malmogiense (Scrippsiella hangoei)*: comparison with *Chimonodinium lomnickii* and description of *Apocalathium* gen. nov . (Dinophyceae)', *Phycologia*, 56(1), pp. 21–35. doi: 10.2216/16-20.1.

Daday, E. von (1888) 'Systematische Übersicht der Dinoflagellaten des Golfes von Neapel', *Természetrajzi Füzetek* 11: 98-104

Daugbjerg, N., Hansen, G., Larsen, J. and Moestrup, Ø. (2000) 'Phylogeny of some of the major genera of dinoflagellates based on ultrastructure and partial LSU rDNA sequence data, including the erection of three new genera of unarmoured dinoflagellates', *Phycologia*, 39(4), pp. 302–317. doi: 10.2216/i0031-8884-39-4-302.1.

Daugbjerg, N., Jensen, M. H. and Hansen, P. J. (2013) 'Using Nuclear-encoded LSU and SSU rDNA Sequences to Identify the Eukaryotic Endosymbiont in Amphisolenia bidentata (Dinophyceae)', *Protist*, 164(3), pp. 411–422. doi: 10.1016/j.protis.2012.10.001.

Decelle, J., Romac, S., Sasaki, E., Not, F. and Mahe, F. (2014) 'Intracellular diversity of the V4 and V9 regions of the 18S rRNA in marine protists (radiolarians) assessed by high-throughput sequencing', *PLoS ONE*, 9(8). doi: 10.1371/journal.pone.0104297.

Decelle, J., Colin, S. and Foster, R. A. (2015) 'Photosymbiosis in marine planktonic protists', in *Marine Protists: Diversity and Dynamics*, pp. 465–500. doi: 10.1007/978-4-431-55130-0_19.

Decelle, J., Romac, S., Stern, R. F., Bendif, E. M., Zingone, A., Audic, S., Guiry, M. D., Guillou, L.,

Tessier, D., Le Gall, F., Gourvil, P., Dos Santos, A. L., Probert, I., Vaulot, D., de Vargas, C. and Christen, R. (2015) 'PhytoREF: A reference database of the plastidial 16S rRNA gene of photosynthetic eukaryotes with curated taxonomy', *Molecular Ecology Resources*, 15(6), pp. 1435–1445. doi: 10.1111/1755-0998.12401.

Dodge, J. D. and Crawford, R. M. (1971) 'A fine-structural survey of dinoflagellate pyrenoids and food-reserves', *Botanical Journal of the Linnean Society*, 64(2), pp. 105–115. doi: 10.1111/j.1095-8339.1971.tb02138.x.

Dodge, J. D. and Marshall, H. G. (1994) 'Biogeographic analysis of the armored planktonic dinoflagellate *Ceratium* in the north Atlantic and adjacent seas', *Journal of Phycology*, 30(6), pp. 905–922. doi: 10.1111/j.0022-3646.1994.00905.x.

Dolan, J. R. (2005) 'Biogeography of aquatic microbes', *Aquatic Microbial Ecology*, 41, pp. 39–48.

Doucette, G. J., Cembella, A. D. and Boyer, G. L. (1989) 'Cyst formation in the red tide dinoflagellate *Alexandrium tamarense* (Dinophyceae): Effects of iron stress', *Journal of Phycology*, 25(4), pp. 721–731. doi: 10.1111/j.0022-3646.1989.00721.x.

Drebes, G. and Schnepf, E. (1988) '*Paulsenella* Chatton (Dinophyta), ectoparasites of marine diatoms: development and taxonomy', *Helgolander Meeresuntersuchungen*, 42(3–4), pp. 563–581. doi: 10.1007/BF02365627.

Duff, R. J., Ball, H. and Lavrentyev, P. J. (2008) 'Application of combined morphological-molecular approaches to the identification of planktonic protists from environmental samples', *Journal of Eukaryotic Microbiology*, 55(4), pp. 306–312. doi: 10.1111/j.1550-7408.2008.00328.x.

Dunthorn, M., Otto, J., Berger, S. A., Stamatakis, A., Mahé, F., Romac, S., de Vargas, C., Audic, S., BioMarKs Consortium, Stock Frank Kauff, A. and Stoeck, T. (2014) 'Placing environmental next-generation sequencing amplicons from microbial eukaryotes into a phylogenetic context', *Molecular Biology and Evolution*, 31(4), pp. 993–1009. doi: 10.1093/molbev/msu055.

Edgar, R. C. (2004) 'MUSCLE User Guide', *Nucleic Acids Research*, 32(November), pp. 1–15. doi: 10.1093/chemse/bjs128.

Edgar, R. C., Haas, B. J., Clemente, J. C., Quince, C. and Knight, R. (2011) 'UCHIME improves sensitivity and speed of chimera detection', *Bioinformatics*, 27(16), pp. 2194–2200. doi: 10.1093/bioinformatics/btr381.

Edwards, M., Johns, D. G., Leterme, S. C., Svendsen, E. and Richardson, A. J. (2006) 'Regional climate change and harmful algal blooms in the northeast Atlantic', *Limnology and Oceanography*, 51(2), pp. 820–829. doi: 10.4319/lo.2006.51.2.0820.

Ehrenberg, C.G. (1832) 'Beiträge zur Kenntnis der Organisation der Infusorien und ihrer geographischer Verbreitung, besonders in Sibirien' — Abhandlungen der Königlichen Akademie der Wissenschaften zu Berlin. Aus dem Jahre 1830. *Physikalische Abhandlungen,* 1830: 1-88, Pls 1-8.

Elwood, H. J., Olsen, G. J. and Sogin, M. L. (1985) 'The small-subunit ribosomal RNA gene sequences from the hypotrichous ciliates *Oxytricha nova* and *Stylonychia pustulata*.', *Molecular Biology and Evolution*, 2(5), pp. 399–410. doi: 10.1093/oxfordjournals.molbev.a040362.

Fensome, R. A., Taylor, F. J. R., Norris, G., Sarjeant, W. A. S., Wharton, D. I. and Williams, G. L. (1993) *A classification of living and fossil dinoflagellates*. Hanover, Pennsylvania, USA: Sheridan Press.

Fensome, R. A., MacRae, A.R. and Williams, G.L. (1998) 'Evolution des dinoflagellés et variation de leur diversité dans le temps', *Marine Micropaleontology*, 1–6.

Fensome, R. A., Saldarriaga, J. F. and Taylor, F. J. R. (1999) 'Dinoflagellate phylogeny revisited: Reconciling morphological and molecular based phylogenies', *Grana*, 38(2–3), pp. 66–80. doi: 10.1080/00173139908559216.

Figueroa, R. I., Bravo, I. and Garcés, E. (2006) 'Multiple routes of sexuality in *Alexandrium taylori* (Dinophyceae) in culture', *Journal of Phycology*, 42(5), pp. 1028–1039. doi: 10.1111/j.1529-8817.2006.00262.x.

Figueroa, R. I., Rengefors, K., Bravo, I. and Bench, S. (2010) 'From homothally to heterothally: Mating preferences and genetic variation within clones of the dinoflagellate *Gymnodinium catenatum*', *Deep-Sea Research Part II: Topical Studies in Oceanography*, 57(3–4), pp. 190–198. doi: 10.1016/j.dsr2.2009.09.016.

Finkel, Z. V., Sebbo, J., Feist-Burkhardt, S., Irwin, A. J., Katz, M. E., Schofield, O. M. E., Young, J. R. and Falkowski, P. G. (2007) 'A universal driver of macroevolutionary change in the size of marine phytoplankton over the Cenozoic', *Proceedings of the National Academy of Sciences*, 104(51), pp. 20416–20420. doi: 10.1073/pnas.0709381104.

Fistarol, G. O., Legrand, C., Rengefors, K. and Granéli, E. (2004) 'Temporary cyst formation in phytoplankton: A response to allelopathic competitors?', *Environmental Microbiology*, 6(8), pp. 791–798. doi: 10.1111/j.1462-2920.2004.00609.x.

Fitt, W. K. (1985) 'Effect of different strains of the zooxanthellae *Symbiodinium microadriaticum* on growth and survival of their coelenterate and molluscan hosts', *Proceedings of the 5th International Coral Reef Congress*, pp. 131–136.

Flynn, K. J., Stoecker, D. K., Mitra, A., Raven, J. A., Gilbert, P. M., Hansen, P. J., Granéli, E. and Burkholder, J. M. (2013) 'Misuse of the phytoplankton-zooplankton dichotomy: The need to assign organisms as mixotrophs within plankton functional types', *Journal of Plankton Research*, 35(1), pp. 3–11. doi: 10.1093/plankt/fbs062.

Foster, R. A., Carpenter, E. J. and Bergman, B. (2006) 'Unicellular cyanobionts in open ocean dinoflagellates, radiolarians, and tintinnids: Ultrastructural characterization and immuno-localization of phycoerythrin and nitrogenase', *Journal of Phycology*, 42(2), pp. 453–463. doi: 10.1111/j.1529-8817.2006.00206.x.

Flø Jørgensen, M., Murray, S. and Daugbjerg, N. (2004) '*Amphidinium* revisited. I. Redefinition of *Amphidinium* (Dinophyceae) based on cladistic and molecular phylogenetic analyses', *Journal of Phycology*, 40(2), pp. 351–365. doi: 10.1111/j.1529-8817.2004.03131.x.

Fritz, L. and Nass, M. (1992) 'Development of the endoparasitic dinoflagellate *Amoebophrya ceratii* within host dinoflagellate species', *Journal of Phycology*, 28(3), pp. 312–320. doi: 10.1111/j.0022-3646.1992.00312.x.

Fukuda, Y. and Endoh, H. (2008) 'Phylogenetic analyses of the dinoflagellate *Noctiluca scintillans* based on β-tubulin and Hsp90 genes', *European Journal of Protistology*, 44(1), pp. 27–33. doi: 10.1016/j.ejop.2007.07.001.

Gaines, G. and Taylor, F. J. R. (1985) 'Form and function of the dinoflagellate transverse flagellum', *The Journal of Protozoology*, 32(2), pp. 290–296. doi: 10.1111/j.1550-7408.1985.tb03053.x.

Garrido, M., Cecchi, P., Collos, Y., Agostini, S. and Pasqualini, V. (2016) 'Water flux management and phytoplankton communities in a Mediterranean coastal lagoon. Part I: How to promote dinoflagellate dominance?', *Marine Pollution Bulletin*. Elsevier Ltd, 104(1–2), pp. 139–152. doi: 10.1016/j.marpolbul.2016.01.049.

Gavelis, G. S., Hayakawa, S., White III, R. A., Gojobori, T., Suttle, C. A., Keeling, P. J. and Leander, B. S. *et al.* (2015) 'Eye-like ocelloids are built from different endosymbiotically acquired components', *Nature*, 523(7559), pp. 204–207. doi: 10.1038/nature14593.

Gavelis, G. S., Wakeman, K. C., Tillmann, U., Ripken, C., Mitarai, S., Herranz, M., Özbek, S., Holstein, T., Keeling, P. J. and Leander, B. S. (2017) 'Microbial arms race : Ballistic " nematocysts " in dinoflagellates represent a new extreme in organelle complexity', *Science Advances*, 3, pp. 1–7. doi: 10.1126/sciadv.1602552.

Genitsaris, S., Monchy, S., Viscogliosi, E., Sime-Ngando, T., Ferreira, S. and Christaki, U. (2015) 'Seasonal variations of marine protist community structure based on taxon-specific traits using the eastern English Channel as a model coastal system', *FEMS Microbiology Ecology*, 91(5), pp. 1–15. doi: 10.1093/femsec/fiv034.

Gestal, C., Novoa, B., Posasa, D., Figueras, A. and Azevedo, C. (2006) '*Perkinsoide chabelardi* n. gen., a protozoan parasite with an intermediate evolutionary position: Possible cause of the decrease of sardine fisheries?', *Environmental Microbiology*, 8(6), pp. 1105–1114. doi: 10.1111/j.1462-2920.2006.01008.x.

Giardine, B., Riemer, C., Hardison, R. C., Burhans, R., Elnitski, L., Shah, P., Zhang, Yi, Blankenberg, D., Albert, I., Taylor, J., Miller, W., Kent, W. J. and Nekrutenko, A. (2005) 'Galaxy: A platform for interactive large-scale genome analysis', *Genome Research*, 15(10), pp. 1451–1455. doi: 10.1101/gr.4086505.

Godhe, A., Asplund, M. E., Härnström, K., Saravanan, V., Tyagi, A. and Karunasagar, I. (2008) 'Quantification of diatom and dinoflagellate biomasses in coastal marine seawater samples by real-time PCR', *Applied and Environmental Microbiology*, 74(23), pp. 7174–7182. doi: 10.1128/AEM.01298-08.

Goecks, J., Nekrutenko, A. and Taylor, J.(2010) 'Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences', *Genome Biology*, 11(8), p. R86. doi: 10.1186/gb-2010-11-8-r86.

Gómez, F. (2005) 'A list of free-living dinoflagellate species in the world's oceans', *Acta Bot. Croat*, 64(1), pp. 129–212.

Gómez, F. (2006) 'Endemic and Indo-Pacific plankton in the Mediterranean Sea: A study based on dinoflagellate records', *Journal of Biogeography*, 33(2), pp. 261–270. doi: 10.1111/j.1365-2699.2005.01373.x.

Gómez, F. (2012a) 'A checklist and classification of living dinoflagellates (Dinoflagellata, Alveolata)', *CICIMAR Oceánides*, 27(1), pp. 65–140.

Gómez, F. (2012b) 'A quantitative review of the lifestyle, habitat and trophic diversity of dinoflagellates (Dinoflagellata, Alveolata)', *Systematics and Biodiversity*, 10(3), pp. 267–275. doi: 10.1080/14772000.2012.721021.

Gómez, F. (2013) 'Reinstatement of the dinoflagellate genus *Tripos* to replace *Neoceratium*, marine species of *Ceratium* (Dinophyceae, Alveolata)', *CICIMAR Oceánides*, 28(1), pp. 1–22.

Gómez, F. (2014) 'Problematic biases in the availability of molecular markers in Protists: The example of the Dinoflagellates', *Acta Protozoologica*, 53(1), pp. 63–75. doi: 10.4467/16890027AP.13.0021.1118.

Gómez, F. (2017) 'The function of the ocelloid and piston in the dinoflagellate *Erythropsidinium* (Gymnodiniales, Dinophyceae)', *Journal of Phycology*, 53(3), pp. 629–641. doi: 10.1111/jpy.12525.

Gómez, F. and Gorsky, G. (2003) 'Annual microplankton cycles in Villefranche Bay, Ligurian Sea,

NW Mediterranean', *Journal of Plankton Research*, 25(4), pp. 323–339. doi: 10.1093/plankt/25.4.323.

Gómez, F. and Skovgaard, A. (2015) 'The molecular phylogeny of the type-species of *Oodinium Chatton*, 1912 (Dinoflagellata: Oodiniaceae), a highly divergent parasitic dinoflagellate with non-dinokaryotic characters', *Systematic Parasitology*, 90(2), pp. 125–135. doi: 10.1007/s11230-014-9538-8.

Gómez, F., Moreira, D. and López-García, P. (2010a) 'Molecular phylogeny of the dinoflagellates *Podolampas* and *Blepharocysta* (Peridiniales, Dinophyceae)', *Phycologia*, 49(May), pp. 212–220. doi: 10.2216/PH09-29.1.

Gómez, F., Moreira, D. and López-García, P. (2010b) '*Neoceratium* gen. nov., a new genus for all marine species currently assigned to *Ceratium* (Dinophyceae)', *Protist*, 161(1), pp. 35–54. doi: 10.1016/j.protis.2009.06.004.

Gómez, F., Qiu, D., Dodge, J. D., Lopes, R. M. and Lin, S.(2016) 'Morphological and molecular characterization of *Ptychodiscus noctiluca* revealed the polyphyletic nature of the order Ptychodiscales (Dinophyceae)', *Journal of Phycology*, 52(5), pp. 793–805. doi: 10.1111/jpy.12438.

Gornik, S. G., Ford, K. L., Mulhern, T. D., Bacic, A., McFadden, G. I. and Waller, R. F. (2012) 'Loss of nucleosomal DNA condensation coincides with appearance of a novel nuclear protein in dinoflagellates', *Current Biology*, 22(24), pp. 2303–2312. doi: 10.1016/j.cub.2012.10.036.

Gottschling, M., Soehner, S., Zinssmeister, C., John, U., Plötner, J., Schweikert, M., Aligizaki, K. and Elbrächter, M.(2012) 'Delimitation of the Thoracosphaeraceae (Dinophyceae), including the calcareous dinoflagellates, based on large amounts of ribosomal RNA sequence data', *Protist*, 163(1), pp. 15–24. doi: 10.1016/j.protis.2011.06.003.

Gottschling, M., Čalasan, A. Ž., Kretschmann, J. and Gu, H. (2017) 'Two new generic names for dinophytes harbouring a diatom as an endosymbiont, *Blixaea* and *Unruhdinium* (Kryptoperidiniaceae, Peridiniales)', *Phytotaxa*, 306(4), pp. 296–300. doi: 10.11646/phytotaxa.306.4.6.

Goulet, T. L., Simmons, C. and Goulet, D. (2008) 'Worldwide biogeography of *Symbiodinium* in tropical octocorals', *Marine Ecology Progress Series*, 355, pp. 45–58. doi: 10.3354/meps07214.

Gouy, M., Guindon, S. and Gascuel, O. (2010) 'SeaView Version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building', *Molecular Biology and Evolution*, 27(2), pp. 221–224. doi: 10.1093/molbev/msp259.

Graham, L.E., Graham, J. M., Wilcox, L. W. and Cook, M. E. (2016) *Algae (Third Edition – version 3.1.1).* LJLM Press, LLC. Pearson Prentice Hall, Pearson Education, Inc. Upper Saddle River, New Jersey (NJ).

Green, B. R. (2004) 'The chloroplast genome of dinoflagellates - A reduced instruction set?', *Protist*, pp. 23–31. doi: 10.1078/143446100161.

Grzebyk, D. and Berland, B. (1996) 'Influences of temperature, salinity and irradiance on growth of *Prorocentrum minimum* (Dinophyceae) from the Mediterranean Sea', *Journal of Plankton Research*, 18(10), pp. 1837–1849. doi: 10.1093/plankt/18.10.1837.

Grzebyk, D., Audic, S., Lasserre, B., Abadie, E., de Vargas, C. and Bec, B. (2017) 'Insights into the harmful algal flora in northwestern Mediterranean coastal lagoons revealed by pyrosequencing metabarcodes of the 28S rRNA gene', *Harmful Algae*. Elsevier B.V., 68, pp. 1–16. doi: 10.1016/j.hal.2017.06.003.

Gu, H., Liu, T. and Mertens, K. N. (2015) 'Cyst–theca relationship and phylogenetic positions of

Protoperidinium (Peridiniales, Dinophyceae) species of the sections Conica and Tabulata, with description of *Protoperidinium shanghaiense* sp. nov.', *Phycologia*, 54(February), pp. 49–66. doi: 10.2216/14-047.1.

Guillou, L., Viprey, M., Chambouvet, A., Welsh, R. M., Kirkham, A. R., Massana, R., Scanlan, D. J. and Worden, A. Z. (2008) 'Widespread occurrence and genetic diversity of marine parasitoids belonging to Syndiniales (Alveolata)', *Environmental Microbiology*, 10(12), pp. 3349–3365. doi: 10.1111/j.1462-2920.2008.01731.x.

Guillou, L., Bachar, D., Audic, S., Bass, D., Berney, C., Bittner, L., Boutte, C., Burgaud, G., de Vargas, C., Decelle, J., del Campo, J., Dolan, J. R., Dunthorn, M., Edvardsen, B., Holzmann, M., Kooistra, H. C. F., Lara, E., Le Bescot, N., Logares, R., Mahé, F., Massana, R., Montresor, M., Morard, R., Not, F., Pawlowski, J., Probert, I., Sauvadet, A. -L., Siano, R., Stoeck, T., Vaulot, D., Zimmermann, P. and Christen, R. (2013) 'The Protist Ribosomal Reference database (PR^2): A catalog of unicellular eukaryote Small Sub-Unit rRNA sequences with curated taxonomy', *Nucleic Acids Research*, 41(D1). doi: 10.1093/nar/gks1160.

Guiry, M. D. and Guiry, G. M. (2017) *AlgaeBase*, *World-wide electronic publication, National University of Ireland, Galway*. Available at: http://www.algaebase.org

Hackett, J. D., Anderson, D. M., Erdner, D. L. and Bhattacharya, D. (2004) 'Dinoflagellates: A remarkable evolutionary experiment', *American Journal of Botany*, pp. 1523–1534. doi: 10.3732/ajb.91.10.1523.

Hallegraeff, G. M. (2004) *Harmful algal blooms: a global overview*, *Manual on Harmful Marine Microalgae*. doi: 10.1016/B978-0-12-391499-6.00008-6.

Halim, Y. (1963) *Microplankton des eaux Egyptiennes. Le genre Ceratium Schrank (Dinoflagellés). Rapp Comm int Mer Medit* 17: 495-502.

Hansen, G., Daugbjerg, N. and Henriksen, P. (2007) '*Baldinia anauniensis* gen. et sp. nov.: a "new" dinoflagellate from Lake Tovel, N. Italy', *Phycologia*, 46(January), pp. 86–108. doi: 10.2216/PH06-23.1.

Harada, A., Ohtsuka, S. and Horiguchi, T. (2007) 'Species of the parasitic genus *Duboscquella* are members of the enigmatic marine Alveolate Group I', *Protist*, 158(3), pp. 337–347. doi: 10.1016/j.protis.2007.03.005.

Hasle, G. R. and Nordli, E. (1951) *Form variation in Ceratium fusus and tripos populations in cultures and from the sea.* Avh Norske Videnskaps-Akademi i Oslo I. Mat Naturv Klasse 4: 1- 25

Hehenberger, E., Imanian, B., Burki, F. and Keeling, P. J. (2014) 'Evidence for the retention of two evolutionary distinct plastids in dinoflagellates with diatom endosymbionts', *Genome Biology and Evolution*, 6(9), pp. 2321–2334. doi: 10.1093/gbe/evu182.

Heisler, J., Glibert, P. M., Burkholder, J. M., Anderson, D. M., Cochlan, W., Dennison, W. C., Dortch, Q., Gobler, C. J., Heil, C. A., Humphries, E., Lewitus, A., Magnien, R., Marshall, H. G., Sellner, K., Stockwell, D. A., Stoecker, D. K, and Suddleson, M. (2008) 'Eutrophication and harmful algal blooms: A scientific consensus', *Harmful Algae*, 8(1), pp. 3–13. doi: 10.1016/j.hal.2008.08.006.

Hoppenrath, M. (2017) 'Dinoflagellate taxonomy — a review and proposal of a revised classification', *Marine Biodiversity*, pp. 381–403. doi: 10.1007/s12526-016-0471-8.

Hoppenrath, M. and Elbrächter, M. (2015). Center of Excellence for Dinophyte Taxonomy (CEDiT). http://www.dinophyta.org/

Hoppenrath, M. and Leander, B. S. (2007a) 'Character evolution in polykrikoid dinoflagellates',

Journal of Phycology, 43(2), pp. 366–377. doi: 10.1111/j.1529-8817.2007.00319.x.

Hoppenrath, M. and Leander, B. S. (2007b) 'Morphology and Phylogeny of the Pseudocolonial Dinoflagellates *Polykrikos lebourae* and *Polykrikos herdmanae* n. sp.', *Protist*, 158(2), pp. 209–227. doi: 10.1016/j.protis.2006.12.001.

Hoppenrath, M. and Leander, B. S. (2010) 'Dinoflagellate phylogeny as inferred from heat shock protein 90 and ribosomal gene sequences', *PLoS ONE*, 5(10). doi: 10.1371/journal.pone.0013220.

Hoppenrath, M. and Saldarriaga, F. (2008) ToLweb dinoflagellate page. http://tolweb.org/Dinoflagellates/2445

Hoppenrath, M., Chomérat, N. and Leander, B. S. (2013) 'Molecular phylogeny of *Sinophysis* : Evaluating the possible early evolutionary history of dinophysoid dinoflagellates', in *Biological and geological perspectives of dinoflagellates. The Micropalaeontological Society, Special Publications. Geological Society, London*, pp. 207–214.

Hoppenrath, M., Murray, S., Sparmann, S. F. and Leander, B. S. (2012) 'Morphology and molecular phylogeny of *Ankistrodinium* gen. nov. (Dinophyceae), a new genus of marine sand-dwelling dinoflagellates formerly classified within *Amphidinium*', *Journal of Phycology*, 48(5), pp. 1143–1152. doi: 10.1111/j.1529-8817.2012.01198.x.

Hoppenrath, M., Murray, S. A., Chomérat, N. and Horiguchi, T. (2014) *Marine benthic dinoflagellates - unveiling their worlwide biodiversity*. Kleine Sen. Edited by M. Hoppenrath et al. Johannesstraße 3A, 70176 Stuttgart, Germany: Schweizerbart'sche Verlagsbuch- handlung.

Horiguchi, T. (2015) 'Diversity and phylogeny of marine parasitic dinoflagellates', in *Marine Protists: Diversity and Dynamics*, pp. 397–419. doi: 10.1007/978-4-431-55130-0_16.

Hou, Y. and Lin, S. (2009) 'Distinct gene number-genome size relationships for eukaryotes and non-eukaryotes: Gene content estimation for dinoflagellate genomes', *PLoS ONE*, 4(9). doi: 10.1371/journal.pone.0006978.

Howe, C. J., Nisbet, R. E. R. and Barbrook, A. C. (2008) 'The remarkable chloroplast genome of dinoflagellates', in *Journal of Experimental Botany*, pp. 1035–1045. doi: 10.1093/jxb/erm292.

Hu, S. K., Liu, Z., Lie, A. A., Countway, P.D., Kim, D. Y., Jones, A. C., Gast, R. J., Craig Cary, S., Sherr, E. B., Sherr, B. F. and Caron, D. A. (2015) 'Estimating protistan diversity using high-throughput sequencing', *Journal of Eukaryotic Microbiology*, 62(5), pp. 688–693. doi: 10.1111/jeu.12217.

Huelsenbeck, J. P. and Ronquist, F. (2001) 'MRBAYES: Bayesian inference of phylogenetic trees', *Bioinformatics*, 17(8), pp. 754–755. doi: 10.1093/bioinformatics/17.8.754.

Huse, S. M., Welch, D. M., Morrison, H. G. and Sogin, M. L. (2010) 'Ironing out the wrinkles in the rare biosphere through improved OTU clustering', *Environmental Microbiology*, 12(7), pp. 1889–1898. doi: 10.1111/j.1462-2920.2010.02193.x.

Jacobson, D. M. (1999) 'A brief history of dinoflagellate feeding research', *Journal of Eukaryotic Microbiology*, 46(4), pp. 376–381. doi: 10.1007/978-1-4419-9872-9.

Jacobson, D. M. and Anderson, D. M. (1986) 'Thecate heterophic dinoflagellates: feeding behavior and mechanisms', *Journal of Phycology*, 22(3), pp. 249–258. doi: 10.1111/j.1529-8817.1986.tb00021.x.

Janouškovec, J., Gavelis, G. S., Burki, F., Dinh, D., Bachvaroff, T. R., Gornik, S. G., Bright, K. J., Imanian, B., Strom, S. L., Delwiche, C. F., Waller, R. F., Fensome, R. A., Leander, B. S., Rohwer, F. L. and Saldarriaga, J. F. (2017) 'Major transitions in dinoflagellate evolution unveiled by phylotranscriptomics', *Proceedings of the National Academy of Sciences*, 114(2), pp. E171–E180.

doi: 10.1073/pnas.1614842114.

John, U., Fensome, R. A. and Medlin, L. K. (2003) 'The application of a molecular clock based on molecular sequences and the fossil record to explain biogeographic distributions within the *Alexandrium tamarense* "species complex" (Dinophyceae)', *Molecular Biology and Evolution*, 20(7), pp. 1015–1027. doi: 10.1093/molbev/msg105.

John, U., Litaker, R. W., Montresor, M., Murray, S., Brosnahan, M. L. and Anderson, D. M. (2014) 'Formal revision of the *Alexandrium tamarense* species complex (Dinophyceae) taxonomy: The introduction of five species with emphasis on molecular-based (rDNA) classification', *Protist*, 165(6), pp. 779–804. doi: 10.1016/j.protis.2014.10.001.

Jørgensen, E. (1911) *Die Ceratien. Eine kurze Monographie der Gattung Ceratium Schrank*. Int Revue ges Hydrogr 4:1–124.

Katoh, K. and Standley, D. M. (2013) 'MAFFT multiple sequence alignment software version 7: Improvements in performance and usability', *Molecular Biology and Evolution*, 30(4), pp. 772–780. doi: 10.1093/molbev/mst010.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P. and Drummond, A. (2012) 'Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data', *Bioinformatics*, 28(12), pp. 1647–1649. doi: 10.1093/bioinformatics/bts199.

Keeling, P. J. (2010) 'The endosymbiotic origin, diversification and fate of plastids', *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1541), pp. 729–748. doi: 10.1098/rstb.2009.0103.

Keeling, P. J. (2013) 'Elephants in the room: protists and the importance of morphology and behaviour', *Environmental Microbiology Reports*, 5, pp. 5–6.

Ki, J. S. (2010) 'Nuclear 28S rDNA phylogeny supports the basal placement of *Noctiluca scintillans* (Dinophyceae; Noctilucales) in dinoflagellates', *European Journal of Protistology*. Elsevier, 46(2), pp. 111–120. doi: 10.1016/j.ejop.2009.11.001.

Ki, J. S. (2012) 'Hypervariable regions (V1-V9) of the dinoflagellate 18S rRNA using a large dataset for marker considerations', *Journal of Applied Phycology*, 24(5), pp. 1035–1043. doi: 10.1007/s10811-011-9730-z.

Ki, J. S., Jang, G. Y. and Han, M.-S. (2005) 'Integrated method for single-cell DNA extraction, PCR amplification, and sequencing of ribosomal DNA from harmful dinoflagellates *Cochlodinium polykrikoides* and *Alexandrium catenella*.', *Marine biotechnology*, 6(6), pp. 587–93. doi: 10.1007/s10126-004-1700-x.

Kim, D. Y., Countway, P. D., Jones, A. C., Schnetzer, A., Yamashita, W., Tung, C. and Caron, D. A. (2014) 'Monthly to interannual variability of microbial eukaryote assemblages at four depths in the eastern North Pacific', *ISME Journal*. Nature Publishing Group, 8(3), pp. 515–530. doi: 10.1038/ismej.2013.173.

Kim, K. M., Park, J. H., Bhattacharya, D. and Yoon, H. S. (2014) 'Applications of next-generation sequencing to unravelling the evolutionary history of algae', *International Journal of Systematic and Evolutionary Microbiology*, 64(PART 2), pp. 333–345. doi: 10.1099/ijs.0.054221-0.

Klais, R., Tamminen, T., Kremp, A., Spilling, K. and Olli, K. (2011) 'Decadal-scale changes of Dinoflagellates and Diatoms in the Anomalous Baltic Sea spring bloom', *PLoS ONE*, 6(6). doi: 10.1371/journal.pone.0021567.

Kofoid, C. A., and Swezy, O. (1921) *The freeliving unarmored Dinoflagellata.* Memoirs of the University of California 5:1–562.

Koft, A., Bicak, M., Kottmann, R., Schnetzer, J., Kastadinov, I., Lehmann, K., Fernandez-Guerra, A., Jeanthon, C., Rahav, E., Ullrich, M., Wichels, A., Gerdts, G., Polymenakou, P., Kotoulas, G., Siam, R., Abdallah, R. Z., Sonnenschein, E. C., Cariou, T., O'Gara, F., Jackson, S., Orlic, S., Steinke, M., Busch, J., Duarte, B., Caçador, I., Canning-Clode, J., Bobrova, O., Marteinsson, V., Reynisson, E., Magalhães Loureiro, C., Luna, G. M., Quero, G. M., Löscher, C. R., Kremp, A., DeLorenzo, M. E., Øvreås, L., Tolman, J., LaRoche, J., Penna, A., Frischer, M., Davis, T., Katherine, B., Meyer, C. P., Ramos, S., Magalhães, C., Jude-Lemeilleur, F., Aguirre-Macedo, M. L., Wang, S., Poulton, N., Jones, S., Collin, R., Fuhrman, J. A., Conan, P., Alonson, C., Stambler, N., Goodwin, K., Yakimov, M. M., Baltar, F., Bodrossy, L., Van De Kamp, J., Frampton, D. M. F., Ostrowski, M., Van Ruth, P., Malthouse, P., Claus, S., Deneudt, K., Mortelmans, J., Pitois, S., Wallom, D., Salter, I., Costa, R., Schroeder, D. C., Kandil, M. M., Amaral, V., Biancalana, F., Santana, R., Pedrotti, M. L., Yoshida, T., Ogata, H., Ingleton, T., Munnik, K., Rodriguez-Ezpeleta, N., Berteaux-Lecellier, V., Wecker, P., Cancio, I., Vaulot, D., Bienhold, C., Ghazal, H., Chaouni, B., Essayeh, S., Ettamimi, S., Zaid, E. H., Boukhatem, N., Bouali, A., Chahboune, R., Barrijal, S., Timinouni, M., El Otmani, F., Bennani, M., Mea, M., Todorova, N., Karamfilov, V., ten Hoopen, P., Cochrane, G., L'Haridon, S., Bizsel, K. C., Vezzi, A., Lauro, F. M., Martin, P., Jensen, R. M., Hinks, J., Gebbels, S., Rosselli, R., De Pasquale, F., Schiavon, R., dos Santos, A., Villar, E., Pesant, S., Cataletto, B., Malfatti, F., Edirisinghe, R., Herrera Silveira, J. A., Barbier, M., Turk, V., Tinta, T., Fuller, W. J., Salihoglu, I., Serakinci, N., Ergoren, M. C., Bresnan, E., Iriberri, J., Fronth Nyhus, P. A., Bente, E., Karlsen, H. E., Golyshin, P. N., Gasol, J. M., Moncheva, S., Dzhembekova, N., Johnson, Z., Sinigalliano, C. D., Gidley, M. L., Zingone, A., Danovaro, R., Tsiamis, G., Clark, M. S., Costa, A. C., El Bour, M., Martins, A. M., Collins, R. E., Ducluzeau, A. -L., Martinez, J., Costello, M. J., Amaral-Zettler, L. A., Gilbert, J. A., Davies, N., Field, D. and Glöckner, F. O. (2015) 'The ocean sampling day consortium', *GigaScience*, 4(1), p. 25. doi: 10.1186/s13742-015-0065-6.

Kohli, G. S., Neilan, B. A., Brown, M. V., Hoppenrath, M. and Murray, S. A. (2014) 'Cob gene pyrosequencing enables characterization of benthic dinoflagellate diversity and biogeography', *Environmental Microbiology*, 16(2), pp. 467–485. doi: 10.1111/1462-2920.12275.

Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K. and Schloss, P. D. (2013) 'Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the miseq illumina sequencing platform', *Applied and Environmental Microbiology*, 79(17), pp. 5112–5120. doi: 10.1128/AEM.01043-13.

Kremp, A., Tamminen, T. and Spilling, K. (2008) 'Dinoflagellate bloom formation in natural assemblages with diatoms: Nutrient competition and growth strategies in Baltic spring phytoplankton', *Aquatic Microbial Ecology*, 50(2), pp. 181–196. doi: 10.3354/ame01163.

Kremp, A., Lindholm, T., Dreßler, N., Erler, K., Gerdts, G., Eirtovaara., S. and Leskinen, E. (2009) 'Bloom forming *Alexandrium ostenfeldii* (Dinophyceae) in shallow waters of the Åland Archipelago, Northern Baltic Sea', *Harmful Algae*, 8(2), pp. 318–328. doi: 10.1016/j.hal.2008.07.004.

Kumar, S., Stecher, G. and Tamura, K. (2016) 'MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets', *Molecular biology and evolution*, 33(7), pp. 1870–1874. doi: 10.1093/molbev/msw054.

LaJeunesse, T. C. (2001) 'Investigating the biodiversity, ecology, and phylogeny of endosymbiotic dinoflagellates in the genus *Symbiodinium* using the ITS region: in search of a "species" level marker.' *Journal of Phycology*, 37, 866–880. doi: 10.1046/j.1529-8817.2001.01031.x.

LaJeunesse, T. C., Andersen, R. A. and Galbraith, D. W. (2005) '*Symbiodinium* (Pyrrhophyta) genome sizes (DNA content) are smallest among dinoflagellates', 41(4), 880–886. doi:

10.1111/j.1529-8817.2005.00111.x.

LaJeunesse, T. C., Pettay, D. T., Sampayo, E. M., Phongsuwan, N., Brown, B., Obura, D. O., Hoegh-Guldberg, O. and Fitt, W. K. (2010) 'Long-standing environmental conditions, geographic isolation and host-symbiont specificity influence the relative ecological dominance and genetic diversification of coral endosymbionts in the genus *Symbiodinium*', *Journal of Biogeography*, 37(5), pp. 785–800. doi: 10.1111/j.1365-2699.2010.02273.x.

Lallias, D., Hiddink, J. G., Fonseca, V. G., Gaspar, J. M., Sung, W., Neill, S. P., Barnes, N., Ferrero, T., Hall, N., Lambshead, P. J. D., Packer, M., Thomas, W. K. and Creer, S. (2015) 'Environmental metabarcoding reveals heterogeneous drivers of microbial eukaryote diversity in contrasting estuarine ecosystems.', *The ISME journal*, 9(5), pp. 1208–21. doi: 10.1038/ismej.2014.213.

Letunic, I. and Bork, P. (2016) 'Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees', *Nucleic acids research*, 44(W1), pp. W242–W245. doi: 10.1093/nar/gkw290.

Le Bescot, N. (2014) *Patrons de biodiversité à l'échelle globale chez les dinoflagellés planctoniques marins*. Thèse de doctorat de l'Université Pierre et Marie Curie.

Le Bescot, N., Mahé, F., Audic, S., Dimier, C., Garet, M. J., Poulain, J., Wincker, P., de Vargas, C. and Siano, R. (2016) 'Global patterns of pelagic dinoflagellate diversity across protist size classes unveiled by metabarcoding', *Environmental Microbiology*, 18(2), pp. 609–626. doi: 10.1111/1462-2920.13039.

Levy, M. G., Litaker, R. W., Goldstein, R. J., Dykstra, M. J., Vandersea, M. W and Noga, E. J. (2007) '*Piscinoodinium*, a fish-ectoparasitic dinoflagellate, is a member of the class dinophyceae, subclass gymnodiniphycidae: convergent evolution with *Amyloodinium*.', *The Journal of parasitology*, 93, pp. 1006–1015. doi: 10.1645/GE-3585.1.

Li, T., Liu, S., Huang, L., Huang, H., Lian, J., Yan, Y. and Lin, S. (2011) 'Diatom to dinoflagellate shift in the summer phytoplankton community in a bay impacted by nuclear power plant thermal effluent', *Marine Ecology Progress Series*, 424, pp. 75–85. doi: 10.3354/meps08974.

Lie, A. A. Y., Kim, D. Y., Schnetzer, A. and Caron, D. A. (2013) 'Small-scale temporal and spatial variations in protistan community composition at the San Pedro Ocean Time-series station off the coast of southern California', *Aquatic Microbial Ecology*, 70(2), pp. 93–110. doi: 10.3354/ame01652.

Lilly, E. L., Halanych, K. M. and Anderson, D. M. (2007) 'Species boundaries and global biogeography of the *Alexandrium tamarense* complex (Dinophyceae)', *Journal of Phycology*, 43(6), pp. 1329–1338. doi: 10.1111/j.1529-8817.2007.00420.x.

Lin, S., Zhang, H., Hou, Y., Zhuang, Y. and Miranda, L. (2009) 'High-level diversity of dinoflagellates in the natural environment, revealed by assessment of mitochondrial cox1 and cob genes for dinoflagellate DNA barcoding', *Applied and Environmental Microbiology*, 75(5), pp. 1279–1290. doi: 10.1128/AEM.01578-08.

Lin, S., Cheng, S., Song., B., Zhong, X., Lin, X., Li, W., Li, L., Zhang, Y., Zhang, H., Ji, Z., Cai, M., Zhuang, Y., Shi, X., Lin, L., Wang, L., Wang, Z., Liu, X., Yu, S., Zeng, P., Hao, H., Zou, Q., Chen, C., Li, Y., Wang, Y., Xu, C., Meng, S., Xu, X., Wang, J., Yang, H., Campbell, D. A., Sturm, N. R., Dagenais-Bellefeuille, S. and Morse, D. (2015) 'The *Symbiodinium kawagutii* genome illuminates dinoflagellate gene expression and coral symbiosis', *Science*, 350(6261), pp. 691–694. doi: 10.1126/science.aad0408.

Lindberg, K., Moestrup, Ø. and Daugbjerg, N. (2005) 'Studies on woloszynskioid dinoflagellates I: *Woloszynskia coronata* re-examined using light and electron microscopy and partial LSU rDNA

sequences, with description of *Tovellia* gen. nov. and *Jadwigia* gen. nov. (Tovelliaceae fam. nov.)', *Phycologia*, 44(4), pp. 416–440. doi: 10.2216/0031-8884(2005)44[416:sowdiw]2.0.co;2.

Litaker, R. W., Vandersea, M. W., Kibler, S. R., Madden, V. J., Noga, E. J. and Tester, P. A. (2002) 'Life cycle of the heterotrophic dinoflagellate *Pfiesteria piscicida* (Dinophyceae)', *Journal of Phycology*, 38(3), pp. 442–463. doi: 10.1046/j.1529-8817.2002.t01-1-01242.x.

Litaker, R. W., Vandersea, M. W., Kibler, S. R., Reece, K. S., Stokes, N. A., Lutzoni, F. M., Yonish, B. A., West, M. A., Black, M. N. D. and Tester, P. A. (2007) 'Recognizing dinoflagellate species using its rDNA sequences', *Journal of Phycology*, 43(2), pp. 344–355. doi: 10.1111/j.1529-8817.2007.00320.x.

Liu, T., Mertens, K. N., Ribeiro, S., Ellegaard, M., Matsuoka, K. and Gu, H. (2015) 'Cyst-theca relationships and phylogenetic positions of Peridiniales (Dinophyceae) with two anterior intercalary plates, with description of *Archaeperidinium bailongense* sp. nov. and *Protoperidinium fuzhouense* sp. nov', *Phycological Research*, 63(2), pp. 134–151. doi: 10.1111/pre.12081.

Livolant, F. (1982a) 'Dinoflagellate trichocyst ultrastructure I. The shaft. Biol.' *Cell,* 43, 201–210.

Livolant, F. (1982b) 'Dinoflagellate trichocyst ultrastructure II. Existence of a sheath. Biol.' *Cell,* 43, 211–216.

Lobban, C. S., Schefter, M., Simpson, A., Pochon, X., Pawlowski, J. and Foissner, W. (2002) '*Maristentor dinoferus* n. gen., n. sp., a giant heterotrich ciliate (Spirotrichea: Heterotrichida) with zooxanthellae, from coral reefs on Guam, Mariana Islands', *Marine Biology*, 140(2), pp. 411–423. doi: 10.1007/s002270100690.

Logares, R., Audic, S., Bass, D., Bittner, L., Boutte, C., Christen, R., Claverie, J. -M., Decelle, J., Dolan, J. R., Dunthorn, M., Edvardsen, B., Gobet, A., Kooistra, W. H. C. F., Mahé, F., Not, F., Ogata, H., Pawlowski, J., Pernice, M. C., Romac, S., Shalchian-Tabrizi, K., Simon, N., Stoeck, T., Santini, S., Siano, R., Wincker, P., Zingone, A., Richards, T. A., de Vargas, C. and Massana, R. (2014) 'Patterns of rare and abundant marine microbial eukaryotes', *Current Biology*, 24(8), pp. 813–821. doi: 10.1016/j.cub.2014.02.050.

Lowe, C. D., Montagnes, D. J., Martin, L. E. and Watts, P. C. (2010) 'High genetic diversity and fine-scale spatial structure in the marine flagellate *Oxyrrhis marina* (Dinophyceae) uncovered by microsatellite loci', *PLoS ONE*, 5(12). doi: 10.1371/journal.pone.0015557.

Lyakh, A. M. and Bryantseva, Y. V. (2014) 'Seasonal polymorphism of the Black Seas dinoflagellates from genus *Ceratium*: *C. furca*, *C. fusus*, *C. tripos* (Dinophyceae). *Modern Phytomorphology*, *5*, pp. 209-214.

Malviya, S., Scalco, E., Audic, S., Vincent, F., Veluchamy, A., Poulain, J., Wincker, P., Iudicone, D., de Vargas, C., Bittner, L., Zingone, A. and Bowler, C. (2016) 'Insights into global diatom distribution and diversity in the world's ocean', *Proceedings of the National Academy of Sciences*, 113(11), pp. E1516–E1525. doi: 10.1073/pnas.1509523113.

Marcinko, C. L. J., Painter, S. C., Martin, A. P. and Allen, J. T. (2013) 'A review of the measurement and modelling of dinoflagellate bioluminescence', *Progress in Oceanography*, pp. 117–129. doi: 10.1016/j.pocean.2012.10.008.

Massana, R., Gobet, A., Audic, S., Bass, D., Bittner, L., Boutte, C., Chambouvet, A., Christen, R., Claverie, J. -M., Decelle, J., Dolan, J. R., Dunthorn, M., Edvardsen, B., Forn, I., Forster, D., Guillou, L., Jaillon, O., Kooistra, W. H. C. F., Logares, R., Mahé, F., Not, F., Ogata, H., Pawlowski, J., Pernice, M. C., Probert, I., Romac, S., Richards, T., Santini, S., Shalchian-Tabrizi, K., Siano, R., Simon, N., Stoeck, T., Vaulot, D., Zingone, A. and de Vargas, C. (2015) 'Marine protist diversity in

European coastal waters and sediments as revealed by high-throughput sequencing', *Environmental microbiology*, 17(10), pp. 4035–4049. doi: 10.1111/1462-2920.12955.

McCauley, L. A. R., Erdner, D. L., Nagai, S., Richlen, M. L. and Anderson, D. M. (2009) 'Biogeographic analysis of the globally distributed harmful algal bloom species *Alexandrium minutum* (Dinophyceae) based on rRNA gene sequences and microsatellite markers', *Journal of Phycology*, 45(2), pp. 454–463. doi: 10.1111/j.1529-8817.2009.00650.x.

McDonald, S. M., Sarno, D., Scanlan, D. J. and Zingone A. (2007) 'Genetic diversity of eukaryotic ultraphytoplankton in the Gulf of Naples during an annual cycle', *Aquatic Microbial Ecology*, 50(1), pp. 75–89. doi: 10.3354/ame01148.

McNally, K. L., Govind, N. S., Thomé, P. E. and Trench, R. K. (1994) 'Small-subunit ribosomal dna sequence analyses and a reconstruction of the inferred phylogeny among symbiotic dinoflagellates (Pyrrophyta)', *Journal of Phycology*, 30(2), pp. 316–329. doi: 10.1111/j.0022-3646.1994.00316.x.

Mendes, A., Reis, A., Vasconcelos, R., Guerra, P. and da Silva, T. (2009) '*Crypthecodinium cohnii* with emphasis on DHA production: A review', *Journal of Applied Phycology*, 21(2), pp. 199–214. doi: 10.1007/s10811-008-9351-3.

Mertens, K. N., Takano, Y., Yamaguchi, A., Gu, H., Bogus, K., Kremp., A., Bagheri, S., Matishov, G. and Matsuoka, K. (2015) 'The molecular characterization of the enigmatic dinoflagellate *Kolkwitziella acuta* reveals an affinity to the *Excentrica* section of the genus *Protoperidinium*', *Systematics and Biodiversity*, 13(6), pp. 509–524. doi: 10.1080/14772000.2015.1078855.

Milne, I., Lindner, D., Bayer, M., Husmeier, D., McGuire, G., Marshall, D. F. and Wright, F. (2009) 'TOPALi v2: A rich graphical interface for evolutionary analyses of multiple alignments on HPC clusters and multi-core desktops', *Bioinformatics*, 25(1), pp. 126–127. doi: 10.1093/bioinformatics/btn575.

Miranda, L. N., Zhuang, Y., Zhang, H. and Lin, S. (2012) 'Phylogenetic analysis guided by intragenomic SSU rDNA polymorphism refines classification of " *Alexandrium tamarense*" species complex', *Harmful Algae*, 16, pp. 35–48. doi: 10.1016/j.hal.2012.01.002.

Moestrup, Ø. and Daugbjerg, N. (2007) *On dinoflagellate phylogeny and classification*. In: Brodie J, Lewis J (eds) *Unravelling the algae — the past, present, and future of algal systematics.* The Systematics Association Special Volume Series 75, CRC Press, Boca Raton, pp 215–230.

Moestrup, Ø., Hansen, G., Daugbjerg, N., Flaim, G. and D'andrea, M. (2006) 'Studies on woloszynskioid dinoflagellates II: On *Tovellia sanguinea* sp. nov., the dinoflagellate responsible for the reddening of Lake Tovel, N. Italy', *European Journal of Phycology*, 41(1), pp. 47–65. doi: 10.1080/09670260600556682.

Moestrup, Ø., Akselmann, R., Fraga, S., Hoppenrath, M., Iwataki, M., Komárek, J., Larsen, J., Lundholm, N. and Zingone, A. (2009) *IOC-UNESCO Taxonomic Reference List of Harmful Micro Algae*, *Accessed at http://www.marinespecies.org/hab on 2016-05-24*. Available at: http://www.marinespecies.org/hab.

Moestrup, Ø., Hakanen, P., Hansen, G., Daugbjerg, N. and Ellegaard, M. (2014) 'On *Levanderina fissa* gen . & comb. nov . (Dinophyceae) (syn. *Gymnodinium fissum* , *Gyrodinium instriatum* , *Gyr. uncatenum*), a dinoflagellate with a very unusual sulcus', *Phycologia*, 53(3), pp. 265–292. doi: 10.2216/13-254.1.

Mollica, T. (2017) *Morphological and molecular analysis of the dinoflagellate genus Tripos in the Gulf of Naples with a single cell approach*. Master thesis. Università degli Studi di Napoli "Federico II"

Montresor, M. and Zingone, A. (1988) '*Scrippsiella precaria* sp. nov. (Dinophyceae) a marine dinoflagellate from the Gulf of Naples', *Phycologia*, 27, pp. 387–394.

Montresor, M., Sgrosso, S., Procaccini, G. and Kooistra, W. H. C. F. (2003) 'Intraspecific diversity in *Scrippsiella trochoidea* (Dinophyceae): evidence for cryptic species', *Phycologia*, 42(1), pp. 56–70. doi: https://doi.org/10.2216/i0031-8884-42-1-56.1.

Montresor, M., John, U., Beran, A. and Medlin, L. K. (2004) '*Alexandrium tamutum* sp. nov. (Dinophyceae): A new nontoxic species in the genus *Alexandrium*', *Journal of Phycology*, 40(2), pp. 398–411. doi: 10.1111/j.1529-8817.2004.03060.x.

Mordret, S., Romac, S., Henry, N., Colin, S., Carmichael, M., Berney, C., Audic, S., Richter, D. J., Pochon, X., de Vargas, C. and Decelle, J. (2016) 'The symbiotic life of *Symbiodinium* in the open ocean within a new species of calcifying ciliate (*Tiarina* sp.)', *ISME Journal*, 10(6). doi: 10.1038/ismej.2015.211.

Mordret, S., Piredda, R., Vaulot, D., Montresor, M., Kooistra, H. C. F. and Sarno, D. (2018) 'DinoREF: a curated dinoflagellate (Dinophyceae) reference database for the 18S rRNA gene. *Molecular Ecology Resources*. pp. 1-14. doi:10.1111/1755-0998.12781.

Müller, O. F. (1786) *Animalcula Infusoria Fluviatilis et Marina*. Havniae et Lipsiae, Copenhagen.

Mungpakdee, S., Shinzato, C., Takeuchi, T., Kawashima, T., Koyanagi, R., Hisata, K., Tanaka, M., Goto, H., Fujie, M., Lin, S., Satoh, N. and Shoguchi, E. (2014) 'Massive gene transfer and extensive RNA editing of a symbiotic dinoflagellate plastid genome', *Genome Biology and Evolution*, 6(6), pp. 1408–1422. doi: 10.1093/gbe/evu109.

Murray, S. A., Jørgensen, M. F., Ho, S. Y., Patterson, D. J. and Jermiin, L. S. (2005) 'Improving the analysis of dinoflagellate phylogeny based on rDNA', *Protist*, 156(3), pp. 269–286. doi: 10.1016/j.protis.2005.05.003.

Murray, S. A., Suggett, D. J., Doblin, M. A., Kohli, G. S., Seymour, J. R., Fabris, M. and Ralph, P. J. (2016) 'Unravelling the functional genetics of dinoflagellates: a review of approaches and opportunities', *Perspectives in Phycology*, 3(1), pp. 37–52. doi: 10.1127/pip/2016/0039.

Nash, E. A., Nisbet, R. E. R., Barbrook, A. C. and Howe, C. J. (2008) 'Dinoflagellates: a mitochondrial genome all at sea', *Trends in Genetics*, pp. 328–335. doi: 10.1016/j.tig.2008.04.001.

Naustvoll, L. J. (1998) 'Growth and grazing by the thecate heterotrophic dinoflagellate *Diplopsalis lenticula* (Diplopsalidaceae, Dinophyceae)', *Phycologia*, 37(1), pp. 1–9. doi: 10.2216/i0031-8884-37-1-1.1.

Nishimura, T., Sato, S., Tawong, W., Sakanari, H., Uehara, K., Shah, M. M. R., Suda, S., Yasumoto, T., Taira, Y., Yamaguchi, H. and Adachi, M. (2013) 'Genetic diversity and distribution of the ciguatera-causing dinoflagellate *Gambierdiscus* spp. (Dinophyceae) in coastal areas of Japan', *PLoS ONE*, 8(4). doi: 10.1371/journal.pone.0060882.

Not, F., Siano, R., Kooistra, W. H. C. F., Simon, N., Vaulot, D. and Probert, I. (2012) 'Diversity and ecology of eukaryotic marine phytoplankton', *Advances in Botanical Research*, 64, pp. 1–53. doi: 10.1016/B978-0-12-391499-6.00001-3.

Okamoto, N., Horák, A. and Keeling, P. J. (2012) 'Description of two species of early branching dinoflagellates, *Psammosa pacifica* n. g., n. sp. and *P. atlantica* n. sp', *PLoS ONE*, 7(6). doi: 10.1371/journal.pone.0034900.

Oksanen, J., Blanchet, F., Kindt, R., Legendre, P. and O'Hara, R. (2016) 'Vegan: community ecology package', *R package 2.3-3*, p. Available at: https://cran.r-project.org/web/packa. doi: 10.4135/9781412971874.n145.

Onda, D. F. L., Medrinal, E., Comeau, A. M., Thaler, A. M., Babin, M. and Lovejoy, C. (2017) 'Seasonal and Interannual Changes in Ciliate and Dinoflagellate Species Assemblages in the Arctic Ocean (Amundsen Gulf, Beaufort Sea, Canada)', *Frontiers in Marine Science*, 4(February), p. 16. doi: 10.3389/fmars.2017.00016.

Orr, R. J. S., Murray, S. A., Stüken, Rhodes, L. and Jakobsen, K. S. (2012) 'When naked became armored: an eight-gene phylogeny reveals monophyletic origin of theca in dinoflagellates', *PLoS ONE*, 7(11), pp. 1–15. doi: 10.1371/journal.pone.0050004.

Ostenfeld, C. H. (1903) *Phytoplankton from the sea around the Faeröes*. In Botany of the Faeröes. II, Det Nordiske Forlag, Copenhagen, pp 558-612.

Parfrey, L. W., Lahr, D. J., Knoll, A. H. and Katz, L. A. (2011) 'Estimating the timing of early eukaryotic diversification with multigene molecular clocks', *Proceedings of the National Academy of Sciences*, 108(33), pp. 13624–13629. doi: 10.1073/pnas.1110633108.

Pawlowski, J., Audic, S., Adl, S., Bass, D., Belbahri, L., Berney, C., Bowser, S. S., Cepicka, I., Decelle, J., Dunthorn, M., Fiore-Donno, A. M., Gile, G. H., Holzmann, M., Jahn, R., Jirků, M., Keeling, P. J., Kostka, M., Kudryavtsev, A., Lara, E., Lukeš, J., Mann, D. G., Mitchell, E. A. D., Nitsche, F., Romeralo, M., Saunders, G. W., Simpson, A. G. B., Smirnov, A. V., Spouge, J. L., Stern, R. F., Stoeck, T., Zimmermann, J., Schindel, D. and de Vargas, C. (2012) 'CBOL Protist Working Group: Barcoding Eukaryotic Richness beyond the Animal, Plant, and Fungal Kingdoms', *PLoS Biology*, 10(11). doi: 10.1371/journal.pbio.1001419.

Penna, A., Casabianca, S., Guerra, A. F., Vernesi, C. and Scardi, M. (2017) 'Analysis of phytoplankton assemblage structure in the Mediterranean Sea based on high-throughput sequencing of partial 18S rRNA sequences', *Marine Genomics*. Elsevier B.V., pp. 1–7. doi: 10.1016/j.margen.2017.06.001.

Percopo, I., Siano, R., Rossi, R., Soprano, V., Sarno, D. and Zingone, A. (2013) 'A new potentially toxic *Azadinium* species (Dinophyceae) from the Mediterranean Sea, *A. dexteroporum* sp. nov.', *Journal of Phycology*, 49(5), pp. 950–966. doi: 10.1111/jpy.12104.

Persson, A., Smith, B. C., Wikfors, G. H. and Quilliam, M. (2006) 'Grazing on toxic *Alexandrium fundyense* resting cysts and vegetative cells by the eastern oyster (Crassostrea virginica)', *Harmful Algae*, 5(6), pp. 678–684. doi: 10.1016/j.hal.2006.02.004.

Pfiester, L. A. (1989) 'Dinoflagellate Sexuality', *International Review of Cytology*, 114(C), pp. 249–272. doi: 10.1016/S0074-7696(08)60863-3.

Piredda, R., Tomasino, M. P., D'erchia, A. M., Manzari, C., Pesole, G., Montresor, M., Kooistra, W. H. C. F., Sarno, D. and Zingone, A. (2017) 'Diversity and temporal patterns of planktonic protist assemblages at a Mediterranean Long Term Ecological Research site', *FEMS Microbiology Ecology*, 93(1). doi: 10.1093/femsec/fiw200.

Pizay, M. D., Lemée, R., Simon, N., Cras, A. L., Laugier, J. P. and Dolan, J. R. (2009) 'Night and Day Morphologies in a Planktonic Dinoflagellate', *Protist*, 160(4), pp. 565–575. doi: 10.1016/j.protis.2009.04.003.

Pochon, X. and Gates, R. D. (2010) 'A new *Symbiodinium* clade (Dinophyceae) from soritid foraminifera in Hawaii', *Molecular Phylogenetics and Evolution*, 56(1), pp. 492–497. doi: 10.1016/j.ympev.2010.03.040.

Pochon, X., LaJeunesse, T. C. and Pawlowski, J. (2004) 'Biogeographic partitioning and host specialization among foraminiferan dinoflagellate symbionts (*Symbiodinium*; Dinophyta)', *Marine Biology*, 146(1), pp. 17–27. doi: 10.1007/s00227-004-1427-2.

Pochon, X., Putnam, H. M. and Gates, R. D. (2014) 'Multi-gene analysis of *Symbiodinium* dinoflagellates: a perspective on rarity, symbiosis, and evolution', *PeerJ*, 2, p. e394. doi: 10.7717/peerj.394.

Posada, D. (2008) 'jModelTest: Phylogenetic model averaging', *Molecular Biology and Evolution*, 25(7), pp. 1253–1256. doi: 10.1093/molbev/msn083.

Price, M. N., Dehal, P. S. and Arkin, A. P. (2010) 'FastTree 2 - Approximately maximum-likelihood trees for large alignments', *PLoS ONE*, 5(3). doi: 10.1371/journal.pone.0009490.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J. and Glöckner, F. O. (2013) 'The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools', *Nucleic Acids Research*, 41(D1). doi: 10.1093/nar/gks1219.

Radakovits, R., Jinkerson, R. E., Darzins, A. and Posewitz, M. C. (2010) 'Genetic engineering of algae for enhanced biofuel production', *Eukaryotic Cell*, pp. 486–501. doi: 10.1128/EC.00364-09.

Rampi L. and Bernard M. (1980) *Chiave per la determinazione delle peridinee pelagiche mediterranee*. Comitato Nazionale Energia Nucleare CNEN: 1-193.

R Development Core Team (2016) 'R: A Language and Environment for Statistical Computing', *R Foundation for Statistical Computing Vienna Austria*, 0, p. {ISBN} 3-900051-07-0. doi: 10.1038/sj.hdy.6800737.

Rein, K. S. and Snyder, R. V. (2006) 'The biosynthesis of polyketide metabolites by dinoflagellates', *Advances in Applied Microbiology*, pp. 93–125. doi: 10.1016/S0065-2164(06)59004-5.

Remsen, A., Hopkins, T. L. and Samson, S. (2004) 'What you see is not what you catch: A comparison of concurrently collected net, Optical Plankton Counter, and Shadowed Image Particle Profiling Evaluation Recorder data from the northeast Gulf of Mexico', *Deep-Sea Research Part I: Oceanographic Research Papers*, 51(1), pp. 129–151. doi: 10.1016/j.dsr.2003.09.008.

Reñé, A., Camp, J. and Garcés, E. (2015) 'Diversity and phylogeny of Gymnodiniales (Dinophyceae) from the NW Mediterranean Sea revealed by a morphological and molecular approach', *Protist*, 166(2), pp. 234–263. doi: 10.1016/j.protis.2015.03.001.

Ribera d'Alcalà, M., Conversano, F., Corato, F., Licandro, P., Mangoni, O., Marino, D., Mazzocchi, M. G., Modigh, M., Montresor, M., Nardella, M., Saggiomo, V., Sarno, D. and Zingone, A. (2004) 'Seasonal patterns in plankton communities in a pluriannual time series at a coastal Mediterranean site (Gulf of Naples): an attempt to discern recurrences and trends', *Scientia Marina*, 68(S1), pp. 65–83. doi: 10.3989/scimar.2004.68s165.

Robinson, N., Eglinton, G., Brassell, S. C. and Cranwell, P. A. (1984) 'Dinoflagellate origin for sedimentary 4α-methylsteroids and 5α(H)-stanols', *Nature*, 308, pp. 439–442. doi: 10.1038/312601a0.

Rognes, T., Flouri, T., Nichols, B., Quince, C. and Mahé, F. (2016) 'VSEARCH: a versatile open source tool for metagenomics', *PeerJ*, 4, p. e2584. doi: 10.7717/peerj.2584.

Rowan, R. and Powers, D. A. (1991) 'A molecular genetic classification of zooxanthellae and the evolution of animal-algal symbioses', *Science*, 251(4999), pp. 1348–1351. doi: 10.1126/science.251.4999.1348.

Rueckert, S. and Leander, B. S. (2008) 'Morphology and molecular phylogeny of *Haplozoon praxillellae* n. sp. (Dinoflagellata): A novel intestinal parasite of the maldanid polychaete *Praxillella pacifica* Berkeley', *European Journal of Protistology*, 44(4), pp. 299–307. doi: 10.1016/j.ejop.2008.04.004.

Ruggiero, M. V., Sarno, D., Barra, L., Kooistra, W. H. C. F., Montresor, M. and Zingone, A. (2015) 'Diversity and temporal pattern of *Pseudo-nitzschia* species (Bacillariophyceae) through the molecular lens', *Harmful Algae*, 42(1), pp. 15–24. doi: 10.1016/j.hal.2014.12.001.

Ruiz Sebastián, C. and O'Ryan, C. (2001) 'Single-cell sequencing of dinoflagellate (Dinophyceae) nuclear ribosomal genes', *Molecular Ecology Notes*, 1(4), pp. 329–331. doi: 10.1046/j.1471-8278.

de Salas, M. F., Bolch, C. J. S. and Hallegraeff, G. M. (2004) '*Karenia umbella* sp. nov. (Gymnodiniales, Dinophyceae), a new potentially ichthyotoxic dinoflagellate species from Tasmania, Australia', *Phycologia*, 43(2), pp. 166–175. doi: 10.2216/i0031-8884-43-2-166.1.

Salas, R., Tillmann, U. and Kavanagh, S. (2014) 'Morphological and molecular characterization of the small armoured dinoflagellate *Heterocapsa minima* (Peridiniales, Dinophyceae)', *European Journal of Phycology*. Taylor & Francis, 49(4), pp. 413–428. doi: 10.1080/09670262.2014.956800.

Saldarriaga, J. F., Taylor, F. J. R., Keeling, P. J. and Cavalier-Smith, T. (2001) 'Dinoflagellate nuclear SSU rRNA phylogeny suggests multiple plastid losses and replacements', *Journal of Molecular Evolution*, 53(3), pp. 204–213. doi: 10.1007/s002390010210.

Saldarriaga, J. F., Cavalier-Smith, T., Menden-Deuer, S. and Keeling, P. J. (2004) 'Molecular data and the evolutionary history of dinoflagellates', *European Journal of Protistology*, 40(1), pp. 85–111. doi: 10.1016/j.ejop.2003.11.003.

Santos, S. R., Shearer, T. L., Hannes, A. R. and Coffroth, M. A. (2004) 'Fine-scale diversity and specificity in the most prevalent lineage of symbiotic dinoflagellates (*Symbiodinium*, Dinophyceae) of the Caribbean', *Molecular Ecology*, 13(2), pp. 459–469. doi: 10.1046/j.1365-294X.2003.02058.x.

Schiller, J. (1931–1937) *Dinoflagellatae (Peridinineae) in monographischer Behandlung.* In: Rabenhorst, L. (ed.), *Kryptogamen-Flora von Deutschland, Österreichs und der Schweiz. Akad.* Verlag., Leipzig. Vol. 10 (3): Teil 1 (1–3) (1931–1933): Teil 2 (1–4)(1935–1937).

Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., Lesniewski, R. A., Oakley, B. B., Parks, D. H., Robinson, C. J., Sahl, J. W., Stres, B., Thallinger, G. G., Van Horn, D. J. and Weber, C. F. (2009) 'Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities', *Applied and Environmental Microbiology*, 75(23), pp. 7537–7541. doi: 10.1128/AEM.01541-09.

Schnepf, E. and Elbrächter, M. (1992) 'Nutritional strategies in dinoflagellates: A review with emphasis on cell biological aspects', *European Journal of Protistology*, 28(1), pp. 3–24. doi: 10.1016/S0932-4739(11)80315-9.

Schrank, F. von Paula (1793) *Mikroskopische Wahrnehmungen*. Der Naturforscher 27:26–37.

Segata, N., Izard, J., Waldron, L., Gevers, D., Miropolsky, L., Garrett, W. S. and Huttenhower, C. (2011) 'Metagenomic biomarker discovery and explanation', *Genome Biology*, 12(6), p. R60. doi: 10.1186/gb-2011-12-6-r60.

Semina H. J., Levashova S. S. (1993) *The biogeography of tropical phytoplankton species in the Pacific Ocean*. Int. Revue ges. Hydrobiol. Hydrogr. 1993, 78, 243–262.

Seo, K. S. and Fritz, L. (2000) 'Cell-wall morphology correlated with vertical migration in the non-motile marine dinoflagellate *Pyrocystis noctiluca*', *Marine Biology*, 137(4), pp. 589–594. doi: 10.1007/s002270000374.

Shaked, Y. and de Vargas, C. (2006) 'Pelagic photosymbiosis: rDNA assessment of diversity and evolution of dinoflagellate symbionts and planktonic foraminiferal hosts', *Marine Ecology Progress Series*, 325, pp. 59–71. doi: 10.3354/meps325059.

Sherr, E. B. and Sherr, B. F. (2007) 'Heterotrophic dinoflagellates: A significant component of microzooplankton biomass and major grazers of diatoms in the sea', *Marine Ecology Progress Series*, pp. 187–197. doi: 10.3354/meps07161.

Siano, R. and Montresor, M. (2005) 'Morphology, ultrastructure and feeding behaviour of *Protoperidinium vorax* sp. nov. (Dinophyceae, Peridiniales)', *European Journal of Phycology*, 40, pp. 221–232. doi: 10.1080/09670260500128293.

Skovgaard, A. (1996) 'Engulfment of *Ceratium* spp. (Dinophyceae) by the thecate photosynthetic dinoflagellate *Fragilidium subglobosum*', *Phycologia*, 35(6), pp. 490–499. doi: 10.2216/i0031-8884-35-6-490.1.

Skovgaard, A. (2005) 'Infection with the dinoflagellate parasite *Blastodinium* spp. in two Mediterranean copepods', *Aquatic Microbial Ecology*, 38(1), pp. 93–101. doi: 10.3354/ame038093.

Skovgaard, A. and Saiz, E. (2006) 'Seasonal occurrence and role of protistan parasites in coastal marine zooplankton', *Marine Ecology Progress Series*, 327(Chatton 1920), pp. 37–49. doi: 10.3354/meps327037.

Skovgaard, A. and Salomonsen, X. M. (2009) '*Blastodinium galatheanum* sp. nov. (Dinophyceae) a parasite of the planktonic copepod *Acartia negligens* (Crustacea, Calanoida) in the central Atlantic Ocean', *European Journal of Phycology*, 44(3), pp. 425–438. doi: 10.1080/09670260902878743.

Skovgaard, A., Karpov, S. A. and Guillou, L. (2012) 'The parasitic dinoflagellates *Blastodinium* spp. Inhabiting the gut of marine, Planktonic copepods: Morphology, ecology, and unrecognized species diversity', *Frontiers in Microbiology*. doi: 10.3389/fmicb.2012.00305.

Smalley, G. W. and Coats, D. W. (2002) 'Ecology of the red-tide dinoflagellate *Ceratium furca*: distribution, mixotrophy, and grazing impact on ciliate populations of Chesapeake Bay.', *The Journal of eukaryotic microbiology*, 49(1), pp. 63–73. doi: 10.1111/j.1550-7408.2002.tb00343.x.

Smalley, G. W., Coats, D. W. and Adam, E. J. (1999) 'A new method using fluorescent microspheres to determine grazing on ciliates by the mixotrophic dinoflagellate *Ceratium furca*', *Aquatic Microbial Ecology*, 17(2), pp. 167–179. doi: 10.3354/ame017167.

Smalley, G. W., Coats, D. W. and Stoecker, D. K. (2003) 'Feeding in the mixotrophic dinoflagellate *Ceratium furca* is influenced by intracellular nutrient concentrations', *Marine Ecology Progress Series*, 262, pp. 137–151. doi: 10.3354/meps262137.

Smayda, T. J. (1997) 'Harmful algal blooms: Their ecophysiology and general relevance to phytoplankton blooms in the sea', *Limnology and Oceanography*, 42(5part2), pp. 1137–1153. doi: 10.4319/lo.1997.42.5_part_2.1137.

Smayda, T. J. (2002) 'Turbulence, watermass stratification and harmful algal blooms: An alternative view and frontal zones as "pelagic seed banks"', *Harmful Algae*, 1(1), pp. 95–112. doi: 10.1016/S1568-9883(02)00010-0.

Smayda, T. J. and Reynolds, C. S. (2003) 'Strategies of marine dinoflagellate survival and some rules of assembly', in *Journal of Sea Research*, pp. 95–106. doi: 10.1016/S1385-1101(02)00219-8.

Smetacek, V. (2012) 'Making sense of ocean biota: How evolution and biodiversity of land organisms differ from that of the plankton', *Journal of Biosciences*, 37(4), pp. 589–607. doi: 10.1007/s12038-012-9240-4.

Smith, K. F., Kohli, G. S., Murray, S. A. and Rhodes, L. L. (2017) 'Assessment of the metabarcoding approach for community analysis of benthic-epiphytic dinoflagellates using mock communities', *New Zealand Journal of Marine and Freshwater Research*. Taylor & Francis, 0(0), pp. 1–22. doi: 10.1080/00288330.2017.1298632.

Sournia, A. (1967) *Le genre Ceratium (Péridinien planctonique) dans le canal de Mozambique, Contribution à une révision mondiale*. Vie et Milieu, Série A: Biologie marine, 28:375-500.

Sournia, A. (1973) 'Catalogue des espèces et taxons infraspécifiques de dinoflagellés marins actuels publiés depuis la révision de J. Schiller. I. Dinoflagellés libres. Beih', *Nova Hedwigia*. 48, 1–92.

Sournia, A. (1978) 'Catalogue des espèces et taxons infra-spécifiques de dinoflagellés marins actuels publiés depuis la révision de J. Schiller. III (Complément)', *Rev. Algol*, **13**, 3–40.

Sournia, A. (1990) 'Catalogue des espèces et taxons infra-spécifiques de dinoflagellés marins actuels publiés depuis la révision de J. Schiller. V. (Complément)', *Acta Protozool,* 29, 321–346.

Sournia, A. (1993) 'Catalogue des espèces et taxons infra-spécifiques de dinoflagellés marins actuels publiés depuis la révision de J. Schiller. VI. (Complément)', *Cryptog. Algol,* 14, 133–144.

Sournia, A. (1986) *Atlas du Phytoplancton Marin. Vol. I: Introduction, Cyanophycées, Dictyochophycées, Dinophycées et Raphidophycées.* Editions du CNRS, Paris.

Stamatakis, A. (2014) 'RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies', *Bioinformatics*, 30(9), pp. 1312–1313. doi: 10.1093/bioinformatics/btu033.

Stanley, G. D. (2006) 'Photosymbiosis and the evolution of modern coral reefs', *Science*, pp. 857–858. doi: 10.1126/science.1123701.

Steemann Nielsen, E. (1939) *Die Ceratien des Indischen Ozeans und der Ostasiatische Gewässer mit einer allgemeinen Zusammenfassung über die Verbreitung der Ceratien in den Weltmeeren*. Carlsberg Foundation, Copenhagen.

Steidinger, K. and Tangen, K. (1997) 'Dinoflagellates', *Identifying marine diatoms and dinoflagellates*, pp. 387–584. doi: 10.1016/S0025-3227(97)81154-1.

Stentiford, G. D. and Shields, J. D. (2005) 'A review of the parasitic dinoflagellates *Hematodinium* species and *Hematodinium*-like infections in marine crustaceans', *Diseases of Aquatic Organisms*, pp. 47–70. doi: 10.3354/dao066047.

Stern, R. F., Horak, A., Andrew, R. L., Coffroth, M. A., Andersen, R. A., Küpper, F. C., Jameson, I., Hoppenrath, M., Véron, B., Kasai, F., Brand, J., James, E. R. and Keeling, P. J. (2010) 'Environmental barcoding reveals massive dinoflagellate diversity in marine environments', *PLoS ONE*, 5(11). doi: 10.1371/journal.pone.0013991.

Stern, R. F., Andersen, R. A., Jameson, I., Küpper, F. C., Coffroth, M. A., Vaulot, D., Le Gall, F., Véron, B., Brand, J., Skelton, H., Kasai, F., Lilly, E. L. and Keeling, P. J. (2012) 'Evaluating the ribosomal internal transcribed spacer (its) as a candidate dinoflagellate barcode marker', *PLoS ONE*, 7(8). doi: 10.1371/journal.pone.0042780.

Stoddard, S. F., Smth, B. J., Hein, R., Roller, B. R. and Schmidt, T. M. (2014). rrn DB: improved tools for interpreting rRNA gene abundance in bacteria and archaea and a new foundation dor future development. *Nucleic acids research*, 43(1), pp. 593-598.

Stoeck, T., Hayward, B., Taylor, G. T., Varela, R. and Epstein, S. S. (2006) 'A multiple PCR-primer approach to access the microeukaryotic diversity in environmental samples', *Protist*, 157(1), pp. 31–43. doi: 10.1016/j.protis.2005.10.004.

Stoecker, D. K. (1999) 'Mixotrophy among Dinoflagellates', *The Journal of Eukaryotic Microbiology*, 46(4), pp. 397–401. doi: 10.1111/j.1550-7408.1999.tb04619.x.

Stoecker, D. K., Johnson, M. D., de Vargas, C. and Not, F. (2009) 'Acquired phototrophy in

aquatic protists', *Aquatic Microbial Ecology*, 57(3), pp. 279–310. doi: 10.3354/ame01340.

Sweeney, B. M. (1976) '*Pedinomonas noctilucae* (Prasinophyceae), the flagellate symbiotic in *Noctiluca* (Dinophyceae) in southeast Asia', *Journal of Phycology*, 12(4), pp. 460–464. doi: 10.1111/j.1529-8817.1976.tb02874.x.

Takano, Y., Hansen, G., Fujita, D. and Horiguchi, T. (2008) 'Serial replacement of diatom endosymbionts in two freshwater dinoflagellates, *Peridiniopsis* spp. (Peridiniales, Dinophyceae)', *Phycologia*, 47(1), pp. 41–53. doi: 10.2216/07-36.1.

Takano, Y., Yamaguchi, H., Inouye, I., Moestrup, Ø. and Horiguchi, T. (2014) 'Phylogeny of five species of *Nusuttodinium* gen. nov. (Dinophyceae), a genus of unarmoured kleptoplastidic dinoflagellates', *Protist*, 165(6), pp. 759–778. doi: 10.1016/j.protis.2014.09.001.

Taylor, F. J. R. (1987) *General group characteristics; special features of interest; short history of dinoflagellate study. The biology of dinoflagellate*s, 21, 1-23.

Taylor, F. J. R. (2004) *Extraordinary Dinoflagellates: Past and Present.* In *Neo-Science of Natural History: Integration of Geoscience and Biodiversity Studies: Proceedings of International Symposium on" Dawn of a New Natural History-Integration of Geoscience and Biodiversity Studies" March 5-6, 2004, Sapporo* (pp. 61-66). Graduate School of Science, Hokkaido University.

Taylor, F. J. R., Hoppenrath, M. and Saldarriaga, J. F. (2008) 'Dinoflagellate diversity and distribution', *Biodiversity and Conservation*, pp. 407–418. doi: 10.1007/s10531-007-9258-3.

Tennekes, M. and Ellis, P. (2017) 'Package "treemap"'.

Thessen, A. E., Patterson, D. J. and Murray, S. A. (2012) 'The taxonomic significance of species that have only been observed once: the genus *Gymnodinium* (Dinoflagellata) as an example', *PLoS ONE*, 7(8). doi: 10.1371/journal.pone.0044015.

Thornhill, D. J., Lajeunesse, T. C. and Santos, S. R. (2007) 'Measuring rDNA diversity in eukaryotic microbial systems: How intragenomic variation, pseudogenes, and PCR artifacts confound biodiversity estimates', *Molecular Ecology*, 16(24), pp. 5326–5340. doi: 10.1111/j.1365-294X.2007.03576.x.

Tillmann, U., Elbrächter, M., Krock, B., John, U. and Cembella, A. (2009) ' *Azadinium spinosum* gen. et sp. nov. (Dinophyceae) identified as a primary producer of azaspiracid toxins', *European Journal of Phycology*, 44(1), pp. 63–79. doi: 10.1080/09670260802578534.

Tillmann, U., Gottschling, M., Nézan, E., Krock, B. and Bilien, G. (2014) 'Morphological and molecular characterization of three new *Azadinium* species (Amphidomataceae, Dinophyceae) from the Irminger Sea', *Protist*, 165(4), pp. 417–444. doi: 10.1016/j.protis.2014.04.004.

Tragin, M., Zingone, A. and Vaulot, D. (2017) 'Comparison of coastal phytoplankton composition estimated from the V4 and V9 regions of the 18S rRNA gene with a focus on photosynthetic groups and especially Chlorophyta', *Environmental Microbiology*, 0. doi: 10.1111/1462-2920.13952.

Trench, R. K. (1993) 'Microalgal-invertebrate symbioses - a review', *Endocytobiosis And Cell Research*, 9, pp. 135–175. Available at: http://www.citeulike.org/group/894/article/511409.

Trench, R. K. and Thinh, L. (1995) '*Gymnodinium linucheae* sp. nov.: The dinoflagellate symbiont of the jellyfish *Linuche unguiculata*', *European Journal of Phycology*, 30(2), pp. 149–154. doi: 10.1080/09670269500650911.

Tubaro, A., Dell'Ovo, V., Sosa, S. and Florio, C. (2010) 'Yessotoxins: A toxicological overview', *Toxicon*, 56(2), pp. 163–172. doi: 10.1016/j.toxicon.2009.07.038.

Tunin-Ley, A. and Lemée, R. (2013) 'The Genus *Neoceratium* (Planktonic Dinoflagellates) as a

potential indicator of ocean warming', *Microorganisms*, 1(1), pp. 58–70. doi: 10.3390/microorganisms1010058.

Tunin-Ley, A., Labat, J. P., Gasparini, S., Mousseau, L. and Lemée, R. (2007) 'Annual cycle and diversity of species and infraspecific taxa of *Ceratium* (Dinophyceae) in the Ligurian Sea, northwest Mediterranean', *Journal of Phycology*, 43(6), pp. 1149–1163. doi: 10.1111/j.1529-8817.2007.00417.x.

Tunin-Ley, A., Ibañez, F., Labat, J. P., Zingone, A. and Lemée, R. (2009) 'Phytoplankton biodiversity and NW Mediterranean Sea warming: Changes in the dinoflagellate genus *Ceratium* in the 20th century', *Marine Ecology Progress Series*, 375, pp. 85–99. doi: 10.3354/meps07730.

Valiadi, M. and Iglesias-Rodriguez, D. (2013) 'Understanding bioluminescence in dinoflagellates—how far have we come?', *Microorganisms*, 1(1), pp. 3–25. doi: 10.3390/microorganisms1010003.

Vanhöffen, E. (1896) 'Das Genus *Ceratium*', *Zoologischer Anzeiger,* 19(499): 133-134.

de Vargas, C., Audic, S., Henry, N., Decelle, J., Mahé, F., Logares, R., Lara, E., Berney, C., Le Bescot, N., Probert, I., Carmichael, M., Poulain, J., Romac, S., Colin, S., Aury, J. -M., Bittner, L., Chaffron, S., Dunthorn, M., Engelen. S., Flegontova, O., Guidi, L., Horák, A., Jaillon, O., Lima-Mendez, G., Lukeš, J., Malviya, S., Morard, R., Mulot, M., Scalco, E., Siano, R., Vincent, F., Zingone, A., Dimier, C., Picheral, M., Searson, S., Kandels-Lewis, S., Tara Oceans Coordinators, Acinas, S., Bork, P., Bowler, C., Gorsky, G., Grimsley, N., Hingamp, P., Iudicone, D., Not, F., Ogata, H., Pesant, S., Raes, J., Sieracki, M. E., Speich, S., Stemmann, L., Sunagawa, S., Weissenbach, J., Wincker, P. and Karsenti, E. (2015) 'Eukaryotic plankton diversity in the sunlit ocean', *Science*, 348(6237), pp. 1261605–1261605. doi: 10.1126/science.1261605.

Vázquez-Domínguez, E., Vaqué, D. and Gasol, J. M. (2012) 'Temperature effects on the heterotrophic bacteria, heterotrophic nanoflagellates, and microbial top predators of the NW Mediterranean', *Aquatic Microbial Ecology*, 67(2), pp. 107–121. doi: 10.3354/ame01583.

Van de Waal, D. B., John, U., Ziveri, P., Reichart, G. J., Hoins, M., Sluijs, A. and Rost, B. (2013) 'Ocean Acidification Reduces Growth and Calcification in a Marine Dinoflagellate', *PLoS ONE*, 8(6). doi: 10.1371/journal.pone.0065987.

Walker, L.M. (1984) *Life histories, dispersal and survival in marine, planktonic dinoflagellates.* In Steidinger, K. A. and Walker, L. M. (Eds.) Marine Plankton Life Cycle Strategies. CRC press, Florida, USA, pp19-34.

Wall, D. and Dale, B. (1969) 'The hystrichosphaerid resting spore of the dinoflagellate *Pyrodinium bahamense*, Platf, 1906', *Journal of Phycology*, 5(2), pp. 140–149. doi: 10.1111/j.1529-8817.1969.tb02595.x.

Waller, R. F. and Jackson, C. J. (2009) 'Dinoflagellate mitochondrial genomes: Stretching the rules of molecular biology', *BioEssays*, 31(2), pp. 237–245. doi: 10.1002/bies.200800164.

Wang, D. Z. (2008) 'Neurotoxins from marine dinoflagellates: A brief review', *Marine Drugs*, pp. 349–371. doi: 10.3390/md20080016.

Wang, Q., Garrity, G. M., Tiedje, J. M. and Cole, J. R. (2007) 'Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy', *Applied and environmental microbiology*, 73(16), pp. 5261–5267. doi: 10.1128/AEM.00062-07.

Wetherbee, R., Andersen, R.A. and Pickett-Heaps, J.D. (Eds.). (2012) *The protistan cell surface*. Springer Science & Business Media. p.37.

Whitney, S. M., Shaw, D. C. and Yellowlees, D. (1995) 'Evidence that some dinoflagellates

contain a ribulose-1,5-bisphosphate carboxylase/oxygenase related to that of the alpha-proteobacteria.', *Proceedings. Biological sciences / The Royal Society*, 259(1356), pp. 271–5. doi: 10.1098/rspb.1995.0040.

Wickham, H. (2009) *ggplot2 Elegant Graphics for Data Analysis*, *Media*. doi: 10.1007/978-0-387-98141-3.

Wiggan, N. J., Riding, J. B. and Franz, M. (2017) 'Resolving the Middle Jurassic dinoflagellate radiation: The palynology of the Bajocian of Swabia, southwest Germany', *Review of Palaeobotany and Palynology*. The Authors, 238, pp. 55–87. doi: 10.1016/j.revpalbo.2016.11.010.

Wisecaver, J. H., Brosnahan, M. L. and Hackett, J. D. (2013) 'Horizontal gene transfer is a significant driver of gene innovation in dinoflagellates', *Genome Biology and Evolution*, 5(12), pp. 2368–2381. doi: 10.1093/gbe/evt179.

Yamada, N., Tanaka, A. and Horiguchi, T. (2015) 'Pigment compositions are linked to the habitat types in dinoflagellates', *Journal of Plant Research*, 128(6), pp. 923–932. doi: 10.1007/s10265-015-0745-4.

Yamaguchi, A. and Horiguchi, T. (2005) 'Molecular phylogenetic study of the heterotrophic dinoflagellate genus *Protoperidinium* (Dinophyceae) inferred from small subunit rRNA gene sequences', *Phycological Research*, 53(1), pp. 30–42. doi: 10.1111/j.1440-183.2005.00370.x.

Zapata, M., Fraga, S., Rodríguez, F. and Garrido, J. L. (2012) 'Pigment-based chloroplast types in dinoflagellates', *Marine Ecology Progress Series*, 465, pp. 33–52. doi: 10.3354/meps09879.

Zhang, H., Bhattacharya, D. and Lin, S. (2005) 'Phylogeny of dinoflagellates based on mitochondrial cytochrome b and nuclear small subunit rDNA sequence comparisons', *Journal of Phycology*, 41(2), pp. 411–420. doi: 10.1111/j.1529-8817.2005.04168.x.

Zhang, H., Bhattacharya, D. and Lin, S. (2007) 'A three-gene dinoflagellate phylogeny suggests monophyly of Prorocentrales and a basal position for *Amphidinium* and *Heterocapsa*', *Journal of Molecular Evolution*, 65(4), pp. 463–474. doi: 10.1007/s00239-007-9038-4.

Zingone, A., Siano, R., D'Alelio, D. and Sarno, D. (2006) 'Potentially toxic and harmful microalgae from coastal waters of the Campania region (Tyrrhenian Sea, Mediterranean Sea)', *Harmful Algae*, 5(3), pp. 321–337. doi: 10.1016/j.hal.2005.09.002.

Zingone, A., Dubroca, L., Iudicone, D., Margiotta, F., Corato, F., Ribera d'Alcalà, M., Saggiomo, V. and Sarno, D. (2010) 'Coastal phytoplankton do not rest in winter', *Estuaries and Coasts*, 33(2), pp. 342–361. doi: 10.1007/s12237-009-9157-9.

Zonneveld, K. A. F., Marret, F., Versteegh, G. J., Bogus, K., Bonnet, S., Bouimetarhan, I., Crouch, E., de Vernal, A., Elshanawany, R., Edwards, L. , Esper, O., Forke, S., Grøsfjeld, K., Henry, M., Holzwarth, U., Kielt, J. -F., Kim, S. -Y., Ladouceur, S., Ledu, D., Chen, L., Limoges, A., Londeix, L., Lu, S. -H., Mahmoud, M. S., Marino, G., Matsouka, K., Matthiessen, J., Mildenhal, D. C., Mudie, P., Neil, H. L., Pospelova, V., Qi, Y., Radi, T., Richerol, T., Rochon, A., Sangiorgi, F., Solignac, S., Turon, J. -L., Verleye, T., Wang, Y., Wang, Z. and Young, M. (2013) 'Atlas of modern dinoflagellate cyst distribution based on 2405 data points', *Review of Palaeobotany and Palynology*, 191, pp. 1–197. doi: 10.1016/j.revpalbo.2012.08.003.

CHAPTER VII: Appendix

Appendix 1: Summary of sequences represented in the DinoREF database.

	GenBank 18S	Superclade #	ORDER	SPECIES (valid name)
1	JQ639757	Super Clade 1	Gonyaulacales	*Ceratium furcoides*
2	JQ639758	Super Clade 1	Gonyaulacales	*Ceratium furcoides*
3	JQ639759	Super Clade 1	Gonyaulacales	*Ceratium hirundinella*
4	AY443014	Super Clade 1	Gonyaulacales	*Ceratium hirundinella*
5	DQ487192	Super Clade 1	Gonyaulacales	*Ceratium hirundinella*
6	EU025021	Super Clade 1	Gonyaulacales	*Ceratium hirundinella*
7	FJ402948	Super Clade 1	Gonyaulacales	*Tripos pentagonus*
8	FJ402949	Super Clade 1	Gonyaulacales	*Tripos declinatus*
9	FJ402951	Super Clade 1	Gonyaulacales	*Tripos petersenii*
10	FJ402953	Super Clade 1	Gonyaulacales	*Tripos petersenii*
11	FJ402954	Super Clade 1	Gonyaulacales	*Tripos azoricus*
12	FJ824911	Super Clade 1	Gonyaulacales	*Tripos platycornis*
13	FJ402964	Super Clade 1	Gonyaulacales	*Tripos minutus*
14	FJ402955	Super Clade 1	Gonyaulacales	*Tripos candelabrus*
15	FJ402945	Super Clade 1	Gonyaulacales	*Tripos candelabrus*
16	FJ402944	Super Clade 1	Gonyaulacales	*Tripos concilians*
17	FJ402950	Super Clade 1	Gonyaulacales	*Tripos concilians*
18	FJ402962	Super Clade 1	Gonyaulacales	*Tripos limulus*
19	FJ402965	Super Clade 1	Gonyaulacales	*Tripos paradoxides*
20	DQ388462	Super Clade 1	Gonyaulacales	*Tripos longipes*
21	FJ402956	Super Clade 1	Gonyaulacales	*Tripos arietinus*
22	FJ402947	Super Clade 1	Gonyaulacales	*Tripos symmetricus*
23	FJ402946	Super Clade 1	Gonyaulacales	*Tripos euarcuatus*
24	FJ402952	Super Clade 1	Gonyaulacales	*Tripos limulus*
25	FJ402963	Super Clade 1	Gonyaulacales	*Tripos kofoidii*
26	FJ402942	Super Clade 1	Gonyaulacales	*Tripos massiliensis*
27	FJ402959	Super Clade 1	Gonyaulacales	*Tripos contrarius*
28	FJ402960	Super Clade 1	Gonyaulacales	*Tripos horridus*
29	FJ402943	Super Clade 1	Gonyaulacales	*Tripos hexacanthus*
30	FJ402966	Super Clade 1	Gonyaulacales	*Tripos furca*
31	AJ276699	Super Clade 1	Gonyaulacales	*Tripos furca*
32	FJ402961	Super Clade 1	Gonyaulacales	*Tripos gravidus*
33	FJ824910	Super Clade 1	Gonyaulacales	*Tripos digitatus*
34	FJ402957	Super Clade 1	Gonyaulacales	*Tripos extensus*
35	FJ402958	Super Clade 1	Gonyaulacales	*Tripos fusus*
36	AF022153	Super Clade 1	Gonyaulacales	*Tripos fusus*
37	AF022192	Super Clade 1	Gonyaulacales	*Tripos tenuis*
38	AB764275	Super Clade 1	Gonyaulacales	*Gambierdiscus scabrosus*
39	AB764276	Super Clade 1	Gonyaulacales	*Gambierdiscus scabrosus*
40	AB764270	Super Clade 1	Gonyaulacales	*Gambierdiscus scabrosus*
41	AB605806	Super Clade 1	Gonyaulacales	*Gambierdiscus scabrosus*
42	AB605807	Super Clade 1	Gonyaulacales	*Gambierdiscus scabrosus*
43	AB764271	Super Clade 1	Gonyaulacales	*Gambierdiscus scabrosus*
44	AB764274	Super Clade 1	Gonyaulacales	*Gambierdiscus scabrosus*
45	AB605812	Super Clade 1	Gonyaulacales	*Gambierdiscus scabrosus*
46	AB764272	Super Clade 1	Gonyaulacales	*Gambierdiscus scabrosus*
47	AB764273	Super Clade 1	Gonyaulacales	*Gambierdiscus scabrosus*
48	AB764234	Super Clade 1	Gonyaulacales	*Gambierdiscus scabrosus*
49	AB764252	Super Clade 1	Gonyaulacales	*Gambierdiscus scabrosus*
50	AB764238	Super Clade 1	Gonyaulacales	*Gambierdiscus scabrosus*
51	AB764250	Super Clade 1	Gonyaulacales	*Gambierdiscus scabrosus*
52	AB764254	Super Clade 1	Gonyaulacales	*Gambierdiscus scabrosus*
53	AB764265	Super Clade 1	Gonyaulacales	*Gambierdiscus scabrosus*
54	AB764239	Super Clade 1	Gonyaulacales	*Gambierdiscus scabrosus*
55	AB764242	Super Clade 1	Gonyaulacales	*Gambierdiscus scabrosus*
56	AB764264	Super Clade 1	Gonyaulacales	*Gambierdiscus scabrosus*
57	AB764246	Super Clade 1	Gonyaulacales	*Gambierdiscus scabrosus*
58	AB764259	Super Clade 1	Gonyaulacales	*Gambierdiscus scabrosus*

59	AB764253	Super Clade 1	Gonyaulacales	*Gambierdiscus scabrosus*
60	AB764251	Super Clade 1	Gonyaulacales	*Gambierdiscus scabrosus*
61	AB764256	Super Clade 1	Gonyaulacales	*Gambierdiscus scabrosus*
62	AB764245	Super Clade 1	Gonyaulacales	*Gambierdiscus scabrosus*
63	AB764244	Super Clade 1	Gonyaulacales	*Gambierdiscus scabrosus*
64	AB764240	Super Clade 1	Gonyaulacales	*Gambierdiscus scabrosus*
65	AB764236	Super Clade 1	Gonyaulacales	*Gambierdiscus scabrosus*
66	AB764249	Super Clade 1	Gonyaulacales	*Gambierdiscus scabrosus*
67	AB764231	Super Clade 1	Gonyaulacales	*Gambierdiscus scabrosus*
68	AB764237	Super Clade 1	Gonyaulacales	*Gambierdiscus scabrosus*
69	AB764262	Super Clade 1	Gonyaulacales	*Gambierdiscus scabrosus*
70	AB764263	Super Clade 1	Gonyaulacales	*Gambierdiscus scabrosus*
71	AB764241	Super Clade 1	Gonyaulacales	*Gambierdiscus scabrosus*
72	AB764257	Super Clade 1	Gonyaulacales	*Gambierdiscus scabrosus*
73	AB764266	Super Clade 1	Gonyaulacales	*Gambierdiscus scabrosus*
74	AB764229	Super Clade 1	Gonyaulacales	*Gambierdiscus scabrosus*
75	AB764248	Super Clade 1	Gonyaulacales	*Gambierdiscus scabrosus*
76	AB764232	Super Clade 1	Gonyaulacales	*Gambierdiscus scabrosus*
77	AB764235	Super Clade 1	Gonyaulacales	*Gambierdiscus scabrosus*
78	AB764243	Super Clade 1	Gonyaulacales	*Gambierdiscus scabrosus*
79	AB764255	Super Clade 1	Gonyaulacales	*Gambierdiscus scabrosus*
80	AB764233	Super Clade 1	Gonyaulacales	*Gambierdiscus scabrosus*
81	AB764230	Super Clade 1	Gonyaulacales	*Gambierdiscus scabrosus*
82	AB764258	Super Clade 1	Gonyaulacales	*Gambierdiscus scabrosus*
83	AB764247	Super Clade 1	Gonyaulacales	*Gambierdiscus scabrosus*
84	AB764261	Super Clade 1	Gonyaulacales	*Gambierdiscus scabrosus*
85	AB605801	Super Clade 1	Gonyaulacales	*Gambierdiscus scabrosus*
86	AB605800	Super Clade 1	Gonyaulacales	*Gambierdiscus scabrosus*
87	AB764267	Super Clade 1	Gonyaulacales	*Gambierdiscus scabrosus*
88	AB605811	Super Clade 1	Gonyaulacales	*Gambierdiscus scabrosus*
89	AB605799	Super Clade 1	Gonyaulacales	*Gambierdiscus scabrosus*
90	AB764268	Super Clade 1	Gonyaulacales	*Gambierdiscus scabrosus*
91	AB764269	Super Clade 1	Gonyaulacales	*Gambierdiscus scabrosus*
92	AB499536	Super Clade 1	Gonyaulacales	*Gambierdiscus* sp.
93	AB499535	Super Clade 1	Gonyaulacales	*Gambierdiscus* sp.
94	EF202872	Super Clade 1	Gonyaulacales	*Gambierdiscus pacificus*
95	EF202864	Super Clade 1	Gonyaulacales	*Gambierdiscus pacificus*
96	EF202863	Super Clade 1	Gonyaulacales	*Gambierdiscus pacificus*
97	EF202862	Super Clade 1	Gonyaulacales	*Gambierdiscus pacificus*
98	EF202865	Super Clade 1	Gonyaulacales	*Gambierdiscus pacificus*
99	EF202875	Super Clade 1	Gonyaulacales	*Gambierdiscus pacificus*
100	EF202874	Super Clade 1	Gonyaulacales	*Gambierdiscus pacificus*
101	EF202861	Super Clade 1	Gonyaulacales	*Gambierdiscus pacificus*
102	EF202871	Super Clade 1	Gonyaulacales	*Gambierdiscus pacificus*
103	EF202873	Super Clade 1	Gonyaulacales	*Gambierdiscus pacificus*
104	EF202882	Super Clade 1	Gonyaulacales	*Gambierdiscus toxicus*
105	EF202890	Super Clade 1	Gonyaulacales	*Gambierdiscus toxicus*
106	EF202885	Super Clade 1	Gonyaulacales	*Gambierdiscus toxicus*
107	EF202881	Super Clade 1	Gonyaulacales	*Gambierdiscus toxicus*
108	EF202889	Super Clade 1	Gonyaulacales	*Gambierdiscus toxicus*
109	EF202888	Super Clade 1	Gonyaulacales	*Gambierdiscus toxicus*
110	EF202884	Super Clade 1	Gonyaulacales	*Gambierdiscus toxicus*
111	EF202878	Super Clade 1	Gonyaulacales	*Gambierdiscus toxicus*
112	EF202883	Super Clade 1	Gonyaulacales	*Gambierdiscus toxicus*
113	EF202886	Super Clade 1	Gonyaulacales	*Gambierdiscus toxicus*
114	EF202880	Super Clade 1	Gonyaulacales	*Gambierdiscus toxicus*
115	EF202887	Super Clade 1	Gonyaulacales	*Gambierdiscus toxicus*
116	EF202879	Super Clade 1	Gonyaulacales	*Gambierdiscus toxicus*
117	EF202876	Super Clade 1	Gonyaulacales	*Gambierdiscus belizeanus*
118	DQ388463	Super Clade 1	Gonyaulacales	*Gambierdiscus toxicus*
119	EF202877	Super Clade 1	Gonyaulacales	*Gambierdiscus belizeanus*
120	EF202866	Super Clade 1	Gonyaulacales	*Gambierdiscus belizeanus*
121	EF202867	Super Clade 1	Gonyaulacales	*Gambierdiscus belizeanus*

122	EF202868	Super Clade 1	Gonyaulacales	*Gambierdiscus belizeanus*
123	EF202869	Super Clade 1	Gonyaulacales	*Gambierdiscus belizeanus*
124	EF202870	Super Clade 1	Gonyaulacales	*Gambierdiscus belizeanus*
125	EF202893	Super Clade 1	Gonyaulacales	*Gambierdiscus australes*
126	EF202895	Super Clade 1	Gonyaulacales	*Gambierdiscus australes*
127	AB764306	Super Clade 1	Gonyaulacales	*Gambierdiscus australes*
128	AB764307	Super Clade 1	Gonyaulacales	*Gambierdiscus australes*
129	AB764301	Super Clade 1	Gonyaulacales	*Gambierdiscus australes*
130	AB605808	Super Clade 1	Gonyaulacales	*Gambierdiscus australes*
131	AB605809	Super Clade 1	Gonyaulacales	*Gambierdiscus australes*
132	AB764303	Super Clade 1	Gonyaulacales	*Gambierdiscus australes*
133	AB764305	Super Clade 1	Gonyaulacales	*Gambierdiscus australes*
134	EF202894	Super Clade 1	Gonyaulacales	*Gambierdiscus australes*
135	AB764308	Super Clade 1	Gonyaulacales	*Gambierdiscus australes*
136	AB764304	Super Clade 1	Gonyaulacales	*Gambierdiscus australes*
137	AB764302	Super Clade 1	Gonyaulacales	*Gambierdiscus australes*
138	AB605805	Super Clade 1	Gonyaulacales	*Gambierdiscus australes*
139	AB605810	Super Clade 1	Gonyaulacales	*Gambierdiscus australes*
140	EF202891	Super Clade 1	Gonyaulacales	*Gambierdiscus australes*
141	EF202892	Super Clade 1	Gonyaulacales	*Gambierdiscus australes*
142	EF202896	Super Clade 1	Gonyaulacales	*Gambierdiscus australes*
143	EF202985	Super Clade 1	Gonyaulacales	*Gambierdiscus caribaeus*
144	EF202917	Super Clade 1	Gonyaulacales	*Gambierdiscus caribaeus*
145	EF202983	Super Clade 1	Gonyaulacales	*Gambierdiscus caribaeus*
146	EF202919	Super Clade 1	Gonyaulacales	*Gambierdiscus caribaeus*
147	EF202918	Super Clade 1	Gonyaulacales	*Gambierdiscus caribaeus*
148	EF202922	Super Clade 1	Gonyaulacales	*Gambierdiscus caribaeus*
149	EF202928	Super Clade 1	Gonyaulacales	*Gambierdiscus caribaeus*
150	EF202924	Super Clade 1	Gonyaulacales	*Gambierdiscus caribaeus*
151	EF202925	Super Clade 1	Gonyaulacales	*Gambierdiscus caribaeus*
152	EF202926	Super Clade 1	Gonyaulacales	*Gambierdiscus caribaeus*
153	EF202927	Super Clade 1	Gonyaulacales	*Gambierdiscus caribaeus*
154	EF202923	Super Clade 1	Gonyaulacales	*Gambierdiscus caribaeus*
155	EF202920	Super Clade 1	Gonyaulacales	*Gambierdiscus caribaeus*
156	EF202921	Super Clade 1	Gonyaulacales	*Gambierdiscus caribaeus*
157	EF202914	Super Clade 1	Gonyaulacales	*Gambierdiscus caribaeus*
158	EF202915	Super Clade 1	Gonyaulacales	*Gambierdiscus caribaeus*
159	EF202916	Super Clade 1	Gonyaulacales	*Gambierdiscus caribaeus*
160	AB605803	Super Clade 1	Gonyaulacales	*Gambierdiscus* sp.
161	AB764295	Super Clade 1	Gonyaulacales	*Gambierdiscus* sp.
162	AB764289	Super Clade 1	Gonyaulacales	*Gambierdiscus* sp.
163	AB764290	Super Clade 1	Gonyaulacales	*Gambierdiscus* sp.
164	AB764277	Super Clade 1	Gonyaulacales	*Gambierdiscus* sp.
165	AB605802	Super Clade 1	Gonyaulacales	*Gambierdiscus* sp.
166	AB764278	Super Clade 1	Gonyaulacales	*Gambierdiscus* sp.
167	AB764291	Super Clade 1	Gonyaulacales	*Gambierdiscus* sp.
168	AB605804	Super Clade 1	Gonyaulacales	*Gambierdiscus* sp.
169	AB764292	Super Clade 1	Gonyaulacales	*Gambierdiscus* sp.
170	AB764293	Super Clade 1	Gonyaulacales	*Gambierdiscus* sp.
171	AB764294	Super Clade 1	Gonyaulacales	*Gambierdiscus* sp.
172	AB499534	Super Clade 1	Gonyaulacales	*Gambierdiscus* sp.
173	AB499537	Super Clade 1	Gonyaulacales	*Gambierdiscus* sp.
174	HE775087	Super Clade 1	Gonyaulacales	*Gambierdiscus* sp.
175	AB764279	Super Clade 1	Gonyaulacales	*Gambierdiscus* sp.
176	AB764280	Super Clade 1	Gonyaulacales	*Gambierdiscus* sp.
177	AB764281	Super Clade 1	Gonyaulacales	*Gambierdiscus* sp.
178	AB764282	Super Clade 1	Gonyaulacales	*Gambierdiscus* sp.
179	AB764283	Super Clade 1	Gonyaulacales	*Gambierdiscus* sp.
180	AB764284	Super Clade 1	Gonyaulacales	*Gambierdiscus* sp.
181	AB764285	Super Clade 1	Gonyaulacales	*Gambierdiscus* sp.
182	AB764286	Super Clade 1	Gonyaulacales	*Gambierdiscus* sp.
183	AB764287	Super Clade 1	Gonyaulacales	*Gambierdiscus* sp.
184	AB764288	Super Clade 1	Gonyaulacales	*Gambierdiscus* sp.

185	EF202910	Super Clade 1	Gonyaulacales	*Gambierdiscus carpenteri*
186	EF202984	Super Clade 1	Gonyaulacales	*Gambierdiscus carpenteri*
187	EF202909	Super Clade 1	Gonyaulacales	*Gambierdiscus carpenteri*
188	EF202908	Super Clade 1	Gonyaulacales	*Gambierdiscus carpenteri*
189	EF202913	Super Clade 1	Gonyaulacales	*Gambierdiscus carpenteri*
190	KM272970	Super Clade 1	Gonyaulacales	*Gambierdiscus carpenteri*
191	AB764297	Super Clade 1	Gonyaulacales	*Gambierdiscus* sp.
192	AB764300	Super Clade 1	Gonyaulacales	*Gambierdiscus* sp.
193	AB764296	Super Clade 1	Gonyaulacales	*Gambierdiscus* sp.
194	AB764298	Super Clade 1	Gonyaulacales	*Gambierdiscus* sp.
195	AB764299	Super Clade 1	Gonyaulacales	*Gambierdiscus* sp.
196	EF202905	Super Clade 1	Gonyaulacales	*Gambierdiscus polynesiensis*
197	EF202902	Super Clade 1	Gonyaulacales	*Gambierdiscus polynesiensis*
198	EF202907	Super Clade 1	Gonyaulacales	*Gambierdiscus polynesiensis*
199	EF202906	Super Clade 1	Gonyaulacales	*Gambierdiscus polynesiensis*
200	EF202903	Super Clade 1	Gonyaulacales	*Gambierdiscus polynesiensis*
201	EF202904	Super Clade 1	Gonyaulacales	*Gambierdiscus polynesiensis*
202	EF202897	Super Clade 1	Gonyaulacales	*Gambierdiscus carolinianus*
203	EF202898	Super Clade 1	Gonyaulacales	*Gambierdiscus carolinianus*
204	EF202899	Super Clade 1	Gonyaulacales	*Gambierdiscus carolinianus*
205	EF202900	Super Clade 1	Gonyaulacales	*Gambierdiscus carolinianus*
206	EF202901	Super Clade 1	Gonyaulacales	*Gambierdiscus carolinianus*
207	KX384639	Super Clade 1	Gonyaulacales	*Gambierdiscus balechii*
208	KX384638	Super Clade 1	Gonyaulacales	*Gambierdiscus balechii*
209	KM886379	Super Clade 1	Gonyaulacales	*Fukuyoa paulensis*
210	KM272972	Super Clade 1	Gonyaulacales	*Fukuyoa paulensis*
211	EF202846	Super Clade 1	Gonyaulacales	*Fukuyoa yasumotoi*
212	EF202847	Super Clade 1	Gonyaulacales	*Fukuyoa yasumotoi*
213	EF202848	Super Clade 1	Gonyaulacales	*Fukuyoa yasumotoi*
214	EF202849	Super Clade 1	Gonyaulacales	*Fukuyoa yasumotoi*
215	EF202850	Super Clade 1	Gonyaulacales	*Fukuyoa yasumotoi*
216	EF202851	Super Clade 1	Gonyaulacales	*Fukuyoa yasumotoi*
217	EF202852	Super Clade 1	Gonyaulacales	*Fukuyoa yasumotoi*
218	AB764309	Super Clade 1	Gonyaulacales	*Fukuyoa yasumotoi*
219	AB764310	Super Clade 1	Gonyaulacales	*Fukuyoa yasumotoi*
220	AB764311	Super Clade 1	Gonyaulacales	*Fukuyoa yasumotoi*
221	EF202853	Super Clade 1	Gonyaulacales	*Fukuyoa ruetzleri*
222	EF202854	Super Clade 1	Gonyaulacales	*Fukuyoa ruetzleri*
223	EF202855	Super Clade 1	Gonyaulacales	*Fukuyoa ruetzleri*
224	EF202856	Super Clade 1	Gonyaulacales	*Fukuyoa ruetzleri*
225	EF202857	Super Clade 1	Gonyaulacales	*Fukuyoa ruetzleri*
226	EF202858	Super Clade 1	Gonyaulacales	*Fukuyoa ruetzleri*
227	EF202859	Super Clade 1	Gonyaulacales	*Fukuyoa ruetzleri*
228	EF202860	Super Clade 1	Gonyaulacales	*Fukuyoa ruetzleri*
229	AB548851	Super Clade 1	Gonyaulacales	*Fukuyo*a sp.
230	KJ447125	Super Clade 1	Gonyaulacales	*Fukuyo*a sp.
231	JN098309	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
232	JN098286	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
233	JN098332	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
234	JN098305	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
235	JF521624	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
236	AB088291	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
237	JF521629	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
238	DQ785888	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
239	DQ785890	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
240	JF521625	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
241	JF521626	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
242	JF521627	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
243	JF521628	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
244	JN098315	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
245	JN098284	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
246	JN098329	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
247	JN098301	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I

248	JN098267	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
249	JN098292	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
250	JN098307	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
251	JN098312	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
252	JN098318	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
253	JN098330	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
254	JN098293	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
255	JN098270	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
256	JN098269	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
257	JN098294	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
258	JN098277	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
259	JN098299	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
260	JN098319	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
261	JN098279	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
262	JN098313	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
263	JN098272	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
264	JN098288	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
265	JN098266	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
266	JN098274	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
267	JN098303	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
268	JN098268	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
269	JN098316	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
270	JN098328	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
271	JN098283	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
272	JN098280	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
273	AB088304	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
274	AB088305	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
275	JN098237	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
276	JN098221	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
277	JN098226	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
278	JN098231	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
279	JN098259	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
280	AB088329	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
281	AY831407	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
282	JN098240	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
283	AB088328	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
284	AY831406	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
285	KF646523	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
286	JN098246	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
287	JN098243	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
288	JN098215	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
289	U09048	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
290	AB088307	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
291	AB088308	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
292	AB088297	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
293	AB088314	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
294	AB088294	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
295	AB088300	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
296	AB088330	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
297	AB088331	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
298	AB088332	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
299	JF521645	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
300	JF521646	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
301	JF521647	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
302	KF646522	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
303	KF646524	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
304	JF521643	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
305	JF521644	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
306	JF521642	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
307	JN098219	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
308	JN098236	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
309	AB088306	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
310	JN098245	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I

311	HQ710797	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
312	JN098224	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
313	KF646521	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
314	JN098263	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
315	JN098239	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
316	JN098218	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
317	JN098229	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
318	JN098308	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
319	JN098334	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
320	JN098324	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
321	JN098282	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
322	JN098323	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
323	DQ785889	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
324	AB088292	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
325	AB088293	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
326	JN098331	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
327	JN098304	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
328	AY421777	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
329	JN098220	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
330	KF908795	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
331	JQ692035	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
332	KF908796	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
333	JN098271	Super Clade 1	Gonyaulacales	*Alexandrium fundyense* Group I
334	JN626281	Super Clade 1	Gonyaulacales	*Alexandrium mediterraneum* Group II
335	JN626282	Super Clade 1	Gonyaulacales	*Alexandrium mediterraneum* Group II
336	AJ535386	Super Clade 1	Gonyaulacales	*Alexandrium mediterraneum* Group II
337	AJ535387	Super Clade 1	Gonyaulacales	*Alexandrium mediterraneum* Group II
338	KF908797	Super Clade 1	Gonyaulacales	*Alexandrium mediterraneum* Group II
339	JN626283	Super Clade 1	Gonyaulacales	*Alexandrium tamarense* Group III
340	KF646525	Super Clade 1	Gonyaulacales	*Alexandrium tamarense* Group III
341	AJ415510	Super Clade 1	Gonyaulacales	*Alexandrium tamarense* Group III
342	DQ444290	Super Clade 1	Gonyaulacales	*Alexandrium tamarense* Group III
343	AY883004	Super Clade 1	Gonyaulacales	*Alexandrium tamarense* Group III
344	KF908799	Super Clade 1	Gonyaulacales	*Alexandrium tamarense* Group III
345	X54946	Super Clade 1	Gonyaulacales	*Alexandrium tamarense* Group III
346	JF906995	Super Clade 1	Gonyaulacales	*Alexandrium tamarense* Group III
347	JF521640	Super Clade 1	Gonyaulacales	*Alexandrium tamarense* Group III
348	DQ785891	Super Clade 1	Gonyaulacales	*Alexandrium tamarense* Group III
349	AJ535391	Super Clade 1	Gonyaulacales	*Alexandrium tamarense* Group III
350	EU024794	Super Clade 1	Gonyaulacales	*Alexandrium pacificum* Group IV
351	KF733551	Super Clade 1	Gonyaulacales	*Alexandrium pacificum* Group IV
352	AB088280	Super Clade 1	Gonyaulacales	*Alexandrium pacificum* Group IV
353	AB088284	Super Clade 1	Gonyaulacales	*Alexandrium pacificum* Group IV
354	DQ785885	Super Clade 1	Gonyaulacales	*Alexandrium pacificum* Group IV
355	DQ785886	Super Clade 1	Gonyaulacales	*Alexandrium pacificum* Group IV
356	JF906991	Super Clade 1	Gonyaulacales	*Alexandrium pacificum* Group IV
357	JF906992	Super Clade 1	Gonyaulacales	*Alexandrium pacificum* Group IV
358	JF906994	Super Clade 1	Gonyaulacales	*Alexandrium pacificum* Group IV
359	AY347308	Super Clade 1	Gonyaulacales	*Alexandrium pacificum* Group IV
360	DQ785887	Super Clade 1	Gonyaulacales	*Alexandrium pacificum* Group IV
361	JF906993	Super Clade 1	Gonyaulacales	*Alexandrium pacificum* Group IV
362	AB183676	Super Clade 1	Gonyaulacales	*Alexandrium pacificum* Group IV
363	JF906989	Super Clade 1	Gonyaulacales	*Alexandrium pacificum* Group IV
364	AB088289	Super Clade 1	Gonyaulacales	*Alexandrium pacificum* Group IV
365	AB088335	Super Clade 1	Gonyaulacales	*Alexandrium pacificum* Group IV
366	KF646528	Super Clade 1	Gonyaulacales	*Alexandrium pacificum* Group IV
367	KF908800	Super Clade 1	Gonyaulacales	*Alexandrium pacificum* Group IV
368	KF646529	Super Clade 1	Gonyaulacales	*Alexandrium pacificum* Group IV
369	JN626278	Super Clade 1	Gonyaulacales	*Alexandrium pacificum* Group IV
370	JF906990	Super Clade 1	Gonyaulacales	*Alexandrium pacificum* Group IV
371	JF521641	Super Clade 1	Gonyaulacales	*Alexandrium pacificum* Group IV
372	KM091275	Super Clade 1	Gonyaulacales	*Alexandrium pacificum* Group IV
373	KM091276	Super Clade 1	Gonyaulacales	*Alexandrium pacificum* Group IV

374	AY421772	Super Clade 1	Gonyaulacales	*Alexandrium pacificum* Group IV
375	AJ535392	Super Clade 1	Gonyaulacales	*Alexandrium pacificum* Group IV
376	KF908802	Super Clade 1	Gonyaulacales	*Alexandrium australiense* Group V
377	AF022191	Super Clade 1	Gonyaulacales	*Alexandrium australiense* Group V
378	JQ991015	Super Clade 1	Gonyaulacales	*Alexandrium australiense* Group V
379	JF521639	Super Clade 1	Gonyaulacales	*Alexandrium australiense* Group V
380	AB088318	Super Clade 1	Gonyaulacales	*Alexandrium tamiyavanichi*
381	AB088323	Super Clade 1	Gonyaulacales	*Alexandrium tamiyavanichi*
382	AB088321	Super Clade 1	Gonyaulacales	*Alexandrium tamiyavanichi*
383	AB088316	Super Clade 1	Gonyaulacales	*Alexandrium tamiyavanichi*
384	AB088322	Super Clade 1	Gonyaulacales	*Alexandrium tamiyavanichi*
385	AB088317	Super Clade 1	Gonyaulacales	*Alexandrium tamiyavanichi*
386	AB088324	Super Clade 1	Gonyaulacales	*Alexandrium tamiyavanichi*
387	AB088325	Super Clade 1	Gonyaulacales	*Alexandrium tamiyavanichi*
388	AF113935	Super Clade 1	Gonyaulacales	*Alexandrium cohorticula*
389	AY421776	Super Clade 1	Gonyaulacales	*Alexandrium fraterculum*
390	JF521622	Super Clade 1	Gonyaulacales	*Alexandrium fraterculum*
391	AB088290	Super Clade 1	Gonyaulacales	*Alexandrium fraterculum*
392	JF521623	Super Clade 1	Gonyaulacales	*Alexandrium fraterculum*
393	AB088282	Super Clade 1	Gonyaulacales	*Alexandrium affine*
394	JF521616	Super Clade 1	Gonyaulacales	*Alexandrium affine*
395	JF906996	Super Clade 1	Gonyaulacales	*Alexandrium affine*
396	JF906997	Super Clade 1	Gonyaulacales	*Alexandrium affine*
397	JF521618	Super Clade 1	Gonyaulacales	*Alexandrium affine*
398	DQ166532	Super Clade 1	Gonyaulacales	*Alexandrium affine*
399	AY421778	Super Clade 1	Gonyaulacales	*Alexandrium affine*
400	AY775286	Super Clade 1	Gonyaulacales	*Alexandrium affine*
401	AY831409	Super Clade 1	Gonyaulacales	*Alexandrium affine*
402	DQ171879	Super Clade 1	Gonyaulacales	*Alexandrium affine*
403	AJ535375	Super Clade 1	Gonyaulacales	*Alexandrium affine*
404	JF906999	Super Clade 1	Gonyaulacales	*Alexandrium minutum*
405	JF906998	Super Clade 1	Gonyaulacales	*Alexandrium minutum*
406	U27499	Super Clade 1	Gonyaulacales	*Alexandrium minutum*
407	AJ535380	Super Clade 1	Gonyaulacales	*Alexandrium minutum*
408	DQ168664	Super Clade 1	Gonyaulacales	*Alexandrium minutum*
409	JF521635	Super Clade 1	Gonyaulacales	*Alexandrium minutum*
410	JF521631	Super Clade 1	Gonyaulacales	*Alexandrium minutum*
411	AY831408	Super Clade 1	Gonyaulacales	*Alexandrium minutum*
412	JF521632	Super Clade 1	Gonyaulacales	*Alexandrium minutum*
413	JF521633	Super Clade 1	Gonyaulacales	*Alexandrium minutum*
414	JF521634	Super Clade 1	Gonyaulacales	*Alexandrium minutum*
415	AY883006	Super Clade 1	Gonyaulacales	*Alexandrium minutum*
416	EU418967	Super Clade 1	Gonyaulacales	*Alexandrium insuetum*
417	JF521630	Super Clade 1	Gonyaulacales	*Alexandrium insuetum*
418	AB088298	Super Clade 1	Gonyaulacales	*Alexandrium insuetum*
419	AJ535383	Super Clade 1	Gonyaulacales	*Alexandrium ostenfeldii*
420	KJ361986	Super Clade 1	Gonyaulacales	*Alexandrium ostenfeldii*
421	KJ361996	Super Clade 1	Gonyaulacales	*Alexandrium ostenfeldii*
422	KJ362003	Super Clade 1	Gonyaulacales	*Alexandrium ostenfeldii*
423	KJ361992	Super Clade 1	Gonyaulacales	*Alexandrium ostenfeldii*
424	KJ362001	Super Clade 1	Gonyaulacales	*Alexandrium ostenfeldii*
425	KJ361998	Super Clade 1	Gonyaulacales	*Alexandrium ostenfeldii*
426	JF521636	Super Clade 1	Gonyaulacales	*Alexandrium ostenfeldii*
427	AJ535382	Super Clade 1	Gonyaulacales	*Alexandrium ostenfeldii*
428	KJ361993	Super Clade 1	Gonyaulacales	*Alexandrium ostenfeldii*
429	AJ535381	Super Clade 1	Gonyaulacales	*Alexandrium ostenfeldii*
430	JF521637	Super Clade 1	Gonyaulacales	*Alexandrium ostenfeldii*
431	AJ535384	Super Clade 1	Gonyaulacales	*Alexandrium ostenfeldii*
432	KJ361972	Super Clade 1	Gonyaulacales	*Alexandrium ostenfeldii*
433	KJ361990	Super Clade 1	Gonyaulacales	*Alexandrium ostenfeldii*
434	KJ361999	Super Clade 1	Gonyaulacales	*Alexandrium ostenfeldii*
435	U27500	Super Clade 1	Gonyaulacales	*Alexandrium ostenfeldii*
436	AB538439	Super Clade 1	Gonyaulacales	*Alexandrium ostenfeldii*

437	KJ361988	Super Clade 1	Gonyaulacales	*Alexandrium ostenfeldii*
438	KJ361997	Super Clade 1	Gonyaulacales	*Alexandrium ostenfeldii*
439	LC056069	Super Clade 1	Gonyaulacales	*Alexandrium ostenfeldii*
440	KF925334	Super Clade 1	Gonyaulacales	*Alexandrium andersonii*
441	JF521619	Super Clade 1	Gonyaulacales	*Alexandrium andersonii*
442	JF521620	Super Clade 1	Gonyaulacales	*Alexandrium andersonii*
443	JF521621	Super Clade 1	Gonyaulacales	*Alexandrium andersonii*
444	AJ535376	Super Clade 1	Gonyaulacales	*Alexandrium tamutum*
445	AJ535377	Super Clade 1	Gonyaulacales	*Alexandrium tamutum*
446	AJ535378	Super Clade 1	Gonyaulacales	*Alexandrium tamutum*
447	AJ535379	Super Clade 1	Gonyaulacales	*Alexandrium tamutum*
448	U27498	Super Clade 1	Gonyaulacales	*Alexandrium margalefii*
449	AY641566	Super Clade 1	Gonyaulacales	*Alexandrium satoanum*
450	KM067435	Super Clade 1	Gonyaulacales	*Alexandrium satoanum*
451	AJ535385	Super Clade 1	Gonyaulacales	*Alexandrium taylori*
452	AJ535389	Super Clade 1	Gonyaulacales	*Alexandrium taylori*
453	AJ535390	Super Clade 1	Gonyaulacales	*Alexandrium taylori*
454	LC056070	Super Clade 1	Gonyaulacales	*Alexandrium hiranoi*
455	LC056068	Super Clade 1	Gonyaulacales	*Alexandrium hiranoi*
456	JF521638	Super Clade 1	Gonyaulacales	*Alexandrium pseudogoniaulax*
457	AY641564	Super Clade 1	Gonyaulacales	*Alexandrium hiranoi*
458	AB088302	Super Clade 1	Gonyaulacales	*Alexandrium pseudogoniaulax*
459	AY883005	Super Clade 1	Gonyaulacales	*Alexandrium monilatum*
460	AY641565	Super Clade 1	Gonyaulacales	*Alexandrium leei*
461	KF251139	Super Clade 1	Gonyaulacales	*Alexandrium diversaporum*
462	LN811348	Super Clade 1	Gonyaulacales	*Alexandrium pohangense*
463	FR846195	Super Clade 1	Gonyaulacales	*Coolia canariensis*
464	HQ897282	Super Clade 1	Gonyaulacales	*Coolia canariensis*
465	FR847220	Super Clade 1	Gonyaulacales	*Coolia monotis*
466	EF492487	Super Clade 1	Gonyaulacales	*Coolia monotis*
467	FR847217	Super Clade 1	Gonyaulacales	*Coolia monotis*
468	EF492488	Super Clade 1	Gonyaulacales	*Coolia monotis*
469	HQ897279	Super Clade 1	Gonyaulacales	*Coolia malayensis*
470	AJ415509	Super Clade 1	Gonyaulacales	*Coolia* sp.
471	HQ897281	Super Clade 1	Gonyaulacales	*Coolia* sp.
472	KF733525	Super Clade 1	Gonyaulacales	*Coolia* sp.
473	HQ897280	Super Clade 1	Gonyaulacales	*Coolia* sp.
474	HE793379	Super Clade 1	Gonyaulacales	*Ostreopsis ovata*
475	KF359996	Super Clade 1	Gonyaulacales	*Ostreopsis* sp.
476	KF359997	Super Clade 1	Gonyaulacales	*Ostreopsis* sp.
477	KF359998	Super Clade 1	Gonyaulacales	*Ostreopsis* sp.
478	KF359999	Super Clade 1	Gonyaulacales	*Ostreopsis* sp.
479	KF360000	Super Clade 1	Gonyaulacales	*Ostreopsis* sp.
480	KF360001	Super Clade 1	Gonyaulacales	*Ostreopsis* sp.
481	KF360002	Super Clade 1	Gonyaulacales	*Ostreopsis* sp.
482	KF360003	Super Clade 1	Gonyaulacales	*Ostreopsis* sp.
483	KF360004	Super Clade 1	Gonyaulacales	*Ostreopsis* sp.
484	KF733537	Super Clade 1	Gonyaulacales	*Ostreopsis* sp.
485	AF244939	Super Clade 1	Gonyaulacales	*Ostreopsis ovata*
486	KT868529	Super Clade 1	Gonyaulacales	*Ostreopsis siamensis*
487	AB936753	Super Clade 1	Gonyaulacales	*Pyrodinium bahamense* var. *bahamense*
488	AB936750	Super Clade 1	Gonyaulacales	*Pyrodinium bahamense* var. *compressum*
489	AB936751	Super Clade 1	Gonyaulacales	*Pyrodinium bahamense* var. *compressum*
490	AF274275	Super Clade 1	Gonyaulacales	*Pyrodinium bahamense* var. *compressum*
491	DQ500120	Super Clade 1	Gonyaulacales	*Pyrodinium bahamense* var. *compressum*
492	DQ500123	Super Clade 1	Gonyaulacales	*Pyrodinium bahamense* var. *compressum*
493	DQ500119	Super Clade 1	Gonyaulacales	*Pyrodinium bahamense* var. *compressum*
494	DQ500121	Super Clade 1	Gonyaulacales	*Pyrodinium bahamense* var. *compressum*
495	DQ500122	Super Clade 1	Gonyaulacales	*Pyrodinium bahamense* var. *compressum*
496	KM886380	Super Clade 1	Gonyaulacales	*Goniodoma polyedricum*
497	AF022155	Super Clade 1	Gonyaulacales	*Gonyaulax spinifera*
498	KF925336	Super Clade 1	Gonyaulacales	*Gonyaulax spinifera*
499	AY775287	Super Clade 1	Gonyaulacales	*Gonyaulax polygramma*

500	AJ833631	Super Clade 1	Gonyaulacales	*Gonyaulax polygramma*
501	AY672702	Super Clade 1	Gonyaulacales	*Gonyaulax fragilis*
502	AY443013	Super Clade 2	Gonyaulacales	*Gonyaulax verior*
503	FR865625	Super Clade 3	Gonyaulacales	*Gonyaulax spinifera*
504	AF052190	Super Clade 1	Gonyaulacales	*Gonyaulax spinifera*
505	EU805590	Super Clade 1	Gonyaulacales	*Gonyaulax spinifera*
506	LC036590	Super Clade 1	Gonyaulacales	*Gonyaulax ellegaardiae*
507	DQ867107	Super Clade 1	Gonyaulacales	*Gonyaulax spinifera*
508	DQ388465	Super Clade 1	Gonyaulacales	*Gonyaulax cochlea*
509	AF274258	Super Clade 1	Gonyaulacales	*Gonyaulax cochlea*
510	EF492489	Super Clade 1	Gonyaulacales	*Fragilidium* sp.
511	FJ405355	Super Clade 1	Gonyaulacales	*Fragilidium mexicanum*
512	AF033869	Super Clade 1	Gonyaulacales	*Fragilidium subglobosum*
513	FJ405356	Super Clade 1	Gonyaulacales	*Fragilidium* sp.
514	FR865628	Super Clade 1	Gonyaulacales	*Pyrocystis lunula*
515	FR865629	Super Clade 1	Gonyaulacales	*Pyrocystis lunula*
516	AF274274	Super Clade 1	Gonyaulacales	*Pyrocystis lunula*
517	LC054939	Super Clade 1	Gonyaulacales	*Pyrocystis* sp.
518	AF022156	Super Clade 1	Gonyaulacales	*Pyrocystis noctiluca*
519	AY443024	Super Clade 1	Gonyaulacales	*Pyrophacus steinii*
520	AB375868	Super Clade 1	Gonyaulacales	*Amylax buxus*
521	JX666361	Super Clade 1	Gonyaulacales	*Amylax triacantha*
522	AF274269	Super Clade 1	Gonyaulacales	*Lingulodinium polyedra*
523	DQ202217	Super Clade 1	Gonyaulacales	*Lingulodinium polyedra*
524	DQ202218	Super Clade 1	Gonyaulacales	*Lingulodinium polyedra*
525	DQ202219	Super Clade 1	Gonyaulacales	*Lingulodinium polyedra*
526	DQ202220	Super Clade 1	Gonyaulacales	*Lingulodinium polyedra*
527	DQ202221	Super Clade 1	Gonyaulacales	*Lingulodinium polyedra*
528	AJ415511	Super Clade 1	Gonyaulacales	*Lingulodinium polyedra*
529	AB693194	Super Clade 1	Gonyaulacales	*Lingulodinium polyedra*
530	EF492507	Super Clade 1	Gonyaulacales	*Lingulodinium polyedra*
531	AF377944	Super Clade 1	Gonyaulacales	*Lingulodinium polyedra*
532	AB693196	Super Clade 1	Gonyaulacales	*Lingulodinium polyedra*
533	JQ616824	Super Clade 1	Gonyaulacales	*Lingulodinium polyedra*
534	AY421788	Super Clade 1	Gonyaulacales	*Lingulodinium polyedra*
535	DQ388456	Super Clade 1	Gonyaulacales	*Ceratocorys horrida*
536	LC054924	Super Clade 1	Gonyaulacales	*Ceratocorys horrida*
537	AF022154	Super Clade 1	Gonyaulacales	*Ceratocorys horrida*
538	AB727654	Super Clade 1	Gonyaulacales	*Protoceratium reticulatum*
539	AB727655	Super Clade 1	Gonyaulacales	*Protoceratium reticulatum*
540	AB727656	Super Clade 1	Gonyaulacales	*Protoceratium reticulatum*
541	AY421790	Super Clade 1	Gonyaulacales	*Protoceratium reticulatum*
542	DQ217789	Super Clade 1	Gonyaulacales	*Protoceratium reticulatum*
543	AF274273	Super Clade 1	Gonyaulacales	*Protoceratium reticulatum*
544	HM853766	Super Clade 2	Dinophysiales	*Amphisolenia schauinslandii*
545	HM853765	Super Clade 2	Dinophysiales	*Amphisolenia globifera*
546	HM853763	Super Clade 2	Dinophysiales	*Amphisolenia bidentata*
547	GU196149	Super Clade 2	Dinophysiales	*Amphisolenia bidentata*
548	HM853764	Super Clade 2	Dinophysiales	*Amphisolenia* sp.
549	HM853767	Super Clade 2	Dinophysiales	*Triposolenia bicornis*
550	HM853768	Super Clade 2	Dinophysiales	*Triposolenia bicornis*
551	HM853769	Super Clade 2	Dinophysiales	*Triposolenia bicornis*
552	HM853770	Super Clade 2	Dinophysiales	*Triposolenia bicornis*
553	FJ869120	Super Clade 2	Dinophysiales	*Dinophysis acuminata*
554	AJ506972	Super Clade 2	Dinophysiales	*Dinophysis acuminata*
555	EU130569	Super Clade 2	Dinophysiales	*Dinophysis acuminata*
556	AB073117	Super Clade 2	Dinophysiales	*Dinophysis acuminata*
557	KJ508017	Super Clade 2	Dinophysiales	*Dinophysis acuminata*
558	AY260470	Super Clade 2	Dinophysiales	*Dinophysis norvegica*
559	AJ506974	Super Clade 2	Dinophysiales	*Dinophysis norvegica*
560	AB073119	Super Clade 2	Dinophysiales	*Dinophysis norvegica*
561	AF239261	Super Clade 2	Dinophysiales	*Dinophysis norvegica*
562	HM853816	Super Clade 2	Dinophysiales	*Dinophysis tripos*

563	EU780644	Super Clade 2	Dinophysiales	*Dinophysis caudata*
564	AB366002	Super Clade 2	Dinophysiales	*Dinophysis infundibulum*
565	AB073118	Super Clade 2	Dinophysiales	*Dinophysis fortii*
566	AJ506973	Super Clade 2	Dinophysiales	*Dinophysis acuta*
567	HM853815	Super Clade 2	Dinophysiales	*Dinophysis caudata*
568	HM853805	Super Clade 2	Dinophysiales	*Dinophysis hastata*
569	HM853808	Super Clade 2	Dinophysiales	*Dinophysis pusilla*
570	HM853809	Super Clade 2	Dinophysiales	*Dinophysis pusilla*
571	HM853807	Super Clade 2	Dinophysiales	*Dinophysis acutissima*
572	HM853806	Super Clade 2	Dinophysiales	*Dinophysis phalacromoides*
573	HM853810	Super Clade 2	Dinophysiales	*Dinophysis monacantha*
574	HM853811	Super Clade 2	Dinophysiales	*Dinophysis odiosa*
575	HM853812	Super Clade 2	Dinophysiales	*Dinophysis odiosa*
576	HM853813	Super Clade 2	Dinophysiales	*Dinophysis hastata*
577	HM853814	Super Clade 2	Dinophysiales	*Dinophysis hastata*
578	HM853803	Super Clade 2	Dinophysiales	*Histioneis longicollis*
579	HM853804	Super Clade 2	Dinophysiales	*Histioneis longicollis*
580	EU780646	Super Clade 2	Dinophysiales	*Histioneis* sp.
581	HM853802	Super Clade 2	Dinophysiales	*Histioneis gubernans*
582	HM853801	Super Clade 2	Dinophysiales	*Histioneis cymbalaria*
583	EU780647	Super Clade 2	Dinophysiales	*Ornithocercus quadratus*
584	HM853800	Super Clade 2	Dinophysiales	*Ornithocercus quadratus*
585	HM853799	Super Clade 2	Dinophysiales	*Ornithocercus quadratus*
586	HM853795	Super Clade 2	Dinophysiales	*Ornithocercus heteroporus*
587	HM853793	Super Clade 2	Dinophysiales	*Ornithocercus heteroporus*
588	HM853794	Super Clade 2	Dinophysiales	*Ornithocercus heteroporus*
589	HM853796	Super Clade 2	Dinophysiales	*Ornithocercus heteroporus*
590	HM853797	Super Clade 2	Dinophysiales	*Ornithocercus magnificus*
591	EU780651	Super Clade 2	Dinophysiales	*Ornithocercus magnificus*
592	JN587287	Super Clade 2	Dinophysiales	*Pseudophalacroma nasutum*
593	JN587286	Super Clade 2	Dinophysiales	*Pseudophalacroma nasutum*
594	JN587285	Super Clade 2	Dinophysiales	*Pseudophalacroma nasutum*
595	JN587289	Super Clade 2	Dinophysiales	*Pseudophalacroma nasutum*
596	JN587288	Super Clade 2	Dinophysiales	*Pseudophalacroma nasutum*
597	JN587290	Super Clade 2	Dinophysiales	*Pseudophalacroma* sp.
598	JQ996372	Super Clade 2	Dinophysiales	*Sinophysis ebriola*
599	JQ996379	Super Clade 2	Dinophysiales	*Sinophysis ebriola*
600	JN587291	Super Clade 2	Dinophysiales	*Sinophysis grandis*
601	JN587292	Super Clade 2	Dinophysiales	*Sinophysis ebriola*
602	JQ996377	Super Clade 2	Dinophysiales	*Sinophysis verruculosa*
603	JQ996381	Super Clade 2	Dinophysiales	*Sinophysis stenosoma*
604	JQ996373	Super Clade 2	Dinophysiales	*Sinophysis stenosoma*
605	JQ996375	Super Clade 2	Dinophysiales	*Sinophysis grandis*
606	JQ996374	Super Clade 2	Dinophysiales	*Sinophysis grandis*
607	JQ996382	Super Clade 2	Dinophysiales	*Sinophysis grandis*
608	JQ996380	Super Clade 2	Dinophysiales	*Sinophysis grandis*
609	JQ996376	Super Clade 2	Dinophysiales	*Sinophysis grandis*
610	JQ996378	Super Clade 2	Dinophysiales	*Sinophysis microcephala*
611	HM853780	Super Clade 2	Dinophysiales	*Phalacroma doryphorum*
612	HM853781	Super Clade 2	Dinophysiales	*Phalacroma doryphorum*
613	HM853779	Super Clade 2	Dinophysiales	*Phalacroma doryphorum*
614	HM853775	Super Clade 2	Dinophysiales	*Phalacroma mitra*
615	HM853776	Super Clade 2	Dinophysiales	*Phalacroma mitra*
616	HM853777	Super Clade 2	Dinophysiales	*Phalacroma mitra*
617	HM853778	Super Clade 2	Dinophysiales	*Phalacroma mitra*
618	HM853783	Super Clade 2	Dinophysiales	*Phalacroma rotundatum*
619	HM853784	Super Clade 2	Dinophysiales	*Phalacroma rotundatum*
620	HM853788	Super Clade 2	Dinophysiales	*Phalacroma porodictyum*
621	HM853789	Super Clade 2	Dinophysiales	*Phalacroma porodictyum*
622	HM853790	Super Clade 2	Dinophysiales	*Phalacroma porodictyum*
623	EU780657	Super Clade 2	Dinophysiales	*Phalacroma rotundatum*
624	AJ506975	Super Clade 2	Dinophysiales	*Phalacroma rotundatum*
625	HM853787	Super Clade 2	Dinophysiales	*Phalacroma favus*

626	HM853786	Super Clade 2	Dinophysiales	*Phalacroma favus*
627	HM853791	Super Clade 2	Dinophysiales	*Phalacroma porodictyum*
628	HM853792	Super Clade 2	Dinophysiales	*Phalacroma porodictyum*
629	HM853785	Super Clade 2	Dinophysiales	*Phalacroma sp.*
630	HM853771	Super Clade 2	Dinophysiales	*Phalacroma parvulum*
631	HM853772	Super Clade 2	Dinophysiales	*Phalacroma parvulum*
632	HM853773	Super Clade 2	Dinophysiales	*Phalacroma parvulum*
633	EU780655	Super Clade 2	Dinophysiales	*Phalacroma rapa*
634	FJ477082	Super Clade 2	Dinophysiales	*Phalacroma rapa*
635	HM853774	Super Clade 2	Dinophysiales	*Phalacroma rapa*
636	AB551248	Super Clade 2	Dinophysiales	*Phalacroma mitra*
637	FJ477084	Super Clade 2	Dinophysiales	*Phalacroma sp.*
638	HM853782	Super Clade 2	Dinophysiales	*Phalacroma oxytoxoides*
639	JQ996385	Super Clade 2	Dinophysiales	*Phalacroma oxytoxoides*
640	JQ996384	Super Clade 2	Dinophysiales	*Phalacroma oxytoxoides*
641	EF052682	Super Clade 3	Suessiales	*Baldinia anauniensis*
642	AY829528	Super Clade 3	Suessiales	*Baldinia* sp.
643	KF446617	Super Clade 3	Suessiales	*Borghiella tenuissima*
644	AY443025	Super Clade 3	Suessiales	*Borghiella tenuissima*
645	KF446616	Super Clade 3	Suessiales	*Borghiella tenuissima*
646	KF446620	Super Clade 3	Suessiales	*Borghiella tenuissima*
647	GU067825	Super Clade 3	Suessiales	*Borghiella tenuissima*
648	EF058235	Super Clade 3	Suessiales	*Cystodinium phaseolus*
649	EF058251	Super Clade 3	Suessiales	*Phytodinium* sp.
650	HG529978	Super Clade 3	Suessiales	*Ansanella granifera*
651	HG792066	Super Clade 3	Suessiales	*Ansanella granifera*
652	LC068836	Super Clade 3	Suessiales	*Asulcocephalium miricentonis*
653	LC068838	Super Clade 3	Suessiales	*Asulcocephalium miricentonis*
654	LC068840	Super Clade 3	Suessiales	*Leiocephalium pseudosanguineum*
655	LC068841	Super Clade 3	Suessiales	*Leiocephalium pseudosanguineum*
656	FR690459	Super Clade 3	Suessiales	*Biecheleria cincta*
657	JF794059	Super Clade 3	Suessiales	*Biecheleria cincta*
658	JN934667	Super Clade 3	Suessiales	*Biecheleria cincta*
659	LC068842	Super Clade 3	Suessiales	*Biecheleria brevisulcata*
660	EF058252	Super Clade 3	Suessiales	*Biecheleria baltica*
661	LC054923	Super Clade 3	Suessiales	*Biecheleria natalensis*
662	KF463288	Super Clade 3	Suessiales	*Biecheleria tirezensis*
663	HG792067	Super Clade 3	Suessiales	*Biecheleriopsis adriatica*
664	U37406	Super Clade 3	Suessiales	*Pelgodinium beii*
665	JX661028	Super Clade 3	Suessiales	*Pelgodinium* sp.
666	KF422623	Super Clade 3	Suessiales	*Pelgodinium beii*
667	U41087	Super Clade 3	Suessiales	*Pelgodinium beii*
668	U37365	Super Clade 3	Suessiales	*Pelgodinium beii*
669	JF791066	Super Clade 3	Suessiales	*Pelgodinium beii*
670	EF492490	Super Clade 3	Suessiales	*Pelagodinium* sp.
671	JX661025	Super Clade 3	Suessiales	*Pelagodinium* sp.
672	JX661026	Super Clade 3	Suessiales	*Pelagodinium* sp.
673	JX661027	Super Clade 3	Suessiales	*Pelagodinium* sp.
674	JX661029	Super Clade 3	Suessiales	*Pelagodinium* sp.
675	EF016917	Super Clade 3	Suessiales	*Piscinoodinium* sp.
676	EF016918	Super Clade 3	Suessiales	*Piscinoodinium* sp.
677	EF016919	Super Clade 3	Suessiales	*Piscinoodinium* sp.
678	EF016920	Super Clade 3	Suessiales	*Piscinoodinium* sp.
679	EF016921	Super Clade 3	Suessiales	*Piscinoodinium* sp.
680	EF016922	Super Clade 3	Suessiales	*Piscinoodinium* sp.
681	EF016923	Super Clade 3	Suessiales	*Piscinoodinium* sp.
682	GQ375263	Super Clade 3	Suessiales	*Polarella glacialis*
683	KF925337	Super Clade 3	Suessiales	*Polarella glacialis*
684	EF434275	Super Clade 3	Suessiales	*Polarella glacialis*
685	EF434276	Super Clade 3	Suessiales	*Polarella glacialis*
686	EF434277	Super Clade 3	Suessiales	*Polarella glacialis*
687	EF417317	Super Clade 3	Suessiales	*Polarella glacialis*
688	AF099183	Super Clade 3	Suessiales	*Polarella glacialis*

689	JF791031	Super Clade 3	Suessiales	*Protodinium simplex*
690	EF492493	Super Clade 3	Suessiales	*Protodinium simplex*
691	U41086	Super Clade 3	Suessiales	*Protodinium simplex*
692	DQ388466	Super Clade 3	Suessiales	*Protodinium simplex*
693	EF492491	Super Clade 3	Suessiales	*Protodinium simplex*
694	AB016539	Super Clade 3	Suessiales	*Symbiodinium* sp.
695	AB055911	Super Clade 3	Suessiales	*Symbiodinium* sp.
696	AB016578	Super Clade 3	Suessiales	*Symbiodinium* sp.
697	L13718	Super Clade 3	Suessiales	*Symbiodinium meandrinae*
698	L13717	Super Clade 3	Suessiales	*Symbiodinium corculorum*
699	M88521	Super Clade 3	Suessiales	*Symbiodinium microadriaticum*
700	AB055918	Super Clade 3	Suessiales	*Symbiodinium* sp.
701	AB055916	Super Clade 3	Suessiales	*Symbiodinium* sp.
702	AB055915	Super Clade 3	Suessiales	*Symbiodinium* sp.
703	AB055914	Super Clade 3	Suessiales	*Symbiodinium* sp.
704	AB055913	Super Clade 3	Suessiales	*Symbiodinium* sp.
705	AB055912	Super Clade 3	Suessiales	*Symbiodinium* sp.
706	AB055917	Super Clade 3	Suessiales	*Symbiodinium* sp.
707	AB016581	Super Clade 3	Suessiales	*Symbiodinium* sp.
708	AB016580	Super Clade 3	Suessiales	*Symbiodinium* sp.
709	AB016579	Super Clade 3	Suessiales	*Symbiodinium* sp.
710	AB016577	Super Clade 3	Suessiales	*Symbiodinium* sp.
711	AB016576	Super Clade 3	Suessiales	*Symbiodinium* sp.
712	AB016575	Super Clade 3	Suessiales	*Symbiodinium* sp.
713	AB016574	Super Clade 3	Suessiales	*Symbiodinium* sp.
714	AB016573	Super Clade 3	Suessiales	*Symbiodinium* sp.
715	AB016572	Super Clade 3	Suessiales	*Symbiodinium* sp.
716	AB085914	Super Clade 3	Suessiales	*Symbiodinium* sp.
717	AB085913	Super Clade 3	Suessiales	*Symbiodinium* sp.
718	AB016538	Super Clade 3	Suessiales	*Symbiodinium* sp.
719	AB085912	Super Clade 3	Suessiales	*Symbiodinium* sp.
720	AB085911	Super Clade 3	Suessiales	*Symbiodinium* sp.
721	AB016594	Super Clade 3	Suessiales	*Symbiodinium* sp.
722	AB016593	Super Clade 3	Suessiales	*Symbiodinium* sp.
723	AB016597	Super Clade 3	Suessiales	*Symbiodinium* sp.
724	AB016596	Super Clade 3	Suessiales	*Symbiodinium* sp.
725	AB016595	Super Clade 3	Suessiales	*Symbiodinium* sp.
726	DQ838542	Super Clade 3	Suessiales	*Symbiodinium* sp. Clade D
727	AB030646	Super Clade 3	Suessiales	*Symbiodinium* sp.
728	AB085915	Super Clade 3	Suessiales	*Symbiodinium* sp.
729	AB016722	Super Clade 3	Suessiales	*Symbiodinium* sp.
730	AB016724	Super Clade 3	Suessiales	*Symbiodinium* sp.
731	AB016723	Super Clade 3	Suessiales	*Symbiodinium* sp.
732	AF271292	Super Clade 3	Suessiales	*Symbiodinium* sp.
733	AF238263	Super Clade 3	Suessiales	*Symbiodinium* sp. Clade E
734	AF238262	Super Clade 3	Suessiales	*Symbiodinium* sp. Clade E
735	AF238261	Super Clade 3	Suessiales	*Symbiodinium* sp. Clade E
736	AB126931	Super Clade 3	Suessiales	*Symbiodinium* sp.
737	AB126930	Super Clade 3	Suessiales	*Symbiodinium* sp.
738	AB126929	Super Clade 3	Suessiales	*Symbiodinium* sp.
739	AB126928	Super Clade 3	Suessiales	*Symbiodinium* sp.
740	AB126927	Super Clade 3	Suessiales	*Symbiodinium* sp.
741	AB126926	Super Clade 3	Suessiales	*Symbiodinium* sp.
742	AF271291	Super Clade 3	Suessiales	*Symbiodinium* sp.
743	AF260260	Super Clade 3	Suessiales	*Symbiodinium* sp.
744	AY630406	Super Clade 3	Suessiales	*Symbiodinium* sp.
745	KU900226	Super Clade 3	Suessiales	*Symbiodinium microadriaticum*
746	KU188515	Super Clade 3	Suessiales	*Symbiodinium* sp. Clade A
747	KU188514	Super Clade 3	Suessiales	*Symbiodinium* sp. Clade A
748	KU188513	Super Clade 3	Suessiales	*Symbiodinium* sp. Clade A
749	KU188512	Super Clade 3	Suessiales	*Symbiodinium* sp. Clade A
750	KU188511	Super Clade 3	Suessiales	*Symbiodinium* sp. Clade A
751	KU188510	Super Clade 3	Suessiales	*Symbiodinium* sp. Clade A

752	KU188509	Super Clade 3	Suessiales	*Symbiodinium* sp. Clade A
753	KU188508	Super Clade 3	Suessiales	*Symbiodinium* sp. Clade A
754	KJ650343	Super Clade 3	Suessiales	*Symbiodinium* sp.
755	KC848881	Super Clade 3	Suessiales	*Symbiodinium* sp. Clade D
756	KC848879	Super Clade 3	Suessiales	*Symbiodinium* sp. Clade A
757	KC848882	Super Clade 3	Suessiales	*Symbiodinium* sp. Clade C
758	KC848880	Super Clade 3	Suessiales	*Symbiodinium* sp. Clade B
759	GU362427	Super Clade 3	Suessiales	*Symbiodinium* sp.
760	GU362425	Super Clade 3	Suessiales	*Symbiodinium* sp.
761	GU362424	Super Clade 3	Suessiales	*Symbiodinium* sp.
762	DQ838543	Super Clade 3	Suessiales	*Symbiodinium* sp. Clade C
763	AF182822	Super Clade 3	Suessiales	*Symbiodinium* sp.
764	EF419291	Super Clade 3	Suessiales	*Symbiodinium* sp. Clade C
765	EF419290	Super Clade 3	Suessiales	*Symbiodinium* sp. Clade C
766	EF419289	Super Clade 3	Suessiales	*Symbiodinium* sp. Clade C
767	EF419288	Super Clade 3	Suessiales	*Symbiodinium* sp. Clade C
768	EF419287	Super Clade 3	Suessiales	*Symbiodinium* sp. Clade C
769	EF419286	Super Clade 3	Suessiales	*Symbiodinium* sp. Clade C
770	EF419285	Super Clade 3	Suessiales	*Symbiodinium* sp. Clade C
771	EF419281	Super Clade 3	Suessiales	*Symbiodinium* sp. Clade D
772	EF419282	Super Clade 3	Suessiales	*Symbiodinium* sp. Clade D
773	EF419283	Super Clade 3	Suessiales	*Symbiodinium* sp. Clade C
774	EF419284	Super Clade 3	Suessiales	*Symbiodinium* sp. Clade C
775	AY937258	Super Clade 3	Suessiales	*Symbiodinium* sp.
776	AF290918	Super Clade 3	Suessiales	*Symbiodinium* sp.
777	AF290917	Super Clade 3	Suessiales	*Symbiodinium* sp.
778	AF255737	Super Clade 3	Suessiales	*Symbiodinium* sp.
779	AY160124	Super Clade 3	Suessiales	*Symbiodinium* sp.
780	AF238260	Super Clade 3	Suessiales	*Symbiodinium* sp. Clade C2
781	AF238259	Super Clade 3	Suessiales	*Symbiodinium* sp. Clade C2
782	AF238258	Super Clade 3	Suessiales	*Symbiodinium* sp. Clade C
783	AF238257	Super Clade 3	Suessiales	*Symbiodinium* sp. Clade B
784	AF238256	Super Clade 3	Suessiales	*Symbiodinium* sp. Clade A
785	M88509	Super Clade 3	Suessiales	*Symbiodinium* sp.
786	KU197083	Super Clade 3	Suessiales	*Symbiodinium* sp. Clade C
787	KT860942	Super Clade 3	Suessiales	*Symbiodinium* sp.
788	KC816670	Super Clade 3	Suessiales	*Symbiodinium* sp. Clade C
789	KC816662	Super Clade 3	Suessiales	*Symbiodinium* sp. Clade C
790	KC816660	Super Clade 3	Suessiales	*Symbiodinium* sp. Clade C
791	KC816659	Super Clade 3	Suessiales	*Symbiodinium* sp. Clade C
792	KC816647	Super Clade 3	Suessiales	*Symbiodinium* sp. Clade C
793	KC816646	Super Clade 3	Suessiales	*Symbiodinium* sp. Clade C
794	KC816645	Super Clade 3	Suessiales	*Symbiodinium* sp. Clade C
795	KC816644	Super Clade 3	Suessiales	*Symbiodinium* sp. Clade C
796	KC816643	Super Clade 3	Suessiales	*Symbiodinium* sp. Clade C
797	KC816641	Super Clade 3	Suessiales	*Symbiodinium* sp. Clade C
798	KC816640	Super Clade 3	Suessiales	*Symbiodinium* sp. Clade C
799	KC816639	Super Clade 3	Suessiales	*Symbiodinium* sp. Clade C
800	KC816638	Super Clade 3	Suessiales	*Symbiodinium* sp. Clade C
801	KC816635	Super Clade 3	Suessiales	*Symbiodinium* sp. Clade C
802	KC816632	Super Clade 3	Suessiales	*Symbiodinium* sp. Clade C
803	KC816631	Super Clade 3	Suessiales	*Symbiodinium* sp. Clade C
804	HM067613	Super Clade 3	Suessiales	*Symbiodinium* sp.
805	HM067612	Super Clade 3	Suessiales	*Symbiodinium* sp.
806	HM067611	Super Clade 3	Suessiales	*Symbiodinium* sp.
807	HM067608	Super Clade 3	Suessiales	*Symbiodinium* sp.
808	JQ320136	Super Clade 3	Suessiales	*Symbiodinium* sp.
809	AY525027	Super Clade 3	Suessiales	*Symbiodinium* sp.
810	AY525020	Super Clade 3	Suessiales	*Symbiodinium* sp.
811	EF492514	Super Clade 3	Suessiales	*Symbiodinium microadriaticum*
812	EF492496	Super Clade 3	Suessiales	*Symbiodinium microadriaticum*
813	AY165766	Super Clade 3	Suessiales	*Symbiodinium* sp.
814	AB183640	Super Clade 3	Suessiales	*Symbiodinium* sp.

815	AF379641	Super Clade 3	Suessiales	*Symbiodinium* sp.
816	AJ271777	Super Clade 3	Suessiales	*Symbiodinium* sp.
817	AJ271776	Super Clade 3	Suessiales	*Symbiodinium* sp.
818	AJ271775	Super Clade 3	Suessiales	*Symbiodinium* sp.
819	AJ271774	Super Clade 3	Suessiales	*Symbiodinium* sp.
820	AJ271773	Super Clade 3	Suessiales	*Symbiodinium* sp.
821	AJ271772	Super Clade 3	Suessiales	*Symbiodinium* sp.
822	AJ271771	Super Clade 3	Suessiales	*Symbiodinium* sp.
823	AJ271770	Super Clade 3	Suessiales	*Symbiodinium* sp.
824	AJ271768	Super Clade 3	Suessiales	*Symbiodinium* sp.
825	AJ271767	Super Clade 3	Suessiales	*Symbiodinium* sp.
826	AJ271766	Super Clade 3	Suessiales	*Symbiodinium* sp.
827	AJ271765	Super Clade 3	Suessiales	*Symbiodinium* sp.
828	AJ271764	Super Clade 3	Suessiales	*Symbiodinium* sp.
829	AJ271763	Super Clade 3	Suessiales	*Symbiodinium* sp.
830	AJ271762	Super Clade 3	Suessiales	*Symbiodinium* sp.
831	AJ271761	Super Clade 3	Suessiales	*Symbiodinium* sp.
832	AJ271760	Super Clade 3	Suessiales	*Symbiodinium* sp.
833	AJ271759	Super Clade 3	Suessiales	*Symbiodinium* sp.
834	AJ271758	Super Clade 3	Suessiales	*Symbiodinium* sp.
835	AJ271757	Super Clade 3	Suessiales	*Symbiodinium* sp.
836	AJ271756	Super Clade 3	Suessiales	*Symbiodinium* sp.
837	AJ271755	Super Clade 3	Suessiales	*Symbiodinium* sp.
838	AJ271754	Super Clade 3	Suessiales	*Symbiodinium* sp.
839	AJ271481	Super Clade 3	Suessiales	*Symbiodinium* sp.
840	U10893	Super Clade 3	Suessiales	*Symbiodinium* sp.
841	U10892	Super Clade 3	Suessiales	*Symbiodinium* sp.
842	AY456111	Super Clade 3	Suessiales	*Symbiodinium microadriaticum*
843	JN717147	Super Clade 3	Suessiales	*Symbiodinium microadriaticum*
844	AY456113	Super Clade 3	Suessiales	*Symbiodinium* sp.
845	AF225965	Super Clade 3	Suessiales	*Symbiodinium californium*
846	AB863031	Super Clade 3	Suessiales	*Symbiodinium* sp.
847	AB863030	Super Clade 3	Suessiales	*Symbiodinium* sp.
848	HQ822131	Super Clade 3	Suessiales	*Symbiodinium kawagutii*
849	JN255743	Super Clade 3	Suessiales	*Symbiodinium* sp.
850	JN255742	Super Clade 3	Suessiales	*Symbiodinium* sp.
851	JN255741	Super Clade 3	Suessiales	*Symbiodinium* sp.
852	JN255740	Super Clade 3	Suessiales	*Symbiodinium* sp.
853	JN255739	Super Clade 3	Suessiales	*Symbiodinium* sp.
854	JN255738	Super Clade 3	Suessiales	*Symbiodinium* sp.
855	JN255737	Super Clade 3	Suessiales	*Symbiodinium* sp.
856	JN255736	Super Clade 3	Suessiales	*Symbiodinium* sp.
857	JN255735	Super Clade 3	Suessiales	*Symbiodinium* sp.
858	JN255734	Super Clade 3	Suessiales	*Symbiodinium* sp.
859	JN255733	Super Clade 3	Suessiales	*Symbiodinium* sp.
860	AY443023	Super Clade 3	Suessiales	*Symbiodinium* sp.
861	AY139195	Super Clade 3	Suessiales	*Symbiodinium* sp. Clade E
862	AY139194	Super Clade 3	Suessiales	*Symbiodinium* sp. Clade E
863	AY139193	Super Clade 3	Suessiales	*Symbiodinium* sp. Clade E
864	AY139192	Super Clade 3	Suessiales	*Symbiodinium* sp. Clade C
865	EF036539	Super Clade 3	Suessiales	*Symbiodinium goreaui*
866	AY051099	Super Clade 3	Suessiales	*Symbiodinium* sp.
867	AY051098	Super Clade 3	Suessiales	*Symbiodinium* sp.
868	AY051097	Super Clade 3	Suessiales	*Symbiodinium* sp.
869	AY051096	Super Clade 3	Suessiales	*Symbiodinium* sp.
870	AY051095	Super Clade 3	Suessiales	*Symbiodinium* sp.
871	AY051094	Super Clade 3	Suessiales	*Symbiodinium* sp.
872	AY051093	Super Clade 3	Suessiales	*Symbiodinium* sp.
873	AY051092	Super Clade 3	Suessiales	*Symbiodinium* sp.
874	AY051091	Super Clade 3	Suessiales	*Symbiodinium* sp.
875	AY051090	Super Clade 3	Suessiales	*Symbiodinium* sp.
876	AY051089	Super Clade 3	Suessiales	*Symbiodinium* sp.
877	AY051088	Super Clade 3	Suessiales	*Symbiodinium* sp.

878	AY051087	Super Clade 3	Suessiales	*Symbiodinium* sp.
879	AY051086	Super Clade 3	Suessiales	*Symbiodinium* sp.
880	X62650	Super Clade 3	Suessiales	*Symbiodinium pilosum*
881	EF417313	Super Clade 4	Peridiniales	*Apocalathium aciculiferum*
882	AY970653	Super Clade 4	Peridiniales	*Apocalathium aciculiferum*
883	EF417314	Super Clade 4	Peridiniales	*Apocalathium aciculiferum*
884	KF446621	Super Clade 4	Peridiniales	*Apocalathium aciculiferum*
885	EF417315	Super Clade 4	Peridiniales	*Apocalathium aciculiferum*
886	KF751923	Super Clade 4	Peridiniales	*Apocalathium malmogiense*
887	EF417316	Super Clade 4	Peridiniales	*Apocalathium malmogiense*
888	DQ241737	Super Clade 4	Peridiniales	*Crypthecodinium cohnii*
889	FJ821501	Super Clade 4	Peridiniales	*Crypthecodinium cohnii*
890	AB811790	Super Clade 4	Peridiniales	*Crypthecodinium* sp.
891	AB871544	Super Clade 4	Peridiniales	*Crypthecodinium* sp.
892	AB871547	Super Clade 4	Peridiniales	*Crypthecodinium* sp.
893	AB871550	Super Clade 4	Peridiniales	*Crypthecodinium* sp.
894	AB871551	Super Clade 4	Peridiniales	*Crypthecodinium* sp.
895	HM483398	Super Clade 4	Peridiniales	*Duboscquodinium collinii*
896	HM483399	Super Clade 4	Peridiniales	*Duboscquodinium collinii*
897	KR362907	Super Clade 4	Peridiniales	*Pernambugia tuberosa*
898	HQ845331	Super Clade 4	Peridiniales	*Scrippsiella sweeneyae*
899	AF274276	Super Clade 4	Peridiniales	*Scrippsiella sweeneyae*
900	KJ189478	Super Clade 4	Peridiniales	*Scrippsiella erinaceus*
901	LC054940	Super Clade 4	Peridiniales	*Scrippsiella* sp.
902	AB183674	Super Clade 4	Peridiniales	*Scrippsiella* sp.
903	JQ246506	Super Clade 4	Peridiniales	*Scrippsiella* sp.
904	KF733540	Super Clade 4	Peridiniales	*Scrippsiella acuminata*
905	HQ845330	Super Clade 4	Peridiniales	*Scrippsiella acuminata*
906	FR865630	Super Clade 4	Peridiniales	*Scrippsiella acuminata*
907	JX661036	Super Clade 4	Peridiniales	*Scrippsiella acuminata*
908	AY421792	Super Clade 4	Peridiniales	*Scrippsiella acuminata*
909	AJ415515	Super Clade 4	Peridiniales	*Scrippsiella acuminata*
910	AB183671	Super Clade 4	Peridiniales	*Scrippsiella acuminata*
911	AF274277	Super Clade 4	Peridiniales	*Scrippsiella acuminata*
912	EF492513	Super Clade 4	Peridiniales	*Scrippsiella acuminata*
913	HM483396	Super Clade 4	Peridiniales	*Scrippsiella* sp.
914	DQ847435	Super Clade 4	Peridiniales	*Scrippsiella precaria*
915	AM494499	Super Clade 4	Peridiniales	*Scrippsiella* sp.
916	HM483397	Super Clade 4	Peridiniales	*Tintinnophagus acutus*
917	AF080096	Super Clade 4	Peridiniales	*Amyloodinium ocellatum*
918	DQ490256	Super Clade 4	Peridiniales	*Amyloodinium ocellatum*
919	DQ490257	Super Clade 4	Peridiniales	*Amyloodinium ocellatum*
920	KR057921	Super Clade 4	Peridiniales	*Amyloodinium ocellatum*
921	DQ991372	Super Clade 4	Peridiniales	*Cryptoperidiniopsis brodyi*
922	DQ991373	Super Clade 4	Peridiniales	*Cryptoperidiniopsis brodyi*
923	DQ991374	Super Clade 4	Peridiniales	*Cryptoperidiniopsis brodyi*
924	DQ991375	Super Clade 4	Peridiniales	*Cryptoperidiniopsis brodyi*
925	DQ991376	Super Clade 4	Peridiniales	*Cryptoperidiniopsis brodyi*
926	DQ991377	Super Clade 4	Peridiniales	*Cryptoperidiniopsis brodyi*
927	DQ991378	Super Clade 4	Peridiniales	*Cryptoperidiniopsis brodyi*
928	DQ991379	Super Clade 4	Peridiniales	*Cryptoperidiniopsis brodyi*
929	DQ991380	Super Clade 4	Peridiniales	*Cryptoperidiniopsis brodyi*
930	AF080097	Super Clade 4	Peridiniales	*Cryptoperidiniopsis brodyi*
931	AY590476	Super Clade 4	Peridiniales	*Cryptoperidiniopsis brodyi*
932	AJ968729	Super Clade 4	Peridiniales	*Paulsenella vonstoschii*
933	AY112746	Super Clade 4	Peridiniales	*Pfiesteria piscicida*
934	AM231033	Super Clade 4	Peridiniales	*Pfiesteria piscicida*
935	DQ991382	Super Clade 4	Peridiniales	*Pfiesteria piscicida*
936	AY121846	Super Clade 4	Peridiniales	*Pfiesteria piscicida*
937	DQ991381	Super Clade 4	Peridiniales	*Pfiesteria piscicida*
938	AF330600	Super Clade 4	Peridiniales	*Pfiesteria piscicida*
939	AF077055	Super Clade 4	Peridiniales	*Pfiesteria piscicida*
940	AY245693	Super Clade 4	Peridiniales	*Pfiesteria piscicida*

941	AF149793	Super Clade 4	Peridiniales	*Pfiesteria piscicida*
942	AM231028	Super Clade 4	Peridiniales	*Pfiesteria piscicida*
943	FJ600090	Super Clade 4	Peridiniales	*Pfiesteria piscicida*
944	AY033488	Super Clade 4	Peridiniales	*Pfiesteria piscicida*
945	AF080098	Super Clade 4	Peridiniales	*Pfiesteria shumwayae*
946	AY245694	Super Clade 4	Peridiniales	*Pfiesteria shumwayae*
947	AF218805	Super Clade 4	Peridiniales	*Pfiesteria* sp.
948	AJ841809	Super Clade 4	Peridiniales	*Stoeckeria algicida*
949	HG005133	Super Clade 4	Peridiniales	*Stoeckeria algicida*
950	HG005134	Super Clade 4	Peridiniales	*Stoeckeria algicida*
951	FN557541	Super Clade 4	Peridiniales	*Stoeckeria changwonensis*
952	HG005132	Super Clade 4	Peridiniales	*Stoeckeria changwonensis*
953	LK934662	Super Clade 4	Peridiniales	*Aduncodinium glandula*
954	AY443018	Super Clade 4	Peridiniales	*Chimonodinium lomnickii* var. *wierzejskii*
955	EU025010	Super Clade 4	Peridiniales	*Chimonodinium lomnickii* var. *wierzejskii*
956	KF446619	Super Clade 4	Peridiniales	*Chimonodinium lomnickii* var. *wierzejskii*
957	AY443017	Super Clade 4	Peridiniales	*Naiadinium polonicum*
958	JQ639764	Super Clade 4	Peridiniales	*Naiadinium polonicum*
959	KC699492	Super Clade 4	Peridiniales	*Theleodinium calcisporum*
960	LC054944	Super Clade 4	Peridiniales	*Thoracosphaera heimii*
961	AF274278	Super Clade 4	Peridiniales	*Thoracosphaera heimii*
962	HQ845327	Super Clade 4	Peridiniales	*Thoracosphaera heimii*
963	JN615412	Super Clade 5	Peridiniales	*Amphidoma languida*
964	KR362880	Super Clade 5	Peridiniales	*Amphidoma languida*
965	KR362881	Super Clade 5	Peridiniales	*Amphidoma languida*
966	HQ324897	Super Clade 5	Peridiniales	*Azadinium poporum*
967	HQ324898	Super Clade 5	Peridiniales	*Azadinium poporum*
968	HQ324899	Super Clade 5	Peridiniales	*Azadinium poporum*
969	FR877580	Super Clade 5	Peridiniales	*Azadinium poporum*
970	FJ217814	Super Clade 5	Peridiniales	*Azadinium spinosum*
971	JX262491	Super Clade 5	Peridiniales	*Azadinium spinosum*
972	JX559885	Super Clade 5	Peridiniales	*Azadinium spinosum*
973	JN680857	Super Clade 5	Peridiniales	*Azadinium spinosum*
974	GQ914935	Super Clade 5	Peridiniales	*Azadinium obesum*
975	JX661035	Super Clade 5	Peridiniales	Azadinium sp.
976	KF543360	Super Clade 5	Peridiniales	*Azadinium dalianense*
977	KJ481803	Super Clade 5	Peridiniales	*Azadinium trinitatum*
978	KJ481808	Super Clade 5	Peridiniales	*Azadinium trinitatum*
979	KJ481813	Super Clade 5	Peridiniales	*Azadinium trinitatum*
980	KJ481815	Super Clade 5	Peridiniales	*Azadinium trinitatum*
981	KJ481817	Super Clade 5	Peridiniales	*Azadinium trinitatum*
982	KJ481822	Super Clade 5	Peridiniales	*Azadinium cuneatum*
983	KJ481819	Super Clade 5	Peridiniales	*Azadinium cuneatum*
984	JQ247707	Super Clade 5	Peridiniales	*Azadinium caudatum*
985	JQ247701	Super Clade 5	Peridiniales	*Azadinium caudatum*
986	KJ481826	Super Clade 5	Peridiniales	*Azadinium concinnum*
987	JX559886	Super Clade 5	Peridiniales	*Azadinium polongum*
988	KR362890	Super Clade 5	Peridiniales	*Azadinium dexteroporum*
989	KR362889	Super Clade 5	Peridiniales	*Azadinium dexteroporum*
990	LC054925	Super Clade 6	Peridiniales	*Durinskia dybowskii*
991	LC054926	Super Clade 6	Peridiniales	*Durinskia dybowskii*
992	GU999528	Super Clade 6	Peridiniales	*Durinskia dybowskii*
993	AF231803	Super Clade 6	Peridiniales	*Durinskia dybowskii*
994	JF514515	Super Clade 6	Peridiniales	*Durinskia agilis*
995	JF514516	Super Clade 6	Peridiniales	*Durinskia agilis*
996	LC054927	Super Clade 6	Peridiniales	*Durinskia* sp.
997	LC054928	Super Clade 6	Peridiniales	*Durinskia* sp.
998	LC054929	Super Clade 6	Peridiniales	*Durinskia* sp.
999	AB195668	Super Clade 6	Peridiniales	*Galeidinium rugatum*
1000	AF231804	Super Clade 6	Peridiniales	*Kryptoperidinium foliaceum*
1001	EF492508	Super Clade 6	Peridiniales	*Kryptoperidinium foliaceum*
1002	AF274268	Super Clade 6	Peridiniales	*Kryptoperidinium foliaceum*
1003	DQ847436	Super Clade 6	Peridiniales	*Kryptoperidinium foliaceum*

1004	LC054936	Super Clade 6	Peridiniales	*Unruhdinium kevei*
1005	AB353770	Super Clade 6	Peridiniales	*Unruhdinium kevei*
1006	LC054935	Super Clade 6	Peridiniales	*Unruhdinium kevei*
1007	HM596542	Super Clade 6	Peridiniales	*Unruhdinium niei*
1008	HM596543	Super Clade 6	Peridiniales	*Unruhdinium penardii*
1009	AB353771	Super Clade 6	Peridiniales	*Unruhdinium penardii*
1010	JQ639767	Super Clade 6	Peridiniales	*Unruhdinium minima*
1011	KM217384	Super Clade 6	Peridiniales	*Unruhdinium jiulongensis*
1012	AB246744	Super Clade 6	Peridiniales	*Blixaea quinquecornis*
1013	HQ845328	Super Clade 7	Peridiniales	*Ensiculifera loeblichii*
1014	KR362906	Super Clade 7	Peridiniales	*Ensiculifera imariensis*
1015	AF274270	Super Clade 7	Peridiniales	*Pentapharsodinium sp.*
1016	JX262492	Super Clade 7	Peridiniales	*Pentapharsodinium dalei*
1017	AF022201	Super Clade 7	Peridiniales	*Pentapharsodinium tyrrhenicum*
1018	HQ845329	Super Clade 7	Peridiniales	*Pentapharsodinium tyrrhenicum*
1019	EF375879	Super Clade 8	Peridiniales	*Peridinium willei*
1020	DQ166211	Super Clade 8	Peridiniales	*Peridinium* sp.
1021	AB232669	Super Clade 8	Peridiniales	*Peridinium willei*
1022	EF058249	Super Clade 8	Peridiniales	*Peridinium willei*
1023	AF274272	Super Clade 8	Peridiniales	*Peridinium willei*
1024	DQ166210	Super Clade 8	Peridiniales	*Peridinium* sp.
1025	EF058250	Super Clade 8	Peridiniales	*Peridinium willei*
1026	EF058248	Super Clade 8	Peridiniales	*Peridinium volzii*
1027	AF022202	Super Clade 8	Peridiniales	*Peridinium* sp.
1028	EF058245	Super Clade 8	Peridiniales	*Peridinium cinctum*
1029	DQ166209	Super Clade 8	Peridiniales	*Peridinium cinctum*
1030	EF058243	Super Clade 8	Peridiniales	*Peridinium cinctum*
1031	EF058244	Super Clade 8	Peridiniales	*Peridinium cinctum*
1032	AB185114	Super Clade 8	Peridiniales	*Peridinium cinctum*
1033	KF446618	Super Clade 8	Peridiniales	*Peridinium cinctum*
1034	EF058246	Super Clade 8	Peridiniales	*Peridinium gatunense*
1035	DQ487197	Super Clade 8	Peridiniales	*Peridinium gatunense*
1036	DQ166208	Super Clade 8	Peridiniales	*Peridinium gatunense*
1037	EF058242	Super Clade 8	Peridiniales	*Peridinium bipes*
1038	GU046392	Super Clade 8	Peridiniales	*Peridinium bipes*
1039	GU046391	Super Clade 8	Peridiniales	*Peridinium bipes*
1040	AY682801	Super Clade 8	Peridiniales	*Peridinium bipes*
1041	GU046390	Super Clade 8	Peridiniales	*Peridinium bipes*
1042	AY682798	Super Clade 8	Peridiniales	*Peridinium bipes*
1043	AY733008	Super Clade 8	Peridiniales	*Peridinium bipes*
1044	AY682799	Super Clade 8	Peridiniales	*Peridinium bipes*
1045	AF231805	Super Clade 8	Peridiniales	*Peridinium bipes*
1046	AY682800	Super Clade 8	Peridiniales	*Peridinium bipes*
1047	JQ639762	Super Clade 8	Peridiniales	*Peridinium bipes*
1048	DQ980484	Super Clade 8	Peridiniales	*Peridinium limbatum*
1049	DQ980483	Super Clade 8	Peridiniales	*Peridinium limbatum*
1050	DQ980482	Super Clade 8	Peridiniales	*Peridinium limbatum*
1051	AB639343	Super Clade 8	Peridiniales	*Amphidiniopsis rotundata*
1052	JN587284	Super Clade 8	Peridiniales	*Amphidiniopsis hirsuta*
1053	JN587283	Super Clade 8	Peridiniales	*Amphidiniopsis hirsuta*
1054	JN587281	Super Clade 8	Peridiniales	*Amphidiniopsis hirsuta*
1055	JN587282	Super Clade 8	Peridiniales	*Amphidiniopsis swedmarkii*
1056	AY238479	Super Clade 8	Peridiniales	*Amphidiniopsis dragescoi*
1057	AB702985	Super Clade 8	Peridiniales	*Archaeperidinium saanichi*
1058	AB702986	Super Clade 8	Peridiniales	*Archaeperidinium saanichi*
1059	AB702987	Super Clade 8	Peridiniales	*Archaeperidinium saanichi*
1060	AB564308	Super Clade 8	Peridiniales	*Archaeperidinium minutum*
1061	AB564309	Super Clade 8	Peridiniales	*Archaeperidinium minutum*
1062	GQ227501	Super Clade 8	Peridiniales	*Archaeperidinium minutum*
1063	AB780999	Super Clade 8	Peridiniales	*Archaeperidinium minutum*
1064	AB781000	Super Clade 8	Peridiniales	*Archaeperidinium minutum*
1065	AB564298	Super Clade 8	Peridiniales	*Herdmania litoralis*
1066	AB564299	Super Clade 8	Peridiniales	*Herdmania litoralis*

1067	AB564300	Super Clade 8	Peridiniales	*Herdmania litoralis*
1068	AB564303	Super Clade 8	Peridiniales	*Herdmania litoralis*
1069	AB564304	Super Clade 8	Peridiniales	*Herdmania litoralis*
1070	AB564305	Super Clade 8	Peridiniales	*Herdmania litoralis*
1071	AB780843	Super Clade 8	Peridiniales	*Islandinium minutum*
1072	JX627340	Super Clade 8	Peridiniales	*Islandinium minutum*
1073	JX627341	Super Clade 8	Peridiniales	*Islandinium minutum*
1074	JX627342	Super Clade 8	Peridiniales	*Islandinium minutum*
1075	JX627343	Super Clade 8	Peridiniales	*Islandinium minutum*
1076	AB716916	Super Clade 8	Peridiniales	*Islandinium tricingulatum*
1077	AB716917	Super Clade 8	Peridiniales	*Islandinium tricingulatum*
1078	AB716918	Super Clade 8	Peridiniales	*Islandinium tricingulatum*
1079	AB716915	Super Clade 8	Peridiniales	*Protoperidinium parthenopes*
1080	AB716911	Super Clade 8	Peridiniales	*Protoperidinium americanum*
1081	AB780842	Super Clade 8	Peridiniales	*Protoperidinium fukuyoi*
1082	AB716913	Super Clade 8	Peridiniales	*Protoperidinium monovelum*
1083	AB716914	Super Clade 8	Peridiniales	*Protoperidinium monovelum*
1084	AB255834	Super Clade 8	Peridiniales	*Protoperidinium depressum*
1085	AB255833	Super Clade 8	Peridiniales	*Protoperidinium claudicans*
1086	AB255837	Super Clade 8	Peridiniales	*Protoperidinium pentagonum*
1087	AB261516	Super Clade 8	Peridiniales	*Protoperidinium pallidum*
1088	AB181899	Super Clade 8	Peridiniales	*Protoperidinium pallidum*
1089	AY443022	Super Clade 8	Peridiniales	*Protoperidinium pellucidum*
1090	AB181902	Super Clade 8	Peridiniales	*Protoperidinium pellucidum*
1091	AB284159	Super Clade 8	Peridiniales	*Protoperidinium bipes*
1092	AB181888	Super Clade 8	Peridiniales	*Protoperidinium crassipes*
1093	AB181889	Super Clade 8	Peridiniales	*Protoperidinium crassipes*
1094	AB261515	Super Clade 8	Peridiniales	*Protoperidinium crassipes*
1095	AB255835	Super Clade 8	Peridiniales	*Protoperidinium elegans*
1096	AB181892	Super Clade 8	Peridiniales	*Protoperidinium divergens*
1097	AY443021	Super Clade 8	Peridiniales	*Protoperidinium excentricum*
1098	AB275355	Super Clade 8	Peridiniales	*Protoperidinium excentricum*
1099	AB261517	Super Clade 8	Peridiniales	*Protoperidinium punctulatum*
1100	AB181904	Super Clade 8	Peridiniales	*Protoperidinium punctulatum*
1101	AB261518	Super Clade 8	Peridiniales	*Protoperidinium thulesense*
1102	AB261519	Super Clade 8	Peridiniales	*Protoperidinium thulesense*
1103	AB261520	Super Clade 8	Peridiniales	*Protoperidinium thulesense*
1104	AB261521	Super Clade 8	Peridiniales	*Protoperidinium thulesense*
1105	AB261522	Super Clade 8	Peridiniales	*Protoperidinium thulesense*
1106	AB181885	Super Clade 8	Peridiniales	*Protoperidinium conicum*
1107	AB181883	Super Clade 8	Peridiniales	*Protoperidinium conicum*
1108	AY443020	Super Clade 8	Peridiniales	*Protoperidinium conicum*
1109	AB181894	Super Clade 8	Peridiniales	*Protoperidinium leonis*
1110	AB181890	Super Clade 8	Peridiniales	*Protoperidinium denticulatum*
1111	AB181891	Super Clade 8	Peridiniales	*Protoperidinium denticulatum*
1112	AB181881	Super Clade 8	Peridiniales	*Protoperidinium abei*
1113	AB181907	Super Clade 8	Peridiniales	*Protoperidinium thorianum*
1114	AB181908	Super Clade 8	Peridiniales	*Protoperidinium thorianum*
1115	AB181898	Super Clade 8	Peridiniales	*Protoperidinium leonis*
1116	AB181897	Super Clade 8	Peridiniales	*Protoperidinium leonis*
1117	AB181895	Super Clade 8	Peridiniales	*Protoperidinium leonis*
1118	AB181896	Super Clade 8	Peridiniales	*Protoperidinium leonis*
1119	AB181884	Super Clade 8	Peridiniales	*Protoperidinium conicum*
1120	AB181882	Super Clade 8	Peridiniales	*Protoperidinium abei*
1121	AB716912	Super Clade 8	Peridiniales	*Protoperidinium fusiforme*
1122	KJ995958	Super Clade 8	Peridiniales	*Diplopsalis caspica*
1123	AB716909	Super Clade 8	Peridiniales	*Diplopsalis lenticula*
1124	AB261513	Super Clade 8	Peridiniales	*Diplopsalopsis bomba*
1125	AB261514	Super Clade 8	Peridiniales	*Gotoius excentricus*
1126	LC075591	Super Clade 8	Peridiniales	*Kolkwitziella acuta*
1127	AB273724	Super Clade 8	Peridiniales	*Niea torta*
1128	AB273725	Super Clade 8	Peridiniales	*Niea torta*
1129	LC005409	Super Clade 8	Peridiniales	*Niea acanthocysta*

1130	AB273721	Super Clade 8	Peridiniales	*Niea acanthocysta*
1131	AB273722	Super Clade 8	Peridiniales	*Niea acanthocysta*
1132	AB273723	Super Clade 8	Peridiniales	*Niea acanthocysta*
1133	AB716910	Super Clade 8	Peridiniales	*Preperidinium meunieri*
1134	AB261512	Super Clade 8	Peridiniales	*Qia lebouriae*
1135	AF274265	Super Clade 9	Peridiniales	*Heterocapsa niei*
1136	EF492499	Super Clade 9	Peridiniales	*Heterocapsa niei*
1137	LC054932	Super Clade 9	Peridiniales	*Heterocapsa circularisquama*
1138	FJ549370	Super Clade 9	Peridiniales	*Heterocapsa* sp.
1139	AF274266	Super Clade 9	Peridiniales	*Heterocapsa pygmaea*
1140	JX661033	Super Clade 9	Peridiniales	*Heterocapsa* sp.
1141	JX661034	Super Clade 9	Peridiniales	*Heterocapsa* sp.
1142	LC054933	Super Clade 9	Peridiniales	*Heterocapsa psammophila*
1143	GU594638	Super Clade 9	Peridiniales	*Heterocapsa triquetra*
1144	AJ415514	Super Clade 9	Peridiniales	*Heterocapsa triquetra*
1145	AY421787	Super Clade 9	Peridiniales	*Heterocapsa triquetra*
1146	AB183670	Super Clade 9	Peridiniales	*Heterocapsa triquetra*
1147	AF022198	Super Clade 9	Peridiniales	*Heterocapsa triquetra*
1148	DQ388464	Super Clade 9	Peridiniales	*Heterocapsa rotundata*
1149	AF274267	Super Clade 9	Peridiniales	*Heterocapsa rotundata*
1150	JX661030	Super Clade 9	Peridiniales	*Heterocapsa* sp.
1151	JX661031	Super Clade 9	Peridiniales	*Heterocapsa* sp.
1152	KF925338	Super Clade 9	Peridiniales	*Heterocapsa arctica*
1153	EF492492	Super Clade 9	Perdiniales	*Heterocapsa* cf. *niei*
1154	EF492494	Super Clade 9	Perdiniales	*Heterocapsa* sp.
1155	FJ888593	Super Clade 10	Peridiniales	*Blepharocysta* sp.
1156	AF521100	Super Clade 10	Peridiniales	*Lessardia elongata*
1157	FJ888594	Super Clade 10	Peridiniales	*Podolampas palmipes*
1158	FJ888595	Super Clade 10	Peridiniales	*Podolampas bipes*
1159	FJ888596	Super Clade 10	Peridiniales	*Podolampas elegans*
1160	FJ888597	Super Clade 10	Peridiniales	*Podolampas spinifera*
1161	AF521101	Super Clade 10	Peridiniales	*Roscoffia capitata*
1162	DQ174089	Super Clade 11	Prorocentrales	*Prorocentrum mexicanum*
1163	Y16232	Super Clade 11	Prorocentrales	*Prorocentrum mexicanum*
1164	EF492510	Super Clade 11	Prorocentrales	*Prorocentrum mexicanum*
1165	HF565183	Super Clade 11	Prorocentrales	*Prorocentrum rhathymum*
1166	JQ616822	Super Clade 11	Prorocentrales	*Prorocentrum rhathymum*
1167	HF565181	Super Clade 11	Prorocentrales	*Prorocentrum rhathymum*
1168	FJ842096	Super Clade 11	Prorocentrales	*Prorocentrum rhathymum*
1169	KF733536	Super Clade 11	Prorocentrales	*Prorocentrum rhathymum*
1170	HF565182	Super Clade 11	Prorocentrales	*Prorocentrum rhathymum*
1171	JQ390504	Super Clade 11	Prorocentrales	*Prorocentrum texanum*
1172	EF492511	Super Clade 11	Prorocentrales	*Prorocentrum micans*
1173	JN717145	Super Clade 11	Prorocentrales	*Prorocentrum micans*
1174	AJ415519	Super Clade 11	Prorocentrales	*Prorocentrum micans*
1175	AY803739	Super Clade 11	Prorocentrales	*Prorocentrum micans*
1176	EU780638	Super Clade 11	Prorocentrales	*Prorocentrum micans*
1177	DQ004735	Super Clade 11	Prorocentrales	*Prorocentrum micans*
1178	AY833514	Super Clade 11	Prorocentrales	*Prorocentrum micans*
1179	EF492512	Super Clade 11	Prorocentrales	*Prorocentrum triestinum*
1180	DQ004734	Super Clade 11	Prorocentrales	*Prorocentrum triestinum*
1181	AB183673	Super Clade 11	Prorocentrales	*Prorocentrum triestinum*
1182	AY443019	Super Clade 11	Prorocentrales	*Prorocentrum gracile*
1183	DQ028763	Super Clade 11	Prorocentrales	*Prorocentrum cordatum*
1184	EU780639	Super Clade 11	Prorocentrales	*Prorocentrum cordatum*
1185	JX402086	Super Clade 11	Prorocentrales	*Prorocentrum cordatum*
1186	JQ616823	Super Clade 11	Prorocentrales	*Prorocentrum cordatum*
1187	AJ415520	Super Clade 11	Prorocentrales	*Prorocentrum cordatum*
1188	JF715165	Super Clade 11	Prorocentrales	*Prorocentrum cordatum*
1189	FJ587221	Super Clade 11	Prorocentrales	*Prorocentrum cordatum*
1190	AY421791	Super Clade 11	Prorocentrales	*Prorocentrum cordatum*
1191	AY803740	Super Clade 11	Prorocentrales	*Prorocentrum cordatum*
1192	Y16238	Super Clade 11	Prorocentrales	*Prorocentrum cordatum*

1193	AJ841810	Super Clade 11	Prorocentrales	*Prorocentrum donghaiense*
1194	KF032443	Super Clade 11	Prorocentrales	*Prorocentrum donghaiense*
1195	AY551272	Super Clade 11	Prorocentrales	*Prorocentrum donghaiense*
1196	AY803742	Super Clade 11	Prorocentrales	*Prorocentrum dentatum*
1197	AY551273	Super Clade 11	Prorocentrales	*Prorocentrum dentatum*
1198	AB781324	Super Clade 11	Prorocentrales	*Prorocentrum shikokuense*
1199	EF657885	Super Clade 11	Prorocentrales	*Prorocentrum tsawwassenense*
1200	JX912167	Super Clade 11	Prorocentrales	*Prorocentrum fukuyoi*
1201	JX912165	Super Clade 11	Prorocentrales	*Prorocentrum fukuyoi*
1202	Y16239	Super Clade 11	Prorocentrales	*Prorocentrum emarginatum*
1203	GU327677	Super Clade 11	Prorocentrales	*Prorocentrum pseudopanamense*
1204	Y16233	Super Clade 11	Prorocentrales	*Prorocentrum panamense*
1205	GU327678	Super Clade 11	Prorocentrales	*Prorocentrum glenanicum*
1206	GU327679	Super Clade 11	Prorocentrales	*Prorocentrum glenanicum*
1207	FJ489617	Super Clade 11	Prorocentrales	*Prorocentrum levis*
1208	FJ160588	Super Clade 11	Prorocentrales	*Prorocentrum levis*
1209	DQ238043	Super Clade 11	Prorocentrales	*Prorocentrum levis*
1210	JX912166	Super Clade 11	Prorocentrales	*Prorocentrum foraminosum*
1211	Y16237	Super Clade 11	Prorocentrales	*Prorocentrum concavum*
1212	HQ890884	Super Clade 11	Prorocentrales	*Prorocentrum consutum*
1213	FJ842379	Super Clade 11	Prorocentrales	*Prorocentrum consutum*
1214	HQ890882	Super Clade 11	Prorocentrales	*Prorocentrum bimaculatum*
1215	JQ638934	Super Clade 11	Prorocentrales	*Prorocentrum belizeanum*
1216	KF885226	Super Clade 11	Prorocentrales	*Prorocentrum belizeanum*
1217	DQ238042	Super Clade 11	Prorocentrales	*Prorocentrum belizeanum*
1218	KF885224	Super Clade 11	Prorocentrales	*Prorocentrum hoffmannianum*
1219	KF885225	Super Clade 11	Prorocentrales	*Prorocentrum hoffmannianum*
1220	Y16236	Super Clade 11	Prorocentrales	*Prorocentrum maculosum*
1221	JQ638940	Super Clade 11	Prorocentrales	*Prorocentrum maculosum*
1222	KF733552	Super Clade 11	Prorocentrales	*Prorocentrum lima*
1223	Y16234	Super Clade 11	Prorocentrales	*Prorocentrum lima*
1224	AB189773	Super Clade 11	Prorocentrales	*Prorocentrum lima*
1225	AB189774	Super Clade 11	Prorocentrales	*Prorocentrum lima*
1226	AB189775	Super Clade 11	Prorocentrales	*Prorocentrum lima*
1227	AB189776	Super Clade 11	Prorocentrales	*Prorocentrum lima*
1228	AB189777	Super Clade 11	Prorocentrales	*Prorocentrum lima*
1229	AB189778	Super Clade 11	Prorocentrales	*Prorocentrum lima*
1230	AB189779	Super Clade 11	Prorocentrales	*Prorocentrum lima*
1231	AB189780	Super Clade 11	Prorocentrales	*Prorocentrum lima*
1232	Y16235	Super Clade 11	Prorocentrales	*Prorocentrum lima*
1233	EF377326	Super Clade 11	Prorocentrales	*Prorocentrum lima*
1234	JN717143	Super Clade 11	Prorocentrales	*Prorocentrum lima*
1235	FJ160591	Super Clade 11	Prorocentrales	*Procentrum cassubicum*
1236	DQ388460	Super Clade 11	Prorocentrales	*Procentrum cassubicum*
1237	DQ388459	Super Clade 11	Prorocentrales	*Procentrum nanum*
1238	LC054937	Super Clade 11	Prorocentrales	*Plagiodinium belizeanum*
1239	LC054938	Super Clade 11	Prorocentrales	*Plagiodinium* sp.
1240	FR877582	UTD	Peridiniales	*Bysmatrum gregarium*
1241	FR877583	UTD	Peridiniales	*Bysmatrum gregarium*
1242	FR877584	UTD	Peridiniales	*Bysmatrum gregarium*
1243	FR877585	UTD	Peridiniales	*Bysmatrum gregarium*
1244	FR877586	UTD	Peridiniales	*Bysmatrum gregarium*
1245	HQ845326	UTD	Peridiniales	*Bysmatrum subsalsum*
1246	LC002839	UTD	Peridiniales	*Adenoides eludens*
1247	LC002840	UTD	Peridiniales	*Adenoides eludens*
1248	LC002841	UTD	Peridiniales	*Adenoides eludens*
1249	LC002842	UTD	Peridiniales	*Adenoides eludens*
1250	LC002843	UTD	Peridiniales	*Pseudadenoides kofoidii*
1251	AF274249	UTD	Peridiniales	*Adenoides eludens*
1252	DQ975473	UTD	Peridiniales	*Sabulodinium undulatum*
1253	DQ975474	UTD	Peridiniales	*Sabulodinium undulatum*
1254	U52357	UTD	Peridiniales	*Zooxanthella nutricula*
1255	JQ446589	UTD	Peridiniales	*Heterodinium scrippsii*

1256	JQ446590	UTD	Peridiniales	*Heterodinium scrippsii*
1257	JQ446591	UTD	Peridiniales	*Heterodinium scrippsii*
1258	JQ446588	UTD	Peridiniales	*Heterodinium rigdeniae*
1259	JQ446586	UTD	Peridiniales	*Heterodinium globosum*
1260	JQ446587	UTD	Peridiniales	*Heterodinium globosum*
1261	JQ446581	UTD	Peridiniales	*Heterodinium doma*
1262	JQ446592	UTD	Peridiniales	*Heterodinium kofoidii*
1263	JQ446582	UTD	Peridiniales	*Heterodinium milneri*
1264	JQ446583	UTD	Peridiniales	*Heterodinium milneri*
1265	JQ446584	UTD	Peridiniales	*Heterodinium milneri*
1266	JQ446585	UTD	Peridiniales	*Heterodinium milneri*
1267	AF274257	UTD	Peridiniales	*Glenodiniopsis uliginosa*
1268	L13716	UTD	Peridiniales	*Gloeodinium viscum*
1269	AY443016	UTD	Peridiniales	*Hemidinium nasutum*
1270	JQ639763	UTD	Peridiniales	*Palatinus apiculatus*
1271	EF058241	UTD	Peridiniales	*Peridiniopsis borgei*
1272	EF058237	UTD	Dinophyceae ordo incertae sedis	*Glenoaulax inaequalis*
1273	KM879217	UTD	Dinophyceae ordo incertae sedis	*Oodinium pouchetii*
1274	KM879218	UTD	Dinophyceae ordo incertae sedis	*Oodinium pouchetii*
1275	KM879219	UTD	Dinophyceae ordo incertae sedis	*Oodinium pouchetii*
1276	LC054942	UTD	Dinophyceae ordo incertae sedis	*Stylodinium littorale*
1277	FR865631	UTD	Peridiniales	*Parvodinium inconspicuum*
1278	EF058247	UTD	Peridiniales	*Parvodinium inconspicuum*
1279	AF274271	UTD	Peridiniales	*Parvodinium inconspicuum*
1280	GU001637	UTD	Peridiniales	*Parvodinium umbonatum*
1281	EF492509	UTD	Peridiniales	*Peridinium sociale*
1282	LC057317	UTD	Dinophyceae ordo incertae sedis	*Amphidiniella sedentaria*
1283	AB212091	UTD	Dinophyceae ordo incertae sedis	*Amphidiniella sedentaria*
1284	AB036837	UTD	Dinophyceae ordo incertae sedis	*Halostylodinium arenarium*
1285	LC054931	UTD	Dinophyceae ordo incertae sedis	*Halostylodinium arenarium*
1286	AB211357	UTD	Dinophyceae ordo incertae sedis	*Pileidinium ciceropse*
1287	KJ187034	UTD	Dinophyceae ordo incertae sedis	*Ailadinium reticulatum*
1288	KJ187035	UTD	Dinophyceae ordo incertae sedis	*Ailadinium reticulatum*
1289	KF751599	UTD	Dinophyceae ordo incertae sedis	*Madanidinium loirii*
1290	KF751603	UTD	Dinophyceae ordo incertae sedis	*Madanidinium loirii*
1291	EF058238	UTD	Dinophyceae ordo incertae sedis	*Rufusiella insignis*
1292	AY238477	UTD	Gonyaulacales	*Thecadinium yashimaense*
1293	EF492515	UTD	Gonyaulacales	*Thecadinium inclinatum*
1294	AY238478	UTD	Gonyaulacales	*Thecadinium kofoidii*
1295	GU295204	UTD	Gonyaulacales	*Thecadinium kofoidii*
1296	AJ415513	Super Clade 12	Gymnodiniales	*Akashiwo sanguinea*
1297	AB183672	Super Clade 12	Gymnodiniales	*Akashiwo sanguinea*
1298	AY831410	Super Clade 12	Gymnodiniales	*Akashiwo sanguinea*
1299	AY831411	Super Clade 12	Gymnodiniales	*Akashiwo sanguinea*
1300	AY831412	Super Clade 12	Gymnodiniales	*Akashiwo sanguinea*
1301	DQ779987	Super Clade 12	Gymnodiniales	*Akashiwo sanguinea*
1302	DQ779988	Super Clade 12	Gymnodiniales	*Akashiwo sanguinea*
1303	EF492486	Super Clade 12	Gymnodiniales	*Akashiwo sanguinea*
1304	KJ728857	Super Clade 12	Gymnodiniales	*Akashiwo sanguinea*
1305	AY421771	Super Clade 12	Gymnodiniales	*Akashiwo sanguinea*
1306	U41085	Super Clade 12	Gymnodiniales	*Akashiwo sanguinea*
1307	AF276818	Super Clade 12	Gymnodiniales	*Akashiwo sanguinea*
1308	AB232670	Super Clade 12	Gymnodiniales	*Akashiwo sanguinea*
1309	FJ473380	Super Clade 13	Gymnodiniales	*Chytriodinium affine*
1310	KM245128	Super Clade 13	Gymnodiniales	*Chytriodinium* sp.
1311	FJ663049	Super Clade 13	Gymnodiniales	*Chytriodinium roseum*
1312	FJ473378	Super Clade 13	Gymnodiniales	*Dissodinium pseudolunula*
1313	FJ473379	Super Clade 13	Gymnodiniales	*Dissodinium pseudolunula*
1314	LC054930	Super Clade 13	Gymnodiniales	*Gymnodinium dorsalicum*
1315	DQ837534	Super Clade 13	Gymnodiniales	*Gymnodinium dorsalicum*
1316	DQ779992	Super Clade 13	Gymnodiniales	*Gymnodinium impudicum*
1317	EU418974	Super Clade 13	Gymnodiniales	*Gymnodinium impudicum*
1318	GU362426	Super Clade 13	Gymnodiniales	*Gymnodinium impudicum*

1319	DQ779993	Super Clade 13	Gymnodiniales	*Gymnodinium impudicum*
1320	DQ785884	Super Clade 13	Gymnodiniales	*Gymnodinium impudicum*
1321	AF022197	Super Clade 13	Gymnodiniales	*Gymnodinium impudicum*
1322	KP790152	Super Clade 13	Gymnodiniales	*Gymnodinium litoralis*
1323	AF022193	Super Clade 13	Gymnodiniales	*Gymnodinium catenatum*
1324	EU418973	Super Clade 13	Gymnodiniales	*Gymnodinium catenatum*
1325	DQ785883	Super Clade 13	Gymnodiniales	*Gymnodinium catenatum*
1326	DQ779990	Super Clade 13	Gymnodiniales	*Gymnodinium catenatum*
1327	DQ779989	Super Clade 13	Gymnodiniales	*Gymnodinium catenatum*
1328	AY421783	Super Clade 13	Gymnodiniales	*Gymnodinium catenatum*
1329	AB265962	Super Clade 13	Gymnodiniales	*Gymnodinium catenatum*
1330	DQ785882	Super Clade 13	Gymnodiniales	*Gymnodinium catenatum*
1331	EU418972	Super Clade 13	Gymnodiniales	*Gymnodinium catenatum*
1332	EU418954	Super Clade 13	Gymnodiniales	*Gymnodinium catenatum*
1333	JQ638928	Super Clade 13	Gymnodiniales	*Gymnodinium catenatum*
1334	AB265964	Super Clade 13	Gymnodiniales	*Gymnodinium microreticulatum*
1335	AB265965	Super Clade 13	Gymnodiniales	*Gymnodinium microreticulatum*
1336	AB265963	Super Clade 13	Gymnodiniales	*Gymnodinium nolleri*
1337	AF022196	Super Clade 13	Gymnodiniales	*Gymnodinium aureolum*
1338	AY999082	Super Clade 13	Gymnodiniales	*Gymnodinium aureolum*
1339	FN392226	Super Clade 13	Gymnodiniales	*Gymnodinium aureolum*
1340	KJ481834	Super Clade 13	Gymnodiniales	*Gymnodinium aureolum*
1341	HQ270472	Super Clade 13	Gymnodiniales	*Gymnodinium* sp.
1342	HQ270473	Super Clade 13	Gymnodiniales	*Gymnodinium* sp.
1343	JQ639761	Super Clade 13	Gymnodiniales	*Gymnodinium* sp.
1344	FJ024298	Super Clade 13	Gymnodiniales	*Gymnodinium* sp.
1345	AY829527	Super Clade 13	Gymnodiniales	*Gymnodinium* sp.
1346	EU025011	Super Clade 13	Gymnodiniales	*Gymnodinium fuscum*
1347	AF022194	Super Clade 13	Gymnodiniales	*Gymnodinium fuscum*
1348	FJ024297	Super Clade 13	Gymnodiniales	*Gymnodinium* sp.
1349	GQ423576	Super Clade 13	Gymnodiniales	*Gymnodinium* sp.
1350	HG005135	Super Clade 13	Gymnodiniales	*Gymnodinium smaydae*
1351	AB860180	Super Clade 13	Gymnodiniales	*Gymnodinium* sp.
1352	EF058239	Super Clade 13	Gymnodiniales	*Gymnodinium impatiens*
1353	FR720082	Super Clade 13	Gymnodiniales	*Gyrodiniellum shiwhaense*
1354	DQ499645	Super Clade 13	Gymnodiniales	*Lepidodinium viride*
1355	AM184122	Super Clade 13	Gymnodiniales	*Lepidodinium chlorophorum*
1356	AF022199	Super Clade 13	Gymnodiniales	*Lepidodinium viride*
1357	AY331681	Super Clade 13	Gymnodiniales	*Lepidodinium chlorophorum*
1358	AB686255	Super Clade 13	Gymnodiniales	*Lepidodinium* sp.
1359	AB686256	Super Clade 13	Gymnodiniales	*Lepidodinium viride*
1360	AB686253	Super Clade 13	Gymnodiniales	*Lepidodinium chlorophorum*
1361	AB686254	Super Clade 13	Gymnodiniales	*Lepidodinium chlorophorum*
1362	AB921315	Super Clade 13	Gymnodiniales	*Nusuttodinium aeruginosum*
1363	AB921313	Super Clade 13	Gymnodiniales	*Nusuttodinium aeruginosum*
1364	LC027037	Super Clade 13	Gymnodiniales	*Nusuttodinium aeruginosum*
1365	AB921312	Super Clade 13	Gymnodiniales	*Nusuttodinium aeruginosum*
1366	AB921317	Super Clade 13	Gymnodiniales	*Nusuttodinium aeruginosum*
1367	LC027038	Super Clade 13	Gymnodiniales	*Nusuttodinium aeruginosum*
1368	AB921311	Super Clade 13	Gymnodiniales	*Nusuttodinium aeruginosum*
1369	LC027039	Super Clade 13	Gymnodiniales	*Nusuttodinium aeruginosum*
1370	AB921316	Super Clade 13	Gymnodiniales	*Nusuttodinium aeruginosum*
1371	JQ639760	Super Clade 13	Gymnodiniales	*Nusuttodinium acidotum*
1372	AB921309	Super Clade 13	Gymnodiniales	*Nusuttodinium acidotum*
1373	AB921308	Super Clade 13	Gymnodiniales	*Nusuttodinium amphidinioides*
1374	AB921306	Super Clade 13	Gymnodiniales	*Nusuttodinium amphidinioides*
1375	AB921307	Super Clade 13	Gymnodiniales	*Nusuttodinium amphidinioides*
1376	LC027040	Super Clade 13	Gymnodiniales	*Nusuttodinium amphidinioides*
1377	LC027041	Super Clade 13	Gymnodiniales	*Nusuttodinium amphidinioides*
1378	LC027034	Super Clade 13	Gymnodiniales	*Nusuttodinium desymbiontum*
1379	LC027035	Super Clade 13	Gymnodiniales	*Nusuttodinium desymbiontum*
1380	LC027036	Super Clade 13	Gymnodiniales	*Nusuttodinium desymbiontum*
1381	LC027042	Super Clade 13	Gymnodiniales	*Nusuttodinium latum*

1382	LC027043	Super Clade 13	Gymnodiniales	*Nusuttodinium latum*
1383	AB921302	Super Clade 13	Gymnodiniales	*Nusuttodinium latum*
1384	AB921304	Super Clade 13	Gymnodiniales	*Nusuttodinium poecilochroum*
1385	LC027044	Super Clade 13	Gymnodiniales	*Nusuttodinium poecilochroum*
1386	LC027045	Super Clade 13	Gymnodiniales	*Nusuttodinium poecilochroum*
1387	LC027046	Super Clade 13	Gymnodiniales	*Nusuttodinium poecilochroum*
1388	LC027047	Super Clade 13	Gymnodiniales	*Nusuttodinium poecilochroum*
1389	AB921301	Super Clade 13	Gymnodiniales	*Nusuttodinium myriopyrenoides*
1390	AM408889	Super Clade 13	Gymnodiniales	*Paragymnodinium shiwhaense*
1391	DQ371295	Super Clade 13	Gymnodiniales	*Pheopolykrikos beauchampii*
1392	DQ371294	Super Clade 13	Gymnodiniales	*Pheopolykrikos beauchampii*
1393	AY421789	Super Clade 13	Gymnodiniales	*Polykrikos hartmannii*
1394	AB466290	Super Clade 13	Gymnodiniales	*Polykrikos kofoidii*
1395	DQ371291	Super Clade 13	Gymnodiniales	*Polykrikos kofoidii*
1396	DQ371292	Super Clade 13	Gymnodiniales	*Polykrikos kofoidii*
1397	AB466294	Super Clade 13	Gymnodiniales	*Polykrikos kofoidii*
1398	AB466291	Super Clade 13	Gymnodiniales	*Polykrikos kofoidii*
1399	AB466292	Super Clade 13	Gymnodiniales	*Polykrikos kofoidii*
1400	EU418966	Super Clade 13	Gymnodiniales	*Polykrikos schwartzii*
1401	AB466287	Super Clade 13	Gymnodiniales	*Polykrikos schwartzii*
1402	AB466288	Super Clade 13	Gymnodiniales	*Polykrikos schwartzii*
1403	AB466286	Super Clade 13	Gymnodiniales	*Polykrikos schwartzii*
1404	KF806599	Super Clade 13	Gymnodiniales	*Polykrikos tanit*
1405	KF806598	Super Clade 13	Gymnodiniales	*Polykrikos tanit*
1406	JX967270	Super Clade 13	Gymnodiniales	*Polykrikos geminatus*
1407	DQ975472	Super Clade 13	Gymnodiniales	*Polykrikos lebourae*
1408	DQ371293	Super Clade 13	Gymnodiniales	*Polykrikos lebourae*
1409	DQ975471	Super Clade 13	Gymnodiniales	*Polykrikos lebourae*
1410	DQ975470	Super Clade 13	Gymnodiniales	*Polykrikos herdmaniae*
1411	DQ822481	Super Clade 13	Gymnodiniales	*Polykrikos herdmaniae*
1412	KP790164	Super Clade 13	Gymnodiniales	*Polykrikos herdmaniae*
1413	KP790165	Super Clade 13	Gymnodiniales	*Polykrikos herdmaniae*
1414	LC027031	Super Clade 13	Gymnodiniales	*Pellucidodinium psammophilum*
1415	LC027032	Super Clade 13	Gymnodiniales	*Pellucidodinium psammophilum*
1416	LC027033	Super Clade 13	Gymnodiniales	*Pellucidodinium psammophilum*
1417	AB921299	Super Clade 13	Gymnodiniales	*Spiniferodinium palustre*
1418	AB921297	Super Clade 13	Gymnodiniales	*Spiniferodinium galeiforme*
1419	GU295203	Super Clade 13	Gymnodiniales	*Spiniferodinium galeiforme*
1420	LC054941	Super Clade 13	Gymnodiniales	*Spiniferodinium galeiforme*
1421	LC027048	Super Clade 13	Gymnodiniales	*Spiniferodinium galeiforme*
1422	AB626150	Super Clade 13	Gymnodiniales	*Spiniferodinium palauense*
1423	FJ467491	Super Clade 13	Gymnodiniales	*Erythropsidinium agile*
1424	FJ947038	Super Clade 13	Gymnodiniales	*Nematodinium* sp.
1425	FJ947039	Super Clade 13	Gymnodiniales	*Nematodinium* sp.
1426	FJ947036	Super Clade 13	Gymnodiniales	*Proterythropsis* sp.
1427	FJ947037	Super Clade 13	Gymnodiniales	*Proterythropsis* sp.
1428	KP790169	Super Clade 13	Gymnodiniales	*Warnowia* sp.
1429	FJ947040	Super Clade 13	Gymnodiniales	*Warnowia* sp.
1430	FJ467492	Super Clade 13	Gymnodiniales	*Warnowia* sp.
1431	FJ947046	Super Clade 13	Gymnodiniales	*Warnowia* sp.
1432	KP790168	Super Clade 13	Gymnodiniales	*Warnowia* sp.
1433	KP790170	Super Clade 13	Gymnodiniales	*Warnowia* sp.
1434	AB920349	Super Clade 13	Gymnodiniales	*Gymnoxanthella radiolariae*
1435	AB920350	Super Clade 13	Gymnodiniales	*Gymnoxanthella radiolariae*
1436	AB920351	Super Clade 13	Gymnodiniales	*Gymnoxanthella radiolariae*
1437	AB698451	Super Clade 13	Gymnodiniales	*Gymnoxanthella radiolariae*
1438	HM066998	Super Clade 14	Gymnodiniales	*Brachidinium capitatum*
1439	AF274259	Super Clade 14	Gymnodiniales	*Karenia brevis*
1440	AF352818	Super Clade 14	Gymnodiniales	*Karenia brevis*
1441	EF492501	Super Clade 14	Gymnodiniales	*Karenia brevis*
1442	FJ587219	Super Clade 14	Gymnodiniales	*Karenia brevis*
1443	EF492502	Super Clade 14	Gymnodiniales	*Karenia brevis*
1444	EF492503	Super Clade 14	Gymnodiniales	*Karenia brevis*

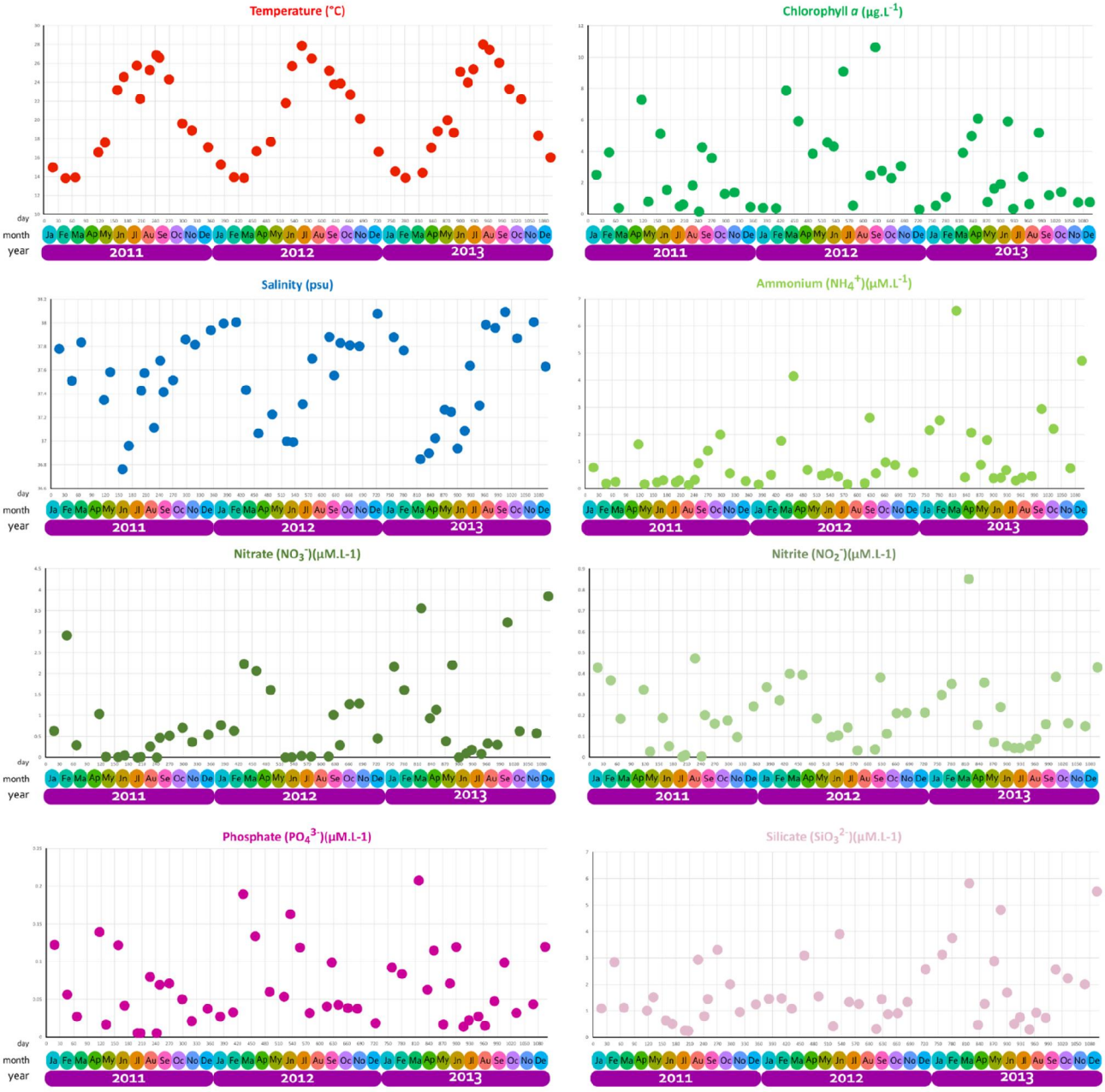
1445	EF492504	Super Clade 14	Gymnodiniales	*Karenia brevis*
1446	DQ847434	Super Clade 14	Gymnodiniales	*Karenia brevis*
1447	AF172714	Super Clade 14	Gymnodiniales	*Karenia brevis*
1448	AJ415518	Super Clade 14	Gymnodiniales	*Karenia brevis*
1449	AF022195	Super Clade 14	Gymnodiniales	*Karenia mikimotoi*
1450	EF492505	Super Clade 14	Gymnodiniales	*Karenia mikimotoi*
1451	AF009131	Super Clade 14	Gymnodiniales	*Karenia mikimotoi*
1452	FR865627	Super Clade 14	Gymnodiniales	*Karenia mikimotoi*
1453	HM067007	Super Clade 14	Gymnodiniales	*Karenia selliformis*
1454	HM067005	Super Clade 14	Gymnodiniales	*Karenia papilionaceae*
1455	HM067002	Super Clade 14	Gymnodiniales	*Karenia bidigitata*
1456	JN986577	Super Clade 14	Gymnodiniales	*Karlodinium veneficum*
1457	AF272045	Super Clade 14	Gymnodiniales	*Karlodinium veneficum*
1458	EF036540	Super Clade 14	Gymnodiniales	*Karlodinium veneficum*
1459	AF272046	Super Clade 14	Gymnodiniales	*Karlodinium veneficum*
1460	AY245692	Super Clade 14	Gymnodiniales	*Karlodinium veneficum*
1461	AF172712	Super Clade 14	Gymnodiniales	*Karlodinium veneficum*
1462	AM494500	Super Clade 14	Gymnodiniales	*Karlodinium veneficum*
1463	AJ415516	Super Clade 14	Gymnodiniales	*Karlodinium veneficum*
1464	EF492506	Super Clade 14	Gymnodiniales	*Karlodinium veneficum*
1465	HQ832504	Super Clade 14	Gymnodiniales	*Karlodinium veneficum*
1466	AF274262	Super Clade 14	Gymnodiniales	*Karlodinium veneficum*
1467	FN357291	Super Clade 14	Gymnodiniales	*Karlodinium* sp.
1468	AY800130	Super Clade 14	Gymnodiniales	*Takayama pulchellum*
1469	HM067010	Super Clade 14	Gymnodiniales	*Takayama acrotrocha*
1470	AJ415517	Super Clade 14	Gymnodiniales	*Karenia* sp.
1471	AF009216	Super Clade 14	Gymnodiniales	*Karenia* sp.
1472	AF172713	Super Clade 14	Gymnodiniales	*Karenia* sp.
1473	DQ779991	Super Clade 14	Gymnodiniales	Kareniaceae_XX_sp.
1474	AY121855	Super Clade 14	Gymnodiniales	*Karlodinium veneficum*
1475	AF274260	Super Clade 14	Gymnodiniales	Kareniaceae_XX_sp.
1476	KP790154	Super Clade 15	Gymnodiniales	*Gyrodinium spirale*
1477	KP790155	Super Clade 15	Gymnodiniales	*Gyrodinium spirale*
1478	KP790156	Super Clade 15	Gymnodiniales	*Gyrodinium spirale*
1479	KP790157	Super Clade 15	Gymnodiniales	*Gyrodinium spirale*
1480	KP790153	Super Clade 15	Gymnodiniales	*Gyrodinium spirale*
1481	AB120002	Super Clade 15	Gymnodiniales	*Gyrodinium spirale*
1482	AB120001	Super Clade 15	Gymnodiniales	*Gyrodinium spirale*
1483	FJ024299	Super Clade 15	Gymnodiniales	*Gyrodinium helveticum*
1484	AB120004	Super Clade 15	Gymnodiniales	*Gyrodinium helveticum*
1485	AB120003	Super Clade 15	Gymnodiniales	*Gyrodinium rubrum*
1486	KP790158	Super Clade 15	Gymnodiniales	*Gyrodinium heterogrammum*
1487	KP790159	Super Clade 15	Gymnodiniales	*Gyrodinium heterogrammum*
1488	HE611580	Super Clade 15	Gymnodiniales	*Gyrodinium moestrupii*
1489	FN669511	Super Clade 15	Gymnodiniales	*Gyrodinium gutrula*
1490	FN669510	Super Clade 15	Gymnodiniales	*Gyrodinium dominans*
1491	FR865623	Super Clade 16	Gymnodiniales	*Amphidinium carterae*
1492	KF733534	Super Clade 16	Gymnodiniales	*Amphidinium carterae*
1493	FR865624	Super Clade 16	Gymnodiniales	*Amphidinium carterae*
1494	AJ415512	Super Clade 16	Gymnodiniales	*Amphidinium carterae*
1495	FR865622	Super Clade 16	Gymnodiniales	*Amphidinium carterae*
1496	AF009217	Super Clade 16	Gymnodiniales	*Amphidinium carterae*
1497	EF057407	Super Clade 16	Gymnodiniales	*Amphidinium carterae*
1498	AF274251	Super Clade 16	Gymnodiniales	*Amphidinium carterae*
1499	JN717139	Super Clade 16	Gymnodiniales	*Amphidinium carterae*
1500	AF274255	Super Clade 16	Gymnodiniales	*Amphidinium carterae*
1501	EU046334	Super Clade 16	Gymnodiniales	*Amphidinium carterae*
1502	EF492485	Super Clade 16	Gymnodiniales	*Amphidinium carterae*
1503	EF057406	Super Clade 16	Gymnodiniales	*Amphidinium carterae*
1504	EU046336	Super Clade 16	Gymnodiniales	*Amphidinium* sp.
1505	EU046337	Super Clade 16	Gymnodiniales	*Amphidinium* sp.
1506	KF733541	Super Clade 16	Gymnodiniales	*Amphidinium klebsii*
1507	KF733526	Super Clade 16	Gymnodiniales	*Amphidinium* sp.

1508	AB103390	Super Clade 16	Gymnodiniales	*Amphidinium* sp.
1509	EU046340	Super Clade 16	Gymnodiniales	*Amphidinium* sp.
1510	AB103389	Super Clade 16	Gymnodiniales	*Amphidinium* sp.
1511	HF674441	Super Clade 16	Gymnodiniales	Amphidinium massartii
1512	HF674442	Super Clade 16	Gymnodiniales	Amphidinium massartii
1513	HF674443	Super Clade 16	Gymnodiniales	Amphidinium massartii
1514	AB092335	Super Clade 16	Gymnodiniales	*Amphidinium* sp.
1515	EU046338	Super Clade 16	Gymnodiniales	*Amphidinium* sp.
1516	EU046339	Super Clade 16	Gymnodiniales	*Amphidinium* sp.
1517	EU035777	Super Clade 16	Gymnodiniales	*Amphidinium* sp.
1518	AB626895	Super Clade 16	Gymnodiniales	*Amphidinium* sp.
1519	L13719	Super Clade 16	Gymnodiniales	Amphidinium gibbosum
1520	EU035776	Super Clade 16	Gymnodiniales	*Amphidinium* sp.
1521	AB107845	Super Clade 16	Gymnodiniales	*Amphidinium* sp.
1522	AB863027	Super Clade 16	Gymnodiniales	*Amphidinium gibbosum*
1523	EU046335	Super Clade 16	Gymnodiniales	*Amphidinium* sp.
1524	LC054920	Super Clade 16	Gymnodiniales	*Amphidinium steinii*
1525	LC054921	Super Clade 16	Gymnodiniales	*Amphidinium steinii*
1526	LC056067	Super Clade 16	Gymnodiniales	*Amphidinium cupulatisquama*
1527	AB704006	Super Clade 16	Gymnodiniales	*Amphidinium operculatum*
1528	AF274254	Super Clade 16	Gymnodiniales	*Amphidinium longum*
1529	GU295202	Super Clade 16	Gymnodiniales	*Amphidinium mootonorum*
1530	AF274253	Super Clade 16	Gymnodiniales	*Amphidinium herdmanii*
1531	KP790166	Super Clade 17	Torodiniales	*Torodinium robustum*
1532	KP790167	Super Clade 17	Torodiniales	*Torodinium robustum*
1533	KR139784	Super Clade 17	Torodiniales	*Torodinium robustum*
1534	KR139781	Super Clade 17	Torodiniales	*Torodinium teredo*
1535	KR139782	Super Clade 17	Torodiniales	*Torodinium teredo*
1536	KR139783	Super Clade 17	Torodiniales	*Torodinium teredo*
1537	KP790160	Super Clade 17	Torodiniales	*Kapelodinium vestifici*
1538	KP790161	Super Clade 17	Torodiniales	*Kapelodinium vestifici*
1539	KP790162	Super Clade 17	Torodiniales	*Kapelodinium vestifici*
1540	JQ439938	Super Clade 18	Dinophyceae ordo incertae sedis	*Esoptrodinium* sp.
1541	JQ439940	Super Clade 18	Dinophyceae ordo incertae sedis	*Esoptrodinium* sp.
1542	JQ439941	Super Clade 18	Dinophyceae ordo incertae sedis	*Esoptrodinium* sp.
1543	JQ439942	Super Clade 18	Dinophyceae ordo incertae sedis	*Esoptrodinium* sp.
1544	JQ439943	Super Clade 18	Dinophyceae ordo incertae sedis	*Esoptrodinium* sp.
1545	JQ439944	Super Clade 18	Dinophyceae ordo incertae sedis	*Esoptrodinium* sp.
1546	EF058240	Super Clade 18	Dinophyceae ordo incertae sedis	*Jadwigia aplanata*
1547	KU359052	Super Clade 18	Dinophyceae ordo incertae sedis	*Tovellia aveirensis*
1548	EF058253	Super Clade 19	Dinophyceae ordo incertae sedis	*Woloszynskia* sp.
1549	JQ639765	Super Clade 20	Dinophyceae ordo incertae sedis	*Woloszynskia* sp.
1550	JN257673	Super Clade 19	Peridiniales	*Blastodinium* sp.
1551	JX473663	Super Clade 19	Peridiniales	*Blastodinium spinulosum*
1552	JN257672	Super Clade 19	Peridiniales	*Blastodinium* sp.
1553	JN257671	Super Clade 19	Peridiniales	*Blastodinium spinulosum*
1554	JN257667	Super Clade 19	Peridiniales	*Blastodinium* sp.
1555	JN257668	Super Clade 19	Peridiniales	*Blastodinium* sp.
1556	FJ228702	Super Clade 19	Peridiniales	*Blastodinium crassum*
1557	HQ226069	Super Clade 19	Peridiniales	*Blastodinium crassum*
1558	HQ226071	Super Clade 19	Peridiniales	*Blastodinium spinulosum*
1559	HQ226070	Super Clade 19	Peridiniales	*Blastodinium spinulosum*
1560	FJ541189	Super Clade 19	Peridiniales	*Blastodinium pruvoti*
1561	JX473666	Super Clade 19	Peridiniales	*Blastodinium oviforme*
1562	JX473667	Super Clade 19	Peridiniales	*Blastodinium contortum*
1563	DQ317536	Super Clade 19	Peridiniales	*Blastodinium contortum*
1564	DQ317537	Super Clade 19	Peridiniales	*Blastodinium contortum*
1565	Fj228701	Super Clade 19	Peridiniales	*Blastodinium contortum*
1566	JN257680	Super Clade 19	Peridiniales	*Blastodinium contortum*
1567	JN257675	Super Clade 19	Peridiniales	*Blastodinium* sp.
1568	JX473656	Super Clade 19	Peridiniales	*Blastodinium mangini*
1569	JN257674	Super Clade 19	Peridiniales	*Blastodinium mangini*
1570	JN257678	Super Clade 19	Peridiniales	*Blastodinium mangini*

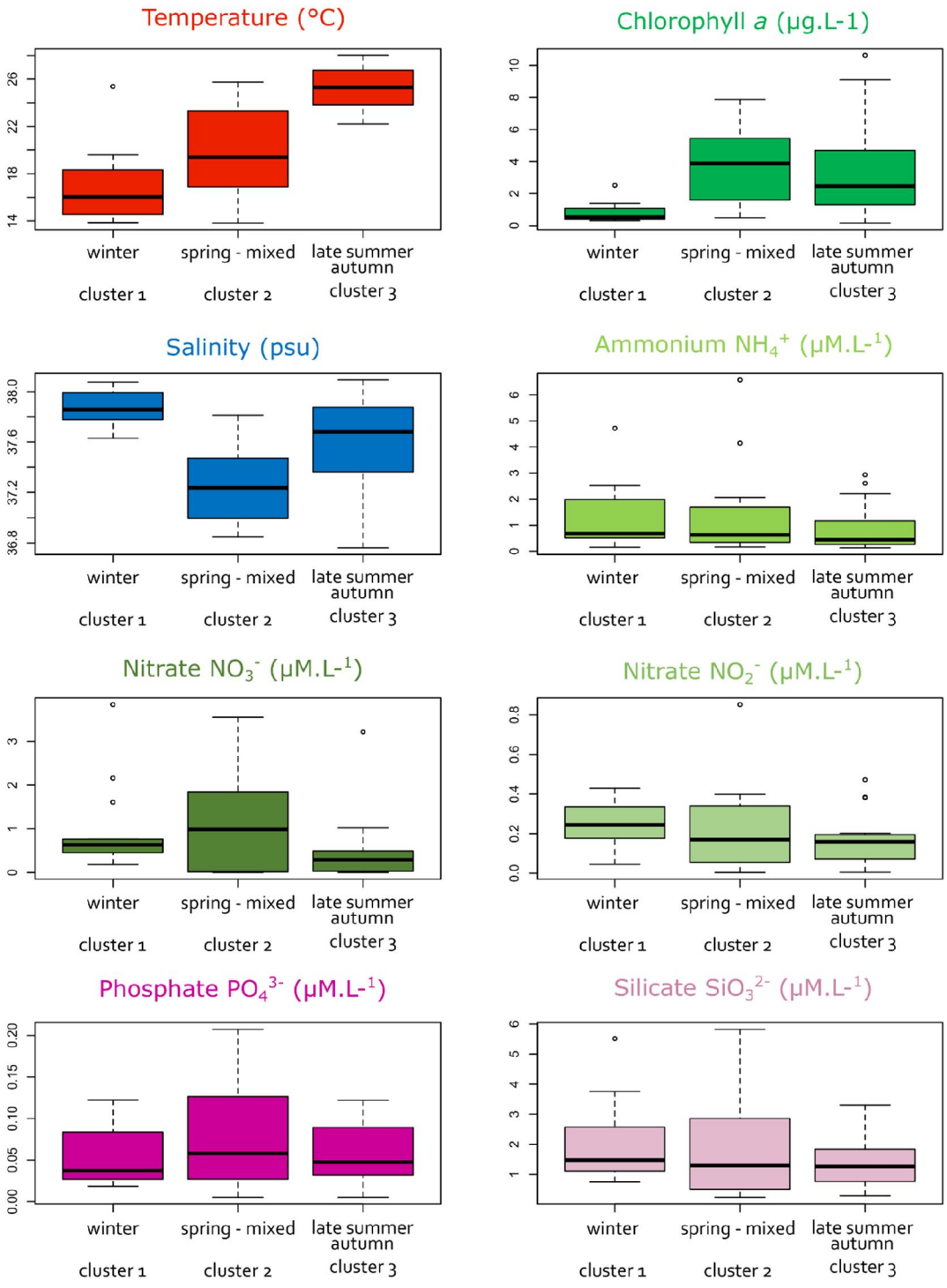
1571	JX473659	Super Clade 19	Peridiniales	*Blastodinium mangini*
1572	JX473664	Super Clade 19	Peridiniales	*Blastodinium mangini*
1573	FJ541188	Super Clade 19	Peridiniales	*Blastodinium galatheanum*
1574	FJ541187	Super Clade 19	Peridiniales	*Blastodinium galatheanum*
1575	JX473662	Super Clade 19	Peridiniales	*Blastodinium navicula*
1576	JN257679	Super Clade 19	Peridiniales	*Blastodinium* sp.
1577	JX473665	Super Clade 19	Peridiniales	*Blastodinium navicula*
1578	DQ317538	Super Clade 19	Peridiniales	*Blastodinium navicula*
1579	JN257676	Super Clade 19	Peridiniales	*Blastodinium mangini*
1580	JX473655	Super Clade 19	Peridiniales	*Blastodinium mangini*
1581	JN257677	Super Clade 19	Peridiniales	*Blastodinium* sp.
1582	KU640194	Super Clade 20	Dinophyceae ordo incertae sedis	*Ptychodiscus noctiluca*
1583	AB704003	UND	Gymnodiniales	*Testudodinium corrugatum*
1584	AB704004	UND	Gymnodiniales	*Testudodinium corrugatum*
1585	LC054943	UND	Gymnodiniales	*Testudodinium* sp.
1586	AB704002	UND	Gymnodiniales	*Testudodinium testudo*
1587	AB704005	UND	Gymnodiniales	*Testudodinium maedaense*
1588	EF492495	UND	Gymnodiniales	*Levanderina fissa*
1589	DQ388457	UND	Gymnodiniales	*Levanderina fissa*
1590	EF492498	UND	Gymnodiniales	*Levanderina fissa*
1591	KP790163	UND	Gymnodiniales	*Levanderina fissa*
1592	AF274261	UND	Gymnodiniales	*Levanderina fissa*
1593	EF492497	UND	Gymnodiniales	*Levanderina fissa*
1594	DQ084522	UND	Gymnodiniales	*Levanderina fissa*
1595	AY721981	UND	Gymnodiniales	*Levanderina fissa*
1596	AF274263	UND	Gymnodiniales	*Levanderina fissa*
1597	AY421786	UND	Gymnodiniales	*Levanderina fissa*
1598	DQ847433	UND	Gymnodiniales	*Levanderina fissa*
1599	AY443010	UND	Gymnodiniales	*Togula britannica*
1600	AF274250	UND	Gymnodiniales	*Togula* sp.
1601	AF274252	UND	Gymnodiniales	*Togula jolla*
1602	JQ179859	UND	Gymnodiniales	*Ankistrodinium semilunatum*
1603	JQ179860	UND	Gymnodiniales	*Ankistrodinium semilunatum*
1604	AF274256	UND	Gymnodiniales	*Ankistrodinium semilunatum*
1605	JQ179861	UND	Gymnodiniales	*Ankistrodinium semilunatum*
1606	AB858349	UND	Gymnodiniales	*Ankistrodinium armigerum*
1607	EU293236	UND	Gymnodiniales	*Apicoporus parvidiaboli*
1608	EU293237	UND	Gymnodiniales	*Apicoporus parvidiaboli*
1609	EU293238	UND	Gymnodiniales	*Apicoporus parvidiaboli*
1610	EU293235	UND	Gymnodiniales	*Apicoporus glaber*
1611	KR139789	UND	Gymnodiniales	*Balechina pachydermata*
1612	KR139790	UND	Gymnodiniales	*Balechina pachydermata*
1613	KR139792	UND	Gymnodiniales	*Balechina pachydermata*
1614	KR139785	UND	Dinophyceae ordo incertae sedis	*Cucumeridinium coeruleum*
1615	KR139786	UND	Dinophyceae ordo incertae sedis	*Cucumeridinium coeruleum*
1616	KR139788	UND	Dinophyceae ordo incertae sedis	*Cucumeridinium cucumis*
1617	KR139787	UND	Dinophyceae ordo incertae sedis	*Cucumeridinium lira*
1618	HQ896315	UND	Gymnodiniales	*Margalefidinium fulvescens*
1619	AB288380	UND	Gymnodiniales	*Margalefidinium fulvescens*
1620	EU418964	UND	Gymnodiniales	*Margalefidinium polykrikoides*
1621	EU418963	UND	Gymnodiniales	*Margalefidinium polykrikoides*
1622	EU418965	UND	Gymnodiniales	*Margalefidinium polykrikoides*
1623	EU418943	UND	Gymnodiniales	*Margalefidinium polykrikoides*
1624	EU418960	UND	Gymnodiniales	*Margalefidinium polykrikoides*
1625	EU418956	UND	Gymnodiniales	*Margalefidinium polykrikoides*
1626	EU418944	UND	Gymnodiniales	*Margalefidinium polykrikoides*
1627	EU418962	UND	Gymnodiniales	*Margalefidinium polykrikoides*
1628	EU418971	UND	Gymnodiniales	*Margalefidinium polykrikoides*
1629	KJ561350	UND	Gymnodiniales	*Margalefidinium polykrikoides*
1630	DQ779984	UND	Gymnodiniales	*Margalefidinium polykrikoides*
1631	DQ915170	UND	Gymnodiniales	*Margalefidinium polykrikoides*
1632	JQ616826	UND	Gymnodiniales	*Margalefidinium polykrikoides*
1633	DQ915169	UND	Gymnodiniales	*Margalefidinium polykrikoides*

1634	JX967271	UND	Gymnodiniales	*Margalefidinium polykrikoides*
1635	AY421779	UND	Gymnodiniales	*Margalefidinium polykrikoides*
1636	EU418940	UND	Gymnodiniales	*Margalefidinium polykrikoides*
1637	EU418946	UND	Gymnodiniales	*Margalefidinium polykrikoides*
1638	EU418955	UND	Gymnodiniales	*Margalefidinium polykrikoides*
1639	EU418958	UND	Gymnodiniales	*Margalefidinium polykrikoides*
1640	EU418957	UND	Gymnodiniales	*Margalefidinium polykrikoides*
1641	EU418959	UND	Gymnodiniales	*Margalefidinium polykrikoides*
1642	EU418961	UND	Gymnodiniales	*Margalefidinium polykrikoides*
1643	EU418952	UND	Gymnodiniales	*Margalefidinium polykrikoides*
1644	EU418951	UND	Gymnodiniales	*Margalefidinium polykrikoides*
1645	EU418950	UND	Gymnodiniales	*Margalefidinium polykrikoides*
1646	EU418941	UND	Gymnodiniales	*Margalefidinium polykrikoides*
1647	EU418953	UND	Gymnodiniales	*Margalefidinium polykrikoides*
1648	EU418942	UND	Gymnodiniales	*Margalefidinium polykrikoides*
1649	LC025891	UND	Gymnodiniales	*Moestrupia oblonga*
1650	LC025894	UND	Gymnodiniales	*Moestrupia oblonga*
1651	LC025889	UND	Gymnodiniales	*Moestrupia oblonga*
1652	LC025888	UND	Gymnodiniales	*Moestrupia oblonga*
1653	LC025887	UND	Gymnodiniales	*Moestrupia oblonga*
1654	LC025886	UND	Gymnodiniales	*Moestrupia oblonga*
1655	LC025892	UND	Gymnodiniales	*Moestrupia oblonga*
1656	LC025893	UND	Gymnodiniales	*Moestrupia oblonga*
1657	LC025879	UND	Gymnodiniales	*Moestrupia oblonga*
1658	LC025880	UND	Gymnodiniales	*Moestrupia oblonga*
1659	LC054934	UND	Gymnodiniales	*Moestrupia oblonga*
1660	LC025895	UND	Gymnodiniales	*Moestrupia oblonga*
1661	LC025896	UND	Gymnodiniales	*Moestrupia oblonga*
1662	LC025897	UND	Gymnodiniales	*Moestrupia oblonga*
1663	LC025890	UND	Gymnodiniales	*Moestrupia oblonga*
1664	LC025883	UND	Gymnodiniales	*Moestrupia oblonga*
1665	LC025884	UND	Gymnodiniales	*Moestrupia oblonga*
1666	LC025885	UND	Gymnodiniales	*Moestrupia oblonga*
1667	LC025881	UND	Gymnodiniales	*Moestrupia oblonga*
1668	LC025882	UND	Gymnodiniales	*Moestrupia oblonga*
1669	KP790150	UND	Gymnodiniales	*Ceratoperidinium falcatum*
1670	KP790151	UND	Gymnodiniales	*Ceratoperidinium falcatum*
1671	AB762397	UND	Gymnodiniales	*Bispinodinium angelaceum*

Appendix 2: Temporal dynamics of environmental variables measured at the LTER-MC station during the studied period (2011-2013). These parameters are representative of long-term variations observed at LTER-MC over 30 years.



Appendix 3: Boxplot grouping the environmental variables measured at the LTER-MC station by seasonal cluster. Cluster 1, 2 and 3 were defined in **Fig.3.3.6**.



Appendix 4: Sequences produced from single cell (SC) or culture (NaD) in LTER-MC during this thesis.

	Code	Name	Superclade #
1	NaD12	*Azadinium* cf. *porporum*	**#5**
2	NaD26	*Biecheleriopsis* cf. *adriatica*	**#3**
3	NaD29	*Biecheleriopsis* cf. *adriatica*	**#3**
4	NaD30	*Biecheleriopsis* cf. *adriatica*	**#3**
5	SC109	*Ceratocorys horrida*	**#1**
6	SC147	*Gyrodinium* cf. *spirale*	**#15**
7	NaD20	*Heterocapsa* cf. *pygmaea*	**#9**
8	NaD18	*Heterocapsa niei*	**#9**
9	NaD46	*Karlodinium* cf. *decipiens*	**#14**
10	NaD47	*Karlodinium* cf. *veneficum*	**#14**
11	NaD22	*Biecheleria* sp.	**#3**
12	Nad24	*Takayama* sp.	**#14**
13	NaD36	*Karlodinium* cf. *veneficum*	**#14**
14	NaD37	*Karlodinium* cf. *decipiens*	**#14**
15	NaD38	*Karlodinium* cf. *decipiens*	**#14**
16	SC75	*Pyrocystis* cf. *noctiluca*	**#1**
17	SC76	*Pyrocystis* cf. *noctiluca*	**#1**
18	SC77	*Pyrocystis* cf. *noctiluca*	**#1**
19	SC80	*Pyrocystis* cf. *noctiluca*	**#1**
20	SC82	*Pyrocystis* cf. *noctiluca*	**#1**
21	SC106	Unknown thecate	**#8**
22	SC108	*Corythodinium constrictum*	**#21**
23	NaD25	*Scripsiella* cf. *acuminata*	**#4**