Available online at www.sciencedirect.com
SCIENCE DIRECT®
Biochemical and Biophysical Research Communications 318 (2004) 176–186
BBRC
www.elsevier.com/locate/ybbrc

Comparative gene expression in the symbiotic and aposymbiotic *Aiptasia pulchella* by expressed sequence tag analysis

Jimmy Kuo,[a] Ming-Chyuan Chen,[a,b] Chorng-Horng Lin,[a] and Lee-Shing Fang[a,*]

[a] *Department of Planning and Research, National Museum of Marine Biology and Aquarium, Pingtung 944, Taiwan, ROC*
[b] *Department of Marine Biotechnology, National Kaohsiung Institute of Marine Technology, Kaohsiung 811, Taiwan, ROC*

Received 24 March 2004

Abstract

Intracellular symbiotic relationships are prevalent between cnidarians, such as corals and sea anemones, and the photosynthetic dinoflagellate symbionts. However, there is little understanding about how the genes express when the symbiotic relationship is set up. To characterize genes involved in this association, the endosymbiosis between sea anemone, *Aiptasia pulchella*, and dinoflagellate zooxanthellae, *Symbiodinium* spp., was employed as a model. Two complementary DNA (cDNA) libraries were constructed from RNA isolated from symbiotic and aposymbiotic *A. pulchella*. Using single-pass sequencing of cDNA clones, a total of 870 expressed sequence tags (ESTs) clones were generated from the two libraries: 474 from symbiotic animal and 396 from aposymbiotic animal. The initial ESTs consisted of 143 clusters and 231 singletons. A BLASTX search revealed that 147 unique genes had similarities with protein sequences available from databases; 120 of these clones were categorized according to their putative function. However, many ESTs could not assign functionally. The putative roles of some of the identified genes relative to endosymbiosis were discussed. This is the first report of the use of EST analysis to examine the gene expression in symbiotic and aposymbiotic states of the cnidarians. The systematic analysis of EST from this study provides a useful database for future investigations of the molecular mechanisms involved in algal-cnidarian symbiosis.

Keywords: Symbiosis; Endosymbiosis; EST analysis; Sea anemone, *Aiptasia pulchella*; *Symbiodinium* spp.; Dinoflagellate

Many marine cnidarians, such as sea anemones and corals, harbor photosynthetically active unicellular algae in their cells. This symbiotic relationship is very peculiar because the host compartment, an animal cell, allows a plant cell of dinoflagellate (mostly belonging to the *Symbiodinium* genus), commonly referred to as zooxanthellae, not only to live intracellular, but also to have complicated physiological interactions [1]. Moreover, the relationship is usually mutually beneficial to the partners; the unicellular dinoflagellate algae provide host nutrition by translocation of reduced organic carbon to the host in return for high concentrations of inorganic nitrogen and an environment safe from herbivores [2]. Endosymbiotic associations between cnidarians and *Symbiodinium* form the highly productive and diversified coral reef ecosystem. These associations are abundant and widespread in many tropical and some temperate marine environments and make *Symbiodinium* an important primary producer and a major contributor to global carbon fixation [3]. However, these associations have been found to break down when animal hosts are under environmental stress including water temperature increment [4,5], solar radiation [6,7], or both [8]. Over the past few decades, coral reefs have experienced increasing rate of decline globally, a result of global warming and pollution of marine environment. In order to preserve and improve the health of remaining reefs, a better understanding of the nature of endosymbiosis between cnidarians and dinoflagellates is necessary.

Endosymbioses between two partners are complex associations that are regulated by the genetic interaction of the partners. Molecular level understanding of these associations is, therefore, crucial to understand

* Corresponding author. Fax: +886-8-882-5086.
E-mail address: lsfang@nmmba.gov.tw (L.-S. Fang).

0006-291X/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.bbrc.2004.03.191

the mechanisms of endosymbiosis. Presently, only very few endosymbiotic systems, e.g., parasitic prokaryotes and their respective mammalian and plant hosts, have been examined at the molecular level [9,10]. However, the genetic interactions between cnidarians and dinoflagellate symbionts are largely unexplored. Recently, sea anemone and *Symbiodinium* have been employed as a model to study endosymbioses between cnidarians and algae, and several host genes have been identified to be associated with cnidarian-algal endosymbiosis [11–13]. By comparing the protein profiles generated from symbiotic sea anemones *Anthopleura elegantissima* and those from non-symbiotic (aposymbiotic) sea anemones with 2D gel electrophoresis, Weis and Levine found several proteins or genes with enhanced expression in symbiotic hosts. These genes included the following: carbonic anhydrase, which supplies CO_2 to symbionts, sym17, a novel gene that is homologous to calcium-binding proteins in other organisms, and sym32, a cell adhesive and/or cell-cell recognition protein that is homologous to fasciclin I (fas I) gene family. Moreover, two host proteins, apARF1 and ApRab7, critical to endosymbiotic interactions were identified in the tropical sea anemone, *Aiptasia pulchella* [14,15]. apARF1 protein was homologous with ARF1 protein and played a negative role in the *Aiptasia–Symbiodinum* endosymbiotic interaction. ApRab7 associated with phagosome maturation was a key protein for algae to survive inside hosts. However, our current molecular level understanding of cnidarian–algal interactions is still very little; therefore, a more comprehensive study to characterize the expressed genes of hosts is needed.

Single-pass sequencing of random cDNA clones to generate the expressed sequence tags is a rapid, powerful, and cost-effective method in massive cloning of cDNA as well as in large scale characterization of cDNA sequences for deciphering genome sequence. Although the sequences generated by this method are incomplete and are prone to error, EST collections (about two-thirds of them of human origin) grow much faster than any other genomic sequence information and become a powerful means of gene discovery. Since its first introduction by Adams et al. [16,17] in the human genome project, EST analysis has been successfully applied to study gene expression in plant, animal, and fungal organisms in various functional or developmental states. In the present study, the sea anemone, *A. pulchella*, and its intracellular symbiont, *Symbiodinium* spp., were chosen as a model system, due to their ease of maintenance and the feasibility of culturing them separately. In order to identify cnidarian host genes that are expressed as a function of symbiotic relationship, two cDNA libraries were constructed from tissues isolated from different symbiotic stages of sea anemones, animals symbiotic with zooxanthellae and aposymbiotic animals. Also, the expressed genes were analyzed on a large scale using ESTs of these two cDNA libraries. A total of 374 clusters composed of 870 ESTs clones were generated in the present study, and most of them were reported in sea anemone for the first time.

Materials and methods

Animal collection and maintenance. Specimens of *A. pulchella* were collected from the outdoor culturing facility of National Museum of Marine Biology and Aquarium (NMMBA), in Pingtung, Taiwan, and maintained indoor in normal seawater in the laboratory tanks with ample access to ambient sunlight. Water temperature of the aquaria was in the range of 26–28 °C. Animals were fed daily with newly hatched brine shrimp, and seawater was changed once every other week.

Preparation of aposymbiotic animals. To generate *A. pulchella* free of zooxanthellae (therefore bleached), symbiotic animals were cultured in a tank with the inhibitor of photosynthetic electron transport DCMU (3,4-dichlorophenyl-1,1-dimethylurea) (Sigma) at 5×10^{-6} M with fresh DCMU seawater replacing half of the tank seawater every other week for two months. At the end of the treatments, the animals appeared white. They were then maintained in seawater pre-filtered through 0.45 μm membranes and fed once everyday with brine shrimp. Using these procedures, the animals have been kept in bleached state for more than six months in our laboratory.

Construction of cDNA library. Expressed sequence tags were obtained from two different cDNA libraries. To prepare total RNA of symbiotic and aposymbiotic *A. pulchella*, approximately 1 g of each animal source was first extracted with Trizol LS reagent (Life Technologies) following the procedure recommended by the manufacturer. The resulting total RNA was then subjected to a filter-based purification procedure using RNAqueous (Ambion) following the protocol detailed in the instruction manual. cDNA libraries were constructed by ZAP Express cDNA Synthesis kit (Stratagene) and SMART cDNA library construction kit (Clonetech) according to the manufacturer's instructions.

Plasmid DNA extraction and sequencing. Plasmid DNA was prepared using QIAprep Spin Miniprep kit (Qiagen) and sequenced using the ABI Prism BigDye Terminator Cycle Sequencing kit (PE Applied Biosystems) on an ABI 377 DNA sequencer (PE Applied Biosystems).

Data handling and bioinformatics analysis. All raw sequence data were analyzed for potential overlaps with the SeqManII module from the Lasergene v. 5.06 software (DNASTAR). Prior to assembly, the vector sequences were removed manually and the sequence ends were trimmed using SeqManII with default parameters. The resulting cDNA sequences were assembled into overlapping contigs using default parameters. A minimum match percentage of 80% for 12 overlapping bases was necessary to merge two sequences as one contig. Consensus sequences from each contig were stored in a database and used in similarity searches of the NCBI GenBank public database. The ESTs were compared with the non-redundant NCBI database using the BLASTX algorithms [18,19]. Queries were performed with TurboBLAST v2.0 (TurboWorx), a commercial, LINUX-based, and distributed parallel processing software developed for batch analysis of multiple sequences. The sequences that showed significant similarity ($E < 10^{-4}$) to characterize proteins were further divided on the basis of their potential cellular function following the categories outlined in the Expressed Gene Anatomy Database (EGAD) [20]. The ESTs described in this paper can be found in the NCBI database under Accession Nos. CK662887–CK663259.

Results

Overview of ESTs from the symbiotic and aposymbiotic animals

A total of 870 clones that are the subject of this research were sequenced from cDNA clones of the symbiotic and aposymbiotic libraries of *A. pulchella*. Of these clones, 474 were from the symbiotic host animal and 396 were from the aposymbiotic host animal. The EST clones consisting of vector sequences and sequences shorter than 150 bp were discarded and ambiguous regions were trimmed before analysis. The average readable sequence length from each library was 567 bp for the symbiotic animal and 556 bp for the aposymbiotic animal. A large fraction of ESTs are between 400 and 700 bp in both libraries. The distributions of the readable sequence length of the two libraries are shown in Fig. 1.

Distribution of ESTs to clusters

To characterize transcript abundance in the libraries, all sequences of both libraries were assembled using the SeqMan II program. Since non-normalized primary cDNA libraries were used, the number of ESTs per cluster could reflect the relative mRNA population. The initial ESTs from the two libraries were grouped into 374 consensus sequences, in which 143 containing more than two ESTs per consensus sequence and 231 were singletons. Therefore, the estimated number of genes identified by the EST database was 374 and the redundancy of EST database was 38.2%. The distribution of

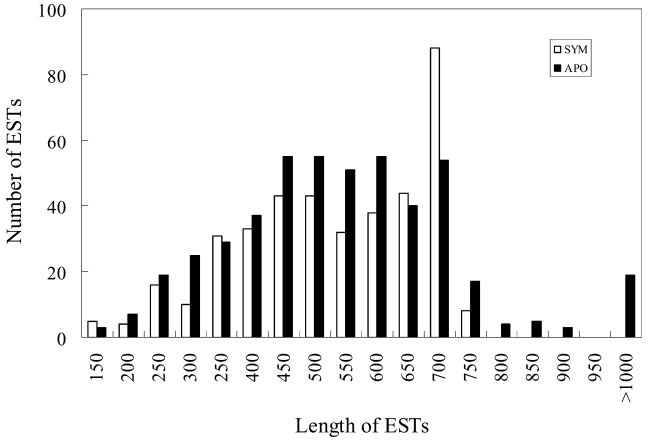

Fig. 1. Distribution of readable sequence length for cDNA clones of symbiotic and aposymbiotic *A. pulchella*. The number of sequence reads is shown for each group of 50 sequence lengths between 150 and 1200 bp. The total number of ESTs clones was 870: 474 from symbiotic library and 396 from aposymbiotic library. The average readable sequence length from each library was 567 bp for the symbiotic library and 556 bp for the aposymbiotic library (SYM, symbiotic sea anemone EST library; APO, aposymbiotic sea anemone EST library).

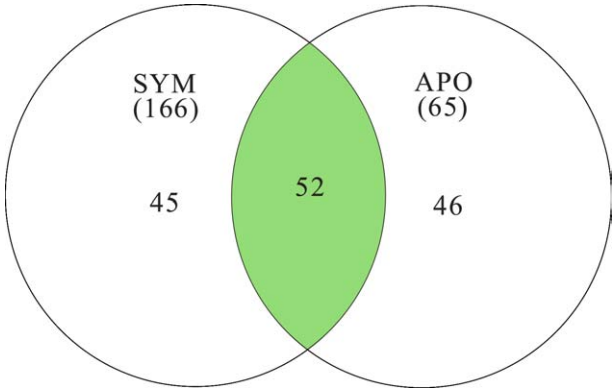

Fig. 2. Distribution of clusters and singletons between symbiotic and aposymbiotic EST libraries. Numbers in shaded area indicate clusters composed of ESTs from two libraries. Numbers in parentheses represent the singletons derived from each cDNA library (SYM, symbiotic sea anemone EST library; APO, aposymbiotic sea anemone EST library).

clusters and singletons between the two EST libraries is shown in Fig. 2. Among the 143 clusters contained more than two ESTs, 52 (36.4%) clusters were commonly expressed in both states and 45 (31.5%) and 46 (32.1%) clusters were only expressed in symbiotic and aposymbiotic state, respectively. Of the 231 singletons, 166 (77.9%) occurred in the symbiotic animal and 65 (22.1%) occurred in the aposymbiotic animal.

Analysis of highly expressed ESTs

The highly expressed genes (as represented by clusters composed of 6 or more ESTs) from each library are listed in Table 1. Among the 21 highly expressed ESTs presented in Table 1, two were ferritin genes, three were actin genes, and 12 were ribosomal components. Ferritin (A001), ribosomal protein L5 (A033), 40S ribosomal protein S6 (A037), and ribosomal protein S10 (A086) occurred at high frequencies in the symbiotic animal than in the aposymbiotic animal. On the other hand, actin (A008), ribosomal protein L19 (A007), ferritin heavy chain polypeptide 1 (A006), qm protein (A099), and ribosomal protein S19 (A101) occurred at high frequencies in the aposymbiotic animal. Cluster A090, a sequence with no match in the database, actin (A102), 40S ribosomal protein S9 (A034), ribosomal protein S14 (A035), cytoplasmic actin (A046), ribosomal protein L17 (A100), Finkel–Biskis–Reilly murine sarcoma virus expressed protein (A182), and ribosomal protein L7 (A124) were found in both libraries and showed high redundancy level among all ESTs.

Analysis of the ESTs by matching with GenBank

The consensus sequences of each cluser were compared with the nr (all nonredundant GenBank CDS

Table 1
Assembled clusters that contain more than six ESTs

Cluster	Gene description	Organism	Blastx *E* value	SYM	APO	Total
A001	Ferritin	*Dermacentor variabilis*	5.00E − 66	57	45	102
A090	Unknown			8	9	17
A008	Actin	*Galaxea fascicularis*	0.00E + 00	3	13	16
A032	Ribosomal protein L26	*Argopecten irradians*	2.00E − 43	6	8	14
A033	Ribosomal protein L5	*Rattus norvegicus*	1.00E − 116	13	0	13
A007	Ribosomal protein L19	*Branchiostoma belcheri tsingtaunese*	2.00E − 63	3	9	12
A102	Actin	*Chlamys farreri*	1.00E − 127	6	6	12
A006	Ferritin heavy chain polypeptide 1	*Branchiostoma lanceolatum*	3.00E − 44	2	7	9
A034	40S ribosomal protein S9	*Ictalurus punctatus*	2.00E − 79	4	4	8
A035	Ribosomal protein S14	*Rattus norvegicus*	9.00E − 56	4	4	8
A046	Cytoplasmic actin	*Pisaster ochraceus*	2.00E − 66	4	4	8
A099	qm protein	*Drosophila melanogaster*	1.00E − 99	2	6	8
A100	Ribosomal protein L17	*Ictalurus punctatus*	1.00E − 71	4	4	8
A101	Ribosomal protein S19	*Branchiostoma belcheri tsingtaunese*	2.00E − 63	2	6	8
A182	Finkel–Biskis–Reilly murine sarcoma virus expressed protein	*Mus musculus*	2.00E − 34	4	4	8
A170	Ribosomal protein S4	*Argopecten irradians*	5.00E − 46	3	4	7
A037	40S ribosomal protein S6	*Aplysia californica*	4.00E − 99	6	0	6
A234	Peptidyl-prolyl *cis–trans* isomerase	*Blattella germanica*	1.00E − 68	2	4	6
A124	Ribosomal protein L7	*Branchiostoma lanceolatum*	1.00E − 57	3	3	6
A036	Ribosomal protein L32	*Argopecten irradians*	3.00E − 47	4	2	6
A086	Ribosomal protein S10	*Bos taurus*	2.00E − 43	6	0	6

SYM, symbiotic sea anemone EST library; APO, aposymbiotic sea anemone EST library.

translations + PDB + Swissport + PIR + PRF) database using the BLASTX program. This program translates the input nucleotide sequence in all possible reading frames, and then searches the resulting protein sequences against protein sequences in the GenBank database. Of the 374 unique sequences, 147, or 44.5%, matched existing sequences with an *E* value of less than 10^{-4} (the *E* value is the expected number of the chance matches of the quality). Sequences with *E* values greater than 10^{-4} were not considered statistically significant. Following the results of previous analysis, ESTs matching the existing protein sequences could be categorized into nine major classes, basically according to a modified classification scheme utilized by the Expressed Gene Anatomy Database, or EGAD, developed at the Institute for Genomic Research (TIGR, Rockville, MD), and available at: http://www.tigr.org/doc/tigr-scripts/egad_scripts/role_report.spl [20].

Table 2 shows the results for functional categories that differ between these two libraries. It should be noted that the functional categories for the sequence were based entirely on inference from BLASTX reports; tests for the gene function have not yet been conducted. Category I, which contained only one EST in the symbiotic library, was associated with cell division. Category II, which contained four ESTs, one in the symbiotic library and three in the aposymbiotic library, was associated with cell signaling/cell communication. Category III, which contained 42 ESTs, 15 in the symbiotic library and 27 in the aposymbiotic library, was associated with cell structure/motility. Category IV, which contained 113 ESTs, 61 in the symbiotic library and 52 in the aposymbiotic library, was associated with cell/organism defense. Category V, which contained nine ESTs in the symbiotic library, was associated with RNA synthesis. Category VI, which contained 243 ESTs, 148 in the symbiotic library and 95 in the aposymbiotic library, was associated with protein synthesis. Category VII, which contained 31 ESTs, 17 in the symbiotic library and 14 in the aposymbiotic library, was associated with metabolism. There were still 45 ESTs, 30 in the symbiotic library and 15 in the aposymbiotic library, classified into Category VIII. These ESTs were matched to ESTs whose functions had not been fully defined due to lack of enough information. The remaining 382 ESTs, 193 in the symbiotic library and 189 in the aposymbiotic library, were categorized into Category IX with no significant similarities to any known sequence. In particular, genes for ribosomal proteins, unclassified, unknown, and carrier proteins/membrane transport were more abundant in the symbiotic sea anemone library. However, gene for cell structure/motility, tRNA synthesis/metabolism, and energy/TCA cycle were more abundant in the aposymbiotic sea anemone library.

To analyze the diversity of gene expression in different symbiotic states of *A. pulchella* and gain a snapshot of gene expression profiles between symbiotic and aposymbiotic anemones, the pattern of gene expression in

Table 2
Distribution of ESTs in different categories and subcategories as defined by a modified classification of EGAD

Category	SYM	APO	Total
I. Cell division	1(0.2%)	0(0%)	1(0.1%)
II. Cell signaling/cell communication			
Cell adhesion	0(0%)	1(0.3%)	1(0.1%)
Channels/transport proteins	0(0%)	2(0.5%)	2(0.2%)
Hormone/growth factors	1(0.2%)	0(0%)	1(0.1%)
III. Cell structure/motility	15(3.2%)	27(6.8%)	42(4.8%)
IV. Cell/organism defense			
Carrier proteins/membrane transport	60(12.6%)	52(13.2%)	112(12.9%)
Stress response	1(0.2%)	0(0%)	1(0.1%)
V. RNA synthesis			
RNA processing	5(1.1%)	0(0%)	5(0.6%)
Transcription factors	3(0.6%)	0(0%)	3(0.3%)
Translation factors	1(0.2%)	0(0%)	1(0.1%)
VI. Protein synthesis			
Post-translational	12(2.5%)	8(2.0%)	20(2.3%)
Protein turnover	3(0.6%)	2(0.5%)	5(0.6%)
Ribosomal proteins	129(27.2%)	77(19.5%)	206(23.7%)
tRNA synthesis/metabolism	1(0.2%)	4(1.0%)	5(0.6%)
Translation factors	3(0.6%)	4(1.0%)	7(0.8%)
VII. Metabolism			
Amino acid	5(1.0%)	0(0%)	5(0.6%)
Energy/TCA cycle	5(1.1%)	8(2.0%)	13(1.5%)
Nucleotide	3(0.6%)	0(0.0%)	3(0.3%)
Protein modification	0(0%)	2(0.5%)	2(0.2%)
Sugar/glycolysis	4(0.8%)	4(1.0%)	8(0.9%)
VIII. Unclassified	30(6.3%)	15(3.8%)	45(5.2%)
IX. Unknown	193(40.6%)	189(47.8%)	382(43.9%)
Total	475(100%)	395(100%)	870(100%)

SYM, symbiotic sea anemone EST library; APO, aposymbiotic sea anemone EST library.

the symbiotic and aposymbiotic anemones is shown in Fig. 3. The pattern of gene expression in the symbiotic and aposymbiotic libraries was similar. Three categories show large difference (>3.5%) in expression patterns between symbiotic and aposymbiotic animals. Unknown ESTs and genes relative to cell structure/motility are higher in the aposymbiotic animal. Genes related to protein synthesis were higher in the symbiotic animal. The mRNA transcripts as presented in each category are shown in Table 3, and most of these were reported in sea anemone for the first time. Of the 147 matched sequences in Table 3, 141 of these matched eukaryotic sequences and six matched bacterial sequences. Of the 141 eukaryotic matches, 133 matched genes from metazoans, one protozoa, two fungi, and five plants. Of the 133 matched genes from metazoans, only four were from cnidarians. None of the sequences matched with sea anemone.

Discussion

Mutualistic endosymbiotic associations between cnidarians and their photosynthetic dinoflagellate symbionts are a biologically and ecologically interesting problem. The environmental factors related to the maintenance and breakdown of these associations have been studied on macroscales by many researchers, but the expression of genes and the interpretation of environmental signals in the host cell during symbiosis are still not well understood. To study the gene expression within the cnidarian as a function of symbiotic state, we randomly sequenced the cDNA clones from cDNA libraries of symbiotic and aposymbiotic tissues of sea anemone *A. pulchella*. Totally, 870 cDNA clones were analyzed by homology with database sequences to identify the gene function; 374 among 870 ESTs were regarded as unigenes and none of the unigenes have been identified in the sea anemone. At this time, only less than 200 sea anemone protein sequences can be found in the GenBank database, and 374 unigenes represent about 1.8-fold of the sea anemone genes in the GenBank database. Therefore, EST sequencing is considered as an efficient method to obtain the coding genes at the whole-genome level. However, it should be noted that the number of genes predicted by clusters and singleton analysis is usually overestimated due to non-overlapping sections of the same genes.

Sequence redundancy analysis between different states or organ ESTs could be used to explain the differential gene expression [21,22]. Therefore, redundancy analysis between symbiotic and aposymbiotic *A. pulchella* ESTs can identify a large number of genes expressed exclusively or predominantly between each state and

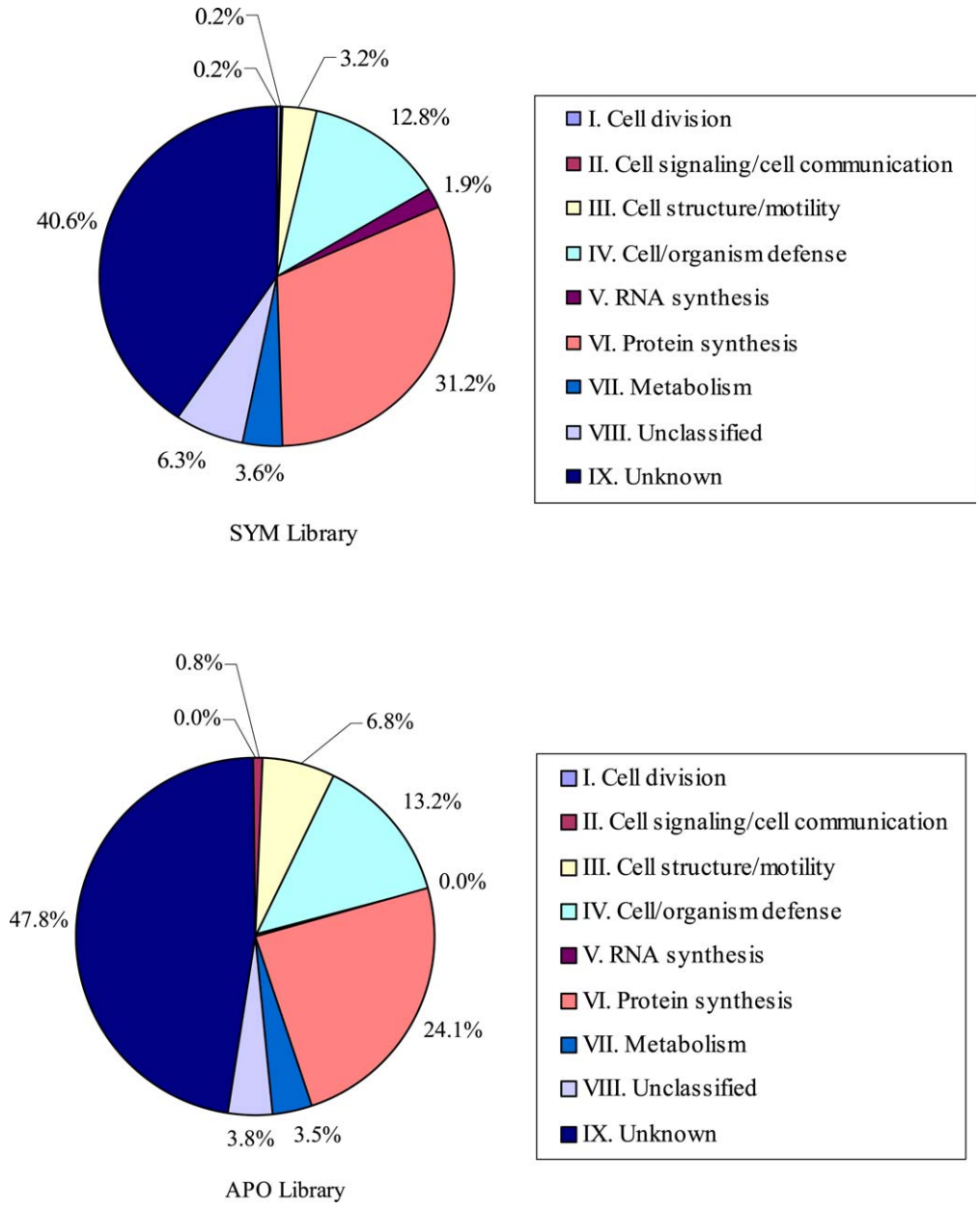

Fig. 3. Distribution of ESTs into modified EGAD categories for symbiotic and aposymbiotic EST libraries. Although initial comparisons were made using clusters and singletons, percentages are calculated from the total number of ESTs, not clusters, in each functional category to reflect the relative levels of expression of genes in each category (SYM, symbiotic sea anemone EST library; APO, aposymbiotic sea anemone EST library).

contribute greatly to the understanding of the cnidarian–algal symbiosis. In comparison of unigenes isolated from the symbiotic and aposymbiotic anemones in Fig. 2, only 52 clusters (13.9% of unigenes) were common and the others were specific to each state in *A. pulchella*. This shows that the gene expression is quite different quantitatively and qualitatively between different states of *A. pulchella*.

To understand gene expression differences between different states, the EST profiles of *A. pulchella* from each library are depicted in Fig. 3. The two libraries have similar pattern of gene expression as determined by functional analysis. It is notable that a large number of sequences display little or no homology with genes of known function. This observation highlights the lack of knowledge of genes expressed in the symbiotic and aposymbiotic animals. Comparing the EST profiles of *A. pulchella* from each library, three categories were found to have difference greater than 3.5%. Unknown ESTs (7.2%, Category IX) and genes relative to cell structure/motility (3.6%, Category III) are higher in the aposymbiotic animal; however, genes related to protein synthesis (7.1%, Category VI) are higher in the symbiotic animal. Based on the previous study on cnidarians

Table 3
List of tentatively identified clones

Cluster	Accession No.	Putative product	Species with best match	Blastx *E* value	Library[a]
I. Cell division					
Apoptosis					
A283	CK663164	mukF protein	*Escherichia coli*	6.00E − 17	S1
II. Cell signaling/cell communication					
Cell adhesion					
A358	CK663239	Thrombospondin	*Drosophila melanogaster*	4.00E − 05	A1
Channels/transport proteins					
A118	CK663002	Translocation protein 1	*Homo sapiens*	1.00E − 09	A2
Hormone/growth factors					
A300	CK663181	Histamine *N*-methyltransferase	*Trichodesmium erythraeum*	1.00E − 05	S1
III. Cell structure/motility					
Cytoskeletal					
A008	CK662892	Actin	*Galaxea fascicularis*	0.00E + 00	S3A13
A102	CK662986	Actin	*Chlamys farreri*	1.00E − 127	S6A6
A255	CK663137	Actinin, α 1	*Homo sapiens*	1.00E − 81	A2
A046	CK662930	Actin, cytoplasmic	*Pisaster ochraceus*	2.00E − 66	S4A4
A303	CK663184	Chain C, arp23 complex	*Bos taurus*	3.00E − 35	S1
A193	CK663076	β-Actin	*Pseudopleuronectes americanus*	3.00E − 29	A2
A174	CK663057	Myosin light chain alkali, smooth-muscle isoform	*Danio rerio*	4.00E − 14	S1
IV. Cell/organism defense					
Carrier proteins/membrane transport					
A001	CK662887	Ferritin	*Dermacentor variabilis*	5.00E − 66	S57A45
A006	CK662890	Ferritin heavy chain polypeptide 1	*Branchiostoma lanceolatum*	3.00E − 44	S2A7
A004	CK662888	Ferritin-like protein	*Pinctada fucata*	2.00E − 05	S1
Stress response					
A070	CK662954	dnaJ protein	*Porphyromonas gingivalis*	1.00E − 17	S1
V. RNA synthesis					
RNA processing					
A246	CK663128	ATP-dependent RNA helicase	*Escherichia coli*	3.00E − 65	S3
A369	CK663250	Splicing factor 3B subunit 3	*Mus musculus*	5.00E − 65	S1
A228	CK663110	Splicing factor	*Homo sapiens*	7.00E − 21	S1
Transcription factors					
A320	CK663201	Adenosylhomocysteinase 2	*Danio rerio*	1.00E − 56	S1
A051	CK662935	CCAAT/enhancer binding protein alpha	*Danio rerio*	2.00E − 17	S2
Translation factors					
A084	CK662968	Translation initiation factor 4A, isoform 1	*Homo sapiens*	8.00E − 53	S1
VI. Protein synthesis					
Post-translational modification/targeting					
A236	CK663118	Chaperonin containing TCP1, subunit 7	*Danio rerio*	1.00E − 85	S1A4
A067	CK662951	Ubiquitin	*Anopheles gambiae*	9.00E − 63	S2
A009	CK662893	Ubiquitin/ribosomal protein S27a fusion protein	*Branchiostoma belcheri*	5.00E − 59	S4
A234	CK663116	Peptidyl-prolyl *cis–trans* isomerase	*Blattella germanica*	1.00E − 68	S2A4
A151	CK663035	Peptidylprolyl isomerase	*Digitalis lanata*	1.00E − 73	S3
Protein turnover					
A208	CK663091	Proteasome subunit β type 3	*Oncorhynchus mykiss*	4.00E − 92	S1
A154	CK663037	Chymotrypsinogen 1	*Paralichthys olivaceus*	3.00E − 49	S1
A126	CK663010	6330590F17Rik protein	*Rattus norvegicus*	2.00E − 27	A2
A231	CK663113	Chymotrypsin	*Xenopus laevis*	3.00E − 09	S1
Ribosomal proteins					
A047	CK662931	60S ribosomal protein L8	*Homo sapiens*	1.00E − 117	S3
A033	CK662917	Ribosomal protein L5	*Rattus norvegicus*	1.00E − 116	S13
A087	CK662971	40S ribosomal protein S2	*Urechis caupo*	1.00E − 115	S4A1
A099	CK662983	qm protein	*Drosophila melanogaster*	1.00E − 99	S2A6
A200	CK663083	40S ribosomal protein S5	*Homo sapiens*	4.00E − 99	S3A2
A037	CK662921	40S ribosomal protein S6	*Aplysia californica*	1.00E − 91	S6
A190	CK663073	Ribosomal protein L5	*Rattus norvegicus*	7.00E − 80	A1
A034	CK662918	40S ribosomal protein S9	*Ictalurus punctatus*	2.00E − 79	S4A4
A158	CK663041	60S ribosomal protein L10a	*Spodoptera frugiperda*	4.00E − 78	S1A2
A179	CK663062	Ribosomal protein L15	*Mus musculus*	3.00E − 77	S1
A014	CK662898	60S ribosomal protein L11	*Drosophila melanogaster*	4.00E − 77	S2

Table 3 (*continued*)

Cluster	Accession No.	Putative product	Species with best match	Blastx *E* value	Library[a]
A159	CK663042	40S ribosomal protein S13	*Homo sapiens*	3.00E − 76	S3
A163	CK663046	40S ribosomal protein S23	*Dermacentor variabilis*	2.00E − 72	S1
A100	CK662984	Ribosomal protein L17	*Ictalurus punctatus*	1.00E − 71	S4A4
A040	CK662924	Ribosomal protein S18	*Branchiostoma belcheri*	2.00E − 67	S4
A098	CK662982	Ribosomal protein S3a	*Felis catus*	9.00E − 67	S2
A050	CK662934	Ribosomal protein L27	*Stichodactyla helianthus*	9.00E − 65	S2
A007	CK662891	Ribosomal protein L19	*Branchiostoma belcheri*	2.00E − 63	S3A9
A349	CK663230	Ribosomal protein L6	*Rattus norvegicus*	1.00E − 59	A1
A124	CK663008	Ribosomal protein L7	*Branchiostoma lanceolatum*	1.00E − 57	S3A3
A106	CK662990	40S ribosomal protein S15A	*Strongylocentrotus purpuratus*	2.00E − 57	S3A2
A111	CK662995	Ribosomal protein S17	*Branchiostoma belcheri*	4.00E − 57	S2A1
A010	CK662894	40S ribosomal protein	*Perinereis aibuhitensis*	7.00E − 57	S2A2
A035	CK662919	Ribosomal protein S14	*Rattus norvegicus*	9.00E − 56	S4A4
A210	CK663093	Ribosomal protein L9, cytosolic	*Homo sapiens*	3.00E − 54	S1
A094	CK662978	60S ribosomal protein L23a	*Rattus norvegicus*	2.00E − 53	S1
A101	CK662985	Ribosomal protein S19	*Branchiostoma belcheri*	5.00E − 52	S2A6
A066	CK662950	Ribosomal protein L44	*Chlamys farreri*	5.00E − 48	S1
A036	CK662920	Ribosomal protein L32	*Argopecten irradians*	3.00E − 47	S4A2
A170	CK663053	Ribosomal protein S4	*Argopecten irradians*	5.00E − 46	S3A4
A019	CK662903	60S acidic ribosomal protein P0	*Rana sylvatica*	1.00E − 45	S1
A038	CK662922	Ribosomal protein S15	*Argopecten irradians*	2.00E − 45	S3A2
A146	CK663030	40S ribosomal protein S8	*Homo sapiens*	3.00E − 45	S1
A032	CK662916	Ribosomal protein L26	*Argopecten irradians*	2.00E − 43	S6A8
A086	CK662970	Ribosomal protein S10	*Bos taurus*	2.00E − 43	S6
A073	CK662957	Ribosomal protein L34	*Branchiostoma belcheri*	1.00E − 41	S1
A082	CK662966	40 S ribosomal protein S24	*Spodoptera frugiperda*	3.00E − 40	A1
A243	CK663125	Ribosomal protein L23	*Branchiostoma belcheri*	3.00E − 40	S1A4
A025	CK662909	Ribosomal protein L35	*Branchiostoma belcheri*	3.00E − 39	A1
A331	CK663212	40S ribosomal protein SA	*Gallus gallus*	3.00E − 39	S1
A274	CK663156	Ribosomal protein L16a	*Emericella nidulans*	5.00E − 39	S1A1
A015	CK662899	60S ribosomal protein L37	*Drosophila melanogaster*	9.00E − 36	S2
A138	CK663022	Ribosomal protein L35A	*Schistosoma japonicum*	3.00E − 35	A1
A216	CK663099	60S ribosomal protein L13	*Danio rerio*	5.00E − 34	S1
A024	CK662908	Ribosomal protein L36	*Branchiostoma belcheri*	3.00E − 33	S1
A088	CK662972	60S ribosomal protein L14	*Lumbricus rubellus*	5.00E − 32	S1A1
A089	CK662973	Ribosomal protein S25 CG6684-PA	*Drosophila melanogaster*	1.00E − 30	S3
A164	CK663047	Ribosomal protein L28	*Ictalurus punctatus*	1.00E − 29	S1
A188	CK663071	Mitochondrial 28 S ribosomal protein S10	*Homo sapiens*	1.00E − 27	S1
A150	CK663034	Ribosomal protein L22	*Branchiostoma belcheri*	4.00E − 23	S1
A011	CK662895	Ribosomal protein P2	*Branchiostoma belcheri*	8.00E − 21	S3A1
A134	CK663018	Ribosomal protein S29	*Branchiostoma belcheri*	2.00E − 20	S2
A045	CK662929	Acidic ribosomal protein P1—hydromedusa	*Polyorchis penicillatus*	8.00E − 20	S3A1
A055	CK662939	Ribosomal protein S21	*Rattus norvegicus*	8.00E − 17	S2
A319	CK663200	40S ribosomal protein S16	*Rattus norvegicus*	2.00E − 16	S1
A110	CK662994	Ribosomal protein L5	*Myxine glutinosa*	1.00E − 11	S1
A337	CK663218	Mitochondrial ribosomal protein S23	*Rattus norvegicus*	2.00E − 11	S1
A361	CK663242	60S acidic ribosomal protein P0	*Sus scrofa*	2.00E − 10	A1
A220	CK663102	Ribosomal protein L8	*Spodoptera frugiperda*	3.00E − 07	A1
tRNA synthesis/metabolism					
A198	CK663081	Histidyl-tRNA synthetase	*Takifugu rubripes*	1.00E − 77	A1
Cluster	Accession number	Putative product	*Species with best match*	Blastx E value	Library[a]
A250	CK663132	Valyl-tRNA synthetase	*Homo sapiens*	2.00E − 67	A2
A298	CK663179	tRNA selenocysteine associated protein	*Homo sapiens*	1.00E − 47	S1
A372	CK663253	Seryl-tRNA synthetase	*Helianthus annuus*	5.00E − 32	A1
Translation factors					
A125	CK663009	Elongation factor-1 α	*Locusta migratoria*	5.00E − 88	A2
A258	CK663140	Elongation factor 2	*Gallus gallus*	1.00E − 77	A2
A058	CK662942	Elongation factor 2	*Euglena gracilis*	3.00E − 75	S2
A301	CK663182	Protein synthesis initiation factor 4A	*Mus musculus*	1.00E − 35	S1
VII. Metabolism					
Amino acid					
A299	CK663180	Phenylethanolamine *N*-methyltransferase	*Xenopus laevis*	2.00E − 12	S1
A213	CK663096	γ-Butyrobetaine hydroxylase	*Drosophila melanogaster*	3.00E − 05	S2
A346	CK663227	2-Aminoethylphosphonate pyruvate aminotransferase	*Leishmania major*	8.00E − 06	S1

Table 3 (*continued*)

Cluster	Accession No.	Putative product	Species with best match	Blastx *E* value	Library[a]
A227	CK663109	Selenophosphate synthetase	*Drosophila melanogaster*	3.00E – 60	S1
Energy/TCA cycle					
A173	CK663056	Proteasome 26S ATPase subunit 6	*Rattus norvegicus*	1.00E – 129	S2
A253	CK663135	NAD(P) transhydrogenase	*Strongylocentrotus purpuratus*	8.00E – 81	A2
A375	CK663256	Oligomycin sensitivity-conferring protein	*Anopheles gambiae*	6.00E – 45	A1
A028	CK662912	Succinate dehydrogenase	*Mus musculus*	3.00E – 23	A2
A061	CK662945	ATP synthase, H^+ transporting	*Danio rerio*	9.00E – 16	S2A1
A031	CK662915	Monooxygenase	*Nicotiana tabacum*	1.00E – 13	A1
A093	CK662977	Ubiquinol-cytochrome *c* reductase subunit	*Bos taurus*	3.00E – 06	S1
A341	CK663222	NADH dehydrogenase 1β subcomplex 8	*Rattus norvegicus*	6.00E – 16	A1
Nucleotide					
A376	CK663257	Methylthioadenosine phosphorylase	*Danio rerio*	3.00E – 90	S1
A165	CK663046	Nucleoside-diphosphate kinase 2	*Homo sapiens*	2.00E – 54	S1
A313	CK663194	dUTP metabolism; Dut1p	*Saccharomyces cerevisiae*	8.00E – 18	S1
Protein modification					
A196	CK663079	Signal peptidase complex	*Mus musculus*	7.00E – 84	A2
Sugar/glycolysis					
A145	CK663029	α-1,3(6)-Mannosylglycoprotein β-1, 6-*N*-acetyl-glucosaminyltransferase	*Homo sapiens*	9.00E – 77	S1
A318	CK663199	Isocitrate dehydrogenase 2 (NADP+)	*Danio rerio*	2.00E – 55	S1
A306	CK663187	Trehalase 6-P hydrolase	*Shigella flexneri*	2.00E – 48	S1
A296	CK663177	UDP glycosyltransferase 1 family	*Bos taurus*	6.00E – 35	S1
A262	CK663144	Pyruvate dehydrogenase β chain 2 precursor	*Danio rerio*	4.00E – 32	A2
A371	CK663252	Galactosidase, β 1	*Mus musculus*	1.00E – 28	A1
A362	CK663243	Thymine DNA glycosylase	*Drosophila melanogaster*	2.00E – 18	A1
VIII. Unclassified					
A142	CK663026	RIKEN cDNA 1110037D14	*Mus musculus*	4.00E – 15	A1
A202	CK663085	Nucleotide-binding protein long form	*Rattus norvegicus*	1.00E – 113	S1
A345	CK663226	FLJ14511 protein	*Danio rerio*	3.00E – 81	S1
A291	CK663172	L-Alanine-DL-glutamate epimerase	*Burkholderia fungorum*	3.00E – 65	S1
A078	CK662962	Dead eye	*Danio rerio*	3.00E – 56	S1
A171	CK663054	Similar to hypothetical protein	*Rattus norvegicus*	6.00E – 55	S2
A263	CK663145	Px19-like protein	*Danio rerio*	5.00E – 46	S2
A309	CK663190	Unnamed protein product	*Homo sapiens*	5.00E – 40	S1
A335	CK663216	Hypothetical protein IMAGE3455200	*Xenopus laevis*	1.00E – 35	S1
A182	CK663065	Finkel-Biskis-Reilly murine sarcoma virus expressed protein	*Mus musculus*	2.00E – 34	S4A4
A251	CK663133	Hypothetical protein FLJ36664	*Homo sapiens*	8.00E – 31	A2
A332	CK663213	Hypothetical protein	*Macaca fascicularis*	1.00E – 27	A1
A180	CK663063	Hypothetical protein KIAA0602	*Homo sapiens*	1.00E – 26	A1
A248	CK663130	ENSANGP00000016831	*Anopheles gambiae*	6.00E – 26	A2
A160	CK663043	Chromosome 11 open reading frame 10	*Homo sapiens*	4.00E – 23	S1
A071	CK662955	CG7291-PA	*Drosophila melanogaster*	6.00E – 23	S1
A273	CK663155	RIKEN cDNA 1700001P01	*Homo sapiens*	4.00E – 18	S2
A097	CK662981	ENSANGP00000021085	*Anopheles gambiae*	9.00E – 12	S1
A148	CK663032	BAC19.9	*Lycopersicon esculentum*	4.00E – 09	A1
A360	CK663241	ENSANGP00000023519	*Anopheles gambiae*	6.00E – 09	A1
A195	CK663078	RIKEN cDNA G431004K08	*Rattus norvegicus*	8.00E – 09	A1
A133	CK663017	CG4692-PB	*Drosophila melanogaster*	1.00E – 06	A1
A194	CK663077	C7orf31 protein	*Homo sapiens*	1.00E – 05	S1
A054	CK662938	ENSANGP00000023016	*Anopheles gambiae*	2.00E – 05	S1
A293	CK663174	L-Isoaspartate (D-aspartate) *O*-methyltransferase	*Drosophila melanogaster*	3.00E – 24	S1
A245	CK663127	Protein for MGC:64448	*Xenopus laevis*	2.00E – 06	S3
A043	CK662927	Ovotransferrin	*Anas platyrhynchos*	3.00E – 32	S5

[a] EST library designations: A, aposymbiotic sea anemone EST library; S, symbiotic sea anemone EST library.

[23], we speculate that greater transcription of genes relative to protein synthesis would be necessary for symbiotic animal with high cell growth rate. Finally, the reasons of high cell structure/motility gene expression in the aposymbiotic animal are unclear, possibly because the host cells have changed their cell structure after the host cells lose their symbiotic algae.

In highly redundant ESTs, it can be considered that the genes were preferentially expressed only in a particular state and EST redundancy is proportional to the

expressed mRNA quantity [24,25]. Most state-specific ESTs occurred 1–3 times but some ESTs occurred several times. In this study, 11 unigenes occurred more than four times only in the symbiotic animal. These unigenes are as follows: ribosomal protein L5 (A033) (13 times), ferritin (A001) (12 times), 40S ribosomal protein S6 (A037) and ribosomal protein S10 (A086) (six times), ovotransferrin (A043) and cluster A042 (five times), and ribosomal protein S18 (A040), ubiquitin/ribosomal protein S27a fusion protein (A009), clusters A214, A377, and A039 (four times). On the other hand, six unigenes occurred more than four times only in the aposymbiotic animal. These unigenes are as follows: actin (A008) (10 times), ribosomal protein L19 (A007) (six times), ferritin heavy chain polypeptide 1 (A006), and A105 (five times), and ribosomal protein S19 (A101) and qm protein (A099) (four times). Although some of these genes, such as clusters A042, A214, A377, A039, and A105, were not identified, it suggests the possibility that these genes were specifically at symbiotic or aposymbiotic state and were possibly symbiosis-related genes.

In Table 1, ferritin (A001) is the most abundant ESTs in *A. pulchella* and is responsible for 11.7% of the total ESTs. Also, there are two different kinds of ferritin, cluster A006 and A004, being identified in Table 3. Ferritin is a widely distributed iron-storage protein thought to be very important to provide protection against the catalysis of deleterious oxidation of biomolecules by iron. Another important function of ferritin is detoxification, protecting cell against oxygen free radical-mediated cell damage [26,27]. The authors suggest that one of the possible roles of this protein is to protect sea anemone from oxygen free radicals produced during photosynthesis by its algal symbionts [28,29]. If this hypothesis were true, then it is not surprising that this kind of genes would be highly expressed in *A. pulchella*.

Gates et al. [30] demonstrated that thermal bleaching in the Hawaiian reef coral *Pocillopora damicronis* and in the sea anemone *A. pulchella* occurred predominantly by the release of intact host cells containing symbiotic algae. That is, thermal stress somehow affected the adhesive properties of the cells. Huang et al. [31] found that heat shock might induce algal release through exocytosis after demonstrating that heat shock alters Ca^{2+} in the coral *Acropora grandis*. As Ca^{2+} could cause loss of adhesive and the cell adhesion moleculars to which they are linked, change in Ca^{2+} could cause loss in adhesion causing collapse of the cytoskeleton. From the above statement, we can expect that some cytoskeleton proteins might be symbiosis-related proteins. Seven proteins associated with cytoskeletal and one protein associated with cell adhesive were found in the cDNA libraries of sea anemone (Table 3). Among these proteins, cluster A008 was identified as actin occurring 13 times in the aposymbiotic animal and three times in the symbiotic animal. It could be inferred that this protein, actin, might play an important role in symbiosis.

A homologue of the dnaJ gene of *Porphyromonas gingivalis* was identified (A070). Mutation of danJ gene in *Rhizobium leguminosarum* bv. *phaseoli* strain P121R results in mutant unable to use glutamate as the sole carbon and nitrogen source and is defective in symbiotic nitrogen fixation. This protein is required for the establishment of an effective symbiosis with *Phaseolus vulgaris* [32].

Since we sequenced randomly chosen cDNA clones, many of the identified sequences represented housekeeping protein. However, many important metabolism genes were identified in this study. Moreover, the metabolic and biochemical requirements between symbiotic partners are important research areas but little is known about their molecular basis. The EST analysis can suggest which metabolic pathways are operating but provide no data on relative rate or fluxes. Several genes involved in amino acid metabolism were identified (Table 3, Catalog VII). Notables are the genes associated with lysine degradation (A213), tyrosine metabolism (A299), aminophosphonate metabolism (A346), and selenoamino acid metabolism (A277). For carbon source metabolism, we identified genes associated with TCA cycle (A318), galactose metabolism (A371), *N*-glycan biosynthesis (A145), and starch and sucrose metabolism (A306).

In conclusion, this is the first report of the use of high-throughput EST analysis to examine the gene expression in symbiotic and aposymbiotic states of the marine cnidarians. The results presented here provide an overview of the transcriptionally and translationally active gene repertoire in two states of *A. pulchella*. The EST analysis exemplifies a potentially fertile approach for the identification of sea anemone genes expressed between symbiotic and aposymbiotic stages in *A. pulchella*. Aboundingly and moderately expressed sequences could be determined and their expression patterns and physiological significances are examined in more detail. This approach, coupled with other methods of gene expression analysis, may eventually unveil the molecular mechanisms of algal–cnidarian symbiosis.

Acknowledgments

This work was supported by a grant from the National Science Council (NSC 92-2611-B-291-001), Taiwan, and by intramural funding from NMMBA.

References

[1] K.L. Lin, J.T. Wang, L.S. Fang, Participation of glycoproteins on zooxanthellal cell walls in the establishment of a symbiotic

relationship with the sea anemone, *Aiptasia pulchella*, Zool. Stud. 39 (2000) 172–178.
[2] P.G. Falkowski, Z. Dubinsky, L. Muscatine, Light and the bioenergetics of a symbiotic coral, BioScience 34 (1984) 705–709.
[3] K.H. Mann, Ecology of Coastal Waters: A Systems Approach, Blackwell Scientific Publications, Oxford, 1982, pp. 160–182.
[4] B.E. Brown, J.C. Ogden, Coral bleaching, Sci. Am. 268 (1993) 64–70.
[5] P.W. Glynn, Coral reef bleaching: facts, hypotheses and implications, Global Change Biol. 2 (1996) 495–509.
[6] D.F. Gleason, G.M. Wellington, Ultraviolet radiation and coral bleaching, Nature 365 (1993) 836–838.
[7] J.M. Shick, M.P. Lesser, P.L. Jokiel, Ultraviolet radiation and coral stress, Global Change Biol. 2 (1996) 527–545.
[8] J.H. Drollet, M. Faucon, P.V.M. Martin, Elevated seawater temperature and solar UVB flux associated with two successive coral mass bleaching events in Tahiti, Mar. Freshwater Res. 46 (1995) 1153–1157.
[9] W. Goebel, R. Gross, Intracellular survival strategies of mutualistic and parasitic prokaryotes, Trends Microbiol. 6 (2001) 267–273.
[10] K. LeVier, R.W. Philips, V.K. Grippe, R.M. Roop II, G.C. Walker, Similar requirements of a plant symbiont and a mammalian pathogen for prolonged intracellular survival, Science 287 (2000) 2492–2493.
[11] V.M. Weis, W.S. Reynolds, Carbonic anhydrase expression and synthesis in the sea anemone *Anthopleura elegantissima* are enhanced by the presence of dinoflagellate symbionts, Physiol. Biochem. Zool. 72 (1999) 307–316.
[12] W.S. Reynolds, J.A. Schwarz, V.M. Weis, Symbiosis-enhanced gene expression in cnidarian–algal associations: cloning and characterization of a cDNA, sym32, encoding a possible cell adhension protein, Comp. Biochem. Physiol. A 126 (2000) 33–44.
[13] V.M. Weis, R.P. Levine, Differential protein profiles reflect the different lifestyles of symbiotic and aposymbiotic *Anthopleura elegantissima*, a sea anemone from temperate waters, J. Exp. Biol. 199 (1996) 883–892.
[14] M.C. Chen, Y.M. Cheng, L.H. Wang, C.H. Lin, X.Y. Huang, M.C. Liu, P.J. Sung, L.S. Fang, Cloning and characterization of the first cnidarian ADP–ribosylation factor, and its involvement in the *Aiptasia–Symbiodium* endosymbiosis, Mar. Biotech., in press.
[15] M.C. Chen, Y.M. Cheng, P.J. Sung, C.E. Kuo, L.S. Fang, Molecular identification of Rab7 (ApRab7) in *Aiptasia pulchella* and its exclusion from phagosomes harboring zooxanthellae, Biochem. Biophys. Res. Commun. 308 (2003) 585–594.
[16] M.D. Adams, J.M. Kelley, J.D. Gocayne, M. Dubnick, M.H. Polymeropoulos, H. Xiao, C.R. Merril, A. Wu, B. Olde, R.F. Moreno, A.R. Kerlavage, W.R. McCombie, J.C. Venter, Complementary DNA sequencing: expressed sequence tags and the human genome project, Science 252 (1991) 1651–1656.
[17] M.D. Adams, M. Dubnick, A.R. Kerlavage, R. Moreno, J.M. Kelley, T.R. Utterback, J.W. Nagle, C. Fields, J.C. Venter, Sequence identification of 2375 human brain genes, Nature 355 (1992) 632–634.
[18] S.F. Altschul, W. Gish, W. Miller, E.W. Myers, D.J. Lipman, Basic local alignment search tool, J. Mol. Biol. 215 (1990) 403–410.
[19] S.F. Altschul, T.L. Madden, A.A. Schaffer, J. Zhang, Z. Zhang, M. Miller, D.J. Lipman, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, Nucleic Acids Res. 25 (1997) 3389–3402.
[20] O. White, R. Kerlavage, TDB: new databases for biological discovery, Method Enzymol. 266 (1996) 24–40.
[21] M.A. Nelson, S. Kang, E.L. Braun, M.E. Crawford, P.L. Dolan, et al., Expressed sequences from conidial, mycelial, and sexual stages of *Neurospora crassa*, Fungal Genet. Biol. 21 (1997) 348–363.
[22] D. Qutob, P.T. Hraber, B.W.S. Sobral, M. Gijzen, Comparative analysis of expressed sequences in *Phytophthora sojae*, Plant Physiol. 123 (2000) 243–253.
[23] K.B. Mobley, A. Daniel, F. Gleason, The effect of light and heterotrophy on carotenoid concentrations in the Caribbean anemone *Aiptasia pallida* (Verrill), Mar. Biol. 143 (2003) 629–637.
[24] D.M. Hwang, A.A. Dempsey, C.Y. Lee, C.C. Liew, Identification of differentially expressed genes in cardiac hypertrophy by analysis of expressed sequence tags, Genomics 66 (2000) 1–14.
[25] M.D. Ospina-Giraldo, P.D. Collopy, C.P. Romaine, D.J. Royse, Classification of sequences expressed during the primordial and basidiome stages of the cultivated mushroom, Fungal Genet. Biol. 29 (2000) 81–94.
[26] K. Orino, L. Lehman, Y. Tsuji, H. Ayaki, S.V. Torti, F.M. Torti, Ferritin and the response to oxidative stress, Biochem. J. 357 (2001) 241–247.
[27] S.D. Aust, Ferritin as a source of iron and protection from iron-induced toxicities, Toxicol. Lett. 82–83 (1995) 941–944.
[28] J.A. Dykens, J.M. Shick, Oxygen production by endosymbiotic algae controls superoxide dismutase activity in their animal host, Nature 297 (1982) 579–580.
[29] J.A. Dykens, J.M. Shick, Photobiology of the symbiotic sea anemone *Anthopleura elegantissima*: defenses against photodynamic effects and seasonal photoacclimatization, Biol. Bull. 167 (1984) 683–697.
[30] R.D. Gates, G. Baghdasarian, L. Muscatine, Temperature stress causes host cell detachment in symbiotic cnidarians: implications for coral bleaching, Biol. Bull. 182 (1992) 324–332.
[31] S.P. Huang, K.K. Lin, L.S. Fang, The involvement of calcium in heat-induced coral bleaching, Zool. Stud. 37 (1998) 89–94.
[32] M. Labidi, S. Laberge, L.P. Vezina, H. Antoun, The dnaJ (hsp40) locus in *Rhizobium leguminosarum* bv. *phaseoli* is required for the establishment of an effective symbiosis with *Phaseolus vulgaris*, Mol. Plant–Microbe Interact. 13 (2000) 1271–1274.