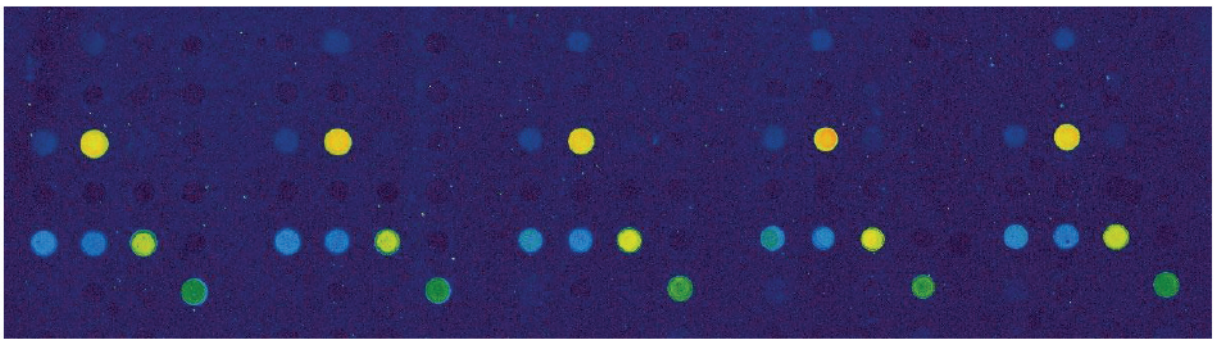
A first prototype of an Invertebrate Chip: Development and Assessment of a DNA microarray as an analytical tool to identify selected marine invertebrates

Srujana Chitipothu

PhD Thesis, University of Bremen

Bremen, July 2008



A first prototype of an Invertebrate Chip: Development and Assessment of a DNA microarray as an analytical tool to identify selected marine invertebrates

Dissertation submitted by

Srujana Chitipothu

In partial fulfilment of the requirements for the degree of
Doctor of natural sciences (Dr. rer. Nat.)

Faculty of Biology / Chemistry

University of Bremen

July 2008

Universität Bremen

UFT
Zentrum für Umweltforschung und Umwelttechnologie

Examination Committee:

First Examiner: Prof. Dr. Dietmar Blohm
Biotechnology and Molecular Genetics, Bremen University

Second Examiner: Prof. Dr. Wolfgang Heyser
Applied Botany, Bremen University

Additional Examiner I: Prof. Dr. Annette Becker
Evolutionary Developmental Genetics Group, Bremen University

Additional Examiner II: Dr. Sascha Todt
Centre of Industrial Mathematics (ZeTeM), Bremen University

Student Member I: Dipl. biol. Kristina Kappel
PhD student, Bremen University

Student Member II: Rohit Nalavade
Master student, Bremen University

Date of public defence: Wednesday, 13th August 2008

Dedicated to my parents

Haribabu and Prameela Rani

and my sister Jwalitha.

their love is my strength…

This dissertation was conceived and written at the Department of Biotechnology and Molecular Genetics, Bremen University, as part of the EC-FP6 grant "fish-and-chip" (Nr 505491), "Towards DNA chip technology as a standard analytical tool for the identification of marine organisms in biodiversity and ecosystem research".

The work was collaborated with the Department of Molecular Genetics for Environmental & Fishery Resources, Bologna University, and supported by the Centre of Industrial Mathematics (ZeTeM), Bremen University, the Alfred-Wegener-Institute, Bremerhaven and the Hellenic Centre for Marine Research, Crete.

Financial support came from the European Commission, grant number GOCE-CT-2003-505491. The writing of the thesis was funded by the Otto-Ritter-Stiftung.

Acknowledgements

I wish to express my sincere thanks to my supervisor "Doktorvater" Prof. Dr. Dietmar Blohm at the first place to have offered me a PhD position in his group, supervising this thesis, providing constructive guidance and inspiration throughout the course of this work. Especially during the last days of writing, in spite of being very busy he took time to correct the thesis and helped me finish on time. His belief in my ability to carry on this work, his constant encouragement and critical discussions during scientific problems in the lab as well as his concern over my personal problems and his caring advices made my life not just better but drove me towards the end without falling down.

I thank Prof. Dr. Wolfgang Heyser for very kindly going through the short summary and immediately agreeing to review this thesis. During the last days of writing when every professor I approached with a request to be the reviewer refused, his words "I will do it" reduced much of the weight from my head.

With warm regards I would like to thank Frau. Prof. Dr. Annette Becker for accepting to be my examiner even during her holiday time and Dr. Sascha Todt for very instantly accepting to be a part of the examination committee. It adds a lot of support and confidence to have you both in the committee.

I want to say a big thank you to Dr. Marc Kochzius for introducing me to the project, explaining me the project, giving timely advices related to experimental and academic queries. I also want to thank him for being so supportive and caring during our conference trip to Taipei.

I would like to say my most sincere thanks to Mr. Haraald Schaaf from the Stifterverband for being very friendly, understanding and helping me twice with the financial support from the Otto-Ritter stiftung. Here, I extend my sincere gratitude to Frau.Dr.Ritter for providing the financial support.

I hereby want to thank Mr. Alex Schöeder from AWI Bremerhaven for allowing me to take part in the "Heincke" cruise to collect the samples and those organisms would have been nameless if Dr. Christos Arvanitidis from the HCMR instititute, Heraklion, Greece did not take up the work and be so kind to teach me the basics of polychaete taxonomy. Along with him I would like to thank Prof. Dr. Andonis Magoulas for giving me the oppurtunity to work in his lab and I extend my warm thanks to Katherina, Sarah, Marian and Maria from the lab for making my stay in Gouves very memorable. At this point I say Danke Schön to the colleagues in the Technomathematics department, the group of Dr. Manfred Nölte including

Hannes, Maria and Janina for being patient with me and working hard to develop the oligonucleotides for this work.

A big part of the credits of this work and a huge thank you is for Ms. Alessia Cariani, University of Bolgna, Italy. This work is incomplete without her. From the start till the end of this work she has an equal share in all the difficulties, agonies, joys, ups and downs of this work. Apart from all this you had been a great friend who was always there for any kind of help and always lent an ear for any kind of talk. Along with her this is the place to show my gratitude and thanks to Prof. Dr. Venugopal Moleyur for sharing part of this work with me, also sharing the lunch for six months and above all being a great help in writing the manuscript of the publication. And then a huge Danke Schön to colleagues in the department each of them personally, especially very big thanks and gratitude and I am totally indebted to Mr. Reinhard Zelm and Mr. Frank Rühle who shared the office room with me and because of that had to bear with all my craziness, and silly questions every minute. I am totally indebted for your help especially during the last hour of submission. You were so kind to take up the responsibilities and do the corrections, formatting, translations what not. I submitted this thesis on time only because of you people. Thank you guys, for being there. And I will not forget the uncountable coffees and chocolates shared with you. Also thank you Kristina for always being patient in listening to me and your discussions during the last few weeks were very useful and encouraging to complete this work. I would like to say special thanks to Mr. Sven Roll and Mr. Frank Meyerjürgens for spotting the microarrays very promptly and for the great help in the lab. I also want to thank Dr. Doris Meyerdierks and Ms. Annemarie Rose for always helping me with the administrative works. I extend this to thank all the others in the department, especially Rijuta, Björn, Janne, Tina, Mrs. Zeng, Christian, Janet, Rohit, Wiebke, Leyla and Claudia. All of you made my work in the lab lively, easy and enjoyable.

Bremen is my second home town now. Had been here for six long years and all my Indian friends have made a home far away from home. I say a big thank you to each and everyone personally. Thank you Aravinda for sharing lunch for 2yrs your care and friendship.

This part is not complete without the mention of my extended family. Budugu nanna, your wake up calls and amma, your care and love have kept me alive. And Budugu without your company, I wouldn´t have stayed so long and without your help I wouldn't have completed this work. I then want to mention my very close aunties Pushpa aunty and Krishna aunty who always encouraged me in what ever I wanted to do and for being there as a huge support. My family's belief in my ability to study and sending me all this far involves tremendous love and sacrifice and I am nothing without them.

Abbreviations

µl	microliter
µm	micrometer
µM	micromolar
A	adenine
a.u.	arbitrary units
bp	basepairs
BSA	bovine serum albumin
C	cytosine
CAG	Centre of Applied Gensensorik
cDNA	copy DNA (from mRNA reverse transcribet DNA)
COI	cytochrome oxidase subunit I
Cy	cyanine
ddH2O	double distilled Water
DNA	deoxyribonucleic acid
dNTPs	deoxinucleotidetriphosphates
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
g	gravitational acceleration
G	guanine
h	hours
HCMR	Hellenic Centre for Marine Research
HOM	homogenate buffer
kb	kilo base pairs
min	minutes
mL	millilitre
mM	millimolar
mRNA	messenger ribonucleic acid
mt	mitochondrial
NaCl	sodiumchloride
nM	nanomolar
nt	nucleotide
PBS	phosphate buffered Saline
PCR	polymerase chain reaction
PDITC	phosphodiisothiocyanat
PMT	photo multiplier tube
rpm	revolutions per minute
S	Svedberg unit
SDS	sodium dodecyl sulfate
sec	seconds
SNP	single nucleotide polymorphism
SSC	sodium chloride sodium citrate
T	thymine
TBS	tris-buffered Saline
TE	buffer with Tris-HCl, EDTA
TETBS	buffer with Tris-HCl, NaCl, EDTA, Tween-20
Tm	melting temperature of complementary nucleic acids
Tris-HCl	2-amino-2-(hydroxymethyl) propane-1,3-diol hydrochloride
UNIBO	University of Bologna
V	Volt
16S	16S rRNA (16S ribosomal Ribonucleicacid)

Table of contents

List of Figures

List of Tables

List of publications

Parts of the results in the following work have been published or presented in:

Publication

Chitipothu S., Cariani A., Venugopal M., Landi M., Mählck M. T., Jäger J., Kochzius M., Arvanitidis C., Weber H., Nölte M., Tinti F., Magoulas A., Blohm D. A first DNA-microarray for the identification of selected groups of marine invertebrates. Submitted.

Poster

Chitipothu S., Cariani A., Venugopal M., Landi M., Mählck M.T., Jäger J., Kochzius M., Arvanitidis C., Weber H., Nölte M., Tinti F., Magoulas A., Blohm D. (2008). „Invertebrate chip" - first microarray to identify selected groups of marine invertebrates in European seas. Statusseminar Chiptechnologien, Frankfurt, Germany.

Chitipothu S., Cariani A., Venugopal M., Landi M., Mählck M. T., Jäger J., Kochzius M., Arvanitidis C., Weber H., Nölte M., Tinti F., Magoulas A., Blohm D. (2007). Towards microarray-based DNA-barcoding of marine invertebrates. 2nd International Barcoding conference, Taipei, Taiwan.

Cariani A., Landi M., Bertasi F., Chitipothu S., Colangelo M., Ceccherelli V.U., Blohm D., Tinti F. (2006). Development and prototyping of an Invertebrate Chip as high-resolution molecular tool to investigate marine food webs. Marine Genomics Conference, Sorrento, Italy.

Presentation

Cariani A., Landi M., Bertasi F., Chitipothu S., Colangelo M., Ceccherelli V.U., Blohm D., Tinti F. (2007). Development of high-resolution molecular technologies to investigate biodiversity and ecosystem changes in the Southern European Seas: the prototyping of DNA microchips to identify invertebrate prey species in marine food webs of Adriatic Sea. 38th CIESM Congress, Istanbul, Turkey.

Kochzius M., Kappel K., Döblitz L., Silkenbeumer N., Nölte M., Weber H., Hjörleifsdottir S., Marteinsson V.P., Hreggvidsson G., Planes S., Tinti F., Magoulas A., Garcia Vazquez E., Turan C., Medlin L., Metfies K., Gescher C., Cariani A., Landi M., Hervet C., Campo Falgueras D., Antoniou A., Bertasi F., Chitipothu S., Blohm D. (2007). The "Fish & Chip" project: Microarrays as a tool for the identification of marine organisms in biodiversity and marine research. Ocean 07 conference, Aberdeen, Scotland.

1. Summary

Apart from gene expression studies DNA microarrays provide a valuable potential for other fields of research where there is a need for highly parallel detection of a large number of organisms such as in marine ecology for food web and biodiversity studies and for species identification in general. This is of special demand with the diminishing number of taxonomical expertise based on morphological characters.

The development of microarrays and their application in environmental studies depends on finding of highly specific capture oligonucleotides and on the optimization of specific experimental conditions for the reliable identification of a single species in presence of large amount of background samples.

For enabling the utilization of this technology various methodological problems involved in doing so were extensively studied in this work ending in the development of a microarray prototype containing 64 oligonucleotide capture probes targeting the 16S and COI genes of 15 common scientifically and economically important invertebrate species in the European seas.

Depending on sequences of target species newly obtained during this work and on sequences of target and non-target species available in the database a reliable set of probes has been designed that satisfies the criteria of the specificity *in silico* and includes redundant probes after the optimisation of hybridisation conditions.

The specificity and the sensitivity of the chip were investigated *in silico* and experimentally under appropriate optimized hybridization conditions. Specific signals were observed from respective captures on the chip resulting in clear distinction of the species analyzed. False positive signals could be significantly reduced during these investigations. Designing the probes from both the 16S and COI genes allowed the improvement of the accuracy of the identification even of very closely related species showing both inter and intra specific variations at the level of individuals.

As a first application of the microarray its performance in terms of multiple target experiments was tested with defined mixtures of various different species in different compositions mimicking the environmental samples. The ability of the chip to distinguish a single low abundant target species in the presence of large amount of background samples was also checked by using various non-target invertebrate DNA. The applicability of the chip to perform gut content analysis of predator fish was investigated by hybridizing DNA extracted from the gut content of the demersal fish *Trigla lucerna*. In concord with estimates using classical methods a mixture of DNA belonging to various invertebrates were identified

using the microarray. These results were validated by cloning and sequencing the mixed gut content sample.

These experiments show that the invertebrate chip can provide semi quantitative information on the important prey species of demersal fishes and also in turn provide information for the food web studies of marine ecosystem.

As marine invertebrates are very tiny and often troublesome tissues for DNA extraction, a small comparative study of classical and commercial DNA extraction methods has been carried out to define a hopefully universal method which can be applied on any kind of tissue and can extract good amount of DNA. “FTA Elute” was found to be a very reliable, cheap and convenient method to extract and store DNA even outside of a laboratory under field conditions.

In summary a fast and relatively easy-to-handle DNA-analytical tool for the identification of morphologically difficult to estimate, important marine invertebrates as well as for analysing complex environmental mixtures of organisms has been worked out in the framework of this thesis.

2. Introduction

2.1. Trends in development of DNA analytics

The discovery of the double helical structure of deoxyribonucleic acid (DNA) by James Watson and Francis Crick (Watson and Crick 1953) has revolutionized our understanding of genetics and cumulated in the decipherment of the complete genome of the human and meanwhile of more than 1000 other organisms. This new field which is called genomics utilizes this sequence information to identify, isolate and investigate thousands of genes simultaneously. Research in genomics is mainly concentrated in biomedical problems as identifying genes contributing to heritable diseases, studying at the molecular level drug actions and the molecular architecture of cancer cells as well as other diseases or to identify pathogens, to name just a few.

Meanwhile the genomics revolution reached the field of biodiversity and uses the new techniques to find previously unknown life forms in the oceans; to learn how species and genes evolved in time; to understand the genetic tools that allow species to adapt to diverse and harsh environments; and to investigate species responses to pollutants. New and sophisticated techniques became increasingly available to gain more and more knowledge at the molecular genetic level (Avise 1994; Burton 1996; Ferraris 1996).

One very new application is the identification of any organism and assigning its taxonomical and phylogenetic relationships via DNA sequences. In fact, this molecular approach to species identification is considered as the most efficient option to document all life forms present on earth for biodiversity conservation and monitoring. It is called “DNA barcoding” (Hebert et al. 2003a) and deals so far with insects, birds, plants and primarily with fish and other marine organisms. Using the new methods of genomics two theoretical problems are under investigation as the genetic variation and understanding of the population structure, levels of gene flow, patterns of historical geography and analysis of parentage and relatedness of marine organisms at one hand, and their identification and traceability as food and feed, their monitoring to watch their stocks to guarantee sustainability of their use and other ecological and biotechnological aspects on the other.

2.2. Molecular identification of species

The classical method of species identification is descriptive and mostly depends on variable morphological characters. Such kind of identification needs a lot of expertise especially if closely related species are to be differentiated. This approach becomes very complicated if the

creatures are very small, if only small fragments of them are available, if the morphology is very variable during ontogenesis. Species identification is an absolute necessity for many marine ecological studies but it is often difficult in day by day investigation especially in case of invertebrates since they can often not be estimated down to the species level by morphological characters alone (Schander 2005). With the decreasing number of classical taxonomists (Mallet 2003) it is essential to utilize the new molecular genetic methods, especially in case of closely related species since classical taxonomic methods are often problematic and error-prone (Dahlgren et al. 2000). If for example only parts or traces of an unknown organism which lack the classic taxonomic characteristics are to be identified, for instance in ecological studies (Pfunder et al. 2004a), forensics (Wan and Fang 2003), quality control of processed food (Jerome et al. 2003; Sebastio et al. 2001), or screening of gut content of fishes for food web studies (Deagle et al. 2005), a molecular-based method is often the only possible way of identification.

As more and more approaches are focussing in this direction, universally usable genes for taxonomy are discovered which are believed to be useful for the classification of the whole animal kingdom (Hebert et al. 2003b; Palumbi and Cipriano 1998). Although some researchers remain sceptical, stating that genetic approaches based on one gene only might not be enough to fully characterize the planet's diverse flora and fauna (Mallet 2003), a combined system preferably based on more than just one gene, can probably contribute to and profit from this field.

2.3. DNA extraction

A prerequisite for the application of any molecular technique is the extraction of DNA from any kind of organisms under investigation. Reliable and time-efficient methods for isolating high-quality nucleic acid are essential for the success of PCR based techniques and in ecological and biodiversity studies. A universal DNA extraction method is a necessity if a wide range of animals especially marine invertebrates are to be investigated since many of them contain substances which are not easy to separate from DNA and can disturb the DNA-analytical techniques.

Traditional methods such as the phenol chloroform extraction, the Chelex method or a range of commercial kits are commonly used on marine invertebrates (Chen and Yu 2000). These conventional techniques use organic solvents such as phenol and chloroform to inactivate the enzymes that degrade the DNA and to denature the proteins and release the pure genomic DNA. Apart from the time consumption for this technique the solvents may be health hazards

and their safe disposal is considerably expensive. Along with this various factors such as the degree of the cell lyses, binding of the DNA to particulate material (silica membranes of commercial kits) and degradation and shearing of the DNA during the extraction steps may adversely affect the DNA recovery (Miller et al. 1999), which influences the downstream processes such as marker gene amplification and sequencing. The collection, transport and the preservation of the tissue itself also has a great influence on the final quality of the extracted DNA.

As this study involves diverse organisms belonging to different phyla, the tissue of the samples varies in structure and hence the treatment for DNA extraction has to be different because usually different tissues require different protocols. Therefore a universal but easy-to-handle extraction method usable in the laboratory as well as in the field to efficiently extract DNA from many different tissues is needed especially in studies involving the analysis of hundreds of environmental samples. This method should reduce nucleic acid degradation during sample collection, transport to the laboratory and long time storage. Modifying the classical methods has not been very successful so far (Aljanabi and Martinez 1997) as shown by a comparative analysis of commercial kits and optimized protocols (Rohland and Hofreiter 2007).

A promising new approach was introduced into the market about two years ago from Whatman Schleicher & Schuell (Whatman GmbH) called “FTA elute” paper. FTA is a chemically treated paper that is designed to collect the samples, store the samples and efficiently lock the DNA in the network of the paper until elution for subsequent DNA analysis. FTA paper have been used efficiently for DNA extraction for biomedical purposes and are widely used in pathological, forensic and diagnostics laboratories for collecting and transporting blood samples (Devost and Choy 2000) as well as for storage and molecular testing of tumour samples (Dobbs et al. 2002). FTA was found to be effective in archiving and processing of mammalian samples and wild life samples (Natarajan et al. 2000; Smith and Burgoyne 2004), since it stores nucleic acids in the dry stage at room temperature and allows easy detection of plant genes (Lin et al. 2000) and those of protozoan and zooxanthellae species whose handling is otherwise cumbersome (Crabbe 2003; Hide et al. 2003). Evaluation studies of FTA paper in comparison with other methods are mostly done with human diagnostic samples with convincing results in terms of being convenient, easy and safe to handle during transport and archiving of the samples. It is used simply by spotting or pressing the sample together with some water or tissue fluid on to the paper and dry it. To elute the DNA a very small part of the paper is removed and heated in water for about 30 min

at 95 °C. The quality of DNA obtained this way is reported to be comparable to the costly commercial kits and can be used for PCR, sequencing or any other downstream processes (Coyne et al. 2004; Subrungruang et al. 2004). However no literature has been published up to now using the FTA elute paper for DNA extraction from marine invertebrate tissue.

2.4. Molecular markers

Genes being unique for similar organisms and useful as "signature sequences" are the basis for the molecular identification of various organisms, so called marker genes. Earlier studies used often proteins for the identification of species but recently DNA is more favoured for this purpose as proteins undergo quick degradation (Sebastio et al. 2001), can not be sequenced easily and the amino acid sequence of two closely related species is not different since the mutation in which they differ often occurs in the wobble position.

The genes for molecular identification have to be very specific and conserved between individuals of that species but as diverse as possible among the various species. Both nuclear and mitochondrial genes have been used for various ecological and phylogenetic studies (Chen et al. 2002; Palumbi and Cipriano 1998; Weinberg et al. 2003). Though few studies showed successful utilization of nuclear ribosomal RNA genes in distinguishing closely related species (Chen et al. 2002), they are mainly used to study relationships of long evolutionary divergences among organisms (Remigio and Hebert 2003). Mitochondrial genome usually shows a evolutionary rate 10 times faster than the nuclear genome (Awadalla et al. 1999), hence they show high levels of polymorphism and are subject to maternal inheritance. This makes mitochondrial genes very useful for DNA-analytical studies of marine organisms specially to identify differences between closely related species and at the same time these genes also show distinction between distantly related taxa (Maggioni et al. 2001; Remigio and Hebert 2003). Many studies have already been published using these genes for DNA based identification of marine organisms as for example for molecular discrimination of planktonic copepods (Bucklin 1999), genetic screening of the gut of giant squid (Deagle et al. 2005), study of marine invertebrate diets (Blankenship 2005), molecular identification of eggs, larvae and adult fishes (Fox et al. 2005; Kochzius and Blohm 2005; Kochzius et al. 2003; Ward et al. 2005) and molecular identification of marine invertebrate larvae (Goffredi et al. 2006).

Three mitochondrial genes are widely used in species identification: 16S rDNA (16S), cytochrome B (Cyt b) and the cytochrome oxidase I (COI). The gene coding for the mitochondrial enzyme cytochrome b (Cytb) is the most popular and commonly used marker

to evaluate the phylogenetics, phylogeography and population analysis of diverse marine organisms (Jerome et al. 2003; Sebastio et al. 2001). An exclusive study by (Parson et al. 2000) showed that the evaluation of the cytb gene from the DNA of 44 different animal species belonging to five major vertebrate phyla is possible and proclaimed that cytb can be used to identify any animal species. However no major investigations using cytb to identify marine invertebrates are done. Mostly this method is used for the phylogeography studies of fishes (Radchenko 2004) since the gene turned out to be diverse among the sub populations of fishes investigated in that study, but it is thought to be too diverse to distinguish very closely related species.

The 16S rRNA gene is found in the highly conserved region of the mtDNA. It has been widely used to study the genetic differences between and among many marine species (Weinberg et al. 2003) and also to study microbial diversity (Amann et al. 1996; Amann 2001; Koizumi et al. 2002; Peplies et al. 2003; Rudi et al. 2000; Small et al. 2001; Woese 1998). Though it is extensively used in phylogenetic studies, it is found to be not very useful for recognising intra specific variation. As the gene sequence is found to be highly similar in closely related species, it is very difficult to use it all alone as an identification system for organisms belonging to the same genus. But the gene has been reported to distinguish the inter-specific variation very well and is therefore used as one of the markers for species identification.

COI gene has been used as a model for identifying organisms and for evaluating evolutionary relationships among closely related species (Hebert et al. 2003b; Remigio and Hebert 2003). Many studies show the successful use of this gene in phylogenetic studies and molecular discrimination of closely related species and many times it is used as molecular evidence for the morphological identification of marine invertebrates (Blair et al. 2006; Dahlgren et al. 2000; Nygren et al. 2005). The use of CO1 sequence data for species differentiation of animals as for example, birds (Hebert et al. 2004), fishes (Ward et al. 2005), moths (Hebert et al. 2003b; Janzen et al. 2005), spiders (Greenstone et al. 2005) and spring tail (Hogg and Hebert 2004) has also been reported. More over an exclusive study was made to develop universal COI primers that can successfully amplify the COI gene from 11 different invertebrate phyla and the sequences thus obtained are useful for phylogenetic analyses at species and higher taxonomical levels (Folmer et al. 1994). Thus among several molecular markers the superiority of mitochondrial cytochrome c oxidase subunit 1 (COI) is proposed as global bio-identification system (DNA barcoding) for animals (Hebert et al 2003) because of being superior over nuclear genes, lack of introns, limited exposure to recombination and

haploid mode of inheritance (Saccone et al. 1999), low variation among individuals of one species and high variability between species. This makes COI particularly useful for population genetics, evolutionary studies and systematics (Harvey et al. 2003).

With the onset of DNA-analytics and the detailed study of the sequence markers a vast variety of techniques have been developed for detecting the genetic variability of organisms to provide an overall understanding of the biodiversity, ecosystems management, environment assessment and marine genomics (Feral 2002).

2.5. PCR techniques

The invention of the polymerase chain reaction (PCR) by Kary Mullis (Mullis et al. 1986) and the first successful use of it has paved way to many modern molecular techniques (Saiki et al. 1988). The PCR technique is basically a primer extension reaction for amplifying specific nucleic acids in vitro. The use of a thermo stable polymerase referred to as Taq, (isolated from *Thermus aquaticus* a National Park hot springs bacterium) allows a short stretch of DNA (usually fewer than 3000 bp) to be amplified to about a million fold so that one can determine its size, nucleotide sequence etc. The particular stretch of DNA to be amplified, called the target sequence, is identified by a specific pair of DNA primers, oligonucleotides usually about 20 nucleotides in length. The quantities of produced DNA are sufficient to be directly visualized on a gel by fluorescence after coloration with stains such as ethidium bromide. The DNA may also be digested by restriction enzymes or can be sequenced.

The main advantage is that it is now possible to work with a very small initial amount of DNA. It is the most commonly used technique for molecular investigations. A major limitation to this approach is the utilization of one specific primer pair per gene detection reaction. Although multiple primer sets may be successfully combined in one reaction (Henegariu et al. 1997), they rarely exceed more than about ten primer sets due to the generation of non specific products or false negatives. Another difficulty with multiplex PCR is that it requires additional post amplification analysis to discriminate the products. Size separation by electrophoresis is frequently used to discriminate multiplex PCR products, but this requires additional labour and that the amplicons of each reaction be significantly different in size which can limit the primer pairs that can potentially be multiplexed (Peplies et al. 2004; Small et al. 2001).

2.6. DNA sequencing

Nucleotide sequences of DNA and RNA offer the greatest genotype resolution as these methods survey the nucleotides themselves. Sequences can be obtained either by direct

sequencing of genomic DNA using specific primers to amplify via PCR targeted regions of the genome or by cloning such regions and sequencing from the cloned fragment. The most commonly used sequencing method is the Sanger di deoxy method (Avise 1994; Ferraris 1996; Sanger et al. 1977). This method is the basis of most automated sequencing equipment and is still the preferred method, although pyrosequencing based devices, probably the next generation of sequencing machines, are already on the market (Hudson 2008). High quality DNA and specific PCR primers for the amplification of the desired fragment of the genome are the prerequisites for obtaining good sequence data. Sequences are the basis for various phylogenetic, ecological, evolutionary, phylogeographic studies and it is the main requisite in the barcoding of life project. Majority of the sequences developed are collected and maintained by the public database of the National Centre for Biotechnology Information (NCBI) based on the GenBank (Benson et al. 2008). Today around 150 thousand polychaete sequences with complete genomes of 8 taxa, around 600 thousands of crustacean sequences with 38 complete genomes and around 927 thousands of molluscan sequences along with 42 complete genomes are present in GenBank. An unknown organism is identified by comparing its DNA sequence with the known database sequences and looking for the similarity criteria thereby deciding the exact species or a close relative. The disadvantage of this method is that the sequence is governed by a prior PCR reaction where there is a possibility of wrong bases getting incorporated by the Taq polymearase (Chen et al. 1991) and the authenticity of the sequences submitted to the public databank is also sceptical (Clayton et al. 1995). Metagenomics, sequencing DNA from whole population is randomly gaining interest to be used in the evolutionary and environmental studies either to identify an unknown species or to understand the adaptation of a species to a particular environment (Chen and Pachter 2005).

Apart from the above mentioned basic techniques for most of the molecular studies a bunch of other methods are also applied for the ecological, environmental and phylogenetic assessment of marine organisms. DNA finger printing technique has been used for the analysis of complex environmental samples showing hyper variability within an individual's genome, by developing a DNA fragment band pattern (Jeffreys et al. 1985). This method has been successfully applied for the identification of individual samples and also to measure the growth of the invertebrate Holothuria in the wild (Uthicke. S. 2002). Reverse transcription (RT) PCR based essays have become more common for the quantification of gene expression data and also for comparing the mRNA levels in different gene populations (Orlando et al. 1998).

As per the demand for various applications this technique is undergoing rapid modifications and many versions of the same are presented in the form of standardized RT-PCR (Willey et al. 1998), real time reverse transcription PCR for the quantification of the mRNA levels (Bustin 2002) and fluorescence based quantitative PCR (qPCR) (Nolan et al. 2006). This technique is very much favourable for the simultaneous analysis of large number of samples like screening of hundreds of invertebrates for viruses (Gomez et al. 2008), identification and quantification of invertebrate larvae (Vadopalas et al. 2006) and also to investigate physiological and genetic structures of marine invertebrates (Clark 2008; Voznesensky et al. 2004).

Though the above methods seemed promising for the genetic evaluation of organisms, they are very expensive in terms of the equipment and the consumables and the application of these techniques in the field is not so feasible. Some of the disadvantages in these methods are overcome by the hybridization techniques.

2.7. Hybridization techniques

Hybridization technique turned in to a fundamental tool of DNA detection in modern molecular genetics, based on the use of labelled nucleic acid probes. Nucleic acid hybridization involves mixing single strands of two nucleic acids, a probe which consists of the known sequence and a target which is complex and unknown nucleic acid molecule that has to be identified. This binding occurs with a prerequisite of sufficient base complementarities of both the molecules. Hybridization technique provides a promising platform to detect huge populations of unknown organisms present in a complex environment, and also the relative frequencies of the abundance of selected population can be determined.

Wide range of hybridization methods (Fig.1) have been used for different applications such as identification of pico-plankton algal groups with dot blot and fluorescent whole cell hybridization was done by designing specific probes for three different groups of taxa (Simon et al. 2000), detection of the sulphate reducing bacteria in coastal sediments was possible by dot blot technique (Sahm et al. 1999), and it is also widely used for the quantization of gene expression and in the screening of alleles differing at a single nucleotide position (Yadetie et al. 2004).

Fluorescent in situ hybridization (FISH) depends on the specific binding of the fluorescent probes to the complementary regions on the DNA sequences and evaluation of the binding by fluorescence microscopy. Fluorescently labelled nucleic acid probes are regarded as promising tools for the environmental assessment of microorganisms (Amann et al. 1996) and

hence this technique is widely used in the filed of microbial ecology, to identify microorganisms and bio films (Amann 2001). As FISH is used to find specific features in the DNA it can also be applied to species identification of organisms. Only a few studies have been published in this scenario but *in situ* hybridization has been successfully used in marine ecology by tracking the transitory of marine invertebrate larvae, an important component of the plankton communities (Le Goff-Vitry 2007).

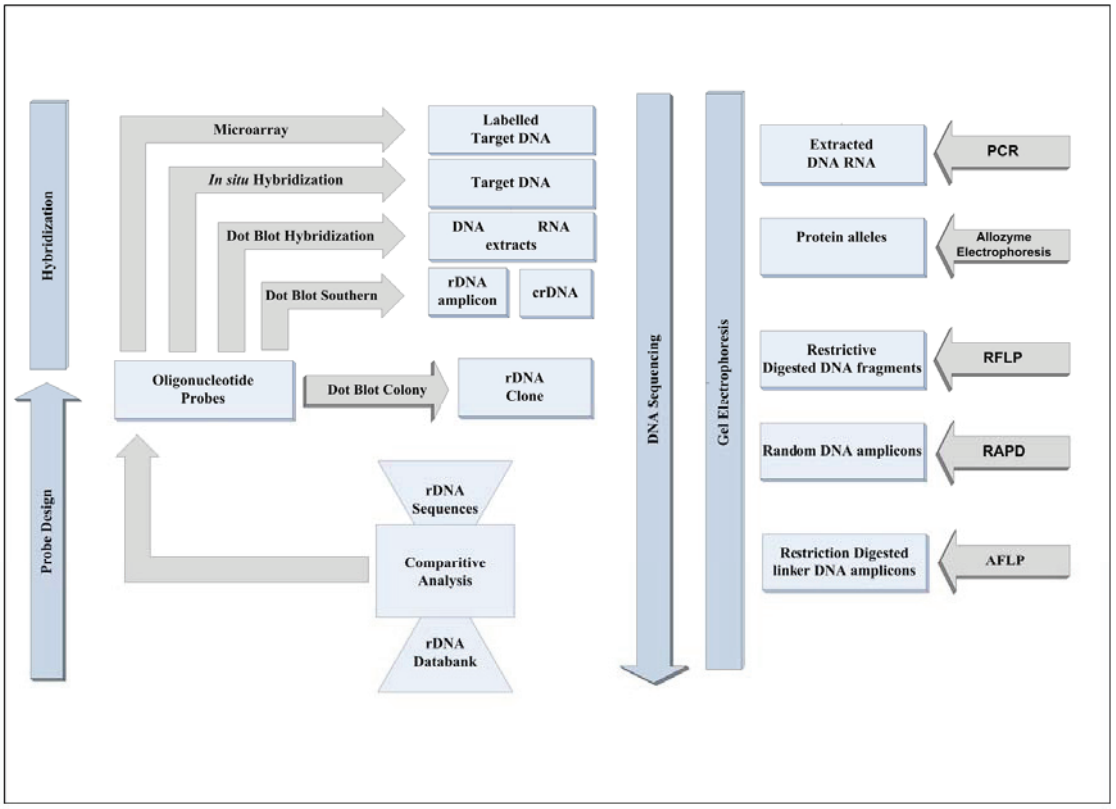


Fig. 1 Scheme showing different molecular techniques, their principle and the functioning.

2.7.1. Origin and definition

DNA microarray technology has a special place in the molecular biological techniques as its introduction brought a totally new dimension to the genetic experiments. Until few years ago, knowing the function and sequence of a single gene was a time consuming and tiresome issue. But now, with the advent of the microarrays, the complete transcriptome of an organism can be analysed with a single experiment. The main advantage of DNA microarrays is that they allow parallel hybridizations of hundreds to thousands of nucleic acids on a small surface area, mostly with a short DNA strand for identification (Service 1998). The first idea of identifying a DNA with radio active labelled short single stranded DNA molecules started in 1975 (Southern 1975), where the DNA is immobilized on a membrane and labelled short strands upon hybridization allowed identification. After fourteen years, in 1989, the reversing of the above design started the microarray revolution where many DNA could be identified

simultaneously by immobilizing many sequence specific short DNA strands on the membrane (Saiki et al. 1989).

DNA arrays found their first major application in quantitative monitoring of gene expression in 1995 (Schena et al. 1995) and the first array with the complete eukaryotic genome was produced in the year 1997 (Lashkari et al. 1997). From then on, there has been a rapid exponential growth (Fig. 2 The increasing significance of the DNA microarrays. A report of the number of publications in the ISI web of science (http://www.isiwebofknowledge.com/) during the years 1995 – 2008, searched with the keywords microarray and filtering gene expression. (Fig. 2) in the microarray research, majority of it being in the field of gene expression profiling, and very little in other areas. More and more applications are being explored till date.

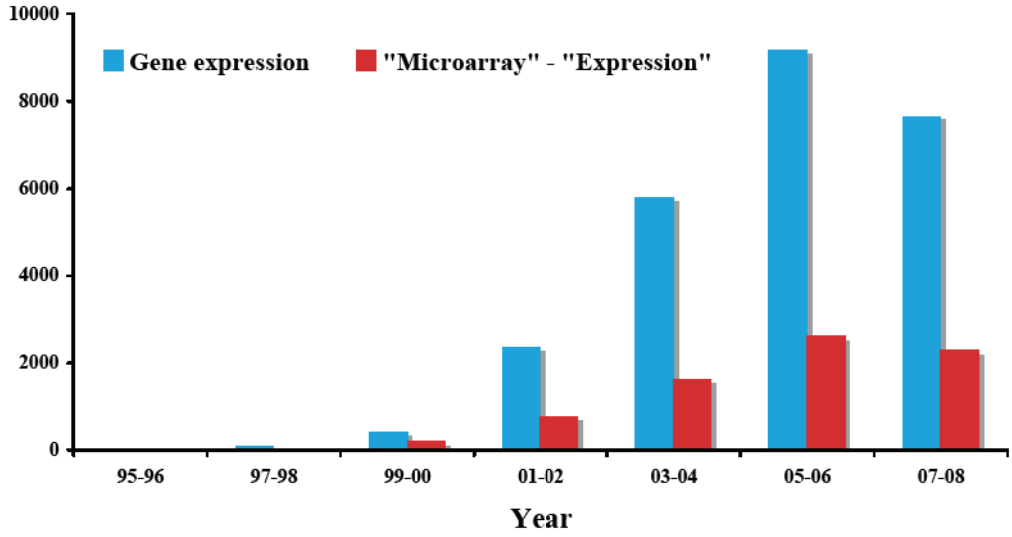


Fig. 2 The increasing significance of the DNA microarrays. A report of the number of publications in the ISI web of science (http://www.isiwebofknowledge.com/) during the years 1995 – 2008, searched with the keywords microarray and filtering gene expression.

Today a basic DNA microarray also designated as a “DNA-Chip“ consists of a firm base, e.g. a glass slide and single stranded nucleic acid molecules are attached at a defined position (“spots“) whose base sequence is complementary to the sequence of the DNA to be identified (“target“). The chip surface thus prepared, when treated with the target molecules in a buffer solution and incubated at controlled conditions, allows specific binding of the probe and target. By the end of the reaction, the detection of the hybridization is facilitated by fluorescent dyes incorporated in to the target molecules which emit light when excited by a laser (Fig. 3). Targets are characterised by detecting the fluorescent signals at specific positions on the chip.

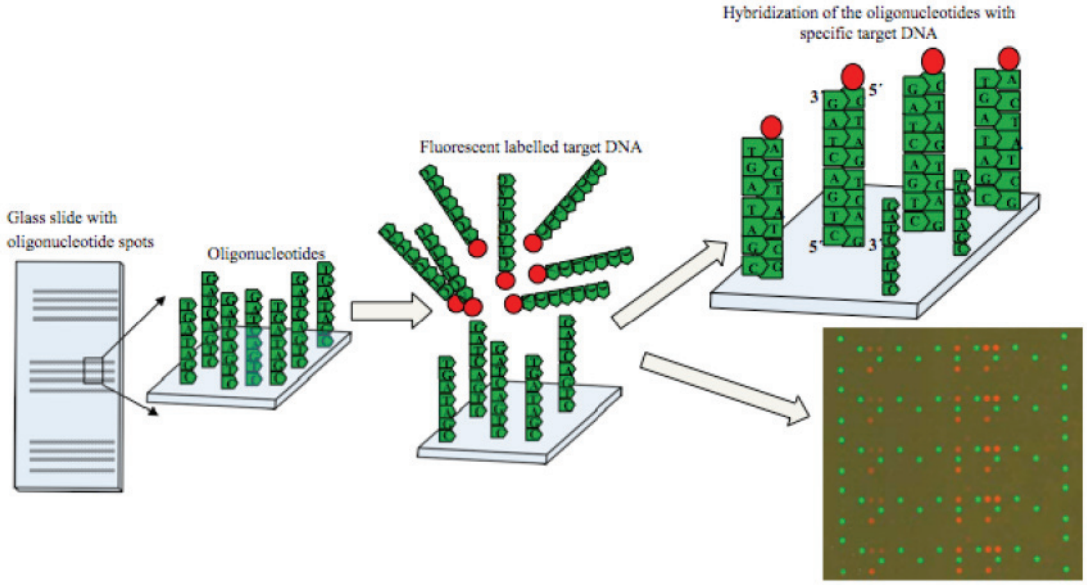


Fig. 3 Principle of DNA microarrays. Various oligonucleotides are immobilized on glass slides at fixed positions (shown in left side scheme). Fluorescently labelled target DNA are then added to the captures in buffer conditions and incubated (centre scheme). The binding of the target to specific captures can be visualized by the fluorescent signal (right side scheme).

2.7.2. Critical factors for the development of microarrays

From "low-density-chips" with a few hundred probes (macroarray) to "high-density-chips" (microarray) (Lipshutz et al. 1999) which can hold as many as a million spots, a wide variety of arrays are available commercially including chips that contain whole genome, chips that use nano crystallized beads (Qbeads) (Xu et al. 2003), fibre-optic gene arrays (Steemers et al. 2000), flow through porous microarrays (Wu et al. 2004).

DNA microarrays have become an essential tool in many areas over the years (detailed in 2.7.3) and the technology continues to grow, being applied to molecules other than DNA. Immobilization of the capture probes on glass surfaces instead of nylon membranes facilitated the use of fluorescent labels for hybridization detection and miniaturized the whole detection area, leading to the use of less sample material. This, along with the development of the process to synthesize oligonucleotides directly on the slide, patented by Affymetrix (Affymetrix, Santa Clara) (Fodor 1995) are considered to be the two important milestones in the evolution of microarray technology (Niemeyer 1999). Spots today may contain proteins, small organic compounds like carbohydrates, tissue samples or even living cells (Howbrook et al. 2003). Similarly, faster and less expensive bead coupled arrays are coming into existence unlike the slides to bind the probes (Kohara et al. 2001). A „lab on a chip" is also being aimed by various researchers around the world using different approaches. (Kuo et al.

2002). High possibility of automation is the critical factor driving interest towards the development of DNA microarray technology even though sometimes the arrays are lacking consistency of results and involve lengthy, complicated protocols. The critical factors that have to be considered during a typical DNA microarray experiment are the design of specific oligonucleotide probes, activation of the glass surface, printing of capture molecules on the surface, labelling of the target molecules before hybridizing them onto the array, optimizing the specific hybridization conditions and scanning of the slides and evaluation of the data (Blohm and Guiseppi-Elie 2001).

2.7.2.1. Capture probe design

The major parameter for the success of any microarray experiment is the design of specific oligonucleotides. The critical challenge in the designing of probes is to find a number of captures, which when used in the hybridization experiments as capture probes, can show high specificity in the same conditions and are able to distinguish between highly similar target molecules. Many factors that affect the specificity and sensitivity have to be considered while designing the probes along with the melting temperature (T_m) (Kaderali and Schliep 2002). Different studies stated a range of criteria that include, total amount of As or Ts should be less than ten, while a total amount of more than six Gs or Cs per probe is recommended. No more than six As and Ts should be in a row, palindrome sequences should be shorter than seven nucleotides, and the probes should not be able to form hairpin loops (Relogio et al. 2002). But increasing the length of the probes significantly reduces the specificity (Relogio et al. 2002). 60mer showed stronger signals but the signals are less specific than 20 mers. Steric hindrance is also an important factor influencing the signal intensity. It is clearly observed in PCR probes of length 200 nt that they give weaker signals than probes of length 20 – 70 nt (Relogio et al. 2002). In contrary, addition of spacers such as poly A and poly C to the probe increase the signal intensity without lowering the specificity (Peplies et al. 2003). A more recent investigation has discussed the influence of the position of the binding of the capture to the target, on the signal intensity. The nearer is the probe binding site to the fluorescent label, the stronger is the signal. (Zhang et al. 2005). The position and the number of mismatches is another important factor influencing the probe specificity. Mismatches present at the terminal positions do not hinder the binding of the target to the specific probes as shown in the study with the poly A linkers, whose presence did not influence the target binding (Peplies et al. 2003). For a good capture, it is preferable to have many mismatches (to the non target sequences) distributed all through the length of the probe because, if a continuous stretch of nucleotides in the capture can bind to the DNA that has not been targeted, unspecific signals

are observed (Bodrossy and Sessitsch 2004; Relogio et al. 2002). A vast number of algorithms have been in use for designing oligonucleotide probes. *In silico* design of captures is a crucial issue, as most of the times the probes developed do not behave in the same way in the experiment. So the algorithms have to be efficient enough to consider all the possible influencing factors before generating the output (Nölte 2002).

2.7.2.2. Immobilization of the captures

Efficient immobilization of the captures to the microarray substrate is important for the success of the whole hybridization process. Wide ranges of pre-activated slides are available these days, a few of them including a coating of poly-L-lysine (Sigma), organoaldehydes (Quantifoil Micro Tools, Telechem), silanes (Quantifoil Micro Tools, Eppendorf) or Gamma-aminopropylsilane (Corning-slides). Oligonucleotides are modified according to the surface chemistry of the slides for efficient binding. The glass slides used in this work are first treated with 3-Aminopropyltrimethoxysilane and then coated with 1, 4 phenylene-diisothiocyanate (PDITC) (Fig. 4). Accordingly the captures are modified at the 3' or 5' ends with an amino group linked by six carbon atoms. By the covalent bonding to the glass surface, the capture nucleotides stand upright on the slide leading to greater density in the spot, reducing the steric hindrance from the surface for the binding of the target (Shchepinov et al. 1997).

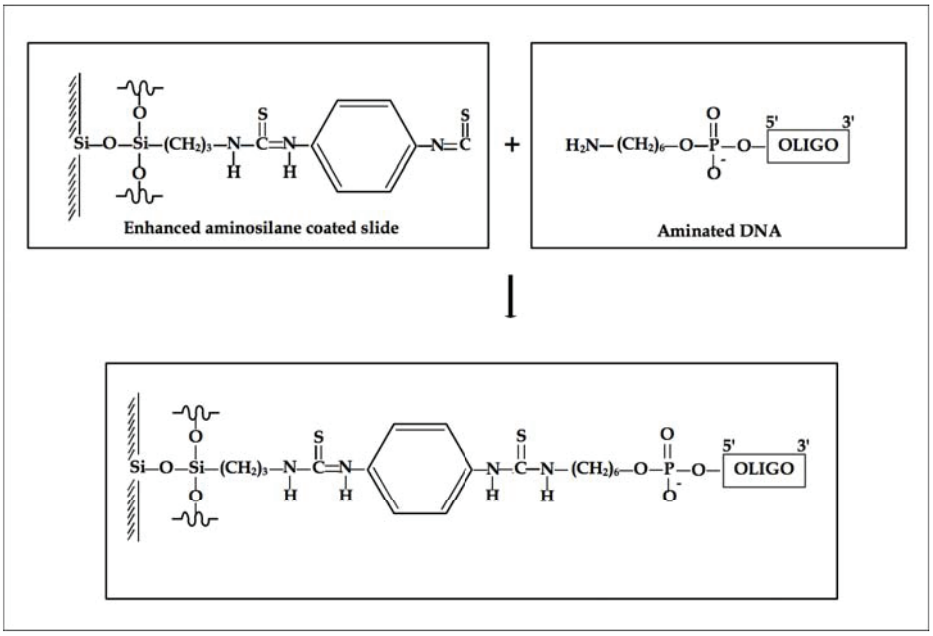


Fig. 4: Amino modified capture with a six carbon linker covalently binds to the glass slides activated with 1, 4 phenylene-diisothiocyanalte (PDITC).

For the printing of the capture molecules on the slide, „spotting" three methods are widely used. Micro spotting works on the principle of adhesion between the slide and the probe

solution injected with a needle. Though it is very easy to manufacture and clean microarray slides by this method, it also has its disadvantages, if spot inhomogeneity and leakage (Bowtell 2003). Another method is micro spraying and it works like an ink jet printer. Next method is the touch-less transmission of the capture solution, where tiniest quantities are sprayed on the surface of the slide in predefined fashion reducing the diameter and increasing the density. In both the procedures the capture sequence accuracies are determined before hand, in contrast to the photolithographic procedure, where the DNA sequence is determined by modifying the nucleotide bases to the photo specific bases and activating them by a mercury lamp. A photolithographic mask determines the activated bases with the help of light. Position of the different bases is determined by using different masks (Lipshutz et al. 1999). The main disadvantage of this method is the time taken for the production of masks and the expense of production.

2.7.2.3. Labelling of the target DNA

To detect the hybridization experiments, either the probe or the target has to be labelled. In case of microarrays, targets labelled as the probes are immobilised on the slide. Labelling is usually done by attaching fluorophore molecules that emit fluorescence when exposed to certain wavelength of light. Different fluorophores generating different colors are also used for parallel detection of many samples. Labelling of the DNA can be done in both sequence independent ways such as radio active labelling or in sequence dependent ways some of which are nick translation using Dnase I to introduce single stranded nicks, DNA polymerase I to introduce labelled dNTPs, end labelling by using T4 polynucleotide kinase or by PCR using Taq DNA polymerase (Dyson 1991). For microarray experiments, labelling is usually done by PCR sometimes, where labelled dNTPs are incorporated into the amplified DNA along side the normal dNTPs by lowering the elongation temperature. However the number of dNTPs and the level of incorporation cannot be controlled and this results in variation in the signal intensities of the fluorescent dyes (Liu et al. 2001). Another method involves attaching the fluorophore to the 5' end of the target DNA by amplifying with labelled primer. A single primer is labelled, as the captures on the chip are specific to only one strand. Every DNA molecule is attached to a fluorophore by this method. To avoid the bias introduced by the labelling procedure, research is being done to develop label free detection system some of which are reflectometric interference spectroscopy (RIFS) (Jung et al. 2001). quartz crystal microbalance (QCM) (Wang 2000) and bioelectronics detection which is a component of the lab-on-a-chip (Liu et al. 2004).

2.7.2.4. Optimization of the hybridization conditions

Optimization of hybridization conditions is the next crucial factor to obtain strong hybridization signals. Even single nucleotide mismatch discriminations can be detected by adjusting the hybridization and washing conditions (Liu et al. 2001; Relogio et al. 2002; Urakawa et al. 2003). The specificity of the hybridization reaction depends mainly on the experimental parameters such as incubation temperature, ionic strength; chemicals used in the washing buffers and probe characteristics such as length, GC content and sequence similarity to the target. It has been shown that, with the increase in the hybridization temperature, the stringency is increased and the T_m yielding the best performance has to be evaluated experimentally as the common calculating algorithms such as wallace rule (Wallace et al. 1979) and the nearest neighbour method (SantaLucia et al. 1996) to name a few, calculate different T_m for the same capture. (Relogio et al. 2002) states that hybridizations should always be carried at 5-15 °C below the calculated T_m. Addition of denaturing formamide (1%) or 6 M urea sometimes reduces the T_m, whereas higher salt concentration increases the T_m, which means the stringency is increased at lower salt concentrations (Dyson 1991). In general, the stringency during the hybridization should enable specific binding, duplex formation and in the washing, higher stringency should remove all the unspecific probe target duplexes retaining the specific duplexes. Other concerns in the microarray experiments that may effect the signal intensity include, the composition of the spotting buffer used for printing (Li 2002; Wrobel et al. 2003), the ion composition of the hybridization buffer that enhances the specificity and the binding of the target (Ku et al. 2004) The thermodynamics of the capture binding to the target influences the overall hybridization efficiency (Todt 2006; Zelm 2004).

2.7.3. Microarray applications

2.7.3.1. Gene Expression and Clinical diagnostics

After their invention, microarrays have found their first application in the gene expression as there was high demand after the whole genome project to study the gene expression of the complete genome in one experiment (Li 2002). Gene expression profiling has turned into a major application of microarrays (Blohm and Guiseppi-Elie 2001) and they also have a vast application in clinical diagnostics to recognise genetic discrepancies , single nucleotide polymorphisms involved in cancer (Van 't Veer and De Jong 2002), drug discovery (Gmuender 2002), routine clinical analysis (Vetter 2005), identification of Influenza viruses (Sengupta et al. 2003), genotyping of Hepatitis C virus (Drutschmann 2004) and very

recently, microarray multiplex assay has been done to detect Hep B, Hep C and HIV (Hsia et al. 2007). The above mentioned studies are just a few examples of the use of microarrays in diagnostics.

2.7.3.2. Ecological and environmental application

Though microarray technology can be used as a potential Its earliest use is for the parallel detection of environmental micro-organisms (Guschin et al. 1997) and based on that, a few more studies were done for the characterization, identification, analysis and biodiversity monitoring studies of marine and fresh water bacterioplankton (Peplies et al. 2006; Peplies et al. 2004; Pfunder et al. 2004b; Rudi et al. 2000). Broader application of the microarray in the environment is also tested by screening compost for the presence of absence of microbes (Franke-Whittle et al. 2005), detection of waterborne pathogens within mixed populations (Maynard et al. 2005), screening of soil extracts for 16S rRNA (Small et al. 2001), monitoring of algal blooms for harmful algae (Gescher 2007b), quantification and taxonomic classification of environmental rDNA of both prokaryotic and eukaryotic origin (Desantis et al. 2005) and detection of sulphate reducing prokaryotes in the environment (Loy et al. 2002). A recent publication even showed successful quantitative profiling of complex microbial populations in the human body along with the information about their taxonomic composition and diversity of the complete microbial flora of the intestine (Palmer et al. 2006). From the above studies microarray surely seems as a potential tool for environmental studies but its broad application seems to be constricted by the methodological deficiency, laborious protocols and insufficient detection limits.

2.7.3.3. Food quality control assessment

High throughput capability, automation and the ability to allow simultaneous analysis of many samples accurately can make microarray a very useful tool in the food quality control department. DNA methods such as RFLP have been shown to be applied in the food industry to identify the fish species in various processed fish samples (Hold et al. 2001). A few steps using microarrays were taken recently in this direction, where pathogenic vibrio spp. were detected in shellfish, a popular seafood (Panicker et al. 2004), harmful fusarium species infecting the cereal grains were identified (Nicolaisen et al. 2005) and marine pathogens infecting the fish were also detected (Gonzalez et al. 2004; Warsen et al. 2004). Identifying the species and checking the legality of the manufacturer poses a challenge for the application of the microarrays in routine analyses.

2.7.3.4. Species identification

Morphological identification of adult marine species is very tedious and time consuming and more over, if morphologically intermediate forms also occur then the determination of exact species needs very high taxonomic expertise, but taxonomic impediments and a decline in the number of systematics is a worldwide phenomena (Hutchings 1999a) which is gradually decreasing in the current generation. Hence molecular biological methods have become the alternative for this purpose and many of the methods mentioned in the section are already in use. But most of the methods are either restricted, i.e. not flexible enough to be adapted to all kinds of samples or have certain inherent shortcomings. A robust, high throughput and sensitive method such as microarray can be a valuable tool for species identification ,an area which has not been explored much, except for a very few studies, like the identification of microorganisms (Call et al. 2003; Loy and Bodrossy 2006; Wang et al. 2006), plants (Ronning et al. 2005) and animals (Pfunder et al. 2004a). Identification of marine organisms by microarrays has not been explored much except for bacteria (Peplies et al. 2003; Peplies et al. 2004) and recently the "Fish & Chips" project aimed at developing three different kinds of chips for the identification of fishes, phytoplankton and invertebrates (Kochzius 2007). A preliminary fish chip has been developed for the identification of pleuronectiformes by the German Bight (Silkenbeumer 2004) and on a large scale, a microarray to identify 11 different fish species from the European Seas, has been developed (Kochzius et al. 2008). Another study has been reported about the development of an ALEX chip which can identify and monitor the harmful algae Alexandrium (Gescher 2007b) and similarly a phytoplankton chip has also been reported for the monitoring of phytoplankton in the European Seas (Gescher 2007a; Metfies 2004; Metfies et al. 2005). Marine invertebrate larvae were identified by hybridization technique (Goffredi et al. 2006), but no studies have been so far reported to be using microarrays to identify marine invertebrates at species level and this study will be one of the most preliminary attempts at developing such a chip.

2.8. Importance of marine invertebrates

95% of the Earth's animal species are invertebrates, animals lacking a notochord. Different species live on land, freshwater and oceans, the latter comprise the highest percentage, a number of them occurring from the sea surface to the sea floor and into the substrate. Invertebrates from the oceans are commonly called marine invertebrates and they make up a huge portion of the life in the sea. They include a very long 18 m squid and also a very tiny, less than 0.25 mm gall mites, and they have an amazing diversity of form. Invertebrates

include varied range of organisms, belonging to eight different phyla and represent a vast majority of marine biodiversity, many of them playing an important role in the functions and processes of the marine ecosystem. Some invertebrates are commercially important for human consumption and aquarium trade, such as oysters, prawns, scallops, abalone, lobsters, squid etc., while some are a major source of tourist attraction like corals and the products from some of them, like pearls are economically important. From the many different kinds of invertebrates occurring crustaceans, molluscs and polychaetes are the major groups of macro invertebrates in terms of diversity, size and numeric abundance.

Polychaetes are the multi segmented worms commonly known as bristle worms, sand worms and tube worms. Approximately 13,000 species of poychaetes belonging to 80 families have been reported so far and these exhibit a great range of morphological and functional diversity, as well as reproductive strategies (Rouse 2001). Polychaetes are economically important, as adult and larvae forms are food for many commercially important fish. The major significance of polychaetes is seen in bio-monitoring marine environmental quality, they being monitors for toxic materials and pollution. Polychaetes show sensitivity to anthropogenic compounds by expressing changes in the growth, reproduction and mortality; hence they are significant as bio-indicator organisms. Being abundant, having a short life cycle and covering a wide range of sizes, polychaetes are suitable for assessing the toxicity of sediments. Studies of the presence or absence of polychaetes in the sediment itself acts as an indicator for the health and condition of the benthic environment.

Molluscs are soft bodied without any internal skeleton, with no standard shape and most of them have a hard protective shell. They are the second largest phylum of animals after arthropods and economically one of the most important groups of invertebrates. About 1,20,000 species recorded to date, they show a great structural and ecological variability. Molluscs are both commercially and economically valued as sea food as well as for their decorative shells. Their role as pests and disease carriers along with the economic importance of commercially exploited species has been studied very extensively (Ponder 2008)

Crustaceans are the largest and extremely diverse group of arthropods found in all depths of marine environment. They vary a great deal in shape and form, making it difficult to define easily recognised traits common to all of them and also some have highly modified bodies. Crustaceans are economically very significant as many species are fished commercially for sea food and they are also ecologically important in a variety of roles as grazers, scavengers, predators and as prey for larger animals such as fish.

Though having a high economical, commercial and ecological importance, this group of organisms is still not very well known due to the lack of taxonomic knowledge, hindering the advances in biological and ecological research. Many groups of invertebrates are still very poorly known and are difficult to work with and cannot be identified even at a morpho species level.

2.9. Aim of the Study

DNA based molecular methods are being used extensively these days in ecological studies, especially for species, difficult to identify morphologically and also for the monitoring of various biological habitats. PCR based techniques though popular are cumbersome, time consuming and the degree of multiplexing is limited to about a dozen. Among the recent molecular biological techniques, microarray is the most advanced and robust high-throughput tool which allows principally the analysis of tens of thousands of samples very fast, accurately and simultaneously. DNA microarrays have been used so far extensively in gene expression analysis and their application as a tool in ecological studies has recently been started to understand the characterization of complex biological communities.

Invertebrates form a major part of marine biota and are an important group of organisms that serve as predators for demersal fish, are main bio-indicator organisms for ecosystem monitoring and also help in the marine environmental assessment. The biodiversity study of invertebrates is often very difficult because of their morphological complexity and also because it requires special knowledge and long years of taxonomical expertise.

As a part of the EU-FP6 project “Fish & Chips”, the aim of this work is to demonstrate that an “Invertebrate microarray” can be developed in terms of its biological and technical dimension. At least a small prototype would be necessary to test its usability for identifying selected groups of ecologically important marine invertebrate species occurring in the European Seas as a prerequisite for routine ecological studies and regular monitoring of the marine environment and also for the quality control in sea food marketing.

Efficient DNA extraction is an important step for any type of DNA analysis. Since many invertebrate species are known to resist different standard DNA extraction protocols and because, for marine research, a wide range of organisms has to be worked on, finding a method for obtaining reasonable amounts of pure genomic DNA from animals belonging to different phyla is an additional aim of this study.

The microarray development involves crucial steps, such as the selection of specific oligonucleotides suitable as species specific microarray capture molecules, preparation of

labelled target DNA from the organisms to be analyzed, optimization of the hybridization conditions and efficient analysis of the signal data. To accomplish these main tasks of this work, the design of the oligonucleotides will focus on the two mitochondrial gene sequences (16S and COI) of at least a dozen invertebrate species from three phyla. The probes will initially be selected and tested for their putative specificity *in silico* by comparing with non-target species used in the work and also by a wide range of closely related species as far as the gene sequences are available in the data bases. Based on this work *in silico,* the microarray will be produced and evaluated in the laboratory by carrying out hybridizations with different target and non-target species to check their specificity. Experimental optimization of various hybridization parameters will be important for trying to overcome the methodological limitations still existing for this technique. After the initial examination of the microarray with the selected specific target organisms, the real application of the microarray for analyzing as an example, the gut content of demersal predator fish was evaluated.

3. Materials and Methods

3.1. Sample collection

The invertebrate sample species used in this work were chosen according to their economical importance as prey to fish predators and as bio-indicator organisms. Examples are *N. hombergii*, important bio-indicators and *L. depurator*, common prey to fish. Around 411 individuals belonging to a total of 27 species of crustaceans, molluscs and polychaetes were collected. As most of the polychaetes chosen for the work are tube dwelling or living in burrows in the sea bottom, they were carefully selected from the sediment and collected with van Veen grabs. Most of the polychaetes were collected from six different locations in the German Bight region of the North Sea during a one week cruise (25.09.04 – 01.10.04) with the research vessel "Heincke" (Alfred Wegener Institut für Polar- und Meeresforschung). Some additional polychaetes collected from the lagoons in the Mediterranean Sea were donated by Dr. Christos Arvanitidis of the HCMR Institute (Crete, Greece). Crustacean and molluscan species were collected during different expeditions in the Adriatic Sea by collaborators Dr. Monica Landi and Ms. Alessia Cariani from the CIRSA, University of Bologna (Ravenna, Italy). Out of the 27 species collected only 15 were used for the development of the chip according to the project requirement. List and species data are presented in Table 1.

The animals collected were preserved in absolute ethanol and stored at 4 °C. The crustacean and the molluscan samples were morphologically identified to species level by the collaborators at UNIBO and the collectors themselves. All the polychaete species were taxonomically identified by polychaete taxonomist Dr. Christos Arvanitidis from HCMR Crete during a three week training course in polychaete identification the author of this thesis experienced in that institute.

Table 1: List of invertebrates collected during the work. *: Species added on the microarray.

Phyllum	Family	Species Name	Species Code	Collection Site	No. of individuals collected
Arthropoda	Alpheidae	**Alpheus glaber*	Ag	Adriatic Sea	10
	Crangonidae	**Crangon crangon*	Cc	North Sea	64
	Goneplacidae	**Goneplax rhomboides*	Gr	Adriatic Sea	10
	Portunidae	**Liocarcinus depurator*	Ld	Adriatic Sea	10
	Portunidae	**Liocarcinus vernalis*	Lv	Adriatic Sea	10
	Euphausiidae	**Meganyctiphanes norvegica*	Mn	Adriatic Sea	10
	Grapsidae	**Pachygrapsus marmoratus*	Pm	Adriatic Sea	10
Mollusca	Octopodidae	**Eledone cirrhosa*	Ec	Adriatic Sea	10
	Corbulidae	**Lentidium mediterraneum*	Lm	Adriatic Sea	10
	Veneridae	**Chamelea gallina*	Cg	Adriatic Sea	10
	Octopodidae	**Illex coindetii*	Ic	Adriatic Sea	10
Annelida	Nereididae	**Hediste diversicolor*	Hd	Mediterranean Sea	22
	Nephtyidae	**Nephtys hombergii*	Nh	Mediterranean Sea & North Sea	56
	Pectinariidae	**Pectinaria koreni*	Pk	North Sea	13
	Sigalionidae	**Sigalion mathildae*	Sm	North Sea	12
	Capitellidae	*Notomastus latericus*	Nl	North Sea	4
	Cibbatulidae	*Chaetozone setosa*	Cs	North Sea	16
	Glyceridae	*Gonieda maculata*	Gm	North Sea	3
	Magelonidae	*Magelona johnstoni*	Mj	North Sea	17
	Orbiniidae	*Scoloplos armiger*	Sa	North Sea	10
	Owenidae	*Owenia fusiformis*	Of	North Sea	21
	Poecillochaetidae	*Poecillochaetus serpens*	Ps	North Sea	5
	Phyllodocidae	*Phyllodoce rosea*	Pr	North Sea	1
	Phyllodocidae	*Phyllodoce lineata*	Pl	North Sea	2
	Scalibregmidae	*Scalibregma inflatum*	Si	North Sea	1
	Spionidae	*Spiophanes bombyx*	Sb	North Sea	58
	Terebellidae	*Lanice conchilega*	Lc	North Sea	6

3.2. Instruments

- Automatic capillary electrophoresis machine (Applied Biosystems DNA-Sequencer prism310)
- Centrifuge (Eppendorf 5403)
- Gel documentation system (CAMAG Reprostar)
- Gel electrophoresis chamber (Biometra Agagel G45)
- Hybridization oven (MWG-Biotech Mini 10; Heraeus B12; Heraeus B5028)
- Laminar flow cabinet (Clean Air CA/REV 6)
- Microarray scanner (Axon 4000B)
- PCR Cycler (Biometra Cycler Trio; Eppendorf Mastercycler Gradient S)
- Spectrophotometer (PeqLab NanoDrop ND-1000)
- Spotter (robot for touch-less spotting based on Top Spot technology, proprietary development of CAG, University Bremen and PicoRapid GmbH)
- Vacuum centrifuge (Savant Speed vac SC110)

3.3. Bioinformatics Software

- BioEdit sequence alignment editor, Version 7.0.0
- Chromas Freeware, Version 1.45
- Clustal X, Version 1.81
- Endnote, XI.0.1
- Genedoc, Version 2.6.002
- Genepix pro, Version 4.1
- MEGA, Version 3.1
- Microsoft Excel, 2003
- Microsoft Visio, 2007
- SeqMAN (DNA star) II, Version 4.05
- Winprobe (Pozhitkov et al., 2002)

3.4. DNA extraction

DNA extraction from crustaceans is considered as not so easy task as it is difficult to digest the exoskeleton in small individuals, and the tough tissue in larger animals. Similarly various methods were mentioned in literature to extract genomic DNA from polychaetes because of the vast diversity of the animals. Therefore various methods were performed which led to an extra investigation of various classical and commercial extraction methods in order to find a standard with which DNA can be extracted successfully from all samples.

3.4.1. Classical methods

Initially classical methods such as CHELEX (Walsh et al. 1991) and phenol chloroform methods (Chomczynski and Sacchi 1987) have been tested for DNA extraction from small tissues on five different kinds of polychaetes since these animals turned out to behave very differently in terms of DNA extraction.

3.4.1.1. CHELEX method protocol

Chemicals :

12 % CHELEX (Bio-Rad)

Dithiothreitol (DTT) (Serva)

Proteinase K (Fluka)

Protocol:

1. Add 150 µl of 5 % CHELEX, 100 µl of 100 mM DTT and 20 µl of proteinase K to 20 -25 mg of the animal tissue.
2. Incubate in a thermomixer at 55 °C, 500 rpm for 4 hrs.
3. Centrifuge at 13,000 x g for 3 - 5 min.
4. Collect the supernatant containing the DNA and incubate at 95 °C for 10 min to inactivate the proteinase K.
5. Take 2 µl of this extract for DNA amplification by PCR.

3.4.1.2. Phenol Chloroform method

Chemicals and buffers:

Phenol chloroform (Sigma)
5 M NH_4-Ac (Riedel-de-Haen)
Isopropanol (Roth)
70% Ethanol (Riedel-de-Haen)
Proteinase K (Fluka)

HOM Buffer for cell lysis:

2 M Tris HCl (Sigma, pH 8.0)
0.5 M EDTA (Roth, pH 8.0)
3 M NaCl (Fluka)
10 % SDS (Sigma)
50 µM DTT (Serva)

TE Buffer for elution:

10 mM Tris HCl
0.1 mM EDTA
pH 8.0

Protocol:

1. Add 500 µl of HOM buffer and 20 µl of proteinase K to 10 - 20 mg of tissue.
2. Incubate in the thermoblock at 54 °C for at least one hour.
3. Add 500 µl of phenol chloroform.
4. Centrifuge at 9000 rpm for 15 min.
5. Transfer the top supernatant in to a new cup carefully.
6. Add 2 volumes of isopropanol (absolute) and 0.1 volumes of NH_4Ac (5 M) and mix thoroughly.

7. Precipitate overnight at -20 °C.
8. Centrifuge samples at full speed (13,200 rpm) for 30 min at 4 °C.
9. Decant and add 600 µl of cold EtOH (70 %) and centrifuge for 15 min at full speed, discard supernatant.
10. Repeat the washing step twice and dry the samples in a speed vac.
11. Re-dissolve the precipitated DNA in 30 µl ddH_2O or low TE buffer.

3.4.2. Commercial kits

Apart from the routine classical methods a wide range of commercial kits, available for DNA extraction from various tissues were tested with the polychaete samples to find the best suited ones. The kits were used according to manufacturer's protocol (unless specified) and the evaluation of about 21 different DNA extraction kits used in this work is mentioned in Table 2.

Table 2: Comparative list of the DNA extraction kits used during the study.

Extraction method	Website	Principle	Processing time	Additional notes &protocol modifications
QIAGEN DNeasy tissue kit	http://www1.qiagen.com/	Proteinase K digestion & silica gel membrane technology	~ 3 hours	DNA is eluted into ddH2O instead of elution buffer in the kit
Mackerey Nagel Nucleospin tissue kit	https://www.macherey-nagel.com	SDS & Proteinase K cell lysis – silica membrane column based method	~ 3 hours	
Cartagen DNA Extraction kit	http://www.cartagen.com	Proteinase K digestion with column based DNA isolation	~ 2 hours	Specific for DNA extraction from Arthropods and Shrimps
SIGMA GenElute mammalian genomic DNA kit	http://www.sigmaaldrich.com	Enzymatic digestion and silicon membrane elution	2 -4 hours	
Peqlab E.Z.N.A tiisue mini kit	http://www.peqlab.de	Silica membrane column based method	2 – 4 hours	
Nexttec enomic DNA isolation kit(Bio Sell)	http://www.bio-sell.de/	Cell lysis and DNA purification with nexttec clean columns	2 hrs to overnight cell lysis and further elution 15 min	
ZR genomic DNA kit(Zymo research)	http://zymoresearch.com/	Cell lysis & Zymo spin column based DNA isolation	overnight lysis & 10min downstream procesing	
Direct PCR Extraction methods				
VIAGEN tissue kit (Peqlab)	http://www.peqlab.de/	Direct PCR lysis	45 min	
Gene Releaser kit (Eurogentec)	http://www.eurogentec.com/	Direct lysis, extraction and amplification	3 – 4 hrs	
Bio Sell tissue mini kit	http://www.bio-sell.de/	Cell lysis and DNA purification in the same tube	2 hrs to overnight cell lysis and further elution 15 min	
Whatman FTA paper	http://www.whatman.com/	DNA bound and purified by matrix on the FTA paper	15 min to1 hour-15 min downstream processing	Centrifugation is done during washing
Picopure DNA extraction kit (Arcturus)	http://www.arctur.com/	Proteinase K lysis and reconstitution buffer elution	3 – 4 hrs cell lysis	
Microlysis (Microzone limited)	http://microzone.co.uk/	Effective lysis of cells by rapid heating & cooling of a thermo cycler	20 min	
Whatman FTA Elute	http://www.whatman.com/	DNA bound and purified by matrix on the FTA paper is eluted into water for further processing	15 min to1 hour-15 min downstream processing	
Other extraction methods				
Easy DNA kit (Invitrogen)	https://www.invitrogen.com	Isolation of DNA by Phenol Chloroform and ethanol precipitation and elution	3 hours to 2 days	
AGOWA DNA extraction kit	http://www.agowa.de	Proteinase K lysis and magnetic separation of DNA.	Overnight lysis and 1hour of downstream procesing	Lysis extended to overnight instead of 10min
Pure gene tissue kit (Gentra systems)	http://www.gentra.com	Cell lysis – Protein precipitation, isopropanol and ethanol DNA precipitation – DNA hydration	3 hours – overnight and DNA hydration 1hour to overnight	
Master pure DNA purification kit (Epicentre Biotechnologies)	http://epibio.com/	Proteinase K lysis and isolation by rapid desalting, precipitation and Isopropanol	45min and 20min downstream processing	
DNA extraction kit (Stratagene)	http://stratagene.com	Enzymatic cell digestion, isolation by salting out, ethanol precipitation	2-24 hours of digestion and 45 min of downstream processing	
Genomic DNA purification kit (Fermentas)	http://fermentas.de/	Chemical cell lysis, detergent precipitation, salting out and ethanol precipitation	overnight cell lysis and 25 min of downstream processing	
DNAMite tissue kit (Microzone)	http://microzone.co.uk/	Proteinase K lysis and chemical DNA precipitation and rehydration	3-24 hours lysis and 50 min of downstream processing	

3.4.3. Whatman FTA cards

Chemicals and Buffers:

FTA purification reagent (Whatman)

1X PBS buffer for homogenate:

0.13 M NaCl

2.6 mM KCl (Fluka)

10 mM $Na_2HPO_4.7H_2O$ (Fluka)

1.76 mM KH_2PO_4 (Fluka)

pH 7.4

TE Buffer for washing:

10 mM Tris HCl

0.1 mM EDTA

pH 8.0

Unlike the tested commercial DNA extraction kits a very simple and innovative extraction method is provided by Whatman in the form of FTA cards where fresh samples can be prepared during sampling in the field on a special paper without the use of any ethanol. Ethanol preserved tissues are homogenised in 1X PBS buffer before adding to the paper. Most of the cell content except nucleic acids and many proteins are destroyed by the chemical components of the paper retaining preferentially the DNA for long time in the dried state. A small disc of the paper containing the absorbed DNA of the sample can be used either directly in the PCR reaction or the DNA can be eluted to be used for PCR. Different use of the two types of FTA kits available for DNA extraction from tissues, FTA classic cards and FTA Elute cards are compared in Table 3.

Table 3: Comparison of the two FTA DNA extraction methods.

	FTA Classic	FTA Elute
Sample application	~25 µl of tissue homogenate added on to each circle	~25 µl of tissue homogenate added on to each circle
Drying	15 min to 1 hour	15min to 1 hour
Extraction from	1.5 mm disc (PCR tube)	3.0 mm disc
Washing 1	FTA purification reagent 3X 5min	ddH2O 3X5 sec
Washing 2	TE buffer 2X5min	
Drying	65°C, 20min	
Elution		50 µl ddH2O, 95°C, 30min
PCR	dried disc directly used in PCR	3-5 µl of eluate for 50µl of PCR reaction

3.5. Marker gene amplification

Chemicals and Buffers:

10X PCR Buffer (Molzym)

100 mM $MgCl_2$ (Molzym)

20 mM dNTPs (Molzym)

5 U/µl Taq polymerase (Molzym)

BSA (Sigma)

DNA extracted from the invertebrate species mainly by FTA Elute and Qiagen DNeasy kit was used as template in PCR for marker gene amplification. Fragments between 500 – 600 bp of the 16S rRNA gene and 708 bp in case of the COI gene respectively are amplified in the Biometra Trio-thermoblock and the Eppendorf Mastercycler. PCRs were prepared under laminar flow cabinet according to the profiles mentioned in Table 4-Table 9.

3.5.1. 16S PCR

Primers: (Palumbi 1991)

16S ar L: 5′ CGC CTG TTT ATC AAA AAC AT 3′

16S br H: 5′ CCG GTC TGA ACT CAG ATC ACG T 3′

Table 4: PCR reaction profile for 16S amplification.

Reagent	Concentration of stock	Final concentration in PCR	µl/reaction (Total 50 µl)
10X PCR buffer	10X	1X	5
MgCl2	100 mM	4 mM	2
dNTPs	20 mM	200 µM	1
16S ar L	100 pmol	0.4 µM	2
16S br H	100 pmol	0.4 µM	2
BSA	20 mg/ml	0.8 mg/ml	2
Taq polymerase	5 U/µl	1U	0.2
DNA			2
ddH2O			33.8

Table 5: Temperature profile for 16S amplification.

Initial denaturation	94 °C	2 min	
Denaturation	94 °C	1 min	
Annealing	50 °C	1 min	39 cycles
Elongation	72 °C	1 min	
Final elongation	72 °C	5 min	
Cooling down	4 °C	pause	

3.5.2. COI PCR

Primers: (Folmer et al. 1994)

LCO 1490: 5′ GGT CAA CAA ATC ATA AAG ATA TTG G 3′

HCO 2198: 5′ TAA ACT TCA GGG TGA CCA AAA AAT CA 3′

Table 6: PCR reaction profile for COI amplification.

Reagent	Concentration of stock	Final concentration in PCR	µl/reaction (Total 50 µl)
10X PCR buffer	10X	1X	5
MgCl2	100 mM	2.5 mM	1.25
dNTPs	20 mM	250 µM	1.25
LCO 1490	100 pmol	0.4 µM	2
HCO	100 pmol	0.4 µM	2
BSA	20 mg/ml	0.8 mg/ml	2
Taq polymerase	5 U/µl	2U	0.4
DNA			2
ddH2O			34.1

Table 7: Temperature profile for COI amplification.

Initial denaturation	95 °C	2 min	
Denaturation	95 °C	1 min	
Annealing	40 °C	1 min	39 cycles
Elongation	72 °C	1.5 min	
Final elongation	72 °C	7 min	
Cooling down	4 °C	pause	

3.5.3. "Simplified multiplex" PCR

In order to reduce the number of PCR reactions that has to be performed to carry on all the single target hybridizations a multiplex PCR was optimized combining the 16S and COI protocols considering the highest amount of the chemicals adding two sets of primers in the same PCR cup. PCR is conducted at an optimized annealing temperature of 45 °C.

Table 8: "Simplified" multiplex reaction profile.

Reagent	Concentration of stock	Final concentration in PCR	µl/reaction (Total 100 µl)
10X PCR buffer	10X	1X	10
MgCl2	100 mM	4 mM	4
dNTPs	20 mM	200 µM	1
16S ar L	100 pmol	0.3 µM	3
16S br H	100 pmol	0.3 µM	3
LCO 1490	100 pmol	0.3 µM	3
HCO	100 pmol	0.3 µM	3
BSA	20 mg/ml	0.8 mg/ml	4
Taq polymerase	5 U/µl	2U	0.4
DNA			2
ddH2O			66.6

Table 9: Temperature profile of "simplified" multiplex PCR.

Initial denaturation	95 °C	2 min	
Denaturation	95 °C	1 min	
Annealing	45 °C	1 min	39 cycles
Elongation	72 °C	1.5 min	
Final elongation	72 °C	7 min	
Cooling down	4 °C	pause	

3.6. Standard DNA-analytical methods

3.6.1. Agarose gel electrophoresis

Chemicals and buffers:

Loading dye (Fermentas)

Generuler 1KB DNA ladder (Fermentas)

Ethidium bromide (Serva)

1X TBE Buffer:

0.89 M Tris

0.89 M Boric acid (Roth)

0.02 M Na_2EDTA (Roth)

pH 7.2

Molecular weight of the PCR amplicons is determined by standard 1 % agarose gel by electrophoresis and ethidium bromide staining (Franklin and Locker 1981). 1 µl of the PCR product is run in the gel with 5 µl of 6X loading dye at an voltage of 110 for about 45 min in a

gel chamber containing about 100 ml of 1X TBE buffer. Visualized gel pictures are recorded with a canon controlled gel documentation system.

3.6.2. Purification of the PCR products and sequencing

To separate the PCR reagents that can interfere in the subsequent reactions from the actual PCR product, purification of the amplicons is done with the Qiagen PCR purification kit following the manufacturer's protocol. After the last washing step pure DNA is eluted in to 30 µl of ddH_2O.

3.6.2.1. DNA-sequencing

Chemicals and buffers:

BigDye terminator v3.1 cycle sequencing kit (Applied Biosystems)
Sodium acetate (Merck)
70 % Ethanol (Riedel-de-Haen)

In this work the di-deoxy termination method originally described by Sanger et al (Sanger et al. 1977) was used and all the sequencing reactions are performed in the automatic capillary electrophoretic machine prism310 . The BigDye terminator sequencing reaction was carried out according to the manufacturer's protocol Fig. 10 (Rosenblum et al. 1997). Both L and H strands were sequenced independently using the primers mentioned in the PCR section. A total volume of 10 µl is prepared with 6 µl of pure PCR product as template, 2 µl of BigDye terminator solution, 1 µl of ddH_2O and 1 µM of the respective primer.

Table 10: Temperature profile of the sequencing PCR reaction.

Denaturation	95 °C	15 sec	
Annealing	40 °C, 50 °C	15 sec	26 cycles
Elongation	60 °C	4 min	

The DNA from the sequencing reaction is precipitated and washed with 1/10 of the PCR product volume of 3 M sodium acetate (pH 5.2) and twice the PCR product volume of absolute ethanol. Pelleted DNA is washed with 70 % ethanol and centrifuged again. The DNA pellet is dried under a fume hood before resuspending in 30 µl of ddH2O for sequencing. The DNA fragments are separated according to their size by a POP4 polymer (Applied Biosystems) with an injection time of 15 sec. Most of the 16S and COI sequences are obtained by using this method but the COI sequences from 10 samples collected from the

Adriatic Sea were obtained using the commercial sequencing service of the Macrogen company in Korea. A total number of 159 sequences were obtained from the 15 species and 128 of them were used for the oligonucleotide design. See Table 14.

3.6.2.2. Sequence editing

The sequence chromatograms obtained were carefully checked and edited for gaps and mismatches. Three different sequence editing programs are used during this work namely Chromas (McCarthy 1998), Seqman II (DNA star Inc. 2000) and MEGA 3.1 (Kumar et al. 2004). The forward and the reverse strands are edited separately and then aligned with Genedoc software Version 2.6.002 (Nicholas 1997) to check ambiguities. Sequences of all individuals belonging to one species are then aligned using GeneDoc or Bioedit to check for single nucleotide polymorphisms. After editing, the identity of the sequence with the similar sequences already submitted to the public databank is checked by a nucleotide BLAST search against the NCBI Genebank (Altschul et al. 1990). Sequences of the target species available in the databank are downloaded and aligned with those sequences obtained from the work. A multiple alignment of all the sequences is performed in BioEdit Version 7.0.0 (Hall 1999) to cut them to an equal length and to submit them to the capture design program. Number of sequences obtained from the work and those downloaded from the databank for each species are given in Table 13.

3.7. Microarray

3.7.1. Design of the capture oligonucleotides

A set of sequences belonging to closely related species of the target organisms are downloaded to be used as non-targets during oligonucleotide design. The oligonucleotides used in this work are designed by a Chip designing internet service (CDIS) offered by the Miconet service of the Techno mathematics department of the University of Bremen. Principle of their algorithms is a sequence comparison by alignments and looking for short stretches of oligonucleotides which are specific for a particular species. *In silico* selection is done by calculating of the thermodynamic properties of the oligonucleotides and the potential cross hybridizations with other sequences. Specificity tests are carried out by comparing the oligonucleotides with other target sequences and vast number of non-target sequences downloaded from the databank. The final set of the oligonucleotides generated consists of the sequence of the captures along with their thermodynamic criteria such as length, GC content, hybridization temperature and self dimerization probability. The mismatch criterion of every

single capture is also provided both from the comparison with the target and the non-target sequences. Two sets of oligonucleotides were obtained and used during this work for two different prototypes of the invertebrate microarray.

3.7.1.1. 16S Invertebrate chip

For the design of captures a total of 70 sequences of the 15 targets obtained from this work their sequences already present in the NCBI databank were used (Table 14). Additional 1489 sequences from the NCBI databank representing closely related species from the European seas and belonging to the three groups of invertebrates used in this work were downloaded and used for the capture probe design as non-target background sequences.

239 captures were generated in the initial calculation from the 16S gene of 10 species and for the rest 5 captures were designed from both 16S and COI genes. The 43 captures (Table 15) selected gave an average of 3 probes per species. The criteria of probe design are the length in the range of 23 to 50 bases, GC content 35 - 55 %, specific binding with the target sequences, the hybridization temperature to be at a NaCl concentration of 1 M and according to the above parameters and is in the range of 70 - 80 °C. The probes were checked for any secondary structure formation and self complementarities and no cross reaction with any of the other targets and those non-targets included. A capture was considered for experimental testing if it recognizes the target species with 0 mismatches and shows with other target and non target species at least 3 or more mismatches in the centre of the alignment.

3.7.1.2. COI Chip

COI sequencing of the samples was done with universal COI primers (Folmer et al. 1994). The sequences were edited with DNA star and aligned with their targets with Bioedit. From a total of 58 target sequences and 3170 non target sequences (Table 14) 84 capture probes were designed in the similar way as for the 16S chip but with much narrower criteria of length ranging from 23 - 35 bp and a GC content of 40 - 50 %. Finally 30 captures were used to spot on a separate COI chip (Table 16).

3.7.1.3. Invertebrate chip

To test whether the chip will be more beneficial if both 16S and COI are present those captures which were giving good specific signals in the previous experiments were chosen and an Invertebrate chip was spotted with 64 captures selecting 35 captures from the first set and 29 captures from the second. The captures that are not included in the invertebrate chip are marked with a star * in Table 15 and Table 16.

3.7.2. Microarray spotting

Buffers:

Spotting Buffer:

150 mM Na-phosphate (Fluka)

1 M Betaine (Sigma)

Washing buffer (spotted slides):

1X SSC (Invitrogen)

0.05 % SDS (Sigma)

The oligonucleotides mentioned in the tables were chemically synthesized by Thermo electron with a 5'-amino modification and HPLC purified captures are obtained in concentration of 100 µM. The amino group is attached to the 5'-hydroxyl group of the oligonucleotide with a chain of six carbon atoms (C6-linker). The amino group enables the specific covalent binding of the oligonucleotide to the glass surface which is treated first with 3-Aminopropyltrimethoxysilane and then with 1, 4-Phenylenediisothiocyanate (Fig. 4). These coated slides were purchased as Genorama SAL-1 slides from Asper Biotech, Estonia.
Oligonucleotides were taken up with spotting buffer in a concentration of 20 µM and dispensed contactless on the Genorama SAL-1 slides (Asper Biotech) using the Top Spot technology based CAG robot at 65% relative humidity at 23 °C creating spots of 200 nm in diameter. The spotted slides are incubated overnight in a 75% relative humidity chamber (containing saturated sodium chloride solution) to saturate the binding of the captures to the glass slide. Later the slides are washed in 1X SSC + 0.05% SDS at room temperature for 3 min with gentle agitation followed by three washes with ddH_2O, 1 min each. The slides are then dried by centrifugation (3 min, 2000 rpm), packed under N_2 conditions and stored at 4 °C until further use.

3.7.3. Array configuration

Each array used in this work is a set of 96 spots each arranged in 4 rows and 24 columns on a glass slide. The amino oligonucleotide 5'-TAT GAT CCA TCG ACT ACA ACG-3' (called S1 21) was spotted as positive control and spotting buffer as negative control along with the target specific captures. The arrangement of the captures in the array is carefully planned so that the capture replicate are equally distributed all over the plane of the glass slides. The positive controls are spotted either on the edges of the array or at those positions which makes tracing the spots easy during analysis. For the invertebrate 16S chip a total of 43 captures (31

16S and 12 COI) were spotted twice on the array with positive (S1 21) and negative controls (8.4.1).

In case of the COI chip a total of 30 COI probes were spotted in triplicates on the array with positive and negative controls respectively (8.4.2).

The Invertebrate-chip is composed of 64 captures (34 COI and 30 16S) spotted once on the array along with the positive and negative controls (8.4.3). Each array is spotted 15 times on the slide in three groups each containing 5 arrays, so that three different hybridization experiments can be carried out on one slide as shown in the Fig. 5.

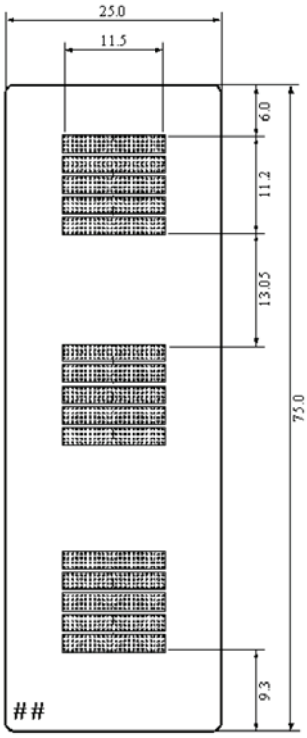


Fig. 5: Layout of the microarray. Dimensions of the slide are given in mm.

3.7.4. DNA-target preparation

The 16S and COI PCR products carrying the fluorescent dye (Cy5) as end label served as DNA targets to be hybridized to the captures on the chips. For preparing these targets DNA from at least three individuals of each species are used for amplification. This target PCR was performed in a similar way as the PCR mentioned above but 100 µl wise in a total volume of 300 - 600 µl to assure good yield of DNA for hybridization. The reaction mixture was similar to the previous PCR except the reverse primers 16S ar-b and H2098 labelled with Cy5 at the 5'-ends are used. These primers are adding the label to the amplified fragment during PCR and produces Cy5 end labelled target PCR products. After testing these PCR products on an agarose gel the target preparation was cleaned with the Qiagen PCR purification kit following

the manufacturer's protocol. 30 - 60 µl of purified PCR product is eluted in the final step. The concentration of this DNA was measured in the NanoDrop photometer.

In order to reduce the number of PCR reactions the multiple target preparations were done in a "simplified multiplex PCR" as described in 3.5.3, which was tested to work with all samples except *C. crangon* for which both 16S and COI PCR were performed separately. Abundant amounts of dNTPs and Mg^{2+} were added along with equal amounts of the primers to make sure that both the genes are amplified to the same extent. Finally, the PCR products are mixed in required amounts.

3.8. Hybridization

Microarray hybridization is carried on with the help of Geneframes (65 µl, 1.5 x 1.6 cm, ABGene). Geneframes are plastic materials forming a square shaped hollow space of definite height surrounded by a double sided sticky gel bands for fixing them in a waterproofed way to the glass surface of the array. After tightly attaching them to the slide around the array and closing them with a cover slip provided in the kit it forms a separated reaction chamber on the slide for hybridization.

3.8.1. Hybridization protocol

Chemicals and buffers:

2X Hybridization buffer:

40 mM Tris-HCl, pH 7.3
2 M NaCl
10 mM EDTA, pH 7.3
0.1 %(w/v) Tween20

Washing buffer 1 (TETBS):

20 mM Tris-HCl, pH 7.3
1 M NaCl
5 mM EDTA, pH 7.3
0.05 % (w/v) Tween20

Washing buffer 2 (TBS):

20 mM Tris-HCl, pH 7.3
1 M NaCl

After attaching the Geneframes to the slides (three per one slide) hybridizations were carried out with 65 µl hybridization solution containing 10 nM target DNA and 1 nM positive control and hybridization buffer each. The solution was prepared by denaturing it at 95 °C for 10 min

and shock cooling by transferring it immediately on to ice until hybridization. Denatured hybridization solution is added carefully and quickly to the array and the gene frames are sealed tightly with cover slips to avoid air bubbles and also evaporation of the hybridization solution. Thus prepared slides are put in a closed humid chamber to avoid the evaporation of the DNA at high hybridization temperatures before transferring them to the incubator. Hybridization is carried on in a hybridization oven at 60 °C for 3 hrs with the slides horizontally placed without any agitation. After hybridization gene frames are removed from the slides and the slides are washed with TETBS buffer 2 times for 60 sec each at room temperature with mild agitation to remove the unbound target DNA molecules. A second wash is done with TBS buffer 2 times for 5 min each at room temperature and mild agitation. The slides are dried by centrifugation at 2000 rpm for 3 min and visualized by scanning in an Axon 4000B Scanner at photo multiplier tube (PMT) values of 450 and 600 for Cy3 as well as for Cy5 labels respectively.

3.8.2. Signal detection and data analysis

The fluorescence scanner Axon 4000B which simultaneously acquires data at two different wavelengths 528 nM (Cy3) and 635 nM (Cy5) measures the positive controls which are labelled with Cy3 dye as well as Cy5-labelled target signals. As oligonucleotides tend to hybridize faster to the oligonucleotide captures than to the long PCR products, the positive control signal is measured at a lower PMT value. The fluorescence is measured as absolute signal intensity which indirectly denotes the hybridization efficiency of the captures to bind to the target DNA. The analysis of the fluorescence signals on the array is done with the help of the Genepix pro, Version 4.1 software. The software allows automatically detecting and assigning features to the spots, so that spots can be marked and individually checked manually for the aligned features and edited if needed. This way artefact originated from salt depositions or air bubbles can be marked for separate evaluation. The means and medians of the pixels in the spots are generated according to the array configuration by the software along with the local background. Local background is the mean of the pixels measured in a ring around the spot two times its diameter. The mean of the means of the individual replicates is later calculated and depicted in Excel diagrams as absolute intensities. The signal intensities of the negative control signals are considered as background and the signals are graphically represented only after subtracting two times the background value from the signal values.

4. Results

4.1. Comparison of DNA extraction methods

The classical DNA extraction methods such as phenol chloroform and chelex were not able to produce good quality DNA from few of the crustaceans and polychaetes. Though there was good DNA band observed on the gel from the extracts obtained by the phenol chloroform method, there were no successful PCR products produced.

Initial experiment with the commercial method Qiagen DNeasy kit produced good amount of the DNA from majority of the samples except the tough tissue of crustaceans. (Fig. 6)

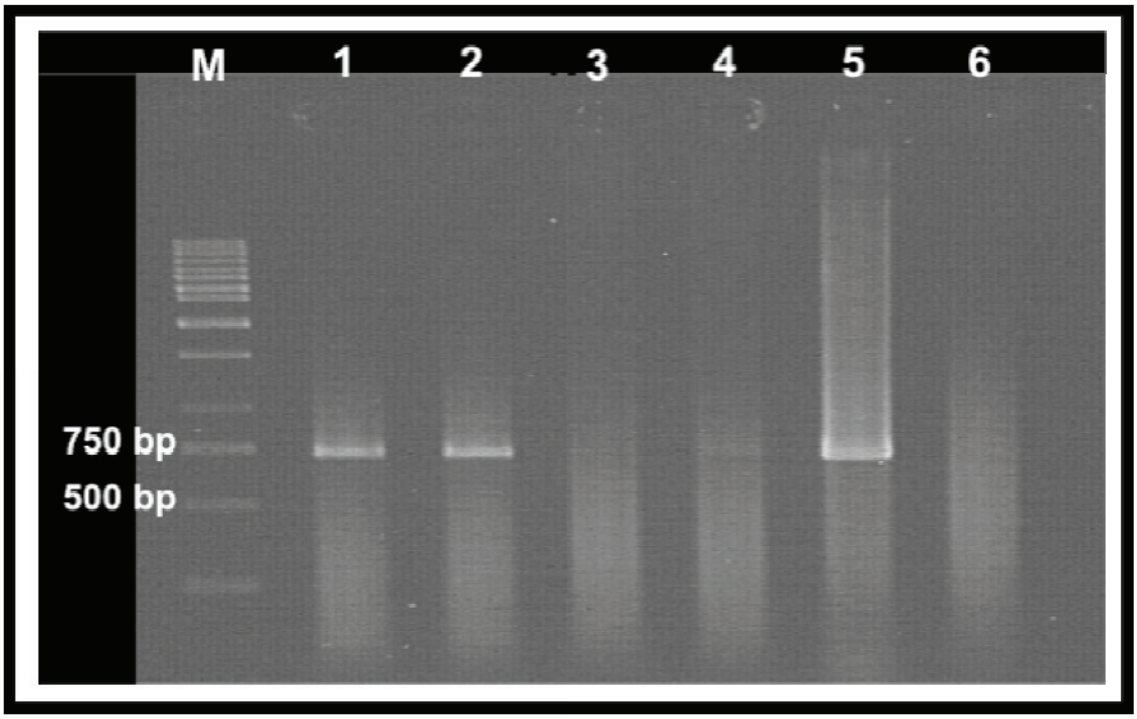


Fig. 6: COI PCR products of DNA extracted with Qiagen DNeasy kit.
M: 1Kb DNA ladder; 1, 2: *N. hombergii*; 3,4: *C. crangon*; 5: positive control; 6: negative control.

2 µl of DNA extract used as a template in PCR showed strong amplification bands on the gel. Both the DNA extracts and PCR products were observed as clear bands on the gels.

An another method, Peqlab DNA extraction kit tested along with Qiagen produced very weak DNA bands from the samples and there were no DNA extracts obtained from many of the tissues.

The paper DNA extraction technology FTA which is simple and less time consuming gave good PCR products from the crustaceans, but the result was not repeatable as the amount of DNA locked in the paper disc cut out each time was not quantifiable. The negative PCR result

can be because of presence of very less DNA in the disc taken or very high amount of DNA locked in the network of the disc. A dilution experiment was performed in order to check if the DNA is overloaded on the paper which may inhibit the PCR reaction. Serially diluted tissue homogenates in the ratios of 1/10, 1/25, 1/50, 1/75 and 1/100 were added to the paper and the PCR products from the respective dilutions showed no bands on the gel. The drawback is that it is not possible to check the DNA extracted by this method on the gel and the result of successful extraction can be observed only after the PCR.

In contrast the DNA extracted from FTA Elute papers is eluted in to solution and can be checked on the gel and so are the PCR products. 3 – 5 µl of the DNA extract is used as a template in the PCR reaction of 50 µl and good bands are observed from almost all the samples tried. A comparison of the three extraction methods tested on 15 species significantly used for the microarray is shown in Table 11.

The PCR results obtained from the Qiagen and the FTA Elute isolations are quite similar except for the crustacean *C. crangon*, which gave a PCR product with DNA from FTA Elute cards whereas no PCR product is seen from Qiagen. FTA classic cards gave no PCR product in any of the target species. The amount of template used in the PCR is slightly higher in case of DNA eluted from FTA Elute as the concentration of DNA obtained by this process is almost 10 times lower than the concentration obtained from Qiagen kit. But FTA Elute cards generated similar amount of DNA even from very less amounts of tissue such as 10 mg whereas Qiagen kit needs at least 20 mg of the tissue to obtain a good quantity of DNA that can be used in PCR. The 260/280 ratio of the DNA from FTA Elute was slightly higher than normal (>1.80) indicating the inability of the cards to remove protein debris completely but this quality assessment did not effect any of the downstream processes starting from PCR amplification, sequencing and microarray hybridisations. The time required for the extraction of DNA with FTA Elute cards is quite low when compared to the Qiagen Dneasy kit and the DNA obtained did not show any significant difference in the results to that of the Qiagen DNA.

Table 11: Comparison of the PCR product yield on an agarose gel from DNA extracted using three different methods. ++: strong band; +: intermediate band; -: no band.

Species Code	Species Name	Extraction method		
		QIAGEN kit	FTA Classic	FTA Elute
Ag	*Alpheus glaber*	++	+	++
Cc	*Crangon crangon*	-	-	++
Gr	*Goneplax rhomboides*	++	+	++
Ld	*Liocarcinus depurator*	++	+	++
Lv	*Liocarcinus vernalis*	-	-	++
Mn	*Meganyctiphanes norvegica*	++	-	++
Pm	*Pachygrapsus marmoratus*	++	-	++
Ec	*Eledone cirrhosa*	++	-	+
Lm	*Lentidium mediterraneum*	+	-	++
Cg	*Chamelea gallina*	++	+	+
Ic	*Illex coindetii*	++	+	+
Hd	*Hediste diversicolor*	++	+	++
Nh	*Nephtys hombergii*	++	-	++
Pk	*Pectinaria koreni*	++	-	+
Sm	*Sigalion mathildae*	++	-	++

Other commercial kits tried gave mixed results. Peqlab DNA kit and nucleospin kit produced DNA from polychaetes and molluscs but not from crustaceans. Even though the DNA extract bands on the gel were strong, PCR applicant bands were found to be very weak. DNA lyses kit from Viagen, PCR based Microlysis kit and the modified protocol of the magnetic bead extraction method AGOWA did not show any kind of result.

The result of ten important extraction methods out of the total 21 checked in the work with the number of samples tried and the successful PCRs is shown in Table 12.

Table 12: DNA extraction profiling of different organisms with the 10 different extraction methods. The total number of individuals used upon the number of individuals giving successful extracts is shown. (-) Indicates that no individual tissue is tried with that particular method.

Name of the Organism	Family	Phyllum	No. of individuals used	FTA Elute	FTA	QIAGEN	Nucleospin	AGOWA	Peqlab	Viagen	Microlysis	Chelex	Phenol-Chloroform
Nephtys hombergii	Nephtyidae	Annelida	18	10/11	3/16	17/24	0/1	1/1	10/18	0/1	0/1	0/2	1/10
Hediste diversicolor	Neriididae		15	3/5	1/5	7/15	0/1	0/1	1/4	0/1	0/1	-	1/7
Pectinaria koreni	Terebellidae		10	1/1	0/1	10/10	-	-	-	-	-	-	0/3
Sigalion mathildae	Sigalionidae		7	6/6	0/8	6/10	1/1	0/1	4/6	0/1	0/1	-	-
Notomastus latericeus	Capitellidae		2	1/2	-	-	-	-	-	-	-	-	-
Chaetozone setosa	Cirratulidae		6	2/3	-	3/3	-	-	-	-	-	-	-
Magelona johnstoni	Magelonidae		6	2/3	-	2/3	-	-	-	-	-	-	-
Scoloplos armiger	Orbiniidae		4	2/2	-	2/2	-	-	-	-	-	-	-
Owenia fusiformis	Oweniidae		3	1/3	-	1/3	-	-	-	-	-	-	-
Poecillochaetus serpens	Pecillochaetidae		2	2/2	-		-	-	-	-	-	-	-
Spiophanes bombyx	Spionidae		6	2/3	-	3/3	-	-	-	-	-	-	-
Lanice conchilega	Terebellidae		3	3/3	-	3/3	-	-	-	-	-	-	-
Meganyctiphanes norvegica	Euphausiidae	Arthropoda	10	1/1	0/1	8/8	-	-	1/1	-	-	-	-
Alpheus glaber	Alpheidae		10	1/1	4/8	10/10	-	-	1/1	-	-	-	-
Pachygrapsus marmoratus	Grapsidae		8	1/1	0/1	5/5	-	-	1/1	-	-	-	-
Goneplax rhomboides	Goneplacidae		10	1/1	4/5	6/10	-	-	1/1	-	-	-	-
Liocarcinus depurator	Portunidae		10	1/1	3/5	7/7	-	-	1/1	-	-	-	-
Liocarcinus vernalis	Portunidae		10	1/1	0/1	5/7	-	-	1/1	-	-	-	-
Crangon crangon	Crangonidae		30	4/6	1/16	16/30	0/3	3/7	4/20	2/9	1/3	0/2	0/17
Thysanoessa longicaudata	Euphausiidae		1	1/1	1/1	1/1	-	-	-	-	-	0/1	-
Themisto abyssorum	Hyperiidae		1	1/1	0/1	1/1	-	-	-	-	-	0/1	-
Unidentified	-		1	1/1	0/1	0/1	-	-	-	-	-	0/1	-
Calanus hyperboreus	Calanidae		3	1/1	1/1	0/1	-	-	-	-	-	0/1	-
Chamelea gallina	Veneridae	Mollusca	10	0/1	4/10	3/10	-	-	1/1	-	-	-	-
Illex coindetti	Ommastrephidae		10	0/1	1/1	7/10	-	-	1/1	-	-	-	-
Eledone cirrhosa	Octopodidae		10	0/1	0/1	7/10	-	-	1/1	-	-	-	-
Lentidium mediterraneum	Corbulidae		10	1/3	1/10	9/10	-	-	0/1	-	-	-	5/11
Tridacna crocea	Tridacnidae		4	4/4	2/4	-	-	-	-	-	-	3/4	-
Gadus morhua	Gadidae	Chordata	2	2/2	2/2	-	-	-	-	-	-	-	-
Cod fish	Gadidae		1	1/1	0/1	1/1	-	-	-	-	-	-	-
Amphiprion ocellaris	Pomacentridae		2	2/2	2/2	-	2/2	-	-	-	-	1/2	1/2
Jellyfish sps.	-	Cnidaria	1	1/1	0/1	0/1	-	-	-	-	-	-	-
Entacmaea quadricolor	Actiniidae		3	0/3	0/3	1/3	-	-	-	-	-	-	-
Linckia leavigata	Ophidiasteridae	Echinodermata	6	6/6	0/6	6/6	-	-	-	-	-	-	-
Eukrohnia sps.	Eukrohniidae	Chaetognatha	1	1/1	1/1	1/1	-	-	-	-	-	-	-
Gut content(Trigla lucerna)				-	0/4	7/35	-	-	-	-	-	-	6/6

4.2. PCR of the 16S and COI genes

Amplification of the 16S and COI fragments worked well with the standard protocol using universal primers as mentioned in the section 2.5. All the reactions are performed either with 2 µl of the Qiagen extracted DNA template or 3 µl of the FTA Elute extracted template. At the optimized annealing temperature the amplified products produced by PCR could be seen on the agarose gel as clear bands and without any unspecific products. 5 µl of PCR product is added to each lane with 1 µl of 6X loading dye. 16S PCR products are about 550 bp long (Fig. 7) and COI products are around 708 bp in length. (Fig. 8)

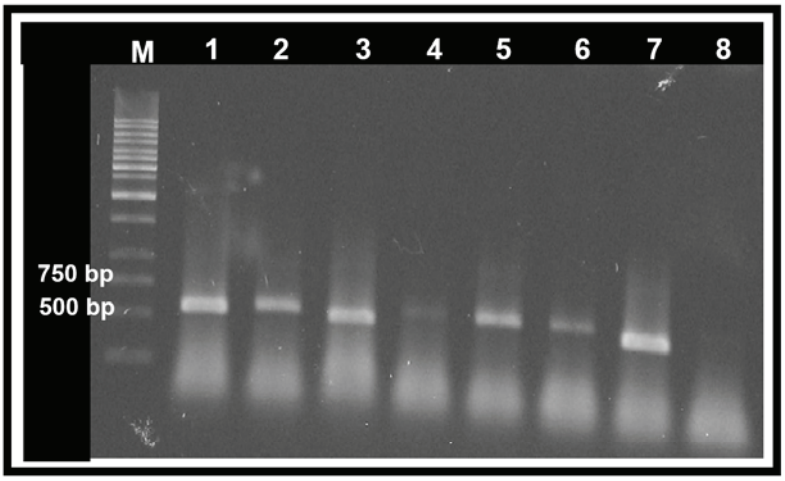


Fig. 7: 16S PCR products of polychaete DNA extracted with the FTA Elute method.
M: 1kb DNA ladder, 1: *N. hombergii*; 3: *H. diversicolor*; 2,4,5,6: *C. crangon*; 7: positive control; 8: negative control

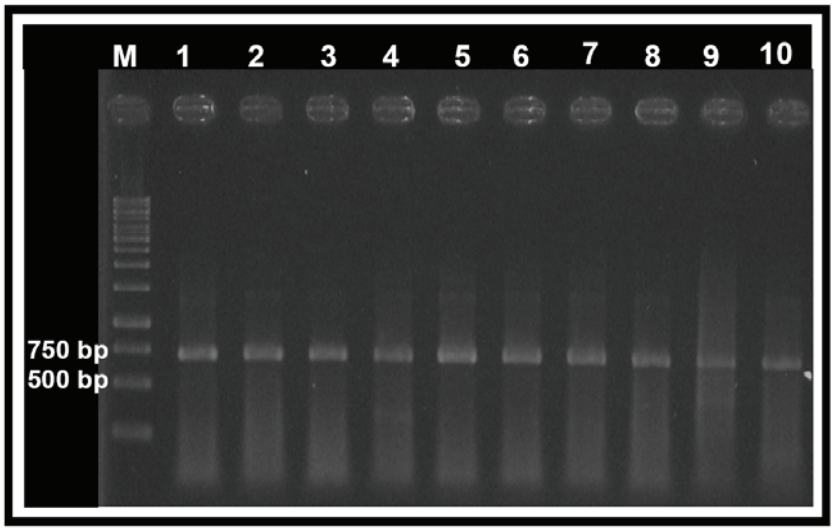


Fig. 8: COI PCR products of different invertebrate DNA extracted with the Qiagen kit.
M: 1Kb DNA ladder; 1: *A. glaber*; 2,3,4: *C. gallina*; 5,6,7: *E. cirrhosa*; 8,9,10: *L. mediterraneum*

4.3. Sequencing of the 16S and COI genes

All the PCR products were purified with Qiagen PCR purification kit and the pure DNA is used as a template for sequencing PCR reaction. The chromatograms obtained from the sequencing reaction are edited with Chromas, Seqman and Mega programs and the final readable sequence is a bit shorter than the actual length of the amplicon. Both the forward and reverse strands are sequenced separately and after editing are aligned together to fill up the missed gaps and a final sequence is obtained. Each individual sequence is then verified with the NCBI BLAST search for the specificity of the species. There are not many sequences of the targets used in the work are available in the databank and in cases where there are no sequences of the exact species, the closest relative was considered and checked for the authenticity of the sequence. The total number of sequences obtained during the work and the number of sequences that are used from the databank can be checked in the following Table 13. And the alignment of all the target sequences obtained can be seen in 8.5.1 and a phylogenetic tree understand the relationships between the targets 8.3.1.

All the sequences obtained in the work can be viewed in the appendix in fasta format along with a phylogenetic tree depicting the cluster formation of the different individuals of the same species showing the intra specific and inter specific variations 8.3.1.

Table 13: Total number of sequences obtained from the work and the databank.

Phyllum	Species name	Species code	16S sequences		COI sequences		NCBI retrieved data		
			Targets		Targets		Family	16S sequences	COI sequences
			From work	NCBI	From work	NCBI			
Arthropoda	*Alpheus glaber*	Ag	4	0	3	0	Alpheidae	142	236
	Crangon crangon	Cc	8	0	9	0	Crangonidae	0	1
	Goneplax rhomboides	Gr	3	0	3	0	Goneplacidae	0	0
	Liocarcinus depurator	Ld	3	0	3	1	Portunidae	51	146
	Liocarcinus vernalis	Lv	4	0	3	0	Portunidae	51	146
	Meganyctiphanes norvegica	Mn	4	1	3	2	Euphausiidae	60	130
	Pachygrapsus marmoratus	Pm	4	2	3	0	Grapsidae	31	28
Mollusca	*Eledone cirrhosa*	Ec	3	3	3	1	Octopodidae	136	35
	Lentidium mediterraneum	Lm	2	0	3	0	Corbulidae	0	1
	Chamelea gallina	Cg	2	3	3	2	Veneridae	154	65
	Illex coindetii	Ic	2	1	3	1	Ommastrephidae	14	68
Annelida	*Hediste diversicolor*	Hd	5	0	7	10	Nereididae	2	12
	Nephtys hombergii	Nh	13	0	15	0	Nephtyidae	0	0
	Pectinaria koreni	Pk	3	3	7	346	Pectinaridae	10	373
	Sigalion mathildae	Sm	3	0	5	0	Sigalionidae	1	7
	Total sequences obtained during the work		63	13	73	363			

4.4. *In silico* characterization of capture oligonucleotides

The total number of 63 16S sequences and 73 COI sequences (Table 15 and Table 16) are generated during the work. The identity of the sequence is confirmed from NCBI BLAST results, those sequences which show at least 95% matching with the existing sequences are considered for the probe designing. All the sequences obtained in the work and also the sequences of the species already present in the NCBI databank (13 16S and 363 COI sequences) are aligned and after eliminating the sequences which are too short in the alignment a total of 70 16S and 58 COI sequences are used for designing the probes. The numbers of sequences per each species that are considered for probe designing are shown in Table 14. The alignment of all the sequences used for probe designing can be seen in the appendix 8.2.

Following the criteria mentioned in the materials and methods section 3.7.1, 239 captures were generated in the calculation from the 16S sequences and 84 from the COI sequences. The probes are then checked Insilco for the specificity and any mismatches with the other target and non-target sequences. The probes are also checked for self-complimentarity, dimer formation and hairpin structure formation. A probe which matches only with the specific target sequence and doesn't show any resemblance with the other targets or non-targets is definitely selected. In case of those probes which match with the specific target completely but also show resemblance to other targets or non targets with certain mismatches, they are selected only if the mismatches are with at least 3 or more base pairs and preferably these mismatch bases are present in the middle of the sequence and not at the edges. After checking every probe thoroughly a total of 43 captures (targeting 16S gene of 15 species and COI gene of 5 species) and 30 probes (targeting COI gene of 15 species) are selected to spot on the respective microarrays. The length and GC content parameters of the selected 16S probes are in the range of 23 – 46 bp and 35 – 52% and of COI probes are 23 – 35 bp and 35 – 54%. All the probes selected are showed in Table 15 and Table 16 along with the parameters that are considered for the selection. Initially two separate chips, one containing majority of the16S probes and the other COI probes were developed. After optimizing the hybridization conditions and evaluating the probes for their specificity, an invertebrate microarray was prepared with the both the 16S and COI probes that gave good signals on individual arrays. The captures that are marked with an asterisk (*) in Table 15 and Table 16 are not considered for the combined invertebrate microarray.

Table 14: Total number of sequences considered for the probe designing.

Phyllum	Species Name	Species Code	Collection Site	16S sequences			COI sequences		
				Targets		Non-targets	Targets		Non-targets
				From work	NCBI		From work	NCBI	
Arthropoda	*Alpheus glaber*	Ag	Adriatic Sea	4	0	601	3	0	1085
	Crangon crangon	Cc	North Sea	5	0		2	0	
	Goneplax rhomboides	Gr	Adriatic Sea	3	0		3	0	
	Liocarcinus depurator	Ld	Adriatic Sea	3	0		3	1	
	Liocarcinus vernalis	Lv	Adriatic Sea	3	0		2	0	
	Meganyctiphanes norvegica	Mn	Adriatic Sea	4	1		3	2	
	Pachygrapsus marmoratus	Pm	Adriatic Sea	4	1		2	0	
Mollusca	*Eledone cirrhosa*	Ec	Adriatic Sea	3	3	726	3	1	1624
	Lentidium mediterraneum	Lm	Adriatic Sea	2	0		2	0	
	Chamelea gallina	Cg	Adriatic Sea	2	3		3	2	
	Illex coindetii	Ic	Adriatic Sea	2	1		3	1	
Annelida	*Hediste diversicolor*	Hd	Mediterranean Sea	4	0	162	5	0	461
	Nephtys hombergii	Nh	Mediterranean Sea & North Sea	13	0		9	0	
	Pectinaria koreni	Pk	North Sea	3	3		4	0	
	Sigalion mathildae	Sm	North Sea	3	0		4	0	
			Total sequences used for probe designing	70		1489	58		3170

Table 15: List of all the oligonucleotides selected for the 16S invertebrate microarray along with their sequence and the chemical parameters checked during the calculation. The captures with the * symbol are not used further for the combined invertebrate chip.

No.	Oligo name	Complete Oligo name	Sequence in 5'- to 3'- direction	Oligo length	TM	%GC	MFE Oligo	Dimer MFE	Ranking
1	Ag_16S_a	Alpheus-glae_l23_p187	AGTTAGGTCATTATGCTGGGGCG	23	82,56	52%	-0,2	-1,75	1,4761
2	Ag_16S_b	Alpheus-glae_l29_p182	GTTTAAGTTAGGTCATTATGCTGGGGCGG	29	83,7	48%	-0,2	-2,77	1,9454
3	Cc_16S_a	Cangon-cangon-16S_l25_p36	GAATGATCGGACAAGGGGCTAACTG	25	82,4	52%	0	-1,33	1,8339
4	Cc_16S_b	Cangon-cangon-16S_l28_p31	GGAATGAATGATCGGACAAGGGGCTAAC	28	83,79	50%	0	-1,58	2,5341
5	Cc_16S_c	Cangon-cangon-16S_l30_p31	GGAATGAATGATCGGACAAGGGGCTAACTG	30	84,85	50%	0	-1,68	2,3477
6	Gr_16S_a	Goneplax-homoides_l46_p26	GAACTTGTATGAATGGTTGGACAAAGGAAAAGCTGTCTCTATTATA	46	83,83	35%	0	-3,7	2,7016
7	Ld_16S_a	Liocacinus-depuato_l29_p184	CTTAGAAAAATGTATTGGGTTGGGGCGAC	29	82,51	45%	0	-1,98	3,3983
8	Ld_16S_b	Liocacinus-depuato_l38_p184	CTTAGAAAAATGTATTGGGTTGGGGCGACTAAGGTATA	38	83,84	39%	0	-3,3	1,996
9	Lv_16S_a	Liocacinus-venalis_l28_p183	ATTCAGAAAAGTGTATTGGGTTGGGGCG	28	84,23	46%	0	-0,93	2,0358
10	Lv_16S_b	Liocacinus-venalis_l32_p177	AGGTTTATTCAGAAAAGTGTATTGGGTTGGGG	32	82,98	41%	0	-0,84	1,8139
11	Mn_16S_a	Meganychtiphanes-novegica_l29_p38	ATGGTCGGACAAGAAACAGACTGTCTTCA	29	84,11	45%	-1,9	-5,39	2,432
12	Mn_16S_b	Meganychtiphanes-novegica_l33_p389	GGTGTAGCAGCTTTAATAGAAGGTCTGTTCGAC	33	83,63	45%	0	-3,52	3,2701
13	Pm_16S_a	Pachygapsus-mamoatus_l23_p168	TCGAGGGCTATAAAGGCTTGGTG	23	82,23	52%	0	-1,43	1,24
14	Pm_16S_b	Pachygapsus-mamoatus_l31_p160	AGTCTAATTCGAGGGCTATAAAGGCTTGGTG	31	83,75	45%	0	-1,58	1,2545
15	Ec_16S_a	Eledone-cihosa_l28_p207	GATTGGGGTGATCAAGGAATAAAAAAGA	28	78,62	36%	0	-1,32	2,8681
16	Ec_16S_b	Eledone-cihosa_l23_p277	GATAAATAAACCGAGAGTTTTGC	23	73,4	35%	0	-1,43	3,0243
17	Lm_16S_a	Lentidium-mediteaneum_l23_p23	TGGGGAAAGGTATGAAGGGACTG	23	81,98	52%	0	-0,29	1,3317
18	Lm_16S_b	Lentidium-mediteaneum_l25_p277	CTGGAGGGCTAATCGAATGGAAAGT	25	81,9	48%	0	-0,88	1,6808
19	Lm_16S_c	Lentidium-mediteaneum_l25_p262	CAGCGTAATTTTCTACTGGAGGGCT	25	81,97	48%	0	-1,61	2,0107
20	Cg_16S_a	Chamelea-gallina_l45_p24	GGAGAATGGTATGAATGGTTTAACGTAGAATAACTGTCTTTGGAA	45	83,93	36%	0	-3,26	2,7763
21	Ic_16S_a	Illex-coindetii_l45_p103	GCTTGAATTTTTTAAAGGGACGAGAAGACCCTAATGAGCTTATAA	45	84,23	36%	-0,9	-4,06	2,8208
22	Hd_16S_a	Hediste-divesicolo-16S_l25_p212	GCTAATTAATCACACACACACCCAA	25	78,96	40%	0	-0,05	2,0853
23	Hd_16S_b	Hediste-divesicolo-16S_l34_p181	TGGGACAACCTAAAGACAAATAAACCTCTTAGCT	34	83,06	38%	-0,1	-1,7	2,2556
24	Nh_16S_a	Nephtys-homegii-16S_l24_p183	AAAAGGAACAAGTTTGGTTGGGGC	24	82,3	46%	0	-2,17	2,3455
25	Nh_16S_b	Nephtys-homegii-16S_l31_p183	AAAAGGAACAAGTTTGGTTGGGGCGACAAAG	29	85,92	45%	-0,2	-3,53	3,1517
26	Pk_16S_a	Pectinaia-koeni-16S_l24_p208	GGGGCGGCTGAGGAAAATTAAATC	24	81,83	50%	0	-2,61	2,5055
27	Pk_16S_b	Pectinaia-koeni-16S_l23_p48	GAGGGCCAAGCTGTCTCTTTAGT	23	82,19	52%	0	-3,1	2,865
28	Pk_16S_c	Pectinaia-koeni-16S_l27_p301	GGATCAAAGAAAATAGCTACCTCGGGG	27	81,73	48%	0	-2,34	3,0267
29	Sm_16S_a	Sigalion-mathildae-16S_l23_p183	CGACCCAGGAGCATTTAAACCCT	23	82,79	52%	0	-0,94	1,255
30	Sm_16S_b	Sigalion-mathildae-16S_l27_p24	GAAGGCTGGAATGAACGGATAAACGAG	27	82,23	48%	-0,5	-2,08	2,6776
31	Lio_16S_a	Liocacinus_l26_p198*	TTGGGTTGGGGCGACTAAGGTATAAT	26	83,07	46%	0	-1,54	4,9113
32	Cc_COI_c	Cangon-cangon-COI_l30_p131	CCTGCTTTAACTCTTCTTCTATCTAGAGGA	30	79,49	40%	0	-1,89	2,633
33	Cc_COI_d	Cangon-cangon-COI_l27_p269*	GCTGGGGTTTCATCTATTCTAGGGGCT	27	84,86	52%	-0,7	-3,69	5,1027
34	Hd_COI_c	Hediste-divesicolo-COI_l26_p231	GTCCGTCAGTAGACCTTGCAATCTTC	26	81,91	50%	0	-2,07	2,6568
35	Hd_COI_d	Hediste-divesicolo-COI_l24_p124*	ATTACCACCATCCCTAACCCTTCT	24	81,05	46%	0	-0,01	3,0896
36	Hd_COI_e	Hediste-divesicolo-COI_l28_p119*	TGATTATTACCACCATCCCTAACCCTTC	28	80,81	43%	0	-0,12	3,1703
37	Nh_COI_c	Nephtys-homegii-COI_l23_p157*	CGCAGCTGTAGAAAAAGGAGTCG	23	81,11	52%	0	-3,86	2,2199
38	Nh_COI_d	Nephtys-homegii-COI_l36_p129*	CCCCTTCTTTAATTCTTCTTGTTATATCCGCAGCTG	36	83,57	42%	0	-2,12	1,4552
39	Pk_COI_c	Pectinaia-koeni-COI_l25_p61	CCCATTAATACTTGCTGCTCCAGAC	25	80,67	48%	0	-0,61	2,3125
40	Pk_COI_d	Pectinaia-koeni-COI_l25_p253	CTTTTCTCTCCACTTAGCCGGGATC	25	81,93	52%	0	-1,75	1,5008
41	Sm_COI_c	Sigalion-mathildae-COI_l38_p233	CCTTCAGTTGATCTCGCTATCTTTTCTCTTCATATTGC	38	82,94	39%	0	-0,94	2,7441
42	Sm_COI_d	Sigalion-mathildae-COI_l38_p128*	CCTCCTTCTCTAATTCTTCTTTTATCTTCAAGAGCCGT	38	83,53	39%	0	-0,62	1,6839
43	Sm_COI_e	Sigalion-mathildae-COI_l24_p158*	AGAGCCGTTGAAAAAGGAGTTGGC	24	83,44	50%	-0,4	-3,11	1,5541

Table 16: List of all the oligonucleotides selected for the COI invertebrate chip.

No.	Oligo Name	Complete oligo name	Sequence in 5'- to 3'- direction	oligo length	TM	%GC	MFE Oligo	Dime MFE	Ranking
1	Ag_COI_a	Alpheus-glae_l23_p394	TCCTACTCAGGCTCCCAGTTCTA	23	82,13	52%	0	-0,88	0,6935
2	Ag_COI_b	Alpheus-glae_l25_p243	TTAGGGATCTTCTCGCTACACCTCG	25	82,49	52%	0	-1,68	1,724
3	Cg_COI_a	Chamelea-gallina_l27_p203	CTTATCTAGGGCTCTGTCTCATTCGGG	27	82,51	52%	0	-0,79	1,7667
4	Cg_COI_b	Chamelea-gallina_l24_p108	CTTAGGTTCTGATTGTTGCCGGTG	24	81,25	50%	0	-2,14	2,2142
5	Cc_COI_a	Cangon-cangon_l33_p126	CCTCCTGCTTTAACTCTTCTTCTATCTAGAGGA	33	81,95	42%	0	-2,18	2,4228
6	Ec_COI_a	Eledone-cihosa_l31_p198	CCTCCCCTATCAAGTAATTTAGCCCACATAG	31	82,11	45%	0	-0,31	1,2419
7	Ec_COI_b	Eledone-cihosa_l35_p230	ACCCTCTGTTGACCTAGCAATTTTCTCTTTACATC	35	83,27	40%	0	-0,93	2,6238
8	Gr_COI_a	Goneplax-homoides_l35_p230	CGCCTCTGTTGATATAGGTATTTTTTCCCTTCATT	35	82,1	37%	0	-2,01	2,1083
9	Gr_COI_b	Goneplax-homoides_l30_p122	ACTTCCTCCCTCTTTAACTCTCTTACTAAT	30	79,22	37%	0	-0,02	2,3578
10	Ic_COI_a	Illex-coindetii_l35_p126	CCACCATCTTTAACTATATTACTAGCCTCTTCAGC	35	81,35	40%	0	-0,34	1,578
11	Ic_COI_b	Illex-coindetii_l33_p201	CCCTTATCTAGAAATTTATCTCATGCTGGACCC	33	81,97	42%	0	-1,31	2,0705
12	Lm_COI_a	Lentidium-mediteaneum_l28_p199	CCCCATTATCCGGTAATACAGCTCACTC	28	82,81	50%	0	-1,44	1,3135
13	Lm_COI_b	Lentidium-mediteaneum_l31_p239	AGACTTTCTTATTTTATCGCTACACCTCGGT	31	81,72	39%	0	-0,64	1,4574
14	Ld_COI_a	Liocacinus-depuato_l30_p120	TTACTCCCACCTTCGCTAACTCTTCTCCTC	30	84,73	50%	0	-0,04	1,2648
15	Lv_COI_a	Liocacinus-venalis_l30_p140	GCTTCTTCTCATGAGAGGTATAGTCGAGAG	30	81,36	47%	0	-8,29	2,2431
16	Lv_COI_b	Liocacinus-venalis_l30_p488	TGACCCAGTTCTCTATCAGCACTTGTTTTG	30	82,96	43%	0	-0,58	2,4052
17	Mn_COI_a	Meganyctiphanes-novegica_l34_p126	CCACCTTCTTTAACTCTTTTATTAGGCAGAGGTC	34	82,07	41%	0	-1,52	1,6287
18	Nh_COI_a	Nephtys-homegii-NS_l35_p128	CCCTTCTTTAATTCTTCTTGTTATATCCGCAGCTG	35	82,4	40%	0	-2,18	1,5635
19	Nh_COI_b	Nephtys-homegii-NS_l26_p151	TATCCGCAGCTGTAGAAAAAGGAGTC	26	81,45	46%	0	-8,45	3,3183
20	Pm_COI_a	Pachygapsus-mamoatus_l31_p125	ACCTCCCTCCTTATCTCTCTTACTTACAAGA	31	81,83	42%	0	-0,32	1,2187
21	Pm_COI_b	Pachygapsus-mamoatus_l33_p317	CAACATACGCTCTTATGGTATGACAATAGACCA	33	81,66	39%	-0,9	-3,18	2,2847
22	Pk_COI_a	Pectinaia-koeni_l28_p200	CCCCTTATCAAGAAACCTTGCACATGCG	28	84,63	50%	-0,5	-2,37	2,2685
23	Pk_COI_b	Pectinaia-koeni_l30_p9*	TTCTTTCTTGTCATACCCGTCTTTATTGGT	30	80,83	37%	0	-0,45	2,3074
24	Sm_COI_a	Sigalion-mathildae_l34_p122	ACTCCCTCCTTCTCTAATTCTTCTTTTATCTTCA	34	80,73	35%	0	-0,04	1,9138
25	Sm_COI_b	Sigalion-mathildae_l27_p150	TCTTCAAGAGCCGTTGAAAAAGGAGTT	27	81,92	41%	-0,4	-3,42	2,2834
26	Mn_COI_b	Meganyctiphanes-novegica_l27_p150	GGCAGAGGTCTTGTAGAAAGAGGAGTC	27	82,75	52%	0	-2,07	2,3385
27	Cc_COI_b	Cangon-cangon_l35_p151	CTAGAGGATTAGTAGAAAGAGGAGTAGGAACTGGA	35	82,36	43%	0	-0,55	3,276
28	Hd_COI_a	Hediste-divesicolo_l35_p185	ATGAACAGTATACCCGCCATTAGCCAGAAATATTG	35	83,44	40%	0	-0,88	2,9891
29	Hd_COI_b	Hediste-divesicolo_l28_p272	GGTCTCATCTATTTTGGGAGCCCTAAAC	28	81,65	46%	-0,6	-5,18	3,1431
30	Ld_COI_b	Liocacinus-depuato_l26_p193	TCTACCCTCCCCTATCGGCTGCTATT	26	85,64	54%	0	-0,68	2,5661

4.5. Target preparation for microarray hybridizations

16S and COI genes were successfully amplified with the universal primers at specific annealing temperatures (Fig. 6). Strong amplification bands were observed from the DNA of the 15 species used in the work. Cy5 labelled targets were prepared by amplifying the DNA with labelled primers. For this purpose the reverse primers of both 16S and COI genes were labelled and used in the PCR reaction.

In order to reduce the experimental steps for microarray hybridization multiplex PCR was designed in which both the genes are simultaneously amplified from the same template. Optimized multiplex PCR reaction and the temperature profile produced very distinct and strong PCR products of both the genes from all the species except the crustacean *C. crangon*. 16S gene from *C. crangon* amplifies at 50 °C and COI at 40 °C annealing temperature. At the optimized multiplex temperature of 45 °C, only 16S PCR product is obtained and there is no amplification of the COI gene. So in case of *C. crangon* both the genes are amplified separately for the microarray experiments. Both the 16S and COI are successfully amplified from all the other species at 45 °C using both sets of universal primers. (Fig. 9)

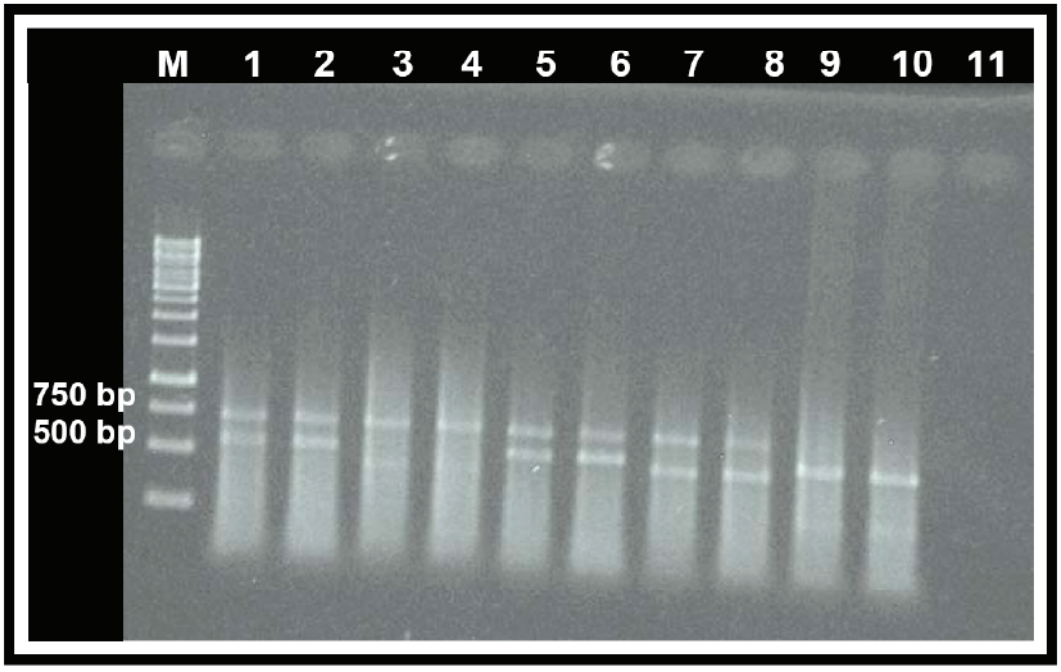


Fig. 9: PCR products of the 16S and COI genes amplified in the same reaction

M: Molecular ladder; 1,2: *A. glaber*; 3,4: *L. mediterraneum*; 5,6: *L. vernalis*; 7,8: *C. gallina*; 9,10: *C. crangon*; 11: NK

4.6. Optimization of microarray conditions

4.6.1. Optimization of hybridization temperature and incubation time:

Initially single target hybridizations were carried on with three different targets at four different hybridization temperatures (50°, 55°, 60° and 65°C) and at three different hybridization durations of 2 hours, 3 hours and 4 hours. Best results were obtained at 60 °C temperature and when carried for 3 hours (Fig. 10). At 50 °C and 55 °C the signal intensity from the specific probes when the target *H. diversicolor* is hybridized is very high, lowering a bit at 60 °C and only one probe showed a signal at 65 °C. Some weak signals of unspecific binding is observed from four probes both at 50 °C and 55 °C, whereas only two unspecific signals are found at 60 °C and at 65 °C (Fig. 11). So, 60 °C is found out to be the optimal temperature for specific hybridizations on this chip. Similarly with the incubation 3 hours is found out to be the optimal time to get good hybridization signals (data not shown).

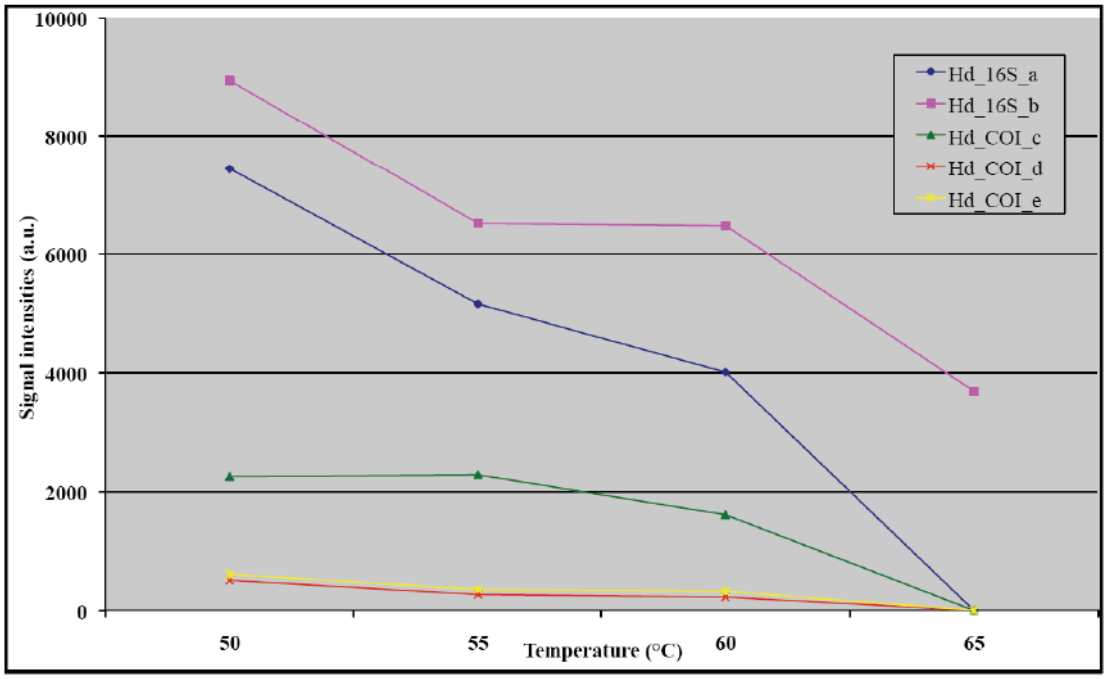


Fig. 10: Signal intensities from the specific probes for target *H. diversicolor* when hybridized at different temperatures.

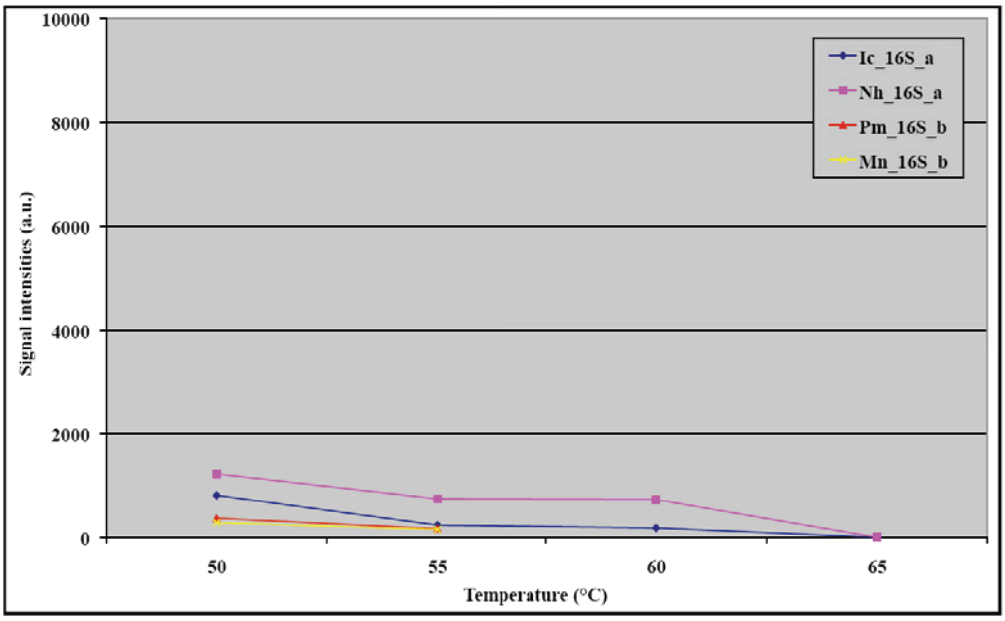


Fig. 11: Signal intensities from the unspecific binding of the target *H. diversicolor* when hybridized at different temperatures.

Similarly the signals observed in Fig. 12 when the target *L. vernalis* is hybridized at four different temperatures also shows that though the specific signal is lowered the unspecific signals are significantly reduced at 60°C. Though the signals at 65 °C shows similar pattern, the positive control is lowered to a great extent. So, further experiments are carried out at 60 °C.

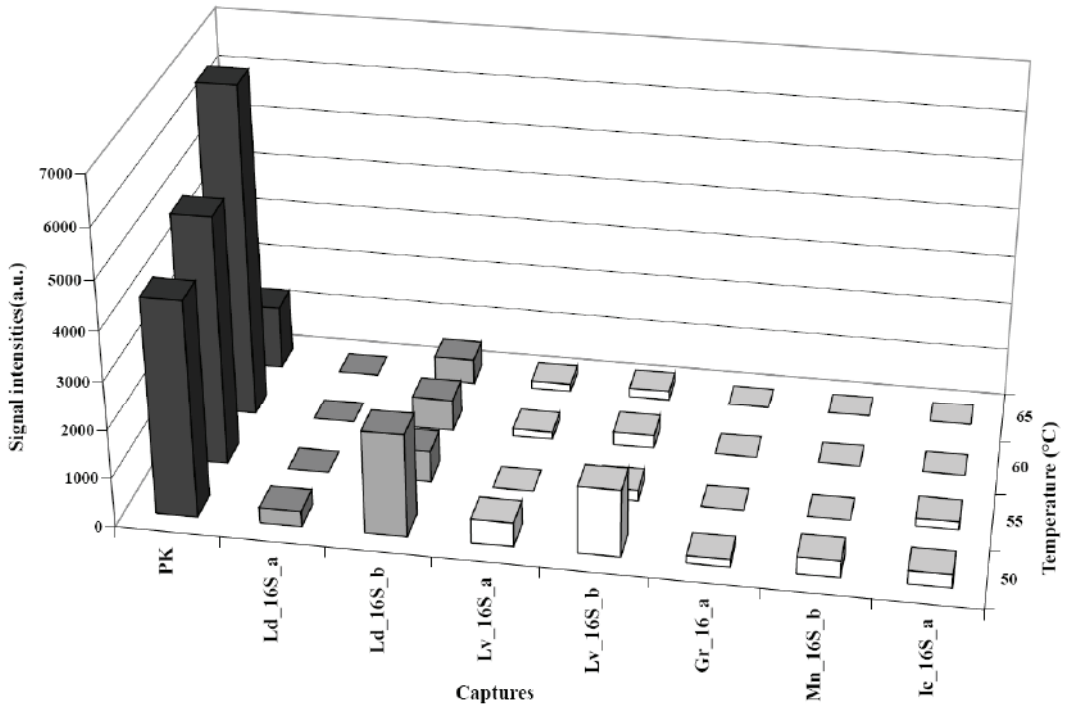


Fig. 12: Signal pattern of the target *L. depurator* when hybridized on to the microarray at different temperatures.

4.7. Specificity test of the microarray

4.7.1. Single target hybridizations

In order to check the behaviour of each target separately on the chip and also the signal patterns obtained by hybridizing different genes is checked by amplifying the two genes separately and hybridizing separately.

After the conditions are optimized single target hybridizations of 16S gene with two individuals of each target species were carried on. Each hybridization was performed thrice and the average signal intensity of both the individuals from a total of six hybridizations was considered for the final calculation.

From the signal pattern (Fig. 13) it can be seen that at least one 16S probe gave a very good specific signal with the target DNA. Probes from two species (*C. crangon* and *G. rhomboides*) gave no specific signals. Only very few probes showed false positive signals with non specific targets which is almost comparable with the negative control signal and the signal from the specific target is much higher than the false positives. The results of all the 16S single target hybridizations are shown in the Fig. 13.

COI gene from all the 15 species was labelled by amplifying with the Cy5 end labelled reverse primer and three individuals of all the 15 species were hybridized separately in a concentration of 10 nM at 60 °C for 3 hrs. The signal intensities shown in the Fig. 14 are the mean of three hybridizations. Each target species has two specific probes from the COI gene and all the probes gave a signal with the target that is not observed from the 16S captures. The signal intensity is also higher compared to the signals from the 16S probes. Four false positive signals are observed from the *H. diversicolor*, *L. vernalis* and *L. depurator* probes hybridizing with *G. rhomboides*, *L. mediterraneum*, *E. cirrhosa* and *S. mathildae* but their intensity was very low compared to the true positive signals.

All the probes from the 16S gene of *L. depurator* always give a cross hybridization signal with the *L. vernalis* target whereas only one probe shows a slight cross hybridization from the COI targets of the respective species.

Some species such as *M. norvegica* and *H. diversicolor* show a very high signal from the 16S probes whereas majority of the species except the *M. norvegica* and *N. hombergii* show high signals from at least one probe of COI. The crustacean *C. crangon* was not detected with the 16S probes but a slight signal is seen from the COI probes.

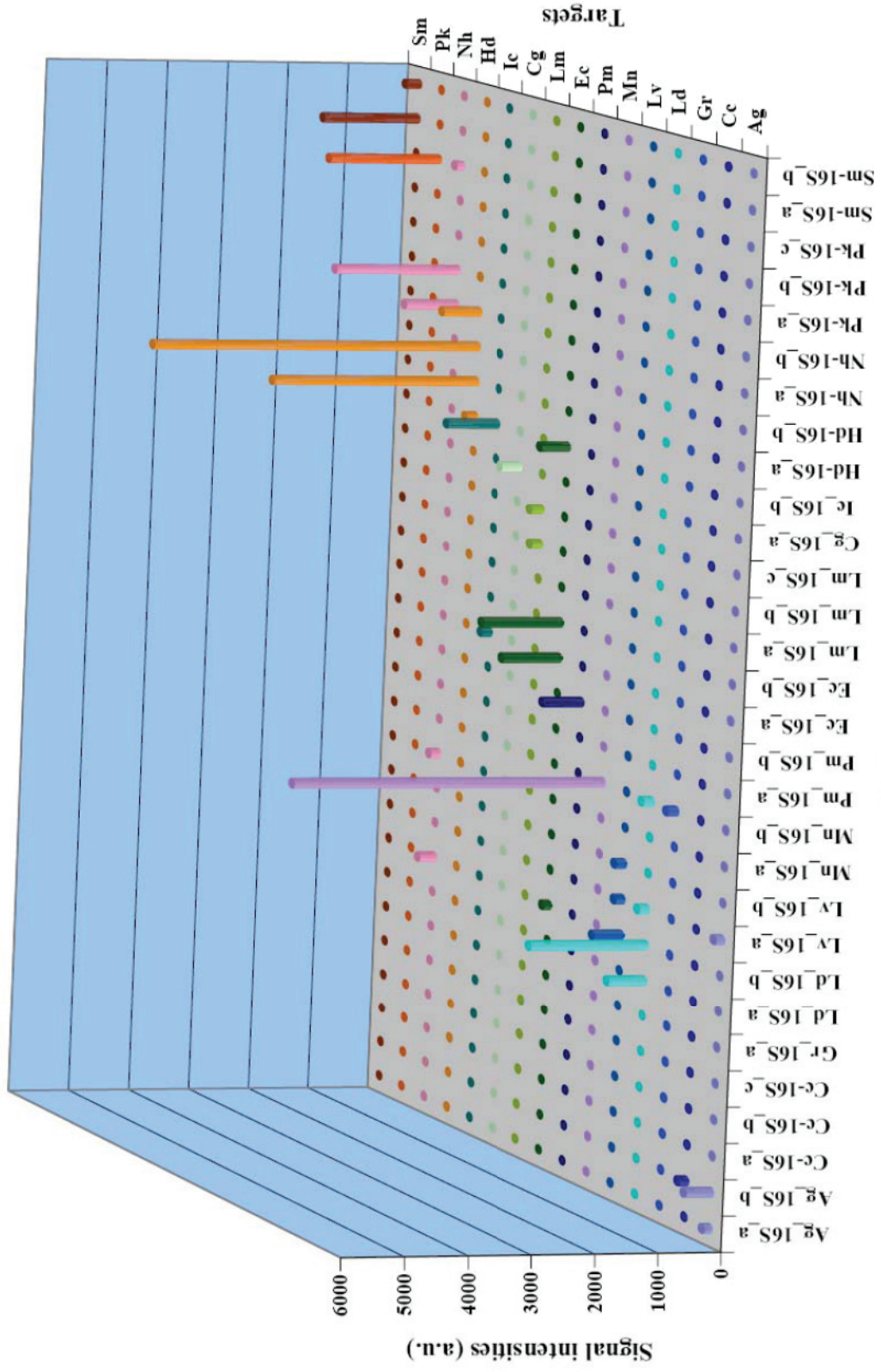


Fig. 13: Signal pattern of all the 15 single target hybridizations on the 16S chip.

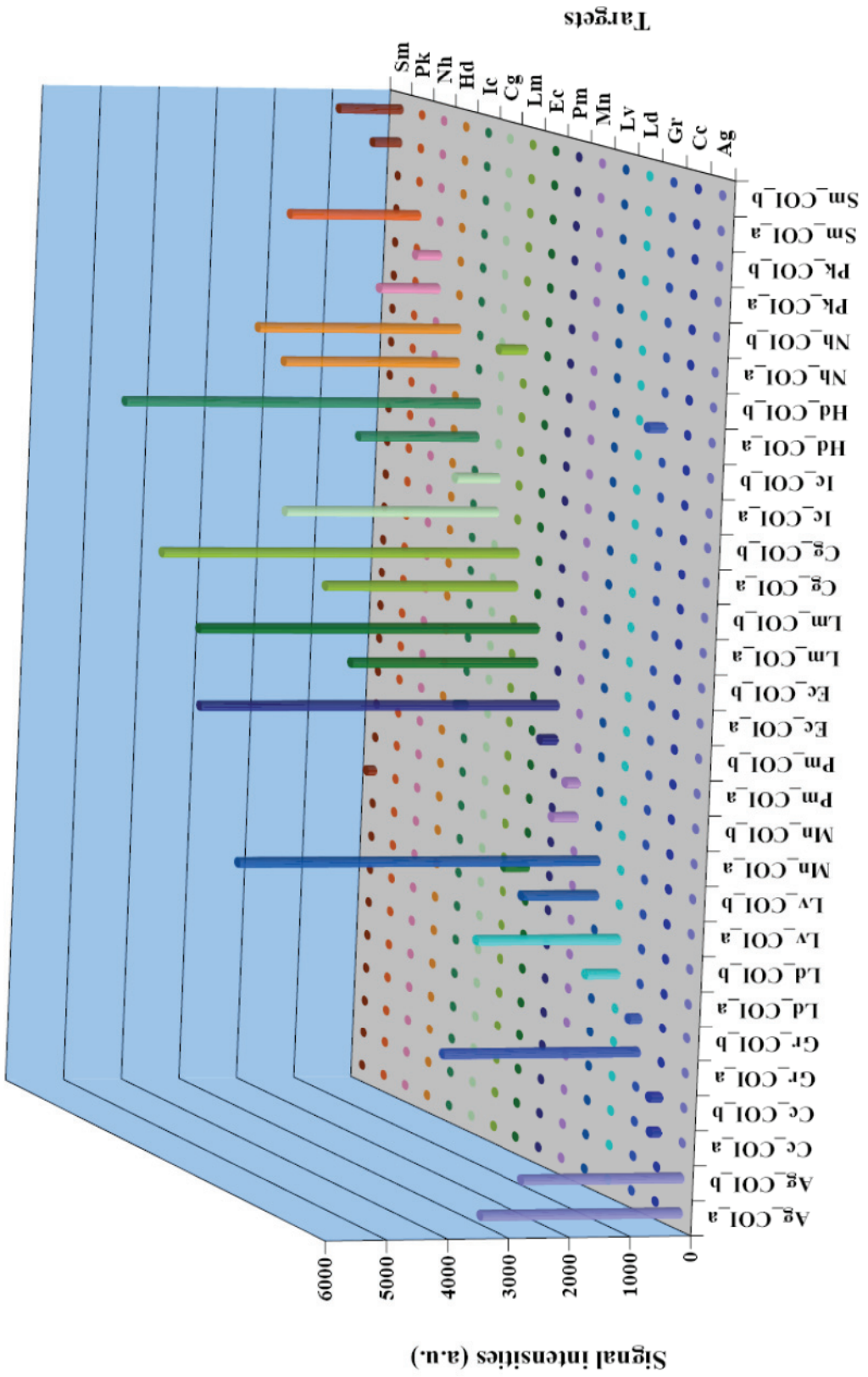


Fig. 14: Signal pattern of all the 15 single target hybridizations to the COI probes.

4.7.2. Two target hybridizations:

A total of 65 captures that gave good signals in the single target hybridizations of both 16S and COI genes are chosen for the combined invertebrate microarray. The captures that are not included in the invertebrate chip are marked with an (*) in Table 15 and Table 16.

Both 16S and COI genes from all the 15 target species except *C. crangon* were amplified in a multiplex PCR with two sets of primers (end labelled reverse primers) as shown in the Fig. 15. As there are two genes amplified in the same solution 20 nM is added to the array instead of 10 nM as in the case of single target hybridizations.

As all the hybridization parameters are checked and the signal pattern of every single target is known, to reduce the number of experiments and time two non-interfering targets were hybridized together on to the microarray each in a concentration of 20 nM. Each experiment was repeated thrice and the final signal intensities are the mean values of three replicate experiments. The intensities shown in the Fig. 15 are those that are two times higher than the negative control. The black bars in the diagram are the 16S signals and the white are COI. Striped bars are the false positive signals. The overall signal patterns of all the single targets along with the values are shown in Table 17.

At least one probe for each of the target gives signal except the 16S probes of *C. crangon* and *G. rhomboides*. For some of the targets 16S probes show very high signal and for some COI probes. So in order to identify a species accurately it is advantageous to have probes from both the genes as even if one of the sets shows low signals, the other give high signals.

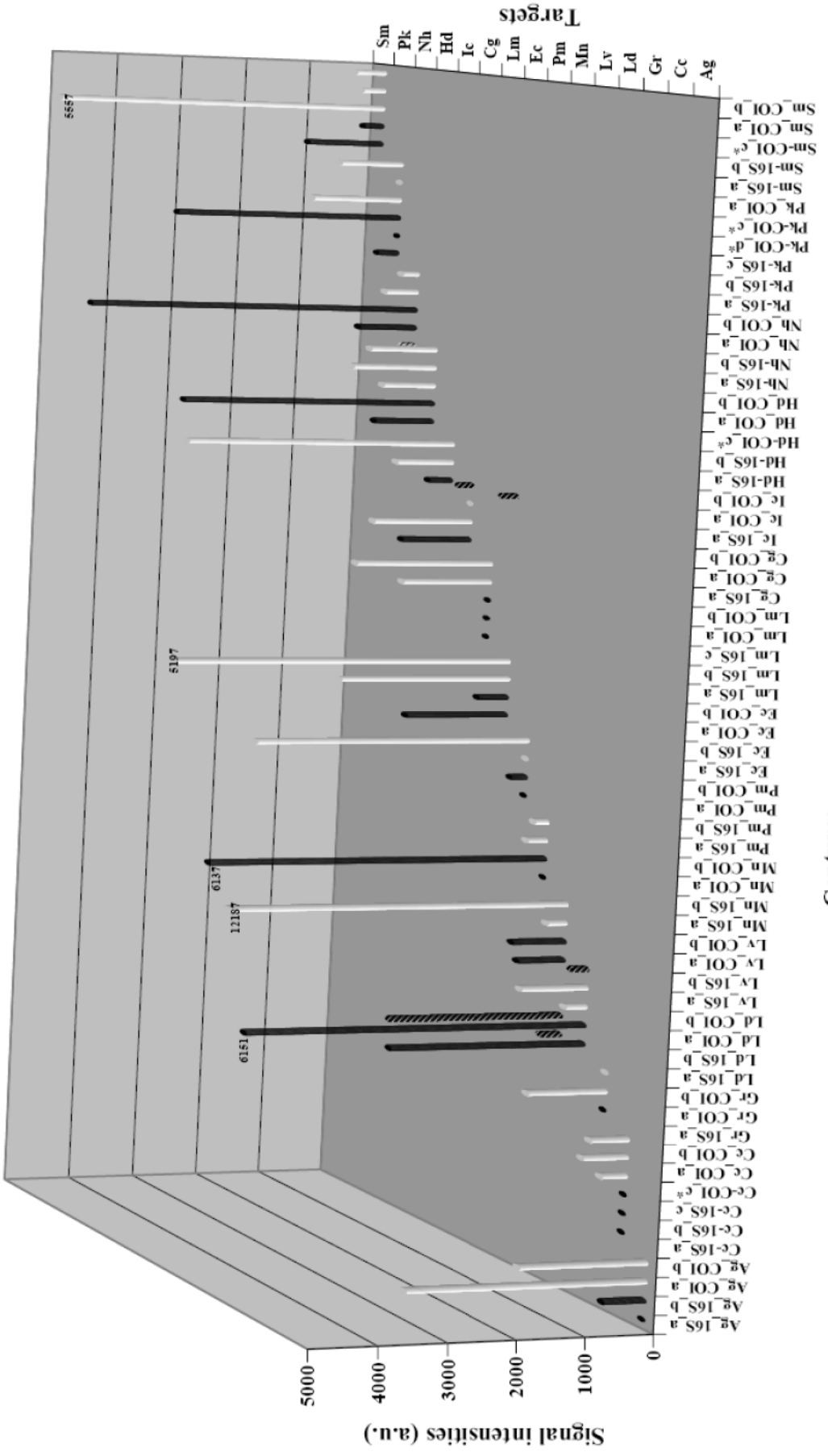


Fig. 15: Signal intensities from the 16S and COI probes present on the array when the 15 target species are added in two target hybridizations.

An overall summary of the two target hybridizations conducted on the invertebrate microarray can be found in the Table 17.

Table 17: Overall result of the hybridizations of the 15 targets on the invertebrate chip.

Green blocks: True positive signals; Blue blocks: True negative signals; Red: False positive signals.

Name	Ag	Cc	Gr	Ld	Lv	Mn	Pm	Ec	Lm	Cg	Ic	Hd	Nh	Pk	Sm
Ag_16S_a															
Ag_16S_b	614														
Ag_COI_a	3512														
Ag_COI_b	1875														
Cc-16S_a															
Cc-16S_b															
Cc-16S_c															
Cc-COI_c*		371													
Cc_COI_a		658													
Cc_COI_b		558													
Gr_16S_a															
Gr_COI_a			1163												
Gr_COI_b															
Ld_16S_a				2880	277										
Ld_16S_b				6151	2559								165		
Ld_COI_a				308											
Ld_COI_b				983											
Lv_16S_a				226	695										
Lv_16S_b					788										
Lv_COI_a					283										
Lv_COI_b					12187										
Mn_16S_a													175		
Mn_16S_b						6137									
Mn_COI_a						290									
Mn_COI_b						191									
Pm_16S_a															
Pm_16S_b							226								
Pm_COI_a															
Pm_COI_b							4077								
Ec_16S_a								1531							
Ec_16S_b								442							
Ec_COI_a								2545							
Ec_COI_b								5197							
Lm_16S_a															
Lm_16S_b															
Lm_16S_c															
Lm_COI_a									1347						
Lm_COI_b									2065						
Cg_16S_a										1049					
Cg_COI_a										1501					
Cg_COI_b															
Ic_16S_a								201		197	343				
Ic_COI_a											857				
Ic_COI_b											4044				
Hd-16S_a												891			
Hd-16S_b												3824			
Hd-COI_c*												792			
Hd_COI_a												1265			
Hd_COI_b												1010	176		
Nh-16S_a													879		
Nh-16S_b	145												4988		
Nh_COI_a													487		
Nh_COI_b	222												241		
Pk-16S_a														305	
Pk-16S_b															
Pk-16S_c														3440	
Pk-COI_d*														1356	
Pk-COI_e*															
Pk_COI_a														951	
Sm-16S_a															1152
Sm-16S_b															290
Sm-COI_c*															5557
Sm_COI_a															335
Sm_COI_b															451

4.7.3. Position effect of the oligonucleotides

Though the potential captures were thoroughly checked for specificity and the mismatch criteria *In silico* before spotting on the microarray, there is a great variation observed in the signal intensities after hybridization of different probes designed for the same target. The probes having completely similar characters Insilco were behaving differently on the microarray. In order to check if the position of the binding of the probe to the fluorescent labelled target i.e. the distance between the end label and the probe binding site is related to the signal intensity, the following analysis was done calculating the distance between the label and probe and the signal intensity of the corresponding hybridization separately from the results of the single target hybridizations of both 16S and COI.

As seen in Fig. 16 and Fig. 17 both 16S and COI results have shown a direct relation between the distance of the fluorescent label to the probe binding site and the signal intensity. The lower the distance the higher the signal and the farther the binding site the signal comparatively got weaker.

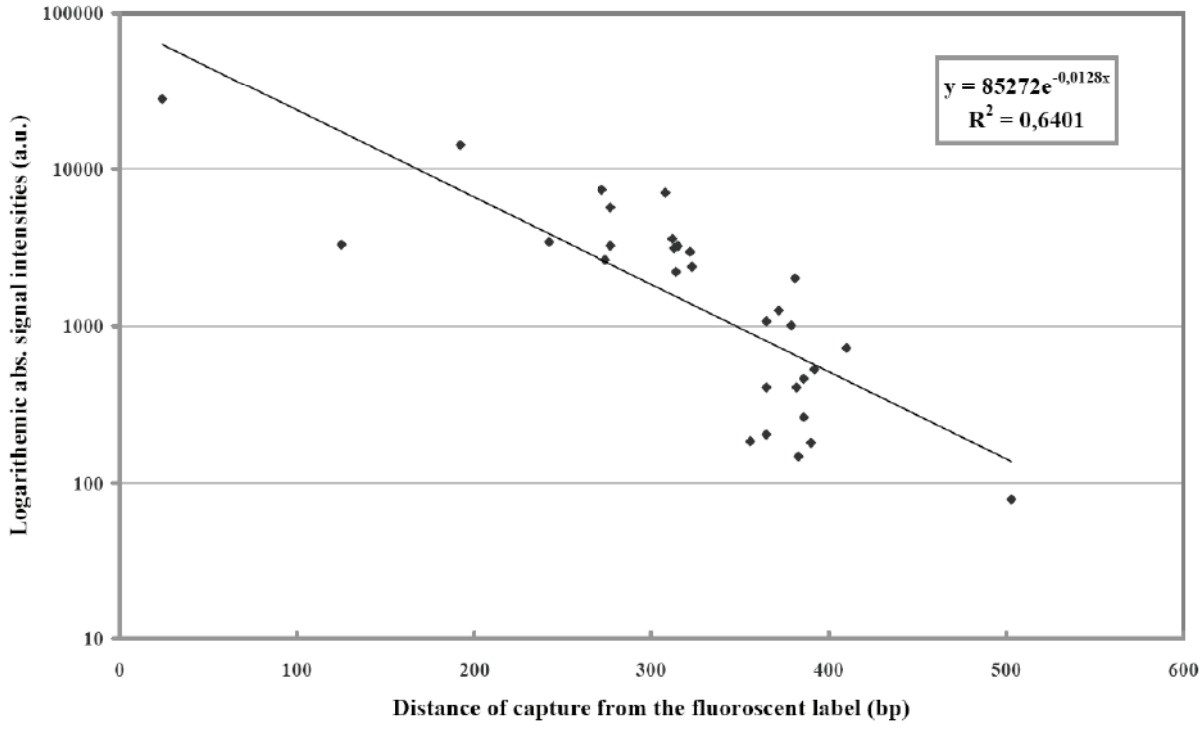


Fig. 16: Effect of the distance between 16S capture binding sites and the fluorescent label on a target.

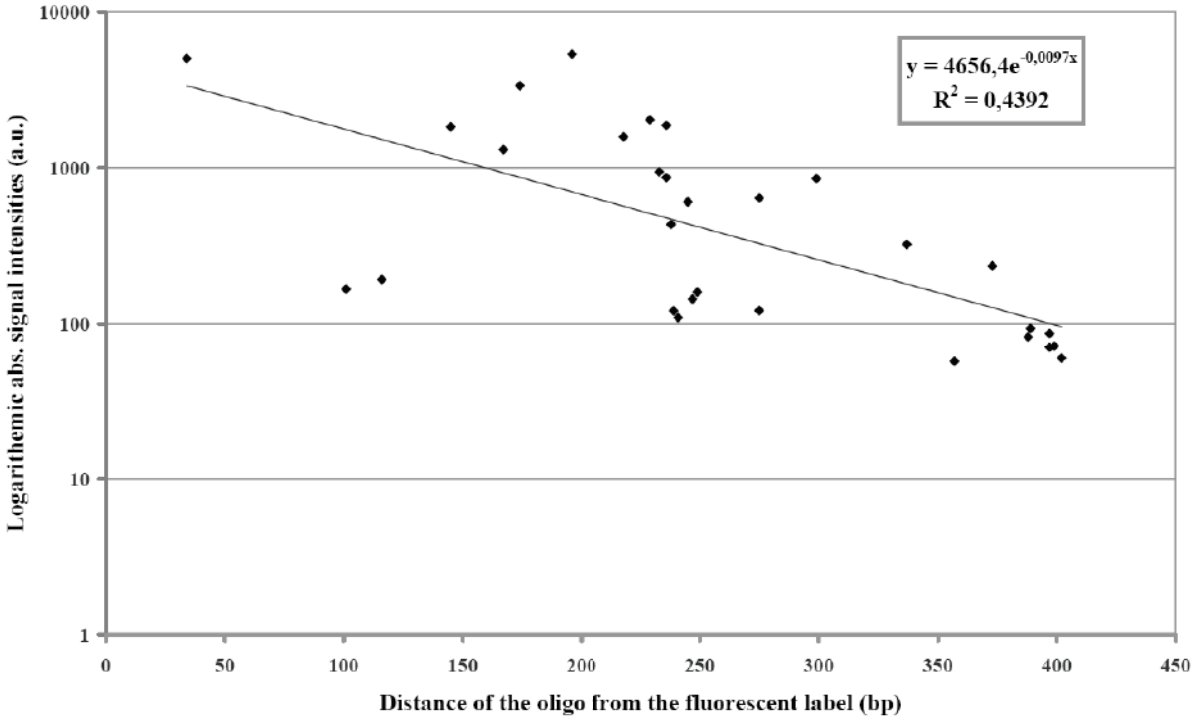


Fig. 17: Effect of the distance between COI capture binding site and the fluorescent label on the target.

4.7.4. Non-target hybridizations

Four different polychaete species (Table 18) from the North Sea which are very closely related to the target annelids on the microarray are collected and tested on the chip to check the specificity of the captures. Both 16S and COI genes are amplified from the DNA of the non-targets and 10 nM of each gene is hybridized on the microarray. Each one of the non-targets is hybridized separately and also a mixture of all the four is added in one experiment (Fig. 18). Two of the non-targets gave a very weak signal with 16S capture of polychaete *N. hombergii*, the third non-target gave a signal with the 16S capture of *L. depurator*. Mixture of all the four added in the concentrations of 10 nM together gave signals from the captures shown in the Fig. 18. The signals obtained are very weak and as the signal intensity is very low it can be neglected and the chip can be considered to be very specific to the selected 15 species of invertebrates.

Table 18: Non-target polychaetes that are used to test the specificity of the chip.

Species Name	**Family**	**Order**
Chaetozone setosa	Cirratulidae	Spionida
Magelona johnstoni	Magelonidae	Spionida
Scoloplos armiger	Orbiniidae	Orbiniida
Lanice conchilega	Terebellidae	Terebellida

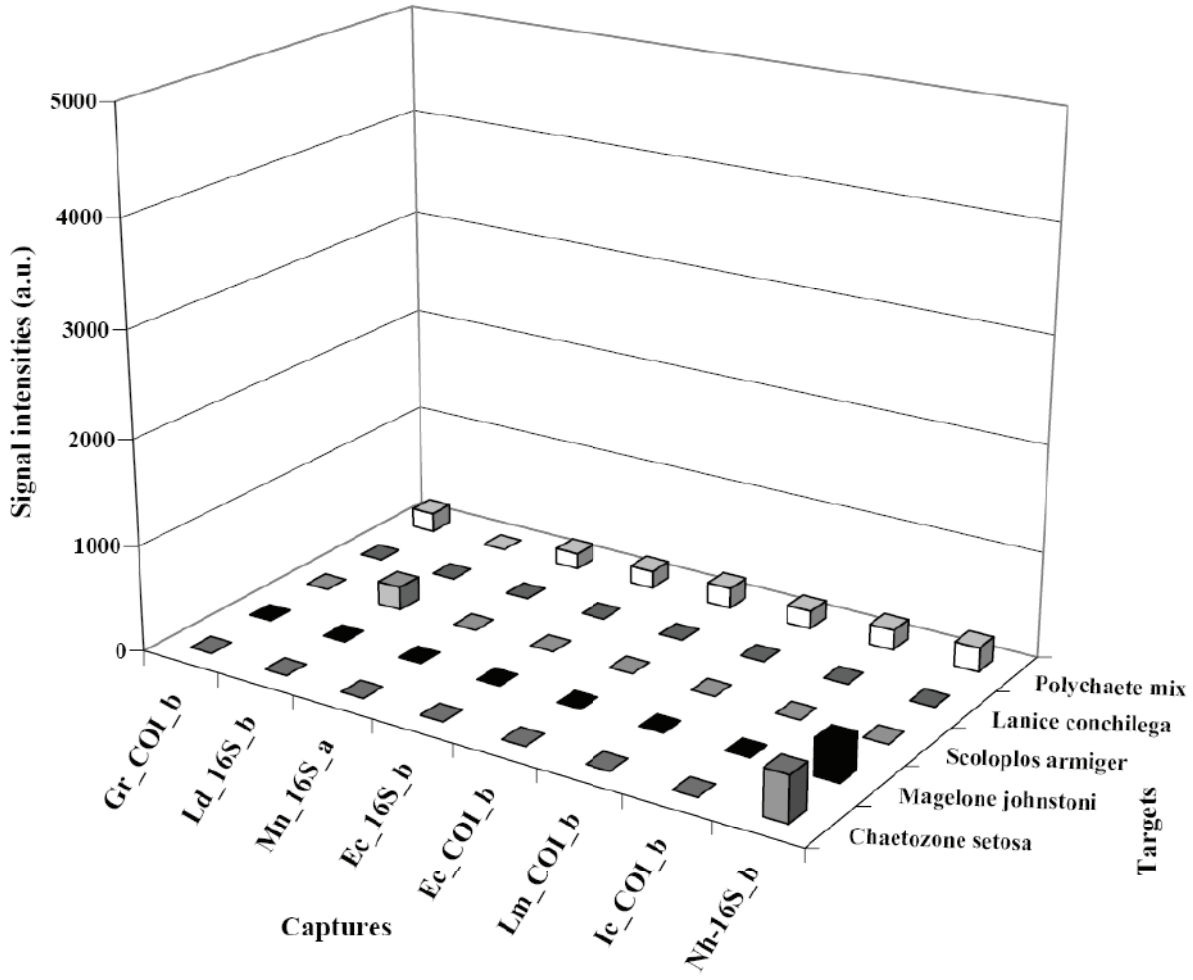


Fig. 18: Signal pattern of the non-target polychaetes when hybridized on the array individually and also a mixture of different kind of DNA.

4.8. Sensitivity test of the microarray

4.8.1. Two target hybridizations

Two target hybridizations were performed to check the influence of two closely related targets upon each other when hybridized together on the same chip and also the effect on the signal intensities of the probes of one target when the other target is present in different concentrations. For this purpose three sets of targets were selected from the three phyla used in the work. *N. hombergii*, *S. mathildae* (Annelida), *G. rhomboides*, *P. marmoratus* (Crustaceans) and *E. cirrhosa*, *C. gallina* (Molluscs). The first two sets are selected according to the criteria that they are closely related but did not show any common cross hybridization in the single target experiments. It was difficult to choose non interfering species as three out of four mollusc species give a cross hybridization with the 16S probe specific for the fourth

target. So two targets that gave signals with both 16S and COI probes during single target hybridizations were selected.

In all the cases each target is amplified with both the 16S and COI primers together resulting in two PCR products. For each set of the targets two kinds of experiments were performed. Target mix and a multiplex PCR. In the target mix the concentration of all the PCR products amplified individually is measured and the mixes are prepared according to the desired ratios just before the hybridization. In the multiplex PCR both the targets are added as templates in the same reaction in the desired ratios and the amplified product is purified and directly used for the hybridization. The hybridization results of the multiplex are checked to find any correlation between the concentration of the template DNA and the signal intensities after the hybridization. And these results are in turn compared to the target mix experiments to check the accuracy and reliability of the multiplex PCR.

The first experiment was conducted with varying mixtures of labelled amplicons of *N. hombergii* and *S. mathildae* (Fig. 19). Five different combinations of the two targets were prepared and hybridized on to the microarray. While keeping the concentration of the one target constant at an average of 10 nM the second target is mixed in increasing amounts and the same is repeated with the second target. Bordered captures in the figure denote the specific probes for the targets added and the rest of the captures are those showing cross hybridization with either of the two target DNA. All the signals shown in the figure are analysed with a threshold of two times the intensity of the negative control.

A significant pattern of the increase in the signal intensity with the increase in the DNA concentration can be observed in both the cases. The rapid increase in the amount of the closely related second target did not effect the signal of the first very much and vice versa except one *S. mathildae* capture Sm_16S_b where the actual signal itself is quite low compared to the other captures and there is no signal seen from the capture when the *N. hombergii* target is added in the highest concentration. Out of the seven captures showing false positive signals all of them except two seem to be because of the high concentration of the *N. hombergii* target DNA and the rest are showing high signals when the *S. mathildae* is added in high amounts.

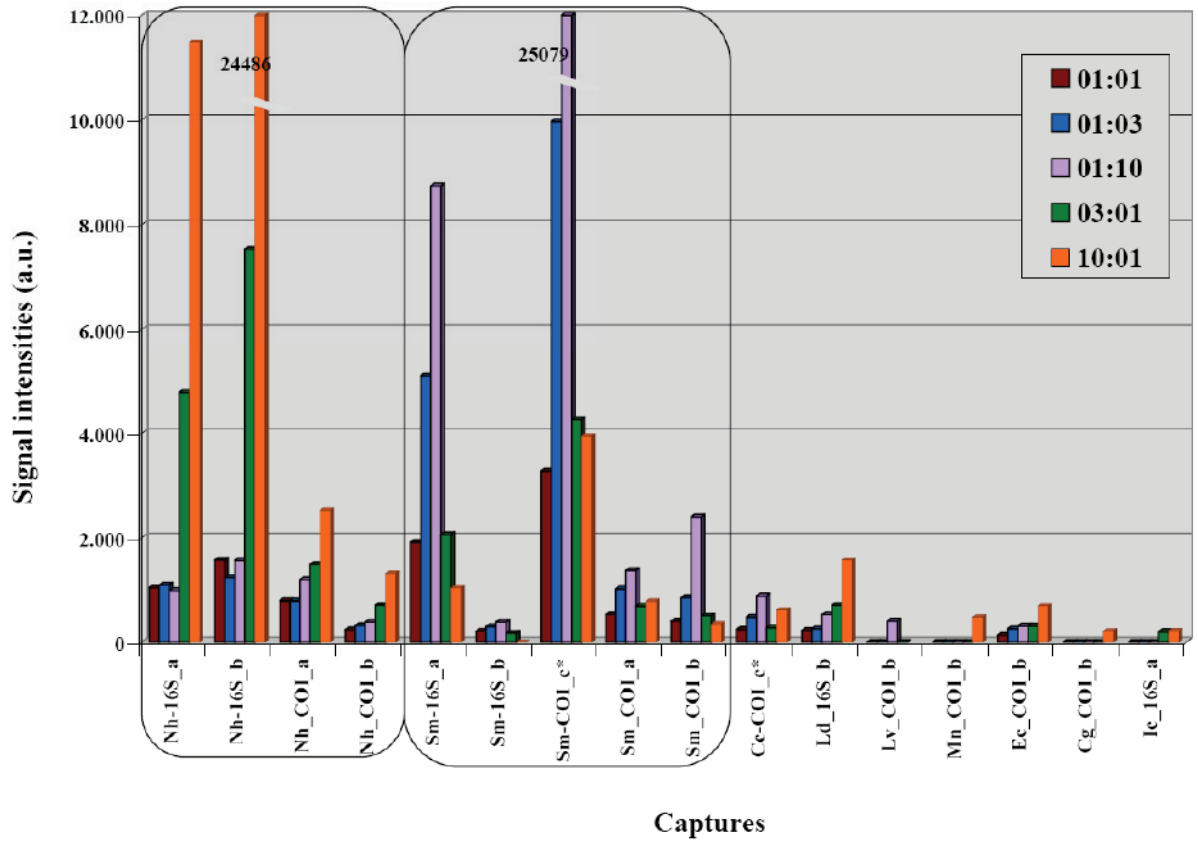


Fig. 19: Signal pattern when the target mix of *N. hombergii* and *S. mathildae* at different concentrations is hybridized in the ratio of *N. hombergii* : *S. mathildae*.

Similar experiments were conducted with a multiplex PCR where both the genes and both are targets are amplified in the same reaction. The two target templates are added to the PCR reaction in different concentrations as mentioned in the above experiment. Keeping the concentration of one of the target constant the other one is increased significantly. Initially the concentration of the DNA extracts is measured and as both of them are found out to be almost the same the first reaction is made with equal amounts of the template (1 µl each) the other combinations are made according to the ratios. The resulting labelled PCR amplicons are purified and are hybridized on the microarray at a concentration of 40 nM.

As seen in the Fig. 20 the absolute signal intensities from these experiments are very low compared to the target mix results. But a gradual increase and decrease patterns are observed from the captures of both the target. There is no significant increase in the signal observed with the increase in the concentration of the target as expected and also as seen in the above experiment. Moreover a decrease in the signal of the *N. hombergii* probes is observed when *S. mathildae* concentration is increasing and there is a gradual decrease in the *S. mathildae* signal with the increase in the *N. hombergii* concentration. Two new captures showed false

positive signals compared to the target mix experiment. The signals from the cross hybridizations are however very low compared to the specific signals.

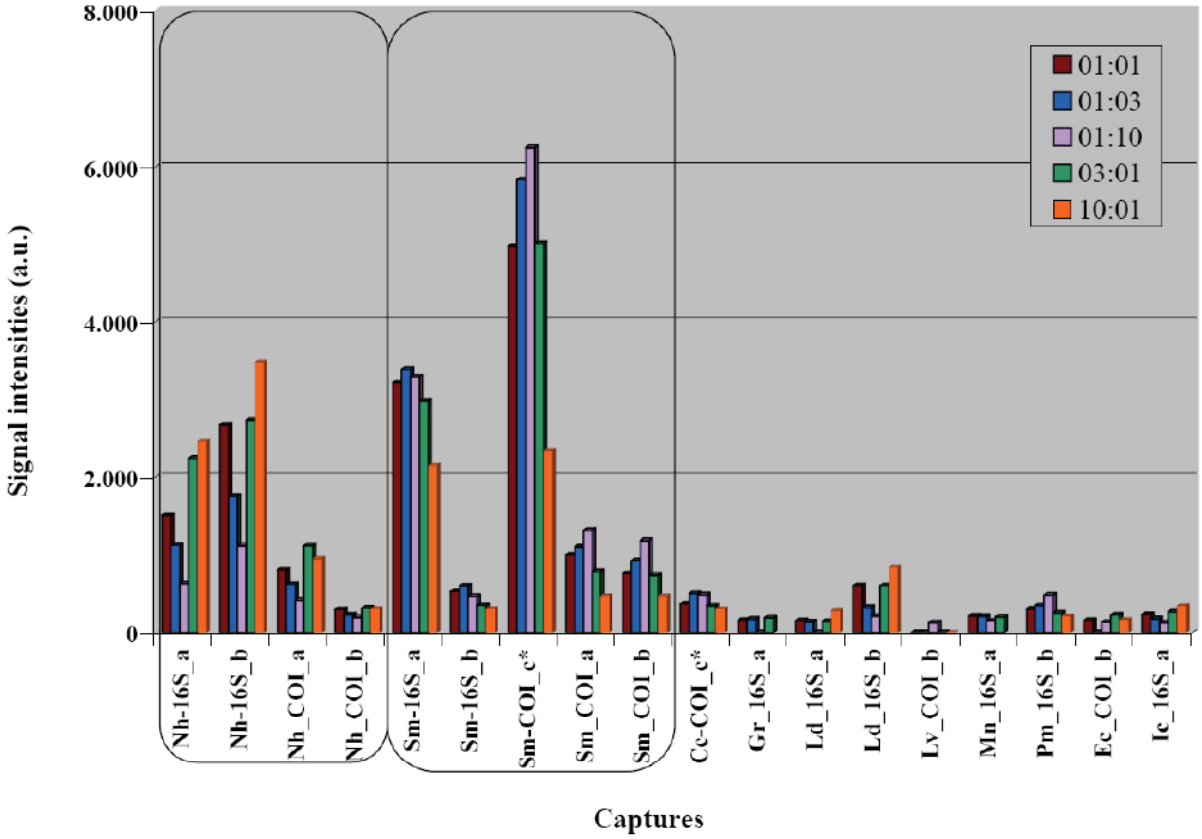


Fig. 20: Signal pattern from the hybridization of the *N. hombergii* and *S. mathildae* on the microarray at the mentioned concentrations in the ratio *N. hombergii*:*S. mathildae*.

The same experiments are also performed with the other two sets of closely related species to check if they show any influence upon each other. Fig. 21 shows the signal pattern when *E. cirrhosa* and *C. gallina* are hybridized together at different proportions. Both *E. cirrhosa* and *C. gallina* captures showed high signal intensities compared to the non-specific captures. The bordered captures are the specific ones, showing a pattern of increased intensity with the increase in the concentration and there is no significant influence of the second target seen except one *C. gallina* capture (Cg_COI_b) where the signal is totally lost when the *E. cirrhosa* target is added 3 times and 10 times higher. The common cross-hybridizing capture Ic_16S_a for both the targets produced a very high signal almost equal to the true positive and the highest is seen when the *E. cirrhosa* target is added in the highest concentration. The signal is completely lost when *C. gallina* is added in the highest amount. The cross hybridizations from three captures of *M. norvegica* and one probe of *L. vernalis* show high signals when the target *E. cirrhosa* is added in highest amounts. *L. mediterraneum* and *G.*

rhomboides probes show signals only when the *C. gallina* target is present in the hybridizations.

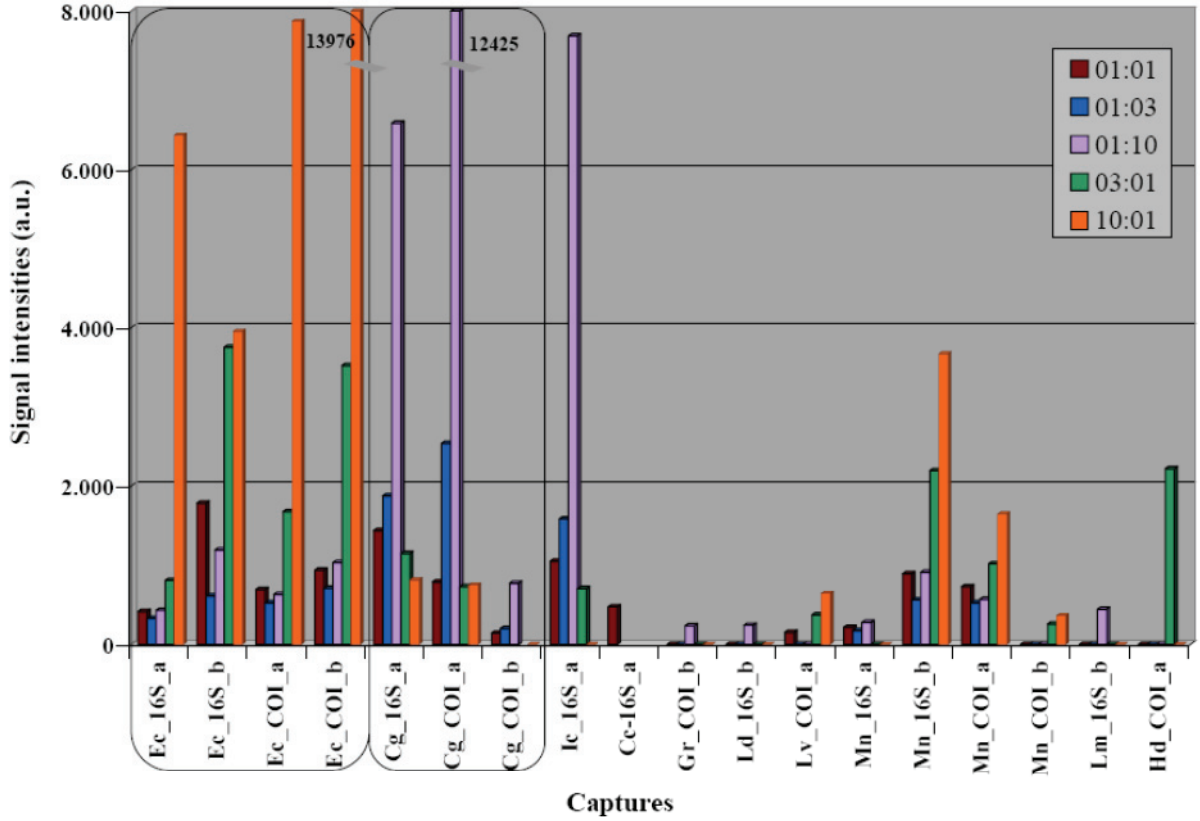


Fig. 21: Signal patterns from the mixture of PCR amplicons of targets *E. cirrhosa* and *C. gallina* hybridized together on the microarray at different concentrations in the ratio *E. cirrhosa*:*C. gallina*.

As seen in the case of *N. hombergii* and *S. mathildae* the multiplex approach of amplifying two targets at different concentrations in the same reaction and hybridizing on the microarray showed very low signal intensities from the targets *E. cirrhosa* and *C. gallina* when compared to the target mix results. A gradual decrease and increase in the signals of one target when the corresponding target is increasing or decreasing is also seen as in the case of *N. hombergii* and *S. mathildae*. Strangely the signal from the specific *E. cirrhosa* signal is very low when the target is added in the highest concentration. In case of *C. gallina* two probes did not show any signals when the other target *E. cirrhosa* is added in higher concentrations. The cross hybridization signals are comparable to those from the target mix results. All the captures of *M. norvegica* showing cross-reaction when the target mix is hybridized are not seen in this experiment

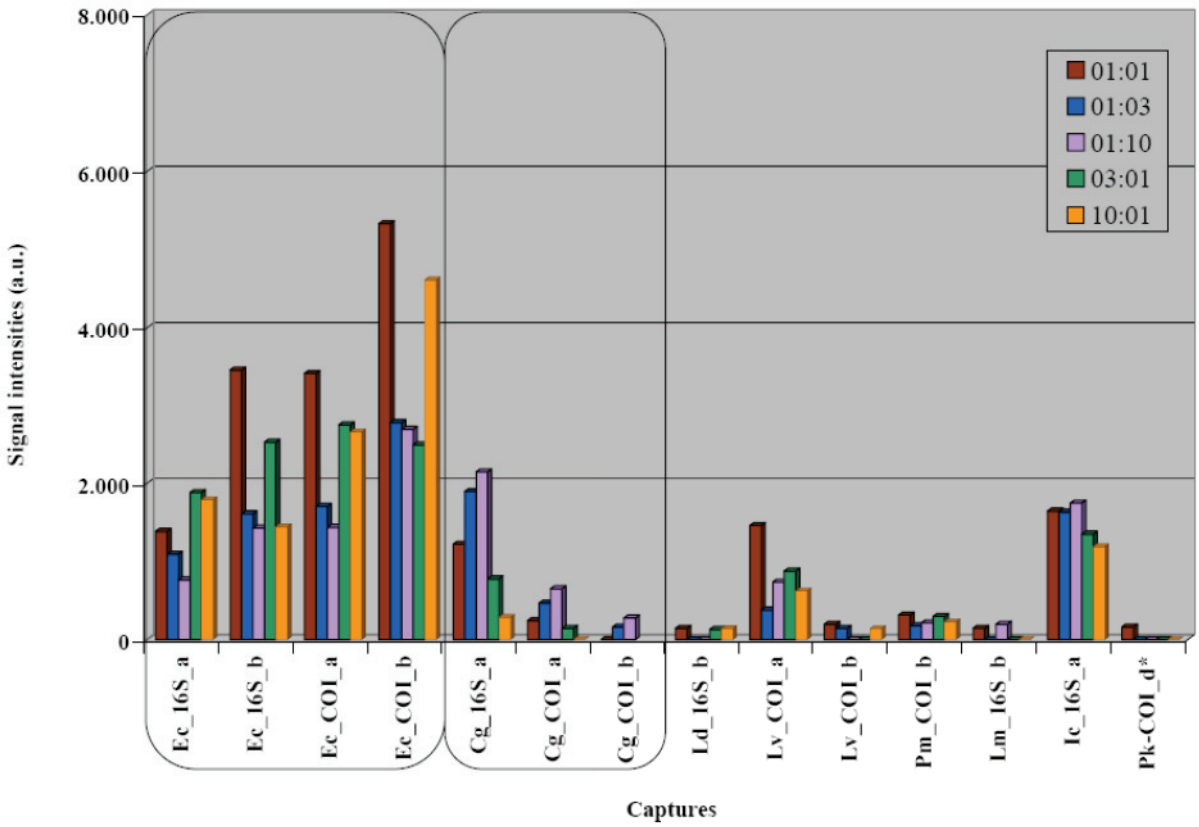


Fig. 22: Signal patterns from the multiplex PCR product of targets *E. cirrhosa* and *C. gallina* hybridized on the microarray at different concentrations.

The last set of the two target hybridizations is the crustaceans *G. rhomboides* and *P. marmoratus*. Similar to the other two sets the target mix hybridizations showed good signals from all the specific probes corresponding to the increase in the concentration of the target. The signals of one target are not influenced when the other target is increased to higher concentrations. The cross-reaction from the *H. diversicolor* and *M. norvegica* captures is high when the concentration of the target *G. rhomboides* is high. There is just one signal seen from the capture *C. gallina* when the concentration of *P. marmoratus* is the highest. Only one 16S probes is added to the microarray and even though the probe is thoroughly checked Insilco before spotting on the microarray. It gave very weak signals when the target is present in lower concentrations and no signals at all when the concentration is increased.

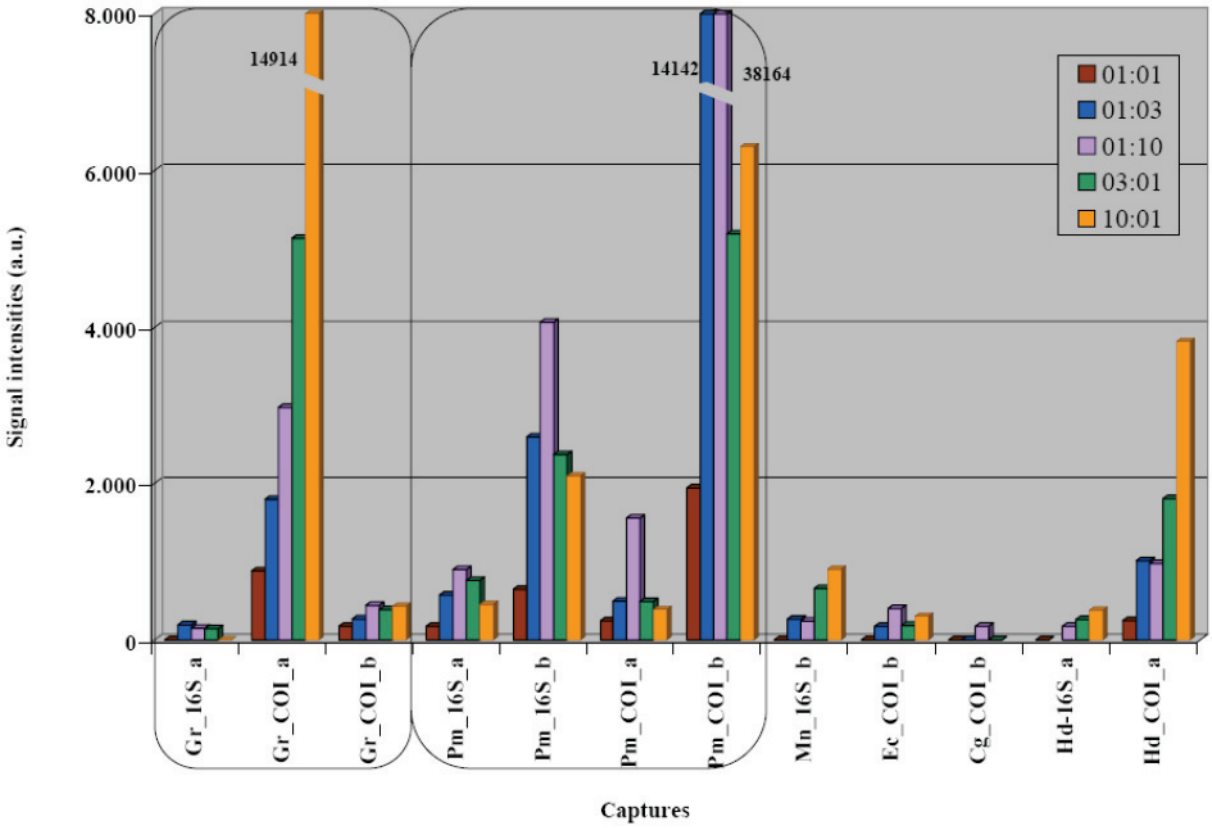


Fig. 23: Signal pattern after the hybridization of the target mix *G. rhomboides* and *P. marmoratus*. The concentrations of the targets are shown in the legend with *G. rhomboides*: *P. marmoratus* ratios.

The results from the multiplex hybridization of *G. rhomboides* and *P. marmoratus* show good pattern of the signals corresponding to the respective concentrations but as seen in the other two sets that overall intensity of the signal is too low when compared to the target mix experiment. The cross hybridizations pattern is also comparable to the target mix. The captures from *H. diversicolor* and *M. norvegica* showed high signals when the *G. rhomboides* target is added in higher amounts. Ic_16S_a capture gave a signal only when the *G. rhomboides* target is high in concentration.

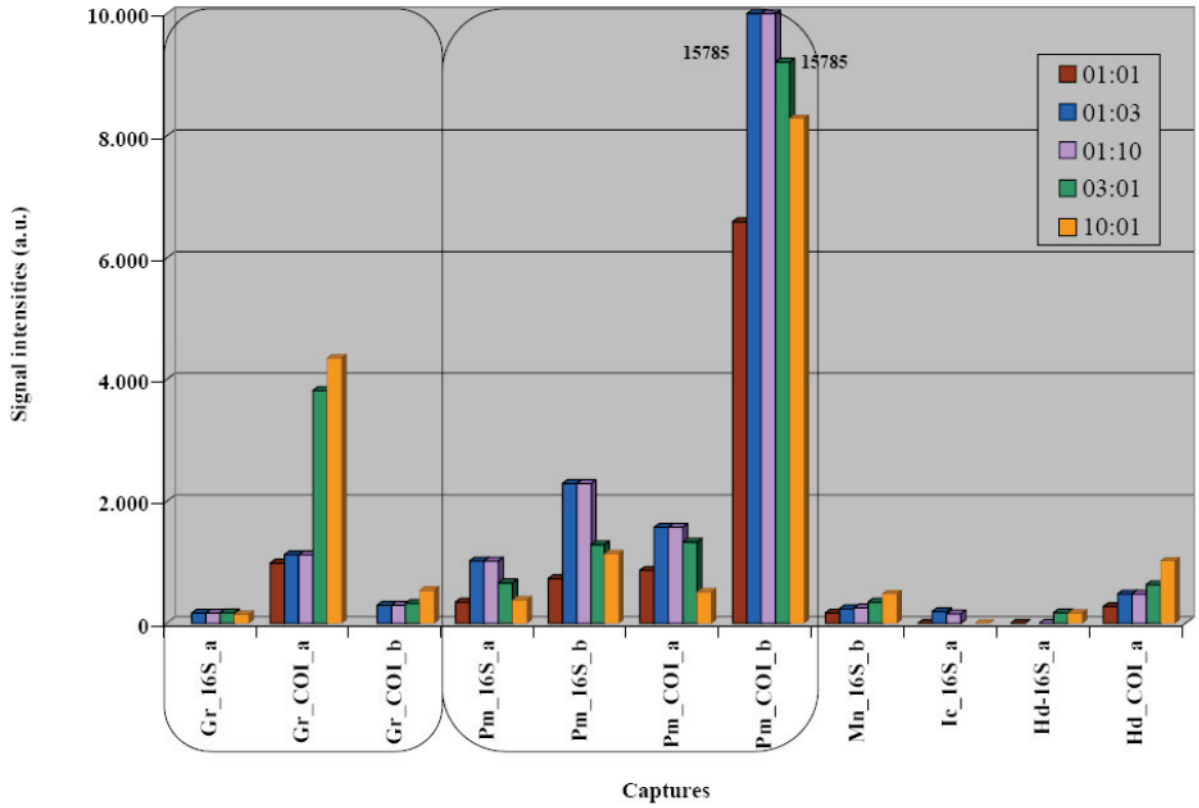


Fig. 24: Signal pattern from the hybridization of the *G. rhomboides* and *P. marmoratus* multiplex targets at different concentrations. The targets are added in the ratios as shown in the legend (GR: PM)

4.8.2. Three target hybridizations

To check the number of targets the microarray can identify without any interference three species belonging to the same phyla are hybridized together in the microarray. In order to carry on the multiple target experiments three targets at different concentrations are amplified in a multiplex PCR with two sets of primers (16S and COI). The resulting PCR product is hybridized on to the chip to check if there is a correlation between the template concentration, the PCR amplification of different targets and the signal intensities.

Three species of annelids (*N. hombergii*, *S. mathildae* and *P. koreni*), three species of molluscs (*E. cirrhosa*, *C. gallina* and *I. coindetii*) and three of crustaceans (*G. rhomboides*, *P. marmoratus* and *L. vernalis*) are used in this experiment. In the first PCR all the targets are added in equal amounts (1:1:1) and the later experiments are carried on keeping one target at the minimum and adding other two targets in the highest concentrations. The concentration of the targets is measured before the PCR in the Nanodrop machine and the amount of the template in the PCR is decided based on the concentration of the template and also and the corresponding ratio in the mixtures.

The set of the polychaetes *N. hombergii*, *S. mathildae* and *P. koreni* are amplified in the required ratios and the signal pattern obtained after the hybridization is seen in the Fig. 25. The ratio of the templates is shown in the legend denoting *N. hombergii*: *S. mathildae*: *P. koreni*. The signals in the figure show that all the three targets are successfully amplified in the multiplex PCR and the signals from all the captures denote that both the genes are amplified. The bordered captures in the Fig. 25 are specific for the targets and the rest of them are captures showing false positive signals. Except one 16S and one COI captures for *P. koreni* all the other specific captures showed signals and they are corresponding to their template concentration in the PCR. Very good signals are seen from the 16S captures of the target *N. hombergii* while the COI signals are very weak. There is no signal seen from the COI captures of *N. hombergii* COI probes when the other two targets are present in higher amounts. Though two of the targets did not show any signals the result from the rest of the probes is comparable to the ratios of the targets. Highest *P. koreni* signal is seen when both *P. koreni* and *S. mathildae* are in higher amounts. In the third combination even though *P. koreni* is present in the highest amount along with *N. hombergii* the signal is reduced slightly compared to the second combination. Specific signals are seen from all the captures from *S. mathildae* and the signal is not lost even when the other two targets are present in higher amounts. Though a few cross hybridizations are seen, their signal strength is very weak compared to the specific signals.

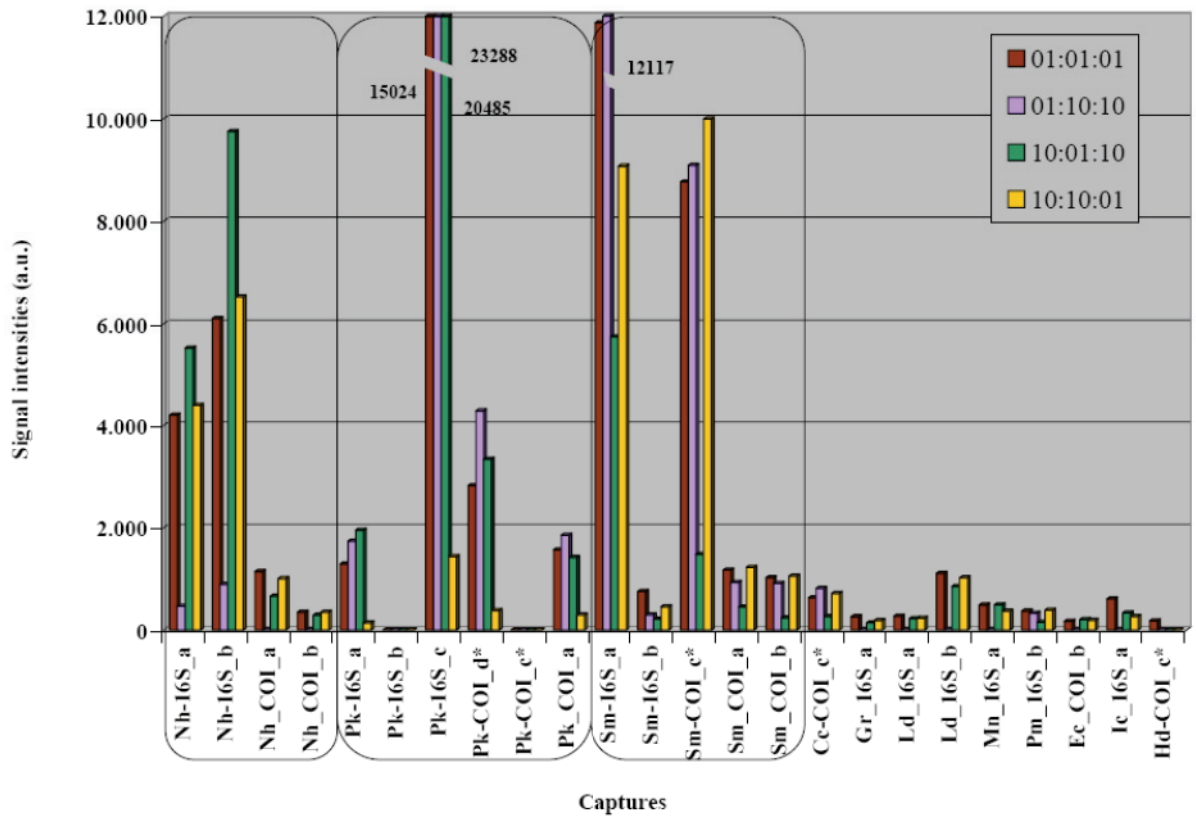


Fig. 25: Signal pattern of the hybridization of three closely related polychaete species when amplified in different combinations.

The second set of multiple targets consists of the molluscs (*E. cirrhosa*: *C. gallina*: *I. coindetii*) hybridized on to the microarray in after amplifying the targets in a multiplex PCR with varied proportions of the templates as shown in the legend (Fig. 26). The signal patterns show that only *E. cirrhosa* target is amplified well and the other two are amplified weakly as the signal intensity from the specific probes is very low compared to the signal target experiments. Nevertheless a specific pattern of the signals is seen corresponding to their ratios in the PCR reaction. All the four probes for the target EC gave good signals and the highest signal is observed when the target is present in 1 nM equally with the other two templates. But strangely the signal decreased when the concentration of the target is very high.

The overall signal from the captures for *C. gallina* is very low. The 16S probe signals corresponded to the concentration of the targets added to the PCR reaction. No signal is observed from the COI probes of *C. gallina* except a small one from Cg_COI_a from the reaction where the concentration of *C. gallina* is very high. Captures specific for *I. coindetii* gave good signals correlating with the template concentrations in the multiplex PCR reaction. Target *E. cirrhosa* gives a strong cross hybridization with the Ic_16S_a probe. So it is expected that the signal from Ic_16S_a will be very high but the signals obtained are quite

low indicating poor amplification of the *I. coindetii* target and also due to competition between the *I. coindetii* and *E. cirrhosa*, the latter could not bind to the captures in presence of the specific target. The cross hybridization of the Lv_COI_a with *E. cirrhosa* target is very high almost equal to the specific signals. Other cross hybridizations is very low and almost negligible.

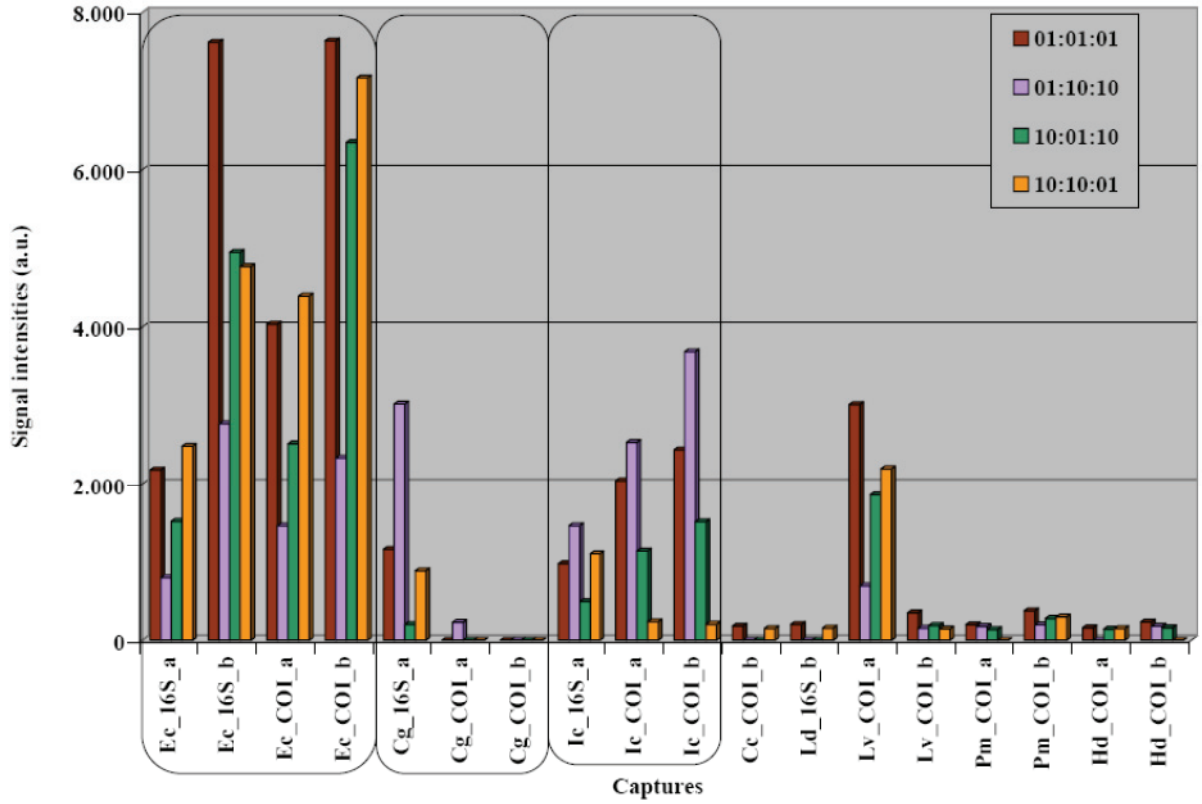


Fig. 26: Signal pattern from the hybridization of multiple mollusc targets in different concentration as shown in the legend in the combination of (*E. cirrhosa*: *C. gallina*: *I. coindetii*)

The third set of organisms; crustaceans are amplified together in a multiplex PCR in the required combinations of the templates. All the three targets seem to be amplified as desired and good signals are seen from all the specific probes. There is no signal seen from the hybridization of the second combination where *G. rhomboides* DNA is very low and both the other targets are in high amounts. The signals from the specific captures of other two targets are very strong and correspond to the concentration of the templates added in to the PCR. Very high cross hybridization signals are observed from the 16S captures of *L. depurator*, which is very closely related to the target *L. vernalis*. The other false positive signals are very low compared to the specific signals.

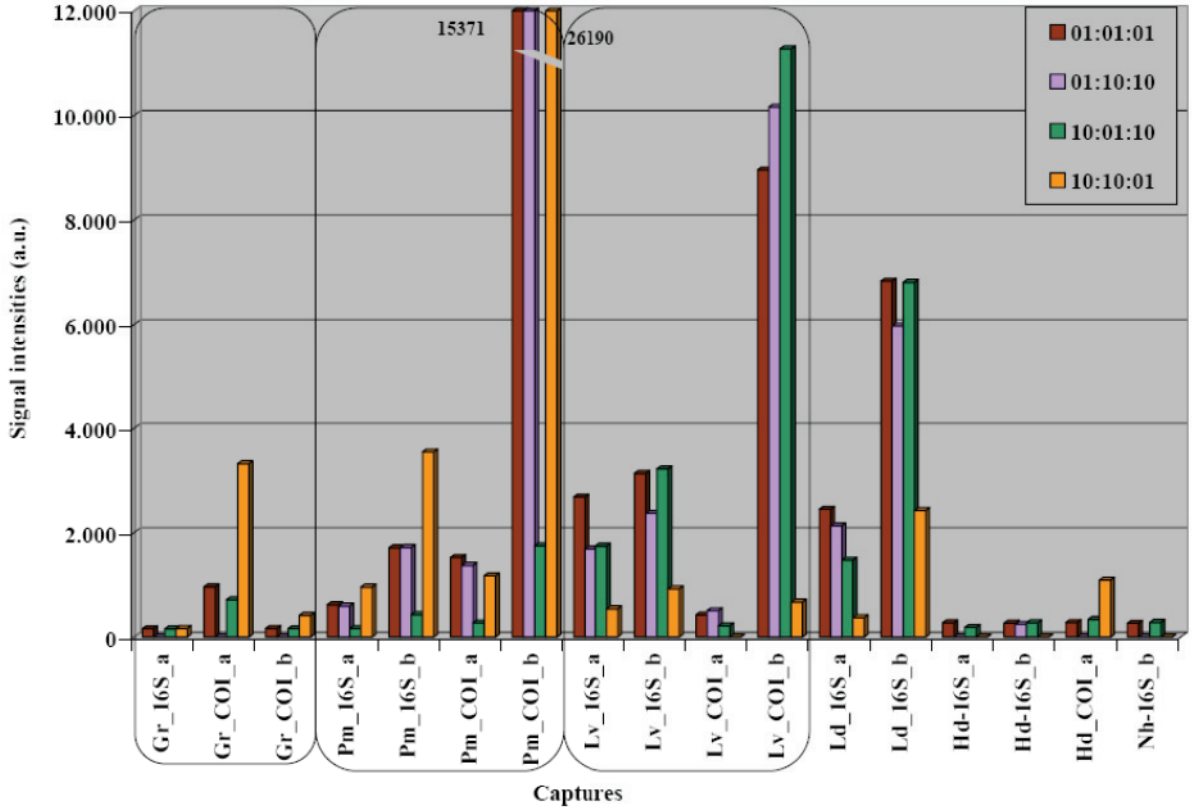


Fig. 27: Signal pattern from the multiplex PCR amplicons of three closely related crustaceans in the ratio *G. rhomboides*: *P. marmoratus*: *L. vernalis*.

4.8.3. Target detection in presence of background samples

The efficiency of the captures in detecting a single target in presence of large number of background samples was checked by hybridizing a mixture of environmental samples containing DNA from gut content of fishes and non-target polychaete samples collected from the Adriatic and the North Sea from where the targets are being collected. The influence of the presence of background samples on the signal intensity has been checked by hybridizing the target DNA along with different concentrations of the environmental samples. Fig. 28 shows the signal patterns of *N. hombergii* and *I. coindetii* when hybridized with different concentrations of the background samples and the signals gradually decreased with increasing concentrations of the background. When the background is 20 times higher than the target DNA the signal of *I. coindetii* is completely lost and *N. hombergii* can still be recognised with a weak signal.

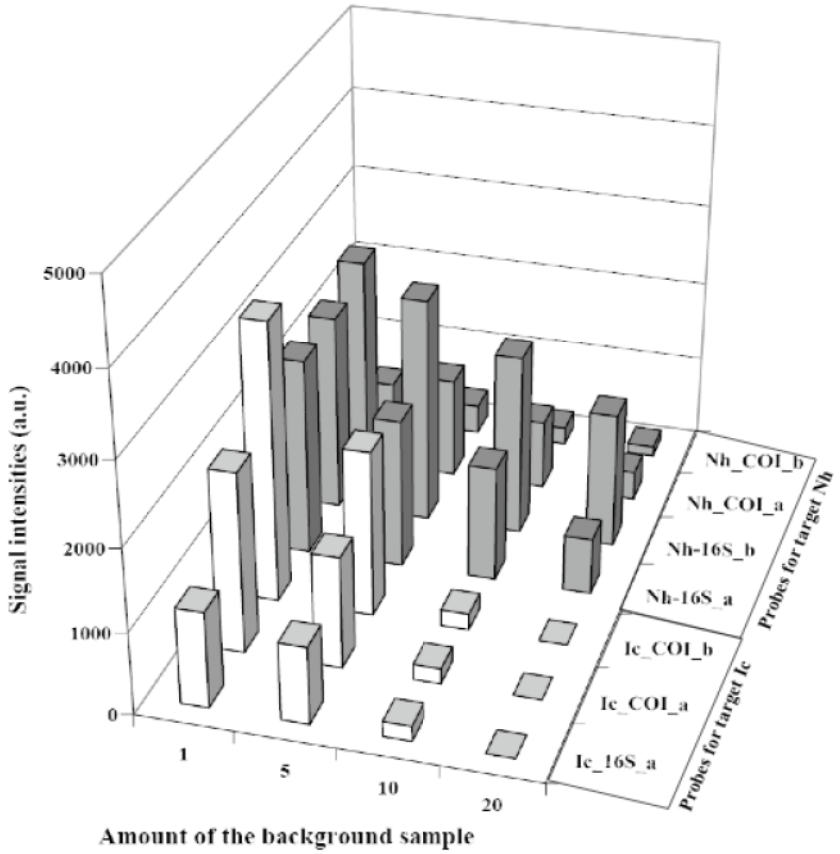


Fig. 28: Signal pattern of the targets *N. hombergii* and *I. coindetii* when hybridized with varying concentrations of the background samples.

4.9. Application of chip in ecological studies

4.9.1. Gut content analysis:

To check if the chip can be applied to environmental samples, gut of demersal fish *Trigla lucerna* is put to analysis. The gut content sample was collected, DNA extracted and both the 16S and COI genes were amplified with labelled reverse primers and hybridized on to the chip. G1 from Fig. 29 is collected from a single prominent material that appeared like a crab under microscope and G3 is the homogenised sample of random picks from the miscellaneous material present in the gut. G5 is a mixture of PCR amplicons from six different samples in the gut. All of samples gave a significant signal from probes specific to the crustacean *G. rhomboides*. Very weak signals are however seen from other probes specific for *M. norvegica*, *H. diversicolor* and *L. depurator*. The gut mix sample gave signals from many of the probes but the signals are very low to decipher the real presence of the target.

Verification of the result is done by sequencing the 16S and the COI gene of all the gut content samples and aligning them to the known sequences. The best fitting alignment can be seen in the appendix 1.1.

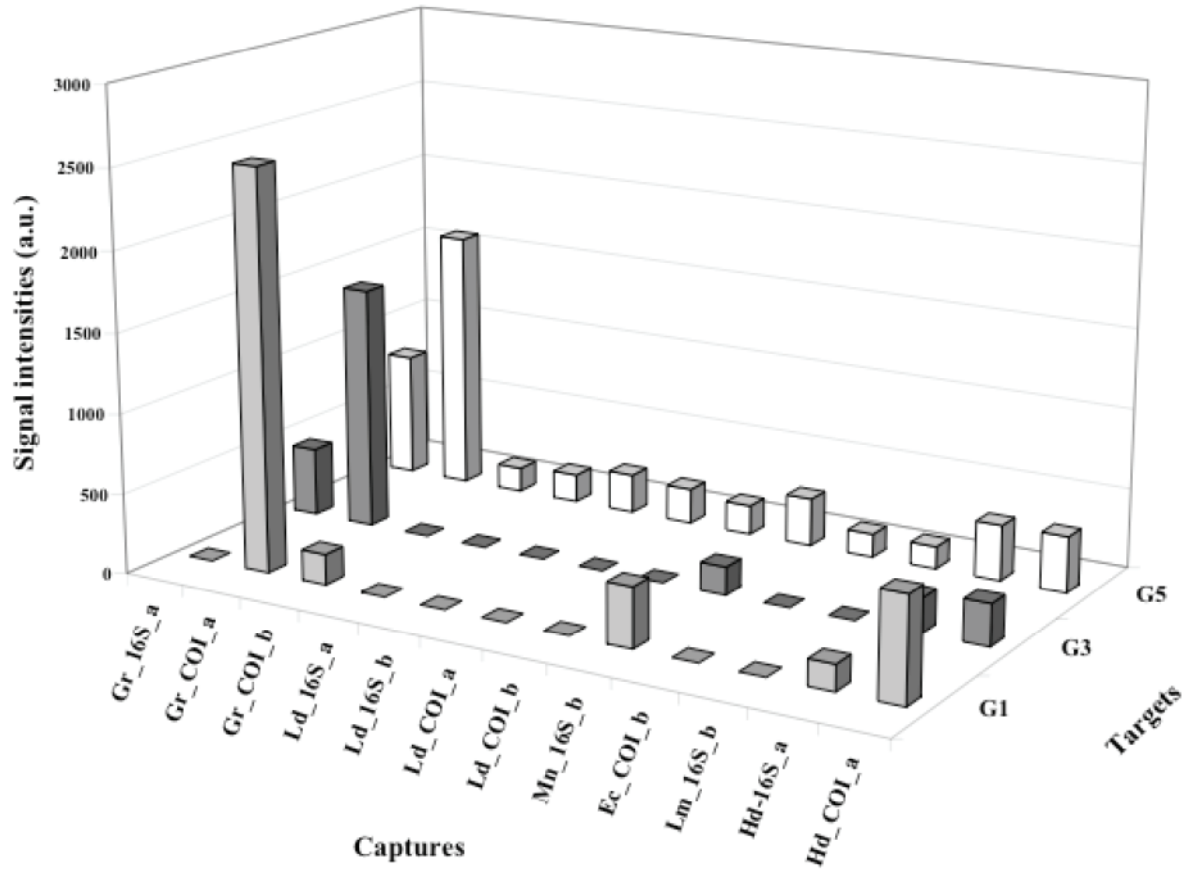


Fig. 29: Signal pattern of the gut content of *Trigla lucerna* hybridized on to the microarray. G1: sample collected from one position and non-homogenised, G3: homogenized sample of the random picks from different places of the gut, G5: mixture of all the gut content samples.

5. Discussion

Essentially, the development of a first prototype of an Invertebrate microarray has been accomplished in the course of this study, its function was successfully tested by identifying a few invertebrates found in the gut of a predator fish and for one of the critical steps, the DNA extraction from these animals, the FTA elute method was adapted which is already a standard in biomedical DNA-analytics. Nevertheless, some technical and biological details could not be worked out satisfactorily as for example the "simplified" multiplex PCR and sensitivity as well as specificity tests under practical conditions. The status reached in terms of these still existing limitations will be discussed in more detail in the following paragraphs.

5.1. Standard DNA extraction method

None of the ten DNA extraction methods comparatively tested in this study gave satisfactory results with all animals included and belonging to three different phyla. The most commonly used commercial kit (Qiagen) and the new technology based on FTA Elute (Whatman) are found to give best results if the amount and quality of the extracted DNA were taken as criteria (Table 12). Whereas no DNA could be obtained from individuals of the species *C. crangon* and *L. mediterraneum* with any of the ten methods tested, in all other cases DNA extraction was successful (Table 11). Out of the two classical methods phenol chloroform and Chelex the former one produced more DNA compared to the latter but the marker gene amplification from this DNA was not successful. Whether chemical components used during the extraction process may act as PCR inhibitors blocking the *Taq* polymerase or whether other reasons are responsible for this failure could not be investigated in the framework of this study. But these classical extraction methods commonly used in many molecular genetics studies of marine invertebrates (Chen and Yu 2000; Sinha 2001), are not considered as suitable since they are very time consuming, are based on tiresome protocols and the chemicals used are toxic or an potential health hazards. Majority of the DNA extraction kits summarized in Table 2 use proteinase K for cell lyses, silica membranes for purification or diverse other materials which makes them sophisticated and expensive (Rohland and Hofreiter 2007). Only the Qiagen and the Peqlab extraction kits and none of the others were able to produce any DNA from the samples used. Since all the samples used in this study are stored in ethanol, it could be that lysing the dehydrated tissue is difficult for some of the methods and not for others and in case of the polychaetes and molluscs the amount of tissue was possibly too less since their tissues are hard and leathery and the crustaceans could partly

not lysed at all because of their shells, but relevant is the comparison and that there are methods which worked. So, the Quigen kit and the Whatman FTA Elute system were found to be suited as standard methods which can overcome these drawbacks and can efficiently extract DNA from any kind of tissue. Since the DNA extraction and storage protocol is much more simple and robust in case of FTA Elute it was finally selected as method of choice. Using this method DNA extraction was successful with all samples (Table 11), it is easy to handle and the purification and elution protocol is simple and quick. The DNA extracted from the new FTA Elute cards is quantifiable as it is eluted into the solution unlike the standard FTA cards, but the concentration of DNA is lower compared to DNA extracted using the Qiagen kit. Hence more volume is used as template in the PCR reactions (Coyne et al. 2004) and there is no problem with degradation of the DNA since the FTA paper can be stored for long time and the DNA can be eluted from the paper when ever necessary which allows the analysis of the same sample many times. Since storage is performed in dryness at room temperature this method avoids both the high costs of freezer and activation of nucleases in water containing samples. Though the Qiagen kit produces more DNA FTA Elute is more preferable to use because the samples can be "extracted" during collection directly on the field without ethanol (Dobbs et al. 2002), the easy transport is very advantageous because of customary air transit problems (Crabbe 2003), for large collections storage at room temperature is cheap (Lin et al. 2000) robust and safe (Smith and Burgoyne 2004) it can be used on any kind of tissue with minimum amounts of tissue and above all, one of its greatest advantages is the extremely low cost per sample. All this makes FTA elute a promising standard technique for microarray based DNA-analysis of environmental studies.

5.2. "Simplified" multiplex PCR

As the invertebrate microarray consist of both types of probes deduced from 16S and COI genes, it is desirable to label both targets in the same PCR reaction to reduce cost and time. Such a multiplex PCR amplification procedure is broadly used and after optimization of the annealing temperature and the concentration of Mg^{2+} ions and dNTPs good results can be obtained (Henegariu et al. 1997). Both 16S and the COI genes have been successfully amplified from all the target species except the crustacean *Crangon crangon* (Fig. 9). COI DNA from this organism amplifies at 40 °C annealing temperature (Fig. 8) and not at 45 °C the optimized temperature for the multiplex reaction. (Fig. 9). Therefore target DNA from these two genes are amplified in separate reactions. With the addition of similar amounts of

the primers, dNTPs and Mg^{2+} ions an equal amplification of both the genes is shown in the gel picture in Fig. 9, but that is the situation if the PCR's are run in separate vials.

A true multiplex PCR can not be carried out, since the fragment size is identical in all targets and there is no analytical technique available to quantify the amount each gene is contributing during the PCR reaction. Therefore this type of PCR is called "simplified multiplex", since it can be assumed but it is not proven that the concentration of the individual target DNA is identical or at least similar or rather different.

This approach was also used to amplify both the genes from two, three and mixed templates for the sensitivity experiments of the array. All the chemicals and the conditions were kept the same as during the "simplified multiplex" PCR mentioned above except that the two and three different DNA templates were added in different concentrations but normalised in such a way that their final amount in the common single reaction was identical. Even the 16S amplicons which vary between different organisms because of the small gaps in their sequences (looped structure), can not be discriminated on the gel. Presuming the amplification to occur in the desired ratios, further assessment could have made at the level of the hybridization results. But the overall intensity of the signal was found to be very low (Fig. 19 and Fig. 20) indicating that the degree of amplification during the PCR was very low in general. One reason would be competition of the primers for binding to the template leading to less amounts of amplicons from each of the templates. In some cases when mixed templates are used in varied amounts, the final product did not seem to maintain the same ratios (Fig. 27). The reason could be the high number of amplification cycles used for labelling. Generally the most obvious variation in the amount of PCR products compared to the amount present at the starting phase occurs around 24 -28 cycles, since at 40 cycles a plateau state is reached which makes it impossible to estimate the amplification rate of different targets. Nevertheless all the targets used in the mixed PCR showed hybridization signals which means that all the templates were amplified successfully. So, for more quantitative results a calibration method is needed to estimate the PCR efficiency of each target, since the advantage of using identical PCR primers for all targets should not be given up. Such a method would be very beneficial for field applications since an efficient true multiplex PCR is very much necessary to amplify the diverse DNA present in the environmental samples.

5.3. Evaluation of the capture oligonucleotides

The success of any microarray experiment depends on the quality of the oligonucleotide captures present on it. The critical issue is that all the probes should exhibit similar

characteristics in terms of binding behaviour, steric hindrance and orientation, so that the final hybridization signal is comparable from all the probes. This greatly depends on the GC content corresponding to the melting temperature of the probes. One of the chip design internet services, CDIS (http://www.miconet.uni-bremen.de/) has the advantage of managing vast sequence data sets and generate probes according to the specified criteria along with the mismatch details to other targets or non-targets and direct links them to the possible cross hybridizing sequences present in the NCBI databank. This reduces the work necessary for checking the specificity of each probe separately first by BLASTn search and then experimentally.

Only a few sequences of the target species are newly obtained during this work (Table 13) compared with the total number of invertebrates present in the European Seas and with the number of them their sequences can already be found in the database (Table 14). Therefore the non target sequences may not represent the whole population of invertebrates. So, DNA from other organisms can produce false positive signals during hybridizations when environmental samples are analysed with the microarray. Therefore the reliability of the set of capture probes which was designed and used during this study was checked *in silico* against all sequences available from organisms of the three phyla. In an attempt to further reduce false positive results redundant probes were selected after optimizing the hybridization conditions and after checking the specificity with a few non target organisms.

Single nucleotide polymorphisms were not observed among the 16S and COI sequences of the species used in this work. But as the sequences used represent only an extreme small part of the individuals found in the environment, unspecific binding from other organisms or false negative signal from specific target sequences not belonging to the species covered by the microarray developed here are to be expected. To minimize this effect more than one probe is designed for each species and is not removed from the microarray even if its signal intensity was found to be very low.

The capture design from the short COI sequences showing a high degree of diversity was easy compared with 16S probes which could not be generated within the narrow range of length and Tm because of their restriction in terms of secondary structure and less sequence diversity. One of the 16S probe (Ic_16S_a) is even 45 bp long showing luckily a moderate Tm of 84 °C because of the GC content of 36% (Table 15). The oligonucleotide binding mainly depends on the hybridization temperature which is primarily dependent on the Tm. Often long captures operate under low stringent conditions (Guschin et al. 1997; Loy et al. 2002; Rudi et al. 2000; Southern et al. 1999; Wilson et al. 2002). To increase their sensitivity concurrently

fragmented target molecules (Guschin et al. 1997; Koizumi et al. 2002; Small et al. 2001) or short PCR fragments (Rudi et al. 2000; Wilson et al. 2002) are used for hybridization analyses. But this increases the danger of false positive signals. By avoiding the fragmentation of the target in combination with optimized hybridization conditions the redundant probes of both 16S and COI used in this study highly specific signal intensities were obtained (Fig. 14). Even though the probes are diverse in their criteria *in silico*, at optimized hybridization temperatures majority of them gave strong signals. But these true positive signals are significantly heterogeneous in their intensities (Fig. 15). These differences in the fluorescence signals from the probes of the same species can be clearly seen on the array picture (Fig. 30). This problem was commonly seen in many hybridization experiments (Kochzius et al. 2008; Peplies et al. 2003; Ronning et al. 2005; Warsen et al. 2004) and it may be overcome by a two color approach (Palmer et al. 2006) or only by adapting extreme methodical effort (Shi et al. 2006). The binding efficiencies of the probes on the microarray may be sequence dependant or effected by the steric hindrance and secondary structures (Southern et al. 1999) (Shchepinov et al. 1997). Another reason for the heterogeneous signals could be the position of label (POL) effect (Zhang et al. 2005), which describes the difference in the distance between the capture binding site of the target DNA molecule and the position of the fluorophore attached to that molecule. In both cases of 16S and COI probes a direct relation was seen between the signal strength and the distance of the capture binding site to the probe (Fig. 16 and Fig. 17). Preventing this effect is possible by using alternative labelling methods during amplification of the target DNA with labelled dNTPs or by adding labelled signal captures into the hybrization solution (Günther 2004), which bind to and label the conserved regions of the target DNA used.

5.4. Specificity of the microarray

Application of microarray based identification of invertebrate species using specific markers was achieved in the present investigation. Earlier reports of suitability of COI as well as 16S rRNA for species differentiation has prompted us to test and compare both these genes for discriminating the invertebrate species. While 16s rRNA marker based probes could differentiate only 12 of the 15 species (Fig. 13 and Fig. 15), the COI markers worked well for all 15 species (Fig. 14 and Fig. 15) thus indicating superiority of COI over ribosomal marker. Furthermore, the overall cross hybridization observed were less which mainly included cross hybridisation by the 16S probes of closely related crustacean and molluscan species which could be due to the conserved nature of the ribosomal gene in a single class of organisms

(*L. depurator* and *L. vernalis*) (Fig. 30) and also due to the relatively long 16S capture designed for the molluscan species *I. coindetii*, all other molluscan targets were binding to it. However the total number of cross hybridizations observed was less and their signal intensity is so low that it is quite negligible compared to the very high true positive signals (Fig. 15). On the other hand COI targets bound very specifically to the probes and out of the 15 species only one was found to bind unspecifically to the capture of other targets which makes the mitochondrial COI gene very useful for species level discrimination of these animals. Further testing of the microarray with the non target species is important to analyse the specificity of the chip because even after extensive *in silico* analysis with comprehensive sequences, there is a chance of the individual non target DNA collected from the same environment that is not sequenced yet or have an SNP with the target gene to bind specifically to the target capture (Kochzius et al. 2008). Hybridization of four non target DNA from polychaetes that are closely related to the targets enhanced the specificity of the microarray as they showed a weak cross reaction from just a single probe specific for the target *N. hombergii* (Fig. 30 and Table 17). Hybridization temperature and the incubation time greatly effect the specificity of the microarray since they vary depending on the annealing temperature of the primers used for target labelling and Tm of the oligonucleotide probes. Optimization of these conditions for high specific signal and low unspecific binding is highly desirable prior to any routine hybridization experiments (Fig. 10 and Fig. 11). Under optimized hybridization conditions the invertebrate microarray developed in this study was found to be highly specific for the 15 target species used and can probably be applied in species identification procedures in ecological studies if used with care. The reduced cross reaction signals and the high specificity of this microarray may be due to the low number of species investigated. But extending the chip for the identification of more species may not retain the specificity to this extent.

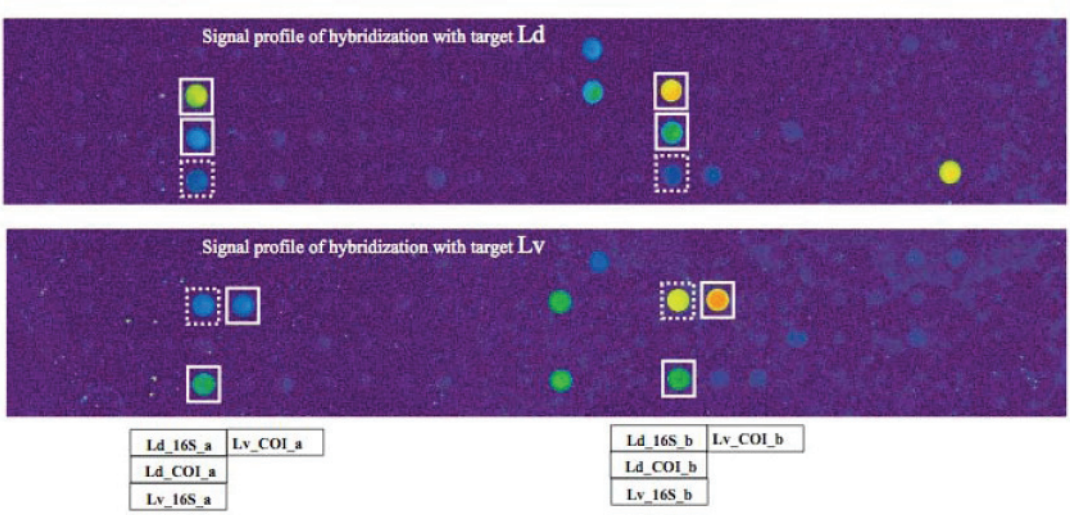


Fig. 30: Example of a microarray showing signal profile of two closely related targets (*L. depurator* and *L. vernalis*). In these experiments both 16S and COI genes are hybridized together. 16S probes of both the species show cross hybridization with both target DNA while the COI probes show signal only with the specific target. Complete thick lines: True-positive signals; Dotted lines: False positive signals.

5.5. Sensitivity of the microarray

Routine utilization of any microarray would desire simultaneous detection of large number of samples in one experiment and efficient discrimination of the target species even in presence of high amount of background samples. The results shown in Fig. 19, Fig. 21 and Fig. 23 representing the hybridizations of two target DNA mixed in varied proportions demonstrate that the signal of one target was not effected at all. Even in the presence of ten fold higher amount of the second target the two targets can be easily be discriminated by the microarray. In case of hybridizations with mixed PCR amplified samples the signal of one target decreased gradually with the increase in the amount of the second target indicating the influence of similar populations upon each other in the hybridization solution. Similarly in three-target hybridizations Fig. 25 - Fig. 27 when two targets are present in very different amounts the third target can still is detected at least in case of one probe. But an overall decrease in the signal intensity from the multiple target amplified products was probably due to the low amplification efficiency during the labelling PCR.

Further, the ability of the microarray to detect the target species in the presence of large amount of background samples in different proportions can be seen in Fig. 28. A variety of DNA from phytoplankton, polychaetes and gut content of fishes were mixed and used as a "background" in this experiment to mimic a natural situation in the oceans. The target signal

gradually decreased as the background concentration was increased but the target can still be detected in the presence of a 20-fold increased concentration of the "background".

5.6. First application

The main drawback of using microarray for the detection and monitoring of environmental samples is their possible cross reaction or unspecific binding resulting in mis-identification of species. This could originate from the many unknown organisms present in the environment since the microarray is designed only for the known species. Additionally, as the microarray carries captures for only a limited number of organisms it is impossible to detect those organisms not represented on the array (Gentry et al. 2006). Further, DNA microarrays used in environmental applications may pose problems in identifying the species lacking exact match, undiscovered haplotype and geographic variants as they cannot anneal to the probes in the array (Hajibabaei et al. 2007). Designing higher level probes such as those specific to the genus and class may probably help to identify unknown organism belonging to the same taxonomic group.

Though having these minor drawbacks microarrays have already been developed for broader application for example in ecological and biodiversity assessment studies of microbial communities and algal blooms (Franke-Whittle et al. 2005; Gescher 2007b; Peplies et al. 2004). The critical step was in these cases the optimizing of the probe design criteria and hybridization conditions for the particular environment. Overcoming the above mentioned problems would further increase the applicability of microarrays as a barcoding tool by identifying any organisms down to species level without extensive sequencing taxonomical investigations (Hajibabaei et al. 2007; Hebert et al. 2003b).

Application in the environment of the invertebrate microarray described in this study could be in ecological and marine food web studies. One example would be analysis of the gut content of demersal fish (predators of invertebrates). This is an important procedure to have an overview on the invertebrate biodiversity and the marine food chain. Hybridization of the DNA extracted from the gut content of the demersal fish *Trigla lucerna* Fig. 29 showed high signals from the probes for the target *G. rhomboides* and few weak signals are also seen from the probes for other crustacean indicating the high presence of crustaceans especially crabs which are known to be common preys of the fish. Though the signal intensities indicate the definite presence of the target in the gut sample, validation of the result with some other analytical techniques is necessary to support the initial observations of the hybridization experiments. As the gut of fish would undergo dynamic changes and the ratio of the prey

species vary randomly quantification would be desirable but is not possible in this stage of development. Nevertheless the microarray can be used for regular monitoring of the gut content for ecological statistics and also to investigate the common preys of commercially important food organisms.

Validation of the gut content analysis was performed by classical microscopy (Mr. M. Stagioni, Laboratory of marine biology, University of Bologna, Italy) of the gut content and/or by cloning and sequencing of the genetic markers from DNA isolated from the gut content (Mrs. A. Cariani, Laboratory of marine biology, University of Bologna). BLASTn search of that sequence provided information about the exact species if the sequences obtained are known from databases or it it is similar to a closer relative with a certain percentage of identity. The morphological identification of the prey in the gut of fish *Trigla lucerna* performed validates that crustaceans comprise majority of fish diet and the crab *Goneplax rhomboides* is present in highest percentage along with less amounts of *Liocarcinus spp*., *Alpheus glaber* and *Pachygrapsus marmoratus a*ppendix 8.6. PCR amplification, cloning and sequencing of DNA isolated from the gut content performed also validates the microarray results obtained. The BLASTn results of the sequences showed an 87 % identity to a crustacean belonging to the class decapoda. But as the microscopical analysis confirms the presence of *G. rhomboides* the clone sequences are aligned with the target sequences obtained in this work and a 100 % identity was found see Appendix 8.5.

5.7. Double marker species identification using microarray

A combination of 16S conserved region and the diverse COI region for designing the oligonucleotides for microarray was advantageous as the signals from two genes provide an additional validity of the species identification. 16S gene being much conserved among the species belonging to a single group, it was hard to find specific regions for capture design. And even though the probes were thoroughly checked for the specificity *in silico,* targets belonging to the same group showed cross hybridizations with the 16S probes of each other. On the other hand COI being very distinct the probe designing was easier and the cross hybridization on the microarray was also not observed. Targets *L. depurator* and *L. vernalis* belonging to the same genus were difficult to distinguish with the 16S probes as they showed cross hybridizations with each other, while no unspecific binding was seen with the COI probes and both the species could be identified accurately in Fig. 15 and Fig. 30. 16S probes targeting the species *C. crangon*, *G. rhomboides* and *L. mediterraneum* showed false negative signals. At least one COI probe for each species gave a signal, though lower in intensity

compared to the 16S. 16S probes for the targets *L. depurator*, *L. vernalis*, *M. norvegica*, *N. hombergii*, *P. koreni* and *H. diversicolor* gave very strong signals with specific targets indicating higher binding capacity of the probe which is beneficial for the exact identification of the species. In some cases the probes from the COI gene showed higher signals than the 16S. So a particular pattern of signals had been seen which can be assigned to each species and can be compared while evaluating unknown samples. As both genes can be simultaneously amplified in the same reaction as shown in this work, the separate amplification cost and time are reduced. Having probes for two genes on the same microarray is beneficial for the correct identification of the species.

5.8. Outlook

The results obtained in this work allow a bunch of further investigations related to methodological improvement and application on real life samples. Signals from different probes specific for the same species were very heterogeneous and it would be beneficial to study the probe characteristics such as binding efficiency (Todt 2006) and spatial arrangement along with the target behaviour in the hybridization solution to obtain more homogenous signals. Though very less the occurrence of false positive signals is not avoided in this work and there are chances of even more such incidents while analysing unknown samples leading to mis-identification. This can be adjusted by optimizing the hybridization conditions following simple modifications of the protocols such as increasing the hybridization temperature by few degrees, using more stringent washing conditions and using different incubation methods. Achieving modulated amplicons from simplified multiplex PCR was not successful during the multi target amplifications and two and three targets together should be worked on to know the exact limitations of the templates than can be amplified together. Similarly the detection limits of the microarray also need further investigation to know the least possibility of target identification amidst diverse DNA in the environment. Furthermore it is important to adjust the design of the microarray and the analysis of the hybridization data for normalization purposes which was not in this work. Specific signal intensity has to be set as a control which determines a positive signal for each probe individually which is more important in analysis mixed samples.

Quantification of the species identification is also needed for any ecological studies and to analyse populations in a defined environment, methods already shown to quantify microbial population (Palmer et al. 2006), can be adopted to marine organisms to investigate the diverse populations of the ocean.

Another chance for using this DNA based system is for the quality control of sea food. There is an urgent requirement for reliable and modern analytical methods to check the quality of all commercial sea food species in Europe and in the USA with the increasing concern of the public about the mislabelled sea food species sold in the market as already seen in case of the fish (Marko et al. 2004). Including species name in the labelling of the sea foods has been a strict regulation by the European Union (EU Council regulation No. 104/2000). This also implies to the invertebrate food species such as mussels and shrimps and microarray which can identify these species will be of potential help for these purposes. But further extension of the chip will be required in terms of developing probes for all the commercially important species so that all of them can be analysed simultaneously. The sequences developed for the targets on the chip and some of the additional sequences developed in this work can be used for the probe designing approach for microarrays or for any taxonomical studies or even for the barcoding purposes. Further the invertebrate chip can be extended to more probes for the identification of more species present in the European Seas. And in far outlook with the development of more probes a separate crustacean, polychaete and mollusc chips might be produced for identification of species of particular phyla on a large scale.

6. Literature review

Aljanabi SM, Martinez I (1997) Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. Nucleic Acids Res 25: 4692-4693

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215: 403-410

Amann R, Snaidr J, Wagner M, Ludwig W, Schleifer KH (1996) In situ visualization of high genetic diversity in a natural microbial community. Journal of Bacteriology 178: 3496-3500

Amann RI, Schleifer, K. H. (2001) Nucleic acid probes and their application in evironmental microbiology. In G. Gravity, D.R. Boone, and R. W. Castenholz (ed.), Bergey's Manual of systemativ Bacteriology, 2nd ed., vol 1.,

Avise JC (1994) Molecular markers, natural history and evolution. Chapman & Hall Newyork

Awadalla P, Eyre-Walker A, Smith JM (1999) Linkage disequilibrium and recombination in hominid mitochondrial DNA. Science 286: 2524-2525

Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Wheeler DL (2008) GenBank. Nucleic Acids Res 36: D25-30

Blair D, Waycott M, Byrne L, Dunshea G, Smith-Keune C, Neil KM (2006) Molecular discrimination of Perna (Mollusca : Bivalvia) species using the polymerase chain reaction and species-specific mitochondrial primers. Marine Biotechnology 8: 380-385

Blankenship LE, Yayanos, A. A. (2005) Universal primers and PCR of gut contents to study marine invertebrate diets. Molecular ecology 14: 891-899

Blohm DH, Guiseppi-Elie A (2001) New developments in microarray technology. Curr Opin Biotechnol 12: 41-47

Bodrossy L, Sessitsch A (2004) Oligonucleotide microarrays in microbial diagnostics. Curr Opin Microbiol 7: 245-254

Bowtell D, Sambrook, J. (2003) DNA microarrays: A molecular cloning manual. Cold spring Harbor Laboratory Press, New York.

Bucklin A, Guarnieri, M., Hill, R. S., Bentley, A. M., Kaartvedt, S. (1999) Taxonomic and systematic assessment of planktonic copepods using mitochondrial COI sequence variation and competitive, species-specific PCR. Hydrobiologia 401: 239-254

Burton RS (1996) Molecular tools in marine ecology. Journal of Experimental Marine Biology and Ecology 200: 85-101

Bustin SA (2002) Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): trends and problems. J Mol Endocrinol 29: 23-39

Call DR, Borucki MK, Loge FJ (2003) Detection of bacterial pathogens in environmental samples using DNA microarrays. J Microbiol Methods 53: 235-243

Chen CA, Yu JK (2000) Universal Primers for Amplification of Mitochondrial Small Subunit Ribosomal RNA-Encoding Gene in Scleractinian Corals. Mar Biotechnol (NY) 2: 146-153

Chen J, Sahota A, Stambrook PJ, Tischfield JA (1991) Polymerase chain reaction amplification and sequence analysis of human mutant adenine phosphoribosyltransferase genes: the nature and frequency of errors caused by Taq DNA polymerase. Mutat Res 249: 169-176

Chen JW, Tauer CG, Huang YH (2002) Nucleotide sequences of the internal transcribed spacers and 5.8S region of nuclear ribosomal DNA in Pinus taeda L. and Pinus echinata Mill. DNA Sequence 13: 129-131

Chen K, Pachter L (2005) Bioinformatics for whole-genome shotgun sequencing of microbial communities. PLoS Comput Biol 1: 106-112

Chomczynski P, Sacchi N (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal Biochem 162: 156-159

Clark MS, Fraser, K. P. P., Peck, L. S. (2008) Lack of an HSP70 heat shock response in two Antarctic marine invertebrates. Polar biology DOI: 10.1007/s00300-008-0447-7

Clayton RA, Sutton G, Hinkle PS, Jr., Bult C, Fields C (1995) Intraspecific variation in small-subunit rRNA sequences in GenBank: why single sequences may not adequately represent prokaryotic taxa. Int J Syst Bacteriol 45: 595-599

Coyne SR, Craw PD, Norwood DA, Ulrich MP (2004) Comparative analysis of the Schleicher and Schuell IsoCode Stix DNA isolation device and the Qiagen QIAamp DNA Mini Kit. J Clin Microbiol 42: 4859-4862

Crabbe MJ (2003) A novel method for the transport and analysis of genetic material from polyps and zooxanthellae of scleractinian corals. J Biochem Biophys Methods 57: 171-176

Dahlgren TG, Lundberg J, Pleijel F, Sundberg P (2000) Morphological and molecular evidence of the phylogeny of Nereidiform polychaetes (Annelida). Journal of Zoological Systematics and Evolutionary Research 38: 249-253

Deagle BE, Jarman SN, Pemberton D, Gales NJ (2005) Genetic screening for prey in the gut contents from a giant squid (Architeuthis sp.). Journal of Heredity 96: 417-423

Desantis TZ, Stone CE, Murray SR, Moberg JP, Andersen GL (2005) Rapid quantification and taxonomic classification of environmental DNA from both prokaryotic and eukaryotic origins using a microarray. FEMS Microbiol Lett 245: 271-278

Devost NC, Choy FY (2000) Mutation analysis of Gaucher disease using dot-blood samples on FTA filter paper. Am J Med Genet 94: 417-420

Dobbs LJ, Madigan MN, Carter AB, Earls L (2002) Use of FTA gene guard filter paper for the storage and transportation of tumor cells for molecular testing. Arch Pathol Lab Med 126: 56-63

Drutschmann DK (2004) Untersuchungen zur Erkennung, zum Nachweis und zur simultanen Genotypisierung von Hepatitis C-viren (HCV) mittels Mikroarrays. PhD thesis, Universität Bremen.

Dyson NJ (1991) Essentail molecular biology: A practical Approach. T. A. Brown. Oxford, Oxford university press. 2.

Feral JP (2002) How useful are the genetic markers in attempts to understand and manage marine biodiversity? Journal of Experimental Marine Biology and Ecology 268: 121-145

Ferraris JD, Palumbi, S.R. (1996) Molecular Zoology: Advances, Strategies and Protocols. Wiley-Liss, Newyork.

Fodor S, Pirrung, M., Read, J., Stryer, L. (1995) Array of oligonucleotides on a solid substrate. USA.

Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Mol Mar Biol Biotechnol 3: 294-299

Fox CJ, Taylor MI, Pereyra R, Villasana MI, Rico C (2005) TaqMan DNA technology confirms likely overestimation of cod (Gadus morhua L.) egg abundance in the Irish Sea: implications for the assessment of the cod stock and mapping of spawning areas using egg-based methods. Mol Ecol 14: 879-884

Franke-Whittle IH, Klammer SH, Insam H (2005) Design and application of an oligonucleotide microarray for the investigation of compost microbial communities. J Microbiol Methods 62: 37-56

Franklin WA, Locker JD (1981) Ethidium-Bromide - a Nucleic-Acid Stain for Tissue-Sections. Journal of Histochemistry & Cytochemistry 29: 572-576

Gentry TJ, Wickham GS, Schadt CW, He Z, Zhou J (2006) Microarray applications in microbial ecology research. Microb Ecol 52: 159-175

Gescher C (2007a) The phytoplankton chip - Development and Assessment of a DNA microarray as a reliable tool for monitoring of phytoplankton. PhD thesis, Universität Bremen.

Gescher C, Metfies, K., Medlin, L. K. (2007b) The ALEX CHIP - Development of a DNA chip for identification and monitoring of Alexandrium. Harmful Algae 360: 1-10

Gmuender H (2002) Perspectives and challenges for DNA microarrays in drug discovery and development. Biotechniques 32: 152-154, 156, 158

Goffredi SK, Jones WJ, Scholin CA, Marin R, 3rd, Vrijenhoek RC (2006) Molecular detection of marine invertebrate larvae. Mar Biotechnol (NY) 8: 149-160

Gomez DK, Baeck GW, Kim JH, Choresca CH, Jr., Park SC (2008) Molecular detection of betanodaviruses from apparently healthy wild marine invertebrates. J Invertebr Pathol 97: 197-202

Gonzalez SF, Krug MJ, Nielsen ME, Santos Y, Call DR (2004) Simultaneous detection of marine fish pathogens by using multiplex PCR and a DNA microarray. J Clin Microbiol 42: 1414-1419

Greenstone MH, Rowley DL, Heimbach U, Lundgren JG, Pfannenstiel RS, Rehner SA (2005) Barcoding generalist predators by polymerase chain reaction: carabids and spiders. Mol Ecol 14: 3247-3266

Günther S (2004) DNA Mikroarray-Technologie zur Klassifizierung von Actinomyceten der Gattung Kitasatospora. PhD thesis, Friedrich-Schiller-Universität, Jena.

Guschin DY, Mobarry BK, Proudnikov D, Stahl DA, Rittmann BE, Mirzabekov AD (1997) Oligonucleotide microchips as genosensors for determinative and environmental studies in microbiology. Appl Environ Microbiol 63: 2397-2402

Hajibabaei M, Singer GA, Clare EL, Hebert PD (2007) Design and applicability of DNA arrays and DNA barcodes in biodiversity monitoring. BMC Biol 5: 24

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis programfor Windows 95/98/NT. Nucleic acids Symposium Series 41: 95-98

Harvey ML, Dadour IR, Gaudieri S (2003) Mitochondrial DNA cytochrome oxidase I gene: potential for distinction between immature stages of some forensically important fly species (Diptera) in western Australia. Forensic Sci Int 131: 134-139

Hebert PD, Cywinska A, Ball SL, deWaard JR (2003a) Biological identifications through DNA barcodes. Proc Biol Sci 270: 313-321

Hebert PDN, Cywinska A, Ball SL, DeWaard JR (2003b) Biological identifications through DNA barcodes. Proceedings of the Royal Society of London Series B-Biological Sciences 270: 313-321

Hebert PDN, Stoeckle MY, Zemlak TS, Francis CM (2004) Identification of birds through DNA barcodes. Plos Biology 2: 1657-1663

Henegariu O, Heerema NA, Dlouhy SR, Vance GH, Vogt PH (1997) Multiplex PCR: critical parameters and step-by-step protocol. Biotechniques 23: 504-511

Hide G, Hughes JM, McNuff R (2003) A rapid and simple method of detection of Blepharisma japonicum using PCR and immobilisation on FTA paper. BMC Ecol 3: 7

Hogg ID, Hebert PDN (2004) Biological identification of springtails (Hexapoda : Collembola) from the Canadian Arctic, using mitochondrial DNA barcodes. Canadian Journal of Zoology-Revue Canadienne De Zoologie 82: 749-754

Hold GL, Russell VJ, Pryde SE, Rehbein H, Quinteiro J, Vidal R, Rey-Mendez M, Sotelo CG, Perez-Martin RI, Santos AT, Rosa C (2001) Development of a DNA-based method aimed at identifying the fish species present in food products. J Agric Food Chem 49: 1175-1179

Howbrook DN, van der Valk AM, O'Shaughnessy MC, Sarker DK, Baker SC, Lloyd AW (2003) Developments in microarray technologies. Drug Discov Today 8: 642-651

Hsia CC, Chizhikov VE, Yang AX, Selvapandiyan A, Hewlett I, Duncan R, Puri RK, Nakhasi HL, Kaplan GG (2007) Microarray multiplex assay for the simultaneous detection and discrimination of hepatitis B, hepatitis C, and human immunodeficiency type-1 viruses in human blood samples. Biochem Biophys Res Commun 356: 1017-1023

Hudson ME (2008) Sequencing breakthroughs for genomic ecology and evolutionary biology. Molecular ecology notes 8: 3-17

Hutchings PA (1999a) Taxonomy of estuarine inverterbates in Australia. Australian Journal of Ecology. 24: 381-394

Janzen DH, Hajibabaei M, Burns JM, Hallwachs W, Remigio E, Hebert PDN (2005) Wedding biodiversity inventory of a large and complex Lepidoptera fauna with DNA barcoding. Philosophical Transactions of the Royal Society B-Biological Sciences 360: 1835-1845

Jeffreys AJ, Wilson V, Thein SL (1985) Individual-Specific Fingerprints of Human DNA. Nature 316: 76-79

Jerome M, Lemaire C, Verrez-Bagnis V, Etienne M (2003) Direct sequencing method for species identification of canned sardine and sardine-type products. J Agric Food Chem 51: 7326-7332

Jung A, Stemmler I, Brecht A, Gauglitz G (2001) Covalent strategy for immobilization of DNA-microspots suitable for microarrays with label-free and time-resolved optical detection of hybridization. Fresenius J Anal Chem 371: 128-136

Kaderali L, Schliep A (2002) Selecting signature oligonucleotides to identify organisms using DNA arrays. Bioinformatics 18: 1340-1349

Kochzius M, Blohm D (2005) Genetic population structure of the lionfish Pterois miles (Scorpaenidae, Pteroinae) in the Gulf of Aqaba and northern Red Sea. Gene. 347: 295-301. Epub 2005 Feb 2019.

Kochzius M, Kappel, K., Doblitz, L., Silkenbaumer, N., Nolte, M., Weber, H., Hjorleifsdottir, S., Marteinsson, V., Hreggvidsson, G., Planes, S., Tinti, F., Magoulas, A., Vazquez, E. G., Turan, C., Medlin, L., Metfies, K., Gescher, C., Cariani, A., Landi, M., Hervet, C., Falgueras, D.C., Antoniou, A., Bertasi, F., Chitipothu, S., Blohm, D. (2007) The "Fish & Chips" project: Microarrays as a tool for the identification of marine organisms in biodiversity and ecosystem research. OCEANS '07 IEEE Aberdeen Conference Proceedings: 1 - 4

Kochzius M, Nolte M, Weber H, Silkenbeumer N, Hjorleifsdottir S, Hreggvidsson GO, Marteinsson V, Kappel K, Planes S, Tinti F, Magoulas A, Garcia Vazquez E, Turan C, Hervet C, Campo Falgueras D, Antoniou A, Landi M, Blohm D (2008) DNA microarrays for identifying fishes. Mar Biotechnol (NY) 10: 207-217

Kochzius M, Soller R, Khalaf MA, Blohm D (2003) Molecular phylogeny of the lionfish genera Dendrochirus and Pterois (Scorpaenidae, Pteroinae) based on mitochondrial DNA sequences. Mol Phylogenet Evol. 28: 396-403.

Kohara Y, Noda H, Okano K, Kambara H (2001) DNA hybridization using "bead-array": probe-attached beads arrayed in a capillary in a predetermined order. Nucleic Acids Res Suppl: 83-84

Koizumi Y, Kelly JJ, Nakagawa T, Urakawa H, El-Fantroussi S, Al-Muzaini S, Fukui M, Urushigawa Y, Stahl DA (2002) Parallel characterization of anaerobic toluene- and ethylbenzene-degrading microbial consortia by PCR-denaturing gradient gel electrophoresis, RNA-DNA membrane hybridization, and DNA microarray technology. Applied and Environmental Microbiology 68: 3215-3225

Ku WC, Lau WK, Tseng YT, Tzeng CM, Chiu SK (2004) Dextran sulfate provides a quantitative and quick microarray hybridization reaction. Biochem Biophys Res Commun 315: 30-37

Kumar S, Tamura K, Nei M (2004) MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. Brief Bioinform 5: 150-163

Kuo WP, Jenssen TK, Butte AJ, Ohno-Machado L, Kohane IS (2002) Analysis of matched mRNA measurements from two different microarray technologies. Bioinformatics 18: 405-412

Lashkari DA, DeRisi JL, McCusker JH, Namath AF, Gentile C, Hwang SY, Brown PO, Davis RW (1997) Yeast microarrays for genome wide parallel genetic and gene expression analysis. Proc Natl Acad Sci U S A 94: 13057-13062

Le Goff-Vitry MC, Jacquelin, S., Comtet, T. (2007) Towards tracking marine larvae with in situ hybridization. Journal of Marine Biological Association of the UK 87: 1077-1080

Li X, Gu, W., Mohan, S., Baylink, J. (2002) "DNA Microarrays: Their Use and Misuse.". Microcirculation 9: 13-22

Lin JJ, Fleming R, Kuo J, Matthews BF, Saunders JA (2000) Detection of plant genes using a rapid, nonorganic DNA purification method. Biotechniques 28: 346-350

Lipshutz RJ, Fodor SP, Gingeras TR, Lockhart DJ (1999) High density synthetic oligonucleotide arrays. Nat Genet 21: 20-24

Liu RH, Yang J, Lenigk R, Bonanno J, Grodzinski P (2004) Self-contained, fully integrated biochip for sample preparation, polymerase chain reaction amplification, and DNA microarray detection. Anal Chem 76: 1824-1831

Liu WT, Mirzabekov AD, Stahl DA (2001) Optimization of an oligonucleotide microchip for microbial identification studies: a non-equilibrium dissociation approach. Environ Microbiol 3: 619-629

Loy A, Bodrossy L (2006) Highly parallel microbial diagnostics using oligonucleotide microarrays. Clin Chim Acta 363: 106-119

Loy A, Lehner A, Lee N, Adamczyk J, Meier H, Ernst J, Schleifer KH, Wagner M (2002) Oligonucleotide microarray for 16S rRNA gene-based detection of all recognized lineages of sulfate-reducing prokaryotes in the environment. Applied and Environmental Microbiology 68: 5064-5081

Maggioni R, Rogers AD, Maclean N, D'Incao F (2001) Molecular phylogeny of western Atlantic Farfantepenaeus and Litopenaeus shrimp based on mitochondrial 16S partial sequences. Mol Phylogenet Evol 18: 66-73

Mallet J, K. Willmott (2003) Taxonomy: renaissance or Tower of Babel? Trends in Ecology & Evolution 18: 57-59

Marko PB, Lee SC, Rice AM, Gramling JM, Fitzhenry TM, McAlister JS, Harper GR, Moran AL (2004) Fisheries: mislabelling of a depleted reef fish. Nature 430: 309-310

Maynard C, Berthiaume F, Lemarchand K, Harel J, Payment P, Bayardelle P, Masson L, Brousseau R (2005) Waterborne pathogen detection by use of oligonucleotide-based microarrays. Appl Environ Microbiol 71: 8548-8557

McCarthy MA, Harte, C. G., Wiggenraad, J. F. M., Michielsen, A. L. P. J., Kohlgrueber, D., Kamoulakos, A. (1998) Computational analysis of energy-absorbing composite aerospace structures under crash loading. International conference on computational engineering science, Atlanta, USA.: 1552-1557

Metfies K, Huljic S, Lange M, Medlin LK (2005) Electrochemical detection of the toxic dinoflagellate Alexandrium ostenfeldii with a DNA-biosensor. Biosens Bioelectron 20: 1349-1357

Metfies K, Medlin, L. (2004) DNA microchips for plytoplankton: the fluorescent wave of the future. Nova Hedwigia 79: 321-327

Miller DN, Bryant JE, Madsen EL, Ghiorse WC (1999) Evaluation and optimization of DNA extraction and purification procedures for soil and sediment samples. Appl Environ Microbiol 65: 4715-4724

Mullis K, Faloona F, Scharf S, Saiki R, Horn G, Erlich H (1986) Specific Enzymatic Amplification of DNA Invitro - the Polymerase Chain-Reaction. Cold Spring Harbor Symposia on Quantitative Biology 51: 263-273

Natarajan P, Trinh T, Mertz L, Goldsborough M, Fox DK (2000) Paper-based archiving of mammalian and plant samples for RNA analysis. Biotechniques 29: 1328-1333

Nicholas KB, Nicholas, H. B. J. (1997) Genedoc: atool for editing and annotation multiple sequence alignements. Distributed by the author. Pittsburgh.

Nicolaisen M, Justesen AF, Thrane U, Skouboe P, Holmstrom K (2005) An oligonucleotide microarray for the identification and differentiation of trichothecene producing and non-producing Fusarium species occurring on cereal grain. Journal of Microbiological Methods 62: 57-69

Niemeyer CF, Blohm, D. H. (1999) "DNA Mikroarrays". Angewandte Chemie 111(19): 3039-3043

Nolan T, Hands RE, Bustin SA (2006) Quantification of mRNA using real-time RT-PCR. Nat Protoc 1: 1559-1582

Nölte M (2002) Optimierung von oligonukelotid-Bibliotheken für DNA-Mikroarrays. PhD thesis, Universität Bremen.

Nygren A, Pleijel F, Sundberg P (2005) Genetic relationships between Nereimyra punctata and N-woodsholea (Hesionidae, Polychaeta). Journal of Zoological Systematics and Evolutionary Research 43: 273-276

Orlando C, Pinzani P, Pazzagli M (1998) Developments in quantitative PCR. Clin Chem Lab Med 36: 255-269

Palmer C, Bik EM, Eisen MB, Eckburg PB, Sana TR, Wolber PK, Relman DA, Brown PO (2006) Rapid quantitative profiling of complex microbial populations. Nucleic Acids Res 34: e5

Palumbi SR, Cipriano F (1998) Species identification using genetic tools: The value of nuclear and mitochondrial gene sequences in whale conservation. Journal of Heredity 89: 459-464

Palumbi SR, Martin, A., Romano, S., McMillan, W. O., Stice, L., Grabowski, G. (1991) The simple fool's guide to PCR. Version 2.0, Department of Zology and Kewalo Marine Laboratory, University of Hawaii, Honolulu, HI.

Panicker G, Call DR, Krug MJ, Bej AK (2004) Detection of pathogenic Vibrio spp. in shellfish by using multiplex PCR and DNA microarrays. Appl Environ Microbiol 70: 7436-7444

Parson W, Pegoraro K, Niederstatter H, Foger M, Steinlechner M (2000) Species identification by means of the cytochrome b gene. Int J Legal Med 114: 23-28

Peplies J, Glockner FO, Amann R (2003) Optimization strategies for DNA microarray-based detection of bacteria with 16S rRNA-targeting oligonucleotide probes. Appl Environ Microbiol 69: 1397-1407

Peplies J, Lachmund C, Glockner FO, Manz W (2006) A DNA microarray platform based on direct detection of rRNA for characterization of freshwater sediment-related prokaryotic communities. Appl Environ Microbiol 72: 4829-4838

Peplies J, Lau SCK, Pernthaler J, Amann R, Glockner FO (2004) Application and validation of DNA microarrays for the 16S rRNA-based analysis of marine bacterioplankton. Environmental Microbiology 6: 638-645

Pfunder M, Holzgang O, Frey JE (2004a) Development of microarray-based diagnostics of voles and shrews for use in biodiversity monitoring studies, and evaluation of

mitochondrial cytochrome oxidase I vs. cytochrome b as genetic markers. Mol Ecol 13: 1277-1286

Pfunder M, Holzgang O, Frey JE (2004b) Development of microarray-based diagnostics of voles and shrews for use in biodiversity monitoring studies, and evaluation of mitochondrial cytochrome oxidase I vs. cytochrome b as genetic markers. Molecular Ecology 13: 1277-1286

Ponder WF, Lindberg, D. R. (2008) Phylogeny and Evolution of the Mollusca. Berkeley: University of California Press. 481 pp.

Radchenko OA (2004) Variability of nucleotide sequences of mitochondrail DNA cytochrome b gene in chars of the genus Salvalinus. Russian journal of genetics 40(3): 244-254

Relogio A, Schwager C, Richter A, Ansorge W, Valcarcel J (2002) Optimization of oligonucleotide-based DNA microarrays. Nucleic Acids Research 30: -

Remigio EA, Hebert PD (2003) Testing the utility of partial COI sequences for phylogenetic estimates of gastropod relationships. Mol Phylogenet Evol 29: 641-647

Rohland N, Hofreiter M (2007) Ancient DNA extraction from bones and teeth. Nat Protoc 2: 1756-1762

Ronning SB, Rudi K, Berdal KG, Holst-Jensen A (2005) Differentiation of important and closely related cereal plant species (Poaceae) in food by hybridization to an oligonucleotide array. J Agric Food Chem 53: 8874-8880

Rosenblum BB, Lee LG, Spurgeon SL, Khan SH, Menchen SM, Heiner CR, Chen SM (1997) New dye-labeled terminators for improved DNA sequencing patterns. Nucleic Acids Research 25: 4500-4504

Rouse GW, Pleijel, F. (2001) Polychaetes. Oxford University Press.

Rudi K, Skulberg OM, Skulberg R, Jakobsen KS (2000) Application of sequence-specific labeled 16S rRNA gene oligonucleotide probes for genetic profiling of cyanobacterial abundance and diversity by array hybridization. Appl Environ Microbiol 66: 4004-4011

Saccone C, De Giorgi C, Gissi C, Pesole G, Reyes A (1999) Evolutionary genomics in Metazoa: the mitochondrial DNA as a model system. Gene 238: 195-209

Sahm K, MacGregor BJ, Jorgensen BB, Stahl DA (1999) Sulphate reduction and vertical distribution of sulphate-reducing bacteria quantified by rRNA slot-blot hybridization in a coastal marine sediment. Environ Microbiol 1: 65-74

Saiki RK, Gelfand DH, Stoffel S, Scharf SJ, Higuchi R, Horn GT, Mullis KB, Erlich HA (1988) Primer-Directed Enzymatic Amplification of DNA with a Thermostable DNA-Polymerase. Science 239: 487-491

Saiki RK, Walsh PS, Levenson CH, Erlich HA (1989) Genetic analysis of amplified DNA with immobilized sequence-specific oligonucleotide probes. Proc Natl Acad Sci U S A 86: 6230-6234

Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci U S A 74: 5463-5467

SantaLucia J, Jr., Allawi HT, Seneviratne PA (1996) Improved nearest-neighbor parameters for predicting DNA duplex stability. Biochemistry 35: 3555-3562

Schander C, Willassen, E. (2005) What can biological barcoding do for marine biology? Marine biology research 1: 79-83

Schena M, Shalon D, Davis RW, Brown PO (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science 270: 467-470

Sebastio P, Zanelli P, Neri TM (2001) Identification of anchovy (Engraulis encrasicholus L.) and gilt sardine (Sardinella aurita) by polymerase chain reaction, sequence of their mitochondrial cytochrome b gene, and restriction analysis of polymerase chain reaction products in semipreserves. J Agric Food Chem 49: 1194-1199

Sengupta S, Onodera K, Lai A, Melcher U (2003) Molecular detection and identification of influenza viruses by oligonucleotide microarray hybridization. J Clin Microbiol 41: 4542-4550

Service RF (1998) DNA chips survey an entire genome. Science 281: 1122

Shchepinov MS, Case-Green SC, Southern EM (1997) Steric factors influencing hybridisation of nucleic acids to oligonucleotide arrays. Nucleic Acids Res 25: 1155-1161

Shi L, Reid LH, Jones WD, Shippy R, Warrington JA, Baker SC, Collins PJ, de Longueville F, Kawasaki ES, Lee KY, Luo Y, Sun YA, Willey JC, Setterquist RA, Fischer GM, Tong W, Dragan YP, Dix DJ, Frueh FW, Goodsaid FM, Herman D, Jensen RV, Johnson CD, Lobenhofer EK, Puri RK, Schrf U, Thierry-Mieg J, Wang C, Wilson M, Wolber PK, Zhang L, Amur S, Bao W, Barbacioru CC, Lucas AB, Bertholet V, Boysen C, Bromley B, Brown D, Brunner A, Canales R, Cao XM, Cebula TA, Chen JJ, Cheng J, Chu TM, Chudin E, Corson J, Corton JC, Croner LJ, Davies C, Davison TS, Delenstarr G, Deng X, Dorris D, Eklund AC, Fan XH, Fang H, Fulmer-Smentek S, Fuscoe JC, Gallagher K, Ge W, Guo L, Guo X, Hager J, Haje PK, Han J, Han T, Harbottle HC, Harris SC, Hatchwell E, Hauser CA, Hester S, Hong H, Hurban P, Jackson SA, Ji H, Knight CR, Kuo WP, LeClerc JE, Levy S, Li QZ, Liu C, Liu Y, Lombardi MJ, Ma Y, Magnuson SR, Maqsodi B, McDaniel T, Mei N, Myklebost O, Ning B, Novoradovskaya N, Orr MS, Osborn TW, Papallo A, Patterson TA, Perkins RG, Peters EH, Peterson R, Philips KL, Pine PS, Pusztai L, Qian F, Ren H, Rosen M, Rosenzweig BA, Samaha RR, Schena M, Schroth GP, Shchegrova S, Smith DD, Staedtler F, Su Z, Sun H, Szallasi Z, Tezak Z, Thierry-Mieg D, Thompson KL, Tikhonova I, Turpaz Y, Vallanat B, Van C, Walker SJ, Wang SJ, Wang Y, Wolfinger R, Wong A, Wu J, Xiao C, Xie Q, Xu J, Yang W, Zhong S, Zong Y, Slikker W, Jr. (2006) The MicroArray Quality Control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements. Nat Biotechnol 24: 1151-1161

Silkenbeumer N (2004) Molecular identification of different Pleuronectiformes of the German Bight using Microarrays. Diploma thesis, Universität Bremen.

Simon N, Campbell L, Ornolfsdottir E, Groben R, Guillou L, Lange M, Medlin LK (2000) Oligonucleotide probes for the identification of three algal groups by dot blot and fluorescent whole-cell hybridization. J Eukaryot Microbiol 47: 76-84

Sinha RP, Dautz, M., Hader, D. P. (2001) A simple and efficient method for the quantitative analysis of thymine Cyanobacteria, Phytoplankton and Macroalgae. Acta protozoologica 40: 187-195

Small J, Call DR, Brockman FJ, Straub TM, Chandler DP (2001) Direct detection of 16S rRNA in soil extracts by using oligonucleotide microarrays. Applied and Environmental Microbiology 67: 4708-4716

Smith LM, Burgoyne LA (2004) Collecting, archiving and processing DNA from wildlife samples using FTA databasing paper. BMC Ecol 4: 4

Southern E, Mir K, Shchepinov M (1999) Molecular interactions on microarrays. Nat Genet 21: 5-9

Southern EM (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. J Mol Biol 98: 503-517

Steemers FJ, Ferguson JA, Walt DR (2000) Screening unlabeled DNA targets with randomly ordered fiber-optic gene arrays. Nat Biotechnol 18: 91-94

Subrungruang I, Mungthin M, Chavalitshewinkoon-Petmitr P, Rangsin R, Naaglor T, Leelayoova S (2004) Evaluation of DNA extraction and PCR methods for detection of Enterocytozoon bienuesi in stool specimens. J Clin Microbiol 42: 3490-3494

Todt S (2006) Untersuchungen zur Hybridisierungseffizienz von Oligonukelotiden in DNA Mikroarray-Experimenten. PhD thesis, Universität Bremen.

Urakawa H, El Fantroussi S, Smidt H, Smoot JC, Tribou EH, Kelly JJ, Noble PA, Stahl DA (2003) Optimization of single-base-pair mismatch discrimination in oligonucleotide microarrays. Appl Environ Microbiol 69: 2848-2856

Uthicke. S. B, J. A. H. (2002) A genetic fingerprint recapture technique for measuirng growth in 'unmarkable' invertebrates: negative growth in commercially fished holothurians (Holothuria nobilis). Marine ecology 241: 221-226

Vadopalas B, Bouma JV, Jackels CR, Friedman CS (2006) Application of real-time PCR for simultaneous identification and quantification of larval abalone. Journal of Experimental Marine Biology and Ecology 334: 219-228

Van 't Veer LJ, De Jong D (2002) The microarray way to tailored cancer treatment. Nat Med 8: 13-14

Vetter C (2005) AmpliChip CYP450: Erster Genchip für die Klinische Routine. Deutsches Ärzteblatt 102 (9)

Voznesensky M, Lenz PH, Spanings-Pierrot C, Towle DW (2004) Genomic approaches to detecting thermal stress in Calanus finmarchicus (Copepoda : Calanoida). Journal of Experimental Marine Biology and Ecology 311: 37-46

Wallace RB, Shaffer J, Murphy RF, Bonner J, Hirose T, Itakura K (1979) Hybridization of synthetic oligodeoxyribonucleotides to phi chi 174 DNA: the effect of single base pair mismatch. Nucleic Acids Res 6: 3543-3557

Walsh PS, Metzger DA, Higuchi R (1991) Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. Biotechniques 10: 506-513

Wan QH, Fang SG (2003) Application of species-specific polymerase chain reaction in the forensic identification of tiger species. Forensic Sci Int 131: 75-78

Wang B, Li F, Dong B, Zhang X, Zhang C, Xiang J (2006) Discovery of the genes in response to white spot syndrome virus (WSSV) infection in Fenneropenaeus chinensis through cDNA microarray. Mar Biotechnol (NY) 8: 491-500

Wang J (2000) From DNA biosensors to gene chips. Nucleic Acids Res 28: 3011-3016

Ward RD, Zemlak TS, Innes BH, Last PR, Hebert PDN (2005) DNA barcoding Australia's fish species. Philosophical Transactions of the Royal Society B-Biological Sciences 360: 1847-1857

Warsen AE, Krug MJ, LaFrentz S, Stanek DR, Loge FJ, Call DR (2004) Simultaneous discrimination between 15 fish pathogens by using 16S ribosomal DNA PCR and DNA microarrays. Appl Environ Microbiol 70: 4216-4221

Watson JD, Crick FH (1953) Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. Nature 171: 737-738

Weinberg JR, Dahlgren TG, Trowbridge N, Halanych KM (2003) Genetic differences within and between species of deep-sea crabs (Chaceon) from the North Atlantic Ocean. Biological Bulletin 204: 318-326

Willey JC, Crawford EL, Jackson CM, Weaver DA, Hoban JC, Khuder SA, DeMuth JP (1998) Expression measurement of many genes simultaneously by quantitative RT-PCR using standardized mixtures of competitive templates. Am J Respir Cell Mol Biol 19: 6-17

Wilson KH, Wilson WJ, Radosevich JL, DeSantis TZ, Viswanathan VS, Kuczmarski TA, Andersen GL (2002) High-density microarray of small-subunit ribosomal DNA probes. Applied and Environmental Microbiology 68: 2535-2541

Woese CR (1998) A manifesto for microbial genomics. Current Biology 8: R781-R783

Wrobel G, Schlingemann J, Hummerich L, Kramer H, Lichter P, Hahn M (2003) Optimization of high-density cDNA-microarray protocols by 'design of experiments'. Nucleic Acids Res 31: e67

Wu Y, de Kievit P, Vahlkamp L, Pijnenburg D, Smit M, Dankers M, Melchers D, Stax M, Boender PJ, Ingham C, Bastiaensen N, de Wijn R, van Alewijk D, van Damme H,

Raap AK, Chan AB, van Beuningen R (2004) Quantitative assessment of a novel flow-through porous microarray for the rapid analysis of gene expression profiles. Nucleic Acids Res 32: e123

Xu H, Sha MY, Wong EY, Uphoff J, Xu Y, Treadway JA, Truong A, O'Brien E, Asquith S, Stubbins M, Spurr NK, Lai EH, Mahoney W (2003) Multiplexed SNP genotyping using the Qbead system: a quantum dot-encoded microsphere-based assay. Nucleic Acids Res 31: e43

Yadetie F, Sandvik AK, Bergum H, Norsett K, Laegreid A (2004) Miniaturized fluorescent RNA dot blot method for rapid quantitation of gene expression. BMC Biotechnol 4: 12

Zelm R (2004) Untersuchung des 'Bindungsprofils' von Oligonukleotiden im Hinblick auf ihr Hybridisierungsverhalten auf DNA-Mikroarrays. Diploma thesis, Universität Bremen.

Zhang L, Hurek T, Reinhold-Hurek B (2005) Position of the fluorescent label is a crucial factor determining signal intensity in microarray hybridizations. Nucleic Acids Res 33: e166

7. Zusammenfassung

Neben Genexpressionsstudien haben DNA Mikroarrays auch großes Potential, in anderen Bereichen der Wissenschaft eingesetzt zu werden, in denen der Bedarf zur hoch parallelen Untersuchung einer großen Anzahl von Organismen besteht. Hierzu zählen in der marinen Ökologie das Studium von Nahrungsnetzen, der Biodiversität oder die Artenidentifikation im Allgemeinen. Auf Grund der sinkenden Zahl von auf morphologische Kennzeichen spezialisierter Fachleute ist dies von großer Wichtigkeit.

Die Entwicklung von Mikroarrays und ihre Anwendbarkeit in Umweltstudien hängt sowohl von der Verfügbarkeit hoch spezifischer Fängeroligonukleotide als auch von der Optimierung der spezifischen experimentellen Bedingungen ab, um eine einzelne Spezies in der Gegenwart einer großen Menge an Hintergrundproben nachweisen zu können.

Um die Nutzung dieser Technologie zu ermöglichen, wurden die damit verbundenen vielfältigen methodischen Probleme in dieser Arbeit intensiv untersucht. Dies führte zur Entwicklung eines Mikroarray Prototypen mit 64 Oligonukleotidfängern gegen 16S und COI Gene von 15 häufigen Invertebratenarten mit wissenschaftlicher und ökonomischer Relevanz der europäischen Meere.

Ausgehend von Sequenzen für Zielspezies, die im Rahmen der Laborarbeit erlangt wurden, und von Sequenzen für Ziel- und Nicht-Zielorganismen aus der Datenbank wurde ein verlässliches Set von Fängern entwickelt, das das Kriterium der Spezifität *in silico* erfüllt und nach der Optimierung der Hybridisierungsbedingungen noch redundante Proben enthält.

Die Spezifität und Sensitivität des Chips wurden sowohl *in silico* als auch experimentell untersucht. Hierzu wurden die Hybridisierungsbedingungen entsprechend optimiert. Spezifische Signale an den jeweiligen Fängern auf dem Chip konnten beobachtet werden, was zu klarer Unterscheidbarkeit der untersuchten Arten führte. Falsch positive Signale konnten während der Untersuchung signifikant reduziert werden. Das Design der Fänger sowohl für 16S als auch für COI Gene erlaubte die Steigerung der Genauigkeit bei der Unterscheidung sogar sehr nah verwandter Arten, die auf Individuenniveau sowohl inter- als auch intraspezifische Unterschiede zeigten.

Als erste Anwendung des Mikroarrays wurde das Verhalten in Multi-Target Experimenten untersucht. Hierbei wurden Unweltproben durch definierte Gemische verschiedener Arten in verschiedenen Zusammensetzungen nachgeahmt. Die Fähigkeit des Chips, eine niedrig konzentrierte Zielart in Anwesenheit eines großen Hintergrunds nachweisen zu können, wurde ebenfalls durch den Einsatz von DNA verschiedener Nicht-Zielinvertebraten

nachgewiesen. Die Fähigkeit des Arrays, den Inhalt des Verdauungstraktes von Raubfischen zu analysieren, wurde untersucht indem extrahierte DNA aus dem Darminhalt des Grundfisches *Trigla lucerna* hybridisiert wurde. In Übereinstimmung mit Abschätzungen durch klassische Methoden wurde ein DNA-Gemisch verschiedener Invertebraten mit Hilfe des Mikroarrays identifiziert. Diese Ergebnisse wurden durch Klonierung und Sequenzierung des Darminhalts verifiziert.

Diese Experimente zeigen, dass der Invertebratenchip semiquantitative Informationen über wichtige Beutearten von Grundfischen und dadurch schlussendlich auch über Nahrungsnetze mariner Ökosystem liefern kann.

Da marine Invertebraten sehr klein und ein für die DNA Extraktion häufig schwieriges Gewebe sind, wurde eine kleine vergleichende Studie von klassischen und kommerziell erhältlichen DNA Extraktionsmethoden durchgeführt. Hierbei sollte eine im Idealfall universelle Methode definiert werden, um ausreichende Mengen von DNA aus allen Arten von Gewebe isolieren zu können. „FTA Elute“ stellte sich als sehr verlässliche, billige und bequem zu handhabe Methode heraus, um DNA zu extrahieren und zu lagern, selbst außerhalb des Labors unter Praxisbedingungen.

Zusammenfassend wurde innerhalb dieser Arbeit ein schnelles und relativ einfach zu handhabendes DNA-Analysesystem zur Identifikation von sowohl morphologisch schwierig zu unterscheidenden, wichtigen marinen Invertebraten als auch von komplexen Organismengemischen ausgearbeitet.

8. Appendix

8.1. Pictures of the target species

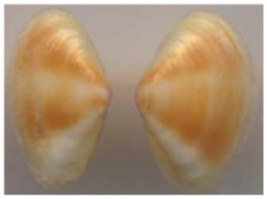
Though they were ugly and untouchable at the first sight, they behaved very well all through the study and they are the most beautiful creatures under microscope.

8.1.1. Mollusca:

Eledone cirrhosa

Source: www.marlin.ac.uk

Lentidium mediterraneum



Source: www.digilander.libero.it

Chamelea gallina

Source: www.marlin.ac.uk

Illex coindetii

Source: www.mareinitaly.it

8.1.2. Annelida:

Hediste diversicolor

Source: www.uv.es

Nephtys hombergii

Source: www.animaldiversity.edu

Pectinaria koreni

Source: www.wikipedia.org

Sigalion mathildae

Source: www.wikipedia.org

8.1.3. Crustacea:

Alpheus glaber

Source: www.marlin.ac.uk

Crangon crangon

Source: www.wikipedia.org

Goneplax rhomboides

Source: www.seaslug.org.uk

Liocarcinus depurator

Source: www.seaslug.org.uk

Liocarcinus vernalis

Source: www.wikipedia.org

Meganyctiphanes norvegica

Source: www.wikipedia.org

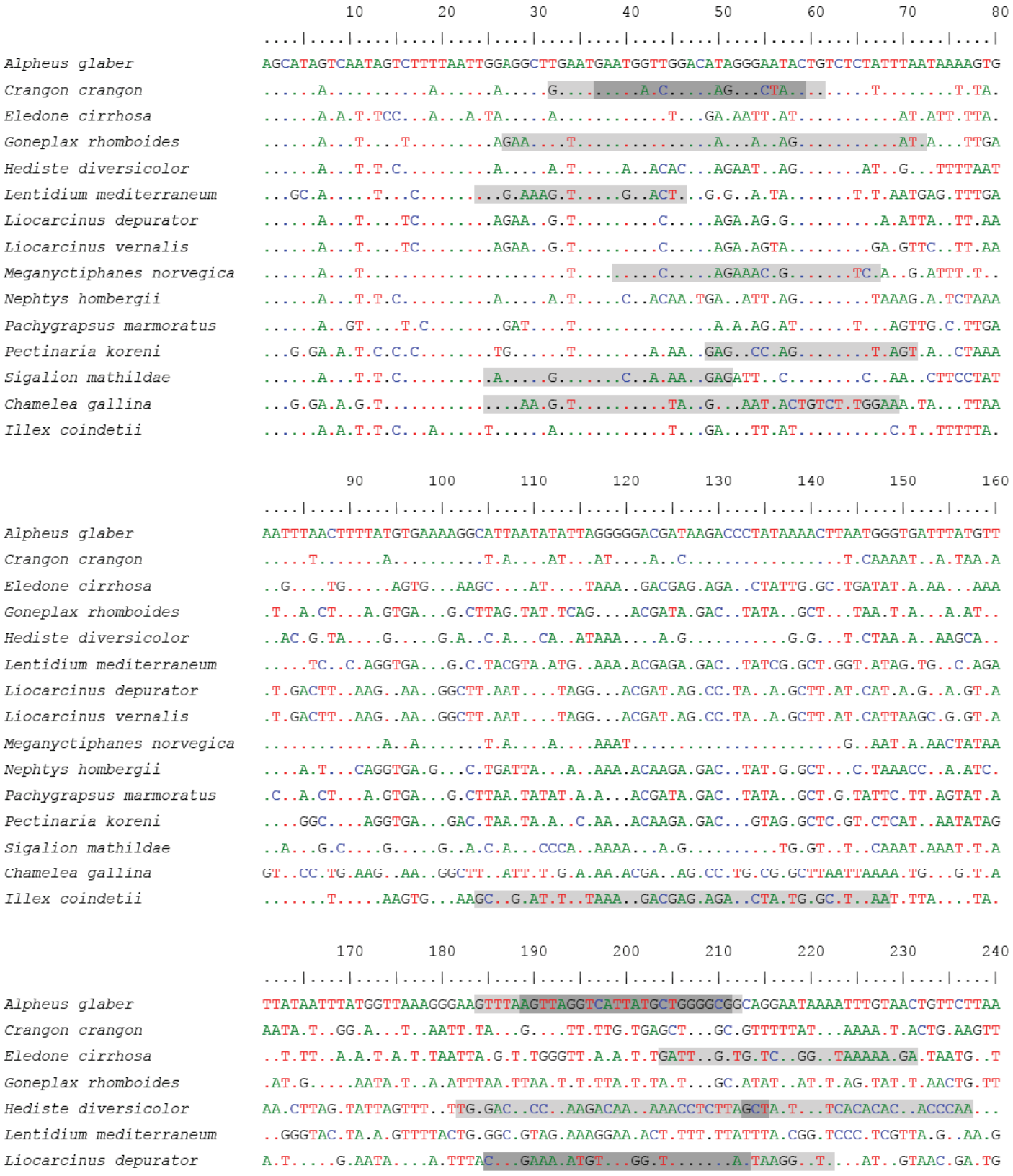
Pachygrapsus marmoratus

Source: www.marlin.ac.uk

8.2. Allignments of the target sequences

8.2.1. 16S alignment and oligonucleotides

All the oligonucleotide probes designed are marked in grey and the dark portions show the overlapping regions.



```
Liocarcinus vernalis        A.T.....G.A.TA.....TTTA..C.GAAA..TGT...GG.T.......A.TAAGG..T....AT.GGTAAC.GA.T.
Meganyctiphanes norvegica   AA........GCTA.TT.AAATTTA..T...GGAT.TTA.T.TA.T...GC.ATAA..AT...A.AAGT.ACTG.T..T.
Nephtys hombergii           .CTA.TAA.TATA..TT.CTATCAAAAGGAAC.A..TTGGT..GG.C.A.AAAG.A..AC...AACC.TC..TCA...TT
Pachygrapsus marmoratus     AGTCT.A..CGA.GGCT.TAA.G.C..GGTGCTTAAAT..TATTATATTG..GC.AT.A..GTAAAA.G.A.AA.AACTG
Pectinaria koreni           GGGG..A.AT.TTA.CTTTTTTCTA.A.TTAA.AAGAGAAC.CAGCT..G.CGGCTG.GG...A..AA.T..TCCA..T.
Sigalion mathildae          .AT.TT.A.CGAT..TGTT..GGCGACCCAGG..CATT.A.AC.CTCCTTTAT.CTC.AG.GC.AATCCCTCA..TCA.T
Chamelea gallina            .ATC....AGAAT..TT.TT.GG.CAAT.TAA.TCAA.A..ACGGTTTAATAACTAG.....GA.CC.TGTC.AAAGA..
Illex coindetii             .AT.T.A.ATATT.A..T.AT..TA.AT.AA.TTTAAT.GGG.TGATTAA..A.TA..G..TTA.ATA.TAAC...C.T.

                                    250       260       270       280       290       300       310       320
                            ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Alpheus glaber              AAAAAAATAATTATAATTAGTATGTGACCCTAAACTTTAGGTTGGAAGATCAAGTTACTTTAGGGATAACAGCGTTATTT
Crangon crangon             ...TT..A.TA..A.TA..TAG.T.A.TAAG.TC.C.......T.G....T..AAG...AAGTTAC.TTAG.GA.A.CAG
Eledone cirrhosa            .T.T.T..T.ACT.CC...TATATAA.TAT.T.TAGA..AA.AAACC..GAGTT..G...A.AA.TA.GTTA.CA..GGG
Goneplax rhomboides         ....TT.AT.CA.A...ATT.GGT.T.AAAGTT.AAGATCC..TTT.TGGATTT.A.GGC..A.TTACTTTAG.GATAAC
Hediste diversicolor        TGTCTT.G...ATAGC.ACC.TA.G..TAAC.GG..AATCC.ACT.GAGAG.TCAC.T.GACA.T.GGGGTTG.CACC.C
Lentidium mediterraneum     G.....G.T.CCG..GGG.TA.CAGCGTAA.TTT..ACT..AG..CTA...G.A.GGAAAGTTT.CG.C.TCGA.GT.GG
Liocarcinus depurator       .G..GT.ATCAATG.TA.TTGG.TAA.G.A...TTGA.CCT..TT..AGATT.AAG.T.AAGTTAC.TTAG.GA.A.CAG
Liocarcinus vernalis        ...G.C.ATCAATA.TAGTTGT.TAAGTTA...TTGA.CCT.CTT..AGATT.AAG.T.AAGTTAC.TTAG.GA.A.CAG
Meganyctiphanes norvegica   .T.TTTTA.TGATAGT.GTCAT.T.A.GAG.G.T.C.GTAT.ATAG.TTAT...A.TAAG.TACTT..GGGATAAC.GCG
Nephtys hombergii           .G.CCT..C.G.TAC.A.GACCC.GT.TTACCG.T.AA.A.AAAA.GCTA.C.TAGGGA.A.CA..CT.ATC.T.CTAGA
Pachygrapsus marmoratus     CTTT.TT.TTGG.C..AATT..ATGATTAG.GTGTAAC.AAAGATCCT.AAT.TAG.T..A.A.AT...GTTAC..TAGG
Pectinaria koreni           ..TC.G.ACT..GA.TAGC.C.AT.AGG.T.C..GCCC.ATC.TA..CTGACCCC.TA.AGGATC.A.GA.AATAGC.AC
Sigalion mathildae          T.G.TCC.TTA..G.TCC.CAGAA.A.A.TACC.TAGGGATAACAGGCT.ATCT.CTTAGAGA.CTC..ATTGA.A.GAA
Chamelea gallina            .T.GC..A.GC..CCGCAG.G..AAC.G.G.T.T.C..CTTAA.AG.TCGA.TTGAGAGAAG..TT.GCG.C.TCG..G.
Illex coindetii             G.TT.....TA....TAG.TA.A..A...AAT..TA.ATT.C.TAT.ATAT.GT.AC.A.AG..ATA.CAGCGTAAT..G

                                    330       340       350       360       370       380       390       400
                            ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Alpheus glaber              CTCTTGAGAGTTCTTATCGACAGAGATGTTTGCGACCTCGATGTTGAATTAAAGTTTTACCTAGGAGTAGTCGTTTAGGT
Crangon crangon             .ATAATTTTCC.TGAGAGTT.TT.TCGACAG.A.TAG.T.GACC.CG..GTN.A..AA.GTCTCCTC.T
Eledone cirrhosa            A.AACAG.GTAAT..T.TT.G....T.C..ATTA.AAAAA.A.A.TGCGACCTCGA.GTTGG.TT.AA.T.ACC..T..G
Goneplax rhomboides         AG.G.T.TTC..T...AGAGTTC.T..CGAAAAA.GAGTTTGCGACCTCG.TGT.GAATTAA.AT.TCTA.ATAG.GCAG
Hediste diversicolor        GATG.TG.CT.AG.GTAACTATT..GC.CAGC..TTAATA.A..N.GT.GGTCAACCNTTAA.TCCT.CCGT.GA.CT.A
Lentidium mediterraneum     A.TAAAGCTCC.A..TGGCGT
Liocarcinus depurator       .GTAATTTTT.CTGAGAGTC..T.TCGAAGAAAA.GT.T.CGACCTCGA.GTT.AA...AAAT.TCT.TA.AA.GC..CA
Liocarcinus vernalis        .GTAATTTTT.CTGAGAGCC..T.TCGAAGAAAA.GT.T.CGACCTCGA.GTT.AA...AAAT.TCT.TA.AA.GC..CA
Meganyctiphanes norvegica   TAA..TTCTT.AAGAG.TCTT.TCT.CAAGAAA..TTG...CC.C..TG.TG.A..AA.AG.TCTTTA..GT..AGCA.C
Nephtys hombergii           GAGCCC.C.T.GA.AGAA.GG.TT.GCACC.CGATGT.G.C.TAGA.TACCCTCCGG.G.AGCA.CCCG.GAAAAGGTTG
Pachygrapsus marmoratus     GATAAC..C...A..T.TTTTGAGAG.TC..ATCGAAAAA.AAG.TTGCG.CCTCGA.GTTG.ATTAA.A.TTC..TACA
Pectinaria koreni           ...GG.GATAACAGAT.AAT.CTTCT..AGA.TTC..ATTGACAGA.GGGCTT.ACACCT.G.T.TTGGC.TAGGATATC
Sigalion mathildae          GGTC..GC.CC..GATGTTGGCTTAGG.CACC.TCTAAATCC.AA.GT.AG.G.GG.AGTTGGTC.ACTT.TAAGGCCC
Chamelea gallina            TGGA.T.T..AAAC.TCTTGGT.CAGCAGC.AT..TGGT..GAC..TTCGTCTT..AATT...T
Illex coindetii             T.TAGAGAGT.CT.AT.GA.A.ACA.GA..GCGAC.TCGATGT.G..T.A..GTAACCTTAAG.TGCAGAAGC.....TA

                                    410       420       430       440       450       460       470       480
                            ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Alpheus glaber              GGTAGGTCTGTTCGACCTTTAAAACTTTACA
Crangon crangon
Eledone cirrhosa            T.C..AAGCTCAGATAGG.G..TCTG..CG.TTTTTAAAATTTTAC
Goneplax rhomboides         CTGCTA.AAAAGAAGGTC.GTTCGACC.TT.AAATTTTACGTGATCTGAGTTCAGACC
Hediste diversicolor        .TCCAA
```

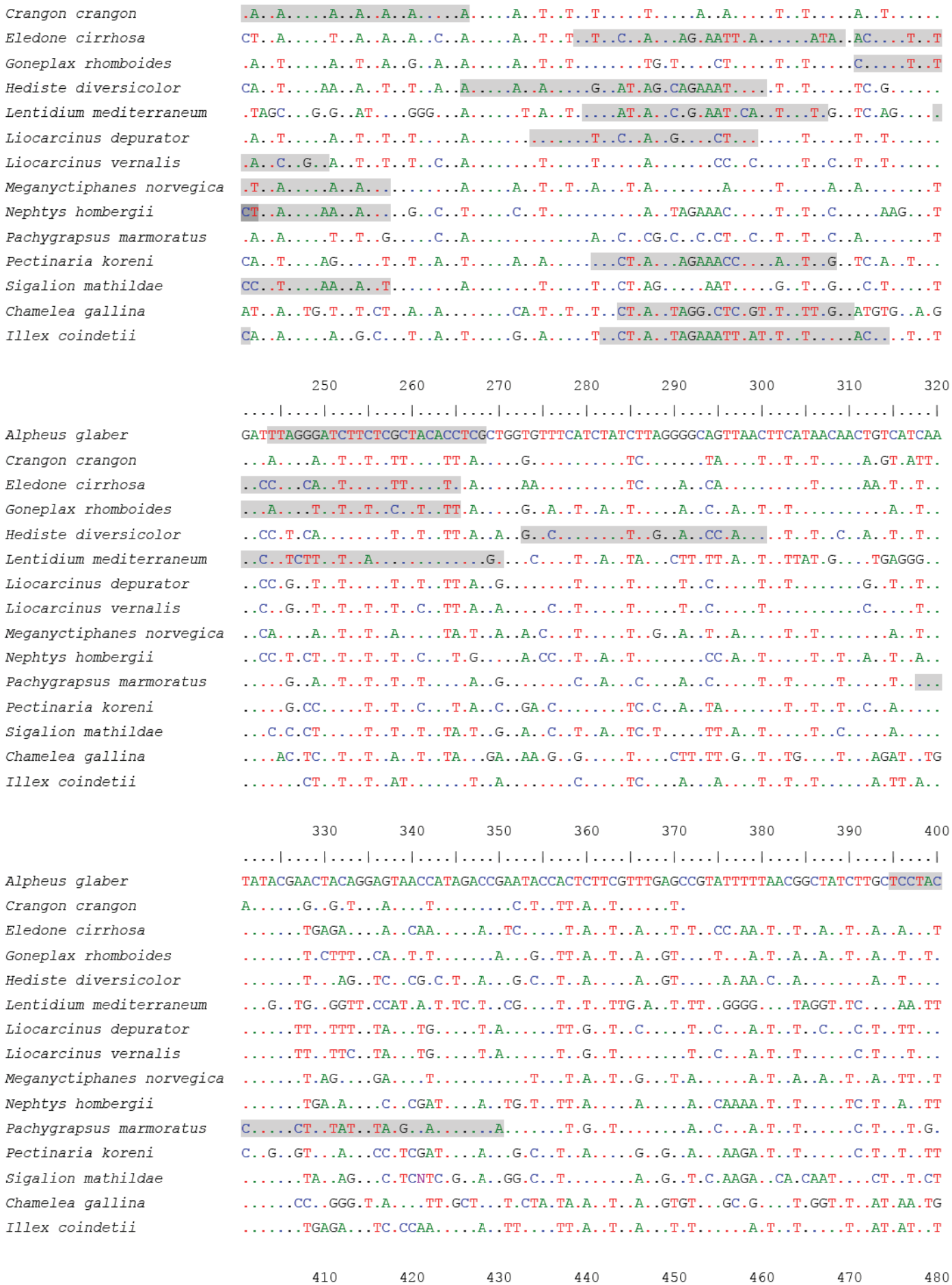
```
Lentidium mediterraneum
Liocarcinus depurator       .T..TAAAA.A
Liocarcinus vernalis        .T..TAAAA.AAA.T.TG..CG.CT....A.TTTTTACGTGATCTGAGTTCAGACCGGA
Meganyctiphanes norvegica   TT..ATAGAAGGTCTGT.CG.CCTT.AA.TTTTTACGTGATCTGAGTTCAGACCGG
Nephtys hombergii           ...T....AACCAT.AAAA.TCTC.G.G.TCTGAGTAAA
Pachygrapsus marmoratus     AT.GCAGGA...GT.AGAGA.GGT..G.T.GACCTTTAAAATTTTACGTGATCTGAGTTCAGACCGGA
Pectinaria koreni           CTGGA.GT..CAGA.G...CC..CGG.GGGTTTGTTCAACCCTTAAAATCCCTACGTGATCTGATCTGAGTTCAGACCGG
Sigalion mathildae
Chamelea gallina
Illex coindetii             A.GTAAATCTG.TCGAT...T......
```

8.2.2. COI alignment and oligonucleotides

```
                                10        20        30        40        50        60        70        80
                                ....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Alpheus glaber                  GTGATAATTTTTTTTATAGTTATGCCAATCATAATTGGGGGGTTTGGAAACTGACTTGTCCCCCTAATATTAGGTGCCCC
Crangon crangon                 ..A........C...........A..T..TC.T.....A..A........T...T.A..G..TT..........A..A..
Eledone cirrhosa                ..A........C...T.......A...G.T........T..A..C.....T...T.A......T..........A..T..
Goneplax rhomboides             ..A..................A.....C..T........A..A........T.....A.....AT..........A.....
Hediste diversicolor            T.A........C..CT.......A...G.A.....C.....A.....T..T......A....AT.....C.G..G.....
Lentidium mediterraneum         ..T...........CT.G........T..A.....A..C...........T...T.A..T..TA.....C..TCC.TT..
Liocarcinus depurator           ..A.....C.........G..A..A..T..T...........A.....T..T..............T..G.....A..T..
Liocarcinus vernalis            ..A.........................T..T........A..A.....T..T...........T......C..........
Meganyctiphanes norvegica       ..A........C.....G.....A.....T........A..A...........GT.A..T..G..T........A..A..
Nephtys hombergii               T.A........C..CT....A..A...G.A.....C..A.....................T..AT.....C....G.....
Pachygrapsus marmoratus         ..T.....C..............A..G..T........T..A...........G.....T..TT.G........A.....
Pectinaria koreni               C.T.......C...C.T..C..A..CG..T.T.....T.....................AT.....C.T.C...T..
Sigalion mathildae              T.A........C...C.T..A..A.....TC.TG.G..A..A..............A....TT.....C.T..G..T..
Chamelea gallina                ..A........C...C....A..A..T..A...........A.....G.....GT.A..T..TT....GC..ACAATA..
Illex coindetii                 A.T...........C........A..T..T........A........T..T...T.A..A..............A..T..

                                90        100       110       120       130       140       150       160
                                ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Alpheus glaber                  GGATATGGCCTTCCCCCGAATAAACAATATAAGATTCTGGCTCCTGCCCCCCTCTCTAACTCTTCTCTTATCTAGGGGGC
Crangon crangon                 T.....A..T.....A..T.....T..............AT.GT.A..T..TG..T..........TC.....A..AT
Eledone cirrhosa                A.....A..A.....A........T..............A..T..C.....T..C..C...T.A..TC.CA.ATCT.CCG
Goneplax rhomboides             T.....A.....T..T..T.....T..C..G.....T..AT.A..T..T......T.......CT.AC..ATA..A..A.
Hediste diversicolor            T.....A..A.....A..CC.T..T..............AT.AT.A..A..A..C.....C...........A..A.CAG
Lentidium mediterraneum         .........T..T..A........T...T....T..T..AA.T..T..TTG....T.CGG.T.AT.AAC..G.TCTCTTT
Liocarcinus depurator           T..C..A...........T.....T........G.....AT.A..C..A..T..G............C.CATA..A..CA
Liocarcinus vernalis            T.....A..T........T.....T..C..G.....T..AT.G..A........G.....G.....TC.CATG..A..TA
Meganyctiphanes norvegica       T.....A..A..T..A.....G..T..C........T..AT.AT.A..A..T...T........T.A...GGC..A..T.
Nephtys hombergii               T.....A..T.....T..TT....T......TCT.....A..T..T.....T...T...T......TG.TATATCC.CAG
Pachygrapsus marmoratus         A.....A..T..T..G........T...........T..A..TT.A..T.....CT..T....CT.AC.TA.A..AA.AA
Pectinaria koreni               A..C.....A..T..............C.....T.....A..T..C..T..GG.C..G.T......TC.GAGCTCC.CTG
Sigalion mathildae              T..C.....T.....A..TC.T..T........G..T..A..A..C..T..T.......T......T.....TCAA.AG
Chamelea gallina                A..C..A..A.....T..T.....T...C.T..G.....AT.GT....GGTG...A..TTAT.A..A...GG.TCT.CTT
Illex coindetii                 A.....A..T.....G..T.....T...........T..AT.A..T..A..A...T.....A.AT.AC..G.CTCTTCAG

                                170       180       190       200       210       220       230       240
                                ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Alpheus glaber                  TGGTGGAAAGCGGCGTCGGCACTGGGTGAACTGTCTACCCCCCTCTGTCAGCTGGGATTGCCCACGCTGGGGCCTCAGTA
```


Crangon crangon
Eledone cirrhosa
Goneplax rhomboides
Hediste diversicolor
Lentidium mediterraneum
Liocarcinus depurator
Liocarcinus vernalis
Meganyctiphanes norvegica
Nephtys hombergii
Pachygrapsus marmoratus
Pectinaria koreni
Sigalion mathildae
Chamelea gallina
Illex coindetii
250 260 270 280 290 300 310 320
Alpheus glaber
Crangon crangon
Eledone cirrhosa
Goneplax rhomboides
Hediste diversicolor
Lentidium mediterraneum
Liocarcinus depurator
Liocarcinus vernalis
Meganyctiphanes norvegica
Nephtys hombergii
Pachygrapsus marmoratus
Pectinaria koreni
Sigalion mathildae
Chamelea gallina
Illex coindetii
330 340 350 360 370 380 390 400
Alpheus glaber
Crangon crangon
Eledone cirrhosa
Goneplax rhomboides
Hediste diversicolor
Lentidium mediterraneum
Liocarcinus depurator
Liocarcinus vernalis
Meganyctiphanes norvegica
Nephtys hombergii
Pachygrapsus marmoratus
Pectinaria koreni
Sigalion mathildae
Chamelea gallina
Illex coindetii
410 420 430 440 450 460 470 480

```
                                ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Alpheus glaber                  TCAGGCTCCCAGTTCTAGCCGGAGCGATTACCATGCTATTAACTGATCGAAATTTAAACACCGCCTTCTTCGACCCTGCT
Crangon crangon
Eledone cirrhosa                .ATCA.........T....T.....A.....A..AT.......C.....T.....T......A.T..T..T..T..AAGA
Goneplax rhomboides             .TTCCT.A......T....A.....T.....T..A..T.....A.................TT.A.....T..T......
Hediste diversicolor            ....AT.A..C..AT....T.....A..C..T..A..TC.C....
Lentidium mediterraneum         .GGC.A.G..C.....T..T..T..TC.C..A..A..GC.C..A.........A.......TT.A..T..T..T....T.
Liocarcinus depurator           .ATCCT.A..T..CT....A..T..C........A..TC.T.........T.....G..T..TT.T...........C...
Liocarcinus vernalis            .ATCC..A..T...T....A..T..T.....T.....TC.T.....C..T.........T..TT.T..............
Meganyctiphanes norvegica       .ATCA..T......T....A.....T.....T..A..TC....A.....T...C.T.....AT.A..T..T.........
Nephtys hombergii               .ATCC
Pachygrapsus marmoratus         .TTCCT.A..T..AT....A..C..T.....T..AT..........C..T..C.....T..TT.A.....T..T......
Pectinaria koreni               .ATCA..T..T...T....A.....C.....T..A
Sigalion mathildae              ATCCCTC..TG..CT.G.AG..G..A.......ATACTCCCT.ACTGAACCGAAATCTT.AAA..GAA...TTTGACCTG
Chamelea gallina                .TGCTA.A..T..AT.G..T.CT..TT.G..A..AT.......A..C........T..T..AT.T.....T.......T.
Illex coindetii                 .ATCC..G..T...T....T.....T.....A..AT..C....A...........T..T...A.T..T..T..T..CAGA

                                        490       500       510       520       530       540
                                ....|....|....|....|....|....|....|....|....|....|....|....|
Alpheus glaber                  GGGGGTGGAGACCCTATCTTATATCAACATTTATTTTGATTTTTTGGTCACCCTGGAAGTTTAAA
Crangon crangon
Eledone cirrhosa                ..A..A........A..TC..............................
Goneplax rhomboides             ..A..A.....T...G.TC.G..C.............................
Hediste diversicolor
Lentidium mediterraneum         ..TTTA...........GC...TCGTC...C.T..C.................G
Liocarcinus depurator           .....A..T......G.TC.T.............C.....................A.GT..A
Liocarcinus vernalis            .....A..T.....AG.TC.C.....G..C..G.....................
Meganyctiphanes norvegica       ..T........T..A..TC...........C.........................A.....
Nephtys hombergii
Pachygrapsus marmoratus         .....G..T......G..C.C..C...........C..................
Pectinaria koreni
Sigalion mathildae              CA..AA..G.GGGACCCA..C.
Chamelea gallina                ...TTA..G..T..AG.G....T.ATT.....G.....................A.A.GT..A
Illex coindetii                 ..T..A...........T...........C.....................
```

8.3. Phylogenetic trees of the target sequences

8.3.1. 16S Phylogenetic tree

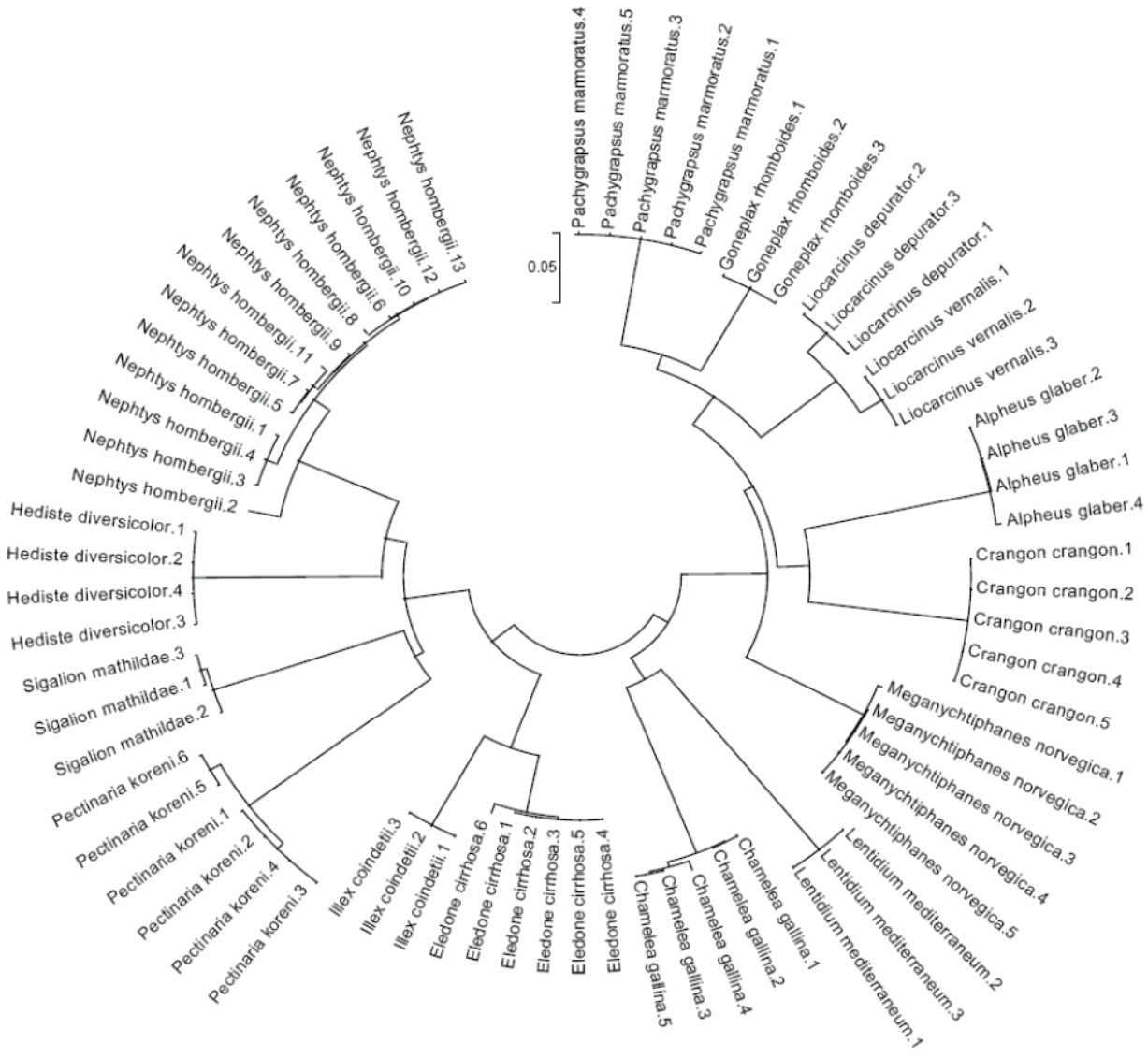


Fig. 31: Phylogenetic tree of the 16S sequences

8.3.2. COI Phylogenetic tree

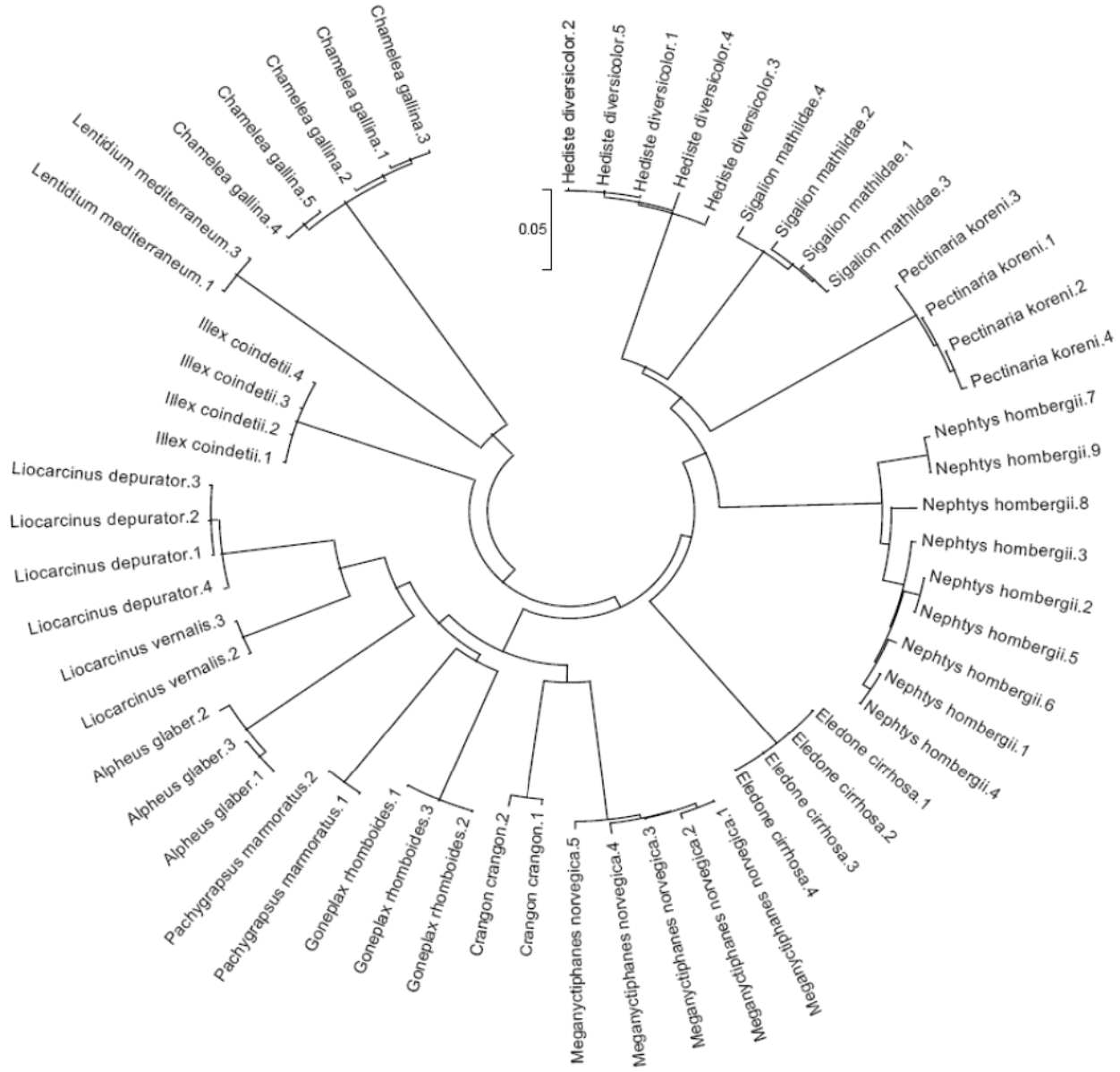


Fig. 32: Phylogenetic tree of the COI sequences

8.4. Layout of the Microarrays

8.4.1. 16S Microarray

PK	NK	NK	PK
Cg_COI_b	Empty	Empty	Ic_COI_b
Ec_COI_b	Gr_COI_b	Lm_16S_b	Lm_COI_b
Empty	Sm_COI_a	PK	Ec_16S_b
Nh_COI_b	Pk_16S_b	Pk_COI_c	Sm_16S_b
Hd_COI_a	PK	Nh_16S_b	Empty
Mn_COI_b	Pm_16S_b	Pm_COI_b	Hd_16S_b
Empty	Lv_COI_b	PK	Mn_16S_b
Cc_COI_a	Ld_16S_b	Ld_COI_b	Lv_16S_b
Ag_COI_b	PK	Cc_16S_b	Empty
Ic_16S_a	Ic_COI_a	Empty	Ag_16S_b
Empty	Cg_16S_a	PK	Cg_COI_a
Gr_COI_a	Lm_16S_a	Lm_COI_a	Lm_16S_c
Ec_COI_a	PK	Gr_16S_a	Empty
Sm_16S_a	Sm_COI_c	Sm_COI_b	Ec_16S_a
Empty	Pk_16S_c	PK	Pk_COI_a
Nh_16S_a	Nh_COI_a	Pk_16S_a	Pk_COI_d
Hd_COI_c	PK	Hd_COI_b	Empty
Mn_COI_a	Pm_16S_a	Pm_COI_a	Hd_16S_a
Empty	Lv_COI_a	PK	Mn_16S_a
Cc_COI_b	Ld_16S_a	Ld_COI_a	Lv_16S_a
Empty	PK	Cc_16S_c	Empty
Ag_16S_a	Ag_COI_a	Cc_16S_a	Cc_COI_c
PK	NK	NK	PK

8.4.2. COI Microarray

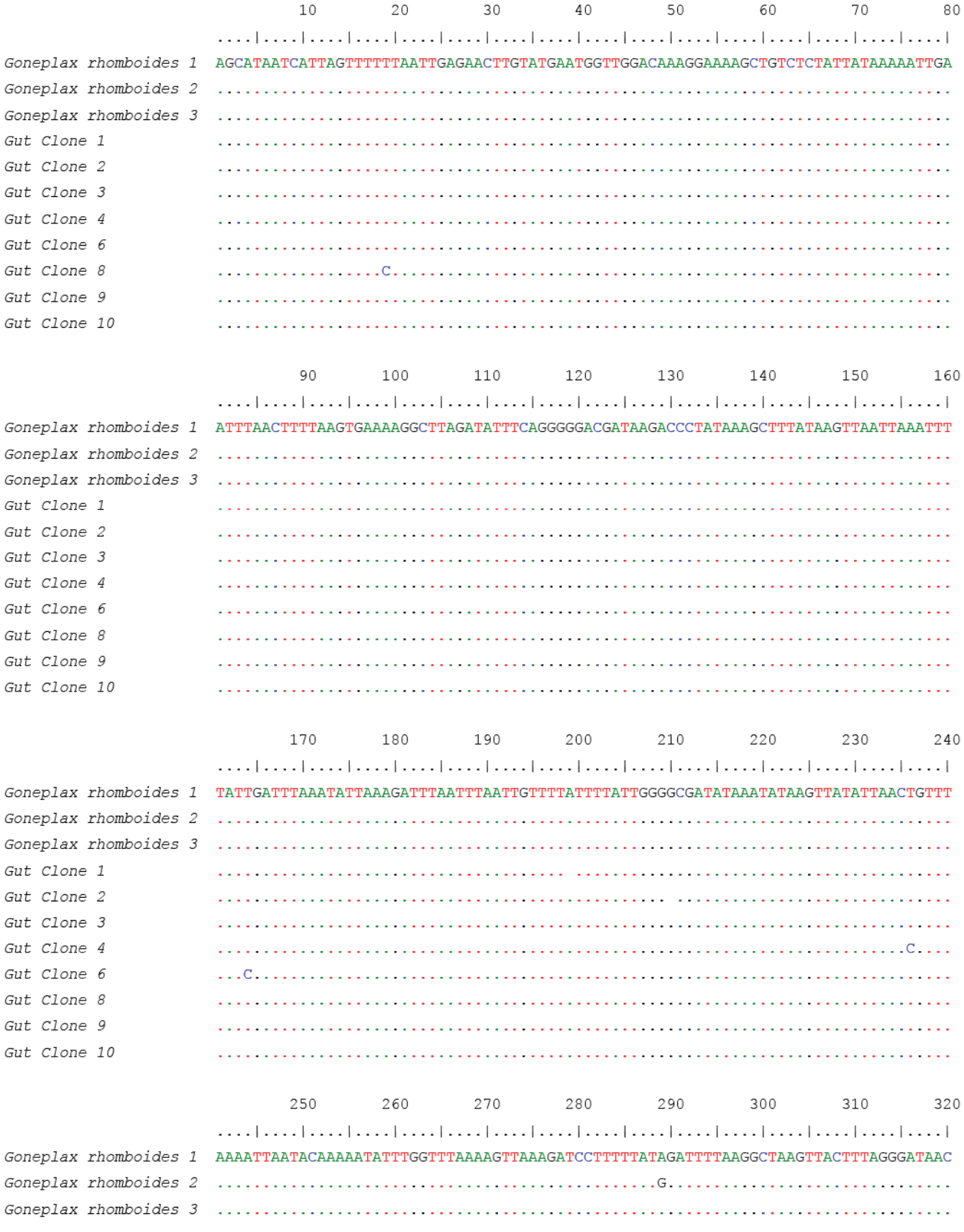
Pm_COI_b	PK	PK	NK
Lv_COI_b	Mn_COI_a	Mn_COI_b	Pm_COI_a
Gr_COI_b	Ld_COI_a	Ld_COI_b	Lv_COI_a
Ag_COI_b	Cc_COI_a	Cc_COI_b	Gr_COI_a
Pk_COI_b	Sm_COI_a	Sm_COI_b	Ag_COI_a
Hd_COI_b	Nh_COI_a	Nh_COI_b	Pk_COI_a
Cg_COI_b	Ic_COI_a	Ic_COI_b	Hd_COI_a
Ec_COI_b	Lm_COI_a	Lm_COI_b	Cg_COI_a
Mn_COI_b	Pm_COI_a	Pm_COI_b	Ec_COI_a
Ld_COI_b	Lv_COI_a	Lv_COI_b	Mn_COI_a
Cc_COI_b	Gr_COI_a	Gr_COI_b	Ld_COI_a
Sm_COI_b	Ag_COI_a	Ag_COI_b	Cc_COI_a
Nh_COI_b	Pk_COI_a	Pk_COI_b	Sm_COI_a
Ic_COI_b	Hd_COI_a	Hd_COI_b	Nh_COI_a
Lm_COI_b	Cg_COI_a	Cg_COI_b	Ic_COI_a
Pm_COI_b	Ec_COI_a	Ec_COI_b	Lm_COI_a
Lv_COI_b	Mn_COI_a	Mn_COI_b	Pm_COI_a
Gr_COI_b	Ld_COI_a	Ld_COI_b	Lv_COI_a
Ag_COI_b	Cc_COI_a	Cc_COI_b	Gr_COI_a
Pk_COI_b	Sm_COI_a	Sm_COI_b	Ag_COI_a
Hd_COI_b	Nh_COI_a	Nh_COI_b	Pk_COI_a
Cg_COI_b	Ic_COI_a	Ic_COI_b	Hd_COI_a
Ec_COI_b	Lm_COI_a	Lm_COI_b	Cg_COI_a
PK	NK	Ec_COI_a	PK

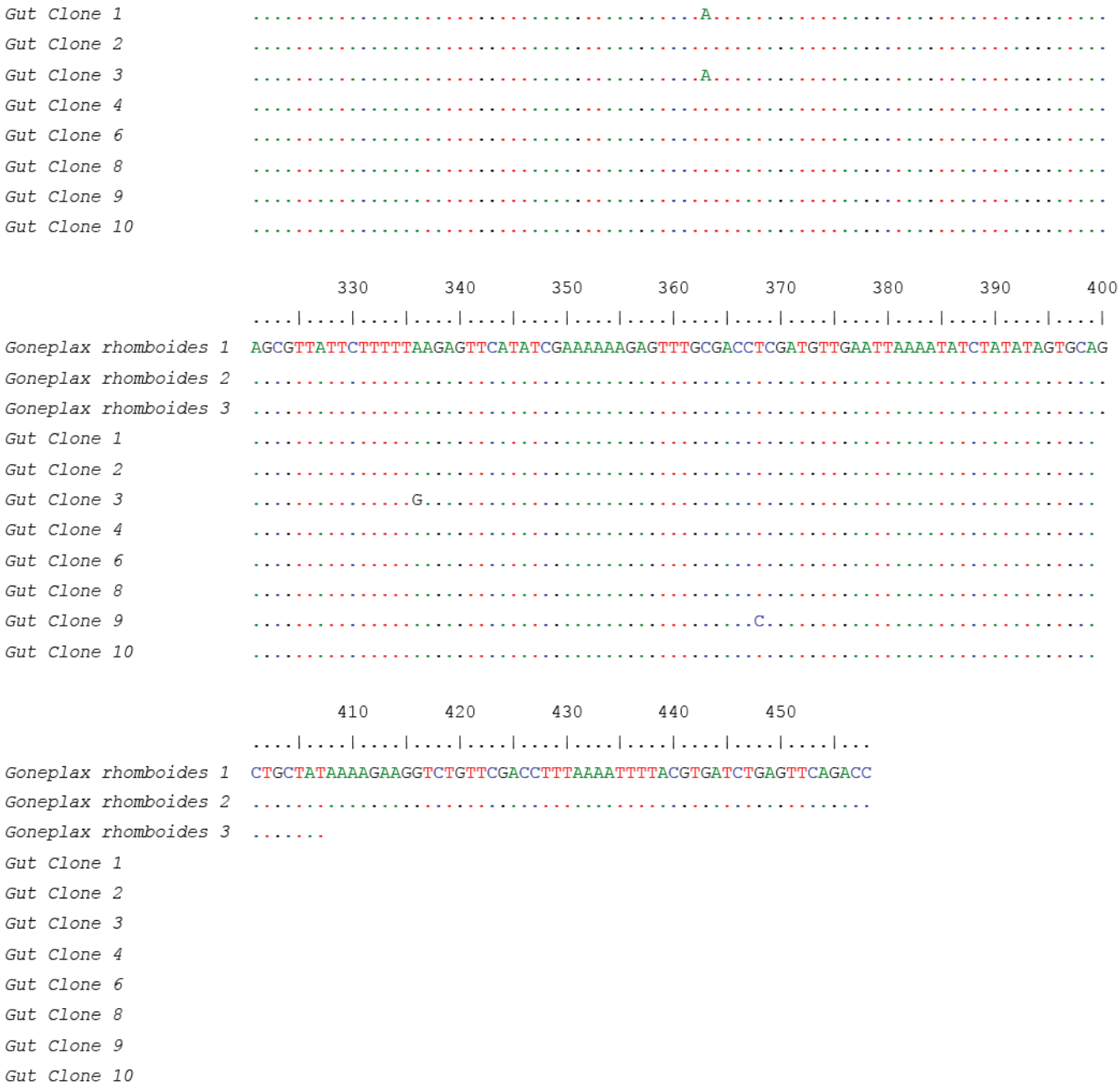
8.4.3. Invertebrate Microarray

NK	PK	PK	NK
Mn_16S_b	Pm_16S_a	Pm_16S_b	Lio_16S_a
Ld_16S_b	Lv_16S_a	Lv_16S_b	Mn_16S_a
Cc_COI_c	Cc_COI_d	Gr_16S_a	Ld_16S_a
Ag_16S_b	Cc_16S_a	Cc_16S_b	Cc_16S_c
Sm_COI_d	Sm_COI_c	PK	Ag_16S_a
Pk_COI_c	Sm_16S_a	Sm_16S_b	Sm_COI_e
Pk_16S_a	Pk_16S_b	Pk_16S_c	Pk_COI_d
Nh_16S_a	Nh_16S_b	Nh_COI_c	Nh_COI_d
Hd_16S_b	Hd_COI_c	Hd_COI_d	Hd_COI_e
Lm_16S_c	Cg_16S_a	Ic_16S_a	Hd_16S_a
Ec_16S_a	Ec_16S_b	Lm_16S_a	Lm_16S_b
Mn_16S_b	Pm_16S_a	Pm_16S_b	Lio_16S_a
Ld_16S_b	Lv_16S_a	Lv_16S_b	Mn_16S_a
Cc_COI_c	Cc_COI_d	Gr_16S_a	Ld_16S_a
Ag_16S_b	Cc_16S_a	Cc_16S_b	Cc_16S_c
Sm_COI_d	Nh_COI_c	PK	Ag_16S_a
Pk_COI_c	Sm_16S_a	Sm_16S_b	Sm_COI_e
Pk_16S_a	Pk_16S_b	Pk_16S_c	Pk_COI_d
Nh_16S_a	Nh_16S_b	Sm_COI_c	Nh_COI_d
Hd_16S_b	Hd_COI_c	Hd_COI_d	Hd_COI_e
Lm_16S_c	Cg_16S_a	Ic_16S_a	Hd_16S_a
Ec_16S_a	Ec_16S_b	Lm_16S_a	Lm_16S_b
PK	NK	NK	PK

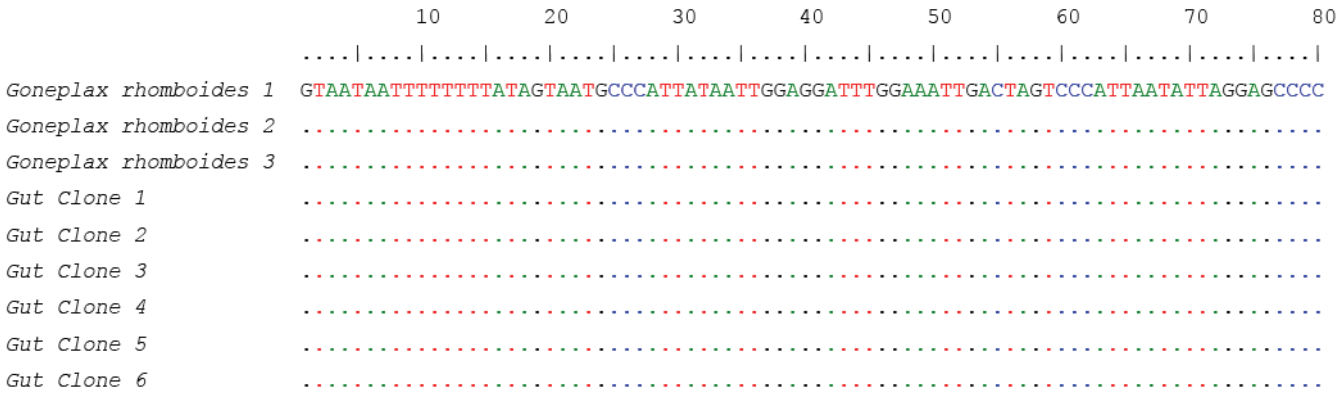
8.5. Alignment of the cloned gut sequences to the target sequences

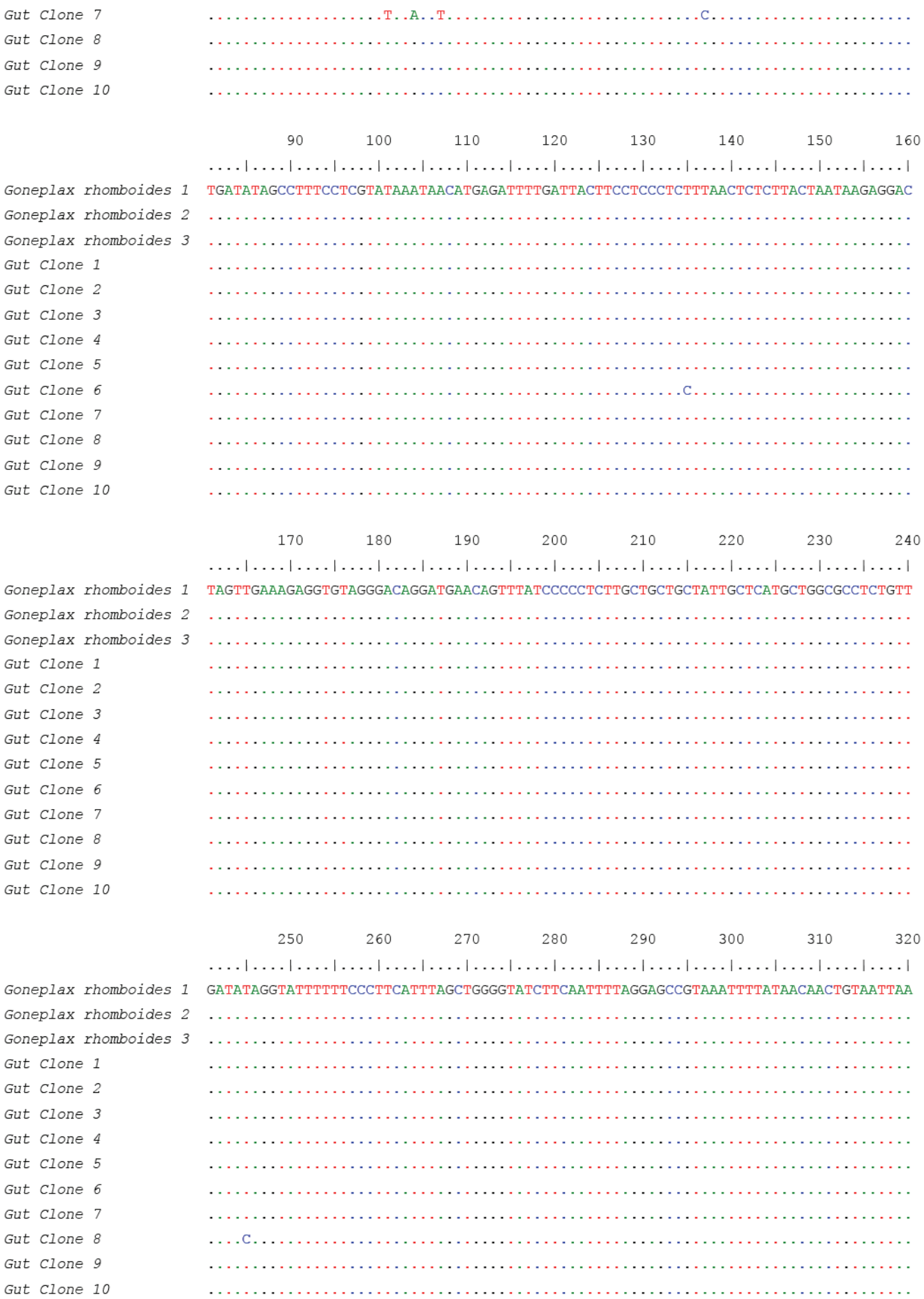
8.5.1. 16S alignment

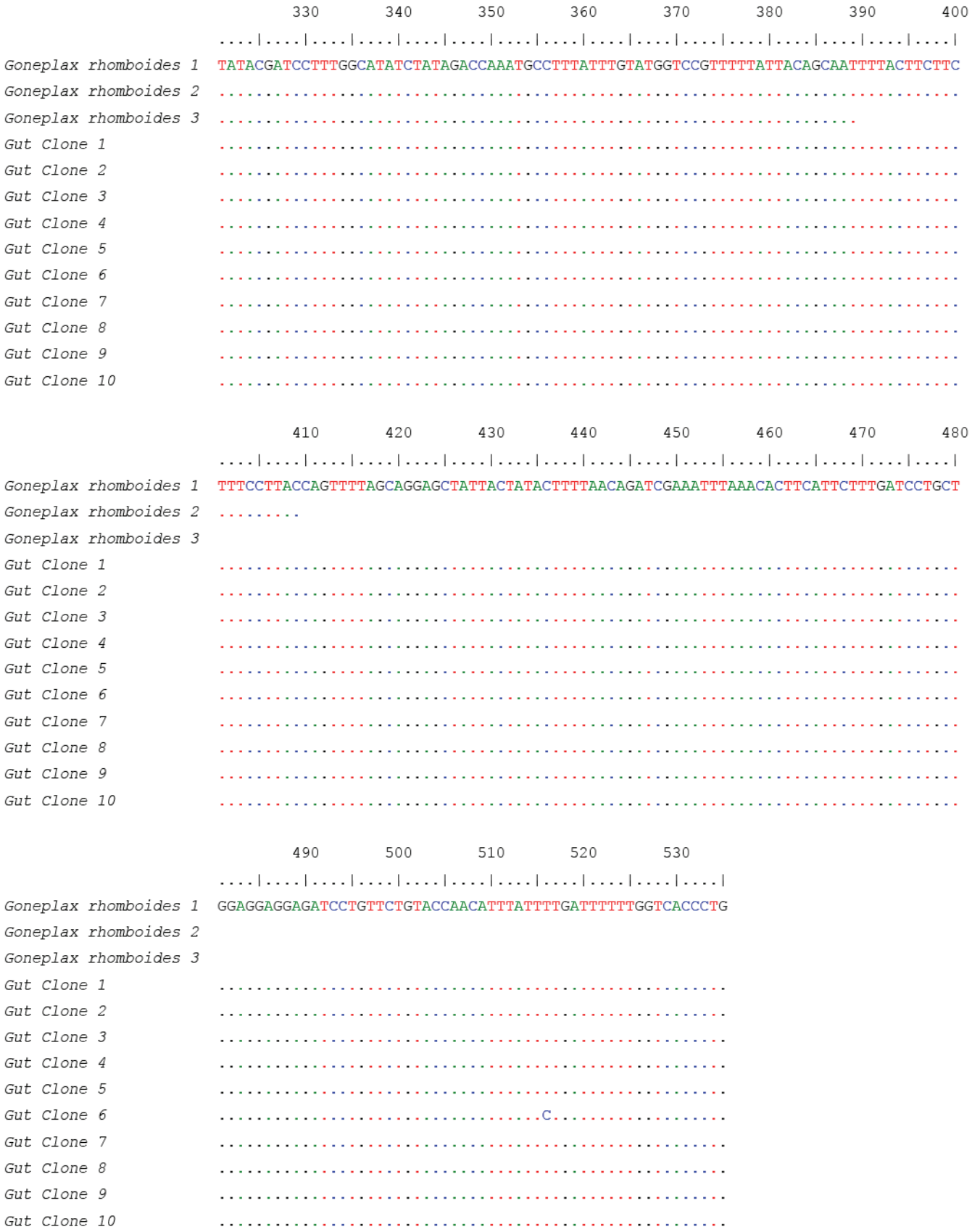




8.5.2. COI alignment




Gut Clone 7 .....T..A..T.....C.....
Gut Clone 8 .....
Gut Clone 9 .....
Gut Clone 10 .....
90 100 110 120 130 140 150 160
Goneplax rhomboides 1 TGATATAGCCTTTCCTCGTATAAATAACATGAGATTTTGATTACTTCCTCCCTCTTTAACTCTCTTACTAATAAGAGGAC
Goneplax rhomboides 2 .....
Goneplax rhomboides 3 .....
Gut Clone 1 .....
Gut Clone 2 .....
Gut Clone 3 .....
Gut Clone 4 .....
Gut Clone 5 .....
Gut Clone 6 .....C.....
Gut Clone 7 .....
Gut Clone 8 .....
Gut Clone 9 .....
Gut Clone 10 .....
170 180 190 200 210 220 230 240
Goneplax rhomboides 1 TAGTTGAAAGAGGTGTAGGGACAGGATGAACAGTTTATCCCCCTCTTGCTGCTGCTATTGCTCATGCTGGCGCCTCTGTT
Goneplax rhomboides 2 .....
Goneplax rhomboides 3 .....
Gut Clone 1 .....
Gut Clone 2 .....
Gut Clone 3 .....
Gut Clone 4 .....
Gut Clone 5 .....
Gut Clone 6 .....
Gut Clone 7 .....
Gut Clone 8 .....
Gut Clone 9 .....
Gut Clone 10 .....
250 260 270 280 290 300 310 320
Goneplax rhomboides 1 GATATAGGTATTTTTTCCCTTCATTTAGCTGGGGTATCTTCAATTTTAGGAGCCGTAAATTTTATAACAACTGTAATTAA
Goneplax rhomboides 2 .....
Goneplax rhomboides 3 .....
Gut Clone 1 .....
Gut Clone 2 .....
Gut Clone 3 .....
Gut Clone 4 .....
Gut Clone 5 .....
Gut Clone 6 .....
Gut Clone 7 .....
Gut Clone 8 ....C.....
Gut Clone 9 .....
Gut Clone 10 .....


330 340 350 360 370 380 390 400
Goneplax rhomboides 1 TATACGATCCTTTGGCATATCTATAGACCAAATGCCTTTATTTGTATGGTCCGTTTTTATTACAGCAATTTTACTTCTTC
Goneplax rhomboides 2
Goneplax rhomboides 3
Gut Clone 1
Gut Clone 2
Gut Clone 3
Gut Clone 4
Gut Clone 5
Gut Clone 6
Gut Clone 7
Gut Clone 8
Gut Clone 9
Gut Clone 10
410 420 430 440 450 460 470 480
Goneplax rhomboides 1 TTTCCTTACCAGTTTTAGCAGGAGCTATTACTATACTTTTAACAGATCGAAATTTAAACACTTCATTCTTTGATCCTGCT
Goneplax rhomboides 2
Goneplax rhomboides 3
Gut Clone 1
Gut Clone 2
Gut Clone 3
Gut Clone 4
Gut Clone 5
Gut Clone 6
Gut Clone 7
Gut Clone 8
Gut Clone 9
Gut Clone 10
490 500 510 520 530
Goneplax rhomboides 1 GGAGGAGGAGATCCTGTTCTGTACCAACATTTATTTTGATTTTTTGGTCACCCTG
Goneplax rhomboides 2
Goneplax rhomboides 3
Gut Clone 1
Gut Clone 2
Gut Clone 3
Gut Clone 4
Gut Clone 5
Gut Clone 6
Gut Clone 7
Gut Clone 8
Gut Clone 9
Gut Clone 10

8.6. Microscopical validation of the gut content analysis

Strategia alimentare di *Trigla lucerna* (Teleostei, Triglidae) in Alto-Medio Adriatico: note di 7 campagne di pesca a strascico

M. Stagioni[1,2], E. Mazzoni[1], S. Montanini[1], M. Vallisneri[2]

[1]*Laboratorio di Biologia Marina e Pesca di Fano (PU), Università di Bologna, Italia.*
[2]*Dipartimento di Biologia Evoluzionistica Sperimentale, Università di Bologna, Italia.*

Introduzione

Nell'ambito della gestione delle risorse alieutiche, di particolare risalto sono gli studi relativi alle strategie alimentari, ai rapporti preda-predatore ed ai livelli trofici, essenziali per quantificare gli effetti della pesca sull'ecosistema. Obiettivo di questo studio, svolto nell'ambito di un progetto ministeriale MIPAAF, è la messa a punto di un modello di analisi della strategia alimentare di una "specie target" del Mediterraneo centrale: *Trigla lucerna* (Linnaeus, 1758), detta cappone o gallinella che presenta una distribuzione omogenea nelle acque costiere del versante italiano dell'Adriatico centro settentrionale.

Materiali e metodi

Nel corso degli anni 2005 e 2006, sono stati effettuati 204 campionamenti in Alto-Medio Adriatico fra 10 e 260 m di profondità (Figura 1). Su un totale di 1096 esemplari sono stati registrati: lunghezza totale (LT, mm), peso corporeo (P, g), sesso (diagnosi macroscopica, distinguendo tra femmine "F", maschi "M", indeterminati "I" e non determinabili "ND"). E' stato valutato lo stato di riempimento degli stomaci e sono state identificate le prede (fino al livello tassonomico più basso possibile), contate, pesate e fotografate digitalmente. I dati sono stati informatizzati mediante database relazionale "Microsoft Access" e analizzati mediante "pacchetto statistico R". E' stata condotta un'analisi dei cluster sulla taglia, utilizzando come misura di similarità le prede. Sono stati calcolati i principali indici di abbondanza, al fine di individuare le variazioni del regime trofico in rapporto al sesso, alla taglia, alle stagioni, alla profondità.

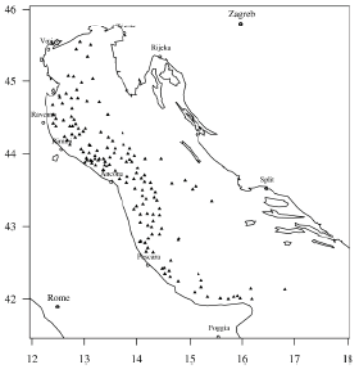


Figura 1 - Cartina dei siti di cattura di *Trigla lucerna*.

Risultati

La taglia è compresa fra 63 e 415 mm, con media di 185 mm (Figura 2).

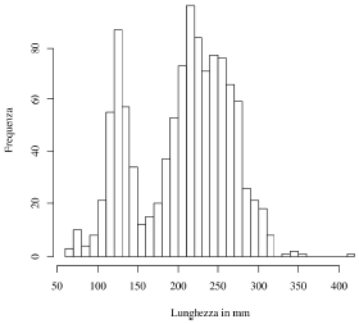


Figura 2 – Distribuzione per taglia di *Trigla lucerna*.

L'attività alimentare è maggiore nei giovanili (<180mm), durante l'estate e a basse profondità, rispecchiando la distribuzione e l'abbondanza delle reclute. L'analisi qualitativa dei contenuti stomacali ha permesso l'identificazione di 55 specie, appartenenti a 6 categorie sistematiche principali (tabella 1). L'analisi di similarità, basata sui contenuti stomacali delle varie classi di taglia, rivela un differente comportamento alimentare, separando gli individui giovani da quelli adulti (Figura 3). L'analisi quantitativa evidenzia che l'alimento preferito a tutte le taglie è costituito da Crustacea (89,7% del numero e 58% in peso) e Teleostea (6,7 % in numero totale e 39,7% in peso). Prede accessorie sono rappresentate da Mollusca (1,6 % in numero totale e 0,6% in peso).

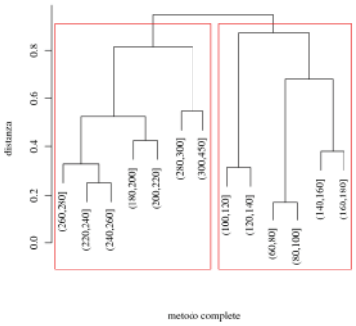


Figura 3 - Analisi dei Cluster sull'alimentazione di *Trigla lucerna* alle diverse taglie.

Tabella 1 - Principali categorie sistematiche predate e relativi indici di abbondanza: N(%) abbondanza numerica, W(%) abbondanza ponderale, O(%) frequenza di occorrenza, IRI(%) indice di importanza relativa espressi in percentuale.

Gruppi	N(%)	W(%)	O(%)	IRI(%)
Crostacei	89,69	58,01	58,01	92,24
Pesci	6,72	39,74	39,74	7,45
Bivalvia	1,13	0,23	0,23	0,03
Gastropoda	0,3	0,22	0,22	0,01
Cephalopoda	0,14	0,1	0,1	<0,01
Polychaeta	0,03	0,01	0,01	<0,01

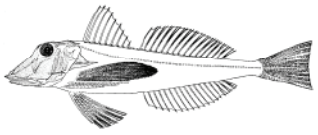
Nell'ambito Crustacea, le specie dominanti sono, tra i Decapoda reptantia i Brachiura *Goneplax rhomboidea* (14,2% del numero e 29,5% in peso) e *Liocarcinus* sp. (13% del numero e 24,7% in peso), tra i Decapoda natantia *Philocheras* sp. (59,5% del numero e 2,5% in peso). (Tabella 2)

Tabella 2 - Principali specie predate e relativi indici di abbondanza: N(%) abbondanza numerica, W(%) abbondanza ponderale, O(%) frequenza di occorrenza, IRI(%) indice di importanza relativa espressi in percentuale.

Specie	N(%)	W(%)	O(%)	IRI(%)
Goneplax rhomboides	14,16	29,52	29,52	50,3
Philocheras sp.	44,38	1,53	1,53	28,51
Liocarcinus depurator	5,66	20,32	20,32	10,47
Liocarcinus sp.	7,09	4,47	4,47	5,04
Philocheras bispinosus	15,1	0,96	0,96	3,04
Processa sp.	6,15	1,8	1,8	1,66
Engraulis encrasicolus	0,6	11,91	11,91	1,14
Gobius niger	0,6	8,77	8,77	0,76
Squilla mantis	0,43	2,95	2,95	0,24
Alpheus glaber	0,63	1,01	1,01	0,24
Solenocera membranacea	0,62	1,51	1,51	0,17
Lesuerigobius friesi	0,49	1,57	1,57	0,13
Cepola macrophthalma	0,13	2,01	2,01	0,05
Corystes cassivelaunus	0,11	1,79	1,79	0,04
Pachygrapsus marmoratus	0,41	0,32	0,32	0,03
Lesuerigobius suerii	0,20	0,00	0,00	0,03
Corbula gibba	0,34	0,19	0,19	0,03
Jaxea nocturna	0,21	0,51	0,51	0,03
Liocarcinus maculatus	0,23	0,43	0,43	0,02
Turritella communis	0,21	0,22	0,22	0,02
Trisopterus minutus	0,08	1,8	1,8	0,01
Arnoglossus laterna	0,08	1	1	0,01
Gobius sp.	0,08	0,52	0,52	0,01
Macropodia sp.	0,28	0,16	0,16	0,01
Serranus hepatus	0,04	0,79	0,79	0,01

Discussione

In relazione alla taglia, si delineano due modelli di "strategie alimentari": il primo, tipico dei giovanili, che si nutrono prevalentemente di crostacei, numerosi e di piccole dimensioni; il secondo, proprio degli adulti, che si nutrono prevalentemente di pesci (in numero minore ma di dimensioni maggiori), parallelamente all'accrescimento, al differenziamento sessuale ed al miglioramento della capacità predatoria. In sintesi, *T. lucerna* si presenta come "consumatore opportunista", caratterizzato da una elevata biodiversità delle prede; tuttavia, la comunità alimentare di cui si nutre, che rispecchia la comunità dell'area, risulta composta da un insieme di nicchie trofiche presumibilmente correlate a differenze ontogenetiche.



Bibliografia

Ardizzone G.D., Colloca F., Gravina M.F. (1994). Trophic ecology of gurnards (Pisces: Triglidae) in the Central Mediterranean Sea. Mar. Life, 4 (2): 45-57.
Colloca, F., G.D. Ardizzone and M.F. Gravina (1994). Trophic ecology of gurnards (Pisces: Triglidae) in the Central Mediterranean Sea. Marine Life 4, 45-57.
Froglia, C. (1976). Osservazioni sull'alimentazione dei giovani di Trigla lucerna della classe di età a nel Medio Adriatico (Pisces, Triglidae). Arch. Oceanogr. Limnol. 18 (suppl. 3), 365-373.

Si ringrazia il prof. Corrado Piccinetti per il supporto logistico e scientifico.

Studio finanziato dal MIPAAF (Dir. Gen. Pesca Marittima) nell'ambito del Progetto: "Struttura trofica di popolazioni demersali"

9. Disclaimer

Gemäß §6 der Promotionsordnung der Universität Bremen für die mathematischen, natur- und ingenieureswissenschaftlichen Fachbereiche vom 14. März 2007 versichere ich, dass

1. die Arbeit ohne unerlaubte fremde Hilfe angefertigt wurde,

2. keine anderen als die angegebenen Quellen und Hilfsmittel benutzt wurden,

3. die den benutzten Werken wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht wurden.

Bremen, 10. Juli 2008-07-09

(S. Chitipothu)