

# DNA Microarrays for Identifying Fishes

M. Kochzius · M. Nölte · H. Weber · N. Silkenbeumer ·  
S. Hjörleifsdóttir · G. O. Hreggvidsson ·  
V. Marteinson · K. Kappel · S. Planes · F. Tinti ·  
A. Magoulas · E. Garcia Vazquez · C. Turan ·  
C. Hervet · D. Campo Falgueras · A. Antoniou ·  
M. Landi · D. Blohm

Received: 26 September 2007 / Revised: 5 November 2007 / Accepted: 7 November 2007  
© The Author(s) 2008

**Abstract** In many cases marine organisms and especially their diverse developmental stages are difficult to identify by morphological characters. DNA-based identification methods offer an analytically powerful addition or even an alternative. In this study, a DNA microarray has been developed to be able to investigate its potential as a tool for the identification of fish species from European seas based on mitochondrial 16S rDNA sequences. Eleven commercially important fish species were selected for a first prototype. Oligonucleotide probes were designed based on the 16S rDNA sequences obtained from 230 individuals of 27 fish species. In addition, more than 1200 sequences of 380 species served as sequence background against which the specificity of the probes was tested *in silico*. Single target hybridisations with Cy5-labelled, PCR-amplified 16S rDNA fragments from each of the 11 species on microarrays containing the complete set of probes confirmed their suitability. True-positive, fluorescence signals obtained were at least one order of magnitude stronger than false-positive

cross-hybridisations. Single nontarget hybridisations resulted in cross-hybridisation signals at approximately 27% of the cases tested, but all of them were at least one order of magnitude lower than true-positive signals. This study demonstrates that the 16S rDNA gene is suitable for designing oligonucleotide probes, which can be used to differentiate 11 fish species. These data are a solid basis for the second step to create a “Fish Chip” for approximately 50 fish species relevant in marine environmental and fisheries research, as well as control of fisheries products.

**Keywords** DNA chips · Species identification · Capture oligonucleotide · Pisces

## Introduction

Compared with terrestrial ecosystems, little is known about marine biodiversity and changes in species richness and

---

M. Kochzius (✉) · H. Weber · N. Silkenbeumer ·  
K. Kappel · D. Blohm  
Centre for Applied Gene Sensor Technology (CAG),  
University of Bremen,  
Bremen, Germany  
e-mail: kochzius@uni-bremen.de

M. Nölte  
Zentrum für Technomathematik (ZeTeM), University of Bremen,  
Bremen, Germany

S. Hjörleifsdóttir · G. O. Hreggvidsson · V. Marteinson  
Prokaria,  
Reykjavík, Iceland

S. Planes · C. Hervet  
UMR CNRS EPHE 8046, Université de Perpignan,  
Perpignan, France

F. Tinti · M. Landi  
CIRSA, University of Bologna,  
Ravenna, Italy

A. Magoulas · A. Antoniou  
Institute of Marine Biology and Genetics,  
Hellenic Centre for Marine Research,  
Heraklion, Greece

E. Garcia Vazquez · D. Campo Falgueras  
Universidad de Oviedo,  
Oviedo, Spain

C. Turan  
College of Fisheries and Aquaculture, Mustafa Kemal University,  
Antakya, Turkey

ecosystem function. This is mainly because of sampling difficulties and problems in taxonomy. Many marine organisms and especially their diverse developmental stages, such as (1) eggs and larvae of fishes and invertebrates, (2) zoo- and phytoplankton, and (3) benthic invertebrates, are cumbersome and difficult to identify by morphological characters. Classical microscopy methods are time-consuming and require a high degree of taxonomic expertise, which is currently falling short. In many cases identifying a species is the major bottleneck in marine biodiversity and ecosystem research, hampering the necessary monitoring of marine biodiversity. As an example, a review of 138 studies on invertebrate diversity in European seas showed that approximately one-third of the specimens could not have been identified to species level (Schander and Willassen 2005).

DNA-based identification methods are meanwhile established (Barlow and Tzotzos 1995) as powerful tools, exhibiting an unprecedented accuracy because of their inherently highest possible resolution, which can reach even the level of single base changes in a whole genome. Using these methods, the following marine animals have been investigated: eggs, larvae, and adults of fishes (Rocha-Olivares 1998; Noell et al. 2001; Fox et al. 2005; Ward et al. 2005), planktonic copepods (Bucklin et al. 1999), invertebrate larvae (Garland and Zimmer 2002; Barber and Boyce 2006), and prey in the gut content or feces of penguins, whales, and fishes (Jarman et al. 2002; Saitho et al. 2003).

Sequences of the small subunit of the rRNA gene are used as a standard method for identifying microbial organisms (Ludwig et al. 2004), and a fragment of the mitochondrial cytochrome oxidase I gene is in use as “DNA-barcode” for the identification of metazoans (Hebert et al. 2002; Ward et al. 2005). As shown above, a growing number of recently published studies are using molecular genetic identification methods. Nevertheless, their application is still restricted mainly because of methodical problems and because special knowledge and experience in molecular genetics is required. This is especially true if DNA-based identification is performed by using microarray platforms that are error-prone and difficult to quantify (Shi et al. 2006). Whereas most of the methods presently in use, such as PCR-based DNA amplification followed by sequencing techniques, allow to handle only single or a few species at the same time, DNA microarrays are believed to have the potential of identifying hundreds of species in parallel and to differentiate them against an even larger number of related species.

Commonly, microarrays are glass microscope slides on which oligonucleotide probes are spotted that are complementary to the DNA target sequences to be analysed (Relógio et al. 2002, Pirrung 2002, Dufva 2005). The

DNA target, which is usually fluorophore-labelled during PCR amplification, hybridises with the oligonucleotide probe on the microarray and can be detected after washing steps by its label. Whereas applying of DNA microarrays for gene expression has already reached the routine level of high throughput systems (Blohm and Guiseppi-Elie 2001, Hoheisel 2006), they have been only recently applied for the identification of organisms, such as microbes (Wang et al. 2002; Call et al. 2003; Korimbocus et al. 2005; Loy and Bodrossy 2005), animals (Pfundner et al. 2004), and plants (Rønning et al. 2005). In terms of identifying marine organisms, microarrays have been used for bacteria (Peplies et al. 2003; Peplies et al. 2004) and phytoplankton (Metfies and Medlin 2004; Metfies et al. 2005; Godhe et al. 2007). Other DNA-hybridisation methods for the identification of higher marine organisms, such as invertebrate larvae (Goffredi et al. 2006), copepods (Kiesling et al. 2002), and larvae of fishes (Rosel and Kocher 2002), have been applied, but microarrays have not been used for this kind of studies. Other applications of microarrays in research on marine organisms are gene expression analysis (Williams et al. 2003; Lidie et al. 2005; Wang et al. 2006; Cohen et al. 2007; Jenny et al. 2007) and genotyping in population genetics (Moriya et al. 2004; Moriya et al. 2007).

One of the methodical limitations for using microarrays is the design of species specific probes. On the one hand, oligonucleotide probes designed *in silico* do not always exhibit the experimental hybridisation properties they were selected for and must be empirically tested. On the other hand, the molecular marker must have highly selective characteristics, such as low intraspecific and a high interspecific variation. One of the most frequently used markers in phylogenetics of fishes is the mitochondrial 16S rRNA gene. This gene has a well-characterized secondary structure (Meyer 1993; Ortí et al. 1996) and especially the loop regions exhibit many insertions, deletions, and substitutions forming highly variable molecular features, which usually allow the design of highly specific probes. This is underlined by a study on lionfishes (Kochzius et al. 2003), which showed that individuals of one species exhibit identical 16S rDNA haplotypes even though they were sampled at sites thousands of kilometres apart, but clear differences could be detected between closely related lionfish species.

In this study, the development of a DNA microarray is described, demonstrating the suitability of the 16S rRNA gene for designing oligonucleotides as microarray probes to differentiate at least 11 fish species from European seas. Based on these data, a “Fish Chip” for approximately 50 fish species is under construction to support the identification of eggs and larval stages from species that are otherwise difficult to identify, and of adult or processed fishes in fisheries industry.

## Material and Methods

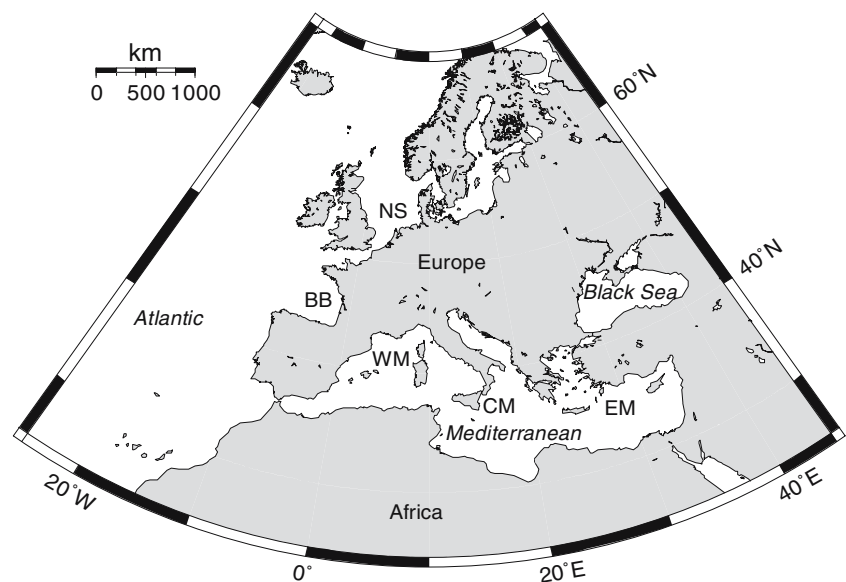
### Sampling and DNA Extraction

To consider possible intraspecies sequence variations, fishes were collected in five different regions of the European seas: North Sea, Bay of Biscay, and Western, Central, as well as Eastern Mediterranean (Fig. 1, Table 1). All together 267 individual samples from 79 fish species were investigated. In addition sequences from the EMBL sequence data base were included. Voucher specimens and tissue samples have been preserved in absolute ethanol and stored at 4°C or were frozen at -20°C. DNA was extracted from gill filaments with the Agowa mag midi DNA isolation kit (AGOWA, Berlin, Germany) or from muscle tissue with the DNeasy tissue kit (Qiagen, Hilden, Germany) according to the instructions of the manufacturer.

### Polymerase Chain Reaction and Sequencing

A fragment of approximately 1380 bp length from the mitochondrial 16S rRNA gene was amplified with the primer 16fiF140 (5'-CGY AAG GGA AHG CTG AAA-3'), which has a single-base modification compared with Palumbi et al. (1991, unpublished manuscript) as well as with the newly designed primer 16fiR1524 (5'-CCG GTC TGA ACT CAG ATC ACG TAG-3'). Polymerase chain reaction (PCR) reactions with a total volume of 15 µl contained 1.5 µl 10 X reaction buffer, 1.5 µl dNTPs (10 mM), 0.05 µl of each primer (100 pmol/µl), 5 µl DNA-extract, 0.3 µl Tg polymerase (3 U/µl; Prokaria, Reykjavik, Iceland), and 6.6 µl deionized water. Thermal profile began at 94°C for 4 min, followed by 35 cycles of 94°C (30 s), 54°C (30 s), 72°C (90 s), with a final step of 7 min at 72°C.

**Fig. 1** Map with the sampling areas; *NS* North Sea; *BB* Bay of Biscay; *WM* Western Mediterranean; *CM* Central Mediterranean; *EM* Eastern Mediterranean



PCR products were purified by using the ExoSAP-IT for PCR clean-up (GE Healthcare, Uppsala, Sweden). The newly designed sequencing primer 16fiseq1463 (5'-TGC ACC ATT AGG ATG TCC TGA TCC AAC-3') was used to sequence one strand of the amplified fragment using the BigDye Terminator Cycle Sequencing Kit (ver. 3.1, PE Biosystems, Foster City, USA). The sequencing reactions were run in an ABI Prism 3730 automated DNA Analyzer (Applied Biosystems, Foster City, USA) according to the manufacturer's instructions.

### Sequence Analysis and Oligonucleotide Probe Design

In the framework of this project, 267 sequences have been acquired and 944 sequences have been obtained from EMBL sequence database and other projects, representing approximately 380 species of fish from European seas. Probes were designed based on 230 sequences obtained from 27 species (Table 1). A multiple alignment of these 230 sequences was performed with the programme Clustal W (Thompson et al. 1994) as implemented in BioEdit (version 7.0.4.1; Hall 1999) to ensure that all sequences represent a homologous fragment of the 16S rDNA. Before probe design gaps have been removed from each sequence. A computer program developed by the bioinformatics group of the Centre for Applied Gene Sensor Technology (CAG) and the Zentrum für Technomathematik (ZeTeM), both at University of Bremen, was used to design species-specific oligonucleotide probes, which ideally cover all sequences of one species and do not match any other species (Nölte 2002). The following criteria have been considered: (1) length of 23 to 27 bp, (2) melting temperature ( $T_m$ ) of 81 to 85°C based on the unified model (SantaLucia 1998), (3) GC content of 52% to

**Table 1** Number of sequences per fish species and sampling region

Species	Family	Order	North Sea	Bay of Biscay	Western Mediterranean	Central Mediterranean	Eastern Mediterranean	EMBL sequence data base and other projects	Total	Accession number of EMBL data base sequences
<b>Target species</b>										
<i>Boops boops</i>	Sparidae	Perciformes			3		4	1	8	AF247396
<i>Engraulis encrasicolus</i>	Engraulidae	Clupeiformes		3	2		1		6	
<i>Helicolenus dactylopterus</i>	Sebastidae	Scorpaeniformes		3	1		6	1	11	AY538975
<i>Lophius budegassa</i>	Lophiidae	Lophiiformes		3			6		9	
<i>Pagellus acarne</i>	Sparidae	Perciformes		3	3		3		9	
<i>Scomber scombrus</i>	Scombridae	Perciformes		3			1	2	6	AB120717, AF055615
<i>Scophthalmus rhombus</i>	Scophthalmidae	Pleuronectiformes	3	3	1		1		8	AY359665
<i>Serranus cabrilla</i>	Serranidae	Perciformes			3		4		7	
<i>Sparus aurata</i>	Sparidae	Perciformes			3		2	1	6	AF247432
<i>Trachurus trachurus</i>	Carangidae	Perciformes	1		2		5	2	10	AB108498, AB096007
<i>Trigla lyra</i>	Triglidae	Scorpaeniformes		3			3		6	
<b>Additional species</b>										
<i>Chelidonichthys lucernus</i>	Triglidae	Scorpaeniformes	1	3			6		10	
<i>Diplodus sargus</i>	Sparidae	Perciformes			2		4		6	
<i>Gadus morhua</i>	Gadidae	Gadiformes	1	2				3	6	X99772, NC_002081, AY850363
<i>Merluccius merluccius</i>	Merlucciidae	Gadiformes		2	3	3	6		14	
<i>Mullus surmuletus</i>	Mullidae	Perciformes	2	5	2		2		11	
<i>Mullus barbatus</i>	Mullidae	Perciformes			2		5		7	
<i>Pagellus erythrinus</i>	Sparidae	Perciformes		1	2		7		10	
<i>Platichthys flesus</i>	Pleuronectidae	Pleuronectiformes	4	2				4	10	AY359670, AB125255, AY157320, AF113180
<i>Pleuronectes platessa</i>	Pleuronectidae	Pleuronectiformes						2	6	AY359673, AY157328
<i>Psetta maxima</i>	Scophthalmidae	Pleuronectiformes		3					4	
<i>Sardina pilchardus</i>	Clupeidae	Clupeiformes		3	3		4		10	
<i>Scorpaena notata</i>	Scorpaenidae	Scorpaeniformes			5		6		11	
<i>Scorpaena porcus</i>	Scorpaenidae	Scorpaeniformes		3			3	1	7	
<i>Serranus hepatus</i>	Serranidae	Perciformes			2	2	3		7	
<i>Sotea sotea</i>	Soleidae	Pleuronectiformes	3	3	2			3	11	AB125247, AF488442, AF112845
<i>Zeus faber</i>	Zeidae	Zeiformes		3	3		2	6	14	NC_003190, AF488474, AF221894-AF221896, AP002941
									Σ	230

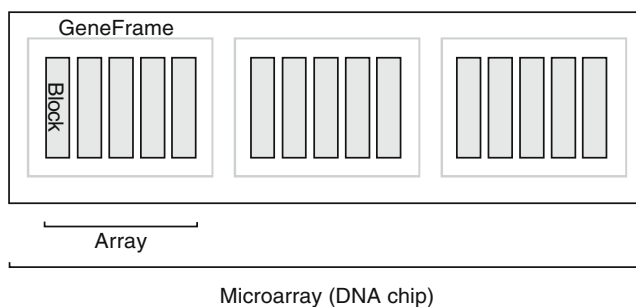
Accession numbers are given for sequences obtained from EMBL sequence database Taxonomy according to FishBase (2007)

54%, (4) secondary structure of the oligonucleotides and the target sequence, (5) possible dimer formation, and (6) the energy content of a bond between the probe and the target sequence. Minimal free energy (mfe) structures are computed by using RNAfold (Hofacker et al. 1994). Probes exhibiting strong secondary structures or binding to a region of the target with such a strong secondary structure were not used. If more than one probe qualified for a species, the one with the highest binding energy between probe and target was chosen. Already during the design-phase the selected oligonucleotide probes were tested *in silico* against 1211 background sequences from approximately 380 fish species.

### Preparation of DNA Microarrays and Hybridisation Experiments

Aminosilane (3-aminopropyltrimethoxysilane)-coated glass slides were used with a PDITC-linker (1,4-phenylendiisothiocyanate) from Asper Biotech (Tartu, Estonia). Oligonucleotide probes (Thermo Hybaid, Ulm, Germany) with a 5'-amino-C6-modification were spotted in 150 mM Na<sub>3</sub>PO<sub>4</sub> buffer (pH 8.5) at a concentration of 30 μM using a spotting robot based on a modified version of the contactless TopSpot® technology. The spotted volume of this oligonucleotides solution was 200 pl, producing a spot diameter of approximately 220 μm. Each probe was spotted in four replicates per block. An array contained five blocks and three arrays were spotted on one microarray slide (Fig. 2). After spotting, the microarrays were incubated for 16 h in a wet chamber to ensure efficient covalent binding of the oligonucleotides. Finally, the microarrays were shrink-wrapped under a nitrogen atmosphere and were storable at 4°C for up to 6 months.

DNA for hybridisation experiments was amplified and labelled with 5'-Cy5-modified primers. The primers 16sar-L (3'-CGC CTG TTT AAC AAA AAC AT-3') and 16sbr-H (5'-CCG GTT TGA ACT CAG ATC ACG T-3') amplify a fragment of approximately 600 bp length from the mitochondrial 16S rRNA gene (Palumbi et al. 1991, unpublished manuscript). PCR reactions with a volume of 100 μl contained



**Fig. 2** Layout of the microarray

**Table 2** Nontarget species tested in hybridisation experiments

Species	Family	Order
<i>Dentex dentex</i>	Sparidae	Perciformes
<i>Diplodus vulgaris</i>	Sparidae	Perciformes
<i>Gadus morhua</i>	Gadidae	Gadiformes
<i>Melanogrammus aeglefinus</i>	Gadidae	Gadiformes
<i>Merlangius merlangus</i>	Gadidae	Gadiformes
<i>Merluccius merluccius</i>	Gadidae	Gadiformes
<i>Micromesistius poutassou</i>	Gadidae	Gadiformes
<i>Mullus surmuletus</i>	Mullidae	Perciformes
<i>Pollachius pollachius</i>	Gadidae	Gadiformes
<i>Pollachius virens</i>	Gadidae	Gadiformes
<i>Psetta maxima</i>	Scophthalmidae	Pleuronectiformes
<i>Serranus hepatus</i>	Serranidae	Perciformes
<i>Trachurus mediterraneus</i>	Carangidae	Perciformes
<i>Trachuru picturatus</i>	Carangidae	Perciformes

Taxonomy according to FishBase (2007)

10 μl 10 X reaction buffer, 4 μl dNTPs (5 mM), 2 μl of each primer (10 μM), 2 μl DNA-extract, 0.4 μl *Taq* polymerase (5 U/μl), 2 μl BSA (20 mg/ml), and 77.6 μl deionized water. The PCR thermal profile began at 95°C for 2 min, followed by 35 cycles of 95°C (30 s), 54°C (45 s), 72°C (60 s), followed by a final step of 10 min at 72°C. The Cy5-labelled PCR amplified DNA was purified using the QIAquick PCR Purification Kit (QIAGEN, Hilden, Germany).

Hybridisation experiments were performed with 11 target and 14 nontarget fish species. As nontarget species fishes were chosen that are closely related to the target species and from which false-positive signals could be expected according to the *in silico* specificity tests performed during the design phase of the probes (Table 2).

A positive control at a concentration of 1 nM (5'-CGT GTG AGT CGA TGG ATC ATA-3'; 5'-Cy3-labelled) and 10 nM of the purified Cy5-labelled PCR product were hybridized to the microarray in a volume of 65 μl using GeneFrames® (ABgene House, Epsom, UK), which were applied to the microarray slides according to the manufacturer's instruction (Fig. 2). Hybridisation was conducted at 50°C in a hybridisation oven. After 2 h hybridisation time the GeneFrames® were removed and the microarrays were washed 5 minutes each with 2×SSC (sodium chloride trisodium citrate) buffer containing 0.05% SDS (sodium dodecyl sulphate), 1×SSC containing 0.05% SDS, and 1×SSC. Finally the microarrays were dried in a centrifuge at 2000 rpm for 3 minutes.

### Measurement of Fluorescence Signals and Data Analysis

Hybridisation signals were measured using an Axon 4000B fluorescence microarray scanner at 635 nm (Cy5) as well as at 528 nm (Cy3). The fluorescence signal analysis was

conducted with the software GenePix 4.1 (Axon, Union City, USA). The fluorescence signals of each probe were measured and the arithmetic mean was calculated. However, data were removed from the analysis if the spots showed artefacts caused during the spotting process (e.g., inhomogeneous spots documented by a monitoring camera during spotting) or experimental artefacts (e.g., air bubbles). Background noise was corrected by subtracting the arithmetic mean of the negative control measurement from the arithmetic mean of the spot measurements. Negative values were set to zero.

## Results

Probes for 11 commercially important fish species have been designed (Table 3). Positions of the oligonucleotide probes in the 16S rDNA fragment used for probe design are given in Fig. 3. Following the nomenclature of Ortí et al. (1996), the binding sites of the microarray probes for *Engraulis encrasicolus*, *Sparus aurata*, and *Trigla lyra* are located in the variable region *j*, whereas the probes for *Boops boops*, *Helicolenus dactylopterus*, *Lophius budegassa*, *Pagellus acarne*, *Scomber scombrus*, *Scophthalmus rhombus*, *Serranus cabrilla*, and *Trachurus trachurus* bind to the variable region *l*.

All single target hybridisations of the Cy5-labelled 16S rDNA fragment gave true-positive fluorescence signals for the corresponding probe (Fig. 4), but the hybridisation efficiency is very different. The signal intensity of the weakest probe-

target pair gave approximately 1,000 arbitrary units and the strongest almost 30,000 (Table 4). Under the experimental conditions selected, single target hybridisations with *B. boops*, *E. encrasicolus*, *H. dactylopterus*, *L. budegassa*, *S. rhombus*, *S. aurata*, *T. trachurus*, and *T. lyra* did not show any false-positive signals. Eight probes gave under these conditions very weak false-positive signals with *P. acarne*, *S. scombrus*, and *S. cabrilla*, representing approximately 7% of all possible cross-hybridisation reactions. The values of these false-positive signals were several orders of magnitude lower than the true positive-signals.

Single nontarget hybridisations showed 48 very weak false-positive signals, representing 31% of 154 possible cross-hybridisations. These false-positive signals were at least one order of magnitude weaker than the true-positive signal. *Merlangius merlangus* and *Merluccius merluccius* did not give any false-positive signal, whereas *Dentex dentex*, *Diplodus vulgaris*, *Gadus morhua*, *Melanogrammus aeglefinus*, *Micromesistius poutassou*, *Mullus surmuletus*, *Pollachius pollachius*, *Pollachius virens*, *Psetta maxima*, *Serranus hepatus*, *Trachurus mediterraneus*, and *Trachurus picturatus* showed two to eight cross-hybridisations. Considering all possible cross-hybridisations, only 18% showed usually very weak false-positive signals.

## Discussion

The results show that the “Fish Chip” described enables the identification of 11 commercially important fish species

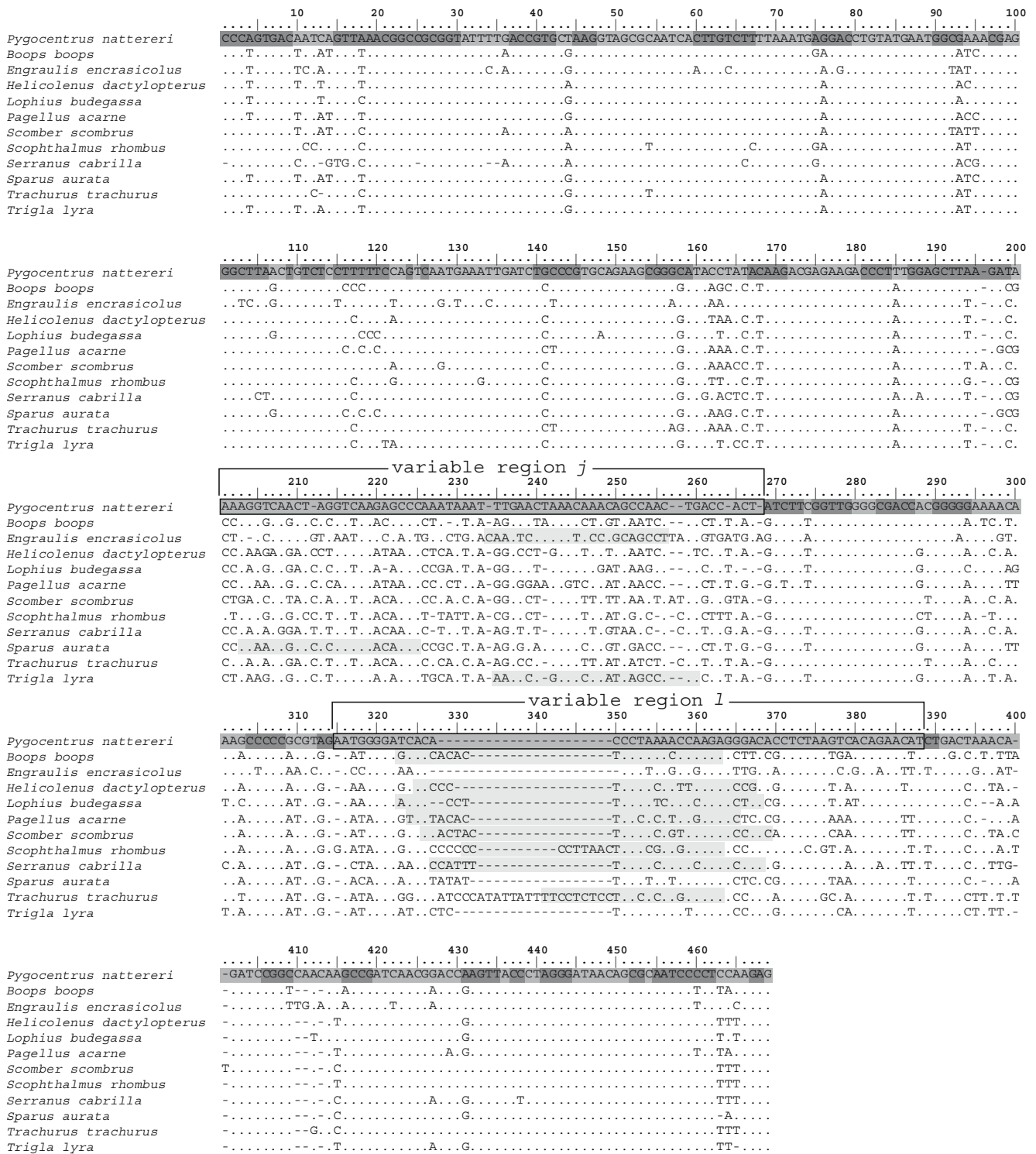
**Table 3** Oligonucleotide probes for the identification of fish species from European seas

Species name	Probe name	Probe sequence (5'-3'), 5'-amino-C6-modified	Length (bp)	T <sub>m</sub> (°C)	GC (%)	Oligo mfe	Dimer mfe	Specificity ( <i>in silico</i> )
<i>Boops boops</i>	Booboo_315	GCACCACACTCCTAAACCCAAGA	23	82.64	52	≥0	-0.07	species
<i>Engraulis encrasicolus</i>	Engenc_213	CAAGTCCTAAATACCCGCAGCCT	23	82.49	52	≥0	-0.17	species
<i>Helicolenus dactylopterus</i>	Heldac_317	ACCCCTCCTACAATTAAGAGCCG	23	81.84	52	≥0	-0.22	species
<i>Lophius budegassa</i>	Lopbud_312	AACACCCCTTCCTATCACCCAGAGCTAC	27	84.39	52	≥0	-0.2	genus
<i>Pagellus acarne</i>	Pagaca_317	TACTACACTCCCACATCCGAGAGC	24	82.77	54	≥0	-0.89	species
<i>Scomber scombrus</i>	Scosco_321	CAACTACTCCTACAGTCAAGAGCCACC	27	82.91	52	≥0	-0.43	species
<i>Scophthalmus rhombus</i>	Scorho_322	CCCCTTAACTCCTCGAAGCAAGA	23	81.88	52	≥0	-0.37	species
<i>Serranus cabrilla</i>	Sercab_313	CCATTTTCCTACAACCCAGAGCGAC	25	82.74	52	≥0	-0.18	species
<i>Sparus aurata</i>	Spaaur_201	AGAACAGCTCACGTCAAACACCC	23	83.02	52	≥0	-0.5	species
<i>Trachurus trachurus</i>	Tratra_333	TTCCTCTCCTCCCACAAGCAAGA	23	83.62	52	≥0	-0.15	genus
<i>Trigla lyra</i>	Trilyr_232	AAGACCGAACCAAATGAGCCCTG	23	83.16	52	≥0	-0.17	family

The number in the probe name indicates the binding site in the 16S rDNA sequence

Oligo mfe minimal free energy of the secondary structure of the oligonucleotide; Dimer mfe minimal free energy of the dimer of two identical oligonucleotide molecules

Values for mfe are given in kcal/mol

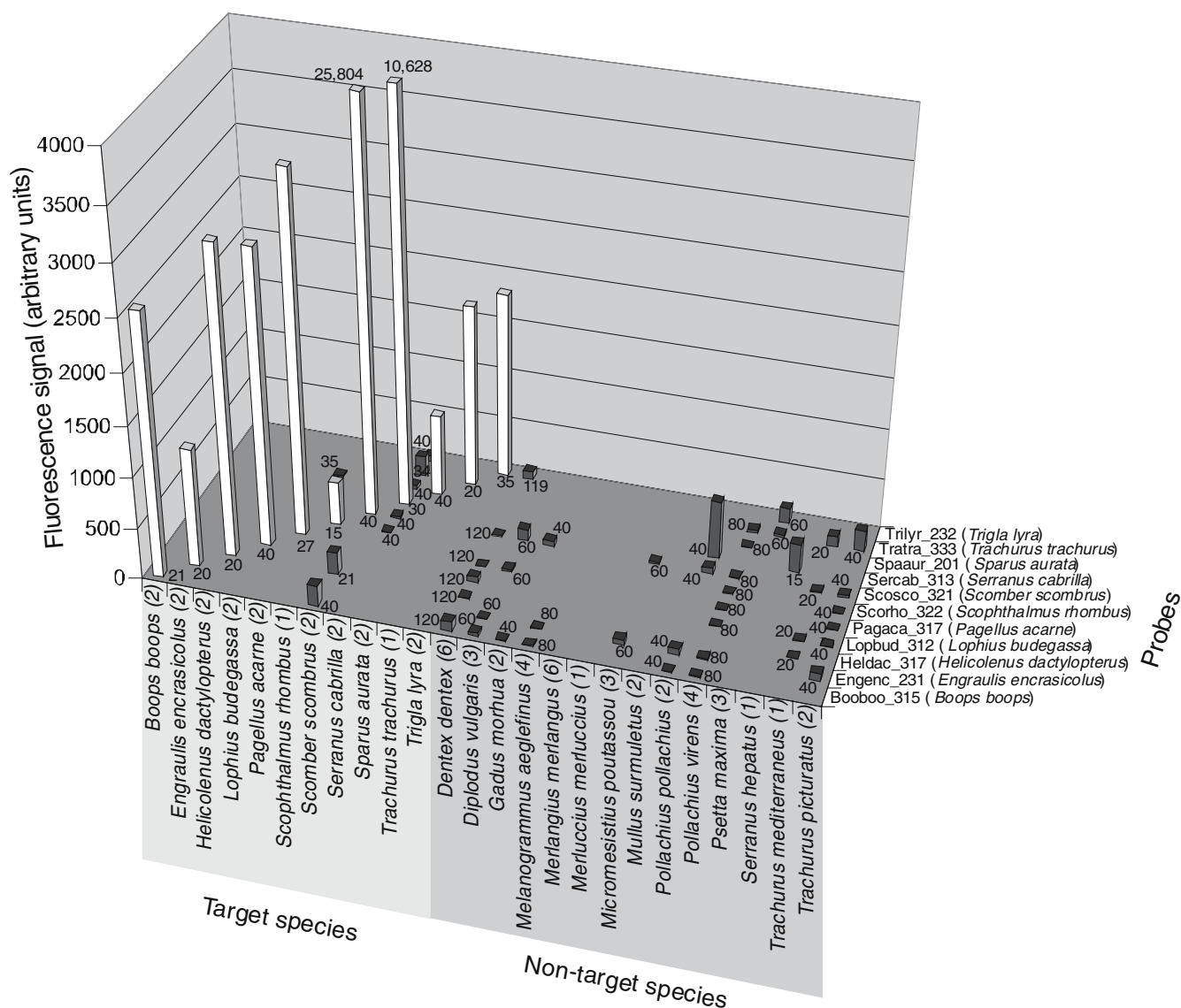


**Fig. 3** Alignment (5'>3') of representative 16S rDNA sequences from the target species with binding sites (light grey) of probes (5'>3'; probes hybridise to the reverse complementary target strand). Double

stranded (dark grey) and single stranded regions (grey) of the secondary structure are indicated in the reference sequence of *Pygoplites nattereri* (Ortí et al. 1996; Accession number: U33590)

from European seas in certain experimental limits. These limits are primarily given by the fact that the fluorescence signal intensities of true-positive hybridisation signals were heterogeneous. This phenomenon is commonly encoun-

tered in DNA microarray experiments (Peplies et al. 2003; Warsen et al. 2004; Korimbocus et al. 2005; Rønning et al. 2005; Tobler et al. 2006) and can probably be overcome only by an extreme methodical effort (Shi et al. 2006). The



**Fig. 4** Signals (background subtracted from absolute signal) of single target and nontarget hybridisations. White bars represent true-positive signals; false-positive signals are shown as grey bars. Numbers at the basis of the bars indicate the amount of measured spots. The number

of hybridisations is given in brackets after target and nontarget names. Replication and absolute signal intensities ( $\pm$  standard deviation) of hybridized targets to the corresponding probe are given in Table 4

**Table 4** Target hybridisations

Hybridized targets	No. of hybridisations	Measurements of specific probes		
		Measured probes/absolute no. of probes	Mean absolute fluorescence signal in arbitrary units	Standard deviation
<i>Boops boops</i>	2	21/40	2991	$\pm$ 1491
<i>Engraulis encrasicolus</i>	2	20/40	1659	$\pm$ 962
<i>Helicolenus dactylopterus</i>	2	20/40	3502	$\pm$ 912
<i>Lophius budegassa</i>	2	40/40	3450	$\pm$ 1515
<i>Pagellus acarne</i>	2	27/40	3727	$\pm$ 1270
<i>Scophthalmus rhombus</i>	1	15/20	1528	$\pm$ 269
<i>Scomber scombrus</i>	2	40/40	27827	$\pm$ 5330
<i>Serranus cabrilla</i>	2	40/40	10814	$\pm$ 4396
<i>Sparus aurata</i>	2	40/40	963	$\pm$ 227
<i>Trachurus trachurus</i>	1	20/20	2015	$\pm$ 880
<i>Trigla lyra</i>	2	35/40	2343	$\pm$ 560



problems encountered when single-colour microarray experiments need to be quantified are severe and in part not yet solved, because complex parameters are influencing the results. These are the sequence dependent hybridisation efficiency, specifically steric hindrance, secondary structures (Southern et al. 1999), and the relative position of the fluorescent label at the target (Zhang et al. 2005).

Sometimes the duplex formation can be favoured by using spacers, which obviously give the captures a greater degree of freedom from their neighbouring molecules and from the surface (Southern et al. 1999), leading to an enhancement of signal intensity with increasing spacer length (Peplies et al. 2003).

An important criterion of probe design was the base composition (Southern et al. 1999) and therefore all but one capture oligonucleotides were designed to have a GC content of 52% and 54% for the exception (Table 3). The different hybridisation efficiencies of the captures and the varying sensitivity for the different fishes is still a severe disadvantage, because it can hamper the estimation of a small amount of fishes of one species in the presence of a larger number of individuals of another species in a mixed sample. These limits are presently under investigation with fish eggs and other biological material.

Because all oligonucleotide probes bind to the variable regions *j* and *l* of the 16S rRNA gene, which represent large single-stranded loops, the secondary structure is unlikely to be a factor contributing to the partially low sensitivity. Also, the POL effect, a phenomenon decreasing the fluorescence signal with increasing distance between the probe binding site and the label on the target (Zhang et al. 2005) seems not to be important in this case. Most of the probes used in this study hybridise to the variable region *l* of the 16S rDNA and their distance from the fluorescence label is more or less identical.

Although cross-hybridisations occur, true-positive signals could clearly be differentiated from false-positive signals because of their generally higher signal. Most false-positive signals occurred when the 16S rDNA fragment of nontarget species was hybridised on the microarray.

Testing of nontarget species is seemingly important in studies using DNA microarrays for the identification of organisms, because even if a comprehensive sequence background is utilised for probe design and even if extensive *in silico* testing has been performed, the specificity of oligonucleotide probes has to be evaluated with closely related species and specifically those that show cross-hybridisations *in silico*.

This study shows that the 16S rRNA gene of fishes is suitable to design oligonucleotide probes that are able to differentiate eleven fish species from European seas by single target hybridisation on a microarray. Such a “Fish Chip” can

hopefully be applied in marine environmental and fisheries research, as well as in fisheries and food control if the uneven hybridisation signal intensities of the different probe-target pairs can be improved or compensated.

The correct identification of fish eggs and larvae is crucial for fish stock assessment based on ichthyoplankton surveys. Genetic identification has shown that the majority of eggs in the Irish Sea, wrongly believed to be from cod, were actually from whiting, leading to an overestimation of cod stocks (Fox et al. 2005). A study on food fish in the United States revealed that three-quarters of fish sold as “red snapper” were mislabelled and belonged to other species (Marko et al. 2004), a situation that needs better analytical tools to be changed. The European Union (EU) also has strict regulations for seafood labelling, which must include, for example, the species name (EU Council Regulation No 104/2000; EU Commission Regulation No 2065/2001). Approximately 420 species of fish are sold in the German market alone, making a reliable identification method urgently necessary to protect the customer. DNA microarrays might have the potential to fulfill these requirements.

Recent efforts in compiling sequences of fishes, such as the European projects “FishTrace” (<http://www.fishtrace.org>) and “Fish & Chips” (<http://www.fish-and-chips.uni-bremen.de>; Kochzius et al. 2007), as well as the international “Fish Barcode of Life Initiative” (<http://www.fishbol.org>; Ward et al. 2005), will provide the necessary sequence background for the design of species specific oligonucleotide probes for the development of DNA microarrays for the identification of fishes.

**Acknowledgements** The authors thank S. Roll and F. Meyerjürgens for producing the microarrays. The “Fish & Chips” project is a Specific Targeted Research Project (STREP) funded by the European Commission under the contract no. 505491.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

## References

- Barlow BA, Tzotzos GT (1995) Biotechnology. In: Heywood VH (ed) Global biodiversity assessment. UNEP, Cambridge University Press, pp. 671–710
- Barber P, Boyce SL (2006) Estimating diversity of Indo-Pacific coral reef stomatopods through DNA barcoding of stomatopod larvae. *Proc Royal Soc London B* 273:2053–2061
- Blohm D, Guiseppe-Elie A (2001) New developments in microarray technology. *Current Opin Biotech* 12:41–47
- Bucklin A, Guarnieri M, Hill RS, Bentley AM, Kaartvedt S (1999) Taxonomic and systematic assessment of planktonic copepods using mitochondrial COI sequence variation and competitive, species-specific PCR. *Hydrobiologia* 401:239–254

- Call DR, Borucki MK, Loge FJ (2003) Detection of bacterial pathogens in environmental samples using DNA microarrays. *J Microbiologic Methods* 53:235–243
- Cohen R, Chalifa-Caspi V, Williams TD, Auslander M, George SG, Chipman JK, Tom M (2007) Estimating the efficiency of fish cross-species cDNA microarray hybridisation. *Marine Biotech* 9:491–499
- Dufva M (2005) Fabrication of high quality microarrays. *Biomolecular Engineer* 22:173–184
- FishBase (2007) FishBase: a global information system on fishes. Available at: <http://www.fishbase.de>
- 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
- Garland ED, Zimmer CA (2002) Techniques for the identification of bivalve larvae. *Marine Ecol Prog Ser* 225:299–310
- Godhe A, Cusack C, Pedersen J, Andersen P, Anderson DM, Bresnan E, Cembella A, Dahl E, Diercks S, Elbrächter M, Edler L, Galluzzi L, Gescher G, Gladstone M, Karlson B, Kulis D, LeGresley M, Lindahl O, Marin R, McDermott G, Medlin LK, Naustvoll L-J, Penna A, Kerstin Töbe (2007) Intercalibration of classical and molecular techniques for identification of *Alexandrium fundyense* (Dinophyceae) and estimation of cell densities. *Harmful Algae* 6:56–72
- Goffredi SK, Jones WJ, Scholin CA, Marin R, Vrijenhoek (2006) Molecular detection of marine invertebrate larvae. *Marine Biotechnol* 8:149–160
- Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser* 41:95–98
- Hebert PDN, Cywinska A, Ball SL, deWaard JR (2002) Biological identifications through DNA barcodes. *Proc Royal Soc London B* 270:313–321
- Hofacker IL, Fontana W, Stadler PF, Bonhoeffer LS, Tacker M, Schuster P (1994) Fast folding and comparison of RNA secondary structures (The Vienna RNA package). *Monatshefte für Chemie (Chemical Monthly)* 125:167–188
- Hoheisel JD (2006) Microarray technology: beyond transcript profiling and genotype analysis. *Nat Rev Genetics* 7:200–210
- Jarman SN, Gales NJ, Tierney M, Gill PC, Elliott NG (2002) A DNA-based method for identification of krill species and its application to analysing the diet of marine vertebrate predators. *Mol Ecol* 11:2679–2690
- Jenny MJ, Chapman RW, Mancia A, Chen YA, McKillen DJ, Trent H, Lang P, Escoubas JM, Bachere E, Boulo V, Liu ZJ, Gross PS, Cunningham C, Cupit PM, Tanguy A, Guo X, Moraga D, Boutet I, Huvet A, De Guise S, Almeida JS, Warr GW (2007) A cDNA Microarray for *Crassostrea virginica* and *C. gigas*. *Mar Biotechnol* 9:577–591
- Kiesling TL, Wilkinson E, Rabalais J, Ortner PB, McCabe, MM, Fell JW (2002) Rapid identification of adult and naupliar stages of copepods using DNA hybridisation methodology. *Marine Biotechnol* 4:30–39
- Kochzius M, Söller R, Khalaf MA, Blohm D (2003) Molecular phylogeny of the lionfish genera *Dendrochirus* and *Pterois* (Scorpaenidae, Pteroinae) based on mitochondrial DNA sequences. *Mol Phylogenetics Evol* 28:396–403
- Kochzius M, Kappel K, Döbitz L, Silkenbeumer N, Nölte M, Weber H, Hjørleifsdottir S, Marteinson V, 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 & Chips” project: microarrays as a tool for the identification of marine organisms in biodiversity and ecosystem research. Oceans '07 IEEE Aberdeen 18–21 June 2007, Aberdeen, Conference Proceedings, ISBN: 1-4244-0635-8, 4 pp
- Korimbocus J, Scaramozzino N, Lacroix B, Crance JM, Garin D, Vernet G (2005) DNA probe array for the simultaneous identification of herpesviruses, enteroviruses, and flaviviruses. *J Clin Microbiol* 43:3779–3787
- Lidie KB, Ryan JC, Barbier M, Van Dolah FM (2005) Gene expression in Florida red tide dinoflagellate *Karenia brevis*: analysis of an expressed sequence tag library and development of DNA microarray. *Marine Biotechnol* 7:481–493
- Loy A, Bodrossy L (2005) Highly parallel microbial diagnostics using oligonucleotide microarrays. *Clinica Chimica Acta* 363:106–119
- Ludwig W, Strunk O, Westram R, Richter L, Meier H, Yadukumar, Buchner A, Lai T, Steppi S, Jobb G, Förster W, Brettske I, Gerber S, Ginhart AW, Gross O, Grumann S, Hermann S, Jost R, König A, Liss T, Lüßmann R, May M, Nonhoff B, Reichel B, Strehlow R, Stamatakis A, Stuckmann N, Vilbig A, Lenke M, Ludwig T, Bode A, Schleifer K-H (2004) ARB: a software environment for sequence data. *Nucleic Acids Res* 32:1363–1371
- Marko PB, Lee SC, Rice AM, Gramling JM, Fitzhenry TM, McAlister JS, Harpert GR, Moran AL (2004) Misslabelling of a depleted reef fish. *Nature* 430:309–310
- Metfies K, Medlin L (2004) DNA microchips for phytoplankton: the fluorescent wave of the future. *Nova Hedwigia* 79:321–327
- Metfies K, Huljic S, Lange M, Medlin LK (2005) Electrochemical detection of the toxic dinoflagellate *Alexandrium ostenfeldii* with a DNA-biosensor. *Biosensors Bioelectronics* 20:1349–1357
- Meyer A (1993) Evolution of mitochondrial DNA in fishes. In: Hochachka PW, Mommsen TP (eds) *Biochemistry and molecular biology of fishes*. 2nd edn. Elsevier, Amsterdam, p 1–38
- Moriya S, Urawa S, Suzuki O, Urano A, Abe S (2004) DNA microarray for rapid detection of mitochondrial DNA haplotypes of chum salmon. *Marine Biotechnol* 6:430–434
- Moriya S, Sato S, Azumaya T, Suzuki O, Urawa S, Urano A, Abe S (2007) Genetic stock identification of chum salmon in the Bering Sea and North Pacific Ocean using mitochondrial DNA microarray. *Marine Biotechnol* 9:179–191
- Noell CJ, Donnellan S, Foster R, Haigh L (2001) Molecular discrimination of garfish *Hyporhamphus* (Beloniformes) larvae in southern Australian waters. *Marine Biotechnol* 3:509–514
- Nölte M (2002) Optimization of oligonucleotide sets for DNA microarrays. Ph.D. thesis, University of Bremen, Germany
- Ortí G, Petry P, Porto JIR, Jégu M, Meyer A (1996) Patterns of nucleotide change in mitochondrial ribosomal RNA genes and the phylogeny of Piranhas. *J Mol Evolution* 42:169–182
- Peplies J, Glöckner 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, Lau SC, Penthaler J, Amann R, Glöckner FO (2004) Application and validation of DNA microarrays for the 16S rRNA-based analysis of marine bacterioplankton. *Environ Microbiol* 6:638–645
- Pfunder M, Holzgang O, Frey JE (2004) 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
- Pirrung MC (2002) How to make a DNA chip. *Angewandte Chemie Int Ed* 41:1276–1289
- Relógio A, Schwager C, Richter A, Ansober W, Valcárcel J (2002) Optimization of oligonucleotide-based DNA microarrays. *Nucleic Acids Res* 30(11):e51
- Rocha-Olivares A (1998) Multiplex haplotype-specific PCR: a new approach for species identification of the early life stages of rockfishes of the species-rich genus *Sebastes* Cuvier. *J Exp Marine Biol Ecol* 231:279–290

- Rønning SB, Rudi K, Berdal KG, Holst-Jensen A (2005) Differentiation of important and closely related cereal plant species (Poaceae) in food by hybridisation to an oligonucleotide array. *J Agricultural Food Chem* 53:8874–8880
- Rosel PE, Kocher TD (2002) DNA-based identification of larval cod in stomach contents of predatory fishes. *J Exp Marine Biol Ecol* 267:75–88
- Saitho K, Takagaki M, Yamashita Y (2003) Detection of Japanese flounder-specific DNA from gut contents of potential predators in the field. *Fisheries Sci* 69:473–477
- SantaLucia J (1998) A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proc Natl Acad Sci USA* 95:1460–1465
- Schander C, Willassen E (2005) What can biological barcoding do for marine biology? *Marine Biol Res* 1:79–83
- 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, Scherf U, Thierry-Mieg J, Wang J, Wilson M, Wolber PK, Zhang L, Amur S, Bao W, Barbacioru CC, Bergstrom Lucas A, Bertholet V, Boysen C, Bromley B, Brown D, Brunner A, Canales R, Megan Cao XM, Cebula TA, Chen JJ, Cheng J, Chu T-M, Chudin E, Corson J, Corton JC, Croner LJ, Davies C, Davison TS, Delenstarr G, Deng X, Dorris D, Eklund AC, Fan X-H, 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 Q-Z, Liu C, Liu Y, Lombardi MJ, Ma Y, Magnuson SR, Maqsoodi 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, Zhang L, Zhong S, Zong Y, Slikker Jr W (2006) The microarray quality control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements. *Nat Biotechnol* 24:1151–1161
- Southern E, Mir K, Shchepinov M (1999) Molecular interactions on microarrays. *Nat Genet* 21:5–9
- Thompson JG, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignments through sequence weighting, position specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22:4673–4680
- Tobler NE, Pfunder M, Herzog K, Frey JE, Altwegg M (2006) Rapid detection and species identification of *Mycobacterium* spp. using real-time PCR and DNA-microarray. *J Microbiol Methods* 66:116–124
- Wang D, Coscoy L, Zylberberg M, Avila PC, Boushey HA, Ganem D, DeRisi JL (2002) Microarray-based detection and genotyping of viral pathogens. *Proc Natl Acad Sci USA* 99:15687–15692
- 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. *Marine Biotechnol* 8:491–500
- Ward RD, Zemlak TS, Innes BH, Last PR, Hebert PDN (2005) DNA barcoding Australia's fish species. *Philos Trans Royal Soc B* 360:1847–1857
- Warsen AE, Krug MK, LaFrentz S, Stanek DR, Loge F, Call DR (2004) Simultaneous discrimination between 15 fish pathogens by using 16S ribosomal DNA PCR and DNA microarrays. *Appl Environ Microbiol* 70:4216–4221
- Williams TD, Gensberg K, Minchin SD, Chipman JK (2003) A DNA expression array to detect toxic stress response in European flounder (*Platichthys flesus*). *Aquatic Toxicol* 65:141–157
- Zhang L, Hurek T, Reinhold-Hurek B (2005) Position of the fluorescent label is a crucial factor determining signal intensity in microarray hybridisations. *Nucleic Acids Res* 33:e166