Available online at www.sciencedirect.com
SCIENCE DIRECT®
Deep-Sea Research I 51 (2004) 1319–1336
DEEP-SEA RESEARCH PART I
www.elsevier.com/locate/dsr

Strategies for molecular genetic studies of preserved deep-sea macrofauna

Elizabeth E. Boyle[a,*], John D. Zardus[a,b], Michael R. Chase[a], Ron J. Etter[a], Michael A. Rex[a]

[a] *Department of Biology, University of Massachusetts, 100 Morrissey Blvd. Boston, MA 02125, USA*
[b] *Kewalo Marine Laboratory, Pacific Biomedical Research Center, University of Hawaii, 41 Ahui St. Honolulu, HI 96813, USA*

Received 23 September 2003; accepted 8 April 2004

Abstract

With the development of new methods to sequence DNA from preserved organisms, existing archival collections can be used to document the population genetic structure of deep-sea species. This has made possible the first direct inferences about patterns of evolutionary diversification in the soft-sediment macrofauna. Here we report protocols and success rates for amplifying and sequencing regions of the mitochondrial 16S rDNA, Cytochrome oxidase I (COI), and Cytochrome b (cytb) genes from formalin-fixed protobranch bivalves and gastropods, major components of the deep-sea benthos. DNA was extracted from 1532 individuals of 12 common bathyal and abyssal species that had been fixed in formalin and preserved in alcohol for up to 36 years. DNA was also extracted from 53 individuals that were dried upon collection, some of which were collected more than 100 years ago. The overall success rate for amplification by PCR was 44%, but this varied considerably among species, stations, and cruises. When DNA amplified, sequencing success was generally high, averaging 85% and ranging from 19% to 100%. The reliability of amplification and sequencing depend strongly on how samples are treated during collection and storage. Amplification success was similar among samples collected from the same station and samples collected on the same cruise. We provide recommendations on strategies for primer design, PCR, and sample selection to improve success rates for genetic analysis of preserved deep-sea organisms. The success rates from different collections, sampling stations, and cruises provide important guidance for selecting material for future genetic work on deep-sea collections examined here.

Keywords: Deep-sea; DNA; Formalin-fixed; Mollusks; Evolution; Population genetics

1. Introduction

Understanding how the deep-sea fauna evolved has remained elusive because, until recently, we lacked the ability to quantify genetic variation within and among populations. New techniques to extract, amplify and sequence DNA from formalin-fixed preserved specimens now make it possible to assess the population genetic structure of deep-sea species by using existing archival collections (France and Kocher, 1996; Chase et al., 1998a;

*Corresponding author. Fax: +1-617-287-6650.
E-mail address: elizabeth.boyle@umb.edu (E.E. Boyle).

0967-0637/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.dsr.2004.04.003

Merritt et al., 1998). This has permitted the first direct insights into the evolutionary origin of faunas that occupy the soft-sediment habitats dominating the deep seascape (France, 1993, 1994; France and Kocher, 1996; Chase et al., 1998b; Etter et al., 1999; Quattro et al., 2001; Etter et al., in prep). Undoubtedly, sequencing fresh-collected material will provide a more complete picture of genetic variation, as it has for megafaunal species living at hydrothermal vents (e.g. Vrijenhoek, 1997; Baco et al., 1999; Van Dover et al., 2002; Goffredi et al., 2003). Deep-sea macrofaunal organisms are much more difficult to collect alive or in sufficient quantity for population-level analysis because they are typically minute, difficult to sort out of sediments and sparsely distributed. Even if a particular site yields abundant collections of fresh material, archived collections of preserved material will remain an indispensable resource for investigating large-scale geographic variation in deep-sea benthic species. A vast amount of archived material exists which is now amenable to genetic study. Improved methods of fixation (Fukatsu, 1999) and extraction (Jung et al., 1991; Walsh et al., 1991; Höss and Pääbo, 1993; Truett, 2000; Schander and Halanych, 2003) will increase the utility of future collections. Our goal here is to describe what we have learned over the past decade working with formalin-fixed deep-sea mollusks.

Recovering sequence information from ancient or preserved material is constrained by the degree to which the DNA is degraded (Pääbo, 1989; Hagelberg et al., 1991; Pääbo and Wilson, 1991). Most archival material has not been collected or stored in ways that are ideal for molecular analysis. Animal tissues are typically fixed in formalin or other formaldehyde-based solutions that cross-link proteins and stabilize tissues (Humason, 1979). After fixation, samples are usually transferred to ethanol for storage. Formalin is thought to inhibit PCR success because it cross-links proteins associated with DNA to each other and to the DNA. This limits the amount of DNA that is accessible to PCR amplification (Karlsen et al., 1994; Schander and Halanych, 2003). Protocols for extraction and amplification of DNA from formalin-fixed material generally involve lysing tissue at elevated temperatures (55°C) in a high concentration of proteinase-K for an extended incubation time (72 h), and targeting relatively short ($\approx$200 bp) PCR fragments (Goelz et al., 1985; France and Kocher, 1996; Shedlock et al., 1997; Chase et al., 1998a; Jalouli et al., 1999; Schander and Halanych, 2003).

When working with archival material to document population structure, it is often necessary to use samples collected by a number of cruises, in different oceans, and by different investigators. Amplification failures can result from a variety of problems that may involve handling procedures (the collection, preservation, or storage of samples), cruise-related events (e.g. a hot deck on sample recovery), or evolutionary alterations of the primer sequence. If the failures are related to handling procedures or conditions on a cruise, the problems may impact multiple taxa and samples. Knowing which samples might be the most useful to work with and, more importantly, which are unlikely to be productive, is essential to planning an efficient genetic study.

As part of our ongoing efforts to study the evolution of the deep-sea macrofauna (Chase et al., 1998b; Etter et al., 1999; Quattro et al., 2001; Etter et al., in prep), we used new molecular genetic methods (Chase et al., 1998a, b; Quattro et al., 2001) to extract, PCR amplify and sequence mitochondrial DNA from six species of deep-sea protobranch bivalves and six species of gastropods that were fixed in formalin and preserved in alcohol for decades. We also sequenced a number of dried museum specimens of the gastropod *Benthonella tenella* that were collected over 100 years ago. In this paper, we quantify overall success rates of PCR amplification within and between collections, quantify the success rate of direct sequencing the PCR products, discuss the limitations of using formalin-fixed tissue for PCR-based studies, and provide advice on working with these and other archived samples.

2. Materials and methods

We selected six species of protobranch bivalves from the Atlantic Ocean: *Ledella ultima* (Smith

1885), *Malletia abyssorum* Verrill & Bush 1898, *Malletia johnsoni* Clarke 1961, *Nuculoma similis* Rhind & Allen 1992, *Nuculoma granulosa* (Verrill 1884), and *Deminucula atacellana* (Schenck 1939); and five species of caenogastropods: *Benthonella tenella* (Jeffreys 1869), *Benthomangelia antonia* (Dall 1881), *Onoba pelagica* (Stimpson 1851), *Fridgidoalvania brychia* (Verrill 1884), and *Mohnia caelata* (Verrill & Smith 1880) and one vitigastropod *Xyloskenea naticiformis* (Jeffreys 1883) based largely on availability of material. The major source of material was the Woods Hole Oceanographic Institution's (WHOI) Benthic Sampling Program, which collected deep-sea samples with epibenthic sleds from 1964 to 1973 (Hessler and Sanders, 1967; Sanders, 1977). A smaller number of specimens were obtained from box-core samples taken from 1984 to 1986 by the Atlantic Continental Slope and Rise (ACSAR) Program (Maciolek et al., 1987a, b) and from 2000 to 2002 by the Northern Gulf of Mexico Continental Slope Habitats and Benthic Ecology Program (DGoMB) (Rowe and Kennicutt, 2002), both for the US Department of Interior's Minerals Management Service. Specimens of *Benthonella tenella* and *Deminucula atacellana* were also obtained from the cruises Incal, BiogasIV, Discovery (Rice et al., 1991) and from the Scottish Association for Marine Science (SAM)-Challenger cruises. In these studies, whole samples were fixed in 10% formalin for 24–48 h and then transferred to 70% ethanol on board ship. Samples were then sorted in the laboratory to species and stored in glass vials with 70% ethanol. SAM samples acquired from the National Museums of Scotland were stored in Industrial Methylated Spirits (IMS). We acquired dried specimens of *Benthonella tenella* collected on the Porcupine 1869 and 1870 cruises (Warén, 1980 pp.57–58) and the US Fisheries Commission's Albatross cruises (1884–1887) (Tanner, 1886, 1887a, b) from the US Natural History Museum. Locality data, date of collection, and the number of DNA extractions carried out for each species are provided in Table 1. We extracted DNA from whole individuals in a buffered proteinase-K solution with an extended lysis time as described in Chase et al. (1998a).

Limited sequence data are available for protobranch bivalves, so initial sequencing was accomplished by using universal primers targeting a portion of the 16S rDNA mitochondrial gene in shallow water protobranchs (Kocher et al., 1989; Palumbi et al., 1991). Species-specific primers were then developed for 16S in both protobranchs and gastropods as described by Chase et al. (1998a). The availability of gastropod COI sequence data is much greater. Thus, for *Benthonella tenella* we developed degenerate primers for COI by aligning 59 caenogastropod sequences obtained from Genbank. We were able to PCR amplify DNA from *Benthonella tenella* with these degenerate primers, and then develop species-specific primers from the resulting sequences.

Initial attempts to amplify DNA from tissue stored in Industrial Methylated Spirits (IMS) were not very successful (see *Benthonella tenella* stations ES255 and ES250). To improve results, we washed subsequent IMS stored samples (ES112, ES252, and ES257) in 90% ETOH, then soaked the samples for 3–4 days in 70% ETOH before extracting the DNA.

To increase PCR yield and reduce contamination, we employed the following thermal cycler profile: 5 cycles [95°C for 1 min; an annealing temperature 3–4°C lower than the Tm of the primers for 1 min, 72°C for 1 min] followed by 25–30 cycles [95°C for 30 s, an annealing temperature at or near the Tm of the primers for 30 s, 72°C for 30 s].

The sizes of PCR products were confirmed on 1.5% agarose gels, and the remaining PCR products were purified with a QIAquick PCR purification kit (Qiagen, Chatsworth, CA, USA), and run out on a 1.5% agarose gel against standards to quantify template. The purified PCR products were sequenced with a *Taq* Dye Deoxy Termination cycle sequencing kit (PE Applied Biosystems, Foster City, CA, USA) or Big Dye kit (PE Applied Biosystems, Foster City, CA, USA), ethanol precipitated, resuspended in formamide loading buffer and run on an Applied Biosystems Model 373 Automated DNA sequencer (ABI) or with conventional isotopic methods. DNA sequences were aligned and edited with

Table 1
Collection localities and success rates of PCR amplification and sequencing by station for 12 species of formalin fixed or dried deep-sea bivalves and gastropods.

Station	Region	Date	Latititude	Longitude	Depth (m)	P/S	cruise	No. Extracted	No. Amplified	% Amplified	No. Sequenced	% Seq.
Ledella ultima 16S 184 base pairs												
200	Ang.	05/22/68	09 43.5′S	10 57.0′E	2699	F/E	AII42	5	4	80	3	75
287	Guy.	02/24/72	13 16.0′N	54 52.2′W	4957	F/E	Knorr25	20	18	90	13	72
288	Guy.	02/25/72	11 02.2′N	55 05.5′W	4423	F/E	Knorr25	30	6	20	3	50
70	N.A.	08/23/64	36 23.0′N	67 58.0′W	4680	F/E	AII12	10	4	40	4	100
77	N.A.	06/30/65	38 00.7′N	69 16.0′W	3806	F/E	Chain50	15	15	100	13	87
78	N.A.	06/30/65	38 00.8′N	69 18.7′W	3828	F/E	Chain50	10	9	90	8	89
84	N.A.	07/04/65	36 24.4′N	67 56.0′W	4749	F/E	Chain50	20	20	100	16	80
92	N.A.	12/13/65	36 20.0′N	67 56.6′W	4694	F/E	AII17	18	15	83	15	100
121	N.A.	08/21/66	35 50.0′N	65 11.0′W	4800	F/E	AII24	17	10	59	8	80
122	N.A.	08/21/66	35 50.0′N	64 57.5′W	4833	F/E	AII24	10	4	40	1	25
123	N.A.	08/22/66	37 29.0′N	64 14.0′W	4853	F/E	AII24	20	5	20	3	60
124	N.A.	08/22/66	37 26.0′N	63 59.5′W	4862	F/E	AII24	10	7	70	3	43
334	N.A.	08/24/72	40 42.6′N	46 13.8′W	4400	F/E	Chain106	17	9	53	9	89
146	S.L.	02/06/67	10 39.5′N	17 44.5′W	2867	F/E	AII31	6	4	67	2	50
147	S.L.	02/06/67	10 38.0′N	17 52.0′W	2934	F/E	AII31	5	3	60	2	67
330	W.E.	08/24/72	50 43.5′N	17 51.7′W	4632	F/E	Chain106	15	12	80	11	92
Total								228	145	64	114	78
Malletiajohnsoni 16S 310 base pairs												
128	N.A.	12/16/66	39 46.5′N	70 36.5′W	1254	F/E	AII30	4	0	0	n/a	n/a
MMSM13.6	N.A.	11/15/85	37 53.33′N	73 45.09′W	1612	F/E	M6	1	1	100	1	100
MMSM6.2	N.A.	08/02/84	39 05.54′N	72 02.97′W	2068	F/E	M2	3	1	33	1	100
MMSN2.1	N.A.	11/09/84	40 57.21′N	67 29.99′W	2100	F/E	N1	2	2	100	1	50
MMSN3.1	N.A.	11/09/84	41 01.40′N	66 20.20′W	1350	F/E	N1	2	2	100	1	50
MMSN3.6	N.A.	07/28/86	41 01.40′N	66 20.20′W	1350	F/E	N6	2	2	100	1	50
MMSN8	N.A.	11/10/84	40 10.32′N	67 37.37′W	2180	F/E	N1	2	1	50	1	100
Total								16	9	69	6	75
Malletia abyssorum 16S 177 base pairs												
243	Arg.	03/14/71	37 36.8′S	52 23.6′W	3819	F/E	AII60	4	0	0	n/a	n/a
72	N.A.	08/24/64	38 16.0′N	71 47.0′W	2864	F/E	AII 12	5	4	80	2	50
77	N.A.	06/30/65	38 00.7′N	69 16.0′W	3806	F/E	Chain50	5	3	60	2	67
78	N.A.	06/30/65	38 00.8′N	69 18.7′W	3828	F/E	Chain50	5	0	0	n/a	n/a
80	N.A.	07/02/65	34 49.8′N	66 34.0′W	4970	F/E	Chain50	9	7	80	3	43
84	N.A.	07/04/65	36 24.4′N	67 56.0′W	4749	F/E	Chain50	20	20	100	14	70
85	N.A.	07/05/65	37 59.2′N	69 26.2′W	3834	F/E	Chain50	20	20	100	19	95
123	N.A.	08/22/66	37 29.0′N	64 14.0′W	4853	F/E	AII24	25	12	46	3	25
331	N.A.	08/29/72	41 13.0′N	41 36.7′W	4793	F/E	Chain106	4	0	0	n/a	n/a
340	N.A.	11/24/73	38 14.1′N	70 20.3′W	3310	F/E	Knorr35	4	0	0	n/a	n/a

149	S.L.	02/21/67	8 03.0′N	18 18.0′W	3861	F/E	AII31	2	1	50	1	100
330	W.E.	08/24/72	50 43.4′N	17 51.7′W	4632	F/E	Chain106	4	0	0	n/a	n/a
Total								107	67	65	44	65
Nuculoma similis 16S 270 base pairs												
MMSN12	N.A.	5/04/85–5/06/86	39 54.3′N	70 55.1′W	550	F/E	N2&N5	10	9	90	9	100
MMSN2	N.A.	11/08/84–7/25/86	40 57.2′N	66 13.9′W	2100	F/E	N1&N6	10	10	100	10	100
MMSN8	N.A.	11/11/84–4/26/86	40 10.3′N	67 37.4′W	2180	F/E	N1&N5	10	10	100	10	100
73	N.A.	08/25/64	39 46.5′N	70 13.3′W	1400	F/E	AII12	15	8	53	5	63
105	N.A.	05/15/66	39 56.6′N	71 03.6′W	530	F/E	ChainSS	24	6	25	6	100
128	N.A.	12/16/66	39 46.5′N	70 45.2′W	1383	F/E	AII30	10	8	80	8	100
207	N.A.	02/21/69	39 51.3′N	7154.3′W	809	F/E	Chain88	15	5	33	5	100
Total								94	56	59	53	95
Nuculoma granulosa 16S 325 base pairs												
MMSN13	N.A.	05/05/85	39 48.4′N	70 55.0′W	1250	F/E	N2	10	10	100	10	100
MMSN3	N.A.	4/26/85–4/27/86	41 01.4′N	66 20.2′W	1350	F/E	N2&N5	10	9	90	9	100
MMSN9	N.A.	11/27/85	39 50.48′N	70 01.73′W	1220	F/E	N4	10	8	80	8	100
73	N.A.	08/25/64	39 46.5′N	70 13.3′W	1400	F/E	AII 12	18	4	22	4	100
87	N.A.	07/06/65	39 48.7′N	70 40.8′W	1102	F/E	Chain50	28	16	57	15	94
209	N.A.	02/22/69	39 47.6′N	70 49.9′W	1500	F/E	Chain88	20	16	80	14	88
Total								96	63	72	60	97
Deminucula atacellana 16S 196 base pairs												
245	Arg.	03/14/71	36 55.7′S	53 0.4′W	2707	F/E	AII60	18	9	50	8	89
256	Arg.	03/24/71	37 40.0′S	52 19.0′W	3910	F/E	AII60	34	20	59	18	90
259	Arg.	03/26/71	37 13.3′S	52 45.0′W	3310	F/E	AII60	34	21	62	4	19
6697	Can.	03/15/68	27 57.0′N	13 46.2′W	1564	F/E	DIS68	2	0	0	n/a	n/a
6701	Can.	03/16/68	27 45.2′N	14 13.0′W	1934	F/E	DIS68	2	0	0	n/a	n/a
303	Guy.	03/01/72	08 28.8′N	56 04.5′W	2848	F/E	Knorr25	4	4	100	0	0
MMSM13	N.A.	11/30/84	37 53.33′N	73 45.09′W	1612	F/E	M3	4	4	100	0	0
MMSM3	N.A.	05/16/85	38 36.84′N	72 51.35′W	2055	F/E	M4	6	5	83	4	80
MMSN13	N.A.	05/05/85	39 48.4′N	70 55.0′W	1250	F/E	N2	2	1	50	1	50
MMSN8	N.A.	11/25/85	40 10.32′N	67 37.37′W	2180	F/E	N4	4	2	50	1	50
Gl	N.A.	05/24/61	39 42′N	70 29′W	2900	F/E	Spring68	5	0	0	n/a	n/a
62	N.A.	08/20/64	39 26.0′N	70 33.0′W	2496	F/E	AII 12	25	12	48	8	67
73	N.A.	08/25/64	39 46.5′N	70 13.3′W	1400	F/E	AII 12	20	13	65	6	46
77	N.A.	06/30/65	38 00.7′N	69 16.0′W	3806	F/E	Chain50	18	18	100	17	94
85	N.A.	07/05/65	37 59.2′N	69 26.2′W	3834	F/E	Chain50	14	7	50	6	86
87	N.A.	07/06/65	39 48.7′N	70 40.8′W	1102	F/E	Chain50	19	19	100	11	58
103	N.A.	05/04/66	39 43.6′N	70 37.4′W	2022	F/E	Chain58	20	18	90	18	100
115	N.A.	08/16/66	39 39.2′N	70 24.5′W	2040	F/E	AII24	20	9	45	6	66
209	N.A.	02/22/69	39 47.6′N	70 49.9′W	1500	F/E	Chain88	20	18	80	14	70
210	N.A.	02/23/69	39 43.0′N	70 46.0′W	2044	F/E	Chain88	11	5	45	5	100

Table 1 (*Continued*)

Station	Region	Date	Latititude	Longitude	Depth (m)	P/S	cruise	No. Extracted	No. Amplified	% Amplified	No. Sequenced	% Seq.
340	N.A.	11/24/73	38 14.1′N	70 20.3′W	3310	F/E	Knorr35	10	5	50	4	80
313	W.E.	08/17/72	51 32.2′N	12 35.9′W	1496	F/E	Chainl06	4	1	25	1	100
DS09	W.E.	07/20/76	55 07.7′N	12 52.6′W	2897	F/E	Incal	5	0	0	n/a	n/a
DS86	W.E.	10/18/74–11/02/74	44 04.8′N	4 18.7′W	1950	F/E	BiogasVI	5	0	0	n/a	n/a
DS87	W.E.	10/18/74–11/02/74	44 05.2′N	4 19.4′W	1913	F/E	BiogasVI	8	1	13	1	100
ES4	W.E.	06/05/73	56 52′N	10 01′W	1993	F/E	ChE/73	5	2	40	1	50
ES14	W.E.	09/22/73	56 45′N	09 46′W	1770	F/E	Ch1 1/73	5	1	20	1	100
ES18	W.E.	09/22/73	56 44′N	09 20′W	1392	F/E	Ch1 1/73	5	1	20	1	100
ES34	W.E.	05/10/75	56 36′N	11 30′W	2515	F/E	Ch7B/75	5	0	0	n/a	n/a
ES176	W.E.	05/28/80	57 15′N	10 26′W	2245	F/E	Ch9A/80	5	0	0	0	0
ES197	W.E.	08/19/81	57 21′N	10 29′W	2200	F/E	Chl2B/81	2	1	50	0	0
ES283	W.E.	04/15/85	54 39′N	12 15′W	2946	F/E	Ch3/85	9	7	78	4	57
ES289	W.E.	04/21/85	57 19′N	10 25′W	2190	F/E	Ch3/85	3	3	100	2	67
Total								353	207	48	149	65
D. atacellana cytb 200 base pairs												
245	Arg.	03/14/71	36 55.7′S	53 0.4′W	2707	F/E	AII60	18	0	0	n/a	n/a
256	Arg.	03/24/71	37 40.0′S	52 19.0′W	3910	F/E	AII60	34	1	5	1	100
259	Arg.	03/26/71	37 13.3′S	52 45.0′W	3310	F/E	AII60	34	0	0	n/a	n/a
303	Guy.	03/01/72	08 28.8′N	56 04.5′W	2848	F/E	Knorr25	4	0	0	n/a	n/a
MMSN13	N.A.	05/05/85	39 48.4′N	70 55.0′W	1250	F/E	N2	2	0	0	n/a	n/a
77	N.A.	06/30/65	38 00.7′N	69 16.0′W	3806	F/E	Chain50	18	5	40	3	60
85	N.A.	07/05/65	37 59.2′N	69 26.2′W	3834	F/E	Chain50	14	0	0	n/a	n/a
103	N.A.	05/04/66	39 43.6′N	70 37.4′W	2022	F/E	Chain58	20	2	13	2	100
DS87	W.E.	10/18/74–11/02/74	44 05.2′N	4 19.4′W	1913	F/E	BiogasVI	8	0	0	n/a	n/a
313	W.E.	08/17/72	51 32.2′N	12 35.9′W	1496	F/E	Chain106	4	0	0	n/a	n/a
ES4	W.E.	06/05/73	56 52′N	10 01′W	1993	F/E	Ch11/73	5	1	20	1	100
ES34	W.E.	06/05/73	56 36′N	11 30′W	2515	F/E	Ch7B/75	5	0	0	n/a	n/a
ES14	W.E.	09/22/73	56 45′N	09 46′W	1770	F/E	Ch11/73	5	0	0	n/a	n/a
ES18	W.E.	9/22/73	56 44′N	09 20′W	1392	F/E	Ch11/73	5	0	0	n/a	n/a
ES283	W.E.	4/15/85	54 39′N	12 15′W	2946	F/E	Ch3/85	9	6	67	6	100
Total								185	15	10	13	92
Benthonella tenella COI 206 base pairs												
242	Arg.	03/13/71	38 16.9′S	51 56.1′W	4392	F/E	AII60	3	1	33	1	100
243	Arg.	03/14/71	37 36.8′S	52 23.6′W	3819	F/E	AII60	1	0	0	n/a	n/a
247	Arg.	03/17/71	43 33.0′S	48 58.1′W	5216	F/E	AII60	1	0	0	n/a	n/a
252	Arg.	03/22/71	38 29.8′S	52 09.1′W	4435	F/E	AII60	1	0	0	n/a	n/a
256	Arg.	03/24/71	37 40.9′S	52 19.3′W	3912	F/E	AII60	5	2	40	2	100
259	Arg.	03/26/71	37 13.3′S	52 45.0′W	3311	F/E	AII60	1	0	0	n/a	n/a
262	Arg.	03/27/71	36 05.2′S	52 17.9′W	2460	F/E	AII60	1	1	100	1	100

167	Br.	02/20/67	7 58.0′S	34 17.0′W	975	F/E	AII31	8	0	0	n/a	n/a
169	Br.	02/21/67	8 03.0′S	34 23.0′W	587	F/E	AII31	2	0	0	n/a	n/a
6697	Can.	03/15/68	27 57.0′N	13 46.2′W	1564	F/E	DIS68	6	0	0	n/a	n/a
191	Cape	05/17/68	23 05.0′S	12 31.5′W	976	F/E	AII42	2	0	0	n/a	n/a
194	Cape	05/17/68	22 54.0′S	11 55.0′W	2864	F/E	AII42	2	0	0	n/a	n/a
BHR1	GoM.	06/18/01	27 28.43′N	91 16.56′W	546	F/E	DGoMBII	1	0	0	n/a	n/a
BHR2	GoM.	06/18/01	27 28.44′N	91 16.55′W	546	F/E	DGoMBII	1	1	100	1	100
C12R3	GoM.	06/02/00	26 13.29′N	89 8.45′W	2924	F/E	DGoMB	1	1	100	1	100
MT5R1	GoM.	06/03/00	27 11.57′N	88 24.2′W	2275	F/E	DGoMB	1	0	0	n/a	n/a
NB4R1	GoM.	05/11/00	26 9.9′N	92 14.13′W	2030	F/E	DGoMB	1	0	0	n/a	n/a
S35R1	GoM.	06/11/00	29 12.2′N	87 02.1′W	658	F/E	DGoMB	1	0	0	n/a	n/a
295	Guy.	02/28/72	8 04.2′N	54 04.3′W	1487	F/E	Knorr25	14	5	36	5	100
299	Guy.	02/29/72	7 55.1′N	55 42.0′W	2009	F/E	Knorr25	14	0	0	n/a	n/a
A2415	N.A.	4/1/1885	30 44′N	79 26′W	805	D	AL1885	10	1	10	1	100
A2654	N.A.	5/2/1886	27 57.30′N	77 27.30′W	1207	D	AL1886	5	2	40	2	100
A2668	N.A.	5/5/1886	30 58.30′N	79 38.30′W	538	D	AL1886	10	4	40	1	25
A2751	N.A.	11/28/1887	16 54′N	63 12′W	1257	D	AL1887	3	0	0	n/a	n/a
70	N.A.	08/23/64	36 23.0′N	67 58.0′W	4680	F/E	AII 12	12	7	60	2	67
77	N.A.	06/30/65	38 00.7′N	69 16.0′W	3806	F/E	Chain50	12	11	92	11	100
80	N.A.	07/02/65	34 49.8′N	66 34.0′W	4970	F/E	Chain50	10	7	70	7	100
84	N.A.	07/04/65	36 24.4′N	67 56.0′W	4749	F/E	Chain50	10	9	90	2	67
85	N.A.	07/05/65	37 59.2′N	69 26.2′W	3834	F/E	Chain50	15	5	33	5	100
93	N.A.	12/14/65	34 39.0′N	66 26.0′W	4967	F/E	AII 17	10	4	100	3	75
118	N.A.	08/18/66	32 19.0′N	64 34.8′W	1144	F/E	AII24	5	1	20	1	100
121	N.A.	08/21/66	35 50.0′N	65 11.0′W	4800	F/E	AII24	11	3	50	3	100
123	N.A.	08/22/66	37 29.0′N	64 14.0′W	4853	F/E	AII24	15	0	0	n/a	n/a
334	N.A.	08/30/72	40 42.6′N	46 13.8′W	4400	F/E	Chain106	15	1	9	0	100
145	S.L.	02/06/67	10 36.0′N	17 49.0′W	2185	F/E	AII31	10	0	0	n/a	n/a
DIS9753	W.E.	04/08/78	50 54.5′N	12 10.9′W	1942	F/E	DIS	10	8	80	8	100
DS01	W.E.	07/15/76	57 59′N	10 40′W	2091	F/E	Incal	10	2	20	2	100
44	W.E.	07/16/67	43 40.8′N	3 35.2′W	1739	F/E	Sarsia76	10	1	10	1	100
313	W.E.	08/17/72	51 32.2′N	12 35.9′W	1496	F/E	Chain106	51	13	25	10	77
318	W.E.	08/19/72	50 26.8′N	13 19.9′W	2506	F/E	Chain106	15	2	13	2	100
ES112	W.E.	10/25/76	55 12′N	15 50′W	1900	F/I	Ch16/76	5	2	40	2	100
ES250	W.E.	07/28/83	59 43′N	12 33′W	1270	F/I	Chl0/83	5	0	0	n/a	n/a
ES252	W.E.	07/30/83	58 52’N	12 53′W	1510	F/I	Chl0/83	5	4	80	4	100
ES255	W.E.	07/31/83	58 26′N	1242′W	1595	F/I	Chl0/83	5	2	40	2	100
ES257	W.E.	07/31/83	57 55′N	12 18′W	1700	F/I	Chl0/83	7	3	60	2	67
J16	W.E.	1870	39 55′N	09 56′W	1818	D	Porc 1870	4	2	50	2	100
J17a	W.E.	1870	39 39′N	09 39′W	1353	D	Porc 1870	3	0	0	n/a	n/a
J39	W.E.	1869	49 01′N	11 56′W	1019	D	Porc 1869	2	2	100	2	100
J40	W.E.	1869	49 01′N	12 05′W	946	D	Porc 1869	4	2	50	2	100
J41	W.E.	1869	49 04′N	12 22′W	1068	D	Porc 1869	4	3	75	2	67
J51	MED	1870	36 55′N	01 10′E	2588	D	Porc 1870	4	1	25	1	100

Table 1 (*Continued*)

Station	Region	Date	Latititude	Longitude	Depth (m)	P/S	cruise	No. Extracted	No. Amplified	% Amplified	No. Sequenced	% Seq.
J54	MED	1870	37 41′N	06 27′E	2758	D	Porc 1870	4	2	50	2	100
Total								368	115	34	94	92
B.tenella 16S 167 base pairs												
242	Arg.	03/13/71	38 16.9′S	51 56.1′W	4392	F/E	AII60	3	0	0	n/a	n/a
256	Arg.	03/24/71	37 40.9′S	52 19.3′W	3912	F/E	AII60	5	2	40	2	100
262	Arg.	03/27/71	36 05.2′S	52 17.9′W	2460	F/E	AII60	1	1	100	1	100
167	Br.	02/20/67	7 58′S	34 17′W	975	F/E	AII31	8	0	0	n/a	n/a
295	Guy.	02/28/72	8 04.2′N	54 04.3′W	1487	F/E	Knorr25	14	2	29	2	100
J54	MED	1870	37 41′N	06 27′E	2758	D	Porc1870	4	0	0	n/a	n/a
A2415	N.A.	4/1/1885	30 44′N	79 26′W	805	D	AL1885	10	1	33	0	0
A2668	N.A.	5/5/1886	30 58.30′N	79 38.30′W	538	D	AL1886	10	1	25	1	100
A2751	N.A.	11/28/1887	16 54′N	63 12′W	1257	D	AL1887	3	0	0	n/a	n/a
70	N.A.	08/23/64	36 23′N	67 58.0′W	4680	F/E	AII12	12	9	90	9	100
77	N.A.	06/30/65	38 00.7′N	69 16.0′W	3806	F/E	Chain50	12	8	89	7	88
84	N.A.	07/04/65	36 24.4′N	67 56′W	4749	F/E	Chain50	10	10	100	10	100
85	N.A.	07/05/65	37 59.2′N	69 26.2′W	3834	F/E	Chain50	15	15	100	11	73
93	N.A.	12/14/65	34 39.0′N	66 26.0′W	4967	F/E	AII17	10	4	100	4	100
123	N.A.	08/22/66	37 29.0′N	64 14.0′W	4853	F/E	AII24	15	0	0	n/a	n/a
334	N.A.	08/30/72	40 42.6′N	46 13.8′W	4400	F/E	Chain106	15	1	33	1	100
145	S.L.	02/06/67	10 36.0′N	17 49.0′W	2185	F/E	AII31	10	0	0	n/a	n/a
44	W.E.	07/16/67	43 40.8′N	3 35.2′W	1739	F/E	DIS68	10	1	33	0	0
313	W.E.	08/17/72	51 32.2′N	12 35.9′W	1496	F/E	Chain106	51	2	50	2	100
318	W.E.	08/19/72	50 26.8′N	13 19.9′W	2506	F/E	Chain106	15	1	25	1	100
Total								233	58	42	51	83
Xyloskenea naticformis 16S 200 base pairs												
200	Ang.	05/22/68	9 41.0′S	10 55.0′W	2699	F/E	AII42	2	0	0	n/a	n/a
242	Arg.	03/13/71	38 16.9′S	51 56. l′W	4392	F/E	AII60	7	0	0	n/a	n/a
159	Br.	02/18/67	7 58.0′S	34.22.0′W	887	F/E	AII31	1	1	100	1	100
167	Br.	02/20/67	7 58.0′S	34 17.0′W	975	F/E	AII31	1	0	0	n/a	n/a
169	Br.	02/21/67	8 03.0′S	34 23.0′W	587	F/E	AII31	1	0	0	n/a	n/a
155	Eq.	02/14/67	00 46.0′S	29 28.0′W	3459	F/E	AII31	2	0	0	n/a	n/a
156	Eq.	02/13/67	00 03.0′S	27 48.0′W	3757	F/E	AII31	1	1	100	1	100
287	Guy.	02/24/72	13 15.8′N	54 52.2′W	4957	F/E	Knorr25	7	0	0	n/a	n/a
70	N.A.	08/23/64	36 23.0′N	67 58.0′W	4680	F/E	AII12	10	5	50	5	100
84	N.A.	07/04/65	36 24.4′N	67 56.0′W	4749	F/E	Chain50	15	13	87	13	100
85	N.A.	07/05/65	37 59.2′N	69 26.2′W	3834	F/E	Chain50	10	4	40	4	100
109	N.A.	08/08/66	36 25′N	68 06.0′W	4750	F/E	AII23	5	5	100	5	100
121	N.A.	08/21/66	35 50′N	65 11′W	4800	F/E	AII24	5	0	0	n/a	n/a
123	N.A.	08/22/66	37 29.0′N	64 14.0′W	4853	F/E	AII24	10	0	0	n/a	n/a

124	N.A.	08/22/66	37 25′N	63 58′W	4862	F/E	AII24	10	1	10	0	0
125	N.A.	08/23/66	37 24′N	65 50′W	4825	F/E	AII24	5	0	0	n/a	n/a
334	N.A.	08/30/72	40 42.6′N	46 13.8′W	4400	F/E	Chain106	2	0	0	n/a	n/a
330	W.E.	08/24/72	50 43.4′N	17 51.7′W	4632	F/E	Chain106	2	0	0	n/a	n/a
Total								96	30	27	29	86
Benthomangelia antonia 16S 325 base pairs												
76	N.A.	06/29/65	39 38.3′N	67 57.8′W	2862	F/E	Chain50	10	2	20	2	100
77	N.A.	06/30/65	38 00.7′N	69 16.0′W	3806	F/E	ChainSO	10	2	20	2	100
340	N.A.	11/24/73	38 14.1′N	70 20.3′W	3310	F/E	Knorr35	10	2	20	0	0
Total								30	6	20	4	67
Onoba pelagica 16S 167 base pairs												
89	N.A.	07/06/65	40 01.6′N	70 40.7′W	196	F/E	ChainSO	15	8	53	8	100
105	N.A.	05/05/66	39 56.6′N	71 03.6′W	530	F/E	Chain58	15	13	87	13	100
172	N.A.	11/27/67	40 12.3′N	70 44.7′W	119	F/E	AII40	15	6	40	6	100
173	N.A	11/28/67	40 10.6′N	70 43.6′W	123	F/E	AII40	15	11	73	11	100
Total								60	38	63	38	100
O.pelagica COI 206 base pairs												
89	N.A.	07/06/65	40 01.6′N	70 40.7′W	196	F/E	Chain50	15	3	100	1	50
105	N.A.	05/05/66	39 56.6′N	71 03.6′W	530	F/E	Chain58	15	4	100	3	100
Total									7	100	4	75
Fridgidoalvania brychia 16S 136 base pairs												
87	N.A.	07/06/65	39 48.7′N	7040.8′W	1102	F/E	Chain50	20	13	65	10	77
96	N.A.	04/27/66	39 55.2′N	70 39.5′W	498	F/E	Chain58	30	15	50	15	100
105	N.A.	05/05/66	39 56.6′N	71 03.6′W	530	F/E	Chain58	20	18	90	16	89
207	N.A.	02/21/69	39 51.3′N	70 54.3′W	808	F/E	Chain88	15	12	80	11	92
346	N.A.	12/03/73	39 54′N	70 10.7′W	457	F/E	Knorr35	25	17	68	11	65
Total								110	75	71	63	84.6
Mohnia caelata 16S 300 base pairs												
73	N.A.	08/25/64	39 46.5′N	70 43.3′W	1400	F/E	AII 12	10	7	70	7	100
87	N.A.	07/06/65	39 48.7′N	70 40.8′W	1102	F/E	Chain50	10	6	60	2	33
103	N.A.	05/04/66	39 43.6′N	70 37.4′W	2022	F/E	Chain58	7	6	86	6	100
Total								27	19	72	15	78

Gene amplified and length of fragment obtained are indicated after the species name. For station numbers, MMS refers to the Atlantic Continenal Slope and Rise Study, ES refers to Scottish Association for Marine Science Stations, DS refers to INCAL or BIOGAS stations, J refers to the Jeffreys collection, A refers to the Albatross cruises, BH, C, MT, NB and S are DGoMB stations, and DIS refers to Discovery cruise material. All other stations are from the Woods Hole Oceanographic Institutions Benthic Sampling Program. Region abbreviations are as follows: Ang. = Angola; Guy. = Guyana; N.A. = North American Basin; S.L. = Sierra Leone Basin; W.E. = West European Basin; Arg = Argentine Basin; Br. = Brazil Basin; MED = Mediterranean Sea; Eq. = Equator and GoM. = Gulf of Mexico. Cruise names are abbreviated-full names appear in Table 3. Primer sequences appear in Table 4. Preservation/Storage conditions (P/S) are Formalin/Ethanol (F/E), Formalin/Industrial Methylated Spirits (F/I), or Dried (D). Latitude and longitude are in degrees and minutes.

Sequencher™, version 3.1 (Gene Codes Corp., Ann Arbor, Michigan) for the Macintosh™.

PCR reactions were scored as positive if a distinct band of expected size was obtained. Reactions that produced a faint band were only scored as positive if a sequence could be obtained through direct sequencing or by cloning the PCR product. Reactions with multiple bands were omitted. A sequencing reaction was scored as positive if unambiguous sequence data emerged. Obvious contamination events were not scored as a positive. Amplification success was determined by the percent of extractions from a station for a given species and gene which yielded positive PCR reactions. Amplification success was averaged for each species among stations, among species within region, within station, and within cruise. Amplification success was also compared between different preservation/storage conditions. Sequencing success was determined by the percent of amplifications per station per gene for a species that yielded a positive sequencing product. Sequencing success was averaged per gene for each species.

3. Results and discussion

3.1. Amplification success (Species, Station, Cruise, Region)

We extracted DNA from 1585 individuals distributed among the twelve species, and amplified a DNA fragment of expected size in 44% of the PCR reactions across all species, collections, and genes (Table 1). Amplification success varied among species 20 to 76%, within species 0 to 100%, among cruises 0 to 100% (Table 2), and among ocean regions 0 to 56% (Table 3). Surprisingly, there was no significant difference in amplification success among samples preserved in formalin stored in ethanol, IMS, or dried (Kruskal–Wallis test $\chi^2 = 0.89$, df = 2, not significant $p = 0.64$).

Amplification success rates are quite variable among species, but are generally good. The gastropods, *Benthomangelia antonia* and *Xyloskenea naticiformis*, are exceptions with amplification success of only 20% and 27%, respectively. The low success in these two species probably reflects a lack of primer specificity. Most of the *X. naticiformis* material, however, came from stations that failed to work for other species (12 out of 18 stations) suggesting the failures may be related to how samples were handled after collection. The small percent of positive results came from stations that worked for other species.

The ability to amplify mtDNA from various species at particular stations can be consistently good, consistently poor or mixed. If specimens of a specific station amplified well for one species, they often amplified well for others. Samples with high amplification success for multiple species were WHOI stations 77, 84, 80, 87, 89, 105, and 209. For example, 5 species averaged 72% amplification success at Station 77. Similarly, some stations seem to yield poor results across all or most species (see *L. ultima, M. abyssorum, B. tenella,* and *X. naticiformis* from station 123) suggesting something might have occurred during sample processing at these stations that degraded the DNA (e.g. a delay in fixation). Mixed results within a station were less common (e.g. *B. tenella* and *M. abyssorum* versus *D. atacellana, X. naticiformis,* from station 85). Notes from the original log for station 85 indicate some of the material was left in cold seawater for 7 h, which may explain the mixed results from this station.

The average amplification success differed between cruises (Table 2) (Kruskal–Wallis test: $\chi^2 = 92.24$, df = 45, $p < 0.001$). For example, in the North American Basin, average success rate was low (25.7%) for samples from Atlantis II cruise 24, but the success rate was higher (68.8%) for samples from Chain 50. In the West European Basin, the average amplification success was low for samples from Biogas IV (4.3%) and Challenger 11/73 (16.67%), but high for Challenger 3/85 (81.67%) and Porcupine 1869 (75%). Within some cruises there were mixed results for different stations. For example, material from Knorr 25 stations 287, 295, and 303 amplified well, but nothing amplified from station 299. The ACSAR material (compare *Nuculoma similis* and *N. granulosa* with other species and other collections) amplified better on average than material from other collections, but even here variation existed among cruises (see ACSAR cruise Mid 2). Some

Table 2
Summary of the mean percent amplification success per cruise. The No. samples is the number of stations sampled/species/gene

Region	Cruise	Mean	No. samples	No. extracted/cruise
North American Basin	Albatross1885	21.5	2	10
	Albatross1886	35	3	15
	Albatross1887	0	2	3
	AtlantisII-17	94.33	3	28
	AtlantisII-12	57.8	10	125
	AtlantisII-23	100	1	5
	AtlantisII-24	25.7	14	163
	AtlantisII-30	40	2	14
	AtlantisII-40	56.5	2	30
	Chain50	68.82	29	320
	Chain58	67.63	8	116
	Chain88	63.6	5	81
	Knorr35	34.5	4	49
	Mid2	33	1	3
	Mid3	100	1	4
	Mid4	83	1	6
	Mid6	100	1	1
	North1	83.33	3	6
	North1&5	100	1	10
	North1&6	100	1	10
	North2	50	3	12
	North2&5	90	2	20
	North4	65	2	14
	North6	100	1	2
	Spring 1961	0	1	5
West European Basin	BIOGAS VI	4.33	3	13
	Challenger10/83	45	4	22
	Challenger11/73	16.67	6	15
	Challenger16/76	40	1	5
	Challenger3/85	81.67	3	12
	Challenger7B/75	0	2	5
	Challenger12B/81	50	1	2
	Challenger9A/80	0	1	5
	Discovery	80	1	10
	INCAL	10	2	15
	Porcupine1869	75	3	10
	Porcupine1870	25	5	15
	Sarsia 1976	21.5	2	10
W. European Basin—N.American Basin	Chain106	22.36	14	129
Gulf of Mexico	DGoMB	25	4	4
	DGoMBII	50	2	2
Canaries	Discovery68	0	3	10
Guyana Basin	Knorr25	34.37	8	89
Argentine Basin	AtlantisII-60	27.17	18	110
Sierra Leone-Brazil Basins	AtlantisII-31	29	13	39
Cape-Angola Basins	AtlantisII-42	20	4	11

The No. extracted is the total number of individuals extracted from a given cruise across all species.

Table 3
Mean percent successful PCR amplification per region (across all sampling programs, stations, genes, and species)

Region	Mean	No. samples	No. extracted/ region
Angola Basin	40	2	7
Argentine Basin	27.16	18	110
Brazil Basin	16.67	6	13
Canary Basin	0	3	10
Cape Basin	0	2	4
Equatorial	50	2	3
Gulf of Mexico	33.33	6	6
Guyana	34.38	8	89
Mediterranean Sea	25	3	8
North American Basin	56.23	108	1090
Sierra Leone	35.4	5	23
West European Basin	31.6	40	222
Total	44	203	1585

The No. samples is the number of stations sampled/species/gene for a given region. The No. extracted is the total number of individuals extracted from a given region across all species.

ocean regions were represented by a single cruise (e.g. Canaries, Brazil) with few individuals, making it difficult to determine if poor results reflect circumstances during a cruise, small sample size or primer specificity problems. Most of the samples from these cruises were stored in the same manner and in the same location, suggesting that pre- or post- fixation handling differences might be responsible for differences in amplification.

We were able to amplify more than one mitochondrial gene for 3 species (*Deminucula atacellana*, *Benthonella tenella* and *Onoba pelagica,* see Table 1)). For *D. atacellana,* amplification success was considerably lower for cytb (10%) than for 16S rDNA (46.7%), but the cytb primers were less specific (Table 4). 16S and COI were similar in amplification success for *B. tenella* (42 versus 33.7%), however the sequence data from 16S possessed little variability. COI primers developed for *B. tenella* were successfully used to amplify COI in a small sample of *O. pelagica,* but

Table 4
List of primers used for each species and gene (R:A/G, Y:C/T, M:A/C, W:A/T, N:A/C/G/T)

Gene	Primer name	Sequence	Species
16s	Led16FB	5′AAG TGA CGA GAA GAC CCT GTC3′	*Ledella ultima*
	LMY16sF	5′GAC GAR AAG ACC CYR TCA AAC3′	*Malletia johnsoni, M. abyssorum, Ledella ultima*
	Ma16F	5′GAC GAG AAG ACC CTG TCA AAC3′	*Malletia abyssorum*
	Proto16R	5′CYC YCA GTT GCC CCA ACT MAA3′	*Malletia johnsoni, M. abyssorum, Ledella ultima*
	NG16F	5′CAT AAA TTA TTA AAA GGG ACG3′	*Nuculoma granulosa*
	NG16R	5′GAA CTC TCA GAA AAG ATT ACG3′	*Nuculoma granulosa, N.similis*
	NS16sF	5′CAT AAA TTG TTA AAA GGG ACG3′	*Nuculoma similis*
	Demi16sF	5′GAG AAG ACC CTA TTG AGT TC3′	*Deminucula atacellana*
	Demi16sR	5′GAT TAC GCT GTT ATC CCT RTG3′	*Deminucula atacellana*
	Bt16f	5′GAA AGA CGA GAA GAC CCT ATC3′	*Benthonella tenella*
	Fb16F	5′AAA TAA ATA TTT AGG TGA AGA AGC3′	*Frigidoalvania brychia*
	16R3	5′GCT GTT ATC CCT RNR GTA ACT3′	*Frigidoalvania brychia, Benthonella tenella, Malletia abyssorum, Ledella ultima, Mohnia caelata*
	BA16F	5′AGA CAA GAA GAC CCT ATC GAG3’	*Benthomangelia antonia*
	BA16R	5’CAT CTT TGG GAT ATT TTG GTC3’	*Benthomangelia antonia*
	Xn16F	5’ATG CTG TTA TCC CTG CGG TAA3’	*Xyloskenia naticisformis*
	Xn16R	5’AGA AGA CCC TRT TGW GCT TTA3’	*Xyloskenia naticisformis*
	NEO16 F	5’GAC RAG AAG ACC CTA TCG AG3’	*Mohnia caelata*
COI	BTCOIP32F	5’CTC ATG CTG GAG GTT CAG TT3’	*Benthonella tenella, Onoba pelagica*
	BTCOIP32R	5’GCA CCA GCT AAA ACA GGA AG3’	*Benthonella, tenella, Onoba pelagica*
Cytb	DemiCytbF3	5’CCT ACA CAT TCT CTT TTT RCA3’	*Deminucula atacellana*
	DemiCytbRB	5’GAA GAG GAA GTA CCA YTC RGG3’	*Deminucula atacellana*

not in other species from the same family (*Fridgidoalvania brychia* and *Pusillina spp.*). These results suggest that some of the primers we have developed for a particular species may work on closely related species.

3.2. Sequencing success

Of the 44% of the PCR reactions that were successful, on average 85% could be reliably sequenced (Table 1). Among species, the average ranged from 65% for *Deminucula atacellana* to 100% for *Xyloskenea natciformis*. Factors that may have contributed to lower sequencing success include low template concentration, lower primer specificity, high secondary structure and admixture of multiple templates. In general, PCR products of concentrations less than or equal to 5 ng/μl were less likely to sequence directly, but often could be amplified by cloning into a T-Vector (pgem®). We only use this approach occasionally, because cloning is a costly and time-consuming process.

Although we used species-specific primers, we encountered sporadic contamination. Contamination represents one of the most insidious problems of working with formalin fixed tissues. Contamination was only encountered when targeting the 16S gene, not COI or Cytb genes. The most common contaminant was human, but we also encountered contamination from chicken, cow and a limpet that had been amplified previously in our research laboratory. The combination of low concentration of template DNA from formalin-fixed tissues, numerous amplification cycles and the dynamics of primer annealing may have contributed to these events. For *Nuculoma granulosa* and *N. similis,* we designed primers upstream and downstream from the 16S sites targeted in the other species, which eliminated contamination events.

Because fresh tissue was not available for these species, indirect evidence must be used to determine if a sequence is reliable. First, sequences were entered into a BLAST search to determine if they were a reasonable phylogenetic match. Sequences were also verified by replicating PCR reactions with DNA extracted on different days and with new reagents on different days to determine if consistent sequences resulted. To verify protobranch sequences, overlapping fragments were amplified for the whole 16Sar/16Sbr fragment and these clustered well with shallow-water protobranch sequences obtained in our laboratory. Further, haplotype diversity was typically low within stations and the sequence data show transition bias. For *Benthonella tenella*, sequences from 4 independent sampling programs in the West European Basin yielded the same or closely related COI sequences, which would be unlikely if there were formalin-induced artifacts or contamination. These observations suggest that formalin or PCR-induced artifacts are unlikely, which is consistent with what others have found (France and Kocher, 1996). Reproducibility with our material and positive phylogenetic matches, suggested by Yang et al. (1997) as a test, affirm the reliability of the data.

3.3. Primer design

The taxon-specific primers used to amplify DNA are presented in Table 4. The ideal approach in primer design is to amplify DNA from fresh or frozen tissue with "universal" primers and develop internal taxon- or species-specific primers, prior to working with formalin-fixed material (Chase et al., 1998a). However, in many studies, as here, archival specimens are the only readily available source of DNA. This does not eliminate them from use in primer design, but it does increase the effort involved. An additional hurdle that we encountered was the occurrence of highly divergent haplotypes among and within protobranch genera (Etter et al., 1999). This high level of divergence made it difficult to design protobranch-specific primers and required the development of species-specific primers (Chase et al., 1998a). To obtain specific primers, we used universal primers on closely related species from shallow water that we could obtain fresh. A series of preliminary primers were then developed using these sequences, which amplified overlapping fragments of the target molecule. Definitive primers were then designed, generally with low degeneracy, and targeted to regions that were taxon-specific and in conserved domains to ensure stable primer binding

(Chase et al., 1998a). Because sequence quality is reduced when degenerate primers are used in conjunction with ABI sequencing, we cloned preliminary products into a T-Vector for sequencing when designing species-level primers.

3.4. Limitations and sample integrity

One of the greatest limitations with archival material in PCR-based applications is the restriction on product size. We found that sequences of approximately 200 bp in length could be reliably amplified, but that amplifying larger fragments was much less consistent. A 450 bp fragment of 16S was amplified from a small number of individuals from the WHOI collection and from approximately 80% of the ACSAR material. Larger fragments can also be obtained from poor quality sources by amplifying independent overlapping fragments, but this substantially increases the amount of effort involved.

Age of the collections, pre-fixation handling and storage can all be factors that influence the success of amplification by PCR (Hagelberg et al., 1991; France and Kocher, 1996). Collection methods differed between studies, and may have influenced sample quality. The WHOI Program, SAM, INCAL and BIOGAS Programs sampled with an epibenthic sled which collects a large amount of material over a distance of about 1 km by skimming off the top few centimeters of sediment (Hessler and Sanders, 1967). The ACSAR and DGoMB samples were collected by using 0.25 and 0.20 m^2 precision box corers, respectively. Retrieval times and processing times on board ship are considerably longer with epibenthic sled sampling. Many deep-sea macrofaunal organisms, including bivalves and gastropods, are minute and vulnerable to temperature extremes and exposure to air. Cell lysis and degradation of DNA begins upon death and prolonged processing times increase the chances of sample deterioration. Storage conditions may also contribute to the loss of sample integrity. Several variables associated with formalin fixation have been shown to affect PCR amplification of DNA: including pH, concentration and length of time in formalin (Wiegand et al., 1996; Bucklin and Allen, 2004). Samples left in formalin for too long due to rough conditions at sea or limited ethanol availability may have had lower amplification success for some stations and cruises.

Amplification of nuclear DNA (nDNA) was much less successful than amplification of mtDNA. In a subset of samples that amplified with the 16S primers for *Deminucula atacellana*, a 270 bp fragment of the 28S rDNA amplified in only 10% of the PCR reactions. The lower success for nuclear DNA may result from the lower copy number, which reduces the probability that a complete fragment of the targeted sequence will occur in the highly degraded DNA of formalin-fixed tissues. It may also result from differences in the way formalin affects nuclear DNA versus mitochondrial DNA, since nuclear DNA is associated with more proteins than is mitochondrial DNA.

3.5. Recommendations

The success rates from different collections, stations, and cruises indicate which samples from these programs might be most productive for genetic studies on other taxa. We found the best material to be from the ACSAR program. This not only amplified well, but also yielded a larger fragment of DNA from some species. The higher rate of amplification success from this collection may be due to the shorter processing time for boxcore samples. Some material from epibenthic sleds also worked quite well. Based on these results, we suggest that material from the WHOI cruises Atlantis II- 17, Chain 50, Chain 58 or Chain 88 would be more useful for genetic analysis in the North American Basin. For the West European Basin samples from Challenger 3/85 and Challenger 10/83 are likely to yield positive results. We have less material from other basins; but more successful amplification was found for station 295 in the Guyana Basin, stations 146 and 147 in the Sierra Leone Basin, station 200 in the Angola Basin, and stations 245 and 262 in the Argentine Basin.

Because of the difficulties of working with formalin preserved tissues, future collections of deep-sea taxa for genetic analysis should consider

alternative preservation methods. Freezing material in liquid nitrogen or dry ice is best for molecular genetic analysis (Dessauer et al., 1996). Another alternative that may be easier in the field is DMSO-NaCl solution, which works well for a variety of marine invertebrate tissues (Dawson et al., 1998). One drawback with these methods is that they do not preserve the tissue in a way that is conducive to species identification, histological studies or museum storage. Ethanol and acetone can be used to preserve tissue and DNA (Fukatsu, 1999); however, both are highly flammable and difficult to transport. DNA preserved in ETOH apparently degrades over time due to shearing (Dean and Ballard, 2000). Tissue stored in ETOH should be kept cold and extracted within a few years (Dessauer et al., 1996; Dean and Ballard, 2000).

The minute size of most deep-sea macrofauna and the difficulty in sorting them out from the similar sized sediments necessitates fixing them in formalin for ecological and systematic research. Several factors should be considered to improve the likelihood of obtaining sequence data from formalin-fixed specimens. Samples preserved in formalin should be fixed in buffered formalin immediately upon collection, and transferred to alcohol shortly after (within 24–48 h). Given the tremendous value of archived deep-sea material for genetic analysis, we suggest controlled experiments to determine what factors account for variability in DNA amplification from small formalin-fixed deep-sea organisms. Relevant variables in such experiments include: time before fixation, time in formalin, temperature and pH of formalin. This information will be valuable for future collections and may be useful for determining whether existing archival collections will be successful for DNA analysis.

Meticulous primer design is necessary for successful amplifications from formalin-fixed material. Taxon-specific primers greatly increase the chance of amplifying the intended target while reducing the chance of spurious sequences resulting from contamination. Universal primers, when successful, can provide a starting point for their development. Sequences can be generated using these less specific primers for multiple relevant taxa and then aligned to develop internal or nested primers. Success with internal primers depends on positioning them at points of true base pair differences between species and amplifying an informative region (Chase et al. 1998a). If possible, use fresh material to design primers before attempting amplifications from formalin-fixed material.

The failure of conspecific samples from some regions to amplify (e.g. South Atlantic) may reflect DNA degradation or molecular evolution of the primer site. Several strategies can be used to discriminate between these alternatives. We developed primers that amplified smaller fragments (<100 bp) and targeted more conserved regions (often at one of the universal sites). Both of these should increase the chances of amplification and allow one to determine if the failure is due to poor quality DNA. We also lowered the stringency of PCR (e.g. lower annealing temperature, increase primer concentration), which increases the chances of amplification even when the primers are not an exact match.

Dried tissue can also be used successfully in PCR and provide sequences that are consistent with formalin-fixed tissues. Results must be interpreted with caution, because a study of dried muscle tissue of an extinct zebra found post-mortem changes in nucleotide bases (deamination) that influenced sequence fidelity (Higuchi et al., 1984, 1987). Our dried samples did not exhibit any signs of sequence alteration, possibly because they were dried more recently.

We made limited attempts to amplify markers such as Simple Sequence Repeats (SSRs) and Randomly Amplified Polymorphic DNA (RAPDs), but based on our results, expect them to be more inconsistent and difficult to work with than mtDNA sequence data. With these markers, one can not easily verify that the bands are from the target taxon, and variability in the ability to amplify DNA from formalin-fixed tissues may yield misleading results. Because contamination is a significant problem when working with formalin-fixed tissues, these markers should only be used with extreme caution.

Amplification of nDNA from formalin-fixed tissues was sufficiently sporadic that it is unlikely

to be useful for population-level studies where large numbers of individuals are needed, even when the targeted gene possesses enough variability for intraspecific resolution. We were unable to consistently amplify the highly repeated nuclear genes 18S/5.8S/28S. However, the small number of amplifications we obtained suggests these genes can probably be used for phylogenetic work on formalin-fixed species. We were unable to amplify single or low copy nuclear genes.

Evolution of the deep-sea macrofauna is poorly understood, in large part, because little is known about genetic variation on sufficiently large geographic and bathymetric scales to reveal patterns of population structure and differentiation. The main impediment to obtaining genetic data from the deep-sea macrofauna is that most collections have been fixed in formalin, which degrades DNA. Over the past decade, we developed techniques to acquire DNA from formalin-fixed mollusks and have begun to explore how evolution unfolds in this enormous and remote ecosystem (Chase et al., 1998b; Etter et al., 1999; Quattro et al., 2001). The results described here should provide useful guidelines to those interested in using molecular genetics to explore evolutionary questions in other deep-sea taxa, and to those working with small formalin-fixed organisms in general.

Acknowledgements

Howard Sanders and John Allen generously provided bivalve samples from the WHOI collection. Fred Grassle made available material from the ACSAR collection. John Gage and Louise Alcott at National Museums of Scotland provided material from the Scottish Association for Marine Science sampling programs. Samples were also provided by Jerry Harasewych from the United States Natural History Museum and by Philippe Bouchet and Philippe Mastrati from the Museum National d'Historie Naturelle, Paris. Gill Rowe generously provided samples from the Northern Gulf of Mexico Continental Slope Habitats and Benthic Ecology Program (DGoMB). Ediane Dutra, Amy Avery, and Alex Plocik provided laboratory assistance. This research was supported by NSF grants OCE-9402855, OCE-9814172, and NSF Academic Research Infrastructure grant 95-12341. The collection of ACSAR material (Table 1) was supported by Minerals Management Service contract no. 14-12-0001-30064 to Battelle New England Marine Research Laboratory and the DGoMB material by the Minerals Management Service contract no. 1435-01-99-CT-30991 to Texas A&M University. A Lerner-Gray Grant to Elizabeth Boyle funded collection of material from museums. This work was partially influenced by a workshop conducted at the National Center for Ecological Analysis and Synthesis, a center funded by NSF (grant DEB-94-21535), the University of California - Santa Barbara, and the State of California. We thank Buzz Wilson and an anonymous reviewer for providing helpful comments on the manuscript.

References

Baco, A.R., Smith, C.R., Peek, A.S., Roderick, G.K., Vrijenhoek, R.C., 1999. The phylogenetic relationships of whale-fall vesicomyid clams based on mitochondrial COI DNA sequences. Marine Ecology Progress Series 182, 137–147.

Bucklin, A., Allen, L.D., 2004. MtDNA sequencing from zooplankton after long-term preservation in buffered formalin. Molecular Phylogenetics and Evolution 30, 879–882.

Chase, M.R., Etter, R.J., Rex, M.A., Quattro, J.M., 1998a. Extraction and amplification of mitochondrial DNA from formalin-fixed deep-sea mollusks. BioTechniques 24, 243–246.

Chase, M.R., Etter, R.J., Rex, M.A., Quattro, J.M., 1998b. Bathymetric patterns of genetic variation in a deep-sea protobranch bivalve, *Deminucula atacellana*. Marine Biology 131, 301–308.

Dawson, M.N., Raskoff, K.A., Jacobs, D.K., 1998. Field preservation of marine invertebrate tissue for DNA analysis. Molecular Marine Biology and Biotechnology 7, 145–152.

Dean, M.D., Ballard, W.O., 2000. Factors affecting mitochondrial DNA quality from museum preserved *Drosophila simulans*. Entomologia Experimentalis et Applicata 98, 279–283.

Dessauer, H.C., Cole, C.J., Hafner, M.S., 1996. Collection and storage of tissues. In: Hillis, D.M., Mortiz, C., Mable, B.K. (Eds.), Molecular Systematics. Sinauer Associates, Sunderland, MA, pp. 29–45.

Etter, R.J., Rex, M.A., Chase, M.R., Quattro, J.M., 1999. A genetic dimension to deep-sea biodiversity. Deep-Sea Research Part I 46, 1095–1099.

France, S.C., 1993. Geographic variation among three isolated populations of the hadal amphipod *Hirondellea gigas* (Crustacea.Amphipoda. Lysianassoidea). Marine Ecology Progress Series 92, 277–287.

France, S.C., 1994. Genetic population structure and gene flow among deep-sea amphipods, *Abyssorchomene* spp., from six California continental borderland basins. Marine Biology 118, 67–77.

France, S.C., Kocher, T.D., 1996. DNA sequencing of formalin-fixed crustaceans from archival research collections. Molecular Marine Biology and Biotechnology 5, 304–313.

Fukatsu, T., 1999. Acetone preservation: a practical technique for molecular analysis. Molecular Ecology 8, 1935–1945.

Goelz, S.E., Hamilton, S.R., Vogelstein, B., 1985. Purification of DNA from formaldehyde fixed and paraffin embedded human tissue. Biochemical and Biophysical Research Communications 130, 118–126.

Goffredi, S.K., Hurtado, L.A., Hallam, S., Vrijenhoek, R.C., 2003. Evolutionary relationships of deep-sea vent and cold seep clams (Mollusca: Vesicomyidae) of the "*pacificalepta*" species complex. Marine Biology 142, 311–320.

Hagelberg, E., Gray, I.C., Jeffreys, A.J., 1991. Identification of the skeletal remains of a murder victim by DNA analysis. Nature 352, 427–429.

Hessler, R.R., Sanders, H.L., 1967. Faunal diversity in the deep-sea. Deep-Sea Research 14, 65–78.

Higuchi, R.G., Bowman, B., Freiberger, M., Ryder, O.A., Wilson, A.C., 1984. DNA sequences from the Quagga, an extinct member of the horse family. Nature 312, 282–284.

Higuchi, R.G., Wrischnik, L.A., Oakes, E., George, M., Tong, B., Wilson, A.C., 1987. Mitochondrial DNA of the extinct quagga: relatedness and extent of post mortem change. Journal of Molecular Evolution 25, 283–287.

Höss, M., Pääbo, S., 1993. DNA extraction from Pleistocene bones by a silica-based purification method. Nucleic Acids Research 21, 3913–3914.

Humason, G.L., 1979. Animal Tissue Techniques. W.H. Freeman & Co., San Francisco, CA.

Jalouli, J., Sand, L., Gustavsson, B., Hirsch, J.M., Larsson, P.-A., 1999. High-throughput DNA extraction from old paraffin-embedded biopsies. BioTechniques 27, 334–338.

Jung, J.M., Comey, C.T., Baer, D.B., Budowle, B., 1991. Extraction strategy for obtaining DNA from bloodstains for PCR amplification and typing of the HLA-DQα gene. International Journal of Legal Medicine 104, 145–148.

Karlsen, F., Kalantari, M., Chitemerere, M., Johansson, B., Hagmar, B., 1994. Modifications of human and viral deoxyribonucleic acid by formaldehyde fixation. Laboratory Investigation 71, 604–611.

Kocher, T.D., Thomas, W.K., Meyer, A., Edwards, S.V., Pääbo, S., Villablanca, F.X., Wilson, A.C., 1989. Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. Proceedings of the National Academy of Sciences USA 86, 6196–6200.

Maciolek, N.J., Grassle, J.F., Hecker, B., Boehm, P.D., Brown, B., Dade, B., Steinhaur, W.G., Baptiste, E., Ruff, R.E., Petrecca, R., 1987a. Study of biological processes on the US Mid-Atlantic Slope and Rise. Phase 2. Final report prepared for US Department of Interior, Minerals Management Service, Washington, DC.

Maciolek, N.J., Grassle, J.F., Hecker, B., Brown, B., Blake, J.A., Boehm, P.D., Petrecca, R., Duffy, S., Baptiste, E., Ruff, R.E., 1987b. Study of biological processes on the US North Atlantic Slope and Rise. Final report prepared for US Department of Interior, Minerals Management Service, Washington, DC.

Merritt, T.J.S., Shi, L., Chase, M.R., Rex, M.A., Etter, R.J., Quattro, J.M., 1998. Universal cytochrome *b* primers facilitate intraspecific studies in molluscan taxa. Molecular Marine Biology and Biotechnology 7, 7–11.

Pääbo, S., 1989. Ancient DNA: extraction, characterization, molecular cloning, and enzymatic amplification. Proceedings of the National Academy of Sciences USA 86, 1939–1943.

Pääbo, S., Wilson, A.C., 1991. Rearrangements of mitochondrial transfer-RNA genes in marsupials. Journal of Molecular Evolution 33, 426–430.

Palumbi, S.R., Martin, A., Romano, S., McMillan, W.O., Stice, L., Grabowsky, G., 1991. The simple fool's guide to PCR, Version 2.0 edition. University of Hawaii, Honolulu, HI.

Quattro, J.M., Chase, M.R., Rex, M.A., Greig, T.W., Etter, R.J., 2001. Extreme mitochondrial DNA divergence of the deep-sea gastropod *Frigidoalvania brychia*. Marine Biology 139, 1107–1113.

Rice, A.L., Billett, D.S.M., Thurston, M.H., Lampitt, R.S., 1991. The Institute of Oceanographic Sciences Biology Programme in the Porcupine Seabight: background and general introduction. Journal of the Marine Biological Association, UK 71, 281–310.

Rowe, G.T., Kennicutt, M.C., 2002. Deepwater program: Northern Gulf of Mexico Continental Slope Habitats and Benthic Ecology, Interim Report-Year 2. US Department of the Interior Minerals Management Service, OCS study MMS 2002-063.

Sanders, H.L., 1977. Evolutionary ecology and the deep-sea benthos. In: Goulden, C.E. (Ed.), The Changing Scenes in Natural Sciences 1776–1976. Academy of Natural Sciences Special Publication, Philadelphia, pp. 223–243.

Schander, C., Halanych, K.M., 2003. DNA, PCR and formalinized animal tissue—a short review and protocols. Organisms Diversity and Evolution 3, 195–205.

Shedlock, A.M., Haygood, M.G., Pietsch, T.W., Bentzen, P., 1997. Enhanced DNA extraction and PCR amplification of mitochondrial genes from formalin-fixed museum specimens. BioTechniques 22, 394–400.

Tanner, Z.L., 1886. Report on the work of the USF.C. steamer ALBATROSS for the year ending December 31, 1884. Rep. USF.C. 1884, part 12, pp. 3–116.

Tanner, Z.L., 1887a. Report on the work of the USF.C. steamer ALBATROSS for the year ending December 31, 1885. Rep. USF.C. 1885, part 13, pp. 3–89.

Tanner, Z.L., 1887b. Record of hydrographic soundings and dredging stations occupied by the steamer ALBATROSS in 1886. Bull. USF.C. 1886, vol. 6, pp 277–285.

Truett, 2000. Preparation of PCR-quality mouse genomic DNA with hot sodium hydroxide and Tris (HotSHOT). BioTechniques 29, 52–54.

Van Dover, C.L., German, C.R., Speer, K.G., Parson, L.M., Vrijenhoek, R.C., 2002. Evolution and biogeography of deep-sea vent and seep invertebrates. Science 295, 1253–1257.

Vrijenhoek, R.C., 1997. Gene flow and genetic diversity in naturally fragmented metapopulations of deep-sea hydrothermal vent animals. Journal of Heredity 88, 285–293.

Wiegand, P., Domhover, J., Brinkmann, B., 1996. Dna degradation in formalin fixed tissues. Pathologe 17, 451–454.

Walsh, P.S., Metzger, D.A., Higuchi, R.G., 1991. Chelex® 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. BioTechniques 10, 506–513.

Warén, A., 1980. Marine Mollusca described by John Gwyn Jeffreys, with the location of the type material. Conchological Society of Great Britain and Ireland, Special publication No.1.

Yang, H., Golenberg, E.M., Shoshani, J., 1997. A blind testing design for authenticating ancient DNA sequences. Molecular Phylogenetics and Evolution 7, 261–265.