GENETIC MARKERS FOR THE STUDY OF SOME MARINE MICROALGAE

SARAVANAN V., *M.F.Sc*

DEPARTMENT OF FISHERY MICROBIOLOGY
COLLEGE OF FISHERIES, MANGALORE
KARNATAKA VETERINARY, ANIMAL AND FISHERIES SCIENCES UNIVERSITY, BIDAR

2007

GENETIC MARKERS FOR THE STUDY OF SOME MARINE MICROALGAE

SARAVANAN V., M.F.Sc

DEPARTMENT OF FISHERY MICROBIOLOGY

COLLEGE OF FISHERIES, MANGALORE

KARNATAKA VETERINARY, ANIMAL AND FISHERIES SCIENCES UNIVERSITY, BIDAR

2007

GENETIC MARKERS FOR THE STUDY OF SOME MARINE MICROALGAE

SARAVANAN V., M.F.Sc

Thesis submitted to the KARNATAKA VETERINARY, ANIMAL AND FISHERIES SCIENCES UNIVERSITY, Bidar in partial fulfillment of the requirement of the

Degree of

DOCTOR OF PHILOSOPHY

IN

FISHERY MICROBIOLOGY

MANGALORE APRIL 2007

Dedicated to my loving

Parents and Brother

DEPARTMENT OF FISHERY MICROBIOLOGY

COLLEGE OF FISHERIES, MANGALORE

KARNATAKA VETERINARY, ANIMAL & FISHERIES SCIENCES UNIVERSITY

****

Certificate

This is to certify that the thesis entitled **"Genetic markers for the study of some marine microalgae"** submitted by **Mr. Saravanan, V.** for the award of Doctor of Philosophy in **Fishery Microbiology** of the Karnataka Veterinary, Animal and Fisheries Sciences University, Bidar is a record of research done by him during the period of study in the University under my supervision and the thesis has not previously formed the basis of the award of any degree, diploma, associateship or other similar titles.

Dr. (Mrs.) Indrani Karunasagar
Professor & Head
Dept. of Fishery Microbiology

Mangalore
April 2007

Approved by:

Chairman: _____

Dr. (Mrs.) Indrani Karunasagar

Members : _____

Dr. I. Karunasagar

_____

Dr. M.N. Venugopal

_____

Shri. M.M. Maragal

_____

Prof. K.S. Udupa

Acknowledgements

I take this opportunity to express my deep sense of gratitude and heartfelt thanks to my major advisor, ***Dr. (Mrs.) Indrani Karunasagar, Professor and head,*** *Department of Fishery Microbiology, for her inspiring encouragement, constant support, pragmatic guidance and kind counsel throughout the period of my study and in the preparation of this manuscript. It has been my pleasure and privilege to work under her guidance.*

I express my heartfelt gratitude to ***Dr. I. Karunasagar, National Professor and Director of Research, KVAFSU, Bidar*** *for his pragmatic guidance and kind counsel throughout the period of my study and also for critically going through this manuscript inspite of his busy schedule.*

My sincere thanks are due to my advisory committee members, ***Dr. M.N. Venugopal, Prof. K.S. Udupa*** *and* ***Shri. M. M. Maragal*** *for their valuable suggestions and guidance during the preparation of this manuscript.*

I am grateful to ***Dr. Anna Godhe. Research Scientist, Goteborg University, Sweden,*** *For her timely help in the preparation of this manuscript.*

I take this opportunity to thank ***everyone in the department*** *for all their love, help, guidance, support and encouragement rendered throughout my stay.*

Above all, the real guiding force behind all these efforts is ***my mother, father, brothers, sisters,*** *and* ***sisters-in-law*** *without whose blessings and sacrifice this endeavour would not have been possible.*

Finally, ***travel grant*** *provided by* ***SIDA*** *to carry out part of* ***my Ph.D research studies*** *in* ***Goteborg University, Sweden,*** *is greatly acknowledged.*

Mangalore
April 2007

Saravanan. V.
(SARAVANAN. V)

CONTENTS

1. INTRODUCTION

I. INTRODUCTION

Phytoplankton comprises the base of food chain in the marine environment. On certain occasions, excessive growth of microalgae occurs in the aquatic environment and this is generally referred to as “bloom”. Most phytoplankton blooms in the coastal zones are harmless, but certain harmful algal blooms cause damage both to the environment and also the species inhabiting the particular marine environment (Hallegraeff, 1995). Some planktonic cells secrete mucus which mechanically damage or clog gills and other delicate tissues of animals, other species produce hemolytic toxins which affect sessile organisms including seaweeds (Gjosacter *et al.,* 2000). Certain toxins produced by the harmful algal blooms pass through the food chain ultimately causing human illness or even death (Godhe *et al.,* 2002). Among more than 5000 known species of marine phytoplankton, only a minor fraction has the capacity to produce potent toxins that can find their way through fish and shellfish to humans (Sournia *et al.,* 1991). The apparent increase in the algal blooms may be due to the eutrophication of waters or introduction of new species via transport of microalgae as resting stages either in ship’s ballast water or associated with the translocation of shellfish stocks from one area to another (Anderson, 1989). The rising awareness and economic interest in harmful algal blooms have led to an increased knowledge of different toxins produced by these microalgae (Hallegraeff, 1995).

Many of the harmful bloom-forming microalgae have a resting stage called “cyst” in the marine sediment. It is often thought that the resting cysts act as “seeds” for initiation of blooms whenever ecological conditions are favorable for the planktonic status. Whenever blooms occur, it is not clear whether the source is the cyst in the sediment or whether the bloom has migrated from another location. Recent studies have shown that phytoplankers have extensive intraspecific variation within a geographic location, these variations can be measured by genetic markers such as randomly amplified polymorphic DNA (RAPD) (Medlin *et al.*, 1996; Shankle

et al., 2004) and microsatellite DNA markers (Rynearson and Ambrust, 2004).

Microsatellites consist of tandem repeats of sequence units, each generally 2-10 bp in length, such as (CA)n or (AAT)n. Microsatellites are widely dispersed in eukaryotic genomes and are often highly polymorphic in nature, due to the variation in the number of repeat units (Litt and Luty, 1989; Weber and May, 1989; Tautz, 1989). The existence of simple repetitive elements, the building blocks of microsatellite loci in eukaryotic genomes has been documented since the 1970s, though the large number and almost ubiquitous distribution of such sequences throughout the eukaryotes was first highlighted by Hamada *et al.*, (1982), who found hundreds of copies of poly (dT-dG)n sequences in yeast and thousands in vertebrates (Tautz and Renz, 1984). High numbers of "cryptic" repeats, scrambled arrangement of repetitive sequences were shown to occur in the genome and these were found to vary both within and between species. The existence of these scrambled motifs led to the speculation that replication slippage was the major genomic mechanism involved in the propagation of complex microsatellite sequences (Tautz *et al.*, 1986). The discovery in the human genome of hypervariable tandem repeats that contained longer repeat units, often called as minisatellites and their use in DNA fingerprinting led to the widespread use of repetitive sequences for the individual identification, parentage testing and genome mapping (Jeffreys *et al.*, 1985). As with microsatellite loci, minisatellite alleles vary in the number of tandemly repeated elements and hence a general designation for both is "Variable number of tandem repeat loci (VNTRs)". PCR primers are designed to recognize sequences flanking the tandem repeats and the polymorphic amplified products are separated on polyacrylamide gels (Weber and May, 1989: Tautz, 1989).

Microsatellite loci have been widely used and have proven to be highly polymorphic, their abundance and ubiquitous distribution have made them very valuable genetic markers (Amos *et al.*, 1993). Microsatellite sequences have become important tools for linkage mapping in diverse organisms from

humans to mosquitoes (Weissenbach *et al.*, 1992; Dietrich *et al.*, 1992; Zheng *et al.*, 1993) and their use has enabled the identification of several quantitative trait loci (Todd *et al.*, 1991). Microsatellites have proved valuable in forensic cases to identify human remains (Hagelberg *et al.*, 1991) and to analyze museum specimens (Ellegren, 1991). In addition, microsatellites have been successfully applied to parentage and relatedness testing in humans and other species (Morin *et al.*, 1993) and this technique is proved to be especially valuable where non-invasive sampling methods are necessary (Woodruff, 1993). An additional, potentially valuable characteristic of Microsatellites is that primers developed in one species can be used in related taxa.

Against this background, the objectives of the investigation were as follows,

1) To evaluate the potential application of molecular techniques like Random Amplification of Polymorphic DNA (RAPD) to study genetic differences in microalgae using *Skeletonema marinoi* as a model

2) To develop some microsatellite markers for *S. marinoi* and study the potential of this marker to study genetic variation in different populations of *S. marinoi*.

II. REVIEW OF LITERATURE

II. REVIEW OF LITERATURE

2.1. Phytoplankton biodiversity

Phytoplankters form the first tropic level in the aquatic environment along with aquatic weeds. Their role in trapping solar energy and conversion of inorganic nutrients to organic matter is unquestionably clear. The growth of phytoplankton with respect to nutrients and light energy is fairly understood. An outburst of planktonic growth, which may be a consequence of nutrient-rich environment, is termed as ‘bloom’. Occurrence of such bloom in the sea causing seawater discoloration, irrespective of the causative organisms, has been called red tide, though always not restricted to red discoloration. Apart from the nutrients, physical parameters also are considered to be important factors contributing to the formation and maintenance of blooms.

It has not always been possible to correlate the bloom formation with different physico-chemical parameters. Moreover, similar nutrient status and physical condition do not always induce bloom formation. Many scientists consider a group of factors influencing each other to be involved in the induction of bloom formation. The absolute dependence of a few phytoplankton on some vitamins which they are not capable to produce has been demonstrated by various workers. For the explosive growth of such phytoplankton to form a bloom, they must depend on other organisms for vitamins, growth initiators and other biotic factors, thus suggesting an intricate relationship among organisms. Bacteria form one such group of organisms which play a vital role in growth and metabolism of phytoplankton. Change of phytoplankton population from blue-green algae which do not require vitamin B_{12} to vitamin B_{12} requiring diatoms and green algae due to production of vitamin B_{12} by bacteria has been reported by Stewart *et al.* (1977). Massive fish kills and different types of toxic outbreaks are related to the formation of blooms by both flagellate and non-flagellate phytoplankton. Thus, phytoplankton blooms have become both ecologically and economically important.

Dinoflagellates are capable of producing toxins that accumulate in the body of filter feeding clams, oysters, scallops and other lamellibranchs. Based on the produced symptoms of outbreak, the toxins produced are classified. Paralytic shellfish poisoning (PSP) is caused by the toxins produced by dinoflagellates like *Alexandrium tamarensis*, *A. catenella*, *Gonyaulax* sp and *Protogonyaulax tamarensis*. The toxic principle involved in PSP is saxitoxin (STX), named after the organism from which it was first isolated namely, the Alaska butter clam *Saxidomus giganteus*. As many as 17 derivatives of saxitoxin have been identified so far and many of them are named as gonyautoxins (GTX) and numbered GTX1 to GTX8. Diarrhetic shellfish poisoning (DSP) has been recognized since its discovery by Yasumoto in 1978. DSP is caused by accumulation of toxins known to be produced by *Dinophysis fortii*, *D. acuminata*, *D. acuta*, *D. mitra*, *D. norvegica*, *D. rotundata*, *D. tripos* and *Prorocentrum lima*. The DSP symptoms are similar to gastroenteritis associated with consumption of polluted shellfish and often been overlooked (Yasumoto, 1990).

About 12 toxins which may be involved in shellfish poisoning have been identified including okadoic acid polyether macrolides. Ciguatera intoxication is a common phenomenon occurring due to ingestion of coral reef fishes. A range of toxicity symptoms are involved in this poisoning. An epiphytic dinoflagellate, *Gambierdiscus toxicus*, is implicated to be the primary producer of ciguatoxin (CTX) and maitotoxin (MTX), the toxic principles involved in ciguatera poisoning. Neurotoxic shellfish poisoning (NSP) is caused by the red tide forming dinoflagellate, *Ptychodiscus brevis* which is one of the most harmful species causing fish kills and destruction of other marine life. Health problems are caused by inhaling the wind sprayed cells on the beaches and by consumption of infested shellfish. One of the toxins involved is named brevitoxin after the dinoflagellate. Amnesic shellfish poisoning (ASP), unlike other shellfish poisoning, does not involve dinoflagellates. A pennate diatom, *Nitzschia pungens* was found to be responsible for this type of poisoning. Later, *Pseudonitzchia australis* was

recorded to be an ASP producer. Domoic acid, an excitatory amino acid, is implicated to be the toxic principle, the consumption of which can lead to gastroenteritis and neurological disorders including amnesia. Venerupin shellfish toxin (VSP), isolated from the dinoflagellate *Prorocentrum minimum,* caused the death of 114 people in Japan in 1942. The vector in this case was the bivalve *Venerupis semidecussata* after which the toxin is named (Nakazima, 1965). Tetrodotoxin (TTX), which is a potent neurotoxin like saxitoxin, responsible for pufferfish poisoning and is widely distributed among various animals including few bacterial species, has been reported to be present in small quantities in the dinoflagellate *Alexandrium tamarensis* (Kodama *et al.*, 1993). Hemolysin is believed to be involved in fish kills during blooms of *Gymnodinium aureolum* and *Chrysochromulina polylepis.* Hemolysin and its derivatives are low molecular weight compounds isolated from the dinoflagellate *Amphidinum carterae.*

Phytoplankton and shellfish associated toxicity have been subjected to study world wide. A harmful phytoplankton bloom does not always consist of a toxic species. It can be harmful by virtue of oxygen and ammonia production during its bloom and on collapse, this can be equally lethal to marine life as toxins. Besides, blooms of several species cause choking of gills in fishes obstructing respiration and by increasing viscosity of water, physically obstructing their movement. Moreover, fish kills by *Chaetoceros* spp. are believed to be due to mechanical damage to fish by spines.

2.2. Diatoms

Many diatom species are known to have global distributions and some of these are able to bloom under a wide variety of environmental conditions. In order for diatoms with extensive geographic ranges to successfully meet selective pressure in a variety of ecological niches, it is hypothesized that they display high phenotypic and genetic diversity (Rynearson and Armbrust, 2000). Genetic and physiological variations in a given environment should allow them to thrive over much broader ranges of hydrographic conditions than would be possible for a homogeneous population, and this undoubtedly contributes to their ecological success throughout the marine environment. Several studies have shown high levels of genetic diversity within a phytoplankton species on local or regional scales (Gallagher, 1980;

Orsini *et al*., 2004; Shankle *et al*., 2004; Nagai *et al*., 2007). Results indicate that despite the potential for widespread dispersal by water currents, distinct populations exist on both large (Evans *et al*., 2005), and small geographic scales (Rynearson and Armbrust, 2005; Rynearson *et al*., 2006), and they are maintained in the specific area for long times (Rynearson and Armbrust, 2004). Several experiments have revealed that clonal cultures kept under identical conditions display a large variety of physiologic response (Brand, 1982; Liang *et al*., 2005). In a few studies, the differentiated physiological characteristics of particular clones have also been linked to distinct geographic or temporal restricted populations (Gallagher, 1982; Moore and Chrisholm, 1999; Rynearson and Armbrust, 2004). Repeated monitoring of particular clones after years in culture, has identified the response as endogenous and independent of external condition (Gallagher, 1982). The genetically differentiated populations have thus adapted to particular physiologic characteristics which are most suitable to their indigenous environment, and it is therefore suggested that advection from that particular environment is a loss for the species as such, as the advected cells are probably less well adapted to a new environment (Brand, 1982; Moore and Chrisholm, 1999).

2.2.1. Diatom blooms

In many estuaries along west coast of India, diatom blooms are observed in upwelled waters rich in nutrients, during June-October (Ramamirthan and Jayaraman, 1963; Devassy, 1983). Standing crops of phytoplankton reach peaks by the end of October with diatoms like *Coscinodiscus* spp. *Leptocylinders* spp. *Navicula* spp. *Pleurosigma* spp. *Rhizosolonia* spp and *Skeletonema* spp. predominating. In a mixed diatom-blue green algal bloom which occurred in Mangalore during 1972, *Fragilaria oceanica* constituted 92% of cells with *Rhizosolenia* sp. *Pleurosigma* sp. *Coscinodiscus* sp. and *Richelia intracelluaris* being the others. Phosphate levels were high during the bloom (Devassy, 1974). Nair and Subrahmanyan (1955) suggested that *F. oceanica* acts as an indicator species for the abundance of oil sardine. Blooms of *Asterionella japonica* have been recorded in the east coast off Waltair (Subba Rao, 1969). Blooms of *Nitzschia sigma, Skeletonema costatum* and mixed blooms of centric and pennate diatoms have been recorded (Ramamurthy *et al.,* 1972; Devassy and Bhattathiri, 1974). Ganapathy and Raman (1979) related *Skeletonema* blooms

in Vishakhapatnam harbor to organic pollution in the area. Shetty and Saha (1971) suggested that the blooms of the centric diatom *Hemidiscus hardmannianus* can be used as indicator species to large scale inland shoaling of *Hilsa ilisha*.

2.3. Genetic diversity in marine phytoplankton

Doyle (1975) hypothesized that planktonic algae consist of a multitude of competing genotypes and it has been assumed that these taxa have little genetic structure over large geographic areas. Palumbi (1992) suggested that speciation and dispersal mechanisms in marine planktonic organisms are very different from those that are present on land. With the advent of nucleic acid based methods, the lack of genetic structure in the marine phytoplankton community has now been seriously challenged.

Gallagher (1982) demonstrated that temporal genetic change can often be greater than spatial change, calling into question the idea of temporal stasis (Hedgecock, 1994). The rate of genetic change often does occur on ecological time scales. The reasons for this are unclear but such changes play an important role in determining how local adaptations and speciation occur in apparently homogeneous populations (Palumbi, 1992).

2.3.1. Phytoplankton biodiversity

Much of our limited knowledge about phytoplankton biodiversity stems from the difficulty of finding polymorphic markers for ecological genetic studies. These problems have been overcome to a large extent, due to the emergence of high resolution DNA fingerprinting techniques. At present most genetic studies in phytoplankton research rely on clonal cultures for their analysis. These single cell isolations, made from natural populations are difficult to perform at sea (Hillis *et al.*, 1996).

In many algal groups, life histories are incomplete and if the algae undergo sexual reproduction during culturing, then this may alter the type of genetic analysis performed. Significant genetic diversity has also been found

to exist both within and between phytoplankton populations, based on physiological or biochemical measurements (Wood and Leatham, 1992). These data have been used to speculate on hidden biodiversity and on temporal and spatial structuring of genetic diversity (Partensky *et al.*, 1993).

2.4. Phylogenetic classification among diatoms (Bacillariophyta)

Diatoms are one of the most successful groups of microalgae, with over 10,000 known species occurring in both aquatic and terrestrial habitats. The cellwall of diatoms made up of silica, which forms the valves and girdle band's and the unusual pattern of reduction in cell size of one of the daughter cells following mitosis have been well documented (Mann and Merchant, 1989). Central to diatom taxonomy and phylogeny has been an assumption that the group contains two forms, based on mode of sexual reproduction, pattern centers or symmetry and plastid structure (Round *et al.*, 1990).

The classification system of diatoms recognizes three classes; Cosinodiscophyceae (centric diatoms), Fragilariophyceae (araphid pennate diatoms) and Bacillariophyceae (raphid pennate diatoms). The raphe and the tubes through the cellwall are probably the most important features used to infer phylogenetic relationships among diatoms (Medlin *et al.*, 1986). Relationships at higher taxonomic level in diatoms are virtually unexplored using modern phylogenetic and morphometric tools. In this regard, RNA sequence comparison has proved to be a powerful tool for resolving phylogenetic relationships, since it is a functionally stable evolutionary marker that is independent of morphometric characters (Woese, 1987; Bhattacharya *et al.*, 1992). Medlin *et al.* (1993) reported for the first time that centric and araphid pennate diatoms had a paraphyletic origin using the rRNA sequence comparison study.

2.5. *Skeletonema marinoi*

S. marinoi cells are 2-12 µm in diameter, forming long often curved or coiled colonies. Each cell contains one or two chloroplasts with the valve face slightly convex with a vertical mantle. The intercalary fultoportula

process (IFPP) of sibling valves are either aligned or displaced with a zig-zag connection line. The terminal rimoportula (TRP) is located close to the central annulus and has a long tubular process with a slightly flared or cup shaped apex. The copulae show the typical central ridge, which is flanked on both sides by transverse ribs interspaced by rows of pores (Sarno *et al.*, 2005).

The planktonic diatom *S. marinoi* has a wide geographic distribution. This species is a common member of coastal plankton bloom and can reach concentrations of millions of cells per liter. *S. marinoi* contributes significantly to phytoplankton blooms in temperate waters, including the Skagerrak-Kattegat and Baltic Sea (Tallberg and Heiskanen, 1998). It is found throughout the year in these waters (Lange *et al.*, 1992), and displays pronounced density peaks, especially during spring bloom. It has a benthic resting stage, which is very abundant in temperate sediments (McQuoid, 2002). Examinations of *S. marinoi* seed bank estimate up to 3.5 million propagules/g sediment, and these cells are capable of surviving for several decades buried in the sediments (McQuoid *et al.*, 2002). In addition, *S. marinoi* is an important primary producer and constitutes a valuable food source for higher tropic levels (Josefson and Hansen, 2003). Thus, the study of this organism is of high ecologic relevance. *S. marinoi* forms long clonal chains and is simple to isolate, and moreover, it is easy to maintain in culture. Single cell isolation and subsequent survival of monoclonal cultures are almost 100% (Godhe *et al.*, 2006). This is an important factor, because bias due to one strain's ability to survive in culture is eliminated. Consequently, it is well suited as a model species for studies on genetic and physiological phytoplankton intra-specific variation and differentiation.

Significant genetic heterogeneity among local and regional populations of *Skeletonema* have previously been demonstrated with several molecular methods such as allozyme studies (Gallagher, 1980), sequencing of ribosomal DNA (Alverson and Kolnick, 2005; Godhe *et al.*, 2006; Ellegaard *et al.*, 2007), and random amplified polymorphic DNA, RAPD (Godhe *et al.*, 2006).

In microcosm experiments previously conducted during multiple seasons in Gullmar Fjord, vegetative populations of *Skeletonema* were established by both planktonic and benthic propagules. In the autumn experiment, *S. marinoi* cells were larger when seeded by benthic propagules and smaller when arising from planktonic cells (McQuoid and Godhe, 2004). Because repeated vegetative division reduces diatom cell size, *S. marinoi* seeded from the plankton may have been dividing for many generations compared to benthic resting stages. However, size variation between treatments and seasons may also reflect the presence of genetically distinct clones at different time of the year in the same fjord. Indeed, clones of *S. marinoi* isolated from different seasons in Gullmar Fjord showed heterogeneity within the LSU rDNA sequence, with differences seen even among clones established from the same plankton net sample (Ellegaard *et al.*, 2007). The differences in sizes and heterogeneity of a conserved part of the genome indicate the presence of several distinct populations with a possible temporal distinct distribution.

2.6. Molecular techniques in phytoplankton research

The evolutionary and ecological questions in phytoplankton are increasingly studied using molecular techniques. The application of molecular techniques in phytoplankton research started during the 1970s. Evolutionary questions mainly focused on phylogenetic relationships among species and its taxonomic implications (Klein *et al.*, 1973; Wilmotte *et al.*, 1992; Neilan *et al.*, 1994; Melkonian and Surek, 1995; Buchheim *et al.*, 1996; Rudi *et al.*, 1997; An *et al.*, 1999; Iteman *et al.*, 2002). The taxonomical identification of species or strains using molecular techniques provides a new set of data that can be used to verify the classification based on morphological and physiological characters (Medlin *et al.*, 1993, 1996; Edvardsen *et al.*, 2000; Lundholm *et al.*, 2002). Ecological studies of phytoplankton applied molecular tools to investigate genetic diversity among phytoplankton populations (van Jannen *et al.*, 1998; van Hannen *et al.*, 1999; Larsen *et al.*, 2001). They were also used to identify and screen phytoplankton populations based on the presence of phytoplankton parasites (Wilson *et al.*, 1993; Chen and Suttle, 1995; Chen *et al.*, 1996) or the ability

to produce toxins or some other physiological trait (Chinain *et al.*, 1997; Tillett *et al.*, 2000, 2001; Dittmann *et al.*, 2001; Kaebernick *et al.*, 2002).

Molecular biology at present provides an array of valuable tools that can be used to study genetic relationships at the strain, population or higher taxonomic level. However the application of molecular tools to a particular research is not straight forward, due to the large number of techniques available. Each molecular technique is based on a number of assumptions such as the homology criterion, marker independence and Hardy-Weinberg equilibrium of populations. In addition, different techniques have their own strengths and weaknesses in aspects like resolution, costs and technical expertise. Furthermore, species differ in their genetic variations at particular loci. Hence, a technique that reveals polymorphism in one species does not necessarily reveal the same level of polymorphism in another species studied.

2.6.1. Allozyme electrophoresis

Allozyme electrophoresis was the first molecular technique used for the genetic research of phytoplankton. Enzymes are proteins that can be separated on a gel based on their charge, size and shape. These are classified as isozymes or a more specific set called allozymes. Isozymes are different molecular forms of an enzyme that are encoded by different loci, and allozymes are molecular forms of an enzyme that are encoded by different alleles at a specific gene locus (Baker, 2000). Earlier molecular studies on phytoplankton were based on allozyme electrophoresis to resolve taxonomical differences resulting from insufficient morphological data at the species or strain level (Klein *et al.*, 1973).

Murphy and Guillard (1976) observed genetic similarities between clones of centric marine diatoms *Thalassiosira pseudonana* and *T. fluviatilis* using allozyme electrophoresis banding patterns. Gallagher (1980, 1982) used allozyme electrophoresis in population genetic studies of diatom *Skeletonema costatum* and compared differences in banding patterns to physiological variation among different strains. Gallagher (1980) demonstrated the

existence of spatial and temporal differences in enzyme banding patterns between populations from different locations. Soudek and Robinson (1983) investigated genetic variation in populations of the freshwater diatom *Asterionella formosa* using allozyme electrophoresis. Allozyme electrophoresis was used to examine the relationship between isozyme variation and the ability to produce toxins by the dinoflagellate *Gambierdiscus toxicus* (Chinain *et al.,* 1997).

2.6.2. Restriction fragment length polymorphism (RFLP)

Restriction fragment length polymorphism referred to as RFLP was one of the first DNA-based molecular techniques applied in phytoplankton studies (Botstein *et al.,* 1980). Restriction endonucleases are enzymes that cleave DNA molecules at specific nucleotide sequences. The cleaved fragments are separated using gel electrophoresis and hybridized using DNA probes via southern hybridization techniques (Wood and Townsend, 1990). Stabile *et al.* (1990) studied the genetic variations that exist between *Skeletonema costatum* populations. The PCR amplified product can be subjected to restriction enzyme digestion and analyzed on ethidium bromide or silver stained gel. This new application of RFLP technique is referred to as cleaved amplified polymorphic site or CAPS (Neilan *et al.,* 1997). Wilson *et al.* (1993) demonstrated the application of PCR-RFLP in the identification of parasites like viruses attacking cyanobacteria and eukaryotic phytoplankton (Chen and Suttle, 1995; Chen *et al.,* 1996).

2.6.3. Denaturing gradient gel electrophoresis (DGGE)

Denaturing gradient gel electrophoresis detect polymorphisms in small (200-700 bp) genomic restriction fragments. DNA fragments as they progress into higher denaturing conditions melt depending on their molecular weight and nucleotide sequence in the melted region (Meyers *et al.,* 1987). van Jannen *et al.* (1998) showed the changes in bacterial and algal community structure after cyanophage-induced lysis of filamentous cyanobacteria using DGGE (van Hannen *et al.,* 1999; Larsen *et al.,* 2001). This technique was

applied to investigate the temporal variation in strain abundance of toxic dinoflagellate *Pfiesteria piscicida* (Coyne *et al.*, 2001).

2.6.4. Random amplified polymorphic DNA (RAPD)

The random amplified polymorphic DNA or RAPD was first described in 1990 (Williams *et al.*, 1990). Primers, 8-10 bp in length are chosen at random without prior knowledge of any primer sites in the genome of the organism (Welsh and McClelland, 1990). The random primers act as both forward and reverse primers, and simultaneously, amplify fragments from three to ten genomic sites. Polymorphisms are detected as the presence or absence of bands of a particular size. Bolch *et al.* (1999) used the RAPD technique to study genetic variation among 19 strains of the genus *Nodularia*, in addition to the sequence information of the phycocyanin intergenic spacer region. Neilan *et al.* (1995) applied the RAPD analysis to identify members of the cyanobacterial genera *Anabaena* and *Microcystis.* Lewis *et al.* (1997) showed the genetic variation both between and within seven populations of freshwater diatom *Fragilaria capurina* using RAPD analysis. Genetic variation in strains of the toxic dinoflagellate *Gymnodinium catenatum* was studied to test the possibility of this flagellate being introduced into Tasmanian waters via ballast water of cargo ships (Bolch *et al.*, 1999). The RAPD is an inexpensive, sensitive technique that can be applied to identify specific markers in population studies, but the only disadvantage is the lack of reproducibility and dominance of markers (Baker, 2000).

2.6.5. Amplified fragment length polymorphisms (AFLP)

The amplified fragment length polymorphisms (AFLP) technique was developed by Vos and co-workers and is more reproducible than RAPD (Vos *et al.*, 1995). AFLP was used to study the mating population types of the green alga *Chlamydomonas reinhardtii* (Werner *et al.*, 2001). Four mating type markers were identified; two were present in the minus mating type and the other two markers were present only in the plus mating type strains (Ferris *et al.*, 1996). Vrieling *et al.* (1997) used the AFLP technique which involves the restriction of genomic DNA followed by ligation of adaptors to

restriction sites and selective PCR amplification of adapted restriction fragments. Gandon (1998) demonstrated the susceptibility of parasite strains on host isolates, he found that virulence tended to be highest on the local hosts, the phenomenon known as local adaptation (Lively and Dybdahl, 2000). Savelkoul *et al.* (1999) reviewed the application of AFLP technique in plant, animal and microbiological genetics. AFLP technique is more expensive, and requires more technical expertise and laboratory equipment than RAPD.

2.6.6. Molecular probes

Rudi *et al.* (1998) evaluated the application of molecular probes together with PCR for the detection of toxin, producing cyanobacteria and dinoflagellates in natural waters (Oldach *et al.*, 2000). Molecular probe or heteroduplex mobility assay is based on the principle that heteroduplex hybridization complex migrates slower through the gel matrix compared to homoduplexes.

2.7. Microsatellites

Microsatellites consist of multiple copies of tandemly arranged simple sequence repeats that range in size from 1 to 6bp (Tautz, 1989; Litt and Luty, 1989). Wright (1993) noted that microsatellites occur once every 10 kb in fishes. Liu *et al.* (2001) reported the occurrence of microsatellite repeat in gene-coding regions or exons, introns and also within the non-gene sequences. Generally speaking, microsatellites containing larger number of repeats are polymorphic in nature, although polymorphisms have been observed in microsatellites containing as few as five repeats. Polymorphisms are based on size differences due to varying numbers of repeat units contained by alleles at a given locus (Karsi *et al.*, 2000). Weber and Wong (1993) reported that microsatellite mutation rates are as high as 10^{-2} per generation and these are believed to be caused by the mechanism of polymerase slippage during DNA replication which results in difference in the number of repeat units (Levinson and Gutman, 1987; Crawford and Cuthbertson, 1996). Balloux and Lugon-Moulin (2002) observed that in a few

fish species alleles with very large differences in repeat numbers were present, which is predictive of an infinite allele model.

Another advantage of microsatellites is that they are inherited in a Mendelian fashion as co-dominant markers, in addition to their abundance, even genomic distribution, small locus size and high polymorphism. But the development of microsatellite marker involves a large amount of investment and effort, since each microsatellite locus has to be identified and its flanking region sequenced for the design of primers (Ostrander *et al.,* 1992; Kijas *et al.,* 1994). Over the past decade, microsatellite markers have been extensively used in research studies comprising genome mapping, parentage, kinships and stock structure analysis (O'Connell and Wright, 1997).

2.7.1. Microsatellite biology

Messier *et al.* (1996) summarized the life cycles of individual microsatellite arrays and observed that these events stretch over tens or even hundreds of million years (Primmer and Ellegren, 1998; Amos, 1999; Estoup and Cornuet, 1999). Brooker *et al.* (1994) reported the existence of 64 microsatellite loci from Atlantic cod *Gadus morhua,* and they observed that certain alleles were found larger than twice the size of even the largest mammalian microsatellites. Amos (1999) found that microsatellite loci in vertebrates are longer and more common than invertebrates and among vertebrates cold-blooded species had longer repeat arrays. Hancock (1999) reported the uniform distribution of microsatellites across intergenic areas with the exception of telomeres and he also observed trinucleotide repeats to be very rare in coding regions.

Microsatellites are reported to be born from regions of "cryptic simplicity" that is regions in which variants of simple repetitive DNA sequence motifs are over-represented (Tautz *et al.,* 1986). Levinson and Gutman (1987a) established that the predominant mechanism in microsatellite tract development is 'slipped-strand' mispairing (Eisen, 1999). Edwards *et al.* (1992) observed that in microorganisms, microsatellite arrays containing

longer tetra-nucleotide repeats evolve faster than those containing shorter di-nucleotide repeats. This is attributed to relatively inefficient repair of larger mismatched segments (Weber and Wong, 1993). The majority of mutations which have been observed directly are due to the gain in repeat mutations (Amos *et al.,* 1996; Primmer *et al.,* 1996; Brinkmann *et al.,* 1998). Callan *et al.* (1993) reported that the pseudo-death of an allele or locus occurs at any stage in the life cycle, due to single or multiple nucleotide substitution, insertion or deletion in the flanking regions, which prevents primer binding and results in the formation of null alleles, is responsible for the majority of failed cross-species amplification experiments and thus referred to as ghost loci. Estoup *et al.* (1995) concluded that size homoplasy is a potentially serious problem when assessing microsatellite array size, since alleles of identical size could be products of different evolutionary lineages (Estoup and Cornuet, 1999). Size homoplasy has already been reported in studies concerning complex tetra-nucleotide repeats in primates (Meyer *et al.,* 1995) and humans (Kimpton *et al.,* 1993). The major impact of size homoplasy is to overestimate relatedness.

2.7.2. Genomic distribution of microsatellites

Microsatellites are short stretches of DNA consisting of tandemly repeated units of 1-6 bp in length (Beckmann and Weber, 1992). Gur-Arie *et al.* (2000) demonstrated that they are widely distributed both in prokaryotes and eukaryotes and are found to be present even in the smallest bacterial genomes. The majority of microsatellites found are di-nucleotide repeats and (AC)n repeats were found to be the most common in the genome of vertebrates (Toth *et al.,* 2000). Beckmann and Weber (1992) reported that the mononucleotide repeats mostly represented by the poly (A/T) traits are the most frequent class of SSR repeats in primates. Tri, tetra, penta and hexa-nucleotide repeats were found to be 1.5-fold less common in genomic DNA of vertebrates when compared to di-nucleotide repeats (Toth *et al.,* 2000). Beckmann and Weber (1992) showed that in case of human genome, one microsatellite repeat was found every 6 kb and one (CA) repeat occurred every 30 kb of the genomic DNA.

According to Toth *et al.* (2000), microsatellites could be found anywhere in the genome, both in the protein encoding and noncoding regions of the DNA. But Metzgar *et al.* (2000) observed that in eukaryotic organisms, SSRs were found to be in excess in the noncoding regions when compared to the coding DNA. A relatively low frequency of SSRs in the coding regions could be explained by the negative selection against frame-shift mutations in the translated sequences (Metzgar *et al.,* 2000; Li *et al.,* 2004). In vertebrates di-nucleotides are 30-40 fold less frequently distributed in exons when compared to intergenic and intronic sequences (Toth *et al.,* 2000). Dokholyan *et al.* (2000) established that long dimeric motifs were highly unstable within expressed sequences, while in noncoding regions most di-nucleotide repeats had surprisingly long stretches probably due to the high tolerance of noncoding DNA to mutations.

The potential size expansion of di or tetra-nucleotide microsatellite repeats in untranslated and intron regions could lead to disruption of native proteins and formation of new genes with frame-shift mutation (Liu *et al.,* 1999a). Wren *et al.* (2000) observed that in contrast to other types of repeat motifs, triplet repeats were found in high frequency both in coding and noncoding genomic regions (Morgante *et al.,* 2002). In all vertebrates, (G+C) rich motifs are most common tri-nucleotide repeats and these repeats were found to dominate in the exon regions when compared to the intronic sequences (Toth *et al.,* 2000).

2.7.3. Microsatellite evolution

An important feature of microsatellites as molecular markers is their hyper mutability and hyper variability in species and populations, the microsatellite mutation rate is estimated to be 10^{-2}-10^{-6} per locus per generation (Li, 1997; Ellegren *et al.,* 2000). Neff and Gross (2001) showed that length was a major factor influencing microsatellite mutation rate and they also observed a directional mutation towards an increase in microsatellite array length.

2.7.4. DNA polymerase slippage and unequal recombination

DNA polymerase slippage and unequal recombination have been suggested to explain microsatellite generation and evolution. Polymerase slippage involves transient dissociation of replicating DNA strands with subsequent reassociation (Schlotterer and Tautz, 1992; Richards and Sutherland, 1994). Sinden (1999) concluded that the slipped structure can be stabilized by hairpin, triplex or quadruplex arrangement of DNA strands. During DNA replication, the repeat regions are preferred target sites for mutations; thus, microsatellite stability at multiple steps *in vivo* is controlled through the DNA mismatch repair (MMR) system (Sia *et al.*, 1997). The mismatch repair system is responsible for correction of replication mistakes and suppression of recombination events between diverged sequences (Kolodner and Marsischky, 1999). Sia *et al.* (1997) noted that if the MMR system is defective, the coding sequences with tandem repeats are subjected to mutations.

Non-reciprocal recombination or gene conversion also plays a vital role in genetic instability of certain microsatellites, particularly the triplet motifs (Jakupciak and Wells, 2000). Richard and Dujon (1997) indicated that replication slippage and recombination events could interact, affecting stability of the microsatellite loci.

2.7.5. Homoplasy

Slipped-strand mispairing during DNA replication represents the predominant mutational mechanism for the development of microsatellites (Schlotterer and Tautz, 1992). Size homoplasy is generated by the mutation process which allows the same microsatellite allele to arise multiple times. Homoplasy represents similarity of traits due to convergent evolution, parallelism, evolutionary reversal, horizontal gene transfer and gene duplications. Sanderson and Hufford (1996) established that homoplasy can violate basic assumptions of analysis of genetic markers, in which variants of similar phenotype are assumed to be derived from a common ancestor.

2.8. Models of microsatellite mutation

Various models describing mutational process that account for homoplasy conditions have been proposed.

2.8.1. Stepwise mutational model (SMM)

Kimura and Ohta (1978) proposed that the stepwise mutational model assumes that all mutational events are caused due to the changes in a single nucleotide repeat only (Bell and Jurka, 1997).

2.8.2. Two phase mutational model (TPM)

The two-phase mutational model allows a proportion of mutations to involve changes greater than single nucleotide repeats (Di Rienzo *et al.*, 1994).

2.8.3. Infinite allele model (IAM)

The infinite-allele model does not consider homoplasy events, since every mutation results in the creation of a new allele (Kimura and Crow, 1964).

2.9. Functions of microsatellites

Microsatellites represent selectively neutral DNA markers and multiple studies have proved the functional relevance of a number of microsatellite markers.

2.9.1. DNA structure

Microsatellites are involved in the formation of a wide variety of unusual DNA structures with simple and complex loop-folding patterns (Rich *et al.*, 1984). Several microsatellite sequences such as (AC)n and (GAA)n exhibit non B-DNA structural properties (Epplen *et al.*, 1993). Canapa *et al.* (2002) showed that various telomeric and centromeric chromosome regions were rich in long arrays of mono, di, tri and hexa-nucleotide motifs. The expanded stretches of simple repeat sequences oriented towards 5’-3’

direction at the end of eukaryotic chromosomes constitute a substantial portion of repetitive DNA in the telomeric regions (Henderson, 1995).

Fang and Sech (1995) found that the (TTAGGG)n hexamer sequence is recognized by a telomerase, which synthesizes telomere repeats on to the chromosome ends to overcome the loss of sequences during DNA replication thus, microsatellites play an important role in the organization of chromosome structure (Martins *et al.,* 2004).

2.9.2. DNA recombination

Microsatellites are considered as hot spots for recombination events (Jeffreys *et al.,* 1998). For recombination events, di-nucleotide motifs are considered as preferential sites due to their high affinity towards recombination enzymes. Some microsatellite sequences, such as GT, CA (etc) may influence recombination events directly through their effects on DNA structure (Biet *et al.,* 1999). Okuda *et al.* (2000) reported that the *rh* gene responsible for the determination of blood groups in higher vertebrates contained multiple microsatellites sequences in their intron regions. Microsatellites were shown to be associated with Rh phenotype in eukaryotes and archaea via replication slippage and gene conversion or recombination mechanisms (Fujiwara *et al.,* 1999; Okuda and Kajii, 2002).

2.9.3. DNA replication

Microsatellites influence DNA replication process. For instance, in certain cases the DNA amplification process is terminated within a specific fragment which consists of $d(GA)_{27}$ x $d(TC)_{27}$ tracts. Such sequences are situated at the end of amplicon-forming loops, which serve as a stop signal for DNA polymerase activity (Li *et al.,* 2002). Johannsdottir *et al.* (2000) noted that MMR deficiency causes frame-shift mutations resulting in insertions and deletions of repeat units within genes, which could be responsible for initiation of tumorigenesis. Chang *et al.* (2001) observed that microsatellites can affect enzymatic reactions controlling mutation rates and cell cycle processes.

2.9.4. Gene expression

Several data show that microsatellites located in the promoter regions can influence gene expression activities (Mansour *et al.,* 1998). In certain fishes like Japanese puffer fish, microsatellites have been found in the transcription factor-binding sites of promoter regions of many important genes (Edwards *et al.,* 1998). Streelman and Kocher (2002) described that microsatellite repeat numbers could significantly influence gene expression levels as described for a di-nucleotide microsatellite repeat (CA/GA)n in the Nile tilapia prolactin1(*PRL1)* promoter site. Transcribed microsatellites located in 5' untranslated regions (UTRs) form specific and unusual DNA structures; thus, the length of the repeat regions affect the translation levels from the target mRNA. Meloni *et al.* (1998) concluded that intronic microsatellites can also affect gene transcription; such an affect was measured for the tetrameric microsatellites located in intron1 of the human tyrosine hydroxylase gene (Gebhardt *et al.,* 1999).

Microsatellite regions in introns serve as target sites for a variety of expression-regulated proteins and these bind nuclear protein regulatory molecules with high affinity (Epplen *et al.,* 1993; 1996). Microsatellites situated in the 3' UTR regions affect gene expression through their influence on the stability of transcribed products (Fabregat *et al.,* 1995). Wang *et al.* (2002) described that long stretches of polypurine and pyrimidine repeat motifs at 3'UTR region could destabilize the 3'end of the mRNA molecule, thereby increasing its availability for degradation by intracellular enzymes. Microsatellite variation can produce either drastic or quantitative variation in gene expression activity. The changes in microsatellite array size serve as a rich source of variation in fitness-related traits of natural populations, due to their genomic overabundance and high mutability (Kashi *et al.,* 1997; Streelman and Kocher, 2002).

The role of SSR array size may be especially important in case of population survival and adaptation to varying environmental conditions (Blankenship *et al.,* 2002). Metzgar *et al,* (2000) reported that the presence

of contingency loci in many bacterial species could explain how microsatellite variability influences the adaptive evolution of microbial pathogens. The contingency loci are specifically associated with genes controlling the antigenic type and phase state of pathogenic microorganisms (Bayliss *et al.*, 2001).

2.10. Type I and Type II markers

Molecular markers are divided into Type I markers associated with genes of known functions and Type II markers associated with anonymous genomic sequences (O'Brein, 1991). Type II markers, which represent microsatellite sequences are usually located in the noncoding intergenic regions. The construction of small insert genomic libraries enriched in tandem repeat arrays are the fastest and simplest way to characterize large number of type II microsatellites (Zane *et al.*, 2002). The enrichment technique usually includes selective hybridization of fragmented genomic DNA with tandem repeat containing oligonucleotide probe and further PCR amplification of the hybridization products. A microsatellite library usually contains 1000-4000 recombinant clones, screening of these clones typically yields about 10-15% unique polymorphic SSRs from a single library (Zane *et al.*, 2002). Bruford *et al.* (1996) showed that Type II markers could be rapidly developed from SSRs isolated previously from closely related species. Cross-species amplification of SSRs provides a possibility to superimpose genetic information from one species to another, their SSR markers could be used for population genetics, parentage analysis and other applications without having a need to isolate polymorphic microsatellites (Cairney *et al.*, 2000; Leclerc *et al.*, 2000).

Type I markers are rather difficult to develop when compared to type II markers (Liu *et al.*, 1999). The sequencing of microsatellite clones from cDNA libraries is the most effective and rapid way of developing Type I markers. Both 5' and 3' ends of the cDNA clones are sequenced to produce expressed sequence tags (ESTs). EST collections provide a good source of sequence data that can be effectively used for gene discovery, genome

annotation and comparative genetic analysis (Dunham, 2004; Ng *et al.*, 2005). The development of EST-derived microsatellite markers includes preliminary analysis of EST sequences from the DNA database in order to remove poly(A) and poly(T) stretches. Following the identification of microsatellite-containing ESTs, flanking primers are designed to amplify microsatellite containing regions (Kantety *et al.*, 2002). Thiel *et al.* (2003) described that in order to hypothesize putative functions of SSR-containing genes, the sequences obtained are compared with amino acid sequences present in the database.

2.11. Applications of microsatellites

SSRs are highly polymorphic due to the variation in the number of repeat units. These are rapidly detected by polymerase chain reaction (PCR) using unique oligonucleotide primers that flank the microsatellite regions, thus, defining the microsatellite locus (Amos and Pemberton, 1992). Microsatellites have been successfully applied in a variety of research fields and disciplines due to their multi-allelic nature, co-dominant inheritance, extensive genome coverage and relative abundance (Powel *et al.*, 1996).

2.11.1. Genetic mapping

Genetic mapping is one the major research fields in which microsatellite markers have been applied. SSRs remain the marker of choice for the construction of linkage maps, due to their high polymorphism and require only small amounts of DNA for each test. According to Cullis (2002), one of the major disadvantages of microsatellite markers is that they are mostly anonymous DNA fragments. Type I coding markers are incorporated within Type II noncoding microsatellites for building dense linkage map framework. Compared to Type II markers, mapping of Type I marker directly shows the location of genes within the linkage map (Poompuang and Hallerman, 1997; Walbieser and Wolter, 1999).

2.11.2. Linkage maps

Linkage maps are known as recombination maps and define the order and distance of loci along a chromosome on the basis of inheritance in families or mapping populations. During linkage map construction, co-segregating markers are placed into linkage groups and the proportion of recombinants detected between linked markers are used as a measure of distance between them (Hartl and Jones, 2001).

Microsatellites and AFLP represent DNA markers, which are extremely useful for constructing primary framework maps that could be further enriched using Type I markers (Woods *et al.*, 2000). Sakamoto *et al.* (2000) proposed that the linkage map length differs between sexes; in species with XY sex determination system, the female map is generally longer than the male map because of higher recombination rates in females compared to males (Singer *et al.*, 2002).

2.11.3. Sex determination

Naruse *et al.* (2000) showed that microsatellites provide a good opportunity to detect the sex-determining locus, due to specific features in the heterogametic sex such as an obvious reduction in recombination between markers linked to the sex determination regions in male compared to female meioses (Peichel *et al.*, 2004).

2.11.4. Individual DNA identification

Microsatellites are widely used in forensic science for individual DNA identification, due to their high stability in degraded DNA (Schneider *et al.*, 2004). In addition, microsatellites are highly unique in nature because of their high degree of allelic variability (Primmer *et al.*, 2000). Burger *et al.* (1999) observed that microsatellite loci remain relatively stable in bone remnants and dental tissues, providing the basis for successful application of ancient DNA for molecular analysis. Microsatellite analysis from ancient human and animal remnants are essential to understand the genetic diversity in current populations and to provide substantial perspectives for the analysis

of archeological issues, heritable diseases and establishment of genealogies in pre-historic populations (Zierdt *et al.,* 1996).

2.11.5. Parentage assignment

Microsatellites represent co-dominant single locus DNA markers; the progeny population inherits one allele from the male parent and another allele from the female parent. This inheritance pattern explains the extreme popularity of polymorphic SSR loci in paternity testing. Microsatellites are extensively exploited for paternity analysis of natural populations and hatchery brood stocks (Liu and Cordes, 2004).

Several mathematical tools are available to evaluate genetic relatedness and inheritance existing in different populations (Luikart and England, 1999; Blouin, 2003). Existing analytical packages combine several different approaches in parentage analysis such as merging likelihood techniques with tools assessing statistical confidence in parental assignment (Jones and Ardren, 2003).

2.12. Phylogeny and conservation genetics

The majority of evolutionary studies attempt to infer phylogenetic relationship from microsatellite data at levels below species level using stretches of tandem repeats, which evolve more rapidly than flanking regions (McCartney *et al.,* 2003; Stamford and Taylor, 2004). Flanking regions of microsatellites have proved valuable in establishing phylogenetic relationships between species and families, because they evolve much more slowly than number of tandem repeats. Koskinen *et al.* (2002) noted that phylogeographical applications of microsatellite are suitable, where population structure is observed over a large geographical scale (Gum *et al.*, 2005). Microsatellite genotypes are particularly helpful to detect structure in closely related populations, regardless of whether they are in evolutionary equilibrium. Some of the common measures of genetic diversity are heterozygosity, allelic diversity that is the number of alleles at a particular

locus in the population and the proportion of polymorphic loci (Pujolar *et al.*, 2005).

Genome-wide scans using microsatellite markers could be applied to search any species for which a high-density genetic map is available (Storz, 2005). The occurrence of highly polymorphic microsatellites in untranslated regions of ESTs is a major source of gene -associated polymorphisms representing genetic signatures of divergent selection (Rise *et al.*, 2004). Ferguson *et al.* (1995) described that the major focus of conservation genetics is on the effects of inbreeding, demography, contemporary genetic structuring and adaptation on the long-term survival of a species (Wan *et al.*, 2004). Genetic analysis often reveals differences between sampled populations with substantial phylogenetic separation, which have minor but statistically significant differences in allele frequency of nuclear or mitochondrial loci (Wan *et al.*, 2004). The high mutation rates of SSR loci also support the use of microsatellite markers in the genetic analysis of very recent events in population dynamics (Zhang and Hewitt, 2003).

2.13. Molecular epidemiology and pathology

Charames and Bapat (2003) have extensively evaluated the genomic instability of microsatellites in the field of carcinogenesis, where chromosomal rearrangements occur. Carcinogenic rearrangements are often associated with loss of heterozygosity in microsatellites located within the affected chromosomal region (Presneau *et al.*, 2003). Microsatellite markers are usually selectively neutral in nature and often represent non- functional sequences. However, SSR markers showing linkage and association with disease have strong linkage disequilibrium with other functional genetic variations which causes the pathological phenotype (Schork *et al.*, 2000). Fornoni *et al.* (2002) noted that genetic markers for disease can be represented by microsatellite markers having direct functional significance (Wang *et al.*, 2002).

Microsatellite-based screening strategies are used in the molecular studies of infectious diseases, which includes mapping and identification of genes responsible for resistance to pathogens and identification of genes controlling drug resistance in pathogens, and identification of genes controlling drug resistance in pathogenic organisms (Behnke *et al.,* 2003; Anderson, 2004). Naidoo and Chetty (1998) reported such microsatellite-based screening approaches in a variety of farmed animals and plants (Yencho *et al.,* 2000).

2.14. Quantitative trait loci mapping

Quantitative trait is one that has measurable phenotypic variation owing to genetic and environmental influences. Quantitative trait locus (QTL) is a genetic locus, the alleles of which affect phenotypic variation (Mackay, 2001). SSR loci are very helpful in coarse and fine linkage mapping approaches, due to the genome-wide distribution and high levels of allelic polymorphism (Glazier *et al.,* 2002). Lander and Kruglyak (1995) defined threshold values to assign significance of linkage of a marker to a given trait in a genome-wide screening process. Nadeau *et al.* (2000) applied the QTL mapping technique in rats and mouse, providing rapid confirmation of loci with modest genetic effect using a limited number of animals.

Generation and analysis of knock-out and knock-in genes, detection and evaluation of polymorphic markers within the candidate gene, examination of gene function and other techniques can be used for identification and confirmation of a gene responsible for quantitative trait (Hentschel and Bonventre, 2005; Salvi and Tuberosa, 2005). QTL mapping in natural populations represents a powerful tool to study the genetic architecture of fitness traits and reproductive isolation. The search for QTLs in wild populations requires the design of appropriate crosses to create a suitable mapping population (Slate, 2005).

2.15. Marker-assisted selection

Marker-assisted selection is based on the concept that the presence of a gene can be inferred by the presence of a marker tightly linked to the gene. Thus, it is important to have high density resolution genetic maps in the vicinity of the target gene or locus that is to be selected. Strategies such as flanking marker analysis and pooled sample mapping are used to find markers tightly linked to the target gene of interest (Churchill *et al.*, 1993; Dixon *et al.*, 1995). In flanking marker analysis, a large segregating population is screened with markers flanking the target interval in order to identify individuals with a crossover within the interval, whereas, in pooled sample mapping, DNA from individuals that share a given phenotype is pooled and DNA from each pool is analyzed with markers flanking the target gene. The inheritance of the gene can be traced, once a tight linkage is established between a molecular marker and the gene of interest, which can be efficiently used in various breeding programmes. The availability of a phenotype, which can be clearly identified and quantified, plays an important role in successful MAS programmes.

Successful implementation of MAS requires well-developed genomic tools, mode of inheritance, interactions with other contributing QTLs and economical magnitude of the QTL studied (Poompuang and Hallerman, 1997). Discovery of genes that control production and reproductive traits greatly benefit MAS in breeding programmes. MAS can be combined with candidate gene analyses that are used to identify important chromosomal regions and individual genes associated with traits of economic interest (Rothschild, 2003). Poompuang and Hallerman (1997) reported that traits difficult to measure, such as disease resistance, low heritability traits, sex-limited traits and traits expressed late in life would be greatly benefited by the use of techniques such as gene technology. Microsatellite markers are useful in early stages of MAS for the primary selection of parents for further crossing and subsequent genetic characterization of progeny individuals.

2.16. In vitro identification of microsatellites

Replication slippages which occur during the in vitro PCR amplification of microsatellite sequences result in the generation of ‘stutter bands’, which become detectable during the electrophoresis of PCR products. Stutter bands represent minor bands that differ in size from the main products by multiples of length of the repeat unit (Hauge and Litt, 1993; Murray *et al.*, 1993). Stutter bands generally appear as products that are slightly shorter than the size of the allele being amplified (Ellegren, 2004).

III. MATERIALS AND METHODS

III. MATERIALS AND METHODS

Experimental protocol

Genomic DNA extraction

↓

Genomic DNA digestion using Sau3AI restriction enzyme

↓

Adaptor preparation

↓

Adaptor ligation

↓

PCR of adaptor ligated DNA

↓

Hybridization of adaptor ligated DNA to biotinylated probes

↓

Capture of microsatellite containing DNA fragments using Streptavidin-coated magnetic beads

↓

PCR amplification following microsatellite capture

↓

Ligation of microsatellite-enriched DNA into plasmid vector

↓

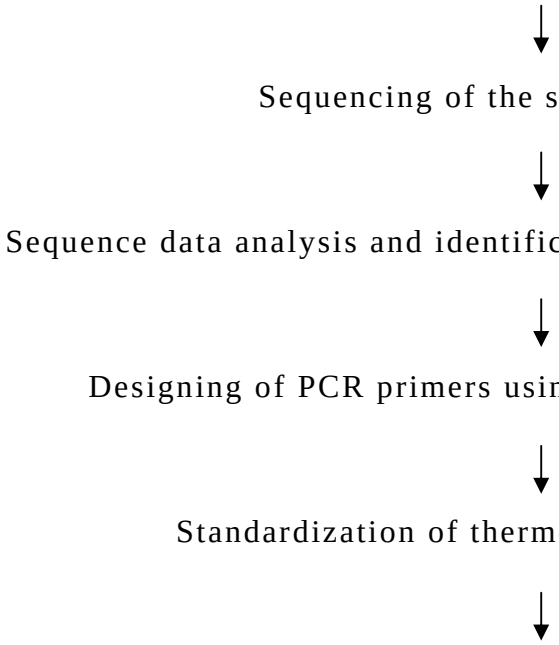
Cloning of the ligated plasmid vector into TOPO (TA) competent cells

↓

PCR screening of positive clones to check for microsatellite repeats

↓

M13 PCR of positive clones to check for the presence of inserts



Flow-chart 1 - Enrichment, cloning and sequencing of genomic DNA fragments containing microsatellite DNA

3.1. Sample collection, processing and storage

The *S. marinoi* cultures were isolated both from Sweden and India. In Sweden, the cultures were isolated from the south-west coast and in India, these cultures were isolated from coastal waters and hatchery samples from places like Goa, Pondicherry and Mangalore by serial dilution method in Guillard's F_2 media (GF_2). The isolated and purified single chains of *S. marinoi* were later transferred to small beakers containing GF_2 media and grown. The cultures were observed daily under the microscope, in order to check both for the density of the blooms and also for flagellate contamination. Later these clones were transferred to Nunc flasks containing 10 ml of sample and grown in the presence of light (12:12, light: dark photoperiod daily). After a day, these Nunc flasks were made up to 50 ml volume using filter-sterilized GF_2 media.

On a daily basis, the Nunc flasks were inverted several times for the purpose of aeration and also for mixing the microalgal cultures, since these cells tend to form brown scum when the density of cells increases. After a

period of 4-5 days, these Nunc flasks which were named and marked individually were checked for contamination and aliquotes were reinoculated into new flasks containing GF2 media for maintenance of algal cultures. Later the algal cells were allowed to sediment to the bottom and the excess supernatant containing media was poured off and the cells were taken into 50 ml centrifuge tubes for cell harvesting purpose. These tubes were centrifuged and the cell pellets were transferred into 1.5 ml microfuge tubes along with 420 µl of distilled water. The algal cells were either stored at -80°C or directly proceeded with extraction of DNA from these cells.

3.1.1. Guillard's F_2 medium (Guillard and Ryther, 1962)

Micronutrients	**(1x)**	**(10x)**
$NaNO_3$	75 mg/l	750 mg
$NaH_2PO_4.H_2O$	5 mg/l	50 mg
$Na_2SiO_3.9H_2O$	30 mg/l	300 mg

Prepared a 10x stock in 100ml distilled water. Added 10 ml to 1000 ml distilled water

3.1.2. Trace elements

$CuSO_4.5H_2O$	9.8 µg/l
$ZnSO_4.7H_2O$	22 µg/l
$CaCl_2.6H_2O$	10 µg/l
$MnCl_2.4H_2O$	18 µg/l
$NaMoO_4.2H_2O$	63 µg/l
Di-sodium EDTA	4.36 mg/l
$FeCl_3.6H_2O$	3.15 mg/l

Prepared a 1000x stock in 100 ml distilled water. Added 0.1ml to 1000 ml sterilized seawater.

3.1.3. Vitamins

Biotin	1 µg
Vitamin B_{12}	1 µg
Thiamine HCl	200 µg

Prepared a 1000x stock in 100ml distilled water. Added 0.1 ml or 0.15ml to 1000 ml of sterilized sea water.

Sea water was autoclaved and cooled. The micronutrients and the trace metals were sterilized separately. Vitamin solution was filter-sterilized. All nutrients, vitamins and trace elements were added to the autoclaved seawater. pH was later adjusted to 7.8-8.2 with hydrochloric acid or sodium hydroxide.

3.2. Genomic DNA extraction

DNA extraction was carried out using two different methods,

a) Phenol-chloroform-isoamyl alcohol method with minor modifications along with a commercial DNA purification kit from Amersham Pharmacia, United Kingdom.

b) Hexadecyl – trimethyl ammonium bromide (CTAB) plus a commercial DNA purification kit from Qiagen, Germany.

3.2.1. Reagents used for DNA extraction

3.2.1.1. SDS (10%)

Sodium do-decyl sulphate (SDS)	10 g
Milli Q water	100 ml

3.2.1.2. Stock proteinase K solution (Bangalore Genei, Bangalore)

Five milli liter distilled water was added to 100 mg of proteinase K powder. The stock prepared would contain 20 mg/ml concentration and stored in -20°C.

3.2.1.3. Lysis buffer

10 mM Tris HCl pH 8.0
1 mM EDTA pH 8.0
0.5% SDS
100 µg/ml proteinase K pH 7.0

3.2.1.4. Ammonium Acetate (10 M)

Ammonium acetate	770.8 g
Distilled water	1000 ml

The reagent was filter-sterilized using 0.45 µ filter and stored at room temperature.

3.2.1.5. Phenol equilibration (Sambrook *et al.*, 1989)

To melted phenol, equal volume of buffer was added (0.5 M Tris HCl, pH 8.0) and stirred using a magnetic stirrer for 15 minutes at room temperature. The upper aqueous phase was discarded and an equal volume of 0.1 M Tris HCl (pH 8.0) was added to phenol and stirred for further 15 min. The process was repeated until a pH of 7.8 was obtained. Hydroxyquinoline was also added to a final concentration of 0.1%. This equilibrated phenol was used for DNA extraction.

3.2.1.6. TE buffer

Tris HCl	10 mM
EDTA	1 mM
pH	8.0

3.3. Phenol chloroform isoamyl alcohol method

Genomic DNA was extracted and purified using phenol-chloroform-isoamyl alcohol method with minor modifications. The *S. marinoi* culture pellets were resuspended in 420 µl of sterile Milli Q water and subsequently 20 µl of 250 mM EDTA (pH 8.0), 50 µl of 10% SDS, 5 µl of 1 M Tris HCl (pH 7.5), 5 µl of 1M NaCl and 4 µl of RNaseA were added to it. This mixture was incubated at 50°C for one hour (Godhe *et al.*, 2001). DNA was extracted using buffered phenol:chloroform:isoamyl alcohol (24:24:1). The mixture was vortexed for one minute and thereafter, centrifuged at 15000 g for 3minutes. The top layer, which contained DNA suspended in water was transferred to a new microfuge tube. The bottom layer which contained

phenol-chloroform solution, proteins and cell debris was discarded. This procedure was repeated twice with phenol-chloroform-isoamyl alcohol solution and a third time with chloroform-isoamyl alcohol solution (24:1). The last step was performed in order to minimize phenol contamination which might inhibit the PCR reaction.

Thereafter, the DNA was precipitated with two volumes of absolute ethanol and 1/10th volume of 3 M sodium acetate (pH 4.5) at -20°C overnight. The samples were later centrifuged at 15000g for 30min to pellet the precipitated DNA. The supernatant was discarded and the pellet was washed with ice cold 70% ethanol and centrifuged again at 15000 g for 10 minutes. After washing the ethanol was evaporated and the genomic DNA was resuspended either in Milli Q water or 1x TE buffer (pH 8.0). In cases where the phenol method didnot yield enough purity, the extract was further purified using flexi-prep kit (Amersham Pharmacia, Germany) following the instructions of the manufacturers.

3.4. Cetyl Tri-Methyl Ammonium Bromide (CTAB) method

To the pelleted *S. marinoi* cultures, 300 µl of lysis buffer was added along with 30 µl of 10% SDS and incubated at 60°C for one hour duration. Thereafter 50 µl of 5 M NaCl, 40 µl of 10% CTAB diluted in 0.7% NaCl and 4 µl of RNase solution was added and incubated at 65°C for 10 minutes. The extraction of DNA from the lysed cells and subsequent purification of DNA were done using the DNeasy plant mini kit (Qiagen Gmbh, Germany) starting at Step 5 in the manufacturer's instructions.

3.5. Estimation of DNA concentration and purity

The concentration and purity of DNA in samples were estimated by spectrophotometric method as detailed by Sambrook *et al.* (1989). Absorbance of samples in distilled water was measured at 260 nm and 280 nm using Shimadzu UV-1601 spectrophotometer. DNA concentration and purity were calculated according to the standard equation.

$$\text{Quantity of DNA } (\mu g/ml) = \text{O.D. at 260 nm} \times 50 \times \frac{\text{Volume of distilled water used } (\mu l)}{\text{Volume of sample used } (\mu l)}$$

$$\text{Purity of DNA} = \frac{\text{Optical density at 260 nm}}{\text{Optical density at 280 nm}}$$

In addition, the genomic DNA samples were checked by running an agarose gel electrophoresis using 1% agarose gel.

3.5.1. Materials used for agarose gel electrophoresis

3.5.2. Agarose (Hi Media, Mumbai)

3.5.2.1. TAE Buffer (50x)

Tris base	242 g
Glacial acetic acid	57.1 ml
EDTA (0.5 M, pH 8.0)	100 ml

The solution was made up to a final volume of 1liter using Milli Q water and sterilized by autoclaving at 121°C for 15 minutes. A 1x TAE buffer for electrophoresis was prepared by diluting 20 ml of stock solution to 1 liter using Milli Q water.

3.5.2.2. Gel loading buffer

Bromophenol blue	0.25 g
Sucrose	40 g
Distilled water	100 ml

3.5.2.3. Ethidium bromide

A stock solution of ethidium bromide was prepared by dissolving 5 mg of ethidium bromide (Sigma Co., USA) in 1 ml of distilled water. The solution was either added to agarose during preparation or to the staining medium (1xTAE buffer) to yield a final concentration of 0.5 µg/ml. In the latter case, the gel was allowed to stain in the staining solution for 20 minutes and destained in distilled water for 10 minutes to remove excess stain.

3.5.2.4. Molecular weight markers

3.5.2.4.1. λ - DNA molecular weight markers

λ DNA double digested with *Eco*Rl and *Hind*III yielding 13 fragments of 21226, 5148, 4973, 4277, 3530, 2027, 1904, 1584, 1330, 983, 831, 564 and 125 base pairs .

3.5.2.4.2. 100 bp DNA Ladder plus (Gene Ruler^{TM})

This marker has size ranges 3000, 2000, 1500, 1200, 1031, 900, 800, 700, 600, 500, 400, 300, 200 and 100 base pairs.

3.5.2.4.3. 1kb DNA Ladder (Gene Ruler^{TM})

This marker has size fragments from 10000, 8000, 6000, 5000, 4000, 3500, 3000, 2500, 2000, 1500, 1000, 750, 500 and 250 base pairs.

3.5.2.4.4. Agarose gel electrophoresis

The agarose gels were prepared in 1x TAE buffer with appropriate percentage (w/v) of agarose and buffer. When the molten agarose had cooled to below 65^{o}C, ethidium bromide was added to a final concentration of 0.5 µg/ml and later the gels were cast. 10 to 15 µl of DNA was mixed with 4 µl of 6x gel-loading buffer and loaded onto the wells. 100 bp DNA ladder plus or 1 kb DNA ladder (Bangalore Genei, Bangalore) was used as a molecular weight marker. Electrophoresis was carried out at 100-120 Volts and the bands were visualized using Gel documentation system (Herolab, Germany).

3.6. Genomic DNA digestion with restriction enzyme

The genomic DNA was digested using Sau3AI restriction enzyme, Sau3AI is a 4 base cutter producing sticky end products and cuts usually once every 256 base pairs. Approximately 0.7 units of Sau3AI (Bangalore Genei, Bangalore) were used for the digestion of 1µg of genomic DNA. The reagents

used for restriction digestion are shown in Table 1. The restriction enzyme is generally added last and the genomic DNA is added last but one in the reaction mixture, in order to prevent any non-specific enzymatic activity from occurring. The digestion mixture was incubated at 37°C for two hours and later heated to 65°C for 10 minutes, in order to inactivate the enzyme. All these steps were performed in a thermocycler.

Table 1. Reagents used for restriction enzyme digestion

Materials used	**Volume (µl)**
Genomic DNA	37.0
Restriction enzyme buffer E	9.0
BSA	0.9
Restriction enzyme	2.1
Milli Q water	41.0
Total volume	90.0

3.6.1. Adaptor preparation

Ten micro liter each of oligoA (5'GGC CAG AGA CCC CAA GCT TCG 3') and oligoB, which were phosphorlyated at the 5' end (5'-PO_4-GAT CCG AAG CTT GGG GTC TCT GGC C-3') with a concentration of 200 pM/µl respectively, were mixed together by vortexing and subsequently spinned briefly. The oligomix was denatured by heating at 80°C for 5 minutes and renatured slowly at room temperature for one hour. 60 µl of Milli Q water was added to the mix to give a final concentration of 25 pM/µl of each oligo. The oligomix prepared was stored at -20°C. The adaptor sequence was as follows,

5'- GGCCAGAGACCCCAAGCTTCG - 3'
3'- CCGGTCTCTGGGGTTCGAAGCCTAG - 5'

3.6.1.1. Adaptor ligation

The reagents that were used in adaptor ligation were mixed gently by tapping on to the sides of the microfuge tubes. The ligation mix was

incubated at 4°C overnight. After the incubation process, the enzymes were inactivated by heating at 65°C for 10 minutes in a thermocycler. The products were cleaned using phenol-chloroform-isoamyl alcohol mixture. The DNA was precipitated using two volumes of absolute ethanol and 1/10th volume of 10 M ammonium acetate at -20°C overnight. The DNA was resuspended in 10 µl of 0.1x TE buffer. At the end of this ligation reaction a 21 nucleotide length adapter containing a PCR primer sequence was ligated to the sticky ends of the restriction-digested genomic DNA fragments. The reagents that were used in adaptor ligation reaction are shown in Table 2.

Table 2. Reagents used for adaptor ligation

Materials used	**Volume (µl)**
Restriction digested mix	90.0
T4 DNA ligase buffer	20.0
T4 DNA ligase	40.0
Adapter mix	2.0
Milli Q water	48.0
Total volume	200.0

3.7. PCR of adaptor-ligated DNA

The adaptor-ligated DNA fragments were subjected to PCR amplification by using one of the oligo (oligoA) in the adaptor sequence mix as the primer sequence. The reagents used for PCR amplification are shown in Table 3. The PCR reagents included 10x buffer (100 mM HCl pH 8.3, 20 mM $MgCl_2$, 500 mM KCl, 1% gelatin), 200 µM concentrations of each dNTPs (dATP, dTTP, dGTP, dCTP) and three units of Taq polymerase (Bangalore Genei, Bangalore). The thermocycling conditions are shown in Table 4. All PCR reactions were performed using DNA DYAD engine thermocycler (MJ Research, USA).

Table 3. Reagents used in PCR of adaptor-ligated DNA

Materials used	**Volume (μl)**
OligoA	1.25
Digested adaptor ligated DNA	1.0
Buffer	2.5
dNTPs	2.0
Taq DNA polymerase	0.25
Milli Q water	18.0
Total volume	25.0

Table 4. Thermocycling conditions for adaptor-ligated DNA

Steps	**Temperature (°C)**	**Time**
1. Intial denaturation	95	5 min
2. Denaturation	95	50 sec
3. Primer annealing	56	1min
4. Primer extension	72	2min
5. Go to step 2 for 30 cycles		
6. Final extension	72	10min

3.7.1. Purification of PCR products

The PCR products were purified to remove excess of dNTPs, enzymes and primers using QIAquick PCR purification kit (Qiagen, USA).

3.7.1.1. Reagents supplied with PCR purification kit

Buffer PB, Buffer PE and Buffer EB

3.7.1.2. Procedure

One hundred microliter of PCR product was mixed with 500 µl of buffer PB and applied on to the QIAquick spin column, and centrifuged to bind the DNA to the column. The column containing the DNA was washed using 750 µl of buffer PE and the bound DNA was subsequently eluted using 30 µl of buffer EB Fig.1.

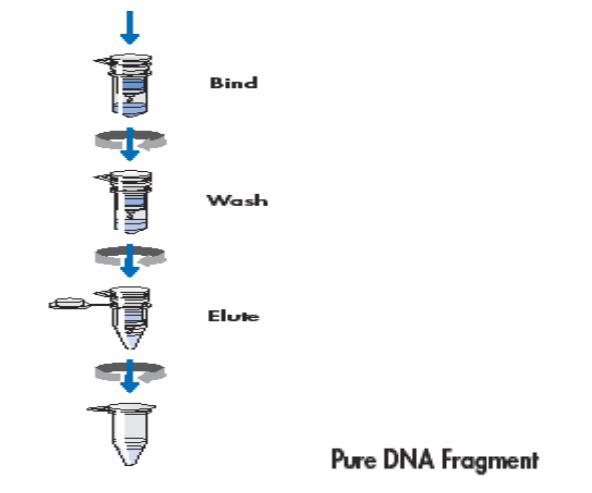


Fig.1. Purification of PCR products using QIAquick spin column

3. 8. Hybridization of adaptor-ligated DNA to biotinylated probes

To identify and enrich PCR fragments containing microsatellite repeats, the PCR products were hybridized to biotinylated $(CA)_{15}$ or $(GA)_{15}$ probes (Bioserve Biotechnologies Pvt. Ltd., Hyderabad). The reagents used for hybridization along with microsatellite $(CA)_{15}$ or $(GA)_{15}$ probes are listed in tables 5 and 6. A 1:1 molar ratio between oligos and PCR products were generally used. The reagents were mixed thoroughly and heated at 95°C for 10 minutes and subsequently, incubated at 55°C overnight for $(CA)_{15}$ and 50°C overnight for $(GA)_{15}$ biotinylated microsatellite probes, respectively. All these steps were performed in DNA engine DYAD thermocycler (MJ Research, USA).

Table 5. Hybridization of ligated DNA to biotinylated $(CA)_{15}$ probe

Materials used	**Volume (µl)**
PCR product	1.0
Oligo (CA)15	2.0
12x SSC, 0.1% SDS	50.0
Milli Q water	47.0
Total volume	100.0

Table 6. Hybridization of ligated DNA to biotinylated $(GA)_{15}$ probe

Materials used	**Volume (µl)**
PCR product	1.0
Oligo (GA)15	2.0
12xSSC, 0.1%SSC	50.0
Milli Q water	47.0
Total volume	100.0

3.8.1. Capture of microsatellite containing DNA fragments using streptavidin coated magnetic beads

3. 8.1.1. Standard saline citrate solution (SSC, 20x)

Sodium chloride	175.4 g
Sodium citrate	88.2 g
Distilled water	1000 ml
pH	7.0

The ingredients were dissolved in 800ml distilled water and pH adjusted to 7.0 with 10 N sodium hydroxide. The volume was made up to 1000 ml and stored at room temperature. The stock solution of 20x SSC was used for preparing 1xSSC and 2xSSC.

3.8.1.2. S1 Reagent

Sodium do-decyl sulphate	1.0 g
1xSSC	1000 ml

3. 8.1.3. S2 Reagent

Sodium do-decyl sulphate	1.0 g
1xSSC	1000 ml

3. 8.1.4. Maleic acid buffer

Maleic acid	11.6 g
Sodium chloride	8.7 g
Distilledwater	1000 ml
pH	7.5

3. 8.1.5. Washing buffer

Washing buffer was prepared by taking 3000 µl of Tween 20 in a 100 ml volumetric flask and made up to the volume using maleic acid buffer. The pH was adjusted to 7.5.

3.8.1.6. Capture of microsatellite-containing DNA fragments

After the hybridization step, the genomic DNA fragments complementary to the biotinylated microsatellite motifs were recovered by the method of streptavidin-coated magnetic bead capture: 100 µl of streptavidin-coated magnetic beads called dynabeads (10 mg/ml M-280 Qiagen, USA) was mixed thoroughly with 100 µl of 1x washing buffer in a microfuge tube. The supernatant was carefully removed without touching the beads by placing the tube in a magnetic stand for 2-3 minutes. This step was repeated twice. Then the magnetic beads were mixed again with 200 µl of 2x washing buffer and resuspended. The hybridization solution was added to the same tube and the reagents used in the reaction mixture were as mentioned earlier.

The reaction mixture was incubated at 42°C for two hours with regular mixing at every 10 minutes interval. After the incubation step, the supernatant was removed and discarded by placing the tube in a magnetic stand. To the pellets, 200 µl of S1 reagent (2xSSC, 0.1% SDS) was added and mixed thoroughly by gentle pipetting. The mixture was incubated at room temperature for five minutes duration and later, the supernatant was removed by placing the tube on a magnetic stand for 1-2 minutes. This step was also carried out twice, and subsequently 200 µl of S2 reagent (1xSSC, 0.1%SDS) was added to the pellets and mixed well by gentle pipetting. The mixture was kept at 45°C for 5 minutes and the supernatant was removed by placing the tube in a magnetic stand for 1-2 minutes. This step was also repeated twice. Finally, the microsatellite enriched DNA present along with pellet was eluted using 120 µl of 1xTE buffer in two steps of 60 µl each by placing the microfuge tube at 95°C for 10 minutes. The supernatant containing the DNA was transferred to a fresh tube by placing the microfuge tube on to a magnetic stand for 1min. The microsatellite enriched DNA sequences collected were stored at 4°C for further analysis.

3.9. PCR following microsatellite capture

The microsatellite-enriched DNA sequence product recovered by magnetic bead capture was amplified by PCR using one of the oligo (OligoA) as the primer sequence in the PCR reaction. The oligoA primer sequence is complementary to the adaptor sequence ligated to the genomic DNA. The reagents used in PCR reaction are shown in Table 7 and the thermocycling conditions for PCR are shown in Table 8 respectively.

Table 7. Reagents for PCR following microsatellite capture

Reagents used	Volume (μl)
OligoA	1.5
Microsatellite-captured DNA	10.0
Buffer	5.0
dNTPs	4.0
Taq DNA polymerase	0.5
Milli Q water	29.0
Total volume	50.0

Table 8. Thermocycling conditions for microsatellite PCR

Steps	Temperature(°C)	Time
1. Initial denaturation	95	3 min
2. Denaturation	95	30 sec
3. Primer annealing	60	30 sec
4. primer extension	72	30 sec
5. Go to step 2 for 5 times		
6. Denaturation	92	30 sec
7. Primer annealing	60	30 sec
8. Primer extension	72	55 sec
9. Go to step 6 for 30 times		
10. Final extension	72	30 min

3.10. Ligation of enriched DNA into plasmid vector

The reagents used in the ligation reaction were mixed gently by tapping on to the sides of the tube. Then incubated at room temperature for 5 minutes and subsequently proceeded with transformation process. The reagents used for ligation are indicated in Table 9. The amplified DNA fragments, which possess the A-overhang site due to the terminal transferase activity of Taq DNA polymerase are ligated to the T-overhang site of the plasmid vector pCR2.1 TOPO10 (Invitrogen, USA).

Table 9. Reagents for ligation of microsatellite DNA into plasmid vector

Materials used	Volume (µl)
Fresh cleaned PCR product	4.0
Salt solution	1.0
Pcr2.1 TOPO vector	1.0
Total reaction volume	6.0

3.11. Transformation

3.11.1. Reagents used for transformation process

3.11.2. Composition of SOC media

Tryptone	2.0 g
Yeast extract	0.5 g
1 M NaCl	1.0 ml
1 M KCl	0.25 ml
2 M $MgCl_2$	1.0 ml
2 M Glucose	1.0 ml

3.11.3. X-gal

Dissolved 40 mg 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal) into 1 ml N, N-di methyl-formamide. Covered with aluminum foil and stored at -20°C. The final concentration of X-gal was 40 mg/ml.

3.11.4. Luria Bertani medium (Hi Media, Mumbai)

Composition	Quantity
Casein enzymatic hydrolysate	10.0 g
Yeast extract	5.0 g
Sodium chloride	10.0 g
Agar agar	20.0 g
Distilled water	1000 ml
pH	7.5 ± 0.2

This medium was autoclaved at 121°C for 15min at 1.055 kg/cm^2 pressure and cooled at 50°C. Antibiotic solution was added to a desired concentration and poured into pre-sterilized Petri plates. The antibiotic

containing LB agar plates were pre-warmed in an incubator at 37°C until use or stored in the refrigerator for not more than a week.

3.11.5. Transformation protocol

The ligated plasmids were transformed into competent *Escherichia coli* cells (Invitrogen, USA). Initially, the competent cells were thawed on ice and mixed gently with 2 µl of ligation mixture. The mixture was incubated on ice for five minutes and later, the cells were subjected to heat shock by transferring the cells into a water bath set at 42°C for 30 seconds. Immediately the cells were transferred on to ice, 250 µl of room temperature SOC medium was added and incubated at 37°C for one hour with agitation using a shaker incubator. Subsequently, aliquots of 50, 100 and 150 µl were spread on to LB agar plates containing ampicillin (100 µg/ml) and X-gal (40 mg/ml) and incubated at 37°C overnight in an incubator.

3.11.6. Screening of microsatellite positive clones

The microsatellite-containing clones were screened by the method of Lac selection using β-galactosidase activity. Screening for β-galactosidase presence or absence involves a lactose analogue called X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside), which is inturn broken down by β-galactosidase activity to a product that is colored ‘blue’. The positive recombinant cells were identified by their ability to synthesize β-galactosidase enzyme, thus they produce ‘white’ colonies. After the overnight growth of the transformed cells on LB agar plates, the plates were transferred to 4°C for one hour, so that false positive colonies change color rapidly from white to blue at lower temperature. Positive clones which were white in color were picked up and transferred on to fresh LB-ampicillin-X-gal plates with numbered grids for further screening in the same method as described above. Finally, only pure white clones were selected from the plates and screened using screening PCR technique.

3.12. Preparation of cell lysates from bacterial isolates for screening PCR

The test cultures were grown at 37°C overnight in 3 ml Luria-Bertani broth (Hi Media, Mumbai) containing ampicillin (100 µg/ml). 50 µl of this culture was suspended in 450 µl of 1xTE buffer (pH 8.0) and lysed at 95°C for 10-15 minutes in a hot bath and further they were snap cooled on ice for five minutes. The cell debris was settled by centrifugation at 10000 g for five minutes and only the clear supernatant was used as the DNA template in PCR amplification.

3.12.1. Screening PCR

The reagents used in screening PCR are shown in Table 10. Here, two individual primers were used in the PCR amplification, one was the un-biotinylated respective microsatellite probe and the other one was oligoA from the adaptor sequence. The thermocycling conditions for screening PCR are listed in Table 11.

Table 10. Reagents used for microsatellite screening PCR

Reagents used	**Volume (µl)**
OligoA	0.25
Un-biotinylated oligo (CA_{15} or AT_{15})	0.5
Buffer	2.0
dNTPs	1.6
Taq DNA polymerase	0.2
Template DNA	10.0
Milli Q water	5.45
Total volume	20.0

Table 11. Thermocycling conditions for screening PCR

Steps	Temperature(°C)	Time
1. Initial denaturation	95	3 min
2. Denaturation	95	30 sec
3. Primer annealing	60	30 sec
4. Primer extension	72	30 sec
5. Go to step 2 for 5 times		
6. Denaturation	92	30 sec
7. Primer annealing	60	30 sec
8. Primer extension	72	55 sec
9. Go to step 6 for 30 times		
10. Final extension	72	30 min

3.13. Long-term storage library construction

The microsatellite-positive clones from screening PCR were grown in LB-ampicillin media at 37°C overnight: 900 µl of each positive clone was mixed with 900 µl of pre-sterilized 30% glycerol broth in 2 ml storage vials. These storage vials were marked individually with respective clone numbers and immediately transferred to -80°C for long term storage. The composition of glycerol broth is as follows;

3.13.1. Glycerol broth

Composition	Quantity
Glycerol	30 ml
LB broth	70 ml
Ampicillin	100 µg/ml
pH	7.0

The medium was sterilized by autoclaving at 110°C for 15 minutes.

3.14. M13 PCR

M13 PCR using M13 universal primers was carried out in order to ascertain the insert sizes of the screening PCR-positive clones. The reagents used for performing M13 PCR are listed in Table 12 and the thermocycling

conditions for PCR are indicated in Table 13. Analysis of the amplified product was carried out by running gel electrophoresis using 2% agarose gel. The positive clones showing insert size in the range of 200-1000 base pair was selected for sequencing process.

Table 12. Reagents used for M13 PCR amplification

Materials used	**Volume (µl)**
M13 forward primer	2.0
M13 reverse primer	2.0
10x assay buffer	3.0
dNTPs	2.4
Taq DNA polymerase	0.3
Plasmid DNA	0.5
Milli Q water	19.8
Total volume	30.0

Table 13. Thermocycling conditions for M13 PCR

Steps	**Temperature (°C)**	**Time**
1. Initial denaturation	94	10 min
2. Denaturation	94	10 min
3. Primer annealing	55	1 min
4. Primer extension	72	1 min
5. Go to step 2 for 30 times		
6. Final extension	72	10 min

3.14.1. Primer used for M13 PCR

M13 forward primer sequence

5’ CTGGCCGTCGTTTTAC 3’

M13 reverse primer sequence

5’ CAGGAAACAGCTATGAC 3’

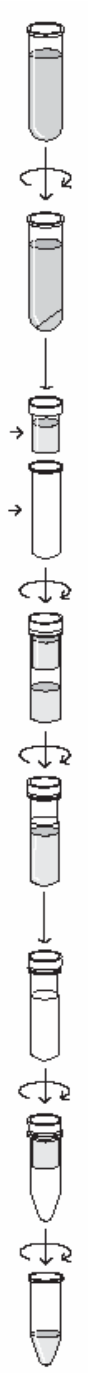
3.15. Plasmid DNA extraction

3.15.1. Reagents supplied with kit (Eppendorf Fast Plasmid Mini Kit)

Lysis solution, RNase solution, Lysozyme, Wash buffer concentrate, Elution buffer and Fast Plasmid Spin Column Assembly.

3.15.2. Plasmid extraction protocol

Young bacterial cultures at the logarithmic growth phase were used for plasmid DNA extraction. 1.5 ml of bacterial culture was centrifuged to sediment the cells to the bottom of the microfuge tube. The clear supernatant was discarded and to the pellet, 400 µl of ice-cold complete lysis solution was added. The pellet was mixed thoroughly and incubated at room temperature for three minutes. Subsequently, the lysate was transferred to the spin column assembly and centrifuged at the maximum speed for 60 seconds. 400 µl of diluted wash buffer was added to the spin column and centrifuged for one minute. Then 50 µl of elution buffer was added to the spin column containing a collection tube at the bottom and briefly centrifuged. Later, the purified plasmid collected in the collection tube was removed and stored.



1.5 ml Bacterial Culture

Centrifuge for 1 minute to pellet the cells

Decant the media
Add 400 µl of ice-cold complete lysis solution
Vortex for a full 30 seconds at the highest setting

Spin Column

Incubate at room temperature for 3 minutes

Transfer the lysate to a Spin Column

Centrifuge for 30–60 seconds

Wash DNA with 400 µl of diluted wash

Centrifuge for 30-60

Decant the filtrate from the waste tube and re-assemble

Centrifuge for 1 minute

Add 50 µl of Elution Buffer to the center of the Spin Column

Centrifuge for 30-60 seconds

Purified Plasmid

Fig. 2. Plasmid purification using Eppendorf Fast Plasmid Mini Kit

3.16. Sequencing

Double-stranded sequencing was performed by using an automated ABI 3100 genetic analyzer that uses fluorescent label dye-terminator (Sanger's sequencing method). Universal M13 forward and M13 reverse primers were used for sequencing the plasmid vector and insert regions. The sequences of M13 forward and reverse primers are given below.

M13 Forward primer sequence
5' CTGGCCGTCGTTTTAC 3'

M13 Reverse primer sequence
5' CAGGAAACAGCTATGAC 3'

3.17. Design of oligonucleotide primers

The sequences of the flanking region on either side of the microsatellite repeat regions obtained from sequencing data were used to design oligonucleotide primers using bioinformatics software program PRIMER3, from the website http://www-genome.wi.edu/cgi-bin/primer/primer3-www.cgi. Several criteria involved in primer designing were taken into account, so that the designed primer sets have approximately equivalent annealing temperatures, a guanine or cytosine nucleotide at the 3'end and no primer dimmer formation. The primers designed were synthesized at (Bioserve Biotechnologies Pvt. Ltd., Hyderabad).

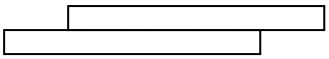
3.18. Standardization of primers

Suitability of the primers designed were checked using cell lysates of the clones that contain the insert sequences, which were used for designing the primers. Gradient PCR was performed in order to ascertain the accurate annealing temperature of the primers, so that the PCR product contains the desired microsatellite region.

1. Genomic DNA Extraction → Check concentration → EtOH Precipitation

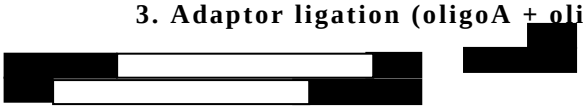
↓

2. Genomic DNA digestion with restriction enzyme (*Sau3AI*)



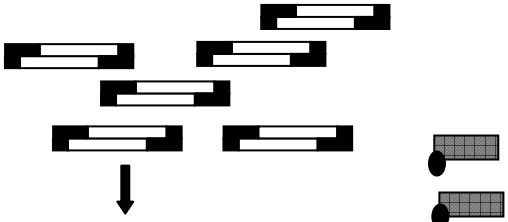
↓

3. Adaptor ligation (oligoA + oligoB)



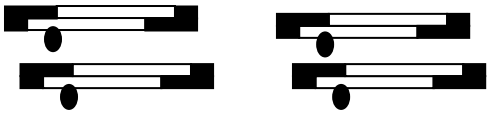
4a. Test – PCR of adaptor-ligated DNA
Check in agarose gel →

4b. PCR of adaptor – ligated DNA



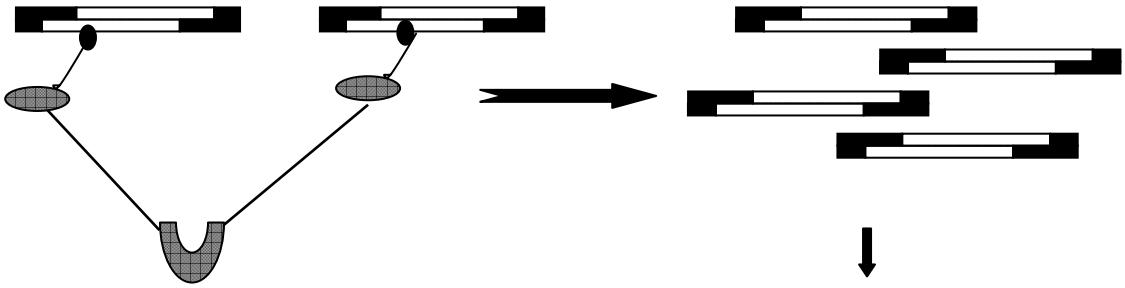
↓

5. Hybridisation of adaptor – ligated DNA to biotinylated probes $(GA)_{15}$, $(CA)_{15}$



↓

6. Capture of microsatellite – containing DNA fragments with streptavidin – magnetic beads

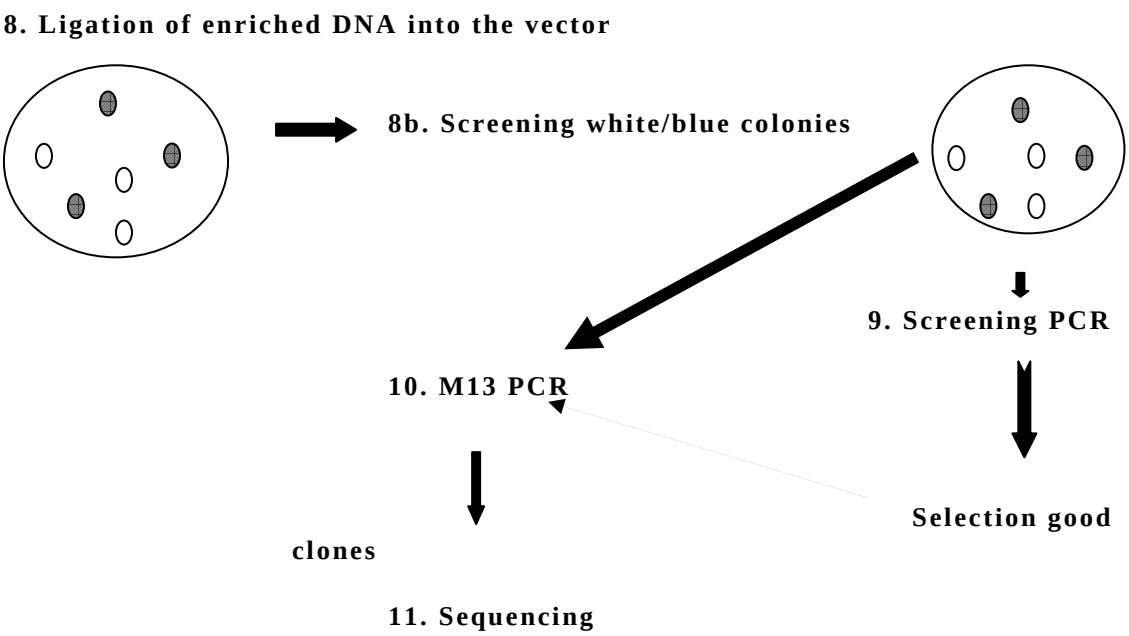


Microsatellite enriched DNA sequences

7. PCR Check in agarose gel (a smear indicate a successful capture of microsatellites fragments)

→ **Clean up**

↓



Flow chart 2 – Schematic representation of outline of work

3.19. Large subunit (LSU), Internal transcriber spacer (ITS) and RAPD analysis

3.19.1. Sample collection and isolation of clones

Sediment cores (Table 14) were collected to obtain undisturbed sediments. The top centimeter of sediment was sampled and kept in dark and cool (4°C) conditions until processed. The sediment sample from a particular site was collected on one occasion only, but the isolation of cells for clonal cultures was not necessarily done at the same time**.** Approximately 1 g of wet sediment was added to 45 ml of f_2 medium (Guillard, 1975) in 50- ml Nunc flasks. The flasks were incubated at 10°C with a 12:12 h light: dark cycle at an irradiance of 100 µmol/cm/s. The flasks were monitored daily with a Zeiss Axiovert 135 inverted microscope and after several days of incubation, when newly formed vegetative cells were first observed, chains of *S. marinoi* were isolated by micropipette. This procedure was followed in order to minimize the risk that two different isolated chains were clonal (i.e., originated from the same resting stage). Individual chains of *S. marinoi* were transferred to a drop of sterile f_2 medium. This was repeated several times to clean the cells. Each clone was then transferred to a small petri dish with f_2 medium and incubated as described above. The clones that grew in the petri dishes were transferred to 50-ml Nunc flasks with f_2 medium and reinoculated at least every third week.

3.19.1. DNA extraction

DNA was extracted from all 53 clones (Table 14). Clonal cultures were kept for a maximum of six months before DNA was extracted. Prior to DNA extraction, cultures in exponential phase were scanned microscopically to ensure that no contaminants were present. The cultures were transferred to 50-ml centrifuge tubes and pelleted by mild centrifugation (7 min at 5000 *g*). Pellets were resuspended in 420 µl Milli-Q water and transferred to Eppendorf tubes. Pellets were processed immediately or frozen at –80° C. Cells were lysed and DNA extracted as described in Godhe *et al.* (2001). Concentration and purity of DNA were measured with a spectrophotometer

(Pharmacia Biotech GeneQuant II, Buckinghamshire, UK). If sample purity was low, i.e., 260/280 ratio < 1.3, the extract was further purified with a Flexi Prep Kit (Amersham Pharmacia, UK) following the manufacturer's instructions. Extracted DNA was stored at -20˚ C.

3.19.3. PCR, Sequencing and RAPD

LSU rDNA domains D1–D3 and the ITS region (ITS-1, 5.8S and ITS-2) were amplified by PCR before sequencing. To eliminate potential Taq errors associated with the original PCR, each region was amplified several times. Reactions were run in total volumes of 50 µl consisting of approximately 100ng template DNA, Taq buffer containing 1.5 mM $MgCl_2$, 0.5 µM of each primer, 200 µM of each deoxynucleotide triphosphate (dNTP), 2.5 U of Taq polymerase (QIAGEN, Valencia, CA, USA), and sterile Milli-Q water. LSU rDNA sequences were amplified by using the primers DIR and D3Ca (Lenaers *et al.*, 1989; Scholin *et al.*, 1994), and the ITS regions were amplified using the primers ITS 5 (White *et al.*, 1990), 1380-Fmod (Lundholm *et al.*, 2003), SSU-R-F (Medlin *et al.*, 1988), ITS055-R (Marin *et al.*, 1998) and DIR-R (Scholin *et al.*, 1994). Amplifications were carried out in a thermal cycler (Perkin Elmer GeneAmp PCR System 2400, Wellesley, MA, USA) as follows: initially 94˚C for five minutes., followed by 30 cycles of denaturation at 94˚C for one minute., annealing between 50 and 55˚C for one minute., extension at 72˚C for one minute. After the cycles, extension was completed at 72˚C for five minutes. The PCR product was loaded on to a 0.8% agarose gel in 1xTBE and ethidium bromide stained gels were studied under UV transillumination. PCR products were purified using the QIAquick PCR Purification Kit (QIAGEN) following the manufacturer's instructions. Final concentration of the PCR product was measured using a spectrophotometer.

Each 10-µl sequencing reaction contained approximately 50ng of PCR product together with the primers. LSU rDNA (D1-D3) fragments were sequenced by using the primers DIR, D2C, D2CF (Scholin *et al.*, 1994) and D3Ca (Lenaers *et al.*, 1989). ITS fragments were sequenced using the primers ITS 1, ITS 3 (White *et al.*, 1990) and DIR-R (Scholin *et al.*, 1994). LSU

rDNA and ITS nucleotide sequences were determined using the CEQ Dye Terminator Cycle Sequencing Kit following the manufacturer's instructions. Reaction products were run on a CEQ 8000 Genetic Analysis System (Beckman Coulter, Fullerton, CA, USA). Amplified DNA from each clone was sequenced, in forward and reverse directions, at least five times. Sequences are deposited in GenBank with accession numbers given in Table 14.

For RAPD analysis, 10 different 10-mer oligonucleotide primers (60–80% GC) (Kresovich *et al.*, 1992; Tassanakajon *et al.*, 1997) were screened in order to identify those best able to produce clear and repeatable banding patterns of the *S. marinoi* genome. Two primers (CRA25 and CRA26) that produced banding patterns (8–20 bands) in all clones were selected (Kresovich *et al.*, 1992). The RAPD-PCR was run in a total volume of 30 µl consisting of 180 to 700 ng of template DNA, Taq buffer containing 1.5 mM $MgCl_2$, 10 µM primer, 200 µM of each dNTP, 1.75U of Taq polymerase (Bangalore Genei, Bangalore, India) and sterile Milli-Q water. Amplifications were carried out in a thermal cycler (PTC-100 Programmable Thermal Controller, MJ Research, Waltham, MA, USA) as follows: initially 94˚C for five minutes., followed by 34 cycles of denaturation at 94˚C for five seconds., annealing at 36˚C for 45 seconds., extension at 72˚C for 1.5 minutes. After the cycles, extension was completed at 72˚C for 5 minutes. The PCR product was loaded on to a 1.6% agarose gel in 1xTBE. On each gel, three lanes (left, centre and right) were loaded with molecular marker (Gene Ruler 100 bp DNA ladder Plus); thus, normalized banding patterns from different gels could be compared by computerized image analysis. Ethidium bromide stained gels were studied under UV transillumination, photographed and the images were transferred to the software program GelCompar II, version 2.5 (Applied Maths, BVBA, Sint-Martems-Latem, Belgium). The reproducibility was verified by replicating each reaction two or, if needed, three times using separate DNA extracts, PCR amplifications and electrophoresis. Using the automatic band search function (min profiling

5%, position tolerance 1%), reproducibility was confirmed when densiometric curves with >85% homology were produced.

3.19.4. Scanning electron microscopy (SEM)

Ten milliliter of culture was fixed in 2% glutaraldehyde. Organic material was removed by oxidation with 2 ml 30% H_2SO_4 and 10 ml saturated aqueous solution of $KMnO_4$. The samples were cleared after 24 hours by addition of 10 ml saturated aqueous solution of oxalic acid and subsequently rinsed several times with distilled water (Christensen, 1988). Pelleted frustules were mounted on slides and glued to aluminum stubs. After drying, they were sputter-coated with 5 nm gold/palladium and examined in a JEOL field emission SEM JSM-6335F.

3.19.5. Data analysis

Blast searches (www.ncbi.nlm.nih.gov/BLAST) were performed on all obtained DNA sequences to verify that the sequences were related to those of other *Skeletonema* species in the database. A minimum of five sequences were obtained for each clone and consensus sequences were produced using Sequencher 4.1.2 (Gene Codes Corporation, Ann Arbor, MI, USA).

Because several *Skeletonema* species have been described based on genetic sequences as well as morphological features (Sarno *et al.*, 2005), we used sequences available in GenBank to determine the taxonomic position of our clones. *Thalassiosira rotula* (AJ633505) was used as outgroup. Maximum likelihood (ML) analyses were performed using the heuristic search in PAUP* 4.0 b10 (Swofford, 2002). Bootstrap analysis was performed in 1000 replicates using strict consensus. The criterion settings for maximum likelihood analyses were according to the Hasegawa-Kisumo-Yano model. A starting tree was obtained via stepwise addition. In order to graphically depict the performance of the investigated molecular tools, DNA sequences were clustered using neighbor joining in PAUP*. This provided an objective picture of the similarity among samples without making phylogenic comparisons.

For RAPD analyses, bands between 200 and 3000 bp were scored for presence or absence in each isolate. We assumed RAPD markers to be inherited in a Mendelian fashion and scored as phenotypes (Hadrys *et al.*, 1992). Each band was treated as a genetic locus and indicated at least one dominant allele (AA or Aa). The same mating pattern (i.e., frequency of sexual propagation) was assumed in all sample locations. Compiled RAPD banding patterns from the replicates were analyzed in GelCompar II, version 2.5, by the band-based Dice similarity coefficient (position tolerance 0.4%) and clustered using neighbor joining.

Analysis of molecular variance (AMOVA) (Excoffier *et al.*, 1992) was used to examine the intraspecific diversity within and among the sample locations. The analysis was performed with the software Arlequin 2.000 using default settings. Significance of the variance components was tested with 10000 permutations (Schneider *et al.*, 2000). Clones were randomly assigned to each molecular method examined to minimize sampling bias. At least 24 clones were examined by each method (Table 14). Common geographic areas were the Swedish west coast (Koljöfjord, KF), the Swedish east coast (Kalmar, KAL), Portugal (PORT), and the Canadian west coast (CAN). In the RAPD analysis, additional samples from the Swedish west coast (Gullmarfjord, GULL and Hakefjord, C) were added. AMOVA was run using all available samples of the molecular methods investigated—LSU rDNA (domains D1–D3), ITS region (ITS 1—5.8S—ITS 2) and RAPD.

Table 14. Sample locations, isolation dates, strain designations, molecular techniques used and GenBank accession numbers

Location (abbreviation)	Isolation date	Strain designation	Molecular technique	GenBank accession number (LSU, ITS)
Swedish west coast				
Koljöfjord (KF)				
58°15'N 11°35'E	Feb-03	KF A2	RAPD	
	Feb-03	KF C1	RAPD	
	Oct-03	KF C23	ITS	AY748222
	Oct-03	KF C27	ITS	AY748223
	Oct-03	KF C33	ITS	AY748224
	Oct-03	KF C36	ITS	AY748225
	Feb-03	KF L2	RAPD	
	Feb-03	KF L3	RAPD, LSU	AY699039
	Feb-03	KF M1	RAPD	
	Feb-03	KF N2	RAPD, LSU, ITS	AY699038, AY748226
	Feb-03	KF P2	RAPD, ITS	AY748227
	Mar-03	KF R2	LSU	AY699040
	Mar-03	KF S1	LSU	AY699041
	Mar-03	KF S2	LSU	AY699042
	Mar-03	KF U1	ITS	AY748228
	Mar-03	KF Y2	RAPD, LSU	AY699043
	Mar-03	KF Z	RAPD	
Hakefjord (C)				
58°03' N 11°48' E	Mar-03	C2 M1	RAPD	
Gullmarfjord (GULL)				
58°16' N 11°26' E	Mar-03	GULL B4	RAPD	
	Mar-03	GULL C3	RAPD	
Swedish east coast				
Kalmar (KAL)				
58°40'N 16°22'E	Mar-03	KAL C3	RAPD, LSU	AY699031
	Mar-03	KAL D2	RAPD, LSU, ITS	AY699032, AY954681
	Mar-03	KAL E1	RAPD, LSU	AY699033
	Mar-03	KAL H1	LSU, ITS	AY699034, AY954682
	Mar-03	KAL H2	RAPD	
	Mar-03	KAL H3	LSU	AY699035
	Mar-03	KAL K1	RAPD	
	Mar-03	KAL M2	ITS	AY954683

	Mar-03	KAL M3	RAPD	
	Mar-03	KAL N2	RAPD, LSU	AY699036
	Mar-03	KAL N3	LSU,ITS	AY699037, AY954684
Portugal				
Lisboa (PORT)				
38°39'N 9°24'W	Mar-03	PORT A1	RAPD, ITS	AY748229
	Mar-03	PORT A3	RAPD, LSU	AY699044
	Mar-03	PORT C2	RAPD	
	Mar-03	PORT C3	LSU, ITS	AY699045, AY748230
	Mar-03	PORT D1	RAPD, ITS	AY748231
	Mar-03	PORT D2	RAPD, LSU, ITS	AY699046, AY748232
	Mar-03	PORT F1	RAPD, ITS	AY748237
	Mar-03	PORT F3	RAPD, LSU	AY699048
	Mar-03	PORT G3	RAPD, LSU, ITS	AY699047, AY748233
	Mar-03	PORT J1	RAPD, ITS	AY748234
Canada				
Strait of Georgia (CAN)				
49°9'N 123°23'W	Mar-03	CAN A1	LSU, ITS	AY699024, AY748236
	Mar-03	CAN B2	LSU, ITS	AY699025, AY748217
	Mar-03	CAN D1	RAPD	
	Mar-03	CAN D2	LSU, ITS	AY699026, AY748218
	Mar-03	CAN E2	RAPD, LSU	AY699027
	Mar-03	CAN G1	LSU, ITS	AY699028, AY748219
	Mar-03	CAN G2	RAPD	
	Mar-03	CAN I2	RAPD	
	Mar-03	CAN J1	LSU, ITS	AY699029, AY748220
	Mar-03	CAN K2	RAPD, LSU, ITS	AY699030, AY748221
	Mar-03	CAN L1	RAPD	
	Mar-03	CAN N2	RAPD	

3.20. Amplified Fragment Length Polymorphisms (AFLP) analysis

3.20.1. Experiment set up

Water samples were collected at the same locations along the coast of Mangalore, Karnataka, India, but separated in time by one year. From these water samples single chains of *S. marinoi* were isolated and cultured in f_2+Si media (Guillard, 1975) in 12:12 light: dark at 27°C. The experiment was set up with two clones of *S. marinoi* isolated from plankton samples collected one year apart, Man 3B (P1) and Bunder G (P2). Small volumes from each of these clonal isolates were mixed. Within a time-period of 1 to 3 weeks, single chains of cells where isolated from the parental-mix, denoted F1A-F1F (Fig 3). For comparison purpose, a clonal isolate (Pond 6A) sampled from Pondicherry, Tamil Nadu, on the east coast of India was also cultured under the same conditions.

3.20.2. Extraction of DNA

Prior to extraction, all cultures were scanned under microscope to confirm that they were not contaminated. The cultures were then transferred to 50-ml centrifuge tubes and centrifuged at 5000 g for seven minutes. The supernatant was discarded and cell pellets were transferred to labeled Eppendorf tubes and processed immediately. DNA was extracted using CTAB protocol (Rynearson and Armbrust, 2000) and Qiagen DNeasy Plant Mini Kit (QIAGEN GmbH, Hilden, Germany) following the instructions of the manufacturer. Concentration of the extracted DNA was measured using an UV-Visible spectrophotometer (Shimadzu Corporation, Japan).

3.20.3. Amplified Fragment Length Polymorphisms analysis

AFLP analysis was performed following certain modifications to the protocol described by Vos *et al.* (1995). Approximately 2000 ng of extracted DNA from each sample was digested with five units of EcoRI and MseI restriction enzymes (Van der Lee *et al.*, 1997) (Table 15) in total reaction volumes of 50 µl for three hours at 37°C. At hourly intervals, the tubes were tapped, and after three hours the enzymes were inactivated at 65°C for 10 minutes. Adaptors were prepared by mixing EcoRI linker1 and EcoRI linker2, MseI linker1 and MseI linker2 (Table 15) in separate tubes and incubating at 65°C for 10 minutes, followed by 37°C for 10 minutes and 25°C for 10 minutes. Ligation to the ends of the DNA fragments was made by adding 10 µl ligation-mix consisting of T4-ligase (1U, Bangalore Genei, Bangalore), 2.5 pmol of EcoRI adaptor, 25 pmol of MseI adaptor and 1x T4 DNA

ligase buffer to the restriction digested DNA and incubated at room temperature for three hours (with tapping at hourly intervals). The enzymes were inactivated at 65°C for 10 minutes. Pre-amplification reactions were set up in total volumes of 20 µl containing 1xPCR buffer, 4 nmol dNTP, 25 ng of primer EcoRI+A and MseI+C (Table 15), 1.2U Taq polymerase (Bangalore Genei, Bangalore) and 1-2 µl of restricted ligated DNA (2 µg). The pre-amplification reactions were performed using a PTC-100, Peltier-Effect Cycling system, (MJ Research Inc., Waltham, MA, USA), and carried out at 92°C for two minutes, 26 cycles of (94°C for 30 s; 58°C for 30 s; 72°C for 1 min) and a final extension at 72°C for five minutes. Duplicates of the pre-amplification samples and negative controls were run on a 2% agarose gel to check for the quality and contamination.

The PCR products from the pre-amplifications were diluted with 100 µl sterile double distilled water and used as template for consecutive PCR reactions. Selective amplifications were set up using 1xPCR buffer, 4 nmol dNTP, 1.2 U Taq polymerase, 25 ng of EcoRI+AAG and EcoRI+ACC, 30 ng of MseI+CTA and MseI+CTT (Table 16) and 10 µl of template DNA from the diluted pre-amplification PCR product. Amplification by touchdown PCR with initial denaturation at 94°C for two minutes and a first cycle at 94°C for 30 seconds, 65°C for 30 seconds, and 72°C for one minute. In the following 12 cycles, the annealing temperature was reduced by 0.7°C each cycle. Touchdown was then followed by 23 cycles of (94°C for 30 sec, 56°C for 30 sec, 72°C for 1 min), 72°C for five minutes and a hold at 4°C. Selective amplification products were loaded on to a 12% Novex® TBE Gel (Invitrogen) using the XCell SureLock™ Mini-Cell and the fragments were separated at 200 V approximately for 40 minutes. For visualization of the bands, the gel was stained using SilverExpress® Silver Staining Kit (Invitrogen), following the protocol supplied by the manufacturer. The results were documented using a digital camera and stored as TIFF-files for further processing.

3.20.4. AFLP data analysis

A consensus for each clone was manually created by measuring fragment sizes from different photographs and then put together to a consensus. It was assumed that bands of the same size where homologous. Data analysis was performed using the software GelComparII (Applied-Maths BVBA, Belgium). The percentage of similarity was calculated using the Dice co-efficient (Van der Lee *et al.*, 1997). For cluster analysis the neighbor joining algorithm (Hagen *et al.*, 2002; Furini and Wunder, 2003) was used to create a dendrogram.

To assess the reproducibility of the method the clonal isolates F1A and F1D were extracted at two occasions 10 days apart. The second extractions were denoted F1A2 and F1D2, respectively and underwent the same protocol as the other DNA extracts.

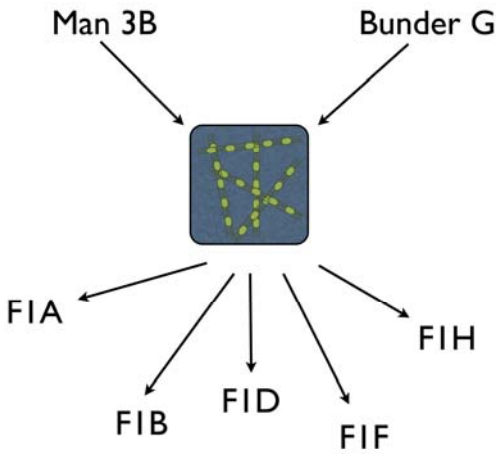


Fig. 3. Schematic representation of the experimental set up

Restriction Enzymes used	Sequence cutting site specificity (5’ – 3’)
Eco R1	G↓AATTC
Mse1	T↓TAA
Adaptors	Sequences
Eco R1 linker1 Eco R1 linker2	5’ – CTC GTA GAC TGC GTA CC – 3’ 3’ – CAT CTG ACG CAT GGT TAA – 5’
Mse1 linker1 Mse1 linker2	5’ – GAC GAT GAG TCC TGA G – 3’ 3’ – TA CTC AGG ACT CAT – 5’
Primers	---------- AS ---------- RS --- SB -----
Eco R1+A	5’ – GACTGCGTACC AATTC A – 3’
Eco R1+AAG	5’ – GACTGCGTACC AATTC AAG – 3’
EcoR1+ACC	5’ – GACTGCGTACC AATTC ACC – 3’
Mse1+C	5’ – GATGAGTCCTGAG TAA C – 3’
Mse1+CTA	5’ – GATGAGTCCTGAG TAA CTA – 3’
Mse1+CTT	5’ – GATGAGTCCTGAG TAA CTT – 3’

Table 15 and Table 16. Oligonucleotide sequences used in AFLP analysis. Restriction enzyme, adaptor and primers designed. AS: adaptor sequence, RS: restriction site sequence, SB: selective bases

3.21. Genetic heterogeneity and physiological variations among *S. marinoi* clones

3.21.1. Field samples and clones

The sampling site was located in the Gullmar Fjord on the Swedish west coast. Two major current systems affect the Swedish west coast; the low saline surface Baltic current running parallel to the coast, and the central Skagerrak water circulation pattern resulting in an inflow of more saline North Atlantic water. Hence, the water is permanently stratified in terms of salinity and a pronounced halocline is present (Rodhe, 1987; Lindahl, 1995). Gullmar Fjord has a sill at 43-meter depth and a maximum depth of 120 metres. The density fluctuation of the coastal water outside the fjord is most important for the water exchange above sill level in Gullmar Fjord, whereas the tidal force in the fjord is of minor importance. The average turnover time for the water above the halocline is 16-28 days. The horizontal exchange of surface water is characterized by incoming water that is only partially mixed with the resident water before it is flushed out of the fjord again. The depth position of the halocline is most stable during May to August (10-15 m) when intermediate water (deeper than 20 metres) very seldom is in contact with the surface, and consequently, the exchange rate with the coastal water outside the fjord is greater. Upwelling and downwelling events have larger amplitudes during the rest of the year, and therefore the position of the halocline is more variable (5-20 m) and the exchange rates with the coastal water is less. The salinity of the surface water is 24-27 PSU throughout the year, but the temperature varies with the lowest recorded temperatures in February-March at 2-4°C, and the highest in August-September at 15-17°C (Arneborg, 2004).

Field sampling was conducted at monthly intervals, during the expected growth season (January-October) of *S. marinoi*. Samples were collected more frequently during expected bloom periods (i.e., February-March and September-October), and not during the summer months (June 3-August 11), when *S. marinoi* is known to be absent or occurring at very low numbers (McQuoid and Godhe, 2004). Water samples were collected from the

surface (0-3 m depth) using an all-plastic 1.6-liter Ruttner water bottle. Simultaneously plankton was obtained by several net hauls (>10 µm) from the surface water. Lugol-fixed water samples were settled in sedimentation chambers overnight (Utermöhl, 1958), and abundances of *S. marinoi* were estimated at 200 or 400x using a Zeiss Axiovert 135 inverted microscope. From each sampling location, approximately 15-20 individual chains of cells were isolated by micropipetting from the fresh plankton-net samples. The chains were transferred to a drop of sterile f_2 medium (Guillard, 1975). This was repeated several times to clean the cells. Each clone was then transferred to a small (5 cm diameter) petri dish with 8 ml of f_2 medium and incubated at 10°C. When growth was confirmed in the petri dishes, cells were transferred to 50-ml Nunc flasks with f_2 medium, and reinoculated every third week. Nine different *S. marinoi* clones isolated from three different seasons of the year (February, June and September) were randomly selected for the experiment (Table 17).

3.21.2. Experiment set-up

The experiment was designed as to expose the clones to typical summer (17°C, 26 PSU), typical winter (3°C, 26 PSU), and various extreme conditions. Each clone was subjected to four different salinities (5, 15, 26 and 35 PSU), and three different temperature conditions (3, 10 and 17°C), yielding 12 different salinity-temperature combinations. The clones were kept in constant light source (50 µmol photons/m^2/s), and a photoperiod of light: dark 12:12 hours for three weeks. Prior to the start of the experiment, the cell numbers of each clone was estimated, in order to know that the volume added would approximately yield similar start concentrations to all clones in all salinity-temperature combinations. At the start of the experiment, 2 ml from each of the clones were inoculated into 50-ml Nunc bottles of each salinity-temperature combination. Later the same day an aliquot of 1.5 ml from the different combinations were fixed in 2% Lugol-Iodine solution to provide start data for cell abundances. Subsequently, all clones in all salinity-temperature combinations were sampled in the same manner once a week until the end of the experiment. Cell abundances

(cells/ml) were estimated by counting the number of cells settled in 10 ml sedimentation chambers overnight (Utermöhl, 1958). Minimum 300 cells were counted at 40x resolution using a Zeiss Axiovert 135 inverted light microscope.

3.21.3. RNA extraction

Extraction of RNA was performed once during the experiment, on Day 7. Fifteen milliliter of each *S. marinoi* clone from each salinity-temperature condition were centrifuged at 5000 *g* for seven minutes in 50-ml centrifuge tubes. The supernatant was decanted and the pelleted cells were transferred to clean 1.5-ml microfuge tubes. Total RNA was extracted using TRI-REAGENT (Sigma) following the instructions of the producers. After extraction, total RNA concentration (µg/ml) was instantly quantified using a spectrophotometer (Pharmacia Biotech GeneQuant II, Buckinghamshire, UK).

3.21.4. DNA extraction, PCR and sequencing

Prior to DNA extraction, the monoclonal cultures in exponential phase were scanned microscopically to ensure that no contaminants were present. The cultures were transferred to 50-ml centrifuge tubes and pelleted as described for RNA extraction. Pellets were resuspended in 420µl milli-Q water and transferred to microfuge tubes. Cells were lysed and DNA was extracted as described by Godhe *et al*. (2001). Concentration and purity of DNA was measured with a spectrophotometer (Pharmacia Biotech GeneQuant II, Buckinghamshire, UK). If sample purity was low, i.e., 260/280 ratio < 1.3, the extract was further purified with a Flexi Prep Kit (Amersham Pharmacia) following the instructions of the manufacturers. Extracted DNA was stored at -20˚C.

LSU rDNA domains D1–D3 were amplified by PCR. To eliminate potential Taq errors associated with the original PCR, the region of each clone was amplified more than once. Reactions were run in total volumes of

50 µl consisting of approximately 100 ng template DNA, Taq buffer containing 1.5 mM $MgCl_2$, 0.5 µM of each primer, 200 µM of each deoxynucleotide triphosphate (dNTP), 2.5 U of Taq polymerase (QIAGEN, Valencia, CA, USA), and sterile Milli-Q water. LSU rDNA sequences were amplified using the primers DIR and D3Ca (Lenaers *et al.*, 1989; Scholin *et al.*, 1994). Amplifications were carried out in a thermal cycler (Perkin Elmer GeneAmp PCR System 2400, Wellesley, MA, USA) as follows: initially, 94˚C for five minutes; followed by 30 cycles of denaturation at 94˚C for one minute; annealing at 55˚C for one minute; extension at 72˚C for one minute. After the cycles, extension was completed at 72˚C for five minutes. The PCR product was loaded on to a 0.8% agarose gel in 1x TBE, and ethidium bromide stained gels were studied under UV-transillumination. PCR products were purified using the QIAquick PCR Purification Kit (QIAGEN) following the manufacturer's instructions. Final concentration of the PCR product was measured using a spectrophotometer. For each clone, a minimum of two PCR products were sequenced. Each 10-µl sequencing reaction contained approximately 50 ng of PCR product together with the primers. LSU rDNA (D1–D3) fragments were sequenced by using the primers DIR, D2C, D2CF (Scholin *et al.*, 1994), and D3Ca (Lenaers *et al.*, 1989). LSU rDNA sequences were determined using the CEQ Dye Terminator Cycle Sequencing Kit following the manufacturer's instructions. Reaction products were run on a CEQ 8000 Genetic Analysis System (Beckman Coulter, Fullerton, CA, USA). Sequences are deposited in GenBank with accession numbers given in Table 17.

3.21.5. Data analysis

Consensus sequences for each clone were produced using Sequencher 4.1.2 (Gene Codes Corporation, Ann Arbor, MI, USA). Alignment provided information on pair-wise base-pair substitutions among all the investigated clones in this experiment. Several *Skeletonema* species have been described based on LSU rDNA sequences and morphological features revealed in electron microscopy (Sarno *et al.*, 2005; 2007). Our clones were identified to

species level by aligning the sequences to previously published *Skeletonema* LSU rDNA sequences (Sarno *et al.*, 2005; Godhe *et al.*, 2006; Sarno *et al.*, 2007). *Thalassiosira rotula* (AJ633505) was used as the outgroup. According to Modeltest 3.7 (Posada and Crandall, 1998), GTR+G model (General Time Reversible) with a Gamma distributed rate of variation across sites, was most appropriate to analyze the dataset. Maximum likelihood (ML) analyses was performed using the heuristic search in PAUP* 4.0 b10 (Swofford, 2002). Growth rate was calculated by $N_t=R^t$ x N_0, were R is the growth rate (doublings/d), N_0 is the cell abundance (cells/ml) at the onset of the experiment, and N_t is the cell abundance on day t. RNA concentration per cell was calculated by dividing the total RNA concentration (µg/ml), by cell abundance (cells/ml) recorded on the same day.

The influence of temperature, salinity, and isolation month on maximum growth rates, maximum cell abundances and RNA concentration per cell, were tested with univariate analysis of variance (ANOVA). Clones were not included as a factor in the ANOVA, since each clone was not replicated in the different salinity-temperature combinations. Significance was defined as $p < 0.05$. Significant interactions were further tested with one-way ANOVA. All data were tested for heterogeneous variances with Levene's test of equality of error variances. Maximum growth rates and RNA concentration per cell was log-transformed prior to analyses. Post-Hoc tests were performed with Student Newman-Keuls (SNK) tests, and Chi-Square tests were used to determine significant comparisons between pairs of variables. All statistic analyses were made in SPSS 11.0.4 for Mac OS X (SPSS Inc.,Chicago IL, USA).

Table 17: Isolation dates and GenBank accession numbers of *S. marinoi* clones used

Monoclonal strains of *S. marinoi*

Strain	Isolation date	GenBank accession number
GF04 1F	20-Feb-04	DQ438874[1]
GF04 1G	20-Feb-04	DQ438873[1]
GF04 1J	20-Feb-04	DQ438872[1]
GF04 7C	2-Jun-04	
GF04 7F	2-Jun-04	
GF04 7J	2-Jun-04	
GF04 9A	23-Sep-04	DQ438871[1]
GF04 9B	23-Sep-04	DQ438870[1]
GF04 9D	23-Sep-04	DQ438869[1]

IV. EXPERIMENTAL RESULTS

IV. RESULTS

4.1. Genomic DNA extraction and analysis

Genomic DNA extraction from *S. marinoi* cells were carried out using two different methods, that is one using phenol-chloroform-isoamyl alcohol extraction method and other using cetyl trimethyl ammonium bromide (CTAB) extraction method. Results suggest that the phenol-chloroform-isoamyl alcohol method was the most effective for extracting genomic DNA from *S. marinoi* cells. The presence of DNA was confirmed by running a gel electrophoresis using 1% Agarose gel. The appearance of thick bright band on the agarose gel indicates the presence of genomic DNA (Fig 4). The concentration and purity of the extracted genomic DNA were measured using UV-spectrophotometric method as indicated (Table 18). The concentration of the extracted genomic DNA was 80.55 µg/ml and the purity (A260/A280) of the genomic DNA was 1.55.

4.2. Genomic DNA digestion with restriction enzyme

Genomic DNA digestion was performed using Sau3A1 restriction enzyme. Sau3AI is a four-base cutter and cuts approximately once every 256 base pairs producing sticky end products. The success of the restriction digestion was analyzed by running a gel electrophoresis of the Sau3AI digested genomic DNA sample on a 2% agarose gel at a constant voltage of 100 V. The appearance of a thin continuous smear of DNA fragments indicated the successful digestion of the genomic DNA by Sau3AI restriction enzyme (Fig 5)

4.3. Adaptor ligation

After the Sau3AI restriction enzyme digestion of genomic DNA, adaptor ligation was performed to the digested genomic DNA end products. The 200 µl of adaptor ligated genomic DNA fragments were purified using phenol-chloroform-isoamyl alcohol and subsequently ethanol-precipitated. The purified DNA fragments were eluted in 10 µl of 0.1xTE buffer, which was used as a DNA template in Test-PCR and PCR of adaptor ligated DNA.

4.4. Test PCR and PCR of adaptor-ligated DNA

The success of the adaptor ligation was confirmed by performing a single Test-PCR reaction using only 1 µl of the adaptor-ligated genomic DNA as the template. The product of the Test-PCR was run on a 2% Agarose gel at a constant power of 120 V. The thick smear on the agarose gel indicated successful amplification of the DNA fragments of varying sizes (Fig. 6). Subsequently, the remaining 9 µl of the adaptor-ligated DNA was used to perform nine separate PCR reactions of the adaptor-ligated DNA using 1 µl ligated DNA as the template in each reaction. The products from only one PCR reaction along with a negative control was again run on a 2% agarose gel. The continuous smear indicated the successful PCR amplification of adaptor-ligated DNA fragments (Fig. 7).

4.5. Hybridization and capture of microsatellite-containing DNA fragments using biotinylated microsatellite probes

The adaptor-ligated genomic DNA fragments after PCR amplification were hybridized to (CA)15 or (GA)15 biotinylated universal microsatellite probes. Microsatellite oligos with the same hybridization temperature can be added at the same time, whereas oligos with different hybridization temperatures have separate reactions. Since $(CA)_{15}$ and $(GA)_{15}$ microsatellite probes had different hybridization temperatures they could not be used in the same hybridization reaction. Microsatellite $(CA)_{15}$ probe had a hybridization temperature of 55°C whereas, 50°C was the hybridization temperature used in case of $(GA)_{15}$ microsatellite probe. The microsatellite containing genomic DNA fragments hybridized to biotinylated $(CA)_{15}$ or $(GA)_{15}$ probes were captured using streptavidin coated magnetic beads and eluted in 120 µl of 1x TE buffer.

4.6. PCR amplification following microsatellite capture

The microsatellite-containing genomic DNA fragments captured were further enriched by microsatellite PCR using primers binding to the adaptor sequences. The PCR products containing microsatellite repeats were electrophoresed using 2% agarose gel at a constant voltage of 120 V. A

bright thick smear on the agarose gel indicated the successful hybridization and capture of microsatellite containing DNA fragments (Fig. 8 and Fig. 9).

4.7. Transformation and screening of microsatellite-containing clones

The microsatellite-enriched DNA fragments were ligated to a plasmid vector, pCR2.1 TOPO10 and subsequently, cloned into competent *E. coli* cells (TOPO10). The colonies were screened for positive and negative clones by blue-white screening method. The pure white colonies were checked for the presence of microsatellite inserts using the screening PCR technique. In screening PCR, two primers one complementary to the adaptors and another complementary to microsatellite regions were used. The PCR products were checked by performing electrophoresis on 2% agarose gel. The appearance of double bands in the gel is the confirmed indication of the presence of microsatellite repeat within the clones (Fig 10, 11, 12 and 13). The very intense, bright single bands also sometimes contained the microsatellite-repeat regions.

4.8. Plasmid extraction and library construction

All the positive clones showing double bands and intense bright single bands were subjected to plasmid extraction using Eppendorf fast plasmid mini kit. All the positive clones were stored in glycerol broth at -80°C for long-term storage and also for genomic library construction.

4.9. M13 PCR

The extracted plasmids were subjected to PCR amplification using M13 universal primers. The M13 PCR was performed in order to check for the presence or absence of inserts and also to determine the insert sizes in all the positive clones. The products of M13 PCR were run on a 2% agarose gel and all the clones possessing the insert size in the range of 200-100 base pairs were selected for sequencing (Fig. 14 and Fig. 15).

4.10. DNA sequencing

From the sequencing results, two loci (AU125 and J33) were obtained which were highly stable and polymorphic in nature. The locus AU125 contained a tri-nucleotide repeat with the core sequence being (GAT) motif, which was repeated nine times consecutively Table 19. In the case of locus J33, the core sequence was a di-nucleotide repeat (CT) motif which repeated 10 times consecutively (Table 20).

4.11. Designing of microsatellite primers

Two microsatellite primer pairs were developed from genomic sequence of *S. marinoi* containing microsatellite motifs $(GAT)_9$ and $(CT)_{10}$ located at the loci (AU125 and J33) respectively. The binding sites for two primer sets designed are located at the flanking sequences encompassing the two microsatellite motifs. The designed forward and reverse primer at locus AU125 along with the primer specifications are shown in Table 21. The forward and reverse primers at locus J33 along with the primer specifications are shown in Table 22.

4.12. Allele frequency position and PCR primer conditions

The *S. marinoi* samples which were used for checking the microsatellite motifs at loci (AU125 and J33) and their allele frequency distribution are shown in Table 23 and 24 respectively. The optimum annealing temperature for microsatellite PCR and other primer characteristics at locus AU125 and locus J33 are shown in Table 25, 25.1 and Table 26, 26.1, respectively. *S. marinoi* is diploid in nature and the allelic diversity at microsatellite loci AU125 and J33 amplified from *S. marinoi* clones isolated from the South-west coast of Sweden during the period from February 18, 2004 to February 21, 2005 are shown in the Fig. 16 and Fig. 17 respectively.

Table 18. UV-Spectrophotometric analysis of the extracted *S. marinoi* genomic DNA

Volume of sample (µl)	Optical density at 260nm	Optical density at 280nm
2	0.0034	0.0022
4	0.0062	0.0040
6	0.0095	0.0061
Ratio (260/280)	1.55	
Concentration of DNA (µg/ml)	80.55	

Table 19. DNA sequencing results at Locus AU125

GGGCCCTCTAGATGCATGCTCGAGCGGCCGCCAGTGTGATGGATATCTG
CAGAATTCGCCCTTGTGCCAGAGACCCCAAGCATTCGGATCGACGAGA
TGGATACTACACAACAACGAGTATCACAAGAAATGAAGAGGAGTTATT
ATCAACTGTG***TTCGGACTTGTTGTATAGGGTGG***AAGATWTGAAGG
AGGAGGGGTACGAGGAGGTTTATACTGGTGGGTTACGAGTACATGATT
GCAGCGCAGCTACCAAACAACAACA**GATGATGATGATGATGATGATG**
ATGATACATGCCGTACTACCCACTGTAGACACA**#CCAACACCCTACAC**
ACCCAA#AGWAAAAGTTACCGTAGCAGATCGAAGCTTGAGGGTCTCTG
GACCAASGGCGAATTCCAGCACACTGGCRGCCGTTACTAGTGGATCCG
AGCTCGGTACCT

Note – * indicates the forward primer sequence region.
# indicates the reverse primer sequence region.
GAT is the microsatellite repeat sequence motif at the Locus AU125.

Table 20. DNA sequence results at Locus J33

CGATGCTCGAGTTGCACTATGCGCGAGCAATACTCCTGGTGRCTAGCAT
GACTCCTCACTGAACGGATCAGACACCATGGGGWTTACTATCTGAGTA
AACTA***AGGCACTCATGTACTAGATGTCA***ATCGTAAACAAGCACCTA
CTTGTACTTACCCAGCTGAAG**CTCTCTCTCTCTCTCTCTCT**GAATCCG
CTACAATAAAT**#ACCCTCTCACTTTCGTCCT#**GCCGGGAAACGGGGTA
AGAGGGAAATGTGGGTATAAGAAATGACCGCGGGTATCTGTTGACATA
TGCCACGCATGCGTGGAATAAAATGGCTGCCAAAGACATGTATGTCTG
TC

Note – * indicates the forward primer sequence region.
# indicates the reverse primer sequence region.
CT is microsatellite repeat sequence motif at the Locus J33.

Table 21. Details of forward (upper) and reverse (lower) primers designed for the
Locus AU125

Table 21.1. Characteristics of the upper primer designed for the Locus AU125

AU125 : 156U23 UPPER PRIMER 5’ **TTC GGA CTT GTT GTA TAG GGT GG** 3’	
Length	23 MER
5’ position	156
Tm	68.3°C (Salt 1000 mM; Oligo 100 pM)
ΔG (25°C)	- 42.5 kcal/mol
Degeneracy	1
P.E.#	439/439
1/E	32.7 µg/A260

Table 21.2. Characteristics of the lower primer designed for the Locus AU125

AU125 : 317L21 LOWER PRIMER 5’ **TTG GGT GTG TAG GGT GTT GGT** 3’	
Length	21 MER
3’ position	317
Tm	67.0°C (Salt 1000 mM; Oligo 100 pM
ΔG (25°C)	- 39.2 kcal/mol
Degeneracy	1
P.E.#	410/410
1/E	33.1 µg/A260

Table 22. Details of the forward (upper) and reverse (lower) primers designed for the Locus J33

Table 22.1. Characteristics of upper primer designed for the Locus J33

J33 : 103U23 UPPER PRIMER 5' **AGG CAC TCA TGT ACT AGA TGT CA** 3'	
Length	23 MER
5' position	103
Tm	61.5°C (Salt 1000 mM; Oligo 100 pM)
ΔG (25°C)	- 37.1 k cal/mol
Degeneracy	1
P.E.#	367/367
1/E	4.35 nmol/A260 31.0 µg/A260

Table 22.2. Characteristics of the lower primer designed for the Locus J33

J33 : 203L19 LOWER PRIMER 5' **AGG ACG AAA GTG AGA GGG T** 3'	
Length	19 MER
3' position	203
Tm	60.2°C (Salt 1000 mM; Oligo 100 pM)
ΔG (25°C)	- 35.3 k cal/mol
Degeneracy	1
P.E.#	380/380
1/E	4.84 nmol/A260 29.4 µg/A260

Table 23. Allele position in *S. marinoi* samples at Locus AU125

Samples	Locus	Allele 1	Allele2
GFO4-1F	AU125	185	
GFO4-1GI	AU125	185	
GFO4-3I	AU125	188	
GFO4-3J	AU125	185	
GFO4-3O	AU125	185	
GFO4-4F	AU125	185	
GFO4-4G	AU125	185	
GFO4-4L	AU125	185	
GFO4-5E	AU125	185	
GFO4-6E2	AU125	185	
GFO4-7B	AU125	185	
GFO4-8B	AU125	ND	
GFO4-9D	AU125	185	
GFO4-9A	AU125	179	185
GFO4-9B	AU125	185	
GFO4-10H	AU125	ND	
GFO4-10K	AU125	185	
GFO4-11A	AU125	185	
GFO4-12G	AU125	185	
GFO4-12S	AU125	185	
GFO4-11K	AU125	ND	
GFO5-3E	AU125	185	
GFO5-5W	AU125	185	
GFO5-6Y	AU125	185	

Table 24. Allele position in *S. marinoi* samples at Locus J33

Samples	Locus	Allele1	Allele 2
GFO4-1F	J33	126	130
GFO4-1GI	J33	124	132
GFO4-3I	J33	124	132
GFO4-3J	J33	124	
GFO4-3O	J33	124	
GFO4-4F	J33	124	
GFO4-4G	J33	124	
GFO4-4L	J33	126	132
GFO4-5E	J33	124	132
GFO4-6E2	J33	132	
GFO4-7B	J33	124	
GFO4-8B	J33	ND	
GFO4-9D	J33	124	
GFO4-9A	J33	126	132
GFO4-9B	J33	ND	
GFO4-10H	J33	126	134
GFO4-10K	J33	124	132
GFO4-11A	J33	ND	
GFO4-12G	J33	124	
GFO4-12S	J33	ND	
GFO4-11K	J33	ND	
GFO5-3E	J33	132	
GFO5-5W	J33	124	132
GFO4-6Y	J33	132	

Table 25. Primer conditions for PCR amplification of the Locus AU125

	Position and length		Tm(°C)	GC (%)	P.E.
Product	182 base pairs		83.8	46.7	---
Upper primer			70.6	47.8	439/439
Lower primer	156	23	70.9	52.4	410/410
	317	21			
Optimal annealing temperature – 54.8°C (Max : 72.0°C)					

Product Tm – Upper primer Tm: 13.1
Primer Tm difference: 0.3

Table 25.1. Primer concentrations

	Concentration
Upper primer	200.0 nM
Lower primer	200.0 nM
Monovalent cation	50.0 mM
Free mg (2+)	0.7 mM

Total Na (+) equivalent – 155.8

Table 26. Primer conditions for PCR amplification of the Locus J33

	Position and length		Tm(*C)	GC (%)	P.E.
Product	119 base pairs		81.4	44.5	---
Upper primer			64.9	43.5	367/367
Lower primer	103	23	64.8	52.6	380/380
	203	19			
Optimal annealing temperature – 51.4°C (Max : 67.8°C)					

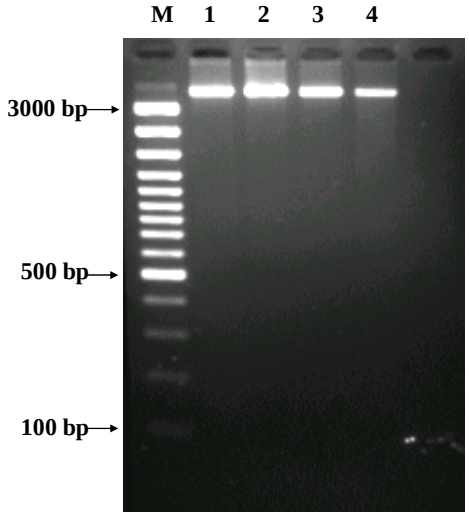
Product Tm – lower primer Tm: 16.7
Primer Tm difference: 0.2

Table 26.1. Primer concentrations

	Concentration
Upper primer	200.0 nM
Lower primer	200.0 nM
Monovalent cation	50.0 mM
Free mg (2+)	0.7 mM

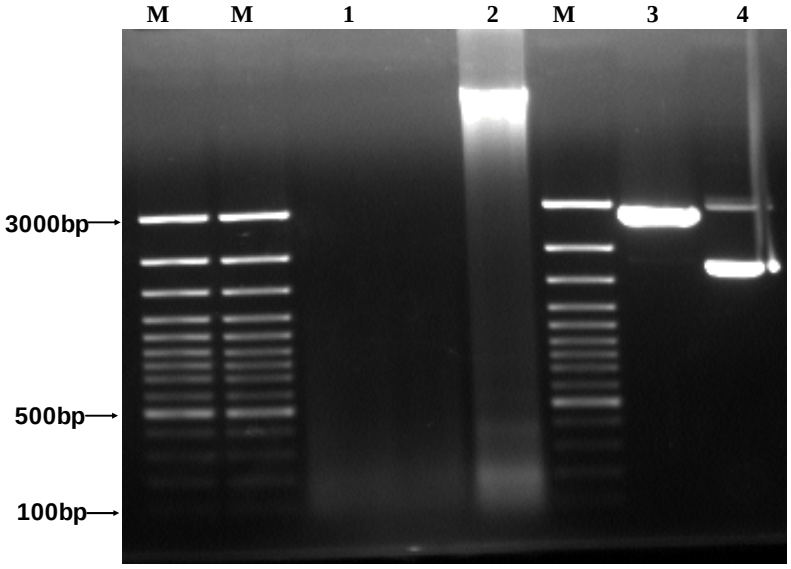
Total Na (+) equivalent – 155.8

Fig. 4. Agarose gel electrophoresis of genomic DNA of *S. marinoi*



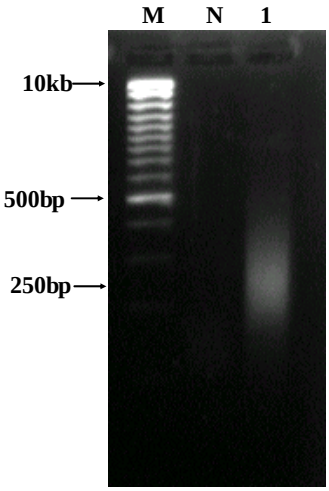
Lanes M: Gene Ruler™100 bp DNA Ladder Plus, Lanes 1-4: *S. marinoi* genomic DNA

Fig 5. Digestion of *S. marinoi* DNA using Sau3AI restriction enzyme



Lanes M: Gene Ruler™100 bp DNA Ladder Plus, Lane 1: Digested DNA, Lane 2: Undigested DNA, Lane 3: Digested PUC18, Lane 4: Undigested PUC18.

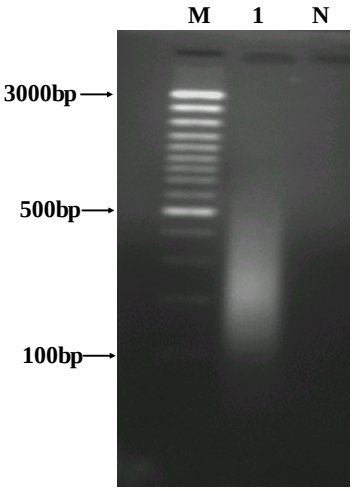
Fig 6. Representative PCR amplification reaction of adaptor-ligated *S. marinoi* DNA



Lanes M: Gene Ruler™ 1kb DNA LADDER, N: Negative control,

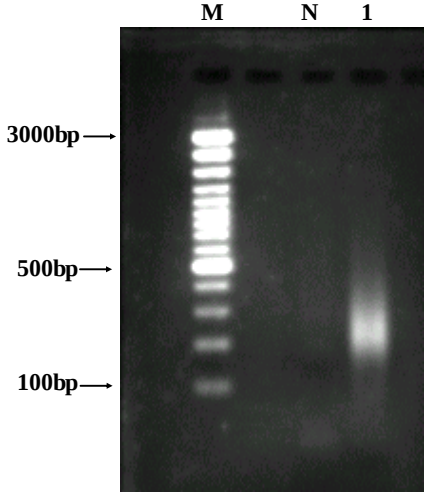
1: Adaptor ligated *S. marinoi* DNA

Fig 7. PCR of the digested, adapter /linker ligated genomic DNA of *S. marinoi*.



1 Lanes M: Gene Ruler™100 bp DNA Ladder Plus, N: Negative control, 1: Digested, Adaptor – Ligated *S. marinoi* Genomic DNA.

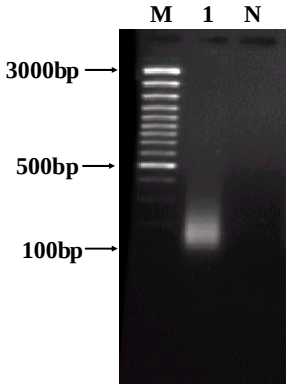
Fig 8. Enrichment PCR for (CA)15 microsatellite repeats of *S. marinoi*



Lanes M: Gene Ruler™ 100 bp DNA Ladder Plus, N: Negative control,

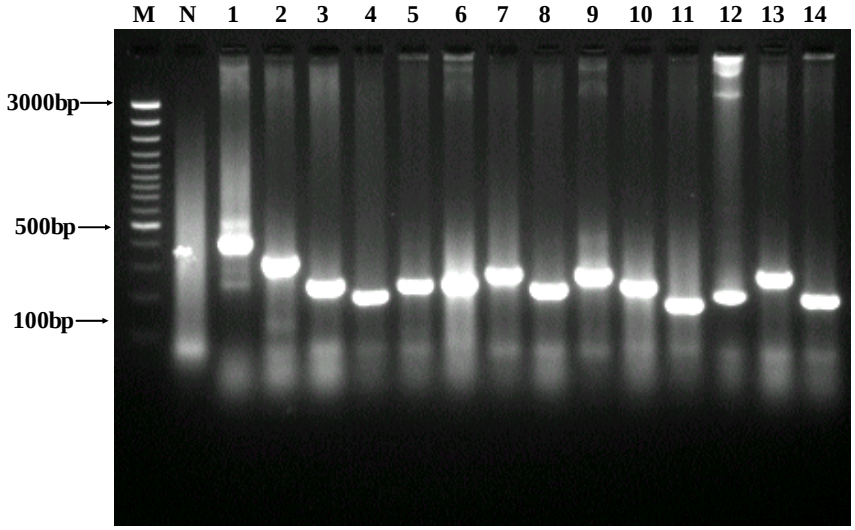
1: *S.* marinoi genomic DNA containing (CA)15 enriched microsatellite repeats.

Fig 9. Enrichment PCR for (GA)15 microsatellite repeats of *S. marinoi*



Lanes M: Gene Ruler™ 100 bp DNA Ladder Plus, N: Negative control, 1: *S. marinoi* Genomic DNA containing (GA)15 enriched microsatellite repeats

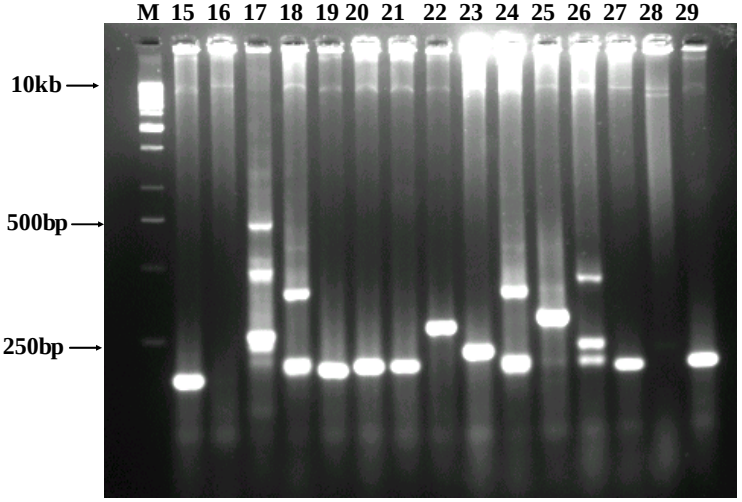
Fig 10. Screening PCR for *Skeletonema marinoi* clones containing (CA)15 repeats



Lanes M: Gene Ruler™ 100 bp DNA Ladder Plus, N: Negative control,

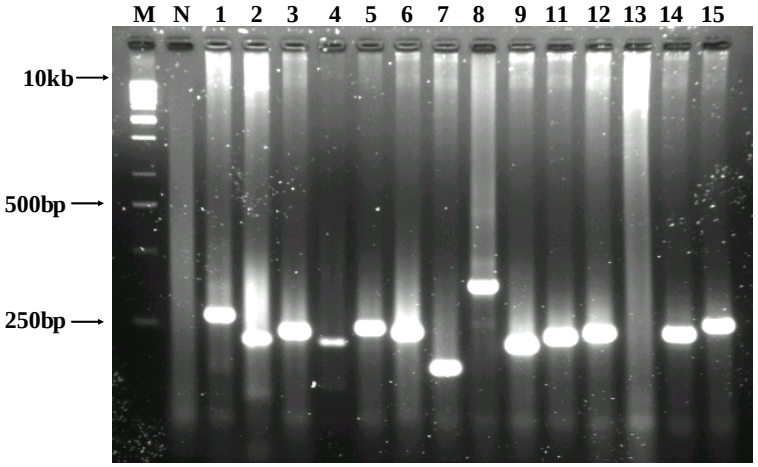
Lanes 1 – 14: *S. marinoi* Clones containing (CA)15 microsatellite repeats

Fig 11. Screening PCR for *Skeletonema marinoi* clones containing$(CA)_{15}$ repeats



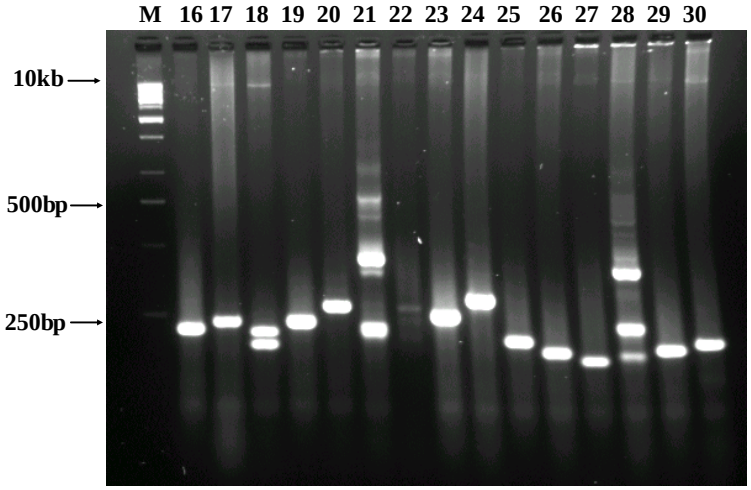
Lanes M: Gene Ruler™ 1kb DNA LADDER, Lanes 15 – 29: *S. marinoi* Clones containing $(CA)_{15}$ microsatellite repeats.

Fig 12. Screening PCR for *Skeletonema marinoi* clones containing $(GA)_{15}$ repeats



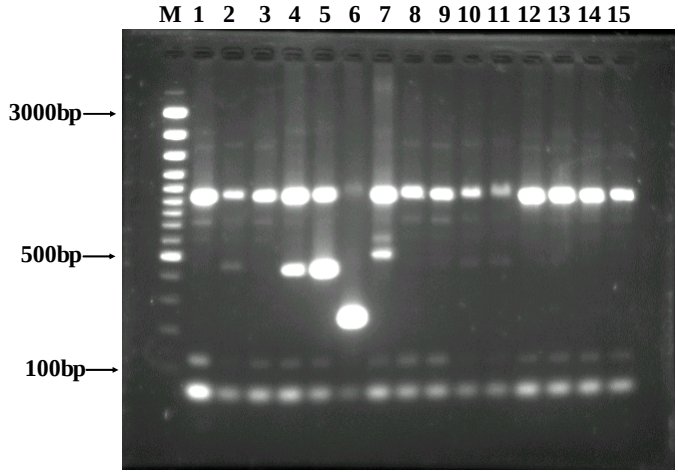
Lanes M: Gene Ruler™ 1kb DNA LADDER, Lanes 1 – 15: Clones containing $(GA)_{15}$ microsatellite repeats.

Fig 13. Screening PCR for *Skeletonema marinoi* clones containing $(GA)_{15}$ repeats



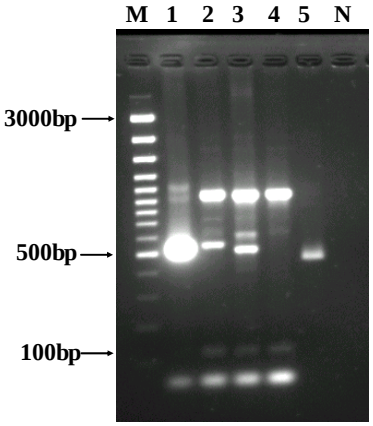
Lanes M: Gene Ruler™ 1kb DNA LADDER, Lanes 16 – 30: *S. marinoi* Clones containing $(GA)_{15}$ microsatellite repeats

Fig 14. M13 PCR of *Skeletonema marinoi* clones containing $(CA)_{15}$ microsatellite inserts



Lanes M: Gene Ruler™ 100 bp DNA Ladder Plus, N: Negative control,
Lanes 1 – 15: *S. marinoi* Clones containing $(CA)_{15}$ microsatellite inserts.

Fig 15. M13 PCR of *Skeletonema marinoi* clones containing $(GA)_{15}$ microsatellite inserts.



Lanes M: Gene Ruler™100 bp DNA Ladder Plus, N: Negative control, Lanes 1 – 5: Clones containing $(GA)_{15}$ microsatellite inserts.

Fig 16. Allelic diversity of *Skeletonema marinoi* clones at Microsatellite loci AU125

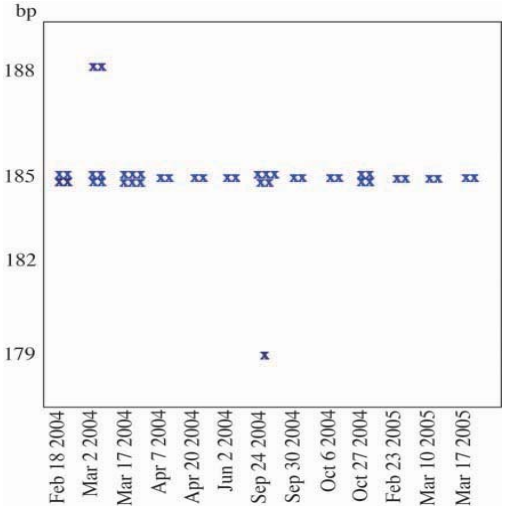


Fig 16: Tri-nucleotide repeats, each (x) indicates one allele.

Fig 17: Allelic diversity of *Skeletonema marinoi* clones at Microsatellite loci J33

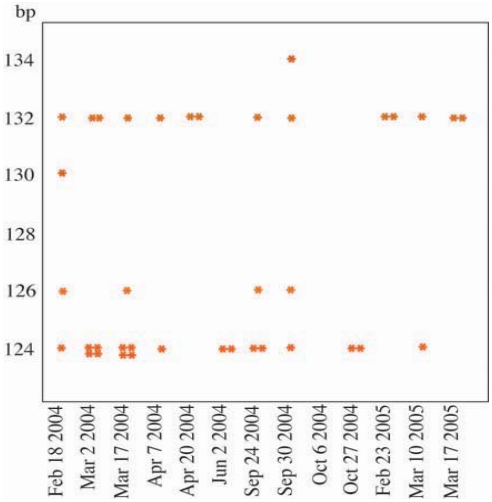


Fig 17: Di-nucleotide repeats, each (*) indicates one allele.

4.13. LSU, ITS and RAPD analysis

More than 90% of the isolated chains of *S. marinoi* grew to densities high enough for DNA extraction. High proportion of survival, and the short period before cell harvest and DNA extraction minimized bias due to culturing conditions. Randomly choosing from a large pool of clones for each method that was investigated, further reduced the probability that the performance of the method was due to the selected suite of clones. This was particularly true for RAPD where the whole genome was analyzed.

The ML tree of LSU rDNA sequences and those found in GenBank showed that all our clones cluster with previously published sequences for *S. marinoi* (Fig. 18). Whereas six of our European clones were identical to the eight previously published sequences for *S. marinoi*, the other 19 sequences showed polymorphism. In these clones, 35 bases of the LSU rDNA sequence were polymorphic (Table 27). All cells analyzed in SEM were loosely connected by fultoportulae with flared tips (Fig. 19).

The power to differentiate among clones of *S. marinoi* was analyzed separately for each method. The full length (approximately 800 bp) of the D1–D3 LSU rDNA was analyzed. The ITS regions (i.e., ITS 1, 5.8S rDNA and ITS 2) were identified by aligning the obtained fragments with ITS sequences available at GenBank. The lengths of ITS 1, 5.8 S, and ITS 2 were approximately 230 bp, 164 bp and 300 bp, respectively (Table 27). The ITS regions were analyzed both separately and as the full fragment (ITS 1—5.8S—ITS 2). In the RAPD analysis two oligonucleotide primers generated repeatable banding patterns in all strains and 38 loci were used for further analyses.

In the cluster analysis inferred from LSU rDNA sequences, clones from Canada were separated from the European strains (Fig. 20). Groups among the European clones, however, were not as clearly separated in the unrooted cladogram. The lowest branch consisted mainly of the Portuguese clones but also contained a few Swedish clones from both the east (KAL H3 and KAL N3) and west (KF N2 and KF R2) coasts. The Canadian clones formed a single

branch to the upper left of the diagram. Two smaller branches were also formed by some of the clones from the Swedish east coast and the Swedish west coast. The number of polymorphic sites along the LSU rDNA sequence varied for the different geographic locations (Table 27). Among the Canadian clones, 18 sites were polymorphic, whereas only one site was polymorphic among the Portuguese isolates. In the AMOVA, 36% of the overall variance was attributed to differences among sample locations, whereas 64% resided within a geographic sample (Table 28).

Genetic signatures of *S. marinoi* from geographically separated samples were best resolved by ITS 2 sequences (Fig. 21a). Branches of the ITS 2 cluster analysis coincided with geographic distribution. The samples from Portugal, the Swedish east coast and Canada formed end branches in the unrooted cladogram. Clones from the Swedish west coast were located at the center. Twenty-one out of the 300 bp ITS 2 sequence were polymorphic (Table 27). Six of the 21 polymorphic sites were attributed to differences among geographic samples only. Fifteen of the 21 sites implied variation within samples. In the AMOVA, more than 79% of the variance could be attributed to differences among geographic samples, whereas less than 21% of the variance originated from within a location (Table 28). In the ITS 1 sequence, 12 out of 230 bp were polymorphic. In the cluster analysis inferred from ITS 1 rDNA sequences, clones from Canada were separated from the European strains. Groups among the European clones, however, were not separated in the unrooted cladogram (Fig. 21b). No polymorphic sites were detected in the 5.8S sequence (Table 27). The variances among samples for the full ITS region (ITS 1—5.8S—ITS 2) and of ITS 1 separately were 73 and 63%, respectively.

The reproducibility of the RAPD banding pattern using different DNA extractions was found to be stable with nearly all bands being reproduced. The differences noted in the densiometric curves were rather based on variation in intensity than absence or presence of bands. Similar to ITS and LSU, the two oligonucleotides used in RAPD separated the Canadian from European clones (Fig. 22). The European clones formed a less distinct pattern. The clones originating from the Swedish west coast (KF, GULL, C2) did not form any discernable branch. Most clones from Portugal and one from the Swedish west coast (KF C1) formed one branch. Other Portuguese clones were interspersed with clones from the Swedish west coast. Five clones from the Swedish east coast (KAL) plus one clone from the

Swedish west coast (KF P2) formed the most distant branch. Two clones from the Swedish east coast (KAL D2 and KAL H2) were more similar to clones originating from the Swedish west coast than to other east coast isolates. Out of 35 clones, only two from Portugal had identical banding patterns (Table 29). Two loci were shared by all 35 clones. For all geographic samples, five loci were unique. The AMOVA based on RAPD banding patterns assigned 26% of the total variance to differences between samples and 74% of the variance within a geographic location (Table 28).

Table 27. Number of polymorphic sites including gaps in the ITS (ITS 1, 5.8S, ITS 2) and LSU rDNA (D1—D3). The number of polymorphic sites are given for all clones, and for clones within each geographic region.

	No. of sequences ITS/LSU	Polymorphic sites			
		ITS 1	5.8S	ITS 2	LSU
Length of sequence (bp)		230	164	300	800
All clones	24/25	12	0	21	35
Swedish west coast	7/6	1	0	6	5
Swedish east coast	4/7	4	0	4	13
Portugal	7/5	4	0	3	1
Canada	6/7	3	0	2	18

Table 28. **Analysis of molecular variance (AMOVA) of ITS, LSU sequences and RAPD. df = degrees of freedom, SS = sum of squares**

Method	Variance component	df	SS	Variance	Percent variation	P value
ITS 2	among populations	3	68.57	3.70	79.50	<0.0001
	within populations	21	19.10	0.95	20.50	<0.0001
	total	24	87.67	4.65		
ITS1—5.8S—ITS2	among populations	3	98.80	5.24	73.01	<0.0001
	within populations	21	38.74	1.94	26.99	<0.0001
	total	24	137.54	7.18		
ITS 1	among populations	3	29.90	1.53	62.90	<0.0001
	within populations	21	18.06	0.90	37.1	<0.0001
	total	24	47.96	2.43		
LSU	among populations	3	36.69	1.52	35.54	<0.0001
	within populations	22	58.03	2.76	64.46	<0.0001
	total	25	94.72	4.28		
RAPD	among populations	3	46.06	1.35	26.19	<0.0001
	within populations	32	117.60	3.79	73.81	<0.0001
	total	35	163.66	5.14		

Table 29. Number of unique and shared RAPD loci for all clones and for samples from the respective geographic locations

	No. of clones /identical clones	Unique loci	Shared loci
All clones	35/2	5	2
Swedish west coast	12/0	13	2
Swedish east coast	7/0	8	2
Portugal	9/2	7	4
Canada	7/0	8	3

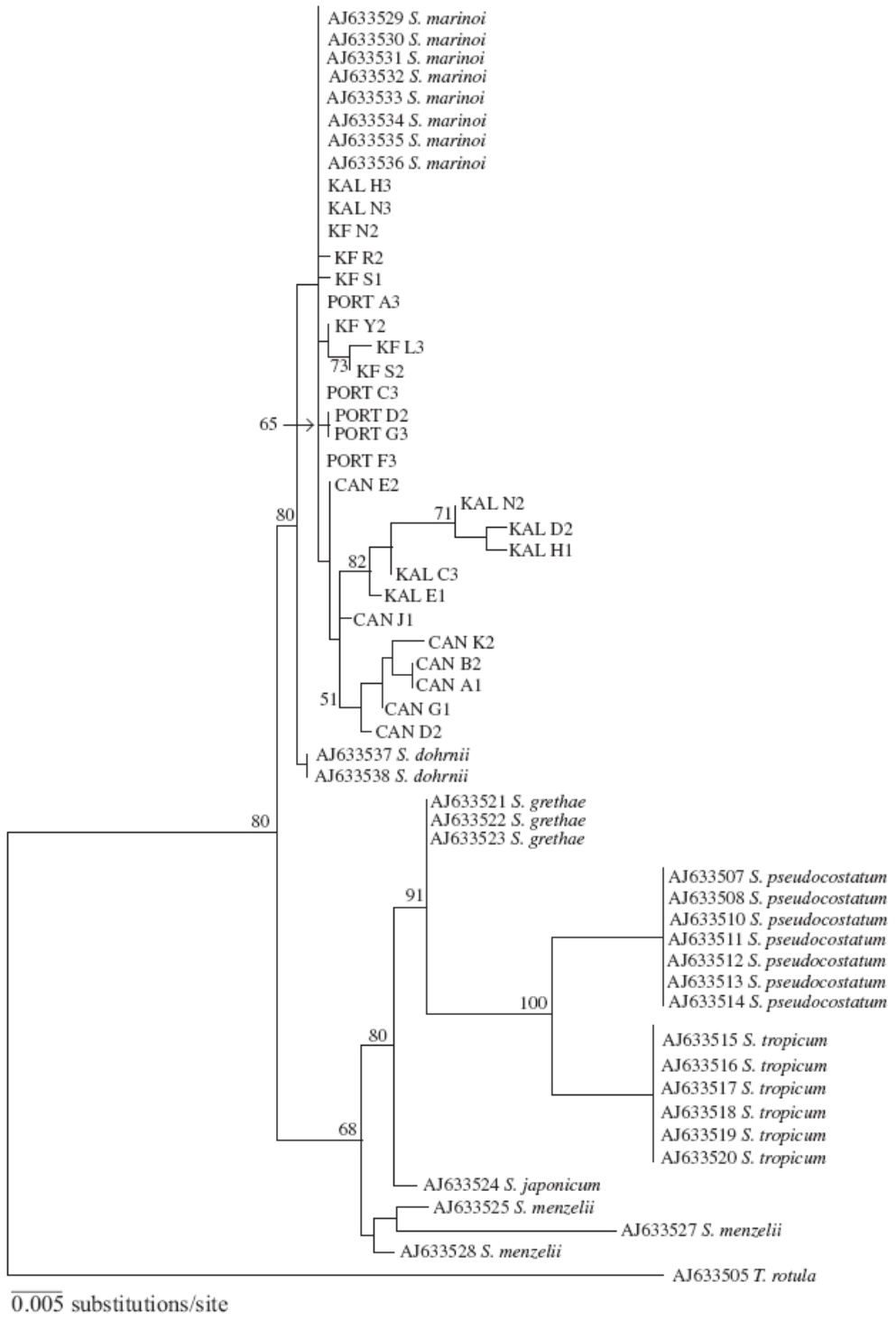


Fig. 18. Maximum likelihood tree inferred for *Skeletonema* clones based on 797 positions at the 5' end of the nuclear rDNA large subunit region. The tree is rooted with the outgroup taxon *Thalassiosira rotula*. Sequences labeled with species names and accession numbers were obtained from GenBank. Bootstrap values are inferred from 1000 replications.

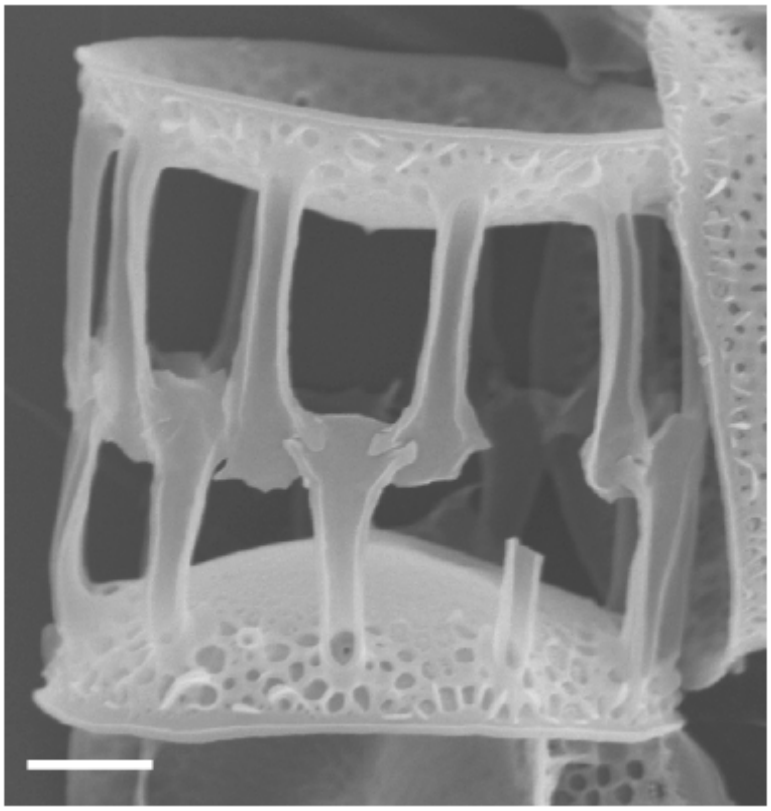


Fig. 19. SEM photograph of *S. marinoi* (clone KF L2), Scale bar, 1 μm

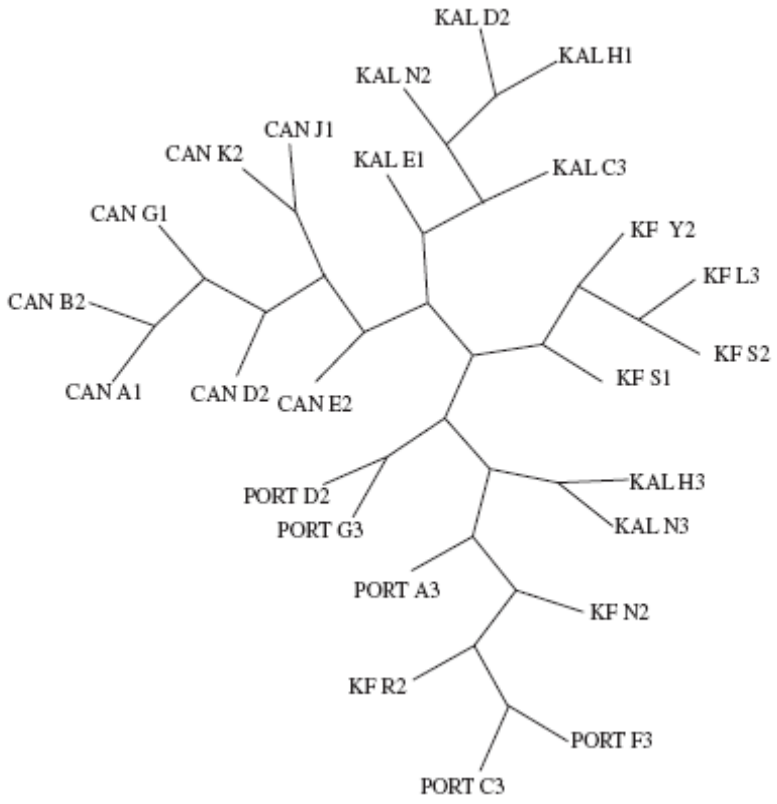


Fig. 20. Neighbor-joining unrooted cladogram of large subunit (D1-D3) rDNA gene sequence of *S. marinoi* (for sample abbreviations, see Table 1)

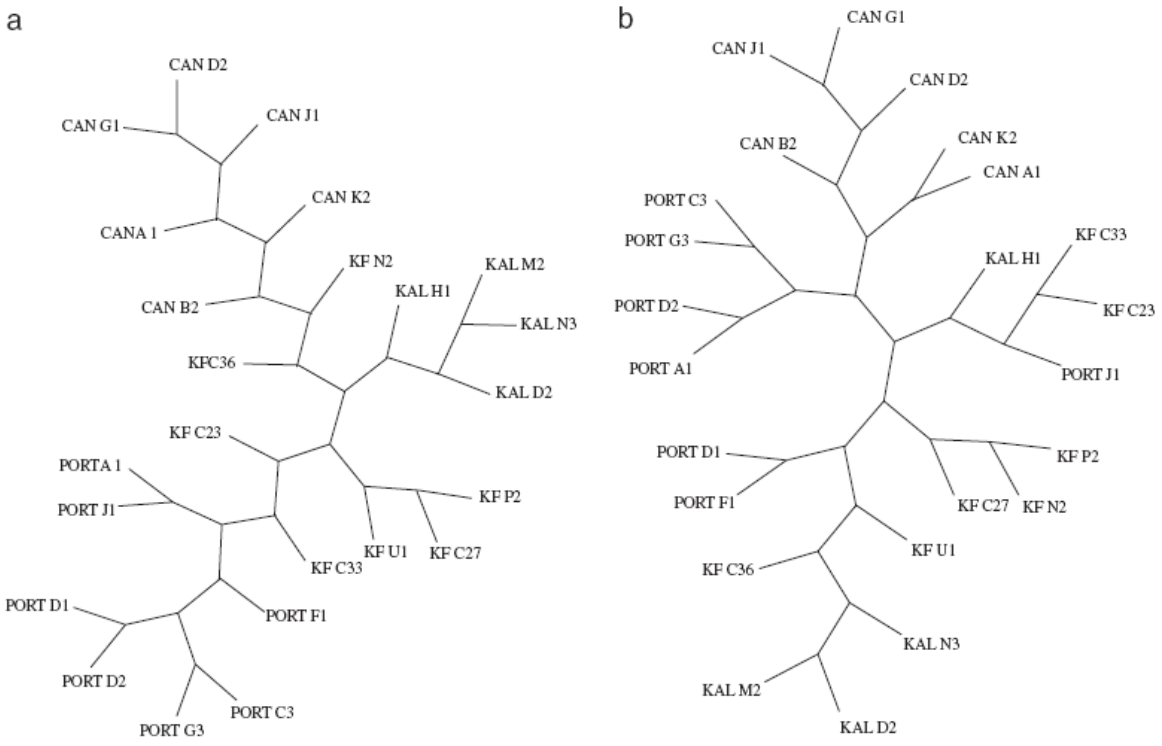


Fig. 21. (a) Neighbour-joining unrooted cladograms of internal transcribed spacer (ITS) 2 sequences of *S. marinoi*. (b) Neighbor-joining unrooted cladogram of ITS1 sequences of *S. marinoi* (for sample abbreviations see table 1)

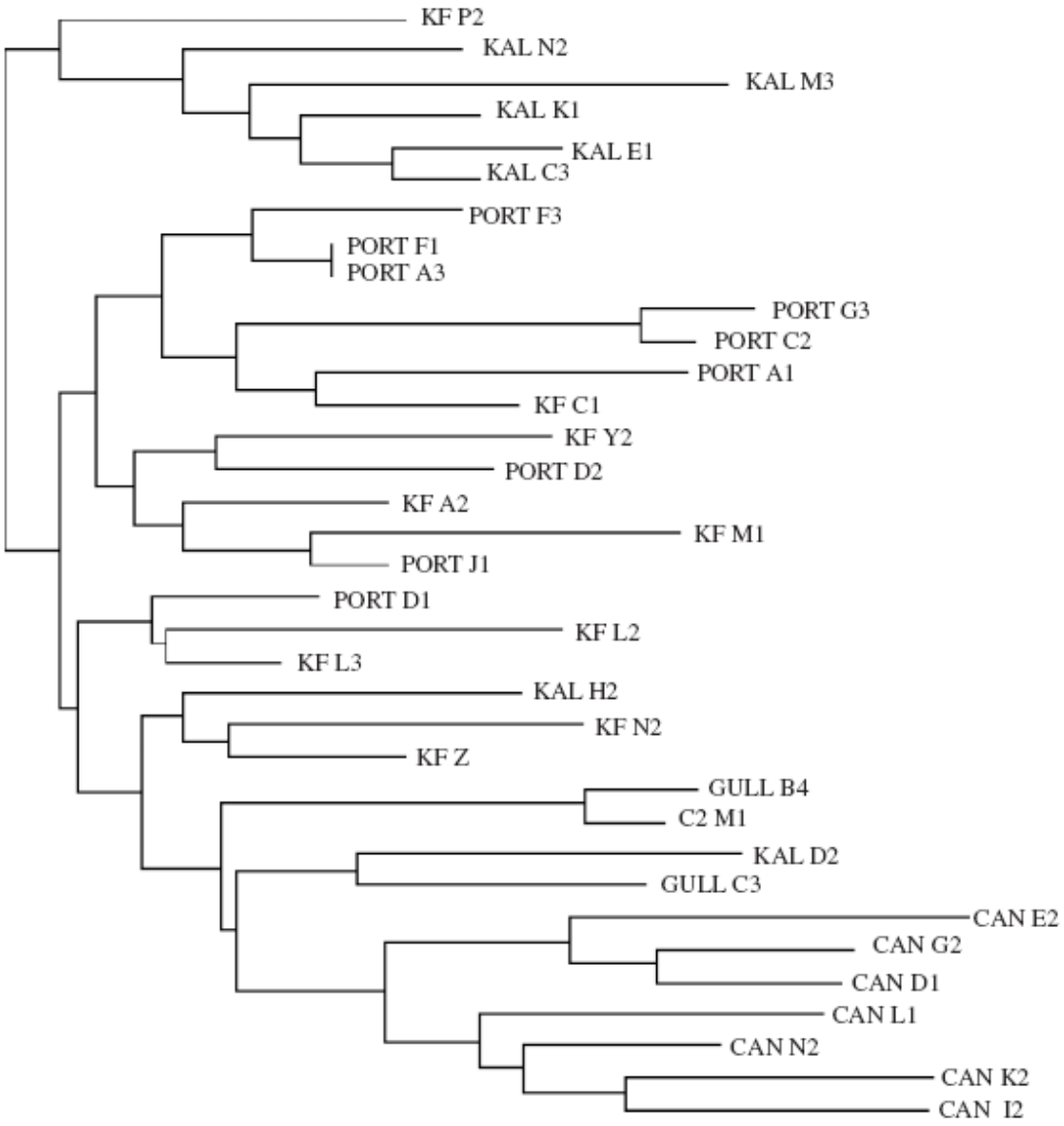


Fig. 22. Neighbor-joining unrooted phylogram of randomly amplified polymorphic DNA patterns of *S. marinoi* (for sample abbreviations, see table 1)

4.14. AFLP analysis

4.14.1. Reproducibility

The samples denoted F1A, F1A2 and F1D, F1D2 originated from the same cultured strain but was extracted on two occasions 10 days apart to assess the reproducibility of the AFLP protocol. Fig. 23 reveals the banding patterns of the replicates of strain F1A and F1D. A dendrogram created using neighbor-joining algorithm (Fig. 24) shows clustering of the replicates. A matrix (Fig. 25) was obtained from analysis using the Dice coefficient that calculated the similarity within the replicates of F1A and F1A2 and within F1D and F1D2 to 93.3% in both cases. The similarity between the replicates F1A, F1A2 and F1D, F1D2 was less than the similarity within F1A and F1A2 and within F1D and F1D2 (Fig. 25). The calculated similarity between F1A, F1A2, F1D, F1D2 and the parent Bunder G ranged from 66.8% to 75%.

4.14.2. Intercrossing experiment

The resulting banding pattern from the intercrossing experiment is seen in (Fig. 26). The separation of the parents Man 3B and Bunder G was with a similarity value of 59.3%. With 73.3% similarity F1A clustered together with Man 3B (Fig. 27 and Fig. 28). The calculated similarity between F1D and F1H was 96.3%, between F1D and F1F it was 88.9% and 92.3% between F1H and F1F (Fig. 28). Similarity values calculated between F1D, F1H, F1F and parental strain Man 3B varies from 64.5% to 66.7% and the similarity to Bunder G varies from 50 to 60.9%. F1B was more similar to the out-group Pond 6A (66.7%) than to the two parental strains Man 3B (33.3%) and Bunder G (47.1%). The similarity between F1B and the other F1-generation (F1A, F1D, F1F and F1H) ranged from 40% to 50% (Fig. 28). Similarity values between Pond 6A and Man 3B, Bunder G, F1A, F1D, F1F and F1H ranged from 31.6% to 44.4%.

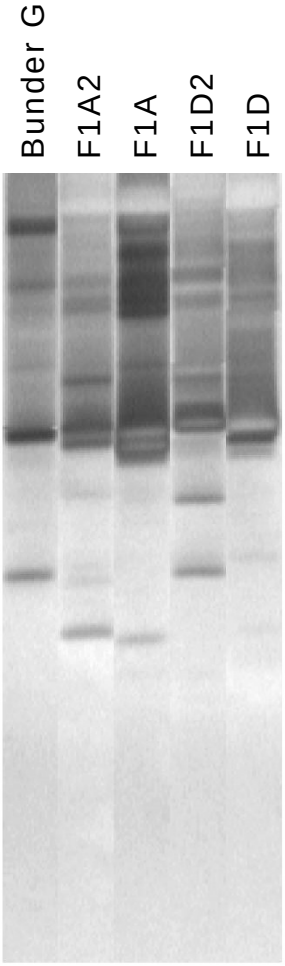


Fig. 23. AFLP banding pattern for replicates of F1A and F1D and parental strain BunderG

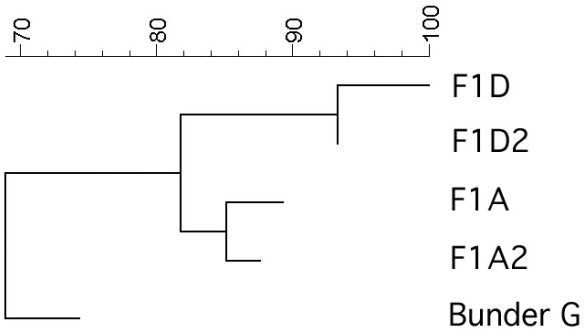


Fig. 24. Dendrogram using neighbor-joining algorithm of two clonal replicates extracted on two different occasions 10 days apart

F1D	100				
F1D2	93.3	100			
F1A	75.9	82.8	100		
F1A2	85.4	82.8	93.3	100	
Bunder G	66.8	69.6	75.0	75.0	100

Fig. 25. Similarity matrix using the Dice coefficient for clonal replicates of F1Aand F1D

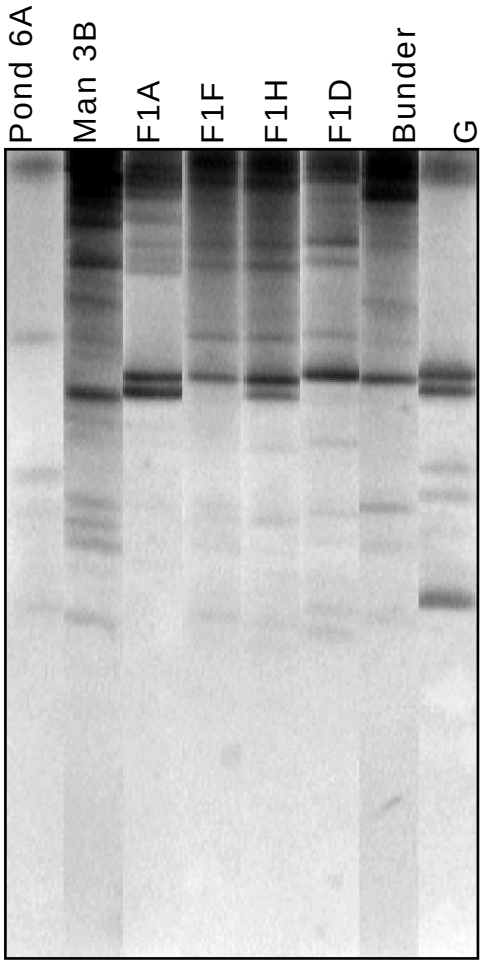


Fig. 26. AFLP banding pattern produced by primer combination EcoR1+AAG, EcoR1+ACC, Mse1+CTA and Mse1+CTT. Parental strains Man3B and Bunder G along with F1 generation and out-group Pond 6A

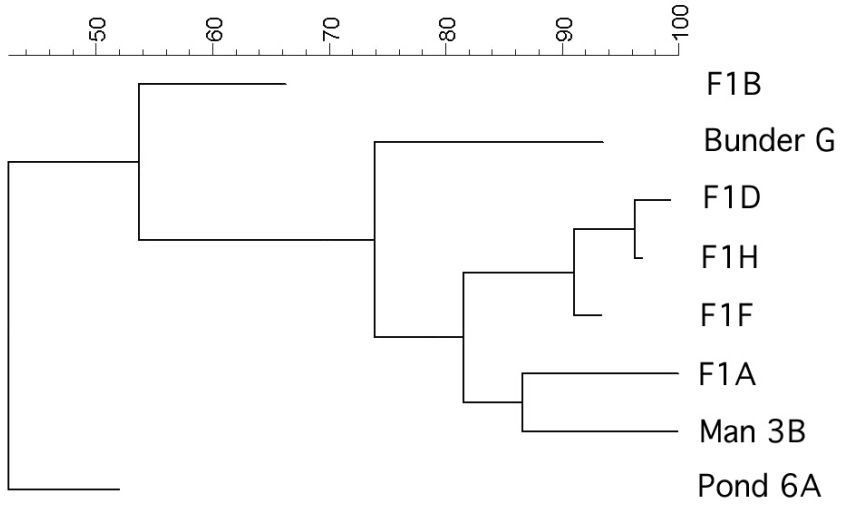


Fig. 27. Dendrogram using Neighbor-joining algorithm of parental strains Man3B and Bunder G and F1 generation F1A, F1B, F1D, F1F and F1H

F1B	100							
Bunder	47.1	100						
G	47.6	50.0	100					
F1D	50.0	52.2	96.3	100				
F1H	50.0	60.9	88.9	92.3	100			
F1F	40.0	52.2	66.7	69.2	69.2	100		
F1A	33.3	59.3	64.5	66.7	66.7	73.3	100	
Man 3B	66.7	40.0	31.6	33.3	44.4	33.3	36.4	100

Fig. 28. Similarity matrix calculated using the Dice co-efficient for parental strains Man 3b and Bunder G and F1 generation F1A, F1B. F1D. F1F and F1H

4.15. Genetic heterogeneity and physiological variation

The abundance of *S. marinoi* in Gullmar Fjord was highest (897 cells/ml) during the spring bloom, i.e., in February and beginning of March (Fig. 29). During the following months March to June, *S. marinoi* was detected in all samples but at very low densities (< 10 cells/ml). At the end of September and beginning of October, another peak in abundance was recorded, reaching a maximum of 200 *S. marinoi* cells/ml. More than 90% of the isolated chains of *S. marinoi* grew to high densities and monoclonal cultures could be established. High proportion of survival and randomly choosing from a large pool of clones for each investigated season, reduced the probability that the outcome of the experiment was due to the selected suite of clones.

Based on abundance data from Gullmar Fjord, nine different clones were selected for the experiment: three clones (GF04 1F, GF04 1G, GF04 1J) from a typical spring bloom sample (Table 17 and Fig. 29), three clones (GF04 7C, GF04 7F, GF04 7J) from a sample displaying low cell abundance, and three clones (GF04 9A, GF04 9B, GF04 9D) from a sample collected just prior to the bloom recorded in September-October. All clones were identified as *S. marinoi* based on the LSU rDNA (D1-D3) sequences alignment and ML analysis. Base-pair substitutions confirmed genetic heterogeneity among clones (Table 30). Five clones had identical LSU sequences (GF04 1F, GF04 7C, GF04 7F, GF04 7J, GF04 9D). In contrast to this, four of the investigated clones showed sequence variation. The maximum variation found among any two pairs of clones was four base-pair substitutions. Considerable variation within the LSU rDNA sequences was seen in clones from single plankton sample. Thus, the three clones from the February 2004 plankton sample (GF04 1F, GF04 1G, GF04 1J) differed in 1–4 base pairs, and the clones from the September 2004 plankton sample (GF04 9A, GF04 9B, GF04 9D) had 2-3 polymorphic sites among them. The three clones from the June 2004 plankton sample (GF04 7C, GF04 7F, GF04 7J) had identical LSU rDNA (D1-D3) sequences. No significant relationship was observed between the

physiological parameters examined, i.e., maximum growth rates, maximum cell abundances or RNA concentration per cell.

Maximum growth rate for the different salinity-temperature combinations ranged from 1.2 (5 PSU, 3°C, GF04 9B) to 1.9 (15 PSU, 17°C, GF04 1G) /doublings(per day). Maximum growth rates were significantly different (ANOVA) with regard to salinity ($p < 0.01$), temperature ($p < 0.01$), isolation month ($p < 0.05$), and the interaction between temperature and isolation month ($p < 0.01$). Cultures grown in the lowest salinity, 5 PSU, had significantly slower maximum growth rate than higher salinity, whereas no difference was found between the higher, 15, 26 and 35 PSU, salinities ($p < 0.05$) (Fig. 30A). Cultures grown in 3°C, had significantly slower maximum growth rate than cultures grown in 10°C, and cultures grown in 17°C had significantly higher maximum growth rate than the two lower temperatures ($p < 0.05$) (Fig. 30B). The clones isolated in February (GF04 1F, GF04 1G, GF04 1J) had significantly higher maximum growth rate (mean 1.40 /doublings (per day), (n=36) than the clones isolated in June (mean 1.27 /doublings (per day), (n=36) and September (mean 1.32 /doublings (per day), (n=36), as calculated by SNK test ($p < 0.05$) (Fig. 30C).

To investigate the effect of the interaction between temperature and isolation month, one-way ANOVA was conducted for each temperature. The test indicated that growth rate was significantly different ($p < 0.01$) only at 17°C. All clones had statistically similar maximum growth rate in the 3°C cultures, whereas in 10 and 17°C, pair-wise comparisons were significantly different between the following pairs of clones: GF04 1F and GF04 1G displayed significantly higher maximum growth rate than GF04 7F, GF04 7J and GF04 9A ($p < 0.05$) (Fig. 31). The variation in growth rates among clones isolated in June was considerably lower than the variation in growth rates among clones isolated in February, irrespective of temperature. The variation in growth rates among the June clones was less than the variation among the September clones in 3 and 17°C (Fig. 32A-C).

Maximum cell abundance for the different salinity-temperature combinations ranged from 6.4 x 10^4 (5 PSU, 3°C, GF04 1F) to 4.6 x 10^6 (26 PSU, 10°C, GF04 9A) cells/ml. There was no significant relation between the initial cell abundances and the maximum cell abundances of respective clone. Maximum cell abundances were significantly different (ANOVA) with regard to salinity ($p < 0.01$), temperature ($p < 0.01$) and sampling month ($p < 0.05$). Cultures grown in the lowest salinity 5 PSU had significantly lower maximum cell density than higher salinity, whereas no significant difference was found between the higher 15, 26 and 35 PSU salinities ($p < 0.05$) (Fig. 33A). Cultures grown at 10°C had significantly higher maximum cell density than cultures grown at 3 and 17°C ($p < 0.05$) (Fig. 33B). The clones isolated in September (GF04 9A, GF04 9B, GF04 9D) displayed significantly higher maximum cell number (mean 2.0 x 10^6 cells/ml, n=36) than the clones isolated in June (mean 1.69 x 10^6 cells/ml, n=36) and February (mean 1.62 x 10^6 cells/ml, n=36) as calculated by SNK test ($p < 0.05$) (Fig. 33C). The variation in maximum cell abundances among clones isolated in September was considerably larger than the variation in growth rates among clones isolated in February and June, at 3 and 17°C. The variation in maximum cell abundances among the February clones was larger than the variation among the September and June clones at 10°C (Fig. 34A-C).

RNA concentration per cell for the different salinity-temperature combinations ranged from 24 (26 PSU, 17°C, GF04 1J) to 4444 (5 PSU, 3°C, GF04 9B) pg/cell. RNA concentrations per cell were significantly different (ANOVA) with respect to salinity ($p < 0.01$), temperature ($p < 0.01$), and month ($p < 0.05$).

Cultures grown in the lowest salinity 5 PSU had significantly higher RNA concentration per cell than cultures grown in more saline media, whereas no significant difference was found among 15, 26 and 35 PSU salinities ($p < 0.05$) (Fig. 35A). Cultures grown in 17°C had significantly lower RNA concentration per cell than cultures grown at 10°C and cultures grown in 5°C had significantly higher RNA concentration per cell than any of

the two higher temperatures ($p < 0.05$) (Fig. 35B). The clones isolated in June had significantly higher RNA concentration per cell than the clones isolated in September. The RNA concentration of the February clones were not statistically different from the June or the September clones ($p < 0.05$) (Fig. 35C). The variation in maximum RNA concentration per cell among clones isolated in September was considerably less than the variation observed among the clones isolated in February and June at 3°C. The variations in maximum RNA concentration per cell among clones isolated in September and February was larger than the variation among the June clones at 10°C. At 17°C, the variation in maximum RNA concentration per cell among clones isolated in February was larger than the variations among the June and September clones (Fig. 36A-C).

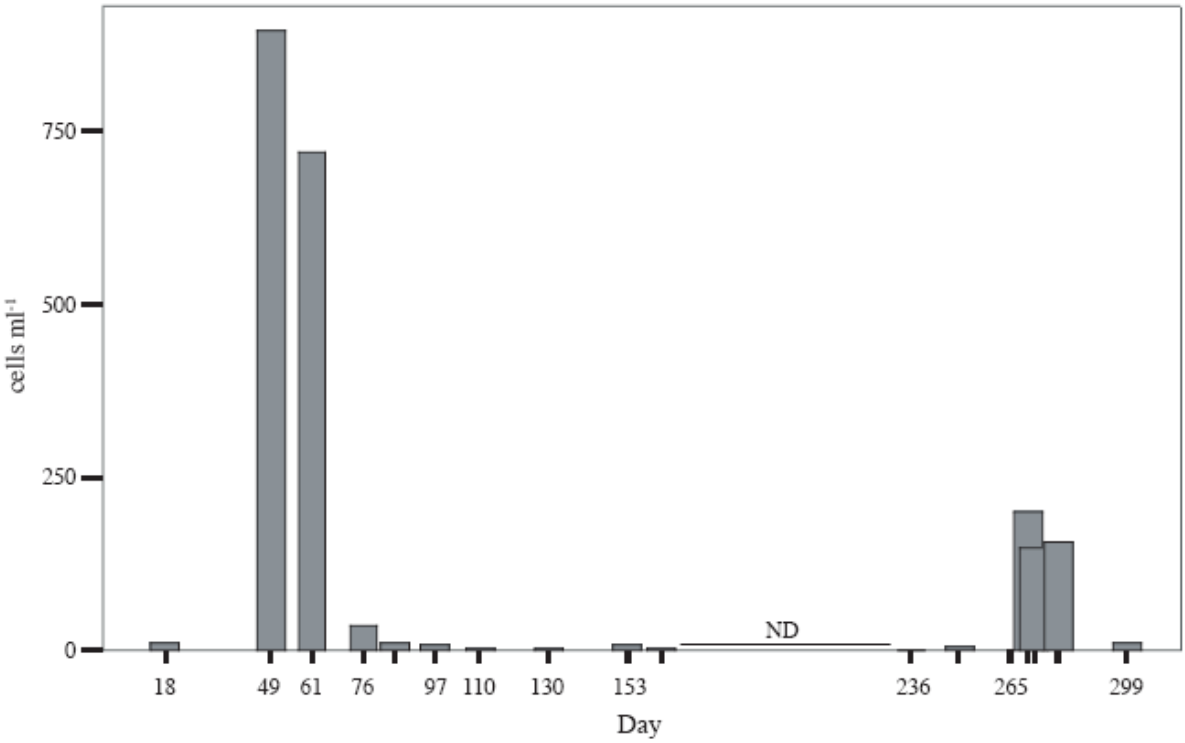


Fig. 29. Abundance of *S. marinoi* cells in discrete water samples collected from Gullmar Fjord
Note = ND – No data. The clones used in the experiment originating from day: 49, 153 and 265.

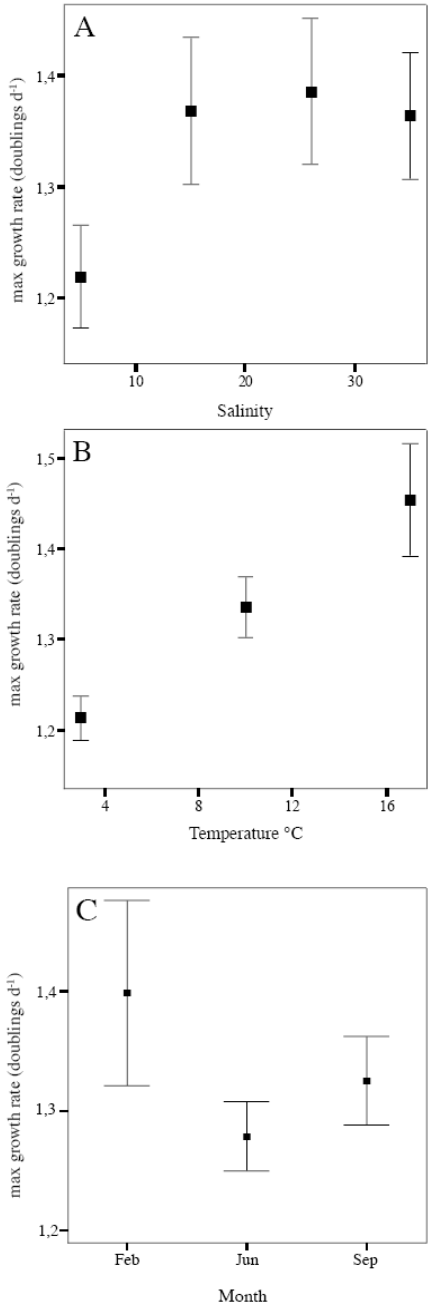


Fig. 30. Maximum growth rates recorded in the different A. Salinities. B. Temperatures. C. Isolation month (Months). Symbols are means (salinity n=27, temperature n=36, month n=36). Error bars represent 95% confidence interval (CI).

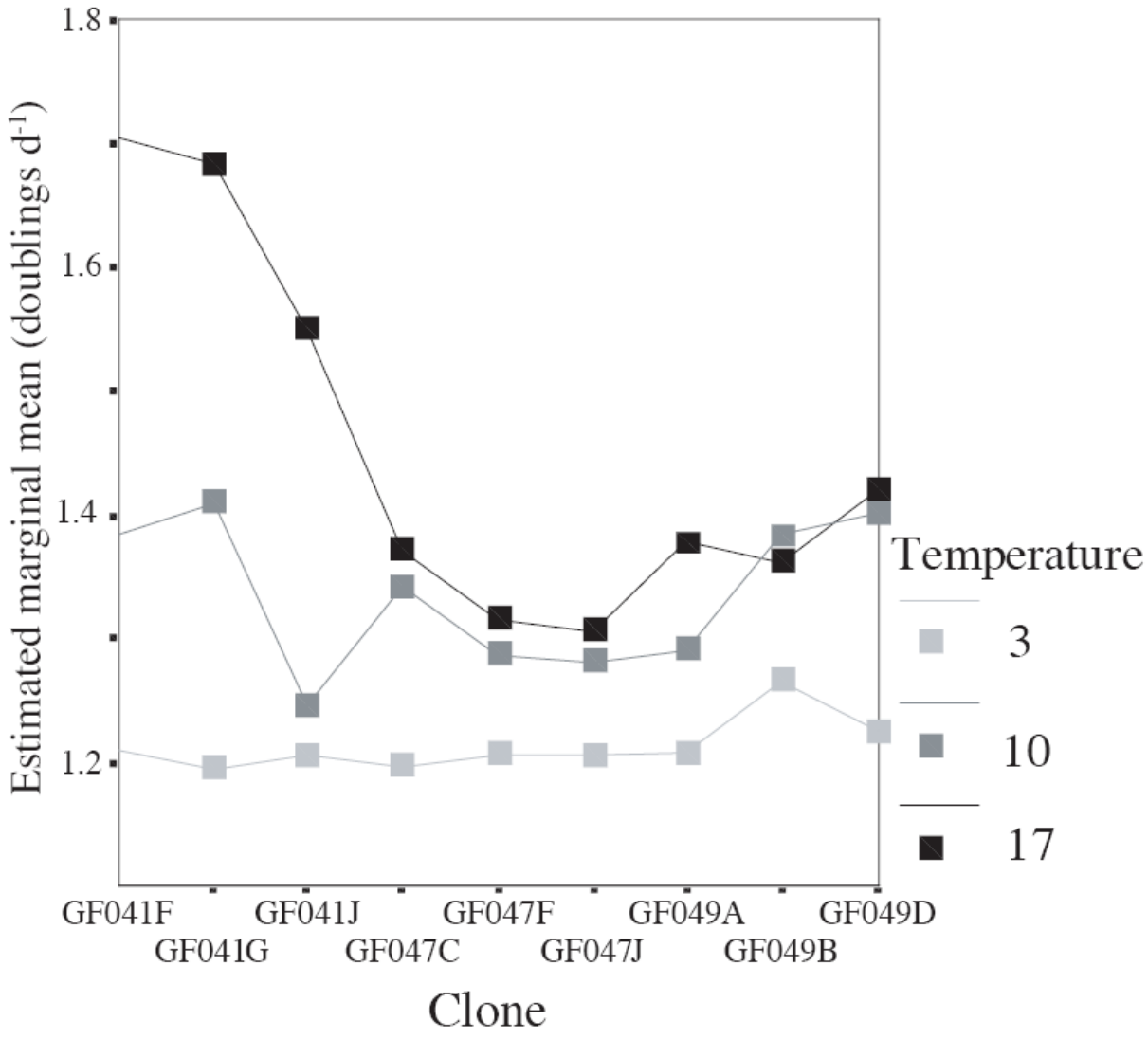


Fig. 31. Estimated marginal means of maximum growth rate per clone in the different temperatures. Symbols represent means (n=4). Lines group clonal growth rate into different temperatures

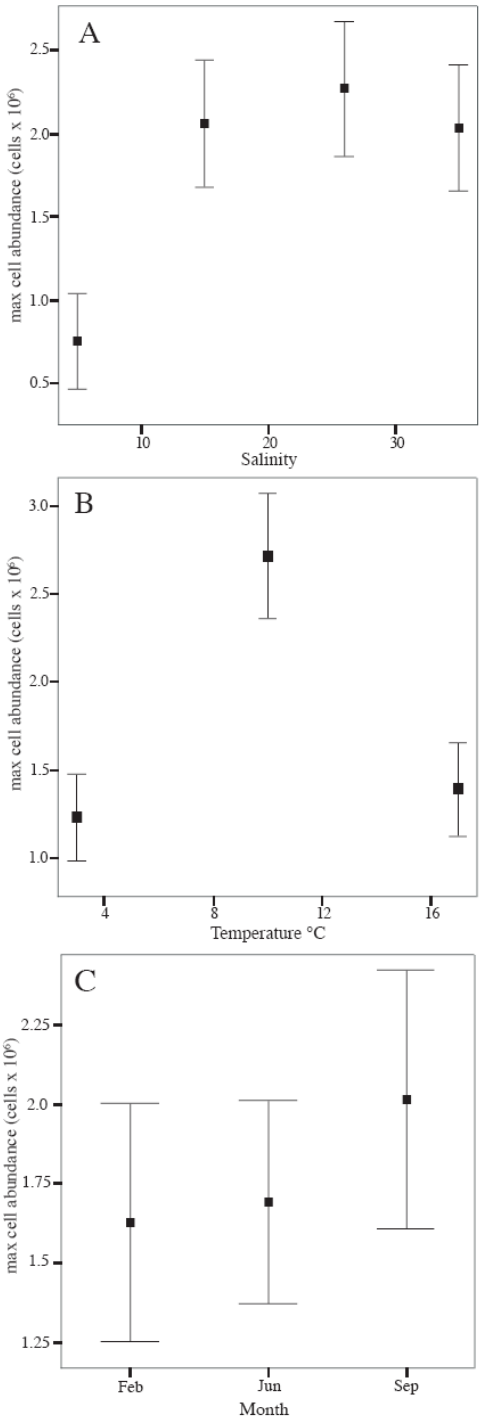


Fig. 32. Maximum cell abundance in different A. Salinities, B. Temperatures, C. Isolation months. Symbols are means (salinity n=27, temperature n=36, month n=36). Error bars represent 95% confidence interval (CI)

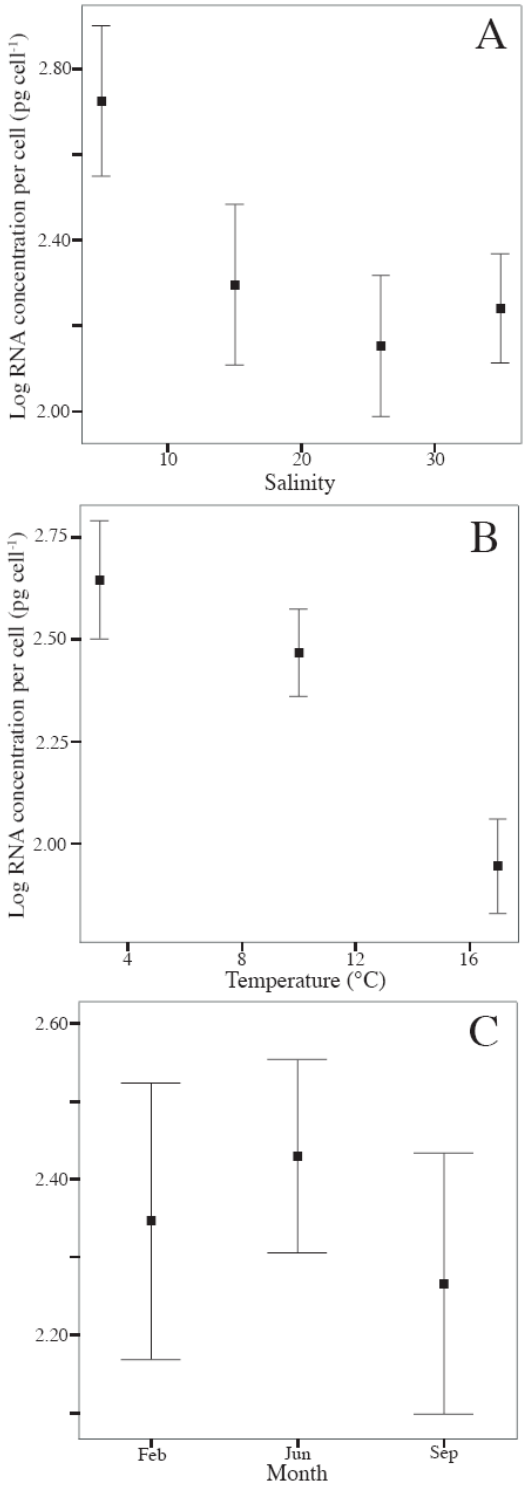
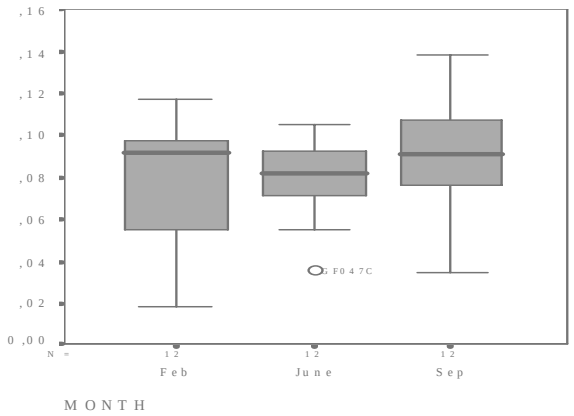


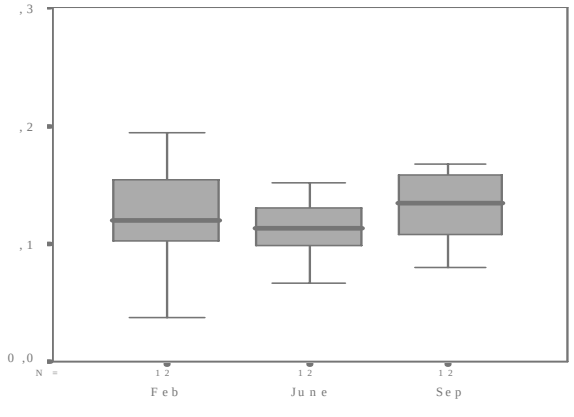
Fig. 33. Maximum RNA concentration per cell in different A. Salinities, B. Temperatures, C. Isolation months. Symbols are means (Salinity n=27, temperature n=36, month n=36) Error bars represent 95% confidence interval (CI)

Max growth rate

A. **3°C**



B. **10°C**



C. **17°C**

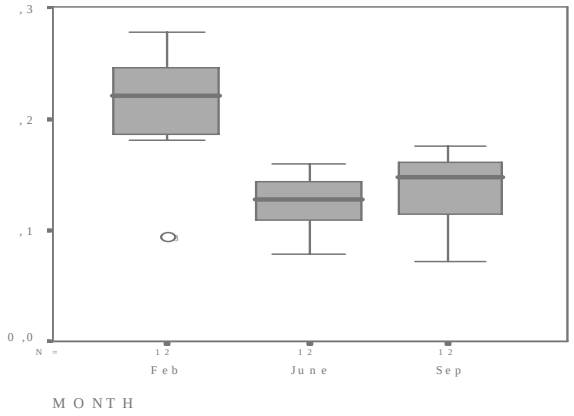
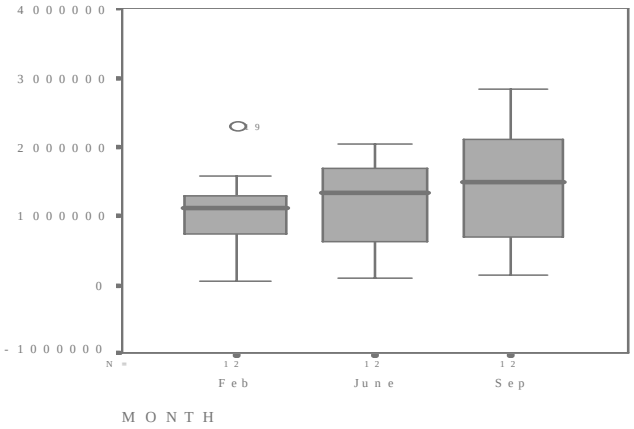


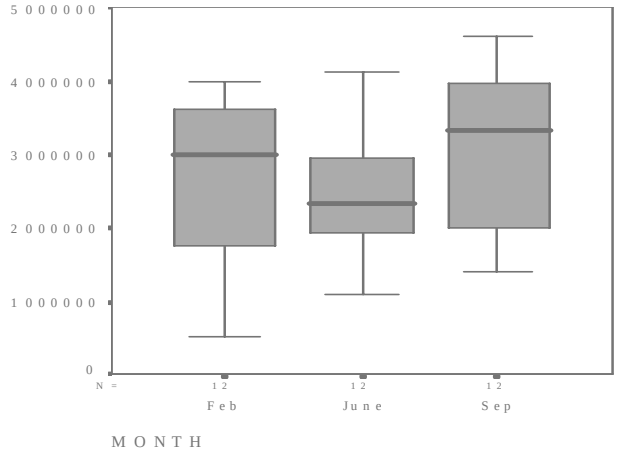
Fig. 34. Variation in maximum growth rates among the clones from the same isolation month. Black bars indicate median value. Error bars represent maximum and minimum observed results. Boxes indicate standard error (N=3). Statistical outliers are represented by o. A. 3°C. B. 10°C. C. 17°C

Max cell abundance

A. **3°C**



B. **10°C**



C. **17°C**

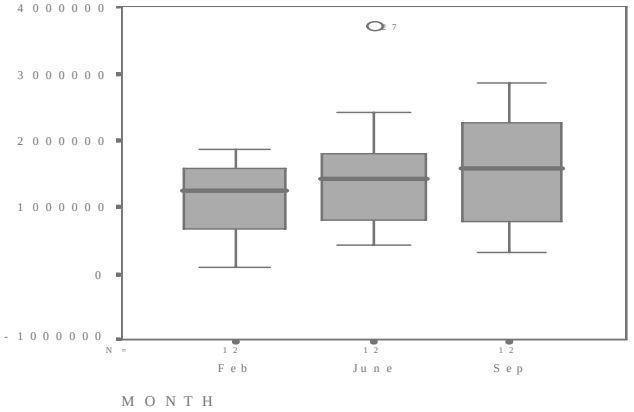
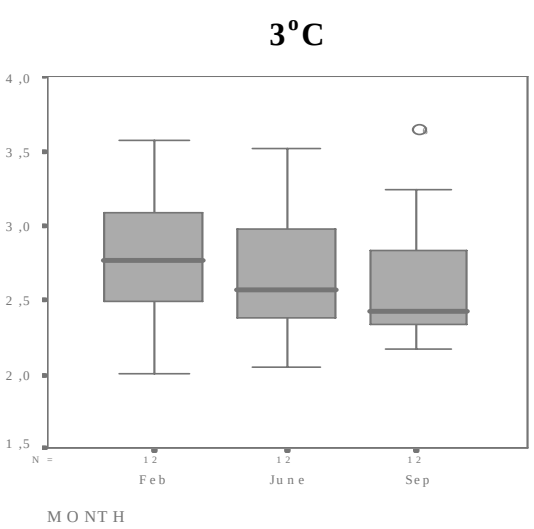
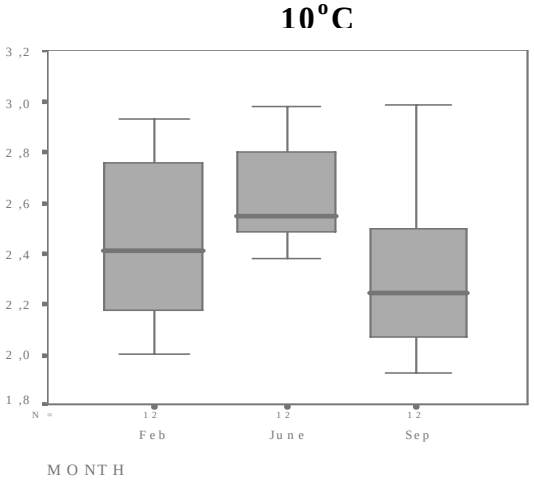


Fig. 35. Variation in maximum cell abundances among the clones from the same isolation month. Black bars indicate median value. Error bars represent maximum and minimum observed results. Boxes indicate standard error (N=3). Statistical outliers are represented by o. A. 3°C. B. 10°C. C. 17°C.

A.



B.



C.

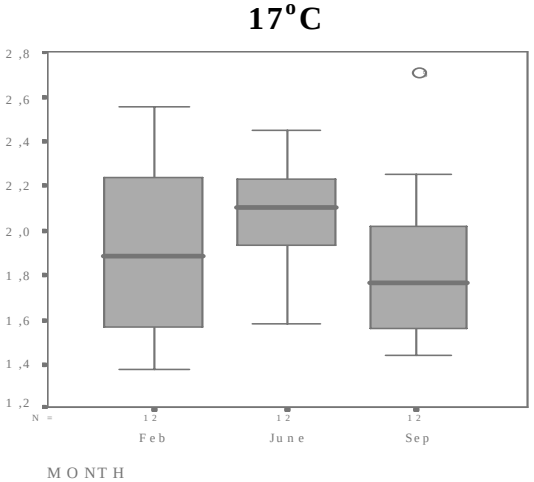


Fig. 36. Variation in in maximum RNA concentration per cell among the clones from the same isolation month. Black bars indicate median value. Error bars represent maximum and minimum observed results. Boxes indicate standard error (N=3). Statistical outliers are represented by o. A. 3°C. B. 10°C. C. 17°C.

V. DISCUSSION

V. DISCUSSION

5.1. Microsatellite genetic markers in *S. marinoi* genome

Marine planktonic diatoms are characterized by their global distribution, with individual species being able to flourish in a range of dramatically different environments (Guillard and Kilham, 1977). One possible explanation for the broad distribution is the fact that few physical boundaries exist over vast areas of the ocean, leading to continuous large-scale dispersal of individuals (Finlay, 2002). Population differentiation has been observed in marine metazoans such as copepods, which like the diatoms spend all or part of their life cycle as plankton (Hillbish, 1996; Bucklin *et al.*, 2000). Genetic diversity within a phytoplankton species is advantageous because it ensures plasticity and ultimately survival of a particular species in a changing environment. Molecular methods have greatly increased our ability to estimate the vast genetic variation within plankton communities and this information would lead to improved prediction and modeling of planktonic blooms (Savin *et al.*, 2004). The usefulness of the genetic information however depends on the quality of the genetic material examined, sample size, frequency and resolution of the molecular method used.

Microsatellites are neutral markers that are scattered throughout the genome of all eukaryotes including diatoms. They have become increasingly popular for detecting genetic variation within populations of multicellular organisms (Bruford and Wayne, 1993). Microsatellites consist of di, tri and tetra-nucleotides that are randomly repeated tens to hundreds of times within the genome (Goldstein and Pollock, 1997). The number of repeat units at a given microsatellite locus varies dramatically between individuals and hence, can be used as a tool for discrimination. This variation in length is predicted to be a result of a process known as strand slippage, which occurs when the DNA synthesizing mechanism slips during replication of the repetitive regions (Schlötterer and Tautz, 1992). The fingerprint of an individual becomes increasingly precise as more and more microsatellite loci are

analyzed since the repeat array or allele size acts as a part of the DNA fingerprint.

One of the major drawbacks of microsatellites is that they need to be isolated *de novo* from most species being examined for the first time. Microsatellites are usually found in the noncoding regions where the nucleotide substitution rate is higher than that in coding regions. Over the last few years, various methods of microsatellite isolation have been developed. Most authors in the past have used the traditional methods of screening microsatellites from partial genomic libraries of species of interest, screening several thousands of clones through colony hybridization with repeat-containing probes (Rassman *et al.*, 1991). However, the traditional method was found to be inefficient for species with less microsatellites or when a large number of microsatellites were required, as in the case of studies involving genetic diversity among populations. Several alternative strategies have been devised to reduce the time involved in microsatellite isolation and also increase the yield obtained. Microsatellite enrichment methods using biotinylated probes to capture repetitive coated magnetic beads were developed, in order to construct genomic libraries (Kandpal *et al.*, 1994).

The diatom *S. marinoi* is present in the planktonic form all along the west coast of Sweden throughout the spring and summer seasons. The ability to multiply and establish under different environmental conditions suggests high genetic diversity. These diatoms are abundant in sediment as resting stages which can survive for several decades; if resuspended they possess the potential to fuel blooms. *S. marinoi* was chosen as a model species to investigate the population genetics of planktonic blooms due to its ability to form resting stages, its wide geographic distribution, dominance under many conditions and the fact that they are easy to isolate and maintain in culture systems. The aim was to obtain clonal molecular signatures and thereby address the question of how important resting stages are for bloom initiation and propagation.

Following the protocol of microsatellite-enrichment libraries, we have identified a tri-nucleotide and a di-nucleotide tandem repeat sequence within the *S. marinoi* genome. The enriched library was constructed using (CA)n and (GA)n biotin-labeled probes. For successful construction of microsatellite-enriched libraries, a few points have to be taken into consideration. One of the oligos (OligoA) in the adaptor sequence must be phosphorlyated at the 5'end; otherwise the adaptor-ligated DNA molecules would carry a nick in one of the two strands resulting in the loss of the adaptor sequence. So the nick should be repaired with Taq DNA polymerase at 72°C for five minutes before DNA denaturation. Biotinylated microsatellite probes bound to the streptavidin-coated magnetic beads greatly enhances the sensitivity of detection by reducing the number of PCR cycles required for amplification. By using this enrichment technique, it is possible to obtain higher numbers of positive recombinant clones that have the desired microsatellite repeats. The probability of the sequenced clones developing into microsatellite markers depends on the number of larger adaptor-ligated DNA fragments selected. Larger DNA fragments easily fulfill the criteria for primer design because of sufficient flanking sequence present (Carleton *et al.*, 2001). The potential for population differentiation in marine planktonic diatoms has been discussed extensively by Brand (1990) and Mann (1999).

Few empirical studies have investigated the spatial and temporal patterns of gene variation with individual diatom species. Gallagher (1980; 1982) found that the genetic and physiological compositions of *Skeletonema costatum* in Narragansett Bay varied over a seasonal cycle. In recent years, evidence has begun to emerge indicating that the genetic composition of a species may vary over large spatial scales. DNA sequence divergence in the β-tubulin gene was observed between clones of *Thalassiosira weissflogii* isolated from different ocean basins (Armbrust and Galindo, 2001). These studies suggest that increasing geographic distance could lead to genetic differentiation.

Development of microsatellite markers involves a large amount of upfront investment and effort. Each microsatellite locus has to be identified and its flanking region sequenced for the design of PCR primers. Microsatellite-enriched genomic DNA libraries are constructed for efficient marker development (Ostrander *et al.,* 1992; Kijas *et al.,* 1994). The success of microsatellite marker development depends on fragment size selection, since every microsatellite clone requires sufficient flanking sequence between the repeat and the putative primer site that would help design primers for PCR. However, if the fragments are too large, then it is considerably difficult to clone and sequence the fragment selected. Thus microsatellite enrichment techniques are used in order to obtain large number of clones containing repeat sequences. However, the stringency of hybridization and washing plays a critical role in obtaining the desired results.

S. marinoi genomic DNA extractions were performed using phenol-chloroform-isoamyl alcohol and CTAB methods. Measurements done after extraction showed that the concentration of DNA was higher in the samples extracted with phenol-chloroform-isoamyl alcohol compared to CTAB. This may be due to the inability of CTAB to break open the silica frustules of *S. marinoi* cells when compared to phenol-chloroform-isoamyl alcohol, which is a strong and hazardous mixture and is therefore able to open up the cells and release the genomic DNA effectively. The partial genomic library of *S. marinoi* was constructed using whole genomic DNA restriction enzyme digestion with Sau3AI restriction enzyme, ligation to suitable adaptor sequence, hybridization with universal $(CA)_{15}$ AND $(GA)_{15}$ biotinylated microsatellite probes and subsequent streptavidin-coated magnetic bead capture and cloning of the microsatellite repeat containing DNA fragments into plasmid vectors. After sequencing several fragments, we obtained two potential loci containing microsatellite repeat sequences, which were highly stable and polymorphic in nature. The locus AU125 contained a tri-nucleotide repeat with GAT motif repeated nine times consecutively (Table 19). The

other microsatellite locus J33 obtained contained a di-nucleotide repeat with CT motif repeated 10 times consecutively (Table 20).

The forward and reverse primers designed amplified the tri-nucleotide repeat with a (GAT)n at the locus AU125 at positions 156/23MER and 317/21MER, respectively. The melting temperature (Tm) for the forward primer was 70.6°C and that of the reverse primer was 70.9°C. The GC (%) of the upper and lower primer was 47.8 and 52.4, respectively. The forward and reverse primers developed for the locus AU125 amplifying a tri-nucleotide repeat motif (GAT) gave a PCR product of 182 base-pair length. The optimum annealing temperature for the PCR amplification using the upper and lower primers developed for locus AU125 was determined to be 54.8°C. The forward and reverse primers designed amplified a di-nucleotide repeat with a (CT)n motif for the locus J33 at a positions of 103/23MER and 203/19MER, respectively. The melting temperature (Tm) for the forward primer was 64.9°C and that of the reverse primer was determined to be 64.8°C. The GC (%) of the upper and lower primer was 43.5 and 52.6, respectively. The upper and lower primers developed for locus J33 amplifying a di-nucleotide repeat motif $(CT)_{10}$ gave a PCR product size of 119 base-pairs. The optimum temperature for PCR amplification using the upper and lower primers synthesized for locus J33 was determined to be 51.4°C.

In this study, from the sequence data that was generated, primers were designed for microsatellite regions and the developed primers were standardized for the identified microsatellite loci. Thus, for the first time we have developed two highly stable and polymorphic microsatellite loci in the *S. marinoi* genome. In this study, we have used several *S. marinoi* samples to determine the number of alleles and also allele positions at the locus AU125 and locus J33 (Table 23 and 24 respectively). The microsatellite techniques open new perspective for studying structure of closely related populations, population samples over a wide geographic scale and less isolated populations (Estoup *et al.*, 1998).

Similar studies using microsatellite markers have shown extremely high levels of genetic diversity within populations of the diatoms *Ditylum brightwelli* (Rynearson and Armbrust, 2000; 2004; 2005) and *Pseudonitzschia multiseries* (Evans *et al.,* 2004). Results from these studies indicate that despite the potential for widespread dispersal by water currents, distinct populations can be maintained on both large and small geographic scales (Rynearson and Armbrust, 2004; 2005). Microsatellite markers have revealed that morphologically identical phytoplankton populations can harbour vast genetic and physiological differences.

Intraspecific genetic and physiological variations over different spatial scales are also indicated among other groups of phytoplankton. Significant differences in maximum growth rates were observed in isolates of the coccolithophorid *Gephyrocapsa oceanica* samples from the Sargasso Sea (Brand, 1982). Some studies involving genetic variation among strains of the dinoflagellate *Gymnodinium catenatum* (Bolch *et al.,* 1999) and the coccolithophorid *Emiliania huxleyi* (Medlin *et al.,* 1996) indicated that isolates originating from the same geographic location were more closely related to each other compared to isolates originating from distant locations.

The identification of genetically distinct diatom populations indicates that the potential for high rates of dispersal does not necessarily lead to genetic homogenization in planktonic organisms. Differential selection in combination with weak physical retention such as recirculation can produce distinct populations with unique physiological characteristics. Large-scale oceanic features such as mid-ocean gyres may also provide conditions suitable for differentiation between phytoplankton populations. The cosmopolitan distribution of many phytoplankton species has been suggested to be a consequence of extensive mixing in the marine environment (Finlay, 2002). Instead the formation of distinct populations, despite the effects of mixing, may be the mechanism allowing individual species to occupy, adapt to and then bloom over a wide range of environments. Finally, the extensive genetic and physiological differentiation indicates that genetic exchange

between populations can be limited thus providing a mechanism for the development of reproductive isolation and eventually speciation in the planktonic organisms.

5.2. LSU, ITS and RAPD analysis

Genetic diversity within a phytoplankton species is advantageous because it ensures plasticity and ultimately, the survival of a particular species in a changing environment. Molecular methods have greatly increased our ability to estimate the vast genetic variation within planktonic communities (Savin *et al.*, 2004), and this information will lead to improved prediction and modeling of planktonic blooms. The usefulness of this genetic information, however, depends on the quality of the genetic material examined, sample size and frequency, and especially the resolution of the molecular method used. All three of the molecular methods examined in this study could separate clones isolated from different oceans, but ITS 2 attributed the highest percentage to among-sample variability, and thus provided the best resolution for separating different geographical isolates of *S. marinoi*.

5.2.1. Taxonomic considerations

Recent studies have examined *Skeletonema costatum*-like species using electron microscopy and the SSU and LSU rRNA genes (Sarno *et al.*, 2005). Morphological and genetic differences observed in the collected material have led to the description of several new species within the genus, and most individuals formerly classified as *S. costatum* have new species designations. The connections between the fultoportulae are important species characters of the emended genus. Our clones were loosely connected by fultoportulae with flared tips. The phylogenetic analysis of LSU rDNA sequences shows that our clones clearly fall within Lineage III, the *S. marinoi* and *S. dohrnii* clade, as described in Sarno *et al.* (2005), and more particularly that they cluster with previously published sequences from *S. marinoi*. The two species in Lineage III are distinguished by morphological differences in the girdle

bands and a 3-bp difference in the LSU sequence (Sarno *et al.*, 2005). All of our isolates had the same three distinguishing basepair as the previously published sequences for *S. marinoi*. The known distribution of *S. marinoi* (Sarno *et al.*, 2005) also fits well with the geographic distribution of our samples, whereas *S. dohrnii* is reported only from the Gulf of Naples. This analysis shows that, according to the current taxonomy, our clones are all *S. marinoi*.

In contrast to the eight previously published LSU rDNA sequences of *S. marinoi*, our clones showed some sequence variation in the D1–D3 domains. Of our 25 clones, six were identical to those in Sarno *et al.* (2005) but 19 showed variations. The number of polymorphic sites ranged from 1–18 within the different sample locations and differences were most common in clones from Canada and the Swedish east coast. This variation in LSU may reflect intraspecies polymorphism, a level of diversity that is important for the establishment of meaningful species concepts in phytoplankton (Orsini *et al.*, 2002). Similar levels of polymorphism have been observed in LSU rDNA sequences from species of foraminifera (Holzmann *et al.*, 1996), the diatoms *Pseudo-nitzschia delicatissima* and *P. pseudodelicatissima* (Orsini *et al.*, 2002, Stehr *et al.*, 2002), and the dinoflagellate *Dinophysis acuminata* (Rehnstam-Holm *et al.*, 2002) and several species of *Protoperidinium* (Gribble and Anderson, 2005). Even greater levels of polymorphism (>100 bp) have been published for the related species *S. menzelii* (Sarno *et al.*, 2005). Because our clones were isolated from resting stages that had accumulated over many months, the differences in LSU sequences may also reflect temporal variation in genetic structure for each sample location.

5.2.2. Collection of material for studies of phytoplankton population genetics

Clones of *Skeletonema* isolated for this study had an extremely high rate of culture success. Isolates originating from different locations are presumably acclimated to the environmental conditions in which they were sampled. Although our sample locations ranged in salinity from 8 to 36, all clones were grown in the same culture medium with a salinity of 26. This

uniform culturing condition may have introduced a bias in our samples; however, dense cultures of almost all isolated cells could be established in these conditions and there was no difference in the percentage of survival among the different sampling areas. Because our material was freshly isolated, there should be no genetic variations due to long-term storage. Our results suggest that compared to some microalgal species (Rynearson and Armbrust, 2004; Evans *et al.*, 2005), it is relatively easy to obtain good quality genetic material from *S. marinoi* and large numbers of samples could be obtained, that make it a very good organism for further population studies.

In this study, sample size ranged from 4 to 12 clones per sample location and method. A greater number of clones is needed to statistically determine if the samples are from genetically separated populations; but to compare the sensitivity of the three methods by AMOVA, this sample size was suitable. We also increased the accuracy of the analyses by randomly selecting clones for each method. In addition to spatial variation in genetic diversity, there may also be variation throughout the growing season. Resting stages are accumulated in coastal sediments throughout the year (Josefson and Hansen, 2003) and, thus, provide an integrated sample of genetic variation. Use of planktonic cells might sometimes show a lower degree of variation for a particular sample and more frequent sampling would be required to quantify the extent of genetic variation in a given locale.

5.2.3. LSU rDNA, ITS and RAPD in phytoplankton genetics studies

The ability of a particular method to detect within-species genetic variation depends on the variability of the gene sequence. The utility of the nuclear rRNA gene for describing intrapopulation diversity decreases for more recently evolved taxa. The most conserved parts of the rRNA gene cluster, the SSU and 5.8S rRNA genes, are not suitable for studying variation within a species or species group. LSU rDNA and ITS regions are the least conserved parts of the rRNA gene cluster (Hillis *et al.*, 1996a), and are therefore more likely to detect differences at the population level.

In our study, LSU rDNA was useful in discriminating among *Skeletonema* species, but sequence differences in the D1–D3 domains could

also separate clones from the same geographic areas. Samples from Canada were clearly separated, whereas the European samples were not. The five clones analyzed from the Portuguese sample displayed only one polymorphic site and a few clones were identical with respect to the LSU rDNA. The same clones were not identical in their ITS sequences and RAPD. This can be explained by the higher mutation rate of the non-coding ITS region (Hillis *et al.*, 1996b). Our results confirm that the LSU gene does not satisfactorily separate clones of *S. marinoi* on small spatial scales.

ITS sequences have been used more extensively than LSU for intraspecific genetic studies of microalgae, although the diversity of this region seems to vary for different microalgal species. For the dinoflagellate *Karenia brevis*, ITS sequences of strains sampled at vast geographical distances in the Gulf of Mexico were identical (Loret *et al.*, 2002). Identical ITS sequences have also been found in isolates of the dinoflagellate *Alexandrium minutum* from the Mediterranean coast (Vila *et al.*, 2005). In contrast, variations in ITS sequences could be used to separate European and Asian clones of two epiphytic dinoflagellates (Penna *et al.*, 2005) and to distinguish among isolates of the limnic diatom *Stephanodiscus niagarae* from interconnected lakes within Yellowstone National Park, USA (Zechman *et al.*, 1994). Moreover, different ITS types of the symbiotic dinoflagellate *Symbiodinium microadriaticum* were not only restricted to certain geographic areas, but also to certain depths and specific host taxa (Rodriguez-Lanetty *et al.*, 2004). Whereas variation in the ITS region is low within some species of *Pseudo-nitzschia* (Fehling *et al.*, 2004), sequence differences can help distinguish among clades of *P. delicatissima*-like species and show that genetic diversity is high during pre-bloom conditions (Lundholm *et al.*, 2003, Orsini *et al.*, 2004). Our analysis suggests that ITS 2 sequences are appropriate to discern among isolates of *S. marinoi* from different ocean basins. The percentage of variance among the sample locations was high when using the full ITS region (ITS 1—5.8S—ITS 2), but higher when only ITS 2 was used. The conservative 5.8S rRNA gene did not contribute to any of the variations, and ITS 1 was less variable than ITS 2. Of the methods

examined, ITS 2 was the most successful at separating clones from different locations within Europe.

RAPD can be used even when a low quantity of DNA is available and requires no forehand knowledge about a particular gene in the target taxon (Hadrys *et al.*, 1992). Despite these advantages, there are few reports using RAPD in phytoplankton ecology, and this is due to poor reproducibility and constraints in inferring RAPD patterns on homology and dominance (Palumbi, 1996). Therefore, RAPD used as a tool for investigating genetic variability in phytoplankton, most probably will be outdated by more reliable and modern methods. However, in this study we could satisfactorily reproduce the RAPD analysis of the 35 clones investigated. Nearly all investigated clones had a unique banding pattern and this finding demonstrates the extensive genetic diversity of phytoplankton populations. The patterns generated from the North American and European strains of *S. marinoi* could be significantly separated. The European strains, however, were not satisfactorily differentiated by this technique. The genetic variation was partitioned mainly within geographic locations (74%) and less (26%) among locations. These values are similar to previous RAPD analyses of phytoplankton (Bolch *et al.*, 1999; Shankle *et al.*, 2004) and macrophytes (Engelen *et al.*, 2001; Faugerson *et al.*, 2001; Van der Strate *et al.*, 2003). In our study two Portuguese clones generated identical banding patterns, but the same clones displayed differences in ITS sequences. These two RAPD patterns might be resolved by increasing the number of primers. Using 5–6 primers, similar analyses of microalgae have been able to identify >40 haplotypes (Lewis *et al.*, 1997; Bolch *et al.*, 1999; Shankle *et al.*, 2004) and show differences in the genetic composition of temporally separated blooms (Shankle *et al.*, 2004) or spatially separated populations (Lewis *et al.*, 1997). Our results for *S. marinoi* show that although the RAPD method could separate samples over large geographic distances, this method does not give satisfactory resolution for intraspecies studies on smaller spatial scales.

Because *S. marinoi* is a geographically widespread taxon that is adapted to varying environmental conditions, we expected it to have high genetic diversity and plasticity. Here we have shown that any of the three molecular methods tested could significantly separate European and non-European strains of *S. marinoi* and that ITS 2 sequences are the most appropriate to discern among clones from different locations within Europe. To resolve genetic differences from samples taken at smaller spatial scales will require a more sensitive technique, such as microsatellite markers (Nagai *et al.*, 2004; Rynearson and Armbrust, 2004). This comparison of three common molecular methods, the nuclear LSU rDNA (D1–D3) and ITS (ITS 1–5.8S–ITS 2) regions and RAPD showed that various levels of genetic variation can be measured in samples of *S. marinoi*. These results provide a basis for further investigations of the genetic structure of different populations, bloom dynamics, and phylogeny in *S. marinoi* and related species.

5.3. Amplified Fragment Length Polymorphism analysis

Reproducibility of the method seemed to be good, the similarity within a single clone extracted with more than one week apart is greater than that between different clones (Fig. 24 and Fig. 25). The calculated similarity for F1A compared with F1A2 and F1D with F1D2 was 93.3% in both cases. That gives an error margin of 6.7% for within clonal strains. The parental strains of Man 3B and Bunder G are separated from each other with a similarity value of 59.3%, in other words, 40.7% difference. When comparing the similarity value of Man 3B and Bunder G, to the value calculated for replicates of the same strain, it is quite safe to make the assumption that Man 3B and Bunder G do not originate from the same strain. Also, since they are sampled one year apart, the likelihood of such close relation is even smaller. F1D, F1H and F1F make up a cluster located in between the parental strains of Man 3B and Bunder G (Fig. 27). Similarity values obtained when comparing F1D, F1H and F1F to Man 3B range from 64.5% to 66.7%. Comparisons of F1D, F1H and F1F to the other parental strain of Bunder G vary from 50% to 60.9%. These similarities indicated that F1D, F1H and F1F were heterothallic recombination of the two parental strains.

The similarity between F1D and F1H is 96.3%; hence, an error margin of 5.7%. A 5.7% difference between F1D and F1H was less than the difference calculated for the replicates of clones and indicated that F1D and F1H were identical clones that are vegetative cells from a single strain. The similarity values calculated for F1F compared with F1D and F1H are high, 88.9% and 92.3%. An error margin of 7.7% to 11.1% is too high to directly state that F1F is also a vegetative propagation along with F1D and F1H, but at the same time so low that the possibility of sexual variation between them is very small. One explanation to the similarity between F1F and F1D/F1H is vegetative propagation of F1D/F1H and a consecutive new sexual recombination that led to F1F. Another explanation is bacteria; since the cultures of *S. marinoi* were not axenic, there is always the possibility that higher concentrations of bacteria in these cultures could have affected the banding pattern to some minor extent.

F1A is 73.3% similar to the parental strain of Man 3B (Fig. 28). The error margin between F1A and Man 3B is 26.7%, which is 20% higher than what would be expected from identical clones; hence, it can be assumed that F1A is not a vegetative propagation of Man 3B. There is a possibility that F1A is a homothallic recombination of Man 3B. A homothallic progeny does not necessarily have to be identical to the parent. Gametes from the same clone fuse together and genetic recombination can occur, but in such cases the deviation is much smaller than in the case of heterothallic recombination, i.e., fusion of gametes from different clones. It seems that similarity value of 93% or higher are strong indications of identical clones while similarity ranging from 73% to 90% indicates homothallic recombination. Heterothallic reproduction seems to give similarity value of 60% to 70% and non-related clones of *Skeletonema* are less than 60% similar. To confirm any interbreeding between clones of *S. marinoi*, it is necessary to do backcross experiment and observe whether genetic distances will decrease.

The genetic pattern of F1B has the highest similarity to the strain Pond 6A. In this study Pond 6A was not part of the intercrossing experiment and is only used as a reference. Speculations about genetic resemblance between F1B and Pond 6A is futile as well as the divergence between F1B and the other strains involved in the experiment. The quality and integrity of the extracted DNA is of great importance when using AFLP for analysis, since the method is based on restriction digestion and ligation of genomic DNA, this can affect the resulting fingerprinting to a great extent. For a few clones the resulting banding pattern was

just a faint smear with two or three bands and this pattern was seen again after a second DNA extraction. Marine diatoms are known to produce many secondary metabolites such as polysaccharides that can have a negative effect on restriction digestion and ligation. Isolation and cultivation of microalgae are a selection method as such and only clones well adjusted to culture conditions will survive. Most probably, this has displayed as a continuum of bands, in which clones well suited for culture conditions yield satisfying DNA extracts and other clones tend to die immediately after isolation. However, previous results suggest that compared to some other microalgal species (Evans *et al*., 2005), it is relatively easy to obtain a large number of clones and subsequently good quality genetic material from *S. marinoi* (Godhe *et al*., 2006).

One of the drawbacks of using AFLP is that this technique does not distinguish between heterozygotes and dominant homozygotes. Another source of error is the assumption that bands of equal size are homologous, which might not be the case always. Genetic variation has previously been observed in asexual lineage of both homothallic and heterothallic clonal isolates of *G. catenatum* and such variations could not be detected in this analysis.

5.4. Genetic heterogeneity and physiological variations among *S. marinoi* clones

In this study, we have demonstrated that clones of *S. marinoi* isolated at different times of the year from the same location display different degrees of heterogeneity within the LSU rDNA sequence (D1-D3), and have significantly different physiological characteristics. Differentiation of physiological response among the different clones was partly attributed to the month of isolation. The relative importance of the isolation month was dependent on physiological character examined and the experimental condition.

In diatoms, the nuclear-encoded large subunit (LSU) rRNA gene clearly contains phylogenetic information (Auwera and Wachter, 1998) and intra-species variations of LSU have been shown for several diatom species

(Lundholm *et al.*, 2002; Orsini *et al.*, 2002; Stehr *et al.*, 2002). Recent works also highlight the importance of this gene for taxonomic assignments within the genus *Skeletonema* (Sarno *et al.*, 2005; 2007). Alignment and subsequent analyses of diatom LSU sequences among global clones might be informative to separate different isolates into geographic clades, however LSU sequences are not ideal to describe *S. marinoi* populations on small temporal or geographical scales (Godhe *et al.*, 2006). By sequencing the LSU rDNA (D1-D3) in this study, we could confirm that all the clones used in the experiment all belonged to one species. Heterogeneity within the LSU sequences was detected among the clones isolated in February and September, but not among the clones isolated in June. Additional LSU sequence data from three clones isolated in February the following year, i.e., 2005, also displayed base-pair substitutions among clones from the same net haul (Ellegaard *et al.*, 2007).

Most likely, none of the isolates belonged to the same clonal lineage and thus, the differentiated physiological response in the investigated clones may indicate genetic origin, rather than environmental control. These assumptions are also applicable for the clones isolated in June, among which no base-pair substitution was detected. However, homogeneity of LSU sequences among these clones suggests that the overall genetic diversity is less compared to the February and September isolates. Rynearson and Armbrust (2004) fingerprinted more than 400 isolates of *Ditylum brightwelli* from the same estuary and found that only two isolates shared the same microsatellite signature. The same authors also investigated 587 isolates of *D. brightwelli* during a bloom lasting eleven days at the same site and found that 87% of the isolates collected during each day were genetically distinct (Rynearson and Armbrust, 2005). Indeed, all the studies designed to investigate the presence or absence of small-scale intra-specific heterogeneity within any phytoplankton species reveal high levels of genetic diversity and low probability to isolate clonal lineages even within a single water sample (Gallagher, 1980; Orsini *et al.*, 2004; Evans *et al.*, 2005; Iglesias-Rodríguez *et al.*, 2006; Nagai *et al.*, 2007).

The observed maximum growth rates were analogous to what has been previously reported for *Skeletonema* (Shi *et al.*, 2004; Takabayashi *et al.*, 2006). Maximum growth rates increased with increased temperatures irrespective of salinity, isolation months or clones (Brand, 1982; Gallagher, 1982). This was obvious from the statistical significance of the interaction between temperature and isolation month. The February isolates had substantially faster growth rates at higher temperatures, whereas no significant difference was observed among the clones at lower temperatures, this suggests that the observed growth rate is partly controlled by environmental cues. The February clones possessed an endogenous ability to considerably increase the growth rate if the conditions were favorable. *S. marinoi* is an important diatom species contributing substantially to the spring blooms in the fjord and coastal waters. When the light and hydrographic settings become favorable, the algal biomass increases several folds. Generally the Chl *a* concentrations increase from less than one to more than 10 µg/l, during a brief time period. Provided that populations possess endogenous characteristics that are advantageous for the seasonal niches they inhabit, the fast response of the February clones appear realistic. In contrast, hydrography and light conditions are more stable during late spring and early autumn, and records of Chl *a* concentrations rarely reach more than 5 µg/l. Changes in algal biomass are gradual and populations adapted for these set of conditions may not have the same advantage of fast growth rates. Intra-specific variation of maximum growth rate in genetically-separated clones pre-adapted to identical conditions has previously been reported from different marine phytoplankton species. Observed growth rates were found to be higher for the coccolithophorid *Gephyrocapsa oceanica* isolated from neritic waters and lower in clones originating in oceanic water (Brand, 1982). *Ditylum brightwelli* isolated from the same locations, but from separate water masses defined by salinity and temperature gradients during different tidal amplitudes also displayed significantly distinct growth rates (Rynearson and Armburst, 2004).

Diverse records of maximum cell densities in different clones with respect to salinity, temperature and isolation month suggest environment control and genetic differentiation as causal factors. All clones irrespective of salinity-temperature combination reached stationary phase before the experiment was terminated. If the maximum cell abundances observed were exclusively a function of the environment provided, all clones exhibiting the same conditions should reach the same cell concentrations since the limiting factors were equal for all clones. Extremely low salinity implied lower maximum cell abundance in all clones, and 5 PSU almost certainly represent the lower range of what the populations can tolerate. In contrast, a rapid shift to a new condition exhibiting approximately 10 PSU more or less did not have any significant effect. Considerably higher records of maximum cell abundances were attained at 10°C, which probably reflects the adaptations of all clones to the conditions they were grown prior to the start of the experiment, i.e., the shift to a new temperature invariably affects the maximum cell number attained negatively. The clones isolated in September reached significantly higher cell abundances compared to the clones isolated in February or June. The phytoplankton community composition and dominance of higher taxa in Gullmar Fjord is less predictable during September. Frequently, the flora is dominated by dinoflagellates, which are generally less competitive compared to diatoms. Therefore, it might be suggested that this seasonal niche provides an opportunity for the *S. marinoi* population to grow and attain a large population size, which perhaps would not be the case with the spring blooms, when the competitors constitute other quickly growing diatom species.

The total RNA concentration per cell reflects the overall activity of the cells. Consequently, isolates subjected to adverse conditions would display high RNA concentrations per cell and engage the cell's machinery at different levels, in order to handle the stress to withstand the situation. The observed effects of temperature on the RNA concentration per cell were inverse to the temperature effects on maximum growth rates and attributed to the experimental design where RNA concentration per cell was measured

only once during the experiment, i.e., after seven days duration. After one week, all clones subjected to 17°C and many of the clones subjected to 10°C had already reached stationary growth phase, whereas the clonal cultures grown at 3°C had barely been initiated or were midway through the exponential growth phase of the culture.

Interestingly RNA concentration per cell was significantly affected by isolation months. The clones isolated in June had significantly higher RNA concentration per cell than the clones isolated in September. The February clones were not statistically differentiated from the June or the September clones. The natural abundance of *S. marinoi* during the summer months is very low or non-existing, probably due to adverse environmental conditions or predation. The impact of grazers on the overall phytoplankton community is most noticeable during this time of the year. However, in a microcosm study previously conducted in Gullmar Fjord during the end of May, no effect of grazing on *S. marinoi* was observed. In contrary, *Skeletonema* generally outgrows the grazers (McQuoid and Godhe, 2004). It is therefore tempting to suggest that the environmental conditions provided in June are non-optimal for *Skeletonema* growth, but the population exploiting this seasonal niche possesses physiological characteristics to handle stressful conditions.

No statistical test was performed to compare the sequence heterogeneity among the clones from different isolation months and the variance in their physiological response. Nevertheless, the clones isolated in June had identical LSU rDNA sequences and displayed less variation in all the investigated variables, i.e., maximum growth rate, maximum cell abundance, and RNA concentration per cell. On the contrary, the February and September clones had a maximum of four and three base-pair substitutions respectively, among the clones isolated from the same plankton net haul. These clones in addition, displayed larger variations in their physiological response. The same was also observed for different ecotypes of *Prochlorococcus*. Surface ecotypes formed a distinct cluster based on SSU. The physiological response, measured as growth rate and Chl *b*/Chl a_2 ratio,

at sub saturated light intensities were uniform among the investigated clones. This contrasted the deepwater ecotypes, which branched from the surface clones, but did not form a distinct cluster in the phylogeny inferred from the SSU and displayed a non-uniform physiological response to the same irradiances (Moore and Chrisholm, 1999).

Temporal and spatial intra-specific physiological differentiation is commonly observed on smaller and larger scales in the marine environment. As a result the total biomass and production may be greater and extend over a longer period of time, than if the species represented a single undifferentiated population (Duffy and Stachowicz, 2006). Our data suggests that physiological differentiation is seen at very small scales, even within the same plankton net, which is supported by earlier studies by Gallagher (1982) and Brand (1982). The plausible causes to the differentiated physiological characteristics are environmental conditions and genetic structure.

All the clones were maintained in exactly the same conditions for approximately two years prior to the experiment and despite this they have retained their physiological characteristics. Clonal cultures kept in nutrient-depleted media are reported to mainly propagate asexually and to restore their cell size through vegetative enlargement rather than through auxospores and sexual propagation (Gallagher, 1982; Hargraves and French, 1983). This is in accordance with regular observations of the clonal cultures isolated from Gullmar Fjord, were swarming cells are extremely rare. Asexual propagation minimizes the risk of genetic drift and random mutations through unequal crossing over. Thus the physiological response coupled to particular isolation months shown in this experiment constitutes a genetic structure endogenous to the clones exploiting the particular seasonal niche. Based on the observations of this study, we would like to raise cautions for using cultures for ecologically conclusions, since the time of year and the particular population isolated might exhibit different traits compared to a clone isolated during other time of the year.

Thus, in this study, we used different molecular techniques like large subunit (LSU) rRNA genes, the internal transcriber spacer (ITS) consisting of ITS-1, 5.8S and ITS-2 regions and random amplified polymorphic DNA (RAPD) to separate *S. marinoi* clones originating from four different geographical areas; though the *S. marinoi* strains were separated by all these methods, the individuals from closely related population were best resolved by ITS-2 sequences. In addition, amplified fragment length polymorphism (AFLP) analysis was used to study genetic variations at the sub-species level by performing an intercrossing experiment. A physiological variation experiment was conducted in order to investigate, if clones of *S. marinoi* isolated from different seasons exhibited physiological differentiation and whether these differentiations could be attributed to genetic heterogeneity previously observed. Finally, two highly stable and polymorphic microsatellite loci were isolated and identified in the genome of *S. marinoi* by constructing partial genomic library using microsatellite-enrichment method. Microsatellite markers are highly sensitive in studying structures of closely related populations, population samples over wide geographic scale and less isolated populations.

VI. SUMMARY

VI. SUMMARY

Skeletonema is a diatom genus which is wide spread all over the world's oceans. It is one of the dominating phytoplankters in temperate regions. In many parts of the world, *Skeletonema* is commonly found throughout the year, but in temperate regions the highest density occurs during spring blooms. Recent studies have examined *S. costatum*-like species using electron microscopy, and the small and large subunit rRNA genes. The diatom *S. costatum* has recently been emended and several new species have been described. One of these is *S. marinoi*, which is a cosmopolitan organism that contributes significantly to phytoplankton blooms in temperate waters. *S. marinoi* is an important primary producer occurring world-wide in high abundance in both plankton and benthos; thus, scientific studies of this organism are of ecological significance.

In this study, we have identified and amplified two microsatellite loci using the primers designed from the sequence data generated. The developed primers were standardized for the identified microsatellite regions. Thus, for the first time we have developed two highly stable and polymorphic microsatellite loci in the *S. marinoi* genome. We have used several *S. marinoi* samples to determine the number of alleles and also allele positions at the two identified microsatellite loci.

In this study, we have also evaluated the utility of three common molecular methods for distinguishing clones of *S. marinoi* from different geographic regions. Clonal cultures were isolated from Canadian west coast, south-west Portugal and the east and west coasts of Sweden. All the strains originated from the resting stages in sediment. Genetic signatures of clones from each sample location were assessed by sequencing variable domains (D1-D3) of the nuclear large subunit (LSU) rRNA gene and internal transcriber spacer (ITS1, 5.8S and ITS2) regions and also by analysis of

RAPD patterns. Analysis of molecular variance showed that strains from four geographic areas were significantly separated by all the three methods but differences in European samples were best resolved by ITS 2 sequences.

Auxospore formation has been observed in single clonal cultures; hence it is known that sexual recombination occurs. But it is not known if *Skeletonema* sp is homothallic or heterothallic. AFLP was used in the analyses of genetic variation at the sub-species level. An intercrossing experiment was set up with two parental clones of *Skeletonema* sp isolated from plankton samples collected off the coast of Mangalore, India one year apart. The parental clones were mixed and from the mix, F1 generation was isolated. There was high reproducibility and results of the experiment showed a similarity value of 93% or higher which was strong indication of identical clones, while similarity ranging from 73% to 90% indicated homothallic recombination. Heterothallic reproduction gave similarity values of 60% to 70% and different clones of *Skeletonema* were less than 60% similar.

In this study, we also report the seasonal distribution of *S. marinoi* in Gullmar Fjord on the Swedish west coast. Monoclonal cultures were established throughout the year and pre-adapted to identical culturing conditions. The purpose was to investigate if clones of *S. marinoi* isolated from different seasons exhibited physiological differentiation and if it could be coupled to genetic heterogeneity. We also investigated if physiological variations among groups of clones are stable when environmental parameters are changed. The physiological parameters examined were maximum growth rates, maximum cell abundance and RNA content per cell.

VII. REFERENCES

VII. REFERENCES

ALVERSON, A. and KOLNICK, l., 2005. Intragenomic nucleotide polymorphism among small subunit (18S) rDNA paralogs in the diatom genus *Skeletonema* (Bacillariophyta). *J. Phycol.,* **41**:1248-1257.

AMOS, B. and PEMBERTON, J., 1992. DNA fingerprinting in non-human populations. *Curr. Opin. Genet. De.,* **2**: 857-860.

AMOS, B., SCHLOTTERER, C. and TAUTZ, D., 1993. Social structure of pilot whales revealed by analytical DNA profiling. *Science,* **260** : 670-672**.**

AMOS, W., SAWEER, S. J., FEAKES, R. W. and RUBINSZTEIN, D. C., 1996. Microsatellies show mutational bias and heterozygote instability. *Nature. Genet.,* **13**: 390-391.

AMOS. W. A., 1999. A comparative approach to the study on microsatellite evolution. **In**: Microsatellites: evolution and applications. GOLDSTEIN, D. B. and SCHLOTTERER, C., (Eds.), Oxford university press, Oxford, pp. 66-79.

AN, S.S., FRIEDL, T. and HEGEWALD, E., 1999. Phylogenetic relationships of *Scenedesmus* and *Scenedesmus* –like coccoid green algae as inferred from ITS – 2 rDNA sequence comparisons. *Plant Biol.,* **1:** 418-428.

ANDERSON, D.M., 1989. Toxic algal blooms and red tides: a global perspective. **In**: *Red tides biology, environmental science and toxicology*. OKAICHI, T., ANDERSON, D.M. and NEMOTO, T. (Eds.), Elsevier science publishing Co. NewYork, pp.11-16.

ANDERSON, T. J., 2004. Mapping drug resistance genes in *Plasmodium falciparum* by genome-wide association. *Curr. Drug. Targets. Infect. Disord.,* **4**: 65-78.

ARMBRUST, E.V. and GALINDO, H.M., 2001. Rapid evolution of a sexual reproduction gene in centric diatoms of the genus *Thalassiosira. Appl. Environ. Microbiol.,* **67:** 3501 – 3513.

ARNEBORG, L., 2004. Turnover times for the water above sill level in Gullmar Fjord. *Continental Shelf Res.,* **24**:443-460.

AUWERA, G. V. D. and WACHTER, R. D., 1998. Structure of the large subunit rDNA from a diatom and comparison between small and large subunit ribosomal RNA for studying Stamenopile evolution. *J. Euk. Microbiol.,* **45**:521-527.

BAKER, A.J., 2000. Molecular methods in ecology. Blackwell Science Ltd., USA.

BALLOUX, F. and LUGON-MOULIN, N., 2002 The estimation of population differentiation with microsatellite markers. *Mol. Ecol.*, **11:** 155-165.

BAYLISS, C. D., FIELD, D. and MOXON, E. R., 2001. The simple sequence contingency loci of *Haemophilis influenzae* and *Neisseria meningitides*. *J. Clin. Inves.*, **107**: 657-662.

BECKMANN, J. S. and WEBER, J. L., 1992. Survey of human and rat microsatellites. *Genomics.*, **12**: 627-631.

BEHNKE, J.M., IRAQI, F., MENGE, D., BAKER, R. L., GIBSON, J. and WAKELIN, D., 2003. Chasing the genes that control resistance to gastrointestinal nematodes. *J. Helminthol.*, **77**: 99-110.

BELL, G. I. and JURKA, J., 1997. The length distribution of perfect dimer repetitive DNA is consistent with its evolution by an unbiased single-step mutation process. *J. Mol. Evol.*, **44**: 414-421.

BHATTACHARYA, D., MEDLIN, L., WAINWRIGHT, P.O., ARIZTIA, E.V., BIBEAU, C., STICKEL, S.K. and SOGIN, M.L., 1992. Algae containing chlorophylls a + c are paraphyletic: molecular evolutionary analysis of the Chromophyta. *Evol.*, **46:** 1808-1817.

BIET, E., SUN, J. and DUTREIX, M., 1999. Conserved sequence preference in DNA binding among recombinant proteins: abnormal effect of ssDNA secondary structure. *Nucleic. Acids. Res.*, **27**: 596-600.

BLANKENSHIP, S. M., MAY, B. and HEDGECOCK, D., 2002. Evolution of a perfect simple sequence repeat locus in the context of its flanking sequence. *Mol. Biol. Evol.*, **19**: 1943-1951.

BLOVIN, M. S., 2003. DNA-based methods for pedigree reconstruction and kinship analysis in natural populations. *Trends Ecol. Evol.*, **18**: 503-511.

BOLCH, C.J.S., BLACKBURN, S.I., HALLEGRAEFF, G.M. and VAILLANCOCURT, R.E., 1999. Genetic variation among strains of the toxic dinoflagellate *Gymnodinium catenatum* (Dinophyceae). *J. Phycol.*, **35:** 356-367.

BOTSTEIN, D., WHITE, R.L., SKOLNICK, M. and DAVIS, R.W., 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.*, **32:** 314-331.

BRAND, L.E., 1982. Genetic variability and spatial patterns of genetic differentiation in the reproductive rates of the marine coccolithophores

Emiliania huxleyi and *Gephyrocapsa oceanica. Limnol. Oceanogr.*, **27:** 236-245.

BRAND, L.E., 1990. Review of genetic variation in marine phytoplankton species and the ecological implications. *Biol. Ocean.*, **6:** 397-409.

BRINKMANN, B., KLINTSCHAR, M., NEUHUBER, F., HUHNE, J. and ROLF, B., 1998. Mutation rate in human microsatellites: influence of the structure and length of the tandem repeat. *Am. J. Hum. Genet.*, **62**: 1408-1415.

BROOKER, A. L., COOK, C., BENTZEN, P., WRIGHT, J. M. and DOYLE, R. W., 1994. Organisation of microsatellites differs between mammals and cold-water teleost fishes. *Can. J. Fish Aquat. Sci.*, **51**: 1959-1966.

BRUFORD, M.W. and WAYNE, R.K., 1993. Microsatellites and their application to population genetic studies. *Curr. Opin. Genet. Dev.*, **3:** 939-943.

BRUFORD, M. W., CHEESMAN, D. J., COOTE, T., GREEN, H.A. A., HAINES, S. and ORYAN, C., 1996. Microsatellites and their application to conservation genetics. **In**: Molecular genetic approaches in conservation. SMITH, T. B. and WAYNE, R. K., (Eds.), Oxford University Press, Oxford, pp. 278-297.

BUCKLIN, A., KAARTVEDT, S., GUARNIERI, M., and GOSWAMI, U., 2000. Population genetics of drifting (*Calanus* spp) and resident (*Acartia clause*) plankton in Norwegian fjords. *J. Plankton Res.*, **22:** 1237-1251.

BURGER, J., HUMMEL, S., HERMANN, B. and HENKE, W., 1999. DNA preservation: a microsatellite-DNA study on ancient skeletal remains. *Electrophoresis.*, **20**: 1722-1728.

CAIRNEY, M., TAGGART, J. B. and HOYHEIM, B., 2000. Characterization of microsatellite and minisatellite loci in Atlantic salmon (*Salmo salar*) and cross-species amplification in other salmonids. *Mol. Ecol.*, **9**: 2175-2178.

CALLAN, D. F., THOMPSON, A. D., SHEN, Y., PHILLIPS, H. A., RICHARDS, R. I., MULLEY, J. C. and SUTHERLAND, G. R., 1993. Incidence and origin of null alleles of the (AC)n microsatellite markers. *Am. J. Hum. Genet.*, **52**: 922-927.

CANAPA, A., CERIONI, P. N., BARUCCA, M., OLMO, E. and CAPUTO, V., 2002. A centromeric satellite DNA may be involved in heterochromatin compactness in gobiid fishes. *Chromosom. Res.*, **10**: 297-304.

CARLETON, K.L., STREELMAN, J.T., LEE, B.Y., GARNHART, N., KIDD, M. and KOCHER, T.D., 2001. Rapid isolation of CA microsatellites from the tilapia genome. *Anim. Genet.,* **33**: 140-144.

CHARAMES, G. S. and BAPAT, B., 2003. Genomic instability and cancer. *Curr. Mol. Med.,* **3**: 589-596.

CHANG, D. K., METZGAR, D., WILLS, C. and BOLAND, C. R., 2001. Microsatellites in the eukaryotic DNA mismatch repair genes as modulators of evolutionary mutation rate. *Genome Res.,* 11: 1145-1146.

CHEN, F. and SUTTLE, C.A., 1995. Amplification of DNA –polymerase gene fragments from viruses infecting microalgae. *Appl. Environ.Microbiol.,* **61:** 1274-1278.

CHEN, F., SUTTLE, C.A. and SHORT, S.M., 1996. Genetic diversity in marine algal virus communities as revealed by sequence analysis of DNA polymerase genes. *Appl. Environ. Microbiol.,* **62:** 2869-2874.

CHINAIN, M., GERMAIN, M., SAKO, Y., PAUILLAC, S. and LEGRAND, A.M., 1997. Intraspecific variation in the dinoflagellate *Gambierdiscus toxicus* (Dinophyceae). I. Isozyme analysis . *J. Phycol.,* **33:** 36-43.

CHURCHILL, G.A., GIOVANNONI, J. J. and TANKSLEY, S. D., 1993. Pooled sampling makes high-resolution mapping practical with DNA markers. *Proc. Natl. Acad. Sci.* U.S.A., **90**: 16-20.

COYNE, K.J., HUTCHINS, D.A., HARE, C.E. and CARY, S.C., 2001. Assessing temporal and spatial variability in *Pfiesteria piscicida* distributions using molecular probing techniques. *Aquat. Microb. Ecol.,* **24:** 275 -285.

CRAWFORD, A.M.and CUTHBERTSON, R.P., 1996. Mutations in sheep microsatellites. *Genome. Res.,* **6:** 876-879.

CULLIS, C.A., 2002. The use of DNA polymorphisms in genetic mapping. *Genet. Eng.,* **24**: 179-189.

DEVASSY, V. P. and BHATTATHIRI, P. M. A., 1974. Phytoplankton ecology of the Cochin backwater. *Indian. J. Mar. Sci.,* **3**: 46.

DEVASSY, V. P., 1974. Observations on the bloom of a diatom *Fragilaria oceanica* cleve. *Mahasagar.,* **7:** 101.

DEVASSY, V. P., 1983. Plankton production associated with cold water incursion into the estuarine environment. *Mahasagar.,* **16**: 221.

DI RIENZO, A. A., PETERSON, J. C., CARZA, A. M., VALDES, M., SLATKIN, M. and FREIMER, N. B., 1994. Mutational processes of simple-sequence repeat loci in human populations. *Proc. Natl. Acad. Sci.* U.S.A., **91**: 3166-3170.

DIETRICH, W., KATZ, H., LINCOLN, S.E., SHIN, H.S., FRIEDMAN, J., DRACOPOLI, N. and LANDER, A.S., 1992. A genetic map of the mouse suitable for typing intraspecific crosses. *Genetics*, **131:** 423-447.

DITTMANN, E., ERHARD, M., KAEBERNICK, M., SCHELER, C., NEILAN, B.A., von DOHREN and BORNER, T., 2001. Altered expression of two light-dependent genes in a microcystin – lacking mutant of *Microcystis aeruginosa* PCC 7806. *Microbiol. Sgm.,* **147:** 3113-3119.

DIXON, M. S., JONES, D. A., HATZIXANTHIS, K., GANAL, M. W., TANKSLEY, S. D. and JONES, J. D. G., 1995. High resolution mapping of the physical location of the tomato Cf-2 gene. *Mol. Plant Microbe Interact.,* **8**: 200-206.

DOKHOLYAN, N. V., BULDYREV, S. V., HAVLIN, S. and STANLEY, H. E., 2000. Distribution of dimeric tandem repeats in non-coding and coding DNA sequences. *J. Theor. Biol.,* **202**: 273-282.

DOYLE, R.W., 1975. Upwelling, clone selection and the characteristic shape of nutrient uptake curves. *Limnol. Oceanogr.,* **20:** 487-489.

DUFFY, J. and STACHOWICZ, J., 2006. Why biodiversity is important to oceanography: potential roles of genetic, species, and trophic diversity in pelagic ecosystem processes. *Mar. Ecol. Prog. Ser.,* **311**:179-189.

DUNHAM, R. A., 2004. Aquaculture and fisheries biotechnology: genetic approaches. *CABI Publishing,* Cambridge, pp. 122-140.

EDWARDS, Y. J., ELGAR, G., CLARK, M. S. and BISHOP, M. J., 1998. The identification and characterization of microsatellites in the compact genome of the Japanese puffer fish, *Fugu rubripes*: perspectives in functional and comparative genomic analysis. *J. Mol. Biol.,* **278**: 843-854.

EDWARDS, A., HAMMOND, H. A., JIN, L., CASKEY, C. T. and CHAKRABORTHY, R., 1992. Genetic variation of five trimeric and tetrameric tandem repeat loci in four human population groups. *Genomics.,* **12**: 241-253.

EDVARDSEN, B., EIKREM, W., GREEN, J.C., ANDERSEN, R.A., DER STAAY, S. and MEDLIN, L.K., 2000. Phylogenetic reconstructions of the Haptophyta inferred from 18S ribosomal DNA sequences and available morphological data. *Phycologia.,* **39:** 19-35.

EISEN, J. A., 1999. Mechanistic basis for microsatellite instability. **In**: Microsatellite: evolution and applications. GOLDSTEIN, D. B. and SCHLOTTERER, C., (Eds.), Oxford university press, Oxford, pp. 34-48.

ELLEGAARD, M., GODHE, A., HÄRNSTRÖM, K. and MCQUOID, M., 2007. The species concept in a marine diatom: LSU rDNA variation in *Skeletonema marinoi/dohrnii* (Bacillariophyceae) is not reflected in morphology. *Phycologia* (submitted).

ELLEGREN, H., 2000. Microsatellite mutations in the germline: implications for evolutionary inference. *Trends Genet.*, **16**: 551-558.

ELLEGREN, H., 2004. Microsatellites: simple sequences with complex evolution. *Nature.*, **5**: 435-445.

ENGELEN, A. H., OLSEN, J. L., BREEMAN, A. M. and STAM, W. T., 2001. Genetic differentiation in *Sargasum polyceratium* (Fucales, Phaeophyceae) around the island of Cuaçao (Netherlands Antilles). *Mar. Biol.*, **139**:267-77.

EPPLEN, C., MELMER, G., SIEDLACZCK, I., SCHWAIGER, F. W., MAUELER, W. and EPPLEN, J. T., 1993. On the essence of “meaningless” simple repetitive DNA in eukaryote genomes. *EXS.*, **67**: 29-45.

EPPLEN, J, T., KYAS, A. and MAUELER, W., 1996. Genomic simple repetitive DNAs are targets for differential binding of nuclear proteins. *FEBS. Lett.*, **389**: 92-95.

ESTOUP, A., TAILLIEZ, C., CORNUET, J. M., SOLIGNAC, M., 1995. Size homoplasy and mutational processes of interrupted microsatellites in two bee species, *Apis mellifera* and *Bombus terrestris* (Apidae). *Mol. Biol. Evol.*, **12**: 1074-1084.

ESTOUP, A., ROUSSETT, F., MICHALAKIS, Y., CORNUET, J.M., ADRIAMANGA, M. and GUYOMARD, R., 1998. Comparative analyses of microsatellite and allozyme markers: a case study investigating microgeographic differentiation in brown trout (*Salmo trutta* L.). *Mol. Ecol.*, **7**: 339-340.

ESTOUP, A. and CORNUET, J. M., 1999. Microsatellite evolution: inferences from population data. **In**: Microsatellite: evolution and applications. GOLDSTEIN, D. B. and SCHLOTTERER, C., (Eds.), Oxford university press, Oxford, pp. 49-65.

EVANS, K., BATES, S., MEDLIN, L. K. and HAYES, P.K., 2004. Microsatellite marker development and genetic variation in the toxic marine diatom Pseudo-nitzschia multiseries (Bacillariophyceae). J. Phycol., 40:911-20.

EVANS, K., KÜHN, S. F. and HAYES, P. K., 2005. High levels of genetic diversity and low levels of genetic differentiation in North Sea *Pseudo-nitzschia pungens* (Bacillariophyceae) populations. *J. Phycol.,* **41**:506-514.

EXCOFFIER, L., SMOUSE, P. E. and QUATTRO, J. M., 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics.,* **131**:479-91.

FABREGAT, I., KOCH, K. S., AOKI, T., ATKINSON, A. E., DANG, H., AMOSOVA, O., FANG, G. and SECH, T. R., 1995. Telomere proteins. In: Telomeres. BLACKBURN, E. H. and GREIDER, C. W., (Eds.), Cold spring harbor laboratory press, Cold Spring Harbor, Newyork, pp. 69-106.

FANG, G. and SECH, T. R., 1995. Telomere proteins. **In**: Telomeres. BLACKBURN, E. H. and GREIDER, C. W. (Eds.), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, Newyork., pp. 69-106.

FAUGERSON, S., VALERO, M., DESTOMBE, C., MARTÍNEZ, E. A. and CORREA, J. A., 2001. Hierarchical spatial structure and discrimination analysis of genetic diversity in the red alga *Mazzaella laminarioides* (Gigartinales, Rhodophyta). *J. Phycol.*, **37**:705-16.

FEHLING, J., GREEN, D. H., DAVIDSON, K., BOLCH, C. J. and BATES, S. S., 2004. Domoic acid production by *Pseudo-nitzschia seriata* (Bacillariophyceae) in Scottish waters. *J. Phycol.*, **40**:622-30.

FERGUSON, A., TAGGART, J. B., PRODOHL, P. A., MC MEEL, O., THOMPSON, C. and STONE, C., 1995. The application of genetic markers to the study of conversion of fish population, with special reference to Salmo. *J. Fish Biol.,* **47**: 103-126.

FERRIS, P.J., WOESSNER, J.P. and GOODENOUGH, U.W., 1996. A sex recognition glycoprotein is encoded by the plus mating-type gene fus1 of *Chlamydomonas reinhardtii. Mol. Biol. Cell.,* **7:** 1235-1248.

FINLAY, B., 2002. Global dispersal of free-living microbial eukaryote species. *Sci.,* **296:** 1061-1063.

FORNONI, A., WANG, Y., LENZ, O., STRIKER, L. J. and STRIKER, G. E., 2002. Association of a decreased number of d(CA) repeats in the

matrix metalloproteinase- 9 promoter with glomerulosclerosis susceptibility in mice. *J. Am. Soc. Nephrol.,* **13**: 2068-2076.

FUJIWARA, H., OKUDA, H., OMI, T., IWAMOTO, S., TANAKA, Y., TAKAHASHI, Y., TANI, Y., MINAKAMI, H., ARAKI, S., SATO, I. and KAJII, E., 1999. The STR polymorphisms in introns 8 may provide information about the molecular evolution of RH haplotypes. *Hum. Genet.,* **104**: 301-306.

FURINI, A. and WUNDER, J., 2003. Analysis of eggplant (*Solanummelongena*)-related germplasm: morphological and AFLP data contribute to phylogenetic interpratations and germplasm utilization. *Theor. Appl. Genet.,* **108**: 197-208.

GALLAGHER, J. C., 1980. Population genetics of *Skeletoneama costatum* (Bacillariophyceae) in Narragansett Bay. *J. Phycol.,* **16**:464-474.

GALLAGHER, J. C., 1982. Physiological variation and electrophoretic banding patterns of genetically different seasonal populations of *Skeletonema costatum* (Bacillariophyceae). *J. Phycol.,* **18**:148-162.

GANAPATHY, P. N. and RAMAN, A. V., 1979. Organic pollution and *Skeletonema blooms* in visakhapatnam harbour. *Indian J. Mar. Sci.,* **8**: 184.

GANDON, S., 1998.Local adaptation and host-parasite interactions. *Trends Ecol. Evol.,* **12:** 214-216.

GEBHARDT, F., ZANKER, K. S. and BRANDT, B., 1999. Modulation of epidermal growth factor receptor gene translation by a polymorphic di-nucleotide repeat in introns 1. *J. Biol. Chem.,* **274**: 13176-13180.

GJOSAETER, J., LEKVE, K., STENSETH, N.C., LEINAAS, H.P., CHRISTIE, H., DAHL, E., DANIELSSEN, D.S., EDVARDSEN, B., OLSGARD, F., OUG, E. and PAASCHE, E., 2000. A long-term perspective on the *Chrysochromulina* bloom on the Norwegian Skageerak coast 1988: a catastrophe or an innocent incident. *Marine Ecology Progress Series.,* **207**: 201-218.

GLAZIER, A. M., NADEAU, J. H. and AITMAN, T. J., 2002. Finding genes that underlie complex traits. *Science.,* **298**: 2345-2349.

GODHE, A., OTTA, S. K., REHNSTAM-HOLM, A-S., KARUNASAGAR, I. and KARUNASAGAR, I., 2001. Polymerase chain reaction in detection of *Gymnodinium mikimotoi* and *Alexandrium minutum* in field samples from southwest India. *Mar. Biotechnol.,* **3**:152-162.

GODHE, A., SVENSSON, S. and REHNSTAM-HOLM, A-S., 2002. Oceanogarphic settings explain fluctuations in *Dinophysis* sp and concentration of dairrhetic shellfish toxin in the plankton community within a mussel farm area on the Swedish west coast. *Marine Ecol. Prog. Series.,* **240**: 71-83.

GODHE, A., MCQOUID, M., KARUNASAGAR, I., KARUNASAGAR, I. and REHNSTAM-HOLM, A-S., 2006. Comparison of three common molecular tools for distinguishing among geographically separated clones of the diatom *Skeletonema marinoi* Sarno *et al* Zingone (Bacillariophyceae). *J. Phycol.,* **42**:280-291.

GOLDSTEIN, D.B. and POLLOCK, D.D., 1997. Launching microsatellites: A review of mutation processes and methods of phylogenetic inference. *J. Hered.,* **88:** 335-342.

GRIBBLE, K. E. and ANDERSON D. M., 2005. Pseudogenes of the LSU rDNA of *Protoperidinium* spp. (Dinophyceae): Interspecific, intraspecific, and intracellular variability. ASLO, Santiago de Compostela, Spain, Conference abstract, pp. 61.

GUILLARD, R. R. L., 1975. Culture of phytoplankton for feeding marine invertebrates. **In**: Culture of Marine Invertebrate Animals SMITH, W., CHANLEY, M., (Eds.), Plenum Press, New York, pp. 29-60.

GUILLARD, R.R.L. and KILHAM, P., 1977. The ecology of marine planktonic diatoms. **In:** The Biology of diatoms. WERNER, D., (ed), Blackwell, London, pp. 327-469.

GUM, B., GROSS, R. and KUEHN., 2005. Mitochondrial and nuclear DNA phylogeography of European grayling (*Thymallus thymallus*): evidence for secondary contact zones in central Europe. *Mol. Ecol.* **14**: 1707-1725.

GUR-ARIE, R., COHEN, C., EITAN, L., SHELEF, E., HALLERMAN, E. and KASHI, Y., 2000. Abundance, non-random genomic distribution, nucleotide composition, and polymorphism of simple sequence repeats in *E .coli. Genome Res.,* **10**: 61-70.

HADRYS, H., BALICK, M. and SCHIERWATER, B., 1992. Applications of random amplified polymorphic DNA (RAPD) to molecular ecology. *Mol. Ecol.,* **1**:55-63.

HAGELBERG, E., GRAY I.C. and FEFFREYS, A. J., 1991. Identification of the skeletal remains of a murder victim by DNA analysis . *Nature,* **352:** 427-429**.**

HAGEN, L. S., KHADARI, B., LAMBERT, P. and AUDERGON, J.-M., 2002. Genetic diversity in apricot revealed by AFLP markers: species and cultivar comparisons. *Theor. Appl. Genet.*, **105**: 298-305.

HALLEGRAEFF, G.M., 1995. Harmful algal blooms; A global overview. **In**: *Manual on harmful marine microalgae*. HALLEGRAEFF, G.M., ANDERSON, D.M. and CEMBELLA, A.D., (Eds.), Ioc Manuals and Guides, no.33. UNESCO, PARIS, pp. 1-22.

HAMADA, H., PETRINO, M. and KAKUNAGA, T., 1982. A novel repeated element with Z-DNA forming potential is widely found in evolutionarily diverse eukaryotic genomes. *Proc. Natl. Acad. Sci.*, USA **76:** 6465-6469.

HANCOCK, J. M., 1999. Microsatellites and other simple sequences: genomic context and mutational mechanisms. **In**: Microsatellites: evolution and applications. GOLDSTEIN, D. B. and SCHLOTTERER, C., (Eds.), Oxfors university press, Oxford, pp. 1-9.

HARGRAVES, P. E. and FRENCH, F. W., 1983. Diatom resting spores: significance and strategies. **In**: Survival strategies of the algae FRYXELL, G. A., (Ed.), Cambridge University Press, Cambridge, pp. 49-68.

HARTL, L. H. and JONES, E. W., 2001. Genetic linkage and chromosome mapping. **In**: genetics: analysis of genes and genomes, 5th ed. Jones and Barlett publishers, London, pp. 175-221.

HAUGE, X.Y. and LITT, M., 1993. A study of the origin of shadow bands seen when typing di-nucleotide repeat polymorphisms by the PCR. *Hm. Mol. Genet.*, **2**: 411-415.

HEDGECOCK, D., 1994. Population genetics of marine organisms. *US Globec News.*, **6:** 1-13.

HENDERSON, E., 1995. Telomere DNA structure. **In**: Telomeres. BLACKBURN, E. H. and GREIDER, C. W., (Eds.), Cold spring harbor laboratory press, Cold Spring Harbor, Newyork, pp. 35-68.

HILLIS, D. M., MABLE, B. K., LARSON, A., DAVIS, S. K. and ZIMMER, E. A., 1996a. **In**:Nucleic acids IV: sequencing and cloning. HILLIS, D. M., MORITZ, C. and MABLE, B. K. (Eds.), *Molecular systematics*. Sinauer Associates, Sunderland, pp. 321-78.

HILLIS, D.M., MORITZ, D. and MABLE, B.K., 1996b. Molecular Systematics. Sunderland, MA : Sinauer Associates, pp.655.

HOLZMANN, M., PILLER, W. and PAWLOWSKI, J., 1996. Sequence variations in the large-subunit ribosomal RNA gene of *Ammonia* (Foraminifera, Protozoa) and their evolutionary implications. *J. Mol. Evol.*, **43**:145-51.

IGLESIAS-RODRIGUEZ, M., SCHOFIELD, O., BATLEY, J., MEDLIN, L. and HAYES, P., 2006. Intraspecific genetic diversity in the marine coccolithophore *Emiliania huxleyi* (Prymnesiophyceae): The use of microsatellite analysis in marine phytoplankton studies. *J. Phycol.*, **42**:526-536.

ITEMAN, I., RIPPKA, R., de MARSAC, N.T. and HERDMAN, M., 2002. rDNA analyses of planktonic heterocystous cyanobacteria, including members of the genra *Anabaenopsis* and *Cyanospira. Microbiol.Sgm.*, **148:** 481-496.

JAKUPCIAK, J. P. and WELLS, R. D., 2000. I. Genetic instabilities of triplet repeat sequences by recombination. *UBMB Life,* **50**: 355-359.

JEFFREYS, A.J., WILSON, V. and THEIN, S.L., 1985. Hypervariable minisatellite regions in human DNA. *Nature*, **314:** 67-73.

JEFFREYS, A. J., NEIL, D. L. and NEWMANN, R., 1998. Repeat instability at human minisatellites arising from meiotic recombination. *EMBO J.*, **17**: 4147-4157.

JOHANNSDOTTIR, J. T., JONASSON, J. G., BERGTHORSSON, J. T., AMUNDADOTTIR, L., MAGNUSSON, J., EGILSSON, V. and INGVARSSON, S., 2000. The effect of mismatch repair deficiency on tumourgenesis; microsatellite instability affecting genes containing short repeated sequences. *Int. J. Oncol.*, **16**: 133-139.

JONES, A. G. and ARDEN, W. R., 2003. Methods of parentage analysis in natural populations. *Mol. Ecol.* **12**: 2511-2523.

JOSEFSON, A.B. and HANSEN, J.L.S., 2003. Quantifying plant pigments and live diatoms in aphotic sediments of Scandinavian coastal waters confirms a major route in the pelagic –benthic coupling. *Mar. Biol.*, **142:** 649-658.

KAEBERNICK, M., DITTMANN, E., BORNER, T. and NEILAN, B.A., 2002. Multiple alternate transcripts direct the biosynthesis of microcystin, a cyanobacterial nonribosomal peptide. *Appl. Environ. Microbiol.*, **68:** 449-455.

KANDPAL, R., KANDPAL, G. and WEISSMAN, S.M., 1994. Construction of libraries enriched for sequence repeats and jumping clones and

hybridization selection for region-specific markers. *Proc. Natl. Acad. Sci.,* **91**: 88-92.

KANTETY, R. V., LA ROTA, M., MATTHEWS, D. E. and SORRELLS, M. E., 2002. Data mining for simple sequence repeats in expressed sequence tags from barley, maize, rice, sorghum and wheat. *Plant Mol. Biol.,* 48: 501-510.

KARSI, A., PATTERSON, A., FENG, J. and LIU, Z.J., 2000. Translational machinery of channel catfish: I. A transcriptomic approach to the analysis of 32 of 40S ribosomal protein genes and their expression. *Gene,* **291:** 177-186.

KASHI, Y., KING, D.C. and SOLLER, M., 1997. Simple sequence repeats as a source of quantitative genetic variation. *Trends Genet.,* **13**: 74-78.

KIJAS, J.M.H., FOWLER, J. C. S., GARBETT, C. A. and THOMAS, M. K., 1994. Enrichment of microsatellite from the citrus genome using biotinylated oligonucleotide sequences bound to streptavidin-coated magnetic particles. *Bio-Techniques.,* **16**: 656-662.

KIMPTON, C. P., GILL, P., WALTON, A., URQUHART, A., MILLICAN, E. S. and ADAMS, M., 1993. Automated DNA profiling employing multiplex amplification of short tandem repeat loci. *PCR. Methods Appl.,* **3**: 13-22.

KIMURA, M. and CROW, J.F., 1964. The number of alleles that can be maintained in a finite population. *Genetics.,* **49**: 725-738.

KIMURA, M. and OHTA, T., 1978. Stepwise mutational model and distribution of allelic frequencies in a finite population. *Proc. Natl. Acad. Sci.* U.S.A., **75**: 2868-2872.

KLEIN, S., CHAPMAN, D.J. and GARBER, E.D., 1973. Chemotaxonomy of the *Oscillatoria-Phormidium* complex. *Biochem. Syst.,* **1:** 173-177.

KODAMA, M., OGATA, T. and SATO, S., 1993. *Alexandrium tamarense* as a source of tetradotoxin in the scallop *Patinopecten yessoensis.* In: Toxic phytoplankton blooms in the sea. SMAYDA, T. J. and SHIMIZU, Y., (Eds.), Elsevier Science Publisher, Tokyo, pp. 401-406.

KOLODNER, R. D. and MARSISCHKY, G. T., 1999. Eukaryotic DNA mismatch repair. *Curr. Opin. Genet. Dev.,* **9**: 89-96.

KOSKINEN, M. T., NILSSON, J., VESELOV, A. J., POTUTKIN, A.G., RANTA, E. and PRIMMER, C. R., 2002. Microsatellite data resolve phylogeographic patterns in European grayling, *Thymallus thymallus,* Salmonidae. *Heredity.,* **88**: 391-401.

KRESOVICH, S., WILLIAMS, J. G. K., MCFERSON, J. R., ROUTMAN, E. J. and SCHAAL, B. A., 1992. Characterisation of genetic identities and relationships of *Brassica oleracea* L. via random amplified polymorphic DNA assay. *Theor. Appl. Genet.,* **85**:190-6.

LANDER, E. and KRUGLYAK, L., 1995. Genetic dissection of complex traits: guidance for interpreting and reporting linkage results. *Nat. Genet.,* **11**: 241-247.

LANGE, C., HASLE, G. and SYVERTSEN, E., 1992. Seasonal cycle of diatoms in the Skagerrak, North Atlantic, with emphasis on the period 1980-1990. *Sarsia.,* **77**:173-187.

LARSEN, A., CASTBERG, T., SANDAA, R.A, BRUSSAARD, C.P.D., EGGE, J., HELDAL, M., PAULINO, A., THYRHAUG, R., vanHANNEN, E.J. and BRATBAK, G., 2001. Population dynamics and diversity of phytoplankton, bacteria and viruses in a seawater enclosure. *Mar. Ecol. Prog. Ser.*, **221:** 47-57.

LECLERC, D., WIRTH, T.and BERNATCHEZ, L., 2000. Isolation and characterization of microsatellite loci in the yellow perch (*Perca flavescens*), and cross-species amplification within the family Percidae. *Mol. Ecol.,* **9**: 995-997.

LENAERS, G., MAROTEAUX, L., MICHOT, B. and HERZOG, M., 1989. Dinoflagellates in evolution. A molecular phylogenetic analysis of large subunit ribosomal RNA. *J. Mol. Evol.,* **29**:40-51.

LEVINSON, G. and GUTMAN, G.A., 1987. High frequency of short frameshifts in poly – CA /GT tandem repeats borne by bacteriophage M13 in *Escherichia coli* K-12. *Nucl. Acid Res.,* **15:** 5323-5338.

LEVINSON, G. and GUTMEN, G. A., 1987a. Slipped-strand mispairing: a major mechanism for DNA sequence evolution. *Mol. Biol. Evol.,* **4**: 203-221.

LEWIS, R.J., JENSEN, S.I., DeNICOLA, D.M., MILLER, V.I., HOAGLAND, K.D. and ERNST, S.G., 1997. Genetic variation in the diatom *Fragilaria capucina* (Fragilariaceae) along the latitudinal gradient across North America. *Plant Syst. Evol.,* **204:** 99-108.

LI, W. H., 1997. Molecular evolution. *Sinauer Associates*, Sunderland, MA, pp. 177-213.

LI, Y. C., KOROL, A. B., FAHIMA, T., BEILES, A. and NEVO, E., 2002. Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. *Mol. Ecol.*, **11**: 2453-2465.

LI, Y. C., KOROL, A. B., FAHIMA, T. and NEVO, E., 2004. Microsatellites within genes: structure, function and evolution. *Mol. Biol. Evol.*, **21**: 991-1007.

LIANG, Y., MAI, K. and SUN, S., 2005. Difference in growth, total lipd contents and fatty acid composition among 60 clones of *Cylindrotheca fusiformis*. *J. Appl. Phycol.*, **17**:61-65.

LINDAHL, O., 1995. Long term studies of primary phytoplankton production in the Gullmar fjord, Sweden. **In**: Ecology of Fjords and Coastal Waters SKJOLDAL, H. R., HOPKINS, C., ERIKSTAD, K. E. and LEINAAS, H. P., (Eds.), Elsevier Science, Amsterdam, pp. 105-112.

LITT, M. and LUTY, J.A., 1989. A hypervariable microsatellite revealed by in vitro amplification of dinucleotide repeat within the cardiac muscle actin gene. *Am. J. Hum. Genet.*, **44:** 397-401.

LIU, Z., TAN, G., LI, P. and DUNHAM, R. A., 1999a. Transcribed di-nucleotide microsatellites and their associated genes from channel catfish *Ictalurus punctatus*. *Biochem. Biophys. Res. Commun.*, **259**: 190-194.

LIU, Z., KARSI, A. and DUNHAM, R. A., 1999b. Development of polymorphic EST markers suitable for genetic linkage mapping of catfish. *Mar. Biotechnol.*, **1**: 437-447.

LIU, Z.J., LI, P., KOCABAS, A., JU, Z., KARSI, A., CAO, D. and PATTERSON, A., 2001. Microsatellite – containing genes from the channel catfish brain: evidence of trinucleotide repeat expansion in the coding region of nucleotide excision repair gene RAD23B. *Biochem. Biophys. Res. Commun.*, **289:** 317-324.

LIU, Z. J. and CORDES, J. F., 2004. DNA marker technologies and their applications in aquaculture genetics. *Aquaculture.*, **238**: 1-37.

LIVELY, C.M. and DYBDAHL, M.F., 2000. Parasite adaptation to locally common host genotypes. *Nature,* **405 :** 679 – 681.

LORET, P., TENGS, T., VILLAREAL, T. A., SINGLER, H., RICHARDSON, B., MCGUIRE, P., MORTON, S., BUSMAN, M. and CAMPBELL, L., 2002. No difference found in ribosomal DNA sequences from physiologically diverse clones of *Karenia brevis* (Dinophyceae) from the Gulf of Mexico. *J. Plankton Res.*, **24**:735-739.

LUIKART, G. and ENGLAND, P. R., 1999. Statistical analysis of microsatellite data. *Trends Ecol. Evol.*, **14**: 253-256.

LUNDHOLM, N., DAUGBJERG, N. and MOESTRUP, O., 2002. Phylogeny of the Bacillariaceae with emphasis on the genus Pseduo-nitzchia

(Bacillariophyceae) based on partial LSU rDNA. *Eur. J. Phycol.*, **37:** 115-134.

LUNDHOLM, N., MOESTRUP, Ø., HASLE, G. R. and HOEF-EMDEN, K., 2003. A study of the*Pseudo-nitzschia pseudodelicatissima/cuspidata* complex (Bacillariophyceae): What is *P. pseudodelicatissima*?. *J. Phycol.*, **39**:797-813.

MACKAY, T. F., 2001. The genetic architecture of quantitative traits. *Annu. Rev. Genet.*, **35**: 303-339.

MANN, D.G. and MERCHANT, H.J., 1989. The origin of the diatom and its life cycle. **In:** The chromophyte algae: problems and perspectives, GREEN, J.C., LEADBEATER, B.S.C. and DIVER, W.L., (Eds.). Clarendon Press, Oxford, UK. pp. 303-307.

MANN, D.G., 1999. The species concept in diatoms. *Phycologia*, **38:** 437-495.

MANSOUR, M., WRIGHT JR, J. R.and POHAJDAK, B., 1998. Cloning, sequencing and characterization of the tilapia insulin gene. *Comp. Biochem. Physiol., B Biochem. Mol. Biol.*, **121**: 291-297.

MARIN, B., KLINGBERG, M. and MELKONIAN, M., 1998. Phylogenetic relationships among the Cryptophyta: analyses of nuclear-encoded SSU rRNA sequences support the monophyly of extant plastid-containing lineage. *Protist.*, **149**:265-276.

MARTINS, C., OLIVERIA, C., WASKO, A. P. and WRIGHT, J. M., 2004. Physical mapping of the Nile tilapia (*Oreochromis niloticus*) genome by fluorescent in-situ hybridization of repetitive DNAs to metaphase chromosomes-a review. *Aquaculture.*, **231**: 37-49.

McCARTNEY, M. A., ACEVEDO, J., HEREDIA, C., RICO, C., QUENOVILLE, B., BERMINGHAM, E. and MC MILLAN, W. O., 2003. Gnetic mosaic in a marine species flock. *Mol. Ecol.*, **12**: 2963-2973.

MCQUOID, M. R., 2002. Pelagic and benthic environmental controls of the spatial distribution of a viable diatom seed bank on the Swedish west coast. *J. Phycol.*, **38**:881-893.

MCQUOID, M. R., GODHE, A. and NORDBERG, K., 2002. Viability of phytoplankton resting stages in the sediments of a coastal Swedish fjord. *Eur. J. Phycol.*, **37**:191-201.

MCQUOID, M. and GODHE, A., 2004. Recruitment of coastal planktonic diatoms from benthic versus pelagic cells: Variation in bloom development and species composition. *Limnol. Oceano.*, **49**:1123-1133

MEDLIN, L.K., CRAWFORD, R.M. and ANDERSEN, R.A., 1986. Histochemical and ultrastructrual evidence for the function of the labiate process in the movement of centric diatoms. *Br. Phycol. J.,* **21:** 297-301.

MEDLIN, L. K., ELWOOD, H. J., STICKEL, S. and SOGIN, M. L., 1988. The characterization of enzymatically amplified eukaryotic 16S-like rRNA-coding regions. *Gene.,* **17**:491-9.

MEDLIN, L. K., WILLIAMS, D.M. and SIMS, P.A., 1993. The evolution of the diatoms (Bacillariophyta). I. Origin of the group and assessment of the monophyly of its major divisions. *Eur. J. Phycol.,* **28:** 261-275.

MEDLIN, L.K., KOOISTRA, W.H.C.F., GERSONDE, R. and WELLBROCK, U., 1996. Evolution of the diatoms (*Bacillariophyta*) . II. Nuclear encoded small subunit rRNA sequence comparisons confirm a paraphyletic origin for the centric diatoms. *Mol. Biol. Evol.,* **13:** 67-75.

MELKONIAN, M. and SUREK, B., 1995. Phylogeny of the chlorophyta congruence between ultrastructural and molecular evidence. *Bull. Soc. Zool. FR- Evol. Zool.,* **120:** 191-208.

MELONI, R., ALBANESE, V., RAVASSARD, P., TREILHOU, F. and MALLET, J., 1998. A tetra-nucleotide polymorphic microsatellite, located in the first introns of the tyrosine hydroxylase gene, acts as a transcription regulatory element in vivo. *Hum. Mol. Genet.,* **7**: 423-428.

MESSIER, W., LI, S. H. and STEWART, C. B., 1996. The birth of microsatellites. *Nature.,* **381**: 483.

METZGAR, D., BYTOF, J. and WILLS, C., 2000. Selection against frameshift mutations limits microsatellite expansion in coding DNA. *Genome Res.,* **10:** 72-80.

MEYERS, R.M., MANIATIS, T. and LERMAN, L.S., 1987. Detection and localization of single base changes by denaturing gradient gel-electrophoresis. *Method Enzymol.,* **155:** 501-527.

MEYER, E., WIEGAND, P., RAND, S. P., KUHLMAN, E., BRACK, M. and BRINKMANN, B., 1995. Microsatellite polymorphisms reveal phylogenetic relationships in primates. *J. Mol. Evol.,* **41**: 10-14.

MOORE, L. and CHRISHOLM, S., 1999. Photophysiology of the marine cyanobacterium *Prochlorococcus*: Ecotyic difference among cultured isolates. *Limnol. Ocean0.,* **44**:628-638.

MORGANTE, M., HANAFEY, M. and POWELL, W., 2002. Microsatellite are preferentially associated with non repetitive DNA in plant genomes. *Nat. Genet.,* **30:** 194-200.

MORIN, P.A., WALLIS, J., MOORE, J., CHAKRABORTY, R. and WOODRUFF, D., 1993. Non-invasive sampling and DNA amplification for paternity exclusion, community structure, and phylogeography in wild chimpanzees. *Primates*, **34:** 347-356.

MURPHY, L.S. and GUILLARD, R.R.L., 1976. Biochemical taxonomy of marine phytoplankton by electrophoresis of enzymes. I. The centric diatoms *Thalassiosira pseudomonana* and *T. fluviatilis* . *J. Phycol.,* **12:** 9-13.

MURRAY, V., MONCHAWN, C. and ENGLAND, P.R. 1993. The determination of the sequences present in the shadow bands of a di-nucleotide repeat PCR. *Nucl. Acids. Res.,* **21**: 2395-2398.

NADEAU, J. H., SINGER, J. B., MATIN, A. and LANDER, E.S., 2000. Analysing complex genetic traits with chromosome substitution strains. *Nat. Genet.*, **24**: 221-225.

NAGAI, S., LIAN, C., HAMAGUCHI, M., MASUYAMA, Y., ITAKURA, S. and HOGETSUS, T., 2004. Development of microsatellite markers in the toxic dinoflagellate *Alexandrium tamarense* (Dinophyceae). *Mol. Ecol. Notes*, **4**:83-85.

NAGAI, S., LIAN, C., SÊYAMAGUCHI, M., MATSUYAMA, Y. S., SHIMADA, H., YAMAUCHI, H., SONDA, Y., NISHIKAWA, T., KIM, C-H. and HOGETSU, T., 2007. Microsatellite markers reveal population genetic structure of the toxic dinoflagellate *Alexandrium tamarense* (Dinophyceae) in Japanese coastal water. *J. Phycol.*, **43**:43-54.

NAIDOO, R. and CHETTY, R., 1998. The application of microsatellites in molecular pathology. *Pathol. Oncol. Res.,* **4**: 310-315.

NAKAZIMA, M., 1965. Studies on the source of shellfish poison in lake Hamana I: Relation of the abundance of species of dinoflagellates *Prorocentrum* sp. To shellfish toxicity. *Bull. Jap. Soc. Sci. Fish.*, **31(3)**: 198-203.

NARUSE, K., FUKAMACHI, S., MITANI, H., KONDO, M., MATSUOKA, T., KONDO, S., HANAMURA, N., MORITA, Y., HASEGAWA, K., NISHIGAKI, R., SHIMADA, A., WADA, H., KUSAKABE, T., SUZUKI, N., KINOSHITA, M., KANAMORI, A., TERADO, T., KIMURA, H., NONAKA, M. and SHIMA, A., 2000. A detailed linkage map of medaka, *Oryzias latipes*: comparative genomics and genome evolution. *Genetics.,* **157**: 1773-1784.

NEFF, B. D. and GROSS, M. R., 2001. Microsatellite evolution in vertebrates: inference from AC dinucleotide repeats. *Evolution.,* **55**: 1717-1733.

NEILAN, B.A., HAWKINS, P.R., COX, P.T. and GOODMAN, A.E., 1994. Towards a molecular taxonomy for the bloom forming cyanobacteria. *Aust. J. Mar. Freshwat. Res.,* **45:** 869-873.

NEILAN, B.A., JACOBS, D. and GOODMAN, A. E., 1995. Genetic diversity and phylogeny of toxic cyanobacteria determined by DNA polymorphisms within the phycocyanin locus. *Appl. Environ. Microbiol.,* **61:** 3875-3883.

NEILAN, B.A., STUART, J.L., GOODMAN, A.E., COX, P.T. and HAWKINS, P.R., 1997. Specific amplification and restriction polymorphisms of the cyanobacterial rRNA operon spacer region. *Syst. Appl. Microbiol.,* **20:** 612-621.

NG, S. H. S., CHANG, A., BROWN, G., KOOP, B. F. and DAVIDSON, W. S., 2005. Data mining ESTs: Atlantic salmon variable microsatellites discovery. Plant and animal genome XIII abstracts. Available on http://www.intl-pag.org/13/abstracts/PAG13_P599.html.

O'BREIN, S. J., 1991. Molecular genome mapping lessons and prospects. *Curr. Opin. Genet. Dev.,* **1**: 105-111.

O'CONNEL, M. and WRIGHT, J.M., 1997. Microsatellite DNA in fishes. *Rev. Fish Biol. Fish.,* **7:** 331-363.

OKUDA, H., SUGANUMA, H., KAMESAKI, T., KUMADA, M., TSUDO, N., OMI, T., IWAMOTO, S. and KAJII, E., 2000. The analysis of nucleotide substitutions, gaps, and recombination events between RHD and RHCE genes complete sequencing. *Biochem. Biophys. Res. Commun.,* **274**: 670-683.

OKUDA, H. and KAJII, E., 2002. The evolution and formation of *RH* genes. *Leg. Med.,* **4**: 139-155.

OLDACH, D.W., DELWICHE, C.F., JAKOBSEN, K.S., ENGS, T., BROWN, E.G., KEMPTON, J.W., SCHAEFER, E.F., BOWERS, H.A., GLASGOW, H.B. and BURKHOLDER, J.M., *et al.*, 2000. Heteroduplex mobility assay-guided sequence discovery: Elucidation of the small subunit (18S) rDNA sequences of *Psiesteria piscicida* and related dinoflagellates from complex algal culture and environmental sample DNA pools. *Proc. Natl. Acad. Sci.,*USA **, 97:** 4303-4308.

ORSINI, L., SARNO, D., PROCACCINI, G., POLETTI, R., DAHLMAN, J. and MONTRESOR, M., 2002. Toxic *Pseudo-nitzschia multistriata*

(Bacillariophyceae) from Gulf of Naples: morphology, toxin analysis and phylogenetic relationships with other *Pseudo-nitzschia* species. *Eur. J. Phycol.*, **37**:247-257.

ORSINI, L., PROCACCINI, G., SARNO, D. and MONTRESOR, M., 2004. Multiple rDNA ITS-types within the diatom *Pseudo-nitzschia delicatissima* (Bacillariophyceae) and their relative abundances across a spring bloom in the Gulf of Naples. *MEPS.*, **271**:87-98.

OSTRANDER, E.A., JONG, P.M., RINE, J. and DUYK, G., 1992. Construction of small – insert genomic DNA libraries highly enriched fro microsatellite repeat sequences. *Proc. Natl. Acad. Sci.,* U.S.A. **89:** 3419-3423.

PALUMBI, S.R., 1992. Marine speciation on a small planet. *Trends Ecol. Evol.*, **7:** 114-118.

PALUMBI, S. R., 1996. Nucleic acids II: the polymerase chain reaction. **In**: Molecular systematics. HILLIS, D. M., MORITZ C. and MABLE, B. K. (Eds.), Sinauer Associates, Sunderland, pp. 205-47.

PARTENSKY, F., HOEPFFNER, N., LI, W.K.W., *et al.*, 1993. Photoacclimation of *Prochlorococcus* sp. (Prochlorophyta) strains isolated from the North Atlantic and the Mediterranean Sea. *Plant Physiol.*, **101:** 295-296.

PEICHEL, C. L., ROSS, J. A., MATSON, C.K., DICKSON, M., GRIMWOOD, J., SCHMUTZ, J., MYERS, R. M., MORI, S., SCHLUTER, D. and KINGSLEY, D. M., 2004. The master sex-determination locus in threespine stickle-backs is on a nascent Y chromosome. *Curr. Biol.*, **14**: 1416-1424.

PENNA, A., VILA, M., FRAGA, S., GIACOBBE, M. G., ANDREONI, F., RIOBO, P. and VERNESI, C., 2005. Characterization of *Ostreopsis* and *Coolia* (Dinophyceae) isolates in the western Mediterranean Sea based on morphology, toxicity and internal transcribed spacer 5.8S rDNA sequences. *J. Phycol.*, **41**:212-25.

POOMPUANG, S. and HALLERMAN, E. M., 1997. Toward selection of quantitative trait loci and marker-assisted selection in fish. *Rev. Fish Biol. Fish.*, **5**: 253-277.

POWEL, W., MACHRAY, G. C. and PROVAN, J., 1996. Polymorphism revealed by simple sequence repeats. *Trends Plant Sci.*, **1**: 215-222.

PRESNEAU, N., MANDERSON, E. N. and TONIN, P. N., 2003. The quest for a tumor suppressor gene phenotype. *Curr. Mol. Med.*, **3**: 605-629.

PRIMMER, C. R., ELLEGREN, H., SAINO, N. and MOLLER, A. P., 1996. Directional evolution in germline microsatellite mutations. *Nat. Genet.* **13**: 391-393.

PRIMMER, C. R. and ELLEGREN, H., 1998. Pattern of molecular evolution in avian microsatellites. *Mol. Biol. Evol.,* **15**: 997- 1008.

PRIMMER, C. R., KOSKINEN, M. T. and PIIRONEN, J., 2000. The one that did not get away: individual assignment using microsatellite data detects a case of fishing competition fraud. *Proc. R. Soc. Lond., B Biol. Sci.* **267**: 1699-1704.

PUJOLAR, J. M., MAES, G. E., VANCOILLIE, C. and VOLCKAERT, F. A. M., 2005. Growth rate correlates to individual heterozygosity in European eel, *Anguilla anguilla* L. *Evolution.,* **59:** 189-199.

RAMAMIRTHAN, C. P. and JAYARAMAN, R., 1963. Some aspects of the hydrological conditions of the backwater around Wellington island (Cochin). *J. Mar. Biol. Assoc. India,* **5**: 170.

RAMAMURTHY, V. D., SELVAKUMAR, A. R. and BHARGAVA, R. M. S., 1972. Studies on the blooms of *Trichodesmium erythraeum* in the waters of the central west coast of India. *Curr. Sci.*, **41**: 803.

RASSMAN, K., SCHLOTTERER, C. and TAUTZ, D., 1991. Isolation of simple sequence loci for use in polymerase chain reaction – based DNA fingerprinting. *Electrophoresis.,* **12**: 113-118.

REHNSTAM-HOLM, A.-S., GODHE, A. and ANDERSON, D. M., 2002. Molecular studies of *Dinophysis* (Dinophyceae) species from Sweden and North America. *Phycologia.,* **41**:348-57.

RICH, A., NORDHEIM, A. and WANG, A. H. J., 1984. The chemistry and biology of left-handed Z-DNA. *Ann. Rev. Biochem.,* **53**: 791-846.

RICHARDS, R. I. and SUTHERLAND, G. R., 1994. Simple repeat DNA is not replicated simply. *Nat. Genet.,* **6**: 114-116.

RICHARD, G. F. and DUJON, B., 1997. Trinucleotide repeats in yeast. *Res. Microbiol.,* **148**: 731-744.

RISE, M. L., VON SCHALBURG, K. R., BROWN, G. D., MAWER, M. A., DEVLIN, R. H., KUIPERS, N., BUSBY, M., BEETZ-SARGENT, M., ALBERTO, R., GIBBS, A. R., HUNT, P., SHUKIN, R., ZEZNIK., J. A., NELSON, C., JONES, S. R. M., SMAILUS, D. E., JONES, S. J. M., SCHEIN, J. E., MARRA, M. A., BUTTERFIELD, Y. S. N., STOOT, J. M., NG, S. H. S., DAVIDSON, W. S. and KOOP, B. F., 2004. Development and application of a salmonid EST database and cDNA

microarray: data mining and interspecific hybridization characteristics. *Genome Res.*, **14**: 478-490.

RODHE, J., 1987. The large scale circulation in the Skagerrak; interpretation of some observations. *Tellus.*, **39**:245-253.

RODRIGUEZ-LANETTY, M., KRUPP, D. A. and WEIS, V. M., 2004. Distinct ITS types of *Symbiodinium* in Clade *C* correlate with cnidarian/dinoflagellate specificity during onset of symbiosis. *Mar. Ecol. Prog. Ser.*, **275**:97-102.

ROTHSCHILD, M.F., 2003. From a sow's ear to a silk purse: real progress in porcine genomics. Cytogenet. *Genome Res.*, **102:** 95-99.

ROUND F.E., CRAWFORD, R.M. and MANN, D.G., 1990. The diatoms : morphology, and biology of the genera. Cambridge University Press, Cambridge, UK.

RUDI, K., SKULBERG, O.M., LARSEN, F. and JAKOBSEN, K.S., 1997. Strain characterization and classification of oxyphotobacteria in clone cultures on the basis of 16S rRNA sequences from the variable regions V6, V7 and V8. *Appl. Environ. Microbiol.*, **63:** 2593-2599.

RUDI, K., SKULBERG, O.M., LARSEN, F. and JAKOBSEN, K.S., 1998. Quantification of toxic cyanobacteria in water by use of competitive PCR followed by sequence-specific labeling of oligonucleotide probes. *Appl. Environ. Microbiol.*, **64:** 2639-2643.

RYNEARSON, T. A. and ARMBURST, E. V., 2000. DNA fingerprinting reveals extensive genetic diversity in a field population of the centric diatom *Ditylum brightwelli*. *Limnol. Oceano.*, **45(6)**: 1329-1340.

RYNEARSON, T. and ARMBURST, E., 2004. Genetic differentiation among populations of the planktonic marine diatom *Ditylum brightwellii* (Bacillariophyceae). *J.Phycol.*, **40**:34-43.

RYNEARSON, T.A. and ARMBRUST, E.V., 2005. Maintenance of clonal diversity during a spring bloom of the centric diatom *Ditylum brightwellii*. *Mol. Ecol.*, **14:** 1631-1640.

RYNEARSON, T., NEWTON, J. and ARMBRUST, A., 2006. Spring bloom development, genetic variation, and population succession in the planktonic diatom *Ditylum brightwelli*. *Limnol. Oceano.*, **51**:1249-1261.

SAKAMOTO, T., DANZMANN, R. G., GHARBI, K., HOWARD, P., OZAKI, A., KHOO, S. K., WORAM, R. A., OKAMOTO, N., FERGUSON, M. M., HOLM, L. E., GUYOMARD, R. and HOYHEIM, B., 2000. A microsatellite linkage map of rainbow trout (*Oncorhynchus mykiss*) characterized by large sex-specific differences in recombination rates. *Genetics*, **155**: 1331-1345.

SALVI, S. and TUBEROSA, R., 2005. To clone or not to clone plant QTLs: present and future challenges. *Trends Plant Sci.*, **10**: 297-304.

SAMBROOK, J., FRITSCH, E.F. and MANIATIS, T. 1989. Molecular cloning: A laboratory manual. Cold spring harbor laboratory press, Newyork.

SARNO, D., KOOISTRA, W., MEDLIN, L., PERCOPO, I. and ZINGONE, A., 2005. Diversity in the genus *Skeletonema* (Bacillariophyceae). II. An assessment of the taxonomy of *S. costatum*-like species with the description of four new species. *J. Phycol.*, **41**:151-176.

SARNO, D., KOOISTRA, W., BALZANO, S., HARGRAVES, P. and ZINGONE, A., 2007. Diversity in the genus Skeletonema (Bacillariophyceae): III. Phylogenetic position and morphological variability of *Skeletonema costatum* and *Skeletonema grevillei*, with the description of *Skeletonema ardens* sp. nov. *J. Phycol.*, **43**:156-170.

SAVIN, M.C., MARTIN, J.L., LeGRESLEY, M., GIEWAT, M. and ROONEY-VARGA, J., 2004. Plankton diversity in the Bay of Fundy as measured by morphological and molecular methods. *Microb. Ecol.*, **48:** 51-65.

SCHLÖTTERER, C. and TAUTZ, D., 1992. Slippage synthesis of simple sequence DNA. *Nucl. Acids Res.*, **20:** 211-215.

SCHNEIDER, S., ROESSLI, D. and EXOFFICER, L., 2000. ARLEQUIN VER. 2.000: A software for population genetics data analysis. Geneva, Switzerland, Genetics and Biometry Laboratory, University of Geneva.

SCHNEIDER, P. M., BENDER, K., MAYR, W. R., PARSON, W., HOSTE, B. and DECORTE, R., 2004. STR analysis of artificially degraded DNA-results of a collaborative European exercise. *Forensic Sci. Int.*, **139**: 123-134.

SCHOLIN, C. A., HERZOG, M., SOGIN, M. and ANDERSON, D. M., 1994. Identification of group- and strain-specific genetic markers for globally distributed *Alexandrium* (Dinophyceae) II. Sequence analysis of a fragment of the LSU rRNA gene. *J. Phycol.*, **30**:999-1011.

SCHORK, N. J., FALLIN, D. and LANCHBURY, J. S., 2000. Single nucleotide polymorphisms and the future of genetic epidemiology. *Clin. Genet.*, **58**: 250-264.

SHANKLE, A.M., MAYALI, X. and FRANKS, P. J. S. 2004. Temporal pattern in population genetic diversity of *Prorocentrum micans* (Dinophyceae). *J. Phycol.*, **40**: 239-247.

SHETTY, H. P.C. and SAHA, S. B., 1971. On the significance of the occurrence of blooms of the diatom *Hemidiscus hardmannianus* (Greville) Mann in relation to *Hilsa* fishery in Bengal. *Curr. Sci.*, **40**: 410.

SHI, Y., HU, H. MA, R., CONG, W. and CAI, Z., 2004. Improved use of organic phosphate by *Skeletonema costatum* through regulation of Zn2+ concentrations. *Biotehnol. Lett.*, **26**:747-751.

SIA, E. A., KOKOSKA, R. J., DOMINSKA, M., GREENWELL, P. and PETES, T. D., 1997. Microsatellite instability in yeast: dependence on repeat unit size and DNA mismatch repair genes. *Mol. Cell. Biol.*, **17**: 2851-2858.

SINDEN, R. R., 1999. Trinucleotide repeats: biological implications of the DNA structures associated with disease-causing triplet repeats. *Am. J. Hum. Genet.*, **64**: 346-353.

SINGER, A., PERLMAN, H., YAN, Y., WALKER, C., CORLEY-SMITH, G., BRANDHORST, B. and POSTLETHWAIT, J., 2002. Sex-specific recombination rates in zebrafish (*Danio rerio*). *Genetics.*, **160**: 649-657.

SLATE, J., 2005. Quantitative traits locus mapping in natural populations: progress, caveats and future directions. *Mol. Ecol.*, **14**: 369-379.

SOUDEK, D. and ROBINSON, G.C., 1983. Electrophoretic analysis of the species and population structure of the diatom *Asterionella Formosa*. *Can. J. Bot. Rev. Can. Bot.*, **61:** 418-433.

SOURNIA, A., CHRETIENNOT, D. and RICARD, M., 1991. Marine phytoplankton: how many species in the world ocean. *J. Plankton. Res.*, **13**: 1093-1099.

STAMFORD, M. D. and TAYLOR, E. B., 2004. Phylogeographical lineages of Arctic grayling (*Thymallus arcticus*) in North America: divergence, origins and affinities with Eurasian *Thymallus*. *Mol. Ecol.* 13: 1533-1549.

STEHR, C., CONNELL, L., BAUGH, K., BILL, B., ADAMS, N. and TRAINER, V., 2002. Morphological, toxicological, and genetic

differences among *Pseudo-nitzschia* (Bacillariophyceae) species in inland embayments and outer coastal waters of Washington State, USA. *J. Phycol.,* **38**:55-65.

STEWART, W. D. P., SINADA, F., CHRISTOFI, N. and DARF, M. J., 1977. Primary production and microbial activity of Scottish freshwater habitats. **In**: Aquatic microbiology. SKINNER, F. A. and SHEWIN, J. M. (Eds.), Academic Press, London, pp. 31-54.

STORZ, J. F., 2005. Using genome scans of DNA polymorphism to infer adaptive population divergence. *Mol. Ecol.,* **14**: 671-688.

STREELMAN, J. T. and KOCHER, T. D., 2002. Microsatellite variation associated with prolactin expression and growth of slat-challenged tilapia. *Physiol. Genomics.,* **9**: 1-4.

SUBBARAO, D. V., 1969. *Asterionella japonica* bloom and discoloration off Waltair, Bay of Bengal. *Limnol. Oceanogr.,* **14**: 632.

SWOFFORD, D. L., 2002. PAUP* phylogenetic analysis using parsimony (*and other methods). version 4.0b10. Sinauer Associates, Sunderland, Massachusetts.

TAKABAYASHI, M., LEW, K., JOHNSON, A., MARCHI, A., DUGDALE, R. and WILKERSON, F., 2006. The effect of nutrient availability and temperature on chain length of the diatom, *Skeletonema costatum. J. Plankton Res.,* **28**:831-840

TALLBERG, P. and HEISKANEN, A., 1998. Species-specific phytoplankton sedimentation in relation to primary production along an inshore-offshore gradient in the Baltic Sea. *J. Plankton Res.,* **20**:2053-2070.

TASSANAKAJON, A., PONGSOMBOON, S. and RIMPHANITCHAYAKIT, V., 1997. Random amplified polymorphic DNA (RAPD) markers for the determination of genetic variation in wild populations of the black tiger prawns (*Penaeus monodon*) in Thailand. *Mol. Mar. Biol. Biotechnol.,* **6**:110-15.

TAUTZ, D. and RENZ, M., 1984. Simple sequences are ubiquitous repetitive components of eukaryote genomes. *Nucl. Acids Res.,* **12:** 4127-4138.

TAUTZ, D., TRICK, M. and DOVER, G. A., 1986. Cryptic simplicity in DNA is a major source of genetic variation. *Nature.,* **322**: 652-656.

TAUTZ, D., 1989. Hypervariablity of simple sequences as a general source for polymorphic DNA markers. *Nucl. Acid. Res.,* **17:** 6463-6471.

THIEL, T., MICHALEK, W., VARSHNEY, R. K. and GRANER, A., 2003. Exploiting EST databases for the development and characterization of

gene-derived SSR-markers in barley (*Hordeum vulgare* L.). *Theor. Appl. Genet.*, **106**: 411-422.

TILLETT, D., DITTMANN, E., ERHARD, M., von DOHRE, H., BORNER, T. and NEILAN, B.A., 2000. Structural organization of microcystin biosynthesis in *Microcystis aeruginosa* PCC7806: an integrated peptide – polyketide synthetase system. *Chem. Biol.*, **7:** 753-764.

TILLETT, D., PARKER, D.L. and NEILAN, B.A, 2001. Detection of toxigenicity by a probe for the microcystin synthetase A gene (mcyA) of the cyanobacterial genus *Microcystis*, comparison of toxicities with 16S rRNA and phycocyanin operon (phycocyanin intergenic spacer) phylogenies. *Appl. Environ. Microbiol.*, **67:** 2810-2818.

TODD, J.A., ATTMAN, T.J., CORNALLI, R.J., GHOSH, S., HALL, J.R., HEARNE, C.M., KNIGHT, A.M., LOVE, J.M., McALEER, M.A., PRINS, J.B., *et al.*,1991. Genetic analysis of autoimmune type I diabetes mellitus in mice. *Nature*, **351:** 542-547.

TOTH, G., GASPARI, Z. and JURKA, J., 2000. Microsatellites in different eukaryotic genomes: survey and analysis. *Genome Res.*, **10**: 967-981.

UTERMÖHL, H., 1958. Zur Vervollkommnung der quantitativen phytoplankton-methodik.

VAN DER LEE, T., DE WITTE, I., DRENTH, A., ALFONSO, C. and GOVERS, F., 1997. AFLP linkage map of the oomycete *Phyophora infestans*. *Fungal Genet. Biol.*, **21**: 278-291.

VAN DER STRATE, H. J., VAN DE ZANDE, L., STAM, W. T., HAROUN, R. J. and OLSEN, J. L., 2003. Within-island differentiation and between-island homogeneity: non-equilibrium population structure in the seaweed *Cladophoropsis membranacea* (Chlorophyta) in the Canary Islands. *Eur. J. Phycol.*, **38**:15-23.

van HANNEN, E.J., van AGTERVELD, M.P., GONS, J.J. and LAANBROEK, H.J., 1998. Revealing genetic diversity of eukaryotic microorganisms in aquatic environments by denaturing gradient gel electrophoresis. *J. Phycol.*, **34:** 206-213.

van HANNEN, E.J., ZWART, G., van AGTERVELD, M.P., GONS, H.J., EBERT, J. and LAANBROEK, H.J., 1999. Changes in bacterial and eukaryotic community structure after mass lysis of filamentous cyanobacteria associated with viruses. *Appl. Environ. Microbiol.*, **65:** 795-801.

VILA, M., GIACOBBE, M. G., MASÓ, M., GANGEMI, E., PENNA, A., SAMPEDRO, N., AZZARO, F., CAMP, J. and GALLUZZI, L., 2005. A

comparative study on recurrent blooms of *Alexandrium minutum* in two Mediterranean coastal areas. *Harmful Algae,* **4**:673-95.

VOS, P., HOGERS, R., BLECKER, M., REIJANS, M., VANDELEE, T., HORNES, M., FRIJTERS, A., POT, J., PELEMAN, J., KUIPER, M., *et al*., 1995. Aflp – a New technique for DNA – fingerprinting. *Nucl. Acid. Res.,* **23 :** 4407-4414.

VRIELING, K., PETERS, J. and SANDBRINK, H., 1997. Amplified fragment length polymorphisms (AFLPs) detected with non-radioactive digoxigenine labeled primers in three plants spcies. *Plant. Mol. Biol. Rep.,* **15:** 255-262.

WALDBIESER, G. C. and WOLTERS, W. R., 1999. Application of polymorphic microsatellite loci in a channel catfish, *Ictalurus punctatus,* breeding program. *J. World Aquacult. Soc.,* **30**: 256-262.

WAN, Q. H., WU, H., FUJIHARA, T. and FANG, S. G., 2004. Which genetic marker for which conservation genetics issue?. *Electrophoresis.,* **25**: 2165-2176.

WANG, X. B., KAKOULIDOU, M., GISCOMBE, R., QIU, Q., HUANG, D. and PIRSKANEN, R., LEFVERT, A. K., 2002. Abnormal expression of CTLA-4 by T cells from patients with *Myasthenia gravis*: effect of an AT-rich gene sequence. *J. Neuroimmunol.,* **130**: 224-232.

WEBER, J.L. and MAY, P.E., 1989. Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *Am. J. Hum. Genet.,* **44:** 388-396.

WEBER, J. L. and WONG, C., 1993. Mutation of human short tandem repeats. *Hum. Mol. Genet.,* **2**: 1123-1128.

WEISSENBACH J., GYAPAY, G., DIB, C., VIGNAL, A., MORISETTE, J., MILLASEAU, P., VAYSSEIX, G. and LATHROP, M., 1992. A second generation linkage map of the human genome. *Nature,* **359:** 794-801.

WELSH, J. and McCLELLAND, M., 1990. Fingerprinting genomes using per with arbitrary primers. *Nucl. Acids Res.,* **18:** 7213-7218.

WERNER, R., OLSCHEWSKI, J. and MERGENHAGEN, D., 2001. Identification and cloning of amplified fragment length polymorphism markers linked to the mating type locus of *Chlamydomonas reinhardtii* (Chlorophyta). *J. Phycol.,* **37:** 427-432.

WHITE, T. J., BRUNS, T., LEE, S. and TAYLOR, J., 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. **In**: PCR Protocols: A Guide to Methods and

Applications. INNIS, M., GELFAND, D., SNINSKY, J. and WHITE, T (Eds.), Academic Press, San Diego, pp. 315-22.

WILLIAMS, J.G.K., KUBELIK, A.R., LIVAK, K.J., RAFALSKI, J.A. and TINGEY, S.V., 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucl. Acid Res.*, **18:** 6521-6535.

WILMOTTE, A., TURNER, S., VANDEPEER, Y. and PACE, N.R., 1992. Taxonomic study of marine oscillatoriacean strains (Cyanobacteria) with narrow trichomes. 2. Nucleotide sequence analysis of the 16s-ribosomal RNA. *J. Phycol.*, **28:** 828-838.

WILSON, W.J., JOINT, I.R., CARR, N.G. and MANN, N.H., 1993. Isolation and molecular characterisation of 5 marine cyanophages propagated on *Synechococcus* sp. Strain Wh7803. *Appl. Environ. Microbiol.*, **59:** 3736-3743.

WOESE, C.R., 1987. Bacterial evolution. *Microbiol. Rev.*, **51:** 221-271.

WOOD, A.M. and TOWNSEND, D., 1990. DNA polymorphism within the Wh7803 serogroup of marine *Synechococcus* spp. (Cyanobacteria). *J. Phycol.*, **26:** 576-585.

WOOD, A.M. and LEATHAM, T., 1992. The species concept in phytoplankton ecology. *J. Phycol.*, **28:** 723-729.

WOODS, I. G., KELLY, P. D., CHU, F., NGO-HAZELETT, P., YAN, Y. L., HUANG, H., POSTLETHWAIT, J. H. and TALBOT, W. S., 2000. A comparative map of the zebrafish genome. *Genome Res.*, **10**: 1903-1914.

WOODRUFF, D.S., 1993. Non-invasive genotyping of primates. *Primates*, **34:** 333-346.

WREN, J. D., FORGACS, E. and FONDON, J. W., 2000. Repeat polymorphism within gene regions: phenotypic and evolutionary implications. *Am. J. Hum. Genet.*, **67**: 345-356.

WRIGHT, J.M., 1993. DNA fingerprinting in fishes. **In:** Biochemistry and Molecular Biology of Fishes. HACHACHKA, P.W., MOMMSEN, T. (Eds.), Elsevier, Amsterdam, pp. 58-91.

YASUMOTO, T., 1990. Marine microorganisms – an overview. **In**: Toxic marine phytoplankton. GRANELI, E., SUNDESTROM, B., EDLER, L. and ANDERSON, D. M. (Eds.), Elsevier Science Publication co. Inc., Newyork, pp. 3-8.

YENCHO, G. C., COHEN, M. B. and BYRNE, P. F., 2000. Application of tagging and mapping insect resistance loci in plants. *Annu. Rev. Entomol.,* **45**: 393-422.

ZANE, L., BARGELLONI, L. and PATARNELLO, T., 2002. Strategies for microsatellite isolation: a review. *Mol. Ecol.,* **11**: 1-16.

ZECHMAN, F. W., ZIMMER, E. A. and THERIOT, E. C., 1994. Use of ribosomal DNA internal transcribed spacers for phylogenetic studies in diatoms. *J. Phycol.,* **30**:507-12.

ZHANG, D. X. and HEWITT, G. M., 2003. Nuclear DNA analyses in genetic studies of population: practice, problems and prospects. *Mol. Ecol.,* **12**: 563-584.

ZHENG, L., COLLINS, F.H., KUMAR, V. and KAFATOS, F.C., 1993. A detailed genetic map for the X chromosome of the Malaria vector, *Anopbeles gambiae. Sci.,* **261:** 605-608.

ZIERDT, H., HUMMEL, S. and HERRMANN, B., 1996. Amplification of human short tandem repeats from medieval teeth and bone samples. *Hum. Biol.,* **68**: 185-199.