MOLECULAR ECOLOGY
RESOURCES
Molecular Ecology Resources (2011) **11**, 842–853 doi: 10.1111/j.1755-0998.2011.03022.x

Molecular profiling of diatom assemblages in tropical lake sediments using taxon-specific PCR and Denaturing High-Performance Liquid Chromatography (PCR-DHPLC)

LAURA S. EPP,*‡ KATHLEEN R. STOOF-LEICHSENRING,* MARTIN H. TRAUTH† and RALPH TIEDEMANN*

*Unit of Evolutionary Biology/Systematic Zoology, Institute of Biochemistry and Biology, University of Potsdam, Karl-Liebknecht-Strasse 24-25, Haus 26, D-14476 Potsdam, Germany, †Institute of Earth and Environmental Science, University of Potsdam, Karl-Liebknecht-Strasse 24-25, Haus 27, Potsdam, Germany, ‡National Centre for Biosystematics, Natural History Museum, University of Oslo, P.O.Box 1172 Blindern, NO-0318 Oslo, Norway

Abstract

Here we present a protocol to genetically detect diatoms in sediments of the Kenyan tropical Lake Naivasha, based on taxon-specific PCR amplification of short fragments (approximately 100 bp) of the small subunit ribosomal (*SSU*) gene and subsequent separation of species-specific PCR products by PCR-based denaturing high-performance liquid chromatography (DHPLC). An evaluation of amplicons differing in primer specificity to diatoms and length of the fragments amplified demonstrated that the number of different diatom sequence types detected after cloning of the PCR products critically depended on the specificity of the primers to diatoms and the length of the amplified fragments whereby shorter fragments yielded more species of diatoms. The DHPLC was able to discriminate between very short amplicons based on the sequence difference, even if the fragments were of identical length and if the amplicons differed only in a small number of nucleotides. Generally, the method identified the dominant sequence types from mixed amplifications. A comparison with microscopic analysis of the sediment samples revealed that the sequence types identified in the molecular assessment corresponded well with the most dominant species. In summary, the PCR-based DHPLC protocol offers a fast, reliable and cost-efficient possibility to study DNA from sediments and other environmental samples with unknown organismic content, even for very short DNA fragments.

Keywords: diatoms, environmental DNA, lake sediments, PCR-DHPLC

Received 27 October 2010; revision received 16 March 2011; accepted 17 March 2011

Introduction

Biodiversity assessments using total DNA isolates of environmental samples are becoming an increasingly important tool in ecological studies (Valentini *et al*. 2009).

Currently, we are witnessing the development of a number of diagnostic genetic markers and protocols (Taberlet *et al*. 2007; Ficetola *et al*. 2008; Soininen *et al*. 2009). The introduction of high-throughput parallel next-generation amplicon sequencing (Binladen *et al*. 2007; Deagle *et al*. 2009; Jumpponen *et al*. 2010) has greatly enhanced the possibilities of conducting large-scale experiments in this field. These techniques provide a very thorough insight into the composition of PCR products, and the experiments can be designed to process a large number of PCR products simultaneously. Retrieving a huge number of different sequences is not always necessary for every experimental setting, though, and the large scale of these techniques is not always adequate for day-to-day molecular genetic screenings. More rapid alternatives that can present a valid estimate of the biotic composition of a sample are still desirable. Such protocols should be robust, reliable, highly sensitive, fast and cost effective at the same time.

DNA from environmental samples may stem from both living and dead organisms (Valentini *et al*. 2009), and degradation of the DNA can be an important issue. Therefore, there is a specific need for techniques that can be used when only short DNA fragments are retrievable. Such short DNA fragments will often display only small sequence differences, complicating physical separation. In the present study, we explore the potential of denaturing high-performance liquid chromatography (DHPLC) to separate and extract very short PCR products (~100 bp excluding primers) from mixed-template reactions.

Correspondence: Laura Epp, Fax: +47 22 85 18 35;
E-mail: laura.epp@nhm.uio.no

DHPLC offers a potentially very rapid possibility to separate different PCR products and prepare them for individual Sanger sequencing, or also to extract a number of different desired amplicons in preparation for next-generation amplicon sequencing. DHPLC protocols were originally developed for mutation analysis (Xiao & Oefner 2001) and used in human clinical medicine. PCR-based DHPLC can also be used to differentiate between a number of diverse amplicons, based on both sequence length and sequence difference. It has been employed to distinguish between PCR products of different bacterial strains and of different crustacean parasites on the basis of sequence difference of PCR products with an identical length (Barlaan *et al.* 2005; Goldenberg *et al.* 2007; Troedsson *et al.* 2008b).

In PCR-based DHPLC, the PCR products are separated by differential elution of partially melted DNA molecules in an ion-pair reversed-phase chromatographic system. The negatively charged DNA molecule interacts between the ion pairing reagent, triethylammonium acetate (TEAA) and the cartridge matrix of the system. Retention time is dependent on the negative charge of the DNA molecule analysed. This negative charge is reduced, when the column temperature is increased and the double-stranded DNA molecule starts to partially melt. A mixture of different DNA fragments, such as PCR products from a mixed-template reaction, will display different melting behaviour owing to sequence differences. Defined temperature conditions and a given flow rate of gradient buffers (containing acetonitrile and TEAA) will allow separation of different DNA molecules. This separation can be either based on the fragment length or on the nucleotide composition, if the sequence length is equal. This should result in different elution times for different PCR products: each eluted compound is visible as a peak and can be collected separately. The size of the peak depends on the amount of the eluted compound (i.e. the amount of the specific PCR product). Here we explore this rapid high-throughput method for use with very short amplicons of diatoms obtained from lake surface sediments of Lake Naivasha in Kenya.

Diatoms are important members of benthic and planktonic algal communities and occur in virtually any aquatic environment worldwide (Mann & Droop 1996). They are good indicators of salinity and pH (Birks *et al.* 1990; Fritz *et al.* 1991; Gasse *et al.* 1995) and are commonly used in ecological and paleoecological research. As their silica frustules are well identifiable by microscopic inspection, they are well suited to compare results obtained by microscopy to molecular genetic approaches and evaluate the potential of these techniques.

To optimize the strategy to detect diatoms in total DNA isolates of environmental samples, we designed sets of primers, which differed in respect to specificity and to the length of the fragments amplified. Currently, for diatoms, the largest and taxonomically most diverse database exists for the small ribosomal subunit (*SSU*) gene. This gene is highly conserved, and while it does not enable optimal taxonomic resolution, it contains variable regions and has been used for detection and identification from environmental samples in other taxa, including diatoms (Coolen *et al.* 2004; Passmore *et al.* 2006). As a multicopy gene, it is well suited for work with degraded samples (Gugerli *et al.* 2005), and it contains conserved sequence stretches, in which group-specific primers can be placed (Valentini *et al.* 2009). Although the 5.8S+ITS2-region (Moniz & Kaczmarska 2009) has been proposed as the most reliable marker to discriminate between different diatom species, the *SSU* gene currently appears superior for environmental screenings, mainly because more diatom reference sequences are available. Primers for the *SSU* gene to screen for diatoms in ancient or otherwise degraded samples have previously been published, but they either produce relatively long amplicons (~500 bp, Coolen *et al.* 2004) or lack specificity to diatoms (Passmore *et al.* 2006). In this study, we designed diatom-specific primers for amplicons of different lengths and a specific primer pair for stramenopiles (the larger phylogenetic group to which diatoms belong, Adl *et al.* 2005) to evaluate differences in the retrieval of diatoms with respect to amplicon length and primer specificity. We used these primers on total DNA isolates from tropical lake sediments, established a protocol for DHPLC to specifically separate very short PCR products, and compared the results obtained by the molecular genetic assessment to a microscopic screening.

Materials and methods

Primer design

Sequences of the *SSU* gene for representatives of the major divisions of diatoms and of nontarget taxa were obtained from GenBank. The nontarget taxa sequences (i.e. those which the primers were designed *not* to amplify) were selected to obtain a large phylogenetic coverage, with special emphasis on three groups of organisms: (i) taxa with a close phylogenetic relationship to diatoms (i.e. nondiatom stramenopiles), (ii) other organisms likely to be present in sediments and soils (i.e. fungi and organisms with a fungal life mode), (iii) taxa detected in initial PCRs using eukaryote-specific primers. The sequences were aligned using ClustalW in BioEdit (Hall 1999) and searched for target-specific primer-binding regions that amplify a sufficiently variable fragment. Suitable primer sequences for five regions (fragments A–E, Table 1) were determined using the program OLIGO® (Rychlik & Rhoads 1989). For most priming

Table 1 *SSU* gene fragments and primers used in this study

Fragment	Primer Sequence 5′–3′	Name	In combination with	Fragment size (bp)	T_A (°C)	Amplified in lab in addition to diatoms
A	TTGTACACACCGCCCGT	A_1618F	A_1725R	107	56	–
	TGTACACACCGCCCGTC	A_1619F*	A_1723R*	104	53	–
	ACTTCACCTTCCTCTARATGATA	A_1725R				
	TTCACCTTCCTCTARATGATAAG	A_1723R*				
B	GTCCCTGCCCTTTGTACAC	A_1607F	A_1741R	134	56	SAG 243.1, SAG 14.87
	CTGCCCTTTGTACACACCG	A_1611F*	A_1739R*	128	53	SAG 243.1, SAG 14.87
	GGAAACCTTGTTACGACTTCAC	A_1741R				
	AAACCTTGTTACGACTTCACCT	A_1739R*				
C	ACCGTAGTAATTCTAGAGCTAAT	A_144F	Dia516R	341	61	SAG 14.87 at T_A < 61°C
	CCGTAGTAATTCTAGAGCTAATA	A_145F*	Dia516R	340	53	SAG 14.87 at T_A < 61°C
	CTCATTCCAATTGCCAGACC	Dia-516R†				
D	CAATAAATAACAATGCCGGG	A_457F	A_1106R	649	63	–
	ATAAATAACAATGCCGGGCC	A_459F*	A_1106R	647	53	SAG 243.1 at T_A < 62.8°C
	TAAGTTTCAGCCTTGCGAC	A_1106R				
E	GAAAGTTAGGGGATCGA AGAT	A_964F	A_1725R	761	56	–
	AAGTTAGGGGATCGA AGATGA	A_966F*	A_1723R*	757	53	–
GC-clamps‡	gcggcccgccgcccccgccg	GC20§				
	gcccgccgcgccccgcgcccggcccgccgcccccgccccg	GC40¶				

Primers marked with an * were designed as nested primers. Ending on F: forward primer, ending on R: reverse primer. Amplicons A, C, D and E were designed to be diatom stramenopiles. T_A: annealing temperature. †Coolen *et al.* (2004), ‡attached to 5′ of A1619F for DHPLC, §courtesy of Transgenomics, ¶slightly modified from van Hannen *et al.* (1998).

sites, internal primers for nested PCR were also developed. We included two sets of newly designed primers (A: A1618/A1725 and A1619/1723; B: A1607/A1741 and A1611/A1739) amplifying a region previously described for diatom detection (Passmore *et al.* 2006). Note that primers A are diatom specific, while primers B generally amplify stramenopiles (including diatoms). A total of 18 different primers and their corresponding amplificates were analysed further (Table 1), including one reverse primer published by Coolen *et al.* (2004). For amplifications to be analysed by PCR-based DHPLC (only performed for fragment A), the primer A1619 was used with one of two GC clamps (20 and 40 bp, Table 1) added to its 5′end.

Samples

To optimize the protocols, we used pure cultures of single diatom species obtained from the Culture Collection of Algae at the University of Goettingen (SAG). Subsequent testing of the protocols on environmental samples was carried out on integrated surface sediment samples from Lake Naivasha in the Kenyan Rift Valley (Table 2). Lake Naivasha is a slightly alkaline (pH 8.5), shallow (maximum depth 4–6 m; excluding Crescent Island Crater), but quite large (~135 km^2) lake located at 1885 m above sea level in the central valley of Kenya's Eastern Rift Valley (Verschuren 2001). Sediment samples from within Lake Naivasha (NSA 14 and NSA 20) were taken from a boat using an Ekman-Birge grabber, subsampling the large Ekman-Birge sample with a clean spatula. Additionally, one sample from the lakeside above the shoreline (NSA 9) was taken with a tulip planter, subsampling again with a cleaned spatula. The samples were chosen according to the following rationale: samples NSA 14 and NSA 20 are below the water column and presumably contain diatom DNA that is more readily retrievable, while sample NSA 9 is above the present shoreline and presumably contains only little diatom DNA.

The samples were stored in Queens Tissue Buffer (Seutin *et al.* 1991, 20% DMSO, 0.25 M EDTA, saturated with NaCl, pH 8.0) in sterile falcon tubes, kept in the dark and stored at 10 °C after arrival in the laboratory.

DNA isolation

DNA from single species was isolated using the Plant Spin Mini Kit (Invitek). Cultures were transferred to

Table 2 DNA template information

Sediment samples from Lake Naivasha		
Sample	Sampling location	Water depth
NSA 14	End of Oloidien Bay (0° 49′ S; 36° 18′ E)	40 cm
NSA 20	Hippo Point Bay (0° 47′ S; 36° 19′ E)	300 cm
NSA 9	Fisherman's Camp (0° 49′ S; 36° 20′ E)	Above shoreline

Single diatom species*	
Strain	Species
SAG 1020-1a	*Cyclotella meneghiniana*
SAG1020-1b	*Thalassiosira pseudonana*
SAG 122.79	*Thalassiosira weisflogii*
SAG 49.91	*Stephanodiscus minutulus*
SAG 40.96	*Navicula salinicola*
SAG 48.91	*Nitzschia curvilineata*
SAG 80.94	*Pinnularia neomaior*
SAG 1032-1	*Gomphonema parvulum*
SAG 1050-3	*Navicula pelliculosa*
SAG 1052-2	*Nitzschia frustulum*
SAG 1052-3a	*Nitzschia palea*

Single nondiatom species*	
Strain	Species
SAG 243.1	*Monoraphidium minutum*
SAG 14.87	*Nannochloris atomus*

Mixed templates from single diatom species*	
Name	Included species (strain numbers)
Mix 3	SAG 40.96, SAG 122.79, SAG 1050-3
Mix 4	SAG 40.96, SAG 122.79, SAG 1050-3, SAG 1020-1b
Mix 5	SAG 40.96, SAG 122.79, SAG 1050-3, SAG 1020-1b, SAG 49.91
Mix 6	SAG 40.96, SAG 122.79, SAG 1050-3, SAG 1020-1b, SAG 49.91, SAG 48.91
Mix 7	SAG 40.96, SAG 122.79, SAG 1050-3, SAG 1020-1b, SAG 49.91, SAG 48.91, SAG 1052-2
Mix 8	SAG 40.96, SAG 122.79, SAG 1050-3, SAG 1020-1b, SAG 49.91, SAG 48.91, SAG 1052-2, SAG 1052-3a
Mix 9-1	SAG 40.96, SAG 122.79, SAG 1050-3, SAG 1020-1b, SAG 49.91, SAG 48.91, SAG 1052-2, SAG 1052-3a, SAG 80.94
Mix 9-2	SAG 40.96, SAG 122.79, SAG 1050-3, SAG 1020-1b, SAG 49.91, SAG 48.91, SAG 1052-2, SAG 1052-3a, SAG 1032-1
Mix 10	SAG 40.96, SAG 122.79, SAG 1050-3, SAG 1020-1b, SAG 49.91, SAG 48.91, SAG 1052-2, SAG 1052-3a, SAG 1032-1, SAG 80.94

*From the Culture Collection of Algae at the University of Goettingen.

the lysis buffer either directly from solid agar medium with a sterile pipette tip, or 1 mL of liquid culture was centrifuged, the culture medium discarded and lysis buffer added. Cell disruption was carried out with a Mini-BeadBeater (BioSpec/Glen Mills Inc., beating 4800 rpm for 50 s after adding approximately 15 0.5-mm glass beads that had been cleaned with HCl and ethanol) after transfer of the cells to the lysis buffer.

DNA from sediment samples was isolated using the PowerSoil™ DNA Isolation Kit (MoBio Laboratories, California). First, the PowerSoil™-bead solution was transferred from the bead solution tube to a sterile Eppendorf tube. The samples and buffer were mixed well just prior to isolation, and 500 μL of each sample/buffer-mixture was transferred to the empty-bead solution tube. The sample was centrifuged at 13 200 *g* for 10 min, and the supernatant was removed (remaining sediment weight

Table 3 Gradient information for the conditions used to analyse PCR products from Lake Naivasha sediment samples on the WAVE®-DHPLC system

Gradient	Time (min)	Buffer A (%)	Buffer B (%)
Loading	0	65	35
Start gradient	1	60	40
Stop gradient	12	42	58
Start clean	12.1	0	100
Stop clean	13.1	0	100
Start equilibrate	13.2	65	35
Stop equilibrate	15.2	65	35

The programme was run at an oven temperature of 64.5 °C and a flow rate of 0.9 mL/min. Amplicons were generated with primers A1619 and A1723, with a GC-clamp (20 or 40 bp) at the 5′ end of the forward primer.

150–200 mg). The PowerSoil™-bead solution and the buffer C1 from the kit were then added, and the sample was vortexed briefly. Release of the DNA from the cells and/or sediment matrix was carried out by incubating the sample with Proteinase K (55 °C, 1 h), then heating it to 70 °C for 10 min and finally bead-beating with the Mini-BeadBeater for 50 s at 3600 rpm. Before continuing with the isolation procedure, 5 μL of RNA solution (1 μL/μg) was added to facilitate precipitation of nucleic acids. The remaining protocol was carried out according to the manufacturer's instruction.

The isolation was carried out in a dedicated DNA isolation laboratory after thorough cleaning of surfaces with DNA-ExitusPlus (AppliChem), with no samples except the ones isolated present at the time of isolation. The sediment samples were isolated on different days each.

PCR, cloning and sequencing

All PCRs were performed in a total volume of 37.5 μL, containing 1 mM Tris–HCl, pH 9.0, 5 mM KCl, 0.15 mM MgCl2, 0.2 mM of each dNTP, 0.13 μM of both forward and reverse primers, and 0.75 U *Taq* polymerase (Qbiogene). PCR buffers in the initial amplification experiments contained BSA. PCR tubes were transferred from ice to a thermocycler (Biometra T-Gradient or T3000) heated to 94 °C. The cycling profile was as follows: 94 °C for 2 min, 35 cycles of 94 °C for 30 s, the primer-specific annealing temperature (T_A in Table 1) for 45 s, 72 °C for 2 min, and a final extension at 72 °C for 15 min.

Specificity of the primer combinations was tested empirically at T_A 53 °C using a set of diatom (10 algae cultures) and nondiatom (two algae cultures) templates (Table 2). If amplification of nondiatom DNA was observed, a temperature-gradient PCR was run, using the respective nondiatom templates and the diatom *Cyclotella meneghiniana* as a positive control. Specificity results and annealing temperatures chosen for sediment samples are shown in Table 1.

PCR products of the sediment DNA isolates were purified using the NucleoSpin®ExtractII-Kit (Macherey-Nagel), and 5 μL of the cleaned PCR product was subjected to a reconditioning PCR (Thompson *et al*. 2002) prior to cloning with the Topo TA cloning kit (Invitrogen). Of each cloning reaction, 48 clones were re-amplified. If possible, 24 positive clones were sequenced, otherwise as many as possible (cf. Table 4).

PCR-based DHPLC analysis

A PCR-based DHPLC analysis was performed for the amplificates retrieved with the primers A: A1619-A1723, using the Transgenomic WAVE® System with a DNA-Sep® cartridge and an integrated FCW-200 fragment collector (Transgenomic, Omaha, NE). The analyses were performed using a two-eluant buffer system. Buffer A consisted of an aqueous solution of 0.1 M TEAA, pH 7.0 and buffer B consisted of an aqueous solution of 0.1 M TEAA, pH 7.0, with 25% (v/v) acetonitrile. Chromatograms were recorded at a wavelength of 260 nm. Forward primers with two different GC clamps (20 and 40 bp, cf. Table 1) added to the 5′end were tested to optimize separation efficiency. Products derived from mixed templates were subjected to a reconditioning PCR prior to the DHPLC run (Thompson *et al*. 2002), using 5 μL of PCR product for each reaction. All PCRs for DHPLC were either set up without BSA or the products were purified prior to analysis to avoid damage to the DNA-Sep® cartridge.

Before beginning the analyses, the optimal oven temperature for the DHPLC for 50% denaturing conditions was theoretically established for the sequence of *Achnanthes bongranii* (GenBank Accession no. AJ535150), using the software Navigator (Transgenomics Inc.). Starting from this (63.5 °C), in the empirical optimization, the temperature was changed in 0.5 °C intervals (both decreased and increased).

Initial optimization of temperature and gradient conditions were carried out using amplificates of ten single diatom species. For these analyses, 5 μL of PCR product was loaded per run. Gradient conditions were varied within a range starting with 35-45% Buffer B and ending with 58% Buffer B, and gradient durations of 8–20 min, with a flowrate of 0.9 mL/min. After these initial tests with single species PCR products, the established conditions were tested on PCR products obtained from mixtures of single species DNA isolates (3–10 different species in each amplificate, Table 2), loading 20 μL of PCR product per run.

Starting with the conditions established for the single species' tests, we then established optimal conditions for

Table 4 Results of cloning and denaturing high-performance liquid chromatography (DHPLC) fragment collection

Sample and colour in DHPLC (Fig. 3)	*I: Aulacoseira I*		*II: Aulacoseira II*		*III: Amphora*		*IV: Sellaphora*		Nondiatom or not determined[1]		Number of authentic clones
	Fragment	Number of clones (%)	Fragment	Number of clones (%)	Fragment	Number of clones (%)	Fragment	Number of clones (%)	Fragment	Number of clones (%)	
NSA 20	A (I)	13 (65)	A (II)	3 (15%)	A (III)	–	A (IV)	–	A	4 (20%)	20
Orange/red	B	7 (32)	B	–	B	–	B	–	B	15 (68%)	22
	C	6 (46)	C	5 (38%)	C	–	C	–	C	2 (15%)	13
	D	19 (91)	D	2 (9%)	D	–	D	–	D	–	21
	E	9 (50)	E	3 (17%)	E	–	E	–	E	6 (33%)	18
NSA14	A (I)	2 (9)	A (II)	6 (27%)	A (III)	10 (46%)	A (IV)	–	A	4 (18%)	22
Green	B	–	B	3 (13%)	B	3 (13%)	B	–	B	17 (74%)	23
	C	1 (4)	C	2 (9%)	C	5 (22%)	C	5 (22%)	C	10 (43%)	23
	D	–	D	–	D	22 (100%)	D	–	D	–	22
	E	–	E	1 (9%)	E	8 (73%)	E	–	E	2 (18%)	11
NSA9	A (I)	–	A (II)	5 (31%)	A (III)	–	A (IV)	–	A	11 (69%)	16
Blue	B	–	B	–	B	–	B	1 (6%)	B	16 (94%)	17
	C	–	C	–	C	–	C	–	C	6 (100%)	6
	D	–	D	1 (100%)	D	–	D	–	D	–	1
	E	–	E	1 (10%)	E	–	E	–	E	9 (90%)	10

Fragments generated with the following primers: (A) A1619-A1723 (length without primers: 104 bp), (B) A1611-A1739 (length: 128 bp), (C) A145-Dia516 (length: 340 bp), (D) A459-A1106 (length: 647 bp), (E) A966-1723 (length: 757 bp). Clones or sequence types that are highlighted were identified as peaks and fragment collected using DHPLC. Number of peak in Fig. 3 is given in parentheses after fragment A.

[1]Sequences were considered 'not determined' if they were <90% identical to sequences in GenBank as found in a BLAST search.

the separation of PCR products from the sediment samples. Twenty microlitres of PCR product was loaded per run. Optimal gradient conditions established for the analysis of the sediment samples are listed in Table 3. We tested the reproducibility in the DHPLC signals, performing runs with independent PCR amplifications of the same environmental samples.

Using the FCW-200 fragment collector, we collected all pronounced peaks of the sediment sample PCR products. Collection intervals were defined manually (using the one-vial-mode in the Navigator Software) after a sample had been run once, and the collection run was carried out directly afterwards. Multiple PCR products of each sample were run (NSA 20: 4 runs, NSA 14: 3 runs, and NSA 9: 2 runs), including the PCR products that had been cloned. The sampled fragments were re-amplified using the same primers, and the products were rerun on the WAVE®. Products resulting in a single peak after re-amplification were sequenced. Additionally, the PCR product of peak IV (cf. Fig. 3) was cloned as described earlier, and 20 positive clones were sequenced.

Postsequencing Analysis

Sequences derived from clones and from the DHPLC fragment collection were aligned using ClustalW in BioEdit (Hall 1999). To exclude PCR artefacts known to appear in such cloning experiments, sequence differences between otherwise highly similar sequences (identity > 95%) were only considered authentic, if they were detected in three or more clones. The sequences retrieved with primers D (approximately 650 bp) and E (approximately 750 bp) were subjected (as complete alignments of all clones of a certain fragment) to chimera analyses using the program Bellerophon (Huber *et al.* 2004), and sequences identified as putative chimeras after an additional manual check were excluded from further analysis. BLAST searches on NCBI (http://blast.ncbi.nlm.nih.gov) were performed for all sequences considered authentic.

Microscopic analysis of diatoms in the sediments

Permanent diatom slides were prepared for the samples NSA 14 and NSA 20 according to Battarbee *et al.* (2001), and the diatoms were checked according to the morphology of their silicon frustules. Diatom valves were counted in transects at 1000× using a light microscope (Leica DM 4000 B) with an oil immersion objective and identified with reference to Hustedt (1949), Gasse (1986) and Krammer & Lange-Bertalot (1986, 1988, 1991a,b). Percentages of diatoms were determined as the relative amount of each species compared to the total amount of diatoms counted.

Results

Success of PCR-based diatom detection

Sequences were obtained from (i) a total of 261 positive clones from the three samples NSA 14, NSA20 and NSA9 and five different *SSU*-fragments, yielding 241 authentic clone sequences and (ii) direct sequencing of peaks collected from the DHPLC profiles of the amplificates obtained with primers A in all three samples (Table 4). A total of 18 sequence types were identified as diatoms (GenBank Accession nos FJ012322–FJ012347, Table 5), and they were inferred to represent four species (i.e. four inferred *SSU* contigs), as the PCR products of the different fragments A–E always revealed a similar or identical species composition (often referring to identical GenBank entries). The inferred contigs were named accordingly: (I) Aulacoseira 1, (II) Aulacoseira 2, (III) Amphora and (IV) Sellaphora (Table 5).

Sequences of the inferred contig IV were found in the samples NSA 14 and NSA 9 with the fragments A, B and C, and identified as different species of Naviculales, either *Sellaphora* or closely related taxa (*Navicula*, *Mayamaea*). For reasons of clarity and parsimony, we include these different sequence types of Naviculales in one single inferred contig IV named Sellaphora (Table 5).

The most diverse sample was NSA 14, which yielded all four different diatom sequence types, when screening the fragments A (by PCR-DHPLC fragment collection) and C (cloning). Sample NSA 20 consistently yielded the two different Aulacoseira sequence types, except for fragment B, which only yielded the otherwise more common Aulacoseira 1. NSA 9 yielded sequence type Aulacoseira 2 for fragments A, D and E, and the Sellaphora sequence type for fragment B.

The diversity found within sample NSA 14 was highest in screens of the fragments A and C, and lowest for the fragment D. Fragment C failed to detect any diatom clones in sample NSA 9, though, and overall amplification and cloning success was substantially lower for fragment C than for fragment A. The sequences of fragment E displayed a high number of putative chimeras (overall: 34%, NSA 20: 25%, NSA 14: 50%, NSA 9: 17%). Fragment B, which was only slightly longer than fragment A but generated using less specific primers (cf. Table 1), displayed the highest number of nondiatom clones overall and failed to detect the otherwise dominant Aulacoseira 2 in sample NSA 9. The number of nondiatom clones was lowest for fragment D.

PCR-based DHPLC analysis

The PCR products of the ten single diatom species eluted as a respective single peak at all conditions tested for DHPLC (cf. Troedsson *et al*. 2008a). Best separation of the

Table 5 Identification of fragments by BLAST relying on best scores

Inferred Contig	Species identified in GenBank	Accession numbers (accessions originating from this study in bold) and per cent identities of different fragments				
		A	B	C	D	E
I Aulacoseira 1		**FJ012322**, **FJ012324**	**FJ012329**	**FJ012333**, **FJ012335**	**FJ012339**	**FJ012343**
	Aulacoseira nyannensis	AJ535187 (100)*	AJ535187 (100)*	AJ535187 (98)	AJ535187 (99)	AJ535187 (99)
	Aulacoseira ambigua	X85404 (100)	X85404 (100)		AY569583 (99)*	
II Aulacoseira 2		**FJ012323**, **FJ012325**, **FJ012328**	**FJ012330**	**FJ012334**, **FJ012336**	**FJ012340**, **FJ012342**	**FJ012344**, **FJ012345**, **FJ012347**
	Aulacoseira granulata v. *angustissima*	AY485493 (100)	AY485493 (100)	AY485493 (100)	AY485493 (99)	AY485493 (99)
	Aulacoseira distans	X85403 (100)		X85403 (100)	X85403 (99)	
III Amphora		**FJ012326**	**FJ012331**	**FJ012337**	**FJ012341**	**FJ012346**
	Amphora lybica	AM501959 (100)	AM501959 (100)	AM501959 (99)	AM501959 (100)	AM501959 (100)
	Amphora pediculus	AM501960 (100)	AM501960 (100)			
	Amphora sp.	AM501957 (100)	AM501957 (100)			
IV Sellaphora		**FJ012327**	**FJ012332**	**FJ012338**		
	Sellaphora auldreekie	EF151965 (97)*	EF151965 (97)*			
	Sellaphora pupula	AJ544652 (97)	AJ544652 (97)			
	Sellaphora cf. *minima*			EF151966 (97)		
	Navicula pelliculosa	AJ544657 (97)	AJ544657 (97)			
	Mayamaea atomus	AM501968 (97)	AM501968 (97)			

*And other accessions of the same species.

single species' PCR products was achieved using a 40-bp GC clamp attached to the forward primer and a temperature of 65 °C: all tested species displayed separate peaks, except for *Gomphonema parvulum* (Peak 3; Fig. 1a), which was undistinguishable from *Nitzschia palea* (Peak 4; Fig. 1a). Pure culture PCR products with a 20-bp GC clamp were best separated at 64.5 °C, but overall distribution of peaks was not as continuous as with 65 °C and the 40-bp GC clamp (Fig. 1b). Separation potential could not be directly attributed to mere sequence divergence: For example, the sequences of *Thalassiosira weisflogii* (Acc. No. FJ012348) and *Pinnularia neomaior* (Acc. No. FJ012349) were identical at only 83.5% of the nucleotide positions, but resulted in a single peak at 64.5 °C in the DHPLC analysis (Fig. 1b). In turn, *Navicula salinicola* (Acc. No. FJ012350) and *Navicula pelliculosa* (Acc. No. FJ012351) shared 91.3% of identical sequence positions, but exhibited the largest difference in retention time (Fig. 1).

When DNA isolates of single species were first mixed and then amplified and separated via DHPLC, resolution was also best at 65 °C with the 40-bp GC clamp (Fig. S1, Supporting information). Resolution of peaks was good when <5 species were amplified together (Fig. S2, Supporting information). When more species were amplified together, peaks seemingly blended and became wider, resulting in fewer peaks than species.

In contrast to the single species' PCR products, separation in the sediment sample PCR products was best at 64.5° for both clamps, as only with the latter temperature a separation between sequence types II and III was achieved (Fig. 2). While the resolution was very similar for both GC clamps, the overall amplification success was higher using a 20-bp GC clamp. Best separation conditions are listed in Table 3. Using these conditions, one run took just over 15 min to complete.

The PCRs and resulting profiles were highly reproducible for each sediment sample. Fig. 3 shows the profiles of different samples, each represented by two different PCR products with nearly identical profiles. The profiles vary between samples, but the dominant peak in NSA 9 occurred at the same retention time as peak II in NSA 14. Overall, four peaks can be recognized, with the following mean retention time differences to peak I: to peak II 0.14 min, to peak III 0.25 min, and to peak IV 0.56 min.

The peaks I-IV correspond to the diatom sequence types I-IV identified in the cloning experiment. They display the following pairwise identities with respect to the sequence of peak I: to peak II 0.98, to peak III 0.88, and to peak IV 0.90.

The sequence types identified as separate peaks and retrieved by fragment collection with DHPLC are highlighted in grey in Table 4. The highest peak in each DHPLC profile always corresponded to the fragment most frequently detected in the clones: NSA 20 yielded only peak I, which contained the diatom sequence type

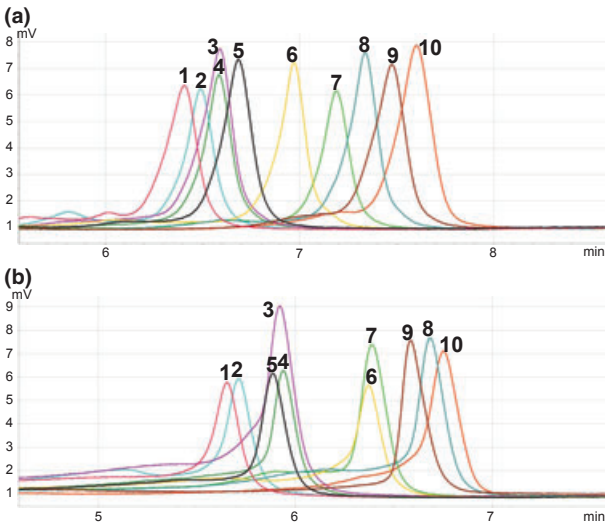

Fig. 1 Profiles of 10 diatom species' PCR products. A: Amplicons with a 40-bp GC clamp attached to the 5′ end of the forward primer, denaturing high-performance liquid chromatography (DHPLC) at 65 °C. Only sample 3 and 4 have an identical peak. B: Amplicons generated with a 20-bp GC clamp, DHPLC at 64.5 °C. The peaks of samples 3/4/5 and 6/7 cannot be discerned. Peaks correspond to the following species: (1) *Navicula salinicola*, (2) *Nitzschia frustulum*, (3) *Gomphonema parvulum*, (4) *Nitzschia palea*, (5) *Thalassiosira pseudonana*, (6) *Thalassiosira weisflogii*, (7) *Pinnularia neomaior*, (8) *Stephanodiscus minutulus*, (9) *Nitzschia curvilineata* and (10) *Navicula pelliculosa*.

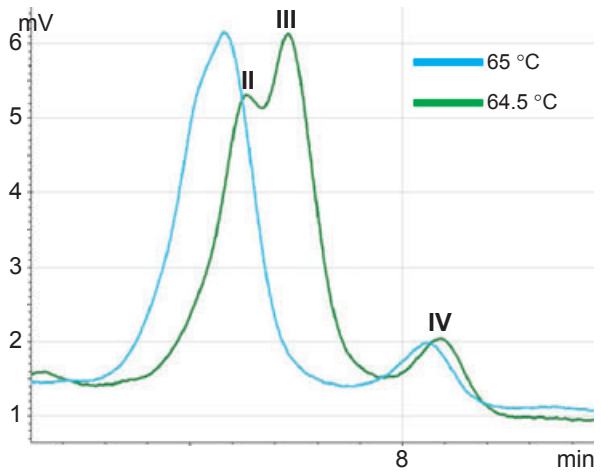

Fig. 2 Two runs of amplificates of the samples NSA14 and a forward primer including a 40-bp GC clamp at 65 and 64.5 °C. Even though the resolution of 40-bp GC clamp-amplificates was higher at 65 °C for single diatom species, in this case, the resolution is higher at 64.5 °C, and the peaks II and III can be separated.

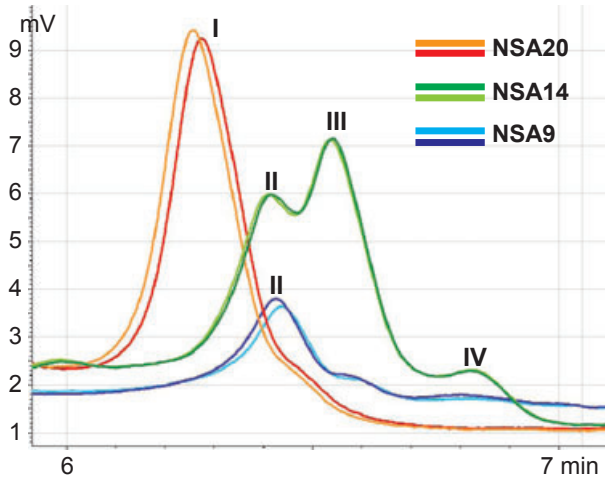

Fig. 3 Denaturing high-performance liquid chromatography profiles of three Lake Naivasha sediment samples. Labelled peaks were sampled and sequenced directly. The sequences I, II and III were also detected in the clone sample set, the sequence IV was not. The sample NSA 20 also contained the sequence of peak II and the sample NSA 14 contained the sequence of peak I, but the profiles of the respective samples do not visibly display these peaks. The two profiles for each sample show the high degree of reproducibility between different PCR products and runs from the same location.

Aulacoseira 1 found in 65% of the clones in the respective cloning experiment. NSA 14 yielded two prominent peaks (II, III), representing sequence types II (Aulacoseira 2) and III (Amphora) found in 27% and 46%, respectively, of the corresponding clones of that sample. It also yielded the small peak IV (Sellaphora), which was not detected in the 22 sequenced clones. NSA9 yielded only peak II, corresponding to Aulacoseira 2, accounting for 31% of the clones of that sample.

Microscopic investigation

The microscopic screening revealed a total of 27 species, with 25 species in NSA 14 and 13 species in NSA 20 (Table 6). Most species were present in only very low numbers. Of all species identified, only 5 were present as more than 5% of total diatoms in either of the samples. The most abundant species were *Aulacoseira ambigua* (NSA 14: 34.31%; NSA 20: 39.44%) and *Aulacoseira granulata* (NSA 14: 25.74%; NSA 20: 44.41%). Other species with abundances over 5% in at least one of the samples were *Amphora lybica* (NSA 14: 5.64%; NSA 20: 3.42%), *Nitzschia intermedia* (NSA 14: 6.86%; NSA 20: 0.31%) and *Synedra acus* (NSA 14: 4.66%; NSA 20: 6.52%). Two species were present with abundances of nearly 4% in sample NSA 14: *Gomphonema gracile* (NSA 14: 3.92%; NSA 20: 0.62%) and *Sellaphora pupula* (NSA 14: 3.92%; NSA 20: 0%).

Table 6 Microscopic analysis of the samples NSA 14 and NSA 20

Species	NSA 14 Number (%)	NSA 20 Number (%)
Amphora lybica	**23 (5.64)**	**11 (3.42)**
Anomoeoneis sphaerophora	1 (0.25)	0 (0)
Aulacoseira ambigua	**140 (34.31)**	**127 (39.44)**
Aulacoseira granulata	**105 (25.74)**	**143 (44.41)**
Aulacoseira granulata var *angustissima*	5 (1.23)	0 (0)
Aulacoseira distans	0 (0)	3 (0.93)
Cyclotella meneghiana	1 (0.25)	0 (0)
Cyclotella stelligera	0 (0)	1 (0.31)
Cymbella muelleri	4 (0.98)	0 (0)
Cymbella ventricosa/turgida	2 (0.49)	0 (0)
Epithemia adnata	1 (0.25)	2 (0.62)
Eunotia pectinalis	2 (0.49)	0 (0)
Gomphonema gracile	**16 (3.92)**	**2 (0.62)**
Gomphonema parvulum	7 (1.72)	3 (0.93)
Navicula pygmea	2 (0.49)	0 (0)
Nitzschia linearis	10 (2.45)	0 (0)
Nitzschia intermedia	**28 (6.86)**	**1 (0.31)**
Nitzschia palea	1 (0.25)	0 (0)
Nitzschia subacicularis	2 (0.49)	0 (0)
Rhopalodia gibba	2 (0.49)	0 (0)
Sellaphora pupula	**16 (3.92)**	**0 (0)**
Surirella engleri	1 (0.25)	2 (0.62)
Synedra acus	**19 (4.66)**	**21 (6.52)**
Synedra acus var *angusstissima*	2 (0.49)	2 (0.62)
Synedra rumpens	14 (3.43)	4 (1.24)
Synedra ulna var *ulna*	3 (0.74)	0 (0)
Thalassiosira faurii	1 (0.25)	0 (0)

Species shaded in grey very likely correspond to the most dominant genetic sequences detected. Species in bold are mentioned in the text.

Discussion

The analysis of molecular genetic data from environmental samples has the potential to provide rapid biodiversity assessments, but still faces methodological challenges. These include the development of taxon-specific markers and increasing the feasibility and speed of such screenings. To this end, we designed and evaluated a set of primers to retrieve genetic data of diatoms from sediments of a tropical, slightly alkaline lake, and we developed a protocol to separate short PCR products (~100 bp without primers) in mixed-template reactions by using denaturing high-performance liquid chromatography (DHPLC). We compared the results retrieved from surface sediment samples employing these molecular genetic techniques to results from a microscopic investigation.

The diversity revealed using the molecular genetic approach mirrored the most dominant taxa identified by microscopic inspection. Only taxa making up over

approximately 5% of the total diatom numbers were detected by molecular genetic means.

The most common species detected microscopically were two species of *Aulacoseira*, which is in accordance with finding two sequence types assigned to the genus *Aulacoseira* in the DNA sequence data. Contrary to this, the dominant sequence in the clones and the DHPLC profiles of the sample NSA 14 is identical with the sequence of a few *Amphora* species, while the *Amphora* species identified by microscopy, *Amphora lybica,* made up only 5.64% of the total diatoms counted. Possibly, the genetic data do not reflect the number of individuals counted, but rather their biomass or biovolume. Using quantitative PCR and group-specific primers, Godhe *et al.* (2008) have shown that there is a significant quantitative correlation between the number of ribosomal DNA copies (i.e. the amount of ribosomal DNA present) and the biovolume of diatom cells in seawater samples. This could lead to DNA from larger celled diatoms being more dominantly represented in the genetic data. Discrepancy between molecular and morphological determinations, even in surface sediments, might also arise from potential differences in DNA degradation between the different taxa (as suggested also by Coolen *et al.* 2006), or from the fact that any step of the genetic analysis potentially constitutes a filter, by which part of the environmental diversity may be lost. Furthermore, the Ekman-Birge sampler used here jointly samples diatoms from different sediment layers (i.e. of different age), and the older cells might contain less DNA. More accurate information could be therefore gained in future studies by analysing different sediment layers in a core.

In our data set, all strictly diatom-specific primers found similar distributions of diatom sequence types for the samples NSA 20 and NSA 14. The primer set C failed to detect any diatom sequences in sample NSA 9, although it revealed maximum diatom diversity in sample NSA 14. Longer fragments amplified with diatom-specific primers still contained a high proportion of diatom sequences, but the diversity was substantially lower. The sequence types revealed were equivalent to the sequence types found most commonly with shorter fragments. The nondiatom-specific primer set amplified a high proportion of nondiatom DNA (68–94%). The diatom diversity revealed by these primers was low, even though the fragment amplified was very short (approximately 130 bp). The importance of primer specificity when searching for representatives of a specific taxon in composite environmental samples has been stressed in other studies (Coolen *et al.* 2004, 2006, 2007) as well. The use of longer fragments seemed to bias the results further towards the taxa most dominant in the molecular genetic assessment, thereby filtering the results.

An additional complication observed in longer fragments was the occurrence of putative chimeric sequences. We found indication for the presence of chimeric amplicons in sequences of approximately 750 bp (fragment E) but not in those of about 100–650 bp (fragments A–D). It has been demonstrated that a high percentage of sequences deposited in GenBank may be compromised by the formation of chimers and other anomalies (Ashelford *et al.* 2005). The appearance of chimeric sequences in PCR products over 650 bp is very likely due to (partial) degradation or shearing of the template DNA over this length, leading to an increase of 'jumping PCR' (Pääbo *et al.* 1990). We only analysed surface samples in this study, but our results are in accordance with previous studies addressing DNA degradation in sediments (Coolen & Overmann 1998; Coolen *et al.* 2006), which found DNA fragments in sediments to be rapidly degraded to a size of approximately 500 bp. Short marker sequences therefore are more reliable, both in recent and in historical environmental screenings.

The results of the PCR-based DHPLC experiment for the shortest amplified fragment corresponded well to the results of the cloning experiment. In initial experiments using single species amplificates, we were able to separate 8 of 10 species at 65 °C using a 40-bp GC clamp, and six of ten single species at 64.5 °C. All in all, optimization of temperature and gradient conditions should be carried out using amplifications of the environmental samples to be analysed and not on amplifications of single species. Resolution in surface sediment samples proved difficult for rare fragments with very little difference in retention times and very little sequence difference to dominant fragments (i.e. peak I: Aulacoseira 1 and peak II: Aulacoseira 2 correspond to sequences of identical length and differing only at two positions). In these cases, the less abundant sequence type could be overlooked. At the same time, sequence type IV, Sellaphora, was not detected in the clones but was detected and collected by PCR-based DHPLC. This amplicon was presumably also infrequent in the PCR product, but had a substantial difference in retention times to the other peaks. The PCR-based DHPLC approach thus displayed a higher overall diversity in the three samples from Lake Naivasha than the cloning experiment.

In short, PCR-based DHPLC analysis on the WAVE®-System proved particularly useful if a limited number of different sequence types were to be separated, and the most dominant sequences in environmental samples could be retrieved for sequencing. Less frequent taxa out of a species-rich group, such as diatoms, might however go undetected. In environmental samples, PCR-based DHPLC seems most suited for less diverse targets or if the retrieval of the most dominant biotic components is sufficient. Such potential applications could include the

analysis of smaller taxonomic groups containing less species or diet analyses from faeces to establish major dietary components.

DHPLC in molecular genetic environmental biodiversity studies can be a rapid method to screen mixed-template amplicons for their major contents and it offers some distinct benefits. For one thing, as we have shown here, DHPLC has the potential to physically separate DNA fragments of very short lengths. Studies to date have typically analysed fragments of 200–500 bp (Barlaan *et al*. 2005; Goldenberg *et al*. 2007; Troedsson *et al*. 2008b), but environmental samples may well contain DNA that is degraded to fragments below this size (Valentini *et al*. 2009).

A further benefit of DHPLC is the possibility to view the PCR product in a semi-quantitative manner by comparing peak heights and overall profiles. This is ideal for environmental monitoring, both on a spatial and on a temporal scale. In addition, this feature allows the researcher to monitor reproducibility and bias within and between samples. This can be especially helpful for mixed-template reactions relying on low DNA concentrations, as these are particularly prone to bias introduced because of stochastic processes in the PCR (Polz & Cavanaugh 1998). Finally, a huge advantage of DHPLC lies in the automation of the procedure, the short analysis time and the ease of fragment collection.

Acknowledgements

This project was carried out in the Graduate School GRK 1364 on 'Shaping Earth's Surface in a Variable Environment' funded by the German Research Foundation (DFG), the State of Brandenburg and the University of Potsdam. The work also received financial support by several grants to M.T. by the German Research Foundation (DFG). We thank the Government of Kenya (Research Permits MOST13/001/30C 59/10, 59/18 and 59/22) and the University of Nairobi for research permits and support. We gratefully acknowledge technical support provided by Anja Schneider. Ursula Gaedke kindly provided algal cultures.

References

Adl SM, Simpson AGB, Farmer MA *et al*. (2005) The new higher level classification of eukaryotes with emphasis on the taxonomy of protists. *Journal of Eukaryotic Microbiology*, **52**, 399–451.

Ashelford KE, Chuzhanova NA, Fry JC, Jones AJ, Weightman AJ (2005) At least 1 in 20 16S rRNA sequence records currently held in public repositories is estimated to contain substantial anomalies. *Applied and Environmental Microbiology*, **71**, 7724–7736.

Barlaan EA, Sugimori M, Furukawa S, Takeuchi K (2005) Profiling and monitoring of microbial populations by denaturing high-performance liquid chromatography. *Journal of Microbiological Methods*, **61**, 399–412.

Battarbee RV, Jones VJ, Flower RJ *et al*. (2001) Diatoms. In: *Tracking Environmental Change Using Lake Sediments—Volume 3: Terrestrial, Algal, and Siliceous Indicators* (eds Smol JP, Birks HJB & Last WM), pp. 155–202. Kluwer Academic Publishers, Dordrecht, Boston, London.

Binladen J, Gilbert MTP, Bollback JP *et al*. (2007) The Use of Coded PCR Primers Enables High-Throughput Sequencing of Multiple Homolog Amplification Products by 454 Parallel Sequencing. *PLoS ONE*, **2**, e197. doi: 110.1371/journal.pone.0000197.

Birks HJB, Line JM, Juggins S, Stevenson AC, Terbraak CJF (1990) Diatoms and pH Reconstruction. *Philosophical Transactions of the Royal Society B-Biological Sciences*, **327**, 263–278.

Coolen MJL, Overmann J (1998) Analysis of subfossil molecular remains of purple sulfur bacteria in a lake sediment. *Applied and Environmental Microbiology*, **64**, 4513–4521.

Coolen MJL, Muyzer G, Rijpstra WIC *et al*. (2004) Combined DNA and lipid analyses of sediments reveal changes in Holocene haptophyte and diatom populations in an Antarctic lake. *Earth and Planetary Science Letters*, **223**, 225–239.

Coolen MJL, Boere A, Abbas B *et al*. (2006) Ancient DNA derived from alkenone-biosynthesizing haptophytes and other algae in Holocene sediments from the Black Sea. *Paleoceanography*, **21**, PA1005, doi: 1010.1029/2005pa001188.

Coolen MJL, Volkman JK, Abbas B *et al*. (2007) Identification of organic matter sources in sulfidic late Holocene Antarctic fjord sediments from fossil rDNA sequence analysis. *Paleoceanography*, **22**, PA2211, doi: 2210.1029/2006pa001309.

Deagle BE, Kirkwood R, Jarman SN (2009) Analysis of Australian fur seal diet by pyrosequencing prey DNA in faeces. *Molecular Ecology*, **18**, 2022–2038.

Ficetola GF, Miaud C, Pompanon F, Taberlet P (2008) Species detection using environmental DNA from water samples. *Biology Letters*, **4**, 423–425.

Fritz SC, Juggins S, Battarbee RW, Engstrom DR (1991) Reconstruction of past changes in salinity and climate using a diatom-based transfer-function. *Nature*, **352**, 706–708.

Gasse F (1986) *East African Diatoms: Taxonomy, Ecological Distribution J*. Cramer, Berlin, Stuttgart.

Gasse F, Juggins S, Khelifa LB (1995) Diatom-based transfer-functions for inferring past hydrochemical characteristics of African Lakes. *Palaeogeography Palaeoclimatology Palaeoecology*, **117**, 31–54.

Godhe A, Asplund ME, Harnstrom K *et al*. (2008) Quantification of diatom and dinoflagellate biomasses in Coastal Marine Seawater samples by real-time PCR. *Applied and Environmental Microbiology*, **74**, 7174–7182.

Goldenberg O, Herrmann S, Marjoram G *et al*. (2007) Molecular monitoring of the intestinal flora by denaturing high performance liquid chromatography. *Journal of Microbiological Methods*, **68**, 94–105.

Gugerli F, Parducci L, Petit RJ (2005) Ancient plant DNA: review and prospects. *New Phytologist*, **166**, 409–418.

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, **41**, 95–98.

van Hannen EJ, van Agterveld MP, Gons HJ, Laanbroek HJ (1998) Revealing genetic diversity of eukaryotic microorganisms in aquatic environments by denaturing gradient gel electrophoresis. *Journal of Phycology*, **34**, 206–213.

Huber T, Faulkner G, Hugenholtz P (2004) Bellerophon: a program to detect chimeric sequences in multiple sequence alignments. *Bioinformatics*, **20**, 2317–2319.

Hustedt F (1949) Süßwasserdiatomeen aus dem Albert-Nationalpark in Belgisch Kongo. In: *Exploration du Parc National Albert (1935–1936)* (ed. Damas H), pp. 117. Hayez, Bruxelles.

Jumpponen A, Jones KL, Mattox D, Yaege C (2010) Massively parallel 454-sequencing of fungal communities in *Quercus spp*. ectomycorrhizas indicates seasonal dynamics in urban and rural sites. *Molecular Ecology*, **19**, 41–53.

Krammer K, Lange-Bertalot H (1986) *Bacillariophyceae—1. Teil: Naviculaceae*. Gustav Fisher, Stuttgart.

Krammer K, Lange-Bertalot H (1988) *Bacillariophyceae – 2. Teil: Bacillariaceae, Epithemiaceae, Suriellaceae*. Gustav Fisher, Stuttgart.

Krammer K, Lange-Bertalot H (1991a) *Bacillariophyceae—3. Teil: Centrales, Fragilariaceae, Eunotiaceae*. Gustav Fisher, Stuttgart.

Krammer K, Lange-Bertalot H (1991b) *Bacillariophyceae—4. Teil: Achnanthaceae, Kritische Ergaenzungen zu Navicula (lineolatae) und Gomphonema, Gesamtliteraturverzeichnis Teil 1-4*. Gustav Fischer, Stuttgart.

Mann DG, Droop SJM (1996) Biodiversity, biogeography and conservation of diatoms. *Hydrobiologia*, **336**, 19–32.

Moniz MBJ, Kaczmarska I (2009) Barcoding diatoms: is there a good marker? *Molecular Ecology Resources*, **9**, 65–74.

Pääbo S, Irwin DM, Wilson AC (1990) DNA damage promotes jumping between templates during enzymatic amplification. *Journal of Biological Chemistry*, **265**, 4718–4721.

Passmore AJ, Jarman SN, Swadling KM *et al.* (2006) DNA as a dietary biomarker in antarctic krill, *Euphausia superba*. *Marine Biotechnology*, **8**, 686–696.

Polz MF, Cavanaugh CM (1998) Bias in template-to-product ratios in multitemplate PCR. *Applied and Environmental Microbiology*, **64**, 3724–3730.

Rychlik W, Rhoads RE (1989) A computer-program for choosing optimal oligonucleotides for filter hybridization, sequencing and invitro amplification of DNA. *Nucleic Acids Research*, **17**, 8543–8551.

Seutin G, White BN, Boag PT (1991) Preservation of avian blood and tissue samples for DNA analyses. *Canadian Journal of Zoology-Revue Canadienne De Zoologie*, **69**, 82–90.

Soininen EM, Valentini A, Coissac E *et al.* (2009) Analysing diet of small herbivores: the efficiency of DNA barcoding coupled with high-throughput pyrosequencing for deciphering the composition of complex plant mixtures. *Frontiers in Zoology*, **6**, 16, doi: 10.1186/1742-9994-1186-1116.

Taberlet P, Coissac E, Pompanon F *et al.* (2007) Power and limitations of the chloroplast trnL (UAA) intron for plant DNA barcoding. *Nucleic Acids Research*, **35**, e14.

Thompson JR, Marcelino LA, Polz MF (2002) Heteroduplexes in mixed-template amplifications: formation, consequence and elimination by 'reconditioning PCR'. *Nucleic Acids Research*, **30**, 2083–2088.

Troedsson C, Lee RF, Stokes V *et al.* (2008a) Development of a denaturing high-performance liquid chromatography method for detection of protist parasites of metazoans. *Applied and Environmental Microbiology*, **74**, 4336–4345.

Troedsson C, Lee RF, Walters T *et al.* (2008b) Detection and discovery of crustacean parasites in blue crabs (*Callinectes sapidus*) by using 18S rRNA gene-targeted denaturing high-performance liquid chromatography. *Applied and Environmental Microbiology*, **74**, 4346–4353.

Valentini A, Pompanon F, Taberlet P (2009) DNA barcoding for ecologists. *Trends in Ecology & Evolution*, **24**, 110–117.

Verschuren D (2001) Reconstructing fluctuations of a shallow East African lake during the past 1800 yrs from sediment stratigraphy in a submerged crater basin. *Journal of Paleolimnology*, **25**, 297–311.

Xiao WZ, Oefner PJ (2001) Denaturing high-performance liquid chromatography: a review. *Human Mutation*, **17**, 439–474.

Data accessibility

DNA sequences: GenBank accessions FJ012322–FJ012347. More detailed information about DNA sequences obtained from respective samples see data deposited at Dryad: doi:10.5061/dryad.90111.

Supporting Information

Additional supporting information may be found in the online version of this article.

Fig. S1 Resolution of PCR products derived from DNA mixes of single species (cf. Table 2) employing different GC clamps and temperatures. The best resolution for the mixes, as for the single species amplified separately is achieved using a 40-bp GC clamp and running the DHPLC at 65 °C.

Fig. S2 DHPLC profiles of PCR products derived from DNA mixes of single species (cf. Table 2). Separation of products was possible, when <5 species were included in the mix. Inclusion of more species, and templates which resulted in more closely eluting peaks resulted in broader and larger peaks that could no longer be discriminated.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.