

Development of genetic probes for the identification of marine microalgae

Progress Report, Part 2, June 2009

Paolo Momigliano, David Blair and Kirsten Heimann
School of Marine and Tropical Biology, James Cook University



Australian Government

**Department of the Environment,
Water, Heritage and the Arts**

Supported by the Australian Government's
Marine and Tropical Sciences Research Facility
Project 2.6.1 Understanding threats and impacts
of invasive pests on the Great Barrier Reef

© James Cook University.

This report should be cited as:

Momigliano, P., Blair, D. and Heimann, K. (2009) *Development of genetic probes for the identification of marine microalgae*. June Interim Report (Part 2) to the Marine and Tropical Sciences Research Facility. Reef and Rainforest Research Centre Limited, Cairns (13pp.).

Published by the Reef and Rainforest Research Centre on behalf of the Australian Government's Marine and Tropical Sciences Research Facility.

The Australian Government's Marine and Tropical Sciences Research Facility (MTSRF) supports world-class, public good research. The MTSRF is a major initiative of the Australian Government, designed to ensure that Australia's environmental challenges are addressed in an innovative, collaborative and sustainable way. The MTSRF investment is managed by the Department of the Environment, Water, Heritage and the Arts (DEWHA), and is supplemented by substantial cash and in-kind investments from research providers and interested third parties. The Reef and Rainforest Research Centre Limited (RRRC) is contracted by DEWHA to provide program management and communications services for the MTSRF.

This publication is copyright. Apart from any use as permitted under the Copyright Act 1968, no part may be reproduced by any process without prior written permission from the Commonwealth. Requests and enquiries concerning reproduction and rights should be addressed to the Commonwealth Copyright Administration, Attorney General's Department, Robert Garran Offices, National Circuit, Barton ACT 2600 or posted at <http://www.ag.gov.au/cca>.

The views and opinions expressed in this publication are those of the authors and do not necessarily reflect those of the Australian Government or the Minister for the Environment, Water, Heritage and the Arts or Minister for Climate Change and Water.

While reasonable effort has been made to ensure that the contents of this publication are factually correct, the Commonwealth does not accept responsibility for the accuracy or completeness of the contents, and shall not be liable for any loss or damage that may be occasioned directly or indirectly through the use of, or reliance on, the contents of this publication.

This report is available for download from the Reef and Rainforest Research Centre Limited website:
http://www.rrrc.org.au/mtsrf/theme_2/project_2_6_1.html



June 2009

Abstract

Molecular methods are being increasingly used to supplement, and in some cases replace, morphological approaches to identification of toxic dinoflagellates. Regions within the nuclear ribosomal RNA gene operon are frequently targeted for such work, and many sequences are available for comparison in public databases. We have amplified and sequenced a portion of the 28S rRNA gene from a subset of dinoflagellate cultures, isolated and held in the North Queensland Algal Identification Facility (NQAIF). In some cases, strong matches were found with known species, providing unambiguous identification. In other cases, matches were less convincing; suggesting that our samples came from a particular genus but that the actual species was not represented in the databases. In rare cases, matches found with publicly available sequences were not with dinoflagellates. It proved possible to amplify and sequence DNA from single cells, aliquots of a few cells, and larger numbers of cells spun down from cultures. The ability to work with single cells will permit us to sequence DNA from individual field-collected dinoflagellates.

Acknowledgements

We wish to acknowledge funding from the Marine and Tropical Sciences Research Facility.

Introduction

Many species of dinoflagellates are capable of producing toxins which can accumulate in the tissues of marine fish and shellfish, creating a risk of poisoning in people consuming seafood. Because of a high degree of morphological similarities between members of different genera and between species of the same genus, accurate identification of dinoflagellate species is a task requiring considerable skills which only a few experts possess and time. Even so, many taxa consist of 'strains' and of clusters of morphologically identical cryptic species. The existence of these has only come to light with the advent of molecular techniques such as the ability of sequence targeted regions of the genome.

Typically, dinoflagellates are identified from clones grown in pure culture. This provides adequate numbers of cells to work with for identification purposes. However, a further difficulty in working with dinoflagellates is that many species cannot be cultured. Thus, an unknown but probably large proportion of dinoflagellate species remain unidentified.

Molecular approaches are making it possible to obtain DNA sequences from cultured dinoflagellates and even from single field-collected cells. The DNA sequences obtained can be compared with the large collections of sequences in public databases (e.g. GenBank: www.ncbi.nlm.nih.gov). Identification to strain, species or genus is possible in many cases. Of course, given the poor baseline of information on dinoflagellates in the Great Barrier Reef (GBR) Province, we might also expect that many sequences would have no matches among published data.

At present, ribosomal RNA genes are the favoured multi-copy targets in the nuclear genome. This is because a) many copies are present, in tandem array, within a single nucleus; b) the process of 'concerted evolution' results in identical copies throughout the array; c) tracts conserved across great phylogenetic range, useful for the design of primers of broad specificity, alternate with highly variable regions that may differ even between 'strains' of the same species. Figure 1 shows the structure of the eukaryotic ribosomal RNA operon.

Our objective, in this report, is to demonstrate the feasibility of using nuclear ribosomal RNA sequences for identification of dinoflagellates in the GBR region. We have initially focused on the 28S gene. To date, all of the material we have worked with has been from cultures, isolated from the Great Barrier Reef and established at the North Queensland Algal Identification/Culturing Facility (NQAIF) at James Cook University, but we will also demonstrate that DNA can be amplified from single cells. We will show that some of the dinoflagellates cultured from GBR waters can be identified to species using this approach. Others belong to species for which there are no matching sequences in the public databases.

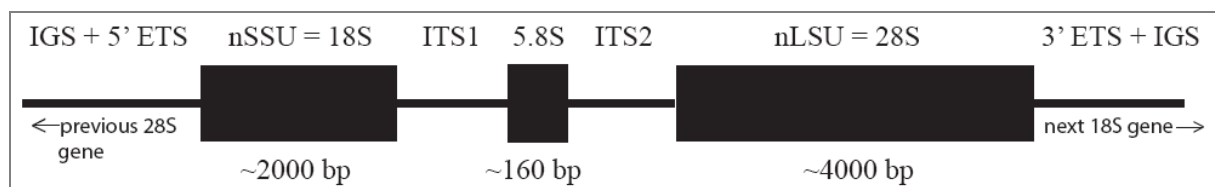


Figure 1: Structure of a single eukaryotic ribosomal RNA gene operon within the tandem array. Boxes indicate the three genes that encode ribosomal RNA used as the structural basis of ribosomes. These are separated by various spacer regions. The 28S gene (and the 18S to a lesser extent) is a mosaic of highly conserved and very variable domains, providing many opportunities for design of probes to whatever level of specificity desired (modified from previous project progress report).

Materials and methods

Dinoflagellate Isolation: Dinoflagellates were isolated from GBR water samples using the microcapillary-capturing-technique at 10x magnification on an inverted light microscope (Olympus CX41). Single cells were dispensed into filtered (0.45µm) and autoclaved seawater (obtained from the outer reef GBR). Cells were allowed to swim for ten minutes and were then recaptured. This procedure was repeated ten times to ensure that nano- and picoplankton, which are invisible at the magnification used to isolate the organisms, were no longer in the vicinity of the dinoflagellate to be isolated (dinoflagellates are strong swimmers). Except for culture NQAIF060, which was donated by S. Murray, all other cultures sequenced were isolated by the NQAIF curator Stanley Hudson.

Dinoflagellate culture maintenance: Cultures were established in L1 medium prepared in natural seawater obtained from the outer GBR and treated as described above. Cultures are maintained at 24°C, a 12:12 h photoperiod and a light intensity of 45 µmol photons m⁻² s⁻¹ in a Contherm cross flow phytoplankton growth chamber. Cultures are subcultured in L1 medium on a four-weekly basis.

Single and 10-cell extraction: For each culture, DNA was extracted separately from single cells and from an aliquot of approximately 10 cells using Chelex® 100 (Walsh *et al.* 1991) following the procedure outlined by Richlen and Barber (2005). Cells, singly or in 10-cell aliquots, were isolated from cultures by micropipetting and washed by transfer into 3 drops of TE buffer (10 mM Tris-HCl, 1mM EDTA, PH 8). Cells in a minimal volume of TE buffer (<2µl) were transferred to 200µl PCR tubes containing 50, 100 and 200µl of 10% Chelex® 100 (Walsh *et al.* 1991). The tubes were quickly spun to push the cells from the supernatant into the Chelex and incubated at 96°C for twenty minutes. Tubes were centrifuged at 10,000 rpm for five minutes and the supernatant was transferred into clean 1.5ml tubes and stored at -20°C prior to PCR amplification. Enzymatic extraction of DNA from single cells using Proteinase K was also attempted, following the procedure illustrated by Ki and others (2005). Single cells were isolated and washed as previously described, and transferred into 200µl PCR tubes in 1µl of TE buffer. 1 µl of Proteinase K (200ug/ml) was added and samples were incubated for fifty minutes at 55°C and Proteinase K was then denatured by incubation at 90°C for fifteen minutes. The tubes were used directly for PCR amplification.

Chelex extraction from pelleted culture: Cultures were pelleted by mild centrifugation (5000 rpm) for five minutes. Pellets of no more than 1mm in diameter were washed three times with 500ul of TE buffer (10 mM Tris-HCl, 1mM EDTA, PH 8) and resuspended in 1.5 ml tubes containing 400 µl of 10% Chelex® 100 and 5 µl of 20 mg/ml Proteinase K. Tubes were shaken vigorously and incubated at 55°C for approximately two hours. Samples were briefly spun at high speed and incubated at 95°C for twenty minutes. Tubes were centrifuged for five minutes at 10,000 rpm and the supernatant was transferred to new microcentrifuge tubes and stored at -20°C.

PCR amplification and sequencing: A fragment of nuclear 28S rDNA encompassing the D1-D6 variable regions (Figure 2) was amplified using the primers D1RFor (Scholin *et al.* 1994) and 1483R shown are locations, orientation and names of many primers designed as part of this work.

PCR reactions were prepared as follows: 1 µl of template DNA (from Chelex extractions), 800µM dNTPs, 5 pmol of each primer, 1x Kapa2G PCR buffer B and 2.5-3.5 mM MgCl₂. Reactions underwent three minutes' initial denaturation at 94°C, and thirty (for template extracted from pelleted cultures) to forty (for template extracted from single cells) cycles of thirty seconds denaturation at 94°C, fifteen seconds annealing at 50°C and one minute extension at 72°C, and a final extension step of three minutes at 72°C. If multiple bands were produced, fragments of the expected size were either excised from agarose gel after

electrophoresis or reamplified by gel-stab hemi-nested PCR (Bjourson and Cooper 1992) using the primers D1RFor and D3B (Nunn *et al.* 1996). Excised bands were cleaned using QIAquick gel extraction kit (Qiagen Ltd) following manufacturer's instructions, and products from hemi-nested PCR were cleaned by isopropanol precipitation. Sequencing reactions were performed at the Australian Genome Research Facility (University of Queensland) using Big Dye Terminator Cycle Kit v3.5 (Applied Biosystems). Capillary separation was performed on a AB 3730xl platform. Sequencing and PCR primers used are listed in Table 1.

Sequence alignment and molecular identification: For each culture the chromatograms of overlapping fragments were aligned using the software Chromas Pro (Technelysium Ltd). The specificity of the sequences obtained was verified using basic nucleotide BLAST (NCBI). Each culture was given a molecular identification at the genus or species level (when possible) based on the result of BLAST search.

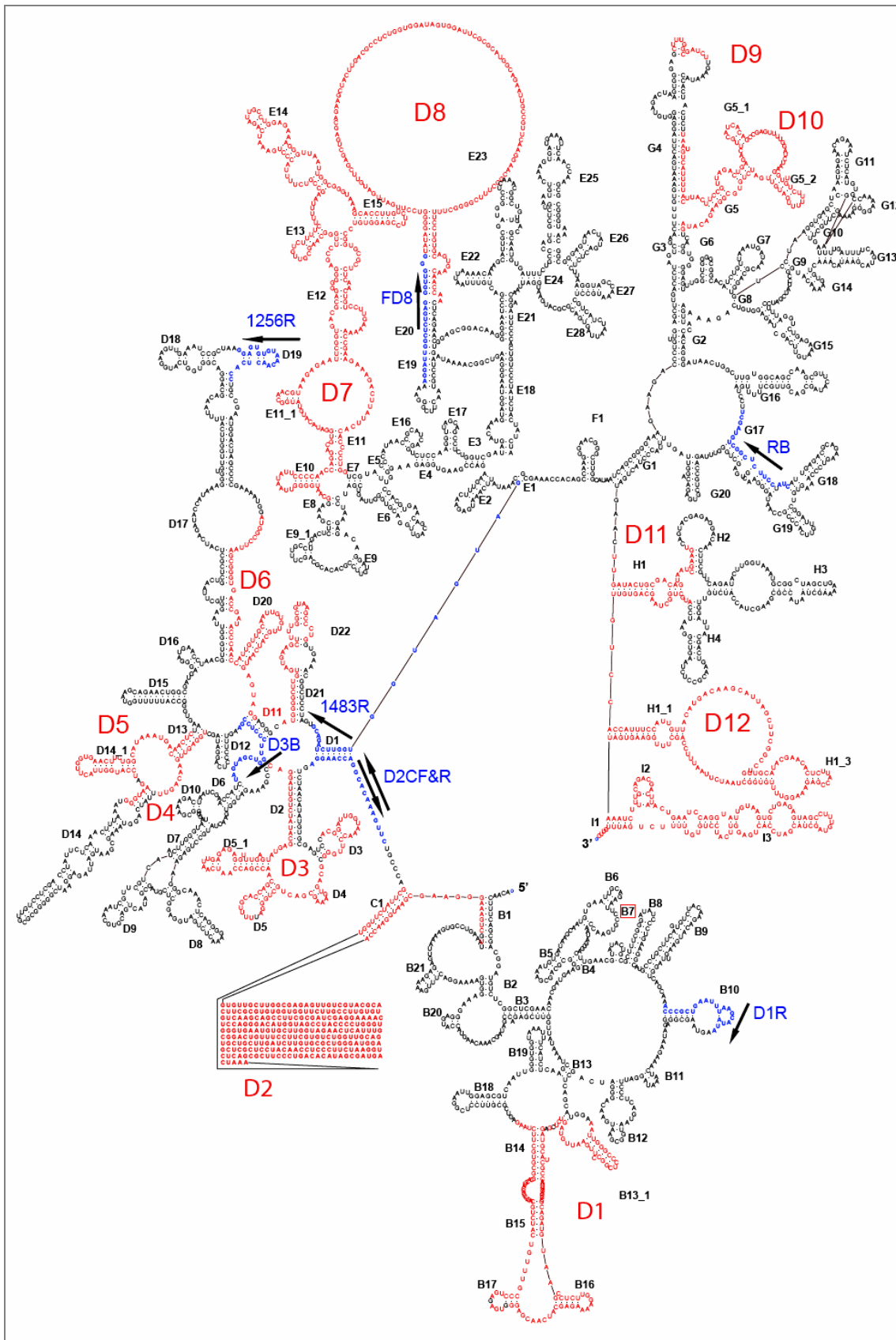


Figure 2: Secondary structure of the 28S rRNA subunit (based on that for *Prorocentrum micans*) showing conserved regions (in black) and variable regions (in red) (Daugbjerg et al. 2000).

Table 1: 28S gene PCR and sequencing primers for the regions D1-D6.

Primer	Sequence	Direction	PCR/ Sequencing	Location (base pairs relative to <i>P.micans</i> 28S)	Reference
D1R For	ACCCGCTGAATTTAAGCATA	Forward	PCR/ Sequencing	21	(Scholin, Herzog <i>et al.</i> 1994)
D2C For	CTTGAAACACGGACCAAGG	Forward	Sequencing	733	New
D3B R	TCGGAGGGAACCAGCTACTA	Reverse	PCR/ Sequencing	1011	(Nunn, Theisen <i>et al.</i> 1996)
1483 R	CTACTACCACCAAGATCTGC	Reverse	PCR	1483	(Daugbjerg, Hansen <i>et al.</i> 2000)
1256 R	GGTGAGTTGTTACACTCC	Reverse	Sequencing	1256	New

Results

(a) Single cell and 10-cell amplifications

As shown in Figure 3, DNA extracted using a chelex protocol from single cells, or from aliquots of ~10 cells, could be successfully amplified using primers D1R and 1483R. A single band of the expected size (~1500 bp) was obtained. Of interest is the fact that DNA extracted from single cells using a proteinase K protocol failed to amplify. For this reason, we standardised on the use of chelex.

b) Pelleted cells from cultures

DNA was extracted from a subset of current viable cultures held in NQAIF. Sequence was obtained from most of these samples using the primers mentioned above. BLAST searches were done against sequences in the public databases (see Table 2 for a summary of results). BLAST searches will always find a best match with something, even if that something is very different from the sequence submitted. Hence caution is required in interpreting the results. In some cases, the results were sufficient to provide an unambiguous identification. For example, cultures NQAIF35, 60 and 90 matched strongly with sequences of *Coolia monotis*. NQAIF56 matched almost completely with *Prorocentrum lima*. The phylogenetic tree in Figure 4 provides a graphical representation of this match.

In other cases, best matches were with a number of different species in one genus, but with no match being convincing (Table 2). Thus, NQAIF71 showed 87% similarity with an *Amphidinium* species and most other matches were with members of this genus. It is likely that this culture is of an *Amphidinium* species as yet unrepresented in the public databases. For NQAIF42, the best match was with *Scripsiella hangoei*, but this match was poor. We cannot say what the affinities are of NQAIF42 at this stage. The sequence obtained from NQAIF263 is of *Nannochloropsis*, a taxon that does not belong to the Dinophyceae. *Prorocentrum micans* was identified by its characteristic morphological features as the originally cultured organism, however, *Nannochloropsis* sp. must have been present as a low-number contaminant not visible at the magnification used for isolation. Over time, this more competitive organism (division rate 4-5x per day) outgrew *Prorocentrum micans*.

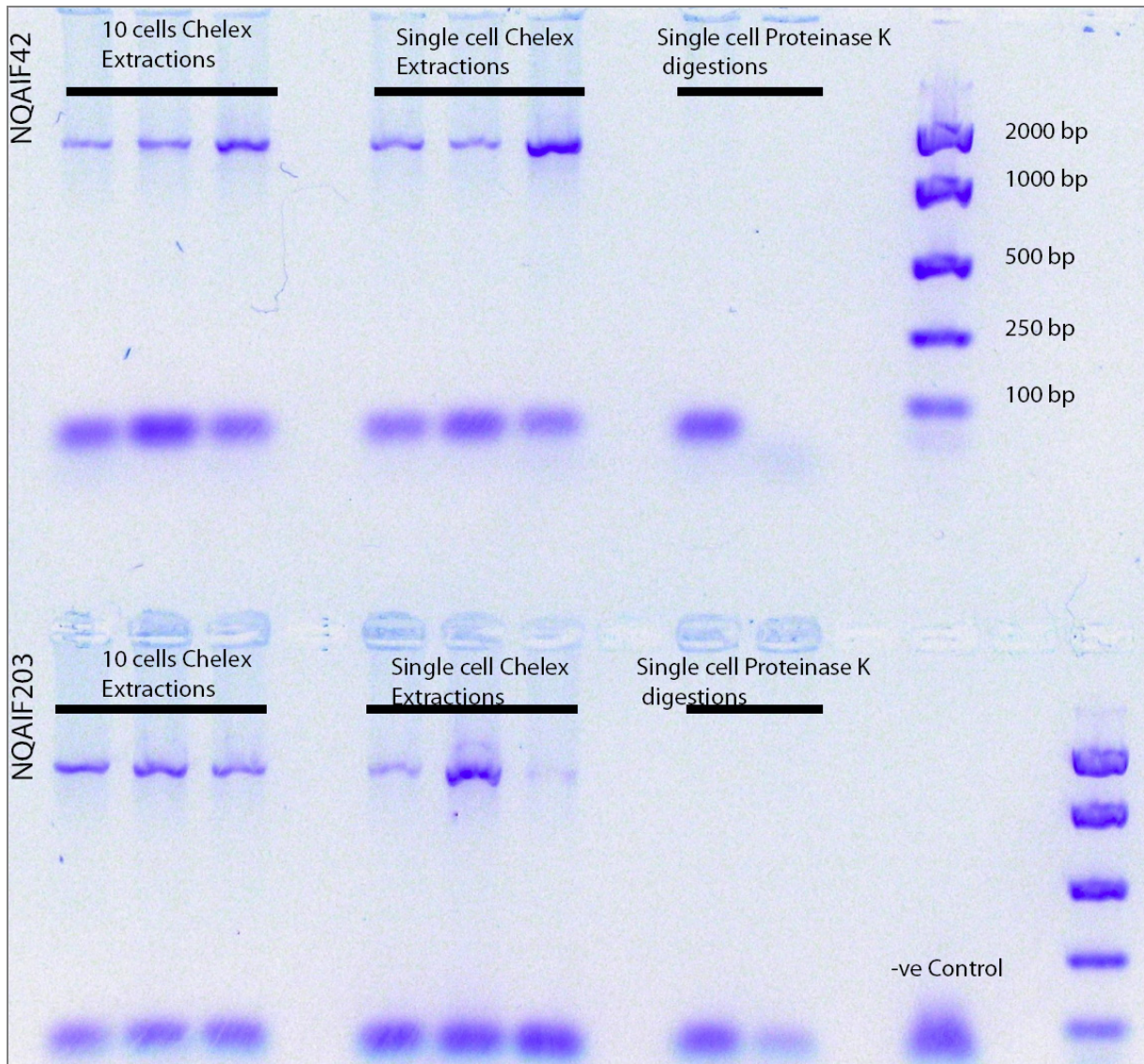


Figure 3: Gel image showing successful amplification of a portion of the 28S rRNA gene from single cells and from aliquots of ~10 cells after chelex extraction of DNA. Cells came from NQAIF cultures 42 and 203. Successful amplification produces a single band of about 1500 bp running to the level shown by the asterisk. Attempts to amplify DNA extracted from single cells following a proteinase K protocol were not successful.

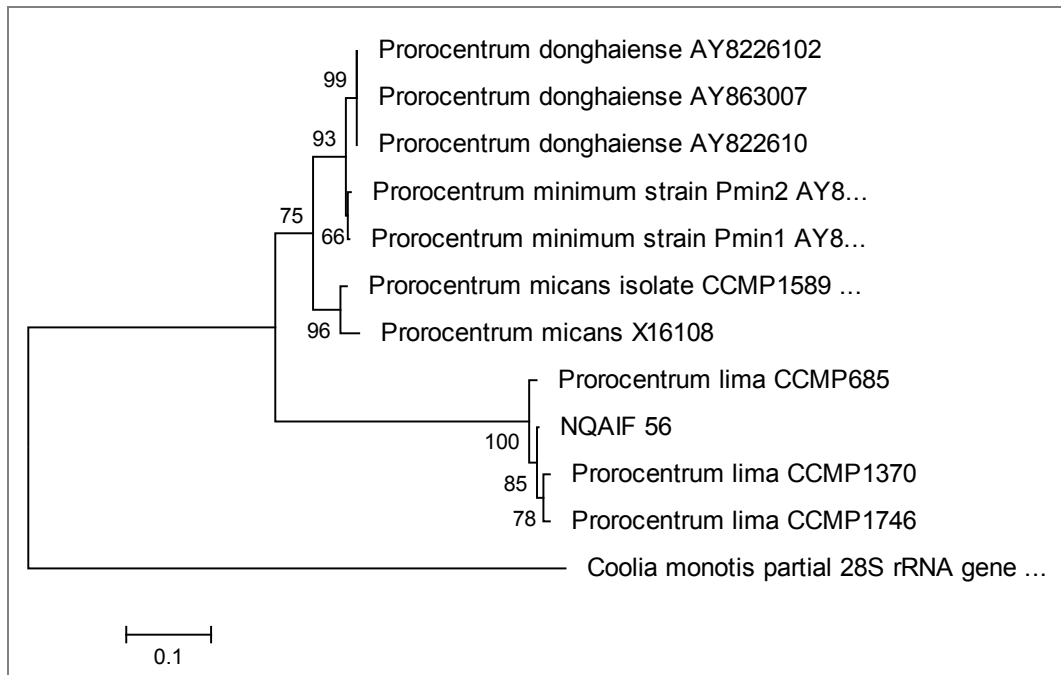


Figure 4: Phylogenetic tree inferred from an alignment of partial 28S sequences obtained from public databases and from NQAIF culture #56. Note that NQAIF56 is nested among published sequences from *Prorocentrum lima*, thus demonstrating that this culture represents that species.

Discussion

Using sequences from the D1-D6 variable regions of the 28S rRNA gene from cultures of marine microalgae held at NQAIF, we were able to identify the organisms in most cultures to genus and/or species. Where exact identification was not possible, this is probably due to that particular species not yet being represented in public databases – at least for 5' half of the 28S gene. Not all studies on dinoflagellates have used sequences from that region. Many have used the 18S gene and/or the ITS2 region (see Figure 1) and a few have used the 3' end of the 28S gene. We have primers allowing us to work with these other gene regions, and are now starting to obtain data from these. This will improve the accuracy and extent of our identifications. Our new data will be deposited in GenBank in the normal course of preparation of these results for publication.

Table 2: Results of molecular identifications of dinoflagellates held in culture at the NQAIF.

NQAIF Culture	Length of 28S sequenced	Length used for alignment	Best match (with GenBank accession number)	% identities	Gaps	Molecular Identification	Notes
NQAIF33	1366 bp	1366 bp	<i>Amphidinium carterae</i> strain SM10 (AY460579)	98%	<1%	<i>Amphidinium carterae</i>	All sequence length used
NQAIF35	1446	682 bp	<i>Coolia monotis</i> U92258	99%	1/582	<i>Coolia monotis</i>	Only short sequences of 28S exist for <i>C.monotis</i> , but the match is convincing
NQAIF40	Bad sequences!						
NQAIF42		750 bp	<i>Scrippsiella hangoei</i> EF205016	72%	7%	Not conclusive	
NQAIF43	1491 bp	1491 bp	<i>Prorocentrum donghaiense</i> AY822610	94%	34/1491	<i>Prorocentrum sp.</i>	Good matches with <i>P. donghaiense</i> , <i>P. mexicanum</i> and <i>P. rathynum</i> .
NQAIF44	1350 bp		No dinoflagellate matches			Not conclusive	Does not seem to match known dinoflagellates
NQAIF56	910 bp	910 bp	<i>Prorocentrum lima</i> NM07 (EF556748)	100%	<1%	<i>Prorocentrum lima</i>	
NQAIF60	1413 bp	651 bp	<i>Coolia monotis</i> U92258	98%	1/651	<i>Coolia monotis</i>	Only short sequences of 28S exist for <i>C.monotis</i> , but the match is convincing
NQAIF70	Bad sequences!						
NQAIF71	950 bp	950 bp	<i>Amphidinium herdmanii</i> (AY455675)	87%	<1%	Related to <i>Amphidinium</i> , but not conclusive	A 200 bp region of the D2 region does not give good matches with anything on the database
NQAIF90	1188 bp	600 bp	<i>Colia monotis</i> isolate CCMP1744 (AM902741)	99%	<1%	<i>Coolia monotis</i>	Only short sequences of 28S exist for <i>C.monotis</i> , but the match is convincing
NQAIF103	1405 bp	650 bp	<i>Coolia monotis</i> AM902743			Probably related to <i>Coolia</i>	<i>Coolia monotis</i> is the only species for which a small section of the D2 region matches, but it is not an extremely good match.

NQAIF Culture	Length of 28S sequenced	Length used for alignment	Best match (with GenBank accession number)	% identities	Gaps	Molecular Identification	Notes
NQAIF116	976 bp	450 bp	<i>Gambierdiscus carpenteri</i>	98% over a 450 bp sequence (includes D2)		<i>Gambierdiscus</i> sp, probably <i>G. carpenteri</i>	Aligned only 450 bp, as the primer D1RFor did not work well. However this fragment includes the D2 region, which aligns very well with <i>Gambierdiscus carpenteri</i> (98%).
NQAIF128	1380 bp	737 bp	<i>Symbiodinium</i> sp. Clade A AF427456	98%	<1%	<i>Symbiodinium</i> sp.	
NQAIF203							
NQAIF252	1417 bp	650 bp	<i>Coolia canariensis</i>	99%	0	<i>Coolia canariensis</i>	
NQAIF260	1480 bp	1480 bp	<i>Heterocapsa</i> sp. AF260399 and <i>H.triquetra</i> EF613355	95%		<i>Heterocapsa</i> sp.	
NQAIF263	1403 bp	1403 bp	<i>Nannochloropsis salina</i> Y07975 and <i>N. gaditana</i> FJ030880	98%	0/1403	<i>Nannochloropsis</i> sp.	Not Dinoflagellate.
NQAIF276	1297 bp	190 bp of D1 and 179 bp of D2	<i>Ostreopsis</i> cf. <i>ovata</i> AF244940	100%(D2) and 97%(D1)		<i>Ostreopsis</i> sp.	D1RFor gave only 190 bp of good sequence. So 2 segments, one with D1 and the second with part of D2 were used, and gave very good matches (100% and 97%) with <i>Ostreopsis</i> .

References

- Bjourson, A. J. and Cooper, J. E. (1992) Band-stab PCR: A simple technique for the purification of individual PCR products. *Nucleic Acids Research* 20: 4675.
- Daugbjerg, N., Hansen, G., Larsen, J. and Moestrup, Ø. (2000) Phylogeny of some of the major genera of dinoflagellates based on ultrastructure and partial LSU rDNA sequence data, including the erection of three new genera of unarmoured dinoflagellates. *Phycologia* 39: 302-317.
- Ki, J-S., Jang, G. Y. and Han, M-S. (2005) Integrated Method for Single-Cell DNA Extraction, PCR Amplification, and Sequencing of Ribosomal DNA from Harmful Dinoflagellates *Cochlodinium polykrikoides* and *Alexandrium catenella*. *Marine Biotechnology* 6: 587-593.
- Nunn, G., Theisen, B., Christensen, B. and Arctander, P. (1996) Simplicity-correlated size growth of the nuclear 28S ribosomal RNA D3 expansion segment in the crustacean order isopoda. *Journal of Molecular Evolution* 42: 211-223.
- Richlen, M. L. and Barber, P. H. (2005) A technique for the rapid extraction of microalgal DNA from single live and preserved cells. *Molecular Ecology Notes* 5: 688-691.
- Scholin, C. A., Herzog, M., Sogin, M. and Anderson, D. M. (1994) Identification of group- and strain-specific genetic markers for globally distributed *Alexandrium* (Dinophyceae). II. sequence analysis of a fragment of the LSU rRNA gene. *Journal of Phycology* 30: 999-1011.
- Walsh, S. P., Metzger, D. A. and Higuchi, R. (1991) Chelex® 100 as a Medium for Simple Extraction of DNA for PCR-Based Typing from Forensic Material *BioTechniques* 10: 506-513.