Biosci. Biotechnol. Biochem., **76** (4), 633–639, 2012

Construction of a Metagenomic Library for the Marine Sponge *Halichondria okadai*

Takahiro Abe,[1,4,†] Fatma Pinar Sahin,[2] Kiyotaka Akiyama,[3] Takayuki Naito,[3] Mizoe Kishigami,[4] Kenji Miyamoto,[4] Yasufumi Sakakibara,[4] and Daisuke Uemura[1,4,†]

[1]*Department of Chemistry, Faculty of Science, Kanagawa University, 2946 Tsuchiya, Hiratsuka, Kanagawa 259-1293, Japan*
[2]*Department of Pharmaceutical Botany, Faculty of Pharmacy, Hacettepe University, 06100, Sihhiye, Ankara, Turkey*
[3]*Molecular Neuroscience Unit, Okinawa Institute of Science and Technology, 12-22 Suzaki, Uruma, Okinawa 904-2234, Japan*
[4]*Department of Biosciences and Informatics, Faculty of Science and Technology, Keio University, 3-14-1 Hiyoshi, Yokohama, Kanagawa 223-8522, Japan*

Received July 19, 2011; Accepted January 11, 2012; Online Publication, April 7, 2012
[doi:10.1271/bbb.110533]

Symbionts of the marine sponge *Halichondria okadai* are promising as a source of natural products. Metagenomic technology is a powerful tool for accessing the genetic and biochemical potential of bacteria. Hence, we established a method of recovering bacterial-enriched metagenomic DNA by stepwise centrifugation. The metagenomic DNA was analyzed by ultrafast 454-pyrosequencing technology, and the results suggested that more than three types of bacterial DNA, Alphaproteobacteria, Actinobacteria, and Cyanobacteria, had been recovered, and that eukaryotic genes comprised only 0.02% of the metagenomic DNA. These results indicate that stepwise centrifugation and real-time quantitative PCR were effective for separating sponge cells and symbiotic bacteria, and that we constructed a bacteria-enriched metagenomic library from a marine sponge, *H. okadai*, selectively for the first time.

Key words: metagenomic library; *Halichondria okadai*; fosmid; 454-pyrosequencing; sponge

Many structurally unique compounds and significant biologically active compounds have been isolated from various marine invertebrates.[1,2] In particular, sponges, members of the porifera, are rich sources of many natural products. The marine sponge *Halichondria okadai* is generally found in tidal pools on the Pacific coast of Japan. It has an irregular round shape with a few large oscules, and is slightly hard. We have isolated halichondrin B, which exhibits strong cytotoxicity toward B16 melanoma cells, from *H. okadai*, and have analyzed its structure.[3,4] Erubrin (E7389), the right-side fragment of halichondrin B, has been accepted as a therapeutic drug (HALAVEN) for the treatment of breast cancer by the U.S. FDA (http://www.fda.gov/default.htm). Halichlorine, an alkaloid, has also been isolated from *H. okadai*. It inhibits the production of blood vessel cell adhesion molecule (VCAM-1).[5–7] Okadaic acid, a polyether that inhibits phosphatase, has also been isolated from this sponge.[8] Since, many other natural products, including Alteramide A, Neohalicholactone and so on, have been isolated from this sponge,[9,10] *H. okadai* is promising as a source of natural products.

Recent research suggests that marine sponges harbor various microbial symbionts, and that the bacterial population may be as high as 40–60% of the sponge biomass.[11,12] Furthermore, many bioactive compounds in sponges are produced by these symbionts. For example, cytotoxic macrolide swinholide A is produced by both the marine sponge *Theonella swinhoei* and its symbiotic Cyanobacteria.[13] Hence, the exploitation of bacterial symbionts of marine sponges might be an effective approach to harvesting large amounts of natural products. Although microorganisms have potential as sources of bioactive compounds, only a small proportion of bacteria have been isolated from the environment.[14,15] Hence, to use symbiotic bacteria efficiently as sources of natural products, a metagenomic approach is appropriate. Recently, several natural products have been isolated using metagenomic libraries derived from soil.[16,17] These metagenomic libraries were used directly as sources of natural products by screening of clones that produce bioactive compounds by heterologous expression of metagenomic DNA. On the other hand, when construct a metagenomic library from a marine sponge, contamination by the eukaryotic genomic DNA of sponge of metagenomic DNA decreases the efficiency of heterologous expression of the symbiotic bacterial genomic DNA in *E. coli*. Hence, strict separation of symbiotic bacteria and sponge cells is very important in the construction of a metagenomic library.

† To whom correspondence should be addressed. Takahiro Abe, Tel: +81-46-359-4111; E-mail: abtaka@kanagawa-u.ac.jp; Daisuke Uemura, Tel: +81-46-359-4111; E-mail: uemurad@kanagawa-u.ac.jp

Abbreviations: PCR, polymerase chain reaction; DNA, deoxyribonucleic acid; RNA, ribonucleic acid; rRNA, ribosomal RNA; COG, clusters of orthologous groups of proteins; Q-PCR, real-time quantitative PCR; VCAM, vessel cell adhesion molecule; PKS, polyketide synthase; NRPS, non-ribosomal peptide synthase; Tris, tris-hydroxymethyl-aminomethane; EDTA, ethylenediaminetetraacetic acid; CTAB, cetyl trimethyl ammonium bromide; NCBI, National Center for Biotechnology Information; nr, non-redundant database

Downloaded from https://academic.oup.com/bbb/article/76/4/633/5954676 by guest on 25 April 2024

In this study, we established a method of extracting bacterial-enriched metagenomic DNA by stepwise centrifugation and real-time quantitative PCR (Q-PCR), and then confirmed the quality of the metagenomic DNA by dataset analysis. Finally, we constructed a fosmid library with metagenomic DNA from a marine sponge, *H. okadai*, for the first time.

Materials and Methods

Sponge collection and separation of bacterial symbionts. The marine sponge *Halichondria okadai* was collected from tidal pools on the coast of Hayama, Kanagawa, Japan. The sponges were stored on ice until used, within 6 h. They were diced into small pieces and crushed manually into a cell suspension on ice with TEN buffer (3.5% sodium chloride, 10 mM tris-hydroxymethyl-aminomethane, 50 mM ethylenediaminetetraacetic acid, pH 8.5). First, the sediment and cell suspension were separated with a large nylon mesh (20 µm). Then the sponge cells and bacteria were separated by step-wise centrifugation (Fig. 1). The cell suspension was transferred to two ultracentrifuge tubes. One cell suspension was centrifuged at 8,000 *g* for 15 min at 4 °C in a HIMAC CR20G2 centrifuge (Hitachi, Tokyo) (fraction 1). The other cell suspension was first centrifuged at 500 *g* for 5 min at 4 °C (fraction 2). The supernatant was then transferred to another tube and centrifuged at 1,000 *g* for 15 min at 4 °C (fraction 3). Next, the supernatant of the 1,000-*g* fraction was transferred to another tube and centrifuged at 3,000 *g* for 15 min at 4 °C (fraction 4). The supernatant of the 3,000-*g* fraction was then transferred to another tube and centrifuged at 8,000 *g* for 15 min at 4 °C (fraction 5). Finally, the various precipitates (fractions 1–5) were resuspended in TEN buffer and centrifuged at 8,000 *g* (20 min, 4 °C) by way of washing, twice each. The genomic DNA of fraction 1 was used for PCR amplification of 16S rRNA, and that of fraction 5 was used in both the construction of a metagenomic library and second-generation ultrafast sequencing.

DNA isolation. Genomic DNA was extracted and purified following the instructions in a commercial genomic DNA extraction kit (for example, Genomic-tips 20/G, Qiagen, Frankfurt), and the protocol used to recover genomic DNA, described by Piel *et al.*, was modified as follows.[18,19] The precipitates were resuspended and lysed in 20 mL of guanidine solution (60% guanidine thiocyanate, 0.5% sodium dodecyl sulfate, 10 mM EDTA) during incubation at 65 °C. An equivalent volume of CTAB buffer (4% cetyl trimethyl ammonium bromide, 20 mM EDTA pH 8.0, 100 mM Tris–HCl pH 8.0, 1.4 M NaCl) and 50 µL of mercaptoethanol were added, and the sample was incubated for 2 h at 65 °C. After lysis, genomic DNA was extracted 2–3 times with phenol-chloroform and chloroform, and precipitated with isopropanol. The precipitate was dissolved in TE buffer and incubated at 37 °C with RNase A (Toyobo, Osaka) overnight, and then precipitated with isopropanol and dissolved in TE again. The amounts of DNA were determined using a UV spectrophotometer at 260 nm (UVmini-1240, Shimadzu, Kyoto).

Bacterial diversity in metagenomic DNA determined by specific PCR amplification of 16S rRNA. PCR amplification of 16S rRNA from metagenomic DNA was carried out with primers 16SrRNAF: 5′-GTGCCAGCAGCCGCGGTAATAC-3′ and 16SrRNAR: 5′-TACAAGGCCCGGGAACGTATTCAC-3′ (Operon, Tokyo), using *Ex Taq* polymerase (Takara-Bio, Kyoto). These primers refer to *Roseovarius nubinhibens* (NR_028728), *Streptomyces* sp. (AB498686), *Staphylococcus capitis* (AB009937), *Synechocystis* sp. (AB364260), and *Escherichia coli* (AB269763), and they were designed so that they would be located outside the V1 and V8 regions of the bacterial 16S rRNA. Thermocycling consisted of 2 min of denaturation at 95 °C, followed by 30 cycles of 30 s at 95 °C, 30 s at 55 °C, and 90 s at 72 °C. The terminal elongation step was extended by 15 min, and the reaction mixtures were cooled to 4 °C upon completion. Amplicon size and integrity were examined by standard agarose gel electrophoresis and ethidium bromide staining. Amplicons were extracted using the GFX PCR DNA and Gel Band Purification Kit (GE Healthcare Japan, Tokyo). The purified PCR products were cloned to pT7-blue T vector (Takara-Bio) with a Mighty mix-DNA ligation kit (Takara-Bio) and transferred into *E. coli*, DH5-alpha competent cells. The 16S rRNA clones were cultured with Luria-Bertani (LB) broth supplemented with ampicillin (100 mg/L), and the clones were extracted using a GenElute Plasmid Miniprep Kit (Sigma-Aldrich, St. Louis, MO) following the manufacturer's instructions. Eighty-three positive clones were randomly selected and sequenced with a BigDye Terminator Cycle Sequencing Kit (Applied Biosystems, Norwalk, CT) with primers pT7blue-SF: 5′-CAGGTCGACTCTAGAGGATC-3′ and pT7blue-SR: 5′-GACGGCCAGTGAATTCGAGC-3′ (Operon), using an ABI 3730xl DNA Analyzer (ABI). The sequence results were analyzed by Genetyx software (Genetyx, Tokyo) and compared with known sequences in the GenBank database using the BLASTN search program (http://www.ncbi.nlm.nih.gov/) to determine approximate phylogenetic affiliations. Chimeric genes were then compared with related 16S rRNA. A phylogenetic tree was constructed by the neighbor-joining method[20] based on distance matrix data by the phylogenetic program ClustalX2 (available at http://www.clustal.org/).[21] Evolutionary distances were calculated using the Kimura model.[22] The topology of the phylogenetic tree was evaluated by bootstrap analysis carried out with 1,000 replications.[23] 16S rRNA alignment was achieved by including the phyla Proteobacteria, Actinobacteria, Cyanobacteria, and Firmicutes. A sequence belonging to archaea was used as out group (Fig. 3).

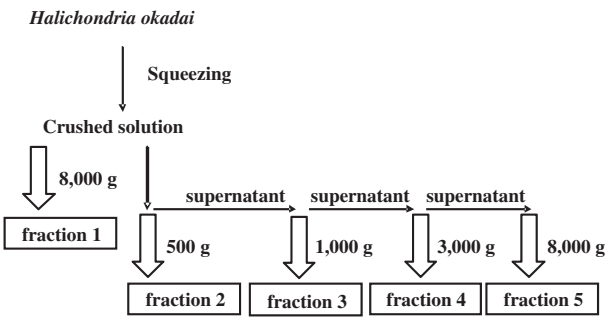

Fig. 1. Scheme for Concentration of the Bacterium from the Sponge Used in This Study.

Cloning of the 18S rRNA of H. okadai. PCR amplification of the 18S rRNA genes of *H. okadai* was performed with primers designed based on the conserved regions of the 18S rRNA of the marine sponges *H. melanodocia* (AY737639) and *Axinella corrugata* (AY737637). Thus 18S rRNA F: 5′-CCTGGTTGATCCTGCCAGTAGTC-3′ corresponded to the 1n to 17n bases of the 18S rRNA of *H. melanodocia*, and to the 1n to 24n bases of the 18S rRNA of *A. corrugate*, and 18S rRNA R: 5′-CTACAGAAACCTTGTTACGAC-3′ corresponded to the 1,759n to 1,779n bases of the 18S rRNA of *H. melanodocia*, and to the 1,769n to 1,789n bases of the 18S rRNA of *A. corrugate* (Operon), using genomic DNA extracted from fraction 1 with *Ex Taq* polymerase. These purified PCR products were cloned, and several positive clones were selected randomly and sequenced with primers pT7blue-SF and pT7blue-SR. The sequence results were analyzed by Genetyx software and compared with known sequences in the GenBank database using the BLASTN search program to determine approximate phylogenetic affiliations.

Quantitative real-time PCR. Cloned 18S rRNA was used as template for a single reaction of Q-PCR. The copy number of the 18S rRNA gene was determined by assuming that based on the molecular weights of the plasmid, which harbored partial 18S rRNA, 1 pg of plasmid equals 1.0×10^5 copies. To prepare a standard curve for Q-PCR, 1, 0.1, 0.01, 0.001, and 0.0001 pg of the purified plasmid was used in identical PCR reactions. Q-PCR was performed using SYBR PremiEx Taq and a Thermal Cycler Dice Real Time System (Takara-Bio).

Pyrosequencing and analysis of metagenomic GS-FLX data. Metagenomic DNA extracted from fraction 5 was used as starting material for pyrosequencing. Approximately 15 µg of genomic DNA was sequenced by three runs on a Roche GS-FLX pyrosequencer

Downloaded from https://academic.oup.com/bbb/article/76/4/633/5954676 by guest on 25 April 2024

Table 1. Closest Phylogenetic Affiliations of Partial 16S rRNAs Retrieved from Selected Clones

Clone	Identity	Closest homolog (BLASTN)	Acessesion no.
2	98.0	Uncultured alphaproteobacteria (HOC32)	AB054166
4	98.5	Mucus bacterium 81	AY654839
47	96.1	Alphaproteobacterium EXT2	AB274734
70	93.5	Uncultured alphaproteobacterium	GQ346733
74	98.6	Uncultured alphaproteobacterium	AB491826
7	99.5	Uncultured bacterium	HM344756
13	99.2	Uncultured bacterium	HM835849
45	99.8	Uncultured bacterium	HM329350

(Roche, Mannheim, Germany). Genomic DNA was dissolved in TE buffer, and its purity was ensured by checking that the A260/A280 wavelength ratio was from 1.8 to 2.0. The resulting data were assembled using Newbler assembly software v1.1.02.15 (Roche). Functional annotation of the protein-coding regions was achieved using the BLASTX algorithm, which was used to query the NCBI non-redundant database (nr) and the clusters of orthologous groups database (COG).[24–27] Comparisons of the metagenomic sequence data to the NCBI-nr and COG databases were performed at a cutoff e-value of 10^{-5}. Manual editing was performed using Genetyx software.

Construction of metagenomic fosmid libraries. A metagenome library was constructed following the manufacturer's instructions using a commercial fosmid library construction kit (CopyControl Fosmid Library Production Kit, Epicentre, Madison, WI). Extracted metagenomic DNA was ligated into the fosmid vector pCC1FOS (Epicentre), and the ligated vectors were packaged into lambda phages and used to transfect *E. coli* EPI300 (Epicentre). The resulting infected cells were spread onto LB medium containing 12.5 µg/mL of chloramphenicol. All the fosmid clones were stored in a deep freezer in LB medium supplemented with a mixture of chloramphenicol (12.5 µg/mL) and 10% glycerol (v/v). Induction of the fosmids to give a high copy number was achieved by the addition of induction solution (Epicentre) and incubation of the cultures at 37 °C for 5 h. After induction, the cells were collected, and the various fosmid DNAs were isolated by the alkaline lysis miniprep method.[28]

Nucleotide sequence accession number. The DNA sequence of the continuous 18S, 28S, and 5.8S rRNAs of *H. okadai* (contig00027) is available from the GenBank database under accession no. AB511881.

Results

Genomic DNA isolation and measurement of 18S rRNA by Q-PCR assay

To measure the concentrations of eukaryotic genomic DNA in the various fractions, the copy numbers of the 18S rRNA were determined using a LightCycler system under the conditions described in "Materials and Methods." First, to design primers for Q-PCR, we extracted genomic DNA from precipitates of *H. okadai* by the CTAB method described in "Materials and Methods." Next, 18S rRNA was cloned. A blast search suggested that the clone was 18S rRNA of *H. okadai*, which shows high homology to that of the marine sponge *Halichondria melanodocia* (AY737639, 99.8%). A 10-fold dilution series of purified 18S rRNA PCR products ranging from 1×10^{-6} to 1 pg/µL was examined by Q-PCR. Genomic DNA samples were then extracted from precipitates collected by graded centrifugation under the conditions described in "Materials and Methods," and the copy numbers of the 18S rRNA segments were determined in duplicate by Q-PCR. The copy number of 18S rRNA per 1 ng of genomic DNA of the fractions collected at high speed decreased dynamically (Fig. 2). The quantity of 18S rRNA clones

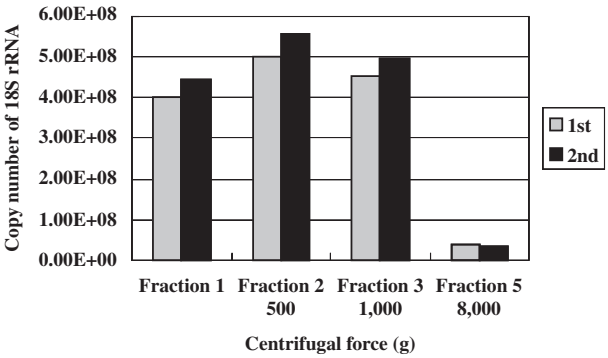

Fig. 2. Quantification and Comparison of 18S rRNA Copies in Centrifuged Pellets of *H. okadai*.

The Q-PCR data were created using four fraction samples (crushed solution, pellets at 500 *g*, pellets at 1,000 *g*, pellets at 8,000 *g*), twice (1st, 2nd).

in the deposition collected at low speed (500 g) (fraction 2) was greater than that in the crushed-sponge solution (fraction 1). This is because the sponge cells were concentrated by centrifugation at low speed. On the other hand, for precipitates collected at high speed (8,000 g), the amount of 18S rRNA clones greatly decreased (fraction 5). This suggests that contaminating sponge cells were largely eliminated by low-speed centrifugation (3,000 g), and that the pellets collected from the supernatant at the higher speed (8,000 g) were enriched bacterial pellets.

Bacterial community composition

Before we evaluated the metagenomic DNA used in library construction, the composition and structure of the bacterial community of *H. okadai* were examined by PCR analysis. We cloned 16S rRNA and sequenced 83 random clones. This revealed that 76 clones (91.6%) showed high homology (97.0–99.0%) to each other, and this group was represented by clone2. First, a blast search in the NCBI database revealed that all of the clones show highest homology to uncultured bacteria (Table 1). Next, to examine the phylogenetic relation between these clones in further detail, cloned 16S rRNA (clones 2, 4, 7, 13, 45, 47, 70, and 74) was aligned with those of representatives of various bacteria using the ClustalX2 program (Fig. 3). This revealed that the dominant phylum was Alphaproteobacteria (clones 2, 4, 47, 70, and 74). Clone2 showed very high homology with the 16S rRNA genes of uncultured Alphaproteobacteria, HOC32 (98.2%), previously isolated from *H. okadai*, at a position apart from the four other clones. Two clones (clone13 and clone45) belonged to Actinobacteria, and clone7 belonged to Firmicutes. This PCR

Downloaded from https://academic.oup.com/bbb/article/76/4/633/5954676 by guest on 25 April 2024

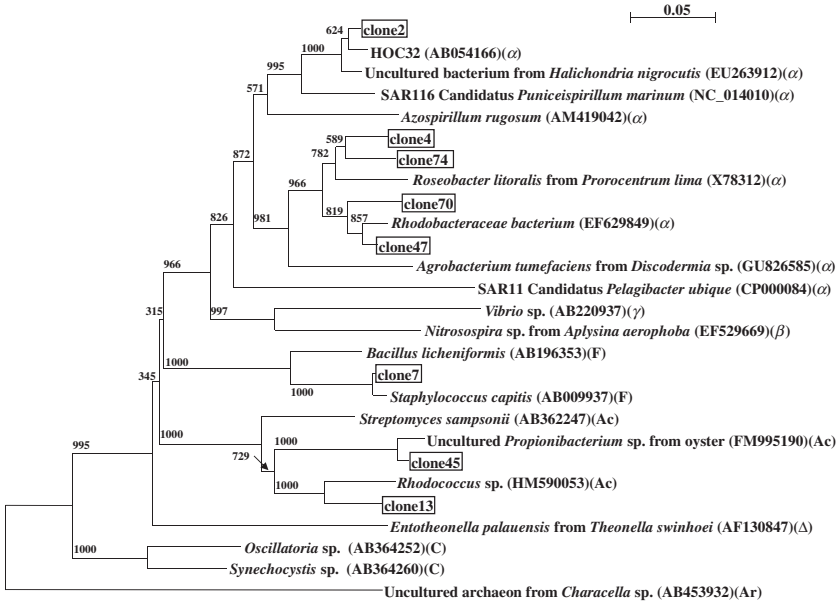

Fig. 3. Neighbor-Joining Tree for Bacterial 16S rRNA.
Bootstrap values calculated from 1,000 resamplings by neighbor-joining are shown at the various nodes. Clones derived from *H. okadai* are boxed. α, alphaproteobacteria; β, betaproteobacteria; γ, gammaproteobacteria; Δ, deltaproteobacteria; Ac, actinobacteria; F, firmicutes; C, cyanobacteria; Ar, archaea.

analysis revealed that Alphaproteobacteria are the overwhelming dominant population in the marine sponge *H. okadai*. Next, to assess the quality of the constructed library, bacteria-enriched metagenomic DNA was estimated by pyrosequencing of the metagenomic DNA, which yielded about 230,000 readings (50 Mb) with an average read length of 220 bases assembled into 17,069 contigs. These contigs were committed to the NCBI-nr database and assigned to either prokaryote (best BLASTX expectation value $\leq 10^{-5}$ to a prokaryotic entry) or eukaryote (best BLASTX expectation value $\leq 10^{-2}$ to a eukaryotic entry). Based on this analysis, at least 50.7% of the contigs in the metagenomic data set showed highest homology to genes that encode prokaryotic proteins. Very few contigs were homologous to eukaryotic proteins (0.02%), and the other contigs were considered not to be assignable. Further analysis revealed that Alphaproteobacteria was also the largest group in the metagenomic library (74.5%), followed by Actinobacteria (3.5%) and Cyanobacteria (3.3%) (Table 2). A further detailed analysis indicated that the Alphaproteobacteria category consisted mainly of uncultured Alphaproteobacteria, Rhodobacterales, Rhizobiales, and Rhodospirillales (Table 3). Rhizobiales are known to be symbionts of plants in general. This indicates that a combination of graded centrifugation and Q-PCR can be used to construct a bacterial-rich genomic library efficiently, and that the metagenomic library is rich in bacterial diversity.

Allocation of assembled contig sequences to microbial genomes

To identify the genes involved in the biosynthesis of natural products, the sequenced data sets were functionally annotated by COG category analysis based on a BLASTX search against the NCBI database, and 8,096 contigs (47.4%) were assigned to functional genes (Fig. 4). The most abundant proteins were associated with metabolism (51.7%), and the biosynthesis, transport, and catabolism of secondary metabolites accounted for 1.8%. To understand better the biosynthetic pathways of natural products derived from symbiotic bacteria, we focused on secondary metabolism, and

Table 2. Prokaryotic Taxa Distribution in the Metagenomic DNA Based on BLAST Search against the Nr-Database

Taxon	%
Alphaproteobacteria	74.5
Betaproteobacteria	3.2
Gammaproteobacteria	6.9
Deltaproteobacteria	1.9
Zetaproteobacteria	0.1
Actinobacteria	3.5
Cyanobacteria	3.3
Planctomycetes	1.2
Firmicutes	1.0
Bacteroidetes	0.9
Chlorobi	0.5
Chloroflexi	0.3
Chlamydiae	0.2
Acidobacteria	0.2
Verrucomicrobia	0.2
Deinococci	0.1
Spirochaetes	0.1
Aquificae	0.1
Tenericutes	0.1
Incertaesedis	1.7

Table 3. Taxonomic Affiliation of Contigs of the Class Alphaproteobacteria

ORDER	%
Rhodobacterales	32.5
Unclassified alphaproteobacteria	27.8
Rhizobiales	19.3
Rhodospirillales	14.9
Sphingomonadales	2.7
Caulobacterales	1.1
Rickettsiales	0.9
Parvularculales	0.7

Downloaded from https://academic.oup.com/bbb/article/76/4/633/5954676 by guest on 25 April 2024

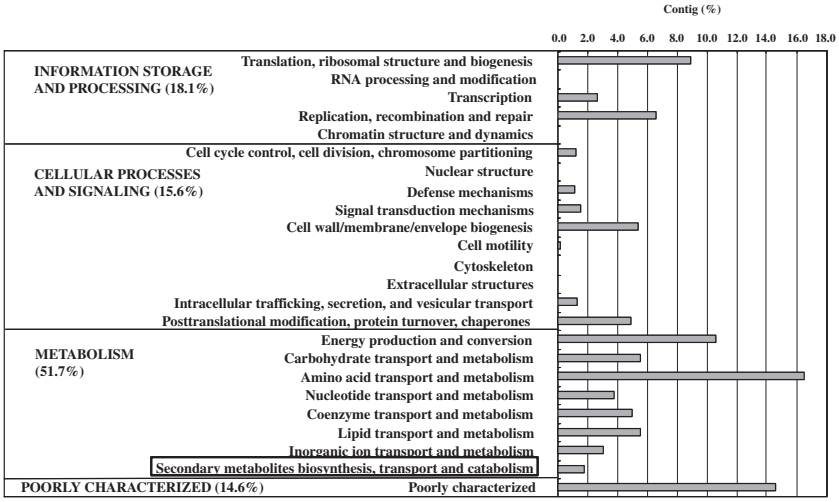

Fig. 4. Categorization of *Halichondria okadai* Metagenome Sequence Contigs According to Clusters of Orthologous Groups of Proteins (COG). The names of subcategories in the COG database are shown at the left, and the corresponding major categories are shown at the right. The numbers of readings assigned to the various major categories and their ratios are shown.

performed an advanced COG analysis using a Swiss-prot database. This revealed that 13 genes showed homology to non-ribosomal peptide synthase modules and related proteins (Table 4). These genes showed similarity to NRPS derived from Proteobacteria (30.8%), Actinobacteria (30.8%), Firmicutes (20.1%), and Cyanobacteria (15.4%).

Genomic library construction

The genomic DNA isolated by the CTAB method was larger than 25 kb, large enough to construct a fosmid library. DNA bands larger than 25 kb were recovered by subjecting the corresponding agarose slices to GFX PCR RNA and using a Gel Band Purification Kit, and the recovered genomic DNA was used to construct a fosmid library. The ligation mixture of digested genomic DNA and fosmid vector was packaged and transferred to EPI300 competent cells. As a result, 150,000 independent clones were obtained. We then confirmed that an appropriate length of genome had been inserted into these fosmids by electrophoresis of restriction enzyme-digested fosmids.

Discussion

Metagenomic analysis can be useful to understand the genetic background of the biosynthesis of natural products. Recently, it was reported that some PKS genes obtained from a metagenomic library of symbiotic bacteria were different from the genes derived from cultivable bacteria.[19] Thus, culture-independent analysis is preferable for the analysis of such symbionts. When one constructs a fosmid library, one must eliminate sponge cells as much as possible, because contamination by genomic DNA from the sponge is an obstacle in screening bacterial natural product-related genes from a metagenomic library. Hence, we established a method of constructing a bacteria-enriched metagenomic library from the marine sponge *H. okadai* by stepwise centrifugation and Q-PCR. Next, we investigated the quality of the metagenomic DNA to determine whether it was suitable for the construction of a metagenomic library and for the screening of natural products. We accessed metagenomic DNA by dataset analysis, and found that the eukaryotic genome accounted for only 0.02% of total metagenomic DNA. Several genome studies have been performed on Porifera, a demosponge, and so the finding that few contigs of the metagenomic database show homology to genes derived from these projects suggests that the sponge's genome was largely removed from the metagenomic DNA.[29] Alphaproteobacteria is the dominant group in the metagenomic library, followed by Actinobacteria and Cyanobacteria (Table 2). Various natural products have been isolated from Alphaproteobacteria, Actinobacteria, and Cyanobacteria,[30–32] and some might be the biogenic source of the non-ribosomal peptides.[32] This suggests that the metagenomic DNA, from which the eukaryotic genome had clearly been eliminated, is a promising genetic resource for the discovery of natural products by functional screening, and that heterologous expression of metagenomic DNA using a multi-host expression system might be useful for exploiting sponges' symbiotic bacteria efficiently.

In this study, we identified 13 contigs that show homology to NRPS genes by dataset analysis (Table 4). For example, contig07303 showed homology with the gene that encodes surfactin synthase, and contig08398 showed homology with the gene that encodes saframycin Mx1 synthase.[33,34] To determine whether these contigs are involved in the biosynthesis of these compounds, heterologous expression of the genes is necessary. Screening of clones whose fosmid contains these contigs by colony hybridization using the contigs as probe might be an efficient approach to compiling sequencing data and to the construction of a whole cluster of NRPS genes. While various natural products have been isolated from *H. okadai*, non-ribosomal peptide has not yet been isolated from it. Hence the expression of whole clusters of NRPS-like genes in a suitable host should lead to the isolation of a non-ribosomal peptide from *H. okadai*.

Marine sponges harbor various natural products derived from secondary metabolites of symbiotic bacteria, and hence the diversity and specificity of symbiotic bacteria is an important issue in the screening

Downloaded from https://academic.oup.com/bbb/article/76/4/633/5954676 by guest on 25 April 2024

Table 4. Predicted Protein Coding Sequences in a Metagenomic Data Set Related to Secondary Metabolite Synthesis

Contig	Length (bp)	e-Value	Closest homolog (BLASTX)	Accession no.	Identity (%)
contig00219	224	6.00E-17	amino acid adenylation domain protein [*Paenibacillus curdlanolyticus* YK9]	ZP_07387014	57
contig04668	256	4.081E-14	amino acid adenylation domain protein [*Streptomyces flavogriseus* ATCC 33331]	ZP_05803229	54
contig06486	261	5.30E-06	amino acid adenylation domain-containing protein [*Salinispora tropica* CNB-440]	YP_00115962	46
contig06528	265	2.889E-12	amino acid adenylation [*Synechocystis* sp. WH 8501]	ZP_00517512	50
contig06841	260	5.00E-21	amino acid adenylation domain protein [*Lyngbya majuscula* 3L]	ZP_08431748	56
contig07303	252	2.00E-04	surfactin synthase subunit 1 [*Bacillus subtilis*]	P27206	47
contig07750	246	9.00E-10	non-ribosomal peptide synthetase [*Pseudomonas brassicacearum*]	YP_00435531	47
contig08398	276	8.00E-27	saframycin Mx1 synthetase B [*Myxococcus xanthus*]	AAC44128	70
contig08571	247	5.00E-08	peptide synthase [*Pseudomonas aeruginosa* PAb1]	ZP_06878666	55
contig08754	263	5.09E-17	non-ribosomal peptide synthetase, putative [*Roseobacter* sp. GAI101]	ZP_05099506	51
contig08780	242	6.00E-20	peptide synthetase [*Streptomyces* sp. Acta 2897]	AEA30273	62
contig08861	243	2.00E-11	hypothetical protein bcere0027_54520 [*Bacillus cereus* AH676]	ZP_04195022	51
contig11049	254	2.03E-13	putative non-ribosomal peptide synthase [*Bradyrhizobium* sp. BTAi1]	YP_00123724	50

of natural products. First, we compared the diversity of the bacterial community of *H. okadai* with those of other marine sponges.[35–38] For sponges of the order Halichondrida, Gammaproteobacteria is the dominant bacteria in *Halichondria* sp., and Acidobacteria and Chloroflexi are the dominant groups in *Svenzea zeai*. For sponges of the order Haplosclerida, Gammaproteobacteria and Firmicutes are the dominant bacteria in *Haliclona simulans* and *Gelliodes carnosa* respectively. On the other hand, Alphaproteobacteria are the dominant bacteria in *H. panicea*, *Rhopaloeides odorabile*, and *Mycale laxissima*, of the orders Halichondrida, Dictyoceratida, and Poecilosclerida respectively.[35,39,40] These results indicate that bacterial composition is highly varied in sponges, and that the dominance of alphaproteobacteria is not a rare characteristic.

We focused on the specificity of symbionts in the marine sponge *H. okadai*. Eighty-three clones of the 16S rRNA of *H. okadai* were analyzed by BLAST search of the NCBI database. This revealed that clone2 showed high homology to other uncultured alphaproteobacteria, HOC32 (98.2%) (AB054166), which had been cloned from *H. okadai*. Furthermore, this clone2 group is phylogenetically distant from the major marine bacteria, including SAR11, for example SAR116 (91.4%) (NC_014010).[41] The fact that 1 certain bacteria were isolated from the sponge extracted separately suggested that these bacteria are sponge-specific. Indeed, it has been reported that the dominant alphaproteobacteria of *H. panicea* was a sponge-specific symbiont by sampling at different stations.[35] These results imply the specificity of alphaproteobacteria clone2. Furthermore, some sponge-phylum-specific symbionts have been detected by comparison of symbionts obtained from various marine sponges,[42] but clone2 derived from *H. okadai* showed low homology to these sponge-phylum-specific alphaproteobacteria (JAWS23, JAWS8) (84.2% and 82.6%). This suggests that this symbiont of the marine sponge *H. okadai* is not a sponge-phylum-specific bacterium.

Recently, it was reported that some sponge symbionts may have co-evolved with their hosts.[43,44] These allied sponges were isolated related species. Because the 5.8S rRNA of *H. okadai* showed high similarity to *H. panicea* (99.0%, AF062607) and sponge-specific clones had been isolated from *H. panicea*, we investigated sponge-order-specific bacteria by homology search. We compared clone2 with the 16S rRNA clones derived from *H. panicea*, and found that the clone2 showed low homology to uncultured alphaproteobacteria in *H. panicea*, for example, HNS27 (Z88567), HNS35 (Z88568), and HNSM50 (Z88569) (86.0%, 69.8%, and 85.7%), the dominant alphaproteobacteria in *H. panicea*. This suggests that clone2 is not a sponge-order-specific bacterium.

Thus there are two methods of transmission of symbiotic bacteria in marine sponges. Usually, marine bacteria are either acquired from the surrounding sea water during filter-feeding by sponges or are transferred from parental sponges to their progeny through reproduction.[45] Sponge-species-specific symbionts are inherited by vertical transmission. Although *H. okadai*, *H. panicea*, and *H. japonica* all belong to the order Halichondrida and live in tidal pools in Japan, different natural products have been isolated from these three sponges. For example, Halichondrin B was not isolated from the latter two sponges. This can be explained by the conjecture that the symbiotic bacterium that produces natural products is sponge-species-specific. Indeed, the 16S rRNA clones derived from *H. panicea* did not show high homology to that of *H. okadai*. Hence a phylogenetic analysis and comparison of symbionts derived from *H. okadai*, *H. panicea*, and *H. japonica* sampled from the same location is necessary to determine the existence of sponge-species-specific symbionts and the relationship between these host sponges, their symbionts, and natural products. To determine this relationship, the metagenomic approach with heterologous expression and DNA analysis is effective.

In this study, we constructed a fosmid library from a marine sponge, *H. okadai*, for the first time. We eliminated the sponge genome from the recovered metagenomic DNA. A dataset analysis indicated that metagenomic DNA is a potential source for screening natural products. An analysis of bacterial diversity suggested that the metagenomic library was constructed by various bacteria, mainly Alphaproteobacteria, Actinobacteria, and Cyanobacteria. An analysis of orthologous proteins suggested that the library contains genes that are involved in non-ribosomal peptide synthesis. This analysis of the genetic profile of the metagenomic DNA should help in using uncultivable and cultivable bacteria as genomic sources in the screening of natural products.

Downloaded from https://academic.oup.com/bbb/article/76/4/633/5954676 by guest on 25 April 2024

Acknowledgments

This study was done mainly at Nagoya University and Keio University. We thank Professor Mamoru Sugita of Nagoya University and Professor Masaya Imoto of Keio University for their support. This study was supported in part by Grants-in-Aid for Creative Scientific Research (Grant no. 16GS0206, to D.U.) and Scientific Research (S) (21221009, to D.U.).

References

1) Uemura D, *Proc. Jpn. Acad. Ser. B, Phys. Biol. Sci.*, **86**, 190–201 (2010).
2) Uemura D, *Chem. Rec.*, **6**, 235–248 (2006).
3) Uemura D, Takahashi K, Yamamoto T, Katayama C, Tanaka J, Okumura Y, and Hirata Y, *J. Am. Chem. Soc.*, **107**, 4796–4798 (1985).
4) Hirata Y and Uemura D, *Pure Appl. Chem.*, **58**, 701–710 (1986).
5) Kuramoto M, Tong C, Yamada K, Chiba T, Hayashi Y, and Uemura D, *Tetrahedron Lett.*, **37**, 3867–3870 (1996).
6) Arimoto H, Hayakawa I, Kuramoto M, and Uemura D, *Tetrahedron Lett.*, **39**, 861–862 (1998).
7) Tsubosaka Y, Murata T, Yamada K, Uemura D, Hori M, and Ozaki H, *J. Pharmacol. Sci.*, **113**, 208–213 (2010).
8) Tachibana K, Scheuer P, Tsukitani Y, Kikuchi H, Engen D, Clardy J, Gopichand Y, and Schmitz F, *J. Am. Chem. Soc.*, **103**, 2469–2471 (1981).
9) Shigemori H, Bae AM, Yazawa K, Sasaki T, and Kobayashi J, *J. Org. Chem.*, **57**, 4317–4320 (1992).
10) Kigoshi H, Niwa H, Yamada K, Stout T, and Clardy J, *Tetrahedron Lett.*, **32**, 2427–2428 (1991).
11) Garson MJ, Zimmermann MP, Hoberg M, Larsen RM, Battershill CN, and Murphy PT, *Lipids*, **28**, 1011–1014 (1993).
12) Kennedy J, Marchesi JR, and Dobson AD, *Appl. Microbiol. Biotechnol.*, **75**, 11–20 (2007).
13) Bewley CA, Holland ND, and Faulkner DJ, *Experientia* (Basel), **52**, 716–722 (1996).
14) Amann RI, Ludwig W, and Schleifer KH, *Microbiol. Rev.*, **59**, 143–169 (1995).
15) Torsvik V, Goksøyr J, and Daae FL, *Appl. Environ. Microbiol.*, **56**, 782–787 (1990).
16) Gillespie DE, Brady SF, Bettermann AD, Cianciotto NP, Liles MR, Rondon MR, Clardy J, Goodman RM, and Handelsman J, *Appl. Environ. Microbiol.*, **68**, 4301–4306 (2002).
17) Brady SF, Chao CJ, Handelsman J, and Clardy J, *Org. Lett.*, **3**, 1981–1984 (2001).
18) Piel J, Hui D, Wen G, Butzke D, Platzer M, Fusetani N, and Matsunaga S, *Proc. Natl. Acad. Sci. USA*, **46**, 16222–16227 (2004).
19) Kim TK and Fuerst JA, *Environ. Microbiol.*, **8**, 1460–1470 (2006).
20) Saitou N and Nei M, *Mol. Biol. Evol.*, **4**, 406–425 (1987).
21) Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, and Higgins DG, *Bioinformatics*, **23**, 2947–2948 (2007).
22) Kimura M, *J. Mol. Evol.*, **16**, 111–120 (1980).
23) Felsenstein J, *Evolution*, **39**, 783–791 (1985).
24) Tatusov RL, Koonin EV, and Lipman DJ, *Science*, **278**, 631–637 (1997).
25) Tatusov RL, Galperin MY, Natale DA, and Koonin EV, *Nucleic Acids Res.*, **28**, 33–36 (2000).
26) Löpez GP, Brochier C, Moreira D, and Rodríguez VF, *Environ. Microbiol.*, **6**, 19–34 (2004).
27) Treusch AH, Kletzin A, Raddatz G, Ochsenreiter T, Quaiser A, Meurer G, Schuster SC, and Schleper C, *Environ. Microbiol.*, **6**, 970–980 (2004).
28) Feliciello I and Chinali G, *Anal. Biochem.*, **212**, 394–401 (1993).
29) Srivastava M, Simakov O, Chapman J, Fahey B, Gauthier ME, Mitros T, Richards GS, Conaco C, Dacre M, Hellsten U, Larroux C, Putnam NH, Stanke M, Adamska M, Darling A, Degnan SM, Oakley TH, Plachetzki DC, Zhai Y, Adamski M, Calcino A, Cummins SF, Goodstein DM, Harris C, Jackson DJ, Leys SP, Shu S, Woodcroft BJ, Vervoort M, Kosik KS, Manning G, Degnan BM, and Rokhsar DS, *Nature*, **466**, 720–726 (2010).
30) Li L, Deng W, Song J, Ding W, Zhao QF, Peng C, Song WW, Tang GL, and Liu W, *J. Bacteriol.*, **190**, 251–263 (2008).
31) Martens T, Gram L, Grossart HP, Kessler D, Müller R, Simon M, Wenzel SC, and Brinkhoff T, *Microb. Ecol.*, **54**, 31–42 (2007).
32) Dunlap WC, Battershill CN, Liptrot CH, Cobb RE, Bourne DG, Jaspars M, Long PF, and Newman DJ, *Methods*, **42**, 358–376 (2007).
33) Arima K, Kakinuma A, and Tamura G, *Biochem. Biophys. Res. Commun.*, **31**, 488–494 (1968).
34) Irschik H, Trowitzsch KW, Gerth K, Höfle G, and Reichenbach H, *J. Antibiot.*, **41**, 993–998 (1988).
35) Wichels A, Würtz S, Döpke H, Schütt C, and Gerdts G, *FEMS Microbiol. Ecol.*, **56**, 102–118 (2006).
36) Lee ZY, He LM, Wu J, and Jiang Q, *J. Exp. Mar. Biol. Ecol.*, **329**, 75–85 (2006).
37) Lee OO, Chui PY, Wong YH, Pawlik JR, and Qian PY, *Appl. Environ. Microbiol.*, **75**, 6147–6156 (2009).
38) Li CQ, Liu WC, Zhu P, Yang JL, and Cheng KD, *Microb. Ecol.*, **62**, 800–812 (2011).
39) Webster NS and Hill RT, *Mar. Biol.*, **138**, 843–851 (2001).
40) Mohamed NM, Cicirelli EM, Kan J, Chen F, Fuqua C, and Hill RT, *Environ. Microbiol.*, **10**, 75–86 (2008).
41) Morris RM, Rappé MS, Connon SA, Vergin KL, Siebold WA, Carlson CA, and Giovannoni SJ, *Nature*, **420**, 806–810 (2002).
42) Hentschel U, Hopke J, Horn M, Friedrich AB, Wagner M, Hacker J, and Moore BS, *Appl. Environ. Microbiol.*, **68**, 4431–4440 (2002).
43) Erpenbeck D, Breeuwer JAJ, van der Velde HC, and van Soest RWM, *Mar. Biol.*, **141**, 377–386 (2002).
44) Thacker RW and Starnes S, *Mar. Biol.*, **142**, 643–648 (2003).
45) Schmitt S, Angermeier H, Schiller R, Lindquist N, and Hentschel U, *Appl. Environ. Microbiol.*, **74**, 7694–7708 (2008).

Downloaded from https://academic.oup.com/bbb/article/76/4/633/5954676 by guest on 25 April 2024