The Journal of Microbiology, December 2007, p. 541-546 Vol. 45, No. 6


# Analysis of Expressed Sequence Tags from the Red Alga *Griffithsia okiensis*

**Hyoungseok Lee[1], Hong Kum Lee[1], Gynheung An[2], and Yoo Kyung Lee[1]***

[1]*Polar BioCenter, Korea Polar Research Institute (KOPRI), KORDI, Incheon 406-840, Republic of Korea*
[2]*Department of Life Science, Pohang University of Science and Technology, Pohang 790-784, Republic of Korea*

(Received October 26, 2007 / Accepted November 9, 2007)

**Red algae are distributed globally, and the group contains several commercially important species. *Griffithsia okiensis* is one of the most extensively studied red algal species. In this study, we conducted expressed sequence tag (ESTs) analysis and synonymous codon usage analysis using cultured *G. okiensis* samples. A total of 1,104 cDNA clones were sequenced using a cDNA library made from samples collected from Dolsan Island, on the southern coast of Korea. The clustering analysis of these sequences allowed for the identification of 1,048 unigene clusters consisting of 36 consensus and 1,012 singleton sequences. BLASTX searches generated 532 significant hits (E-value $<10^{-4}$) and via further Gene Ontology analysis, we constructed a functional classification of 434 unigenes. Our codon usage analysis showed that unigene clusters with more than three ESTs had higher GC contents (76.5%) at the third position of the codons than the singletons. Also, the majority of the optimal codons of *G. okiensis* and *Chondrus crispus* belonging to Bangiophycidae were C-ending, whereas those of *Porphyra yezoensis* belonging to Florideophycidae were G-ending. An orthologous gene search for the *P. yezoensis* EST database resulted in the identification of 39 unigenes commonly expressed in two rhodophytes, which have putative functions for structural proteins, protein degradation, signal transduction, stress response, and physiological processes. Although experiments have been conducted on a limited scale, this study provides a material basis for the development of microarrays useful for gene expression studies, as well as useful information for the comparative genomic analysis of red algae.**

***Keywords*:** codon usage, expressed sequence tags (ESTs), Gene Ontology (GO), *Griffithsia okiensis*, *Porphyra yezoensis*, red alga

Red algae are distributed globally, and the classification encompasses several commercially important species. *Griffithsia* is one of the most extensively studied red algal genera. *Griffithsia* C. Agardh (Ceramiaceae) is a multinucleate red algal genus. *Griffithsia* species have been utilized for physiological (Schechter, 1934; Bauman and Jones, 1986), cytological (Myers *et al.,* 1956; Peyrière, 1970; Goff and Coleman, 1987; Lee *et al*., 1995; Russell *et al.,* 1996; Lee *et al.,* 2001), and developmental studies (Höfler, 1934; Waaland, 1978; Hwang *et al.,* 1991; Lee *et al.,* 2002). Several genes have been reported from *G. japonica* (Lee *et al.,* 1998a; Lee *et al.,* 1998b; Lee *et al.,* 2002; Lee and Lee, 2003; Cho *et al.,* 2005), but this Korean *G. japonica* species was recently identified as *G. okiensis* on the basis of morphological and molecular data (Kim *et al.,* 2006).

EST analyses have been conducted with several species of red algae. ESTs of red algae were initially generated from the commercially important agar-producing alga, *Gracilaria gracilis* (Lluisma and Ragan, 1997). The number of *G. gracilis* EST was quite limited, and only 54 sequences evidenced significant matches with existing genes. Recently, 8,088 ESTs were analyzed from the other agar-producing alga, *Gracilaria changii*, and three genes for mannosyl glycerate synthesis were also identified (Teo *et al*., 2007). The ESTs from protoplasts and thalli from the red alga, *Chondrus crispus,* were compared in order to identify genes involved in cell wall regeneration and stress responses in red algae (Collén *et al*., 2006). *Porphyra* (Kim or Gim in Korean) is an important food alga, which is distributed globally and is economically important. A large number of ESTs (total 20,779 ESTs) was reported from the representative *Porphyra* species, *P. yezoensis*, and this represents 89% of the total number of algal ESTs. Codon usage analysis in the coding regions of non-redundant EST groups showed higher GC contents at the third positions of the codons (Nikaido *et al*., 2000). Differentially expressed genes in the gametophyte and the sporophyte were selected and confirmed via RT-PCR (Asamizu *et al*., 2003). On the basis of this database, SSR (simple sequence repeats) were developed for the analysis of the genetic diversity of *Porphyra* species (Sun *et al*., 2006). Certain gametogenesis-associated genes of *P. yezoensis* were identified using subtracted cDNA libraries (Kakinuma *et al*., 2006).

Despite their long history of academic study, only a very limited amount of molecular data is available regarding *Griffithsia.* Toward the large-scale identification of genes and their characteristic analysis, we conducted EST analysis and synonymous codon usage analysis using cultured *G. okiensis* samples.

## Materials and Methods

### *Sampling site and culture condition*

*Griffithsia okiensis* was collected from Dolsan Island on the

* To whom correspondence should be addressed.
(Tel) 82-32-260-6330; (Fax) 82-32-260-6301
(E-mail) yklee@kopri.re.kr

southern coast of Korea on April 3, 1992, and cultured as described in a previous paper (Lee *et al.,* 1995). The samples were then blotted with a paper towel in order to remove extra moisture, quickly frozen in liquid nitrogen, and stored at -70°C until use.

### *RNA extraction and cDNA library construction*

Total RNA was extracted from *Griffithsia okiensis* using TRI reagent (Molecular Research Center, Inc. USA) in accordance with the manufacturer's recommendations. Poly $(A)^+$ RNA was isolated using a poly-A Ttract mRNA isolation kit (Promega, USA). Following the electrophoresis of 50 μg mRNA, the fragments with sizes ≥500 bp were purified from agarose gel. The cDNA library was constructed from the purified mRNA using a cDNA Synthesis Kit and a ZAP-cDNA Gigapack III Gold Cloning Kit (Stratagene, USA).

### *Single-pass sequencing*

Clones for sequencing were selected randomly from the cDNA library. Each clone was cultured in 3 ml of LB, and the plasmid DNA was extracted with an AccuPrep Plasmid Extraction Kit (Bioneer, Korea). The sequencing reaction with T3 or T7 primers was run on an ABI PRISM 3100 Genetic Analyzer (Applied Biosystems, USA).

### *Sequence processing*

Processing was conducted using GeneMaster Ver. 3.0 (Ensoltech, Korea). Each EST was first individually processed using a multimodule custom pipeline which linked sequence backup, base calling, the elimination of sequences shorter than 50 bp (and low-quality sequences), vector trimming, and sequence assembly. The BLASTX algorithm (Altschul *et al*., 1990) was employed for the assignation of predicted functions to assembled consensus sequences on the basis of homology as previously described (Lee *et al*., 2004). Three categories of sequences were defined: 'Known', 'Unknown', and 'No hit'. They were defined as query sequences with e-values of $<10^{-14}$, $10^{-4}$~$10^{-14}$, and $>10^{-4}$, respectively. Finally, functional categories were defined using the Gene Ontology (GO) Database (www.geneontology.org) and the BLASTX algorithm. In order to compare our data and the ESTs from *Porphyra yezoensis*, a well-known red algal species, we downloaded a dataset from the NCBI database to our local system. Afterwards, homology searches were conducted using our ESTs as queries, and *Porphyra* ESTs as a subject using the BLASTN algorithm.

### *Measures of synonymous codon usage bias and GC frequency*

Among our EST dataset, CDSs of greater than 300 nucleotides in length were determined to avoid sample bias in the calculations of codon usage. Eight unigene clusters comprised of more than three clones and eight singletons were selected. Their synonymous codon usages were analyzed with a general codon usage analyzer, ver 2.0 (McInerney, 1998; http://gcua.schoedl.de/faq/cai.html). Codon usage data for other red algal species were imported from the kazusa website (http://www.kazusa.or.jp/codon/). The relative adaptiveness values for four dataset were calculated using a general codon usage analyzer, ver 2.0, on the basis of the methods described by Sharp and Li (1987). For each amino acid the codon with the highest frequency value was set to 100% relative adaptiveness, all other codons for the same amino acid were accordingly

**Table 1.** EST cluster collection statistics and distribution according to the BLASTX search results

| Strongest BLASTX hit | No. |
|---|---|
| High quality sequence | 1,104 |
| Cluster summary | |
| No. of singletons | 1,012 |
| No. of contigs | 36 |
| No. of unigenes | 1,048 |
| Annotation of unigenes | |
| Known | 418 (39.9%) |
| Unknown | 115 (11%) |
| No Hit | 515 (49.1%) |

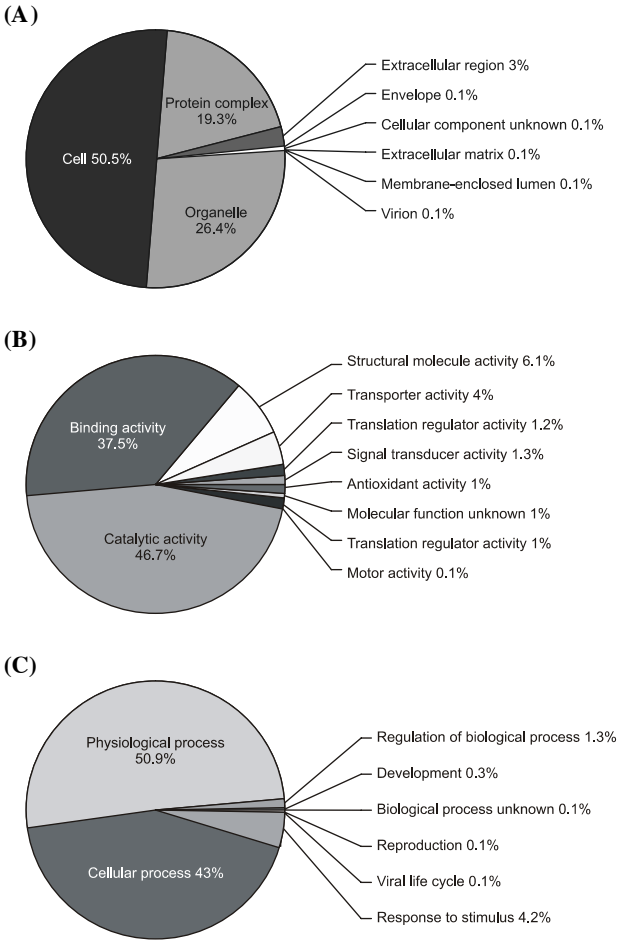


**Fig. 1.** Pie charts of the 1st level GO terms. Overall, 434 unigenes were annotated using GeneMaster software and were included in the graphs. Each of the three GO categories is presented including the cellular component (A), molecular function (B), and biological process (C).

scaled.

***GenBank accession numbers***

EST data of *G. okiensis* were deposited in GenBank.

## Results and Discussion

***General characteristics of G. okiensis ESTs***

A total of 1,123 cDNA clones were selected randomly from the library and single-pass sequences were generated. Poor quality sequences, including sequences of <100 bases were removed, and the final EST number was 1,104 with mean lengths of 506 (Table 1). Cluster analysis was conducted, revealing 36 contigs and 1,012 singletons, with 1,048 unigenes and a redundancy level of 8.3%. The sequences of 1,048 unigenes, generated from the EST clustering, were translated in all possible reading frames, and compared to the NCBI non-redundant protein database. The annotation results were exported into a Microsoft Excel data sheet and sorted manually. From the library, 39.9% and 11% of unigenes evidenced similarity to genes encoding for 'Known' or 'Unknown' proteins, respectively. In addition, 49.1% of the ESTs evidenced no similarity to any sequences in the database, and were deemed 'No hit' (Table 1).

***Functional classification using GO database***

The unigenes were compared to annotations through the Gene Ontology Consortium (Ashburner *et al*., 2000) using the GeneMaster automated software package. We were able to obtain GO terms for 434 unigenes via this method. GO graphs using percentages of 1st level GO terms are presented under the categories of cellular components (Fig. 1A), molecular functions (Fig. 1B), and biological processes (Fig. 1C). Of the cellular component GO terms, 50.5% and 26.4% were directly associated with cellular and organelle components, respectively. In the molecular functions category, the vast majority were associated with catalytic activity (46.7%) and binding activities (37.5%). Under the biological processes category, 50.9% were involved in physiological processes; 43% were involved in cellular processes, 4.2% in responses to stimuli. In a large proportion of the genes identified in this study, approximately 58.6% of 614 unigenes could not be assigned a putative function in the GO database. This observation reflects the specificity of the metabolism of the red algal lineage and the scarcity of relevant information within the protein database.

***Features of protein-coding region***

As the EST database grows, digital northern analyses were made possible through the accumulation of large numbers of ESTs, in which the gene expression level can be deduced *in silico* via the calculation of EST frequencies in different cDNA libraries (Da Silva *et al*., 2005). The statistical significance of such digital expression profiling applied to representative EST datasets has been previously validated in several publications (Stekel *et al*., 2000; Fernandez *et al*., 2003; Susko and Roger, 2004).

In order to characterize the features of the coding regions of *G. okiensis* genes evidencing high expression levels, eight unigene clusters composed of more than 3 ESTs were selected, and were identified to encode for three different light harvesting proteins, carbonic anhydrase, photosystem II 12 kD extrinsic protein, notch-like protein 1a, protein disulfide isomerase 1, and peptidoglycan-binding domain protein. Another 24 singletons were also selected, after which the codon usage and GC contents of these two groups were analyzed. All the selected unigenes evidenced high similarity with known genes with very low E-values (< E-14). One notable feature was that the codons with GC in the third position were significantly more frequently utilized in unigene clusters (76.5%) than in singletons (67.2%). It has been established that a link exists between the third position GC content and gene expression level. Also, highly expressed genes, as compared with weakly expressed ones, selectively utilize 'optimal codons', which correspond to abundant tRNAs, thus improving their translational efficiency (Semon *et al.*, 2005). These results suggest that a positive correlation may exist between GC content at the third position and the gene expression level from *G. okiensis*.

We also assessed the distribution of the optimal codons of *G. okiensis* (Florideophycidae), *P. yezoensis* (Bangiophycidae), and *C. crispus* (Florideophycidae) using datasets from the kazusa website. The optimal codons for *G. okiensis* and *C. crispus* were GCC for Ala, CGC for Arg, GGC for Gly, CUC for Leu, CCC for Pro, UCC for Ser, ACC for Thr, and GUC for Val; however, in the case of *P. yezoensis*, the optimal codons were GCG for Ala, CGC for Arg, GGC for Gly, CUG for Leu, CCG for Pro, UCG for Ser, ACG for

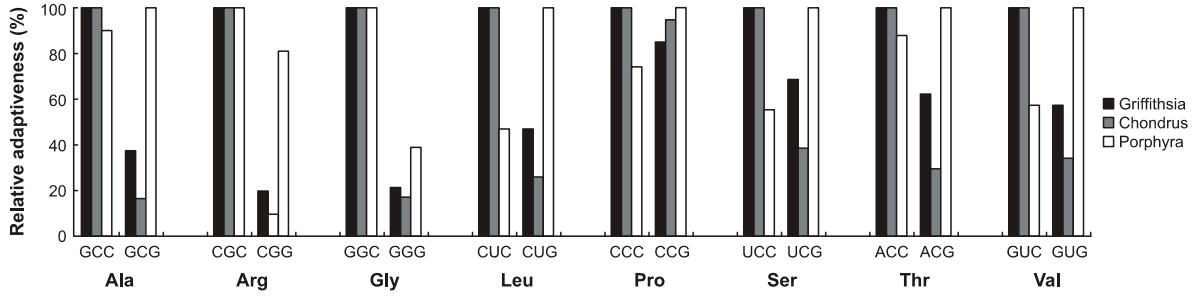


**Fig. 2.** Synonymous codon usage bias for codons of eight amino acids with four or more synonymous codons from three red algal species. Codon usages were calculated with a general codon usage analyzer ver. 2.0 with the data of selected genes from *G. okiensis* of this study and datasets for *P. yezoensis* and *C. crispus* imported from kazusa website (http://www.kazusa.or.jp/codon/). Relative adaptiveness values were calculated as described in the Materials and Methods section. Griffithsia, *G. okiensis*; Porphyra, *P. yezoensis*; Chondrus, *C. crispicus*.

**Table 2.** Annotation and functional classification of 41 unigenes showing high similarity with *P. yezoensis* EST clones

| Query ID | BLASTX | | Porphyra EST BLASTN | | |
|---|---|---|---|---|---|
| | Subject | E-value | Subject | E-value | Similarity (%) |
| **Structural proteins** | | | | | |
| Gj689 | Actin 1 (*Nemalionopsis shawii*, CAI56221.1) | 2.00E-66 | AV430104 | 6.00E-33 | 81 |
| Gj775 | Actin 1 (*N. shawii*, CAI56221.1) | 0 | AV431245 | 3.00E-81 | 84 |
| Gj901 | Histone H2A protein (*Strongylocentrotus purpuratus*, XP_790782.1) | 3.00E-81 | AV433708 | 4.00E-25 | 81 |
| Gj317 | Histone 1 2Bp (*Mus musculus*, AAH61044.1) | 2.00E-61 | AV431967 | 6.00E-30 | 81 |
| Gj488 | Histone 3 (*Conocephalum conicum*, BAD90754.1) | 3.00E-68 | AV432155 | 3.00E-37 | 80 |
| Gj410 | Histone H3 protein, putative (*Danio rerio*, XP_693155.1) | 2.00E-69 | AV432155 | 3.00E-53 | 81 |
| Gj475 | Histone H3.3 (*D. rerio*, XP_689389.1) | 3.00E-68 | AV432155 | 7.00E-73 | 83 |
| Gj458 | Histone protein Hist2h3c1 (*Gallus gallus*, XP_416193.1) | 1.00E-38 | AV431941 | 2.00E-26 | 80 |
| Gj383 | 40S Ribosomal protein S14 (*Oryza sativa*, XP_472410.1) | 3.00E-49 | AV438236 | 3.00E-29 | 78 |
| Gj1108 | Ribosomal protein L10 (*Branchiostoma belcheri* tsingtaunese, AAO31769.1) | 7.00E-79 | AV439248 | 2.00E-88 | 86 |
| Gj1095 | Ribosomal protein S23 (*S. purpuratus*, XP_795223.1) | 6.00E-53 | AV439167 | 8.00E-23 | 80 |
| Gj1118 | Ribosomal protein S7 (*Argopecten irradians*, AAN05602.1) | 5.00E-44 | AV439127 | 1.00E-15 | 87 |
| Gj579 | Ribosomal protein S7 (*G. gallus*, XP_419936.1) | 9.00E-57 | AV439127 | 2.00E-15 | 87 |
| Gj1087 | Ribosomal protein S7 (S8) (*S. purpuratus*, XP_789754.1) | 2.00E-37 | AV439127 | 9.00E-16 | 87 |
| **GTP-binding** | | | | | |
| Gj669 | GTP-binding nuclear protein spi1 (*Schizosaccharomyces pombe*, P28748) | 1.00E-77 | AV431689 | 9.00E-29 | 84 |
| Gj786 | Rab1B protein (*Rattus norvegicus*, 1515250A) | 3.00E-80 | AU189541 | 4.00E-53 | 83 |
| Gj948 | Small GTPase rab11-3 (*G. lemaneiformis*, NP_849539.1) | 1.00E-80 | AV435407 | 3.00E-16 | 81 |
| **Stress response** | | | | | |
| Gj1202 | Heat shock 90 kDa protein (*Cyanidioschyzon merolae*, BAC67671.2) | 1.00E-72 | AV432126 | 4.00E-37 | 84 |
| Gj581 | Heat shock 90 kDa protein (*C. merolae*, BAC67671.2) | 3.00E-82 | AV429526 | 5.00E-18 | 82 |
| Gj354 | Heat shock protein 70 (*Dictyostelium discoideum*, CAA53039.1) | 2.00E-89 | AU193203 | 6.00E-49 | 80 |
| Gj918 | Heat shock protein 82 (*Guillardia theta*, NP_113234.1) | 2.00E-70 | AU187975 | 2.00E-29 | 82 |
| **Protein degradation** | | | | | |
| Gj337 | Polyubiquitin (*Gracilaria lemaneiformis*, AAX56917.1) | 6.00E-49 | AU197049 | 3.00E-62 | 85 |
| Gj949 | Polyubiquitin (*G. lemaneiformis*, AAX56917.1) | 2.00E-78 | AV433190 | 2.00E-91 | 84 |
| Gj532 | Polyubiquitin (*Aglaothamnion neglectum*, AAA72126.1) | 7.00E-93 | AU190729 | 8.00E-94 | 85 |
| Gj418 | 26S proteasome regulatory subunit (*Xenopus laevis*, AAH97594.1) | 1.00E-57 | AU197062 | 2.00E-27 | 80 |
| Gj112 | 26S proteasome subunit 4-like protein (*Brassica napus*, CAC14432.1) | 3.00E-85 | AV435174 | 2.00E-41 | 81 |
| Gj1088 | Proteasome subunit, beta type 7 (*G. gallus*, CAG32014.1) | 2.00E-64 | AV436932 | 5.00E-15 | 81 |
| Gj846 | ATP-dependent Clp protease proteolytic subunit (*G. theta*, NP_113298.1) | 6.00E-68 | AV431813 | 1.00E-37 | 87 |
| **Physiological processes** | | | | | |
| Gj931 | Adenosine 5' phosphosulfate reductase (*Populus alba* x *Populus tremula*, AAQ57202.1) | 3.00E-44 | AU189960 | 2.00E-14 | 88 |
| Gj154 | Alanine-glyoxylate aminotransferase (*Noctoc* sp. PCC7120, NP_485047.1) | 0 | AV430455 | 6.00E-21 | 80 |
| Gj225 | ATP-sulfurylase (*Medicago truncatula*, ABE89666.1) | 0 | AU188628 | 4.00E-16 | 81 |
| Gj331 | DNA repair protein (*M. musculus*, BAE32801.1) | 2.00E-72 | AU187263 | 8.00E-17 | 93 |
| Gj1185 | Fructose-1,6-biphosphate aldolase precursor (*Galdieria sulphuraria*, AAF27641.1) | 3.00E-64 | AV432532 | 1.00E-27 | 83 |
| Gj395 | Fructose-1,6-biphosphate aldolase precursor (*G. sulphuraria*, AAF27641.1) | 3.00E-66 | AV435187 | 5.00E-46 | 83 |
| Gj561 | Glycine decarboxylase p protein (*Anas platyrhynchos*, AAD56281.1) | 0 | AV435132 | 2.00E-15 | 82 |
| Gj917 | NAD dependent epimerase/dehydratase family protein (*O. sativa*, ABG66078.1) | 2.00E-46 | AU194678 | 1.00E-21 | 82 |
| Gj919 | NADH dehydrogenase (ubiquinone) chain PSST precursor (*O. sativa*, NP_916686.1) | 4.00E-49 | AV438819 | 3.00E-22 | 89 |
| Gj1131 | Pyridoxine biosynthesis protein, putative (*O. sativa*, AAX95426.1) | 8.00E-45 | AU188015 | 1.00E-27 | 81 |
| Gj530 | Sedoheptulose-1,7-bisphosphatase (*Bigelowiella natans*, AAP79184.1) | 7.00E-65 | AV433409 | 1.00E-27 | 84 |

Thr, and GUG for Val (Fig. 2). Optimal codons for eight amino acids harboring four or more synonymous codons were all C-ending in *G. okiensis* and *C. crispus,* and six were G-ending and two were C-ending in *P. yezoensis*. These data indicate that different codon usage bias at the third codon position might be a useful key to distinguish Bangiophycidae (G-ending codon) from Florideophycidae (C-ending codon).

***Orthologous genes search for P. yezoensis EST database***
The EST data represent only a fraction of all of the genes of an organism. Thus, a comparison of EST data alone cannot be used to describe unique set of genes from an organism. *P. yezoensis* is the representative species of the red algae, and its EST dataset is composed of 21,954 sequences. This makes it possible to detect potential orthologous genes if present. A complete list of 39 unigenes evidencing only similarity to proteins with known functions present in both organisms can be seen in Table 2. Many genes shared by *G. okiensis* and *P. yezoensis* encoded for the expected structural proteins, such as genes encoding for two actins, six histones, and six ribosomal proteins. Three different GTP-binding protein-encoding genes were identified in both species, suggesting that they may share cellular signaling components and very similar signal transduction mechanisms.

A variety of heat shock proteins (HSPs) have been commonly identified, and this is also the case for *G. gracilis* (Lluisma and Ragan, 1997). HSPs have been shown to perform a crucial function in the recovery of cells from stress (Nollen and Morimoto, 2002) and are usually encoded in eukaryotes by highly conserved multigene families. Under normal conditions, HSP70 evidences chaperone activities and assists in the folding of newly synthesized proteins. Under stress conditions, including oxidative stress, HSP70 prevents protein aggregation and allows for the refolding of denatured proteins (Sheffield *et al*., 1990). On the basis of these reports, the common action of HSPs in red algae may involve the protection of proteins against oxidation damage and the prevention of initiation of cell death under natural conditions.

Polyubiquitin genes encode for direct repeats of ubiquitin contiguously arrayed and covalently joined, and they function in non-lysosomal ATP-dependent selective proteolysis and many other cellular processes (Johnson *et al*., 1992). Zhou and Ragan (1995) reported the presence of the polyubiquitin gene in *Gracilaria verrucosa* for the first time in red algae. Only one polyubiquitin gene was detected in *G. verrucosa;* however, *G. okiensis* and *P. yezoensis* shared at least three different polyubiquitin genes (Table 2). Other than polyubiquitin, protein degradation machinery-associated genes, encoding for 26S proteasome subunit proteins and the ATP-dependent Clp protease proteolytic subunit, were identified. These results indicate that the mechanism underlying protein degradation might be well-conserved in different red algal species.

Several genes involved in the synthesis and modification of biomolecules were shared by two red algal species, which encode for metabolism-associated proteins including fructose-1,6-biphosphate aldolase and sedoheptulose-1,7-bisphosphatase associated with carbohydrate metabolism, pyridoxine biosynthesis protein associated with vitamin synthesis, ATP-sulfurylase associated with sulfate assimilation, and glycine decarboxylase p protein, which is related to glycine metabolism. Their putative orthologous sequences were detected in other organisms from different kingdoms, including *Populus alba* x *Populus tremula*, *Nostoc sp*., *Medicago truncatula*, *Oryza. sativa*, thus showing that they are not specific to red algae.

## Conclusions

In this study, we have generated 1,104 high quality *G. okiensis* EST sequences. Sequence analysis evidenced the presence of 1,048 unigenes in the dataset. Among them, 39.9% of the unigenes showed high similarity with known genes, making them useful for further analyses. Microarrays can now be designed using either cDNA microarrays or oligo-based platforms using the EST information. The unigene dataset provided herein should enhance the effectiveness of molecular studies, including those focused on marker screening and gene expression profiling under a variety of environmental conditions. Finally, the study of different codon usage bias at the third codon position from different red algal species would prove quite useful for the classification of the Bangiophycidae and Florideophycidae.

## Acknowledgements

We thank the reviewers for their critical and useful comments. This work was supported by grant PE07050 from the Korea Polar Research Institute (KOPRI).

## References

Altschul, S.F., W. Gish, W. Miller, E.W. Meyers, and D. Lipman. 1990. Basic local alignment search tool. *J. Mol. Biol*. 215, 403-410.

Asamizu, E., M. Nakajima, Y. Kitade, N. Saga, Y. Nakamura, S. Tabata. 2003. Comparison of RNA expression profiles between the two generations of *Porphyra yezoensis* (Rhodophyta), based on expressed sequence tag frequency analysis. *J. Phycol*. 39, 923-930.

Ashburner, M., C.A. Ball, J.A. Blake, D. Botstein, H. Butler, J.M. Cherry, A.P. Davis, K. Dolinski, S.S. Dwight, J.T. Eppig, M.A. Harris, D.P. Hill, L. Issel-Tarver, A. Kasarskis, S. Lewis, J.C. Matese, J.E. Richardson, M. Ringwald, G.M. Rubin, and G. Sherlock. 2000. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet*. 1, 25-29.

Bauman, Jr., R.W. and B.R. Jones. 1986. Electrophysiological investigations of the red alga *Griffithsia pacifica* KYL. *J. Phycol.* 22, 49-56.

Cho, E.K., Y.K. Lee, and C.B. Hong. 2005. A cyclophilin from *Griffithsia japonica* has thermoprotective activity and is affected by CsA. *Mol. Cells* 20, 142-150.

Collén, J., V. Roeder, S. Rousvoal, O. Collin, B. Kloareg, and C. Boyen. 2006. An expressed sequence tag analysis of thallus and regenerating protoplasts of *Chondrus crispus* (Gigartinales, Rhodophyceae) *J. Phycol*. 42, 104-112.

Da Silva, F.G., A. Iandolino, F. Al-Kayal, M.C. Bohlmann, M.A. Cushman, H. Lim, A. Ergul, R. Figueroa, E.K. Kabuloglu, C. Osborne, J. Rowe, E. Tattersall, A. Leslie, J. Xu, J. Baek, G.R. Cramer, J.C. Cushman, and D.R. Cook. 2005. Characterizing the grape transcriptome. Analysis of expressed sequence tags

from multiple *Vitis* species and development of a compendium of gene expression during berry development. *Plant Physiol*. 139, 574-597.

Fernandez, P., N. Paniego, S. Lew, H.E. Hopp, and R.A. Heinz. 2003. Differential representation of sunflower ESTs in enriched organ-specific cDNA libraries in a small scale sequencing project. *BMC Genomics* 4, 40.

Goff, L.J. and A.W. Coleman. 1987. The solution to the cytological paradox of isomorphy. *J. Cell Biol.* 104, 739-748.

Höfler, K. 1934. Regenerationsvorgänge bei *Griffithsia schousboei*. *Flora* 27, 331-344.

Hwang, M.S., H.S. Kim, and I.K. Lee. 1991. Regeneration and sexual differentiation of *Griffithsia japonica* (Ceramiaceae, Rhodophyta) through somatic cell fusion. *J. Phycol*. 27, 441-447.

Johnson, E.S., B. Bartel, W. Seufert, and A. Varshavsky. 1992. Ubiquitin as a degradation signal. *EMBO J*. 11, 497-505.

Kakinuma, M., I. Kaneko, D. Coury, T. Suzuki, and H. Amano. 2006. Isolation and identification of gametogenesis-related genes in *Porphyra yezoensis* (Rhodophyta) were identified using subtracted cDNA libraries. *J. Appl. Phycol*. 18, 489-496.

Kim, H.S., E.C. Yang, and S.M. Boo. 2006. The occurrence of *Griffithsia okiensis* (Ceramiaceae, Rhodophyta) from Korea on the basis of morphology and molecular data. *Algae* 21, 91-101.

Lee, H., C.G. Hur, C.J. Oh, H.B. Kim, S.Y. Park, and C.S. An. 2004. Analysis of the root nodule-enhanced transcriptome in soybean. *Mol. Cells* 18, 53-62.

Lee, Y.K., H.-G. Choi, C.B. Hong, and I.K. Lee. 1995. Sexual differentiation of *Griffithsia* (Ceramiales, Rhodophyta): nuclear ploidy level of mixed-phase plants in *G. japonica. J. Phycol.* 31, 668-673.

Lee, Y.K., C.B. Hong, N.-K. Lim, J.-S. So, and I.K. Lee. 1998a. Isolation & characterization of a cDNA encoding 40S ribosomal protein S8 from a red alga, *Griffithsia japonica* (Ceramiales, Rhodophyta). *Algae* 13, 157-163.

Lee, Y.K., C.B. Hong, Y. Soh, and I.K. Lee. 2002. A cDNA clone for cyclophilin from *Griffithsia japonica* (Ceramiales, Rhodophyta) and phylogenetic analysis of cyclophilins. *Mol. Cells* 13, 12-20.

Lee, Y.K., M.S. Hwang, and I.K. Lee. 2001. Sexual differentiation of *Griffithsia monilis* (Ceramiales, Rhodophyta) in hybrids between female & male thalli. *Botanica Marina* 44, 547-557.

Lee, Y.K., S.H. Kim, C.B. Hong, O.-K. Chah, G.H. Kim and I.K. Lee. 1998b. Heat-shock protein 90 may be involved in differentiation of the female gametophytes in *Griffithsia japonica* (Ceramiales, Rhodophyta). *J. Phycol.* 34, 1017-1023.

Lee, Y.K. and H.K. Lee. 2003. Putative histone *H2A* genes from a red alga, *Griffithsia japonica*. *Algae* 18, 191-197.

Lluisma, A.O. and M. Ragan. 1997. Expressed sequence tags (ESTs) from the marine red alga *Gracilaria gracilis*. *J. Appl. Phycol*. 9, 287-293.

Myers, A., R.D. Preston, and G.W. Ripley. 1956. Fine structure in the red algae. I. X-ray & electron microscope investigation of *Griffithsia flosculosa*. *Proc. Roy. Soc. B.* 144, 450-459.

Nikaido, I., E. Asamizu, M. Nakajima, Y. Nakamura, N. Saga, and S. Tabata. 2000. Generation of 10,154 expressed sequence tags from a leafy gametophyte of a marine red alga, *Porphyra yezoensis*. *DNA Res*. 7, 223-227.

Nollen, E.A. and R.I. Morimoto. 2002. Chaperoning signaling pathways: molecular chaperones as stress-sensing 'heat shock' proteins. *J. Cell Sci*. 115, 2809-2816.

PeyrièRe, M. 1970. Evolution de l'appareil de Golgi de la tetrasporogenèse de *Griffithsia flosculosa* (Rhodophycée). *C. R. Acad. Sci. Paris, Sér. D.* 270, 2071-2074.

Russell, C.A., M.D. Guiry, A.R. Mcdonald, and D.J. Garbary. 1996. Actin-mediated chloroplast movement in *Griffithsia pacifica* (Ceramiales, Rhodophyta). *Phycol. Res.* 44, 57-61.

Schechter, V. 1934. Electrical control of rhizoid formation in the red alga, *Griffithsia borretiana. J. Gen. Physiol*. 18, 1-22.

Semon, M., D. Mouchiroud, and L. Duret. 2005. Relationship between gene expression and GC-content in mammals: statistical significance and biological relevance. *Hum. Mol. Genet*. 14, 421-427.

Sharp, P.M. and W.H. Li. 1987. The codon Adaptation Index - a measure of directional synonymous codon usage bias, and its potential applications. *Nucleic Acids Res*. 15, 1281-1295.

Sheffield, W.P., G.C. Shore, and S.K. Randall. 1990. Mitochondrial precursor protein. Effects of 70-kilodalton heat shock protein on polypeptide folding, aggregation, and import competence. *J. Biol. Chem*. 265, 11069-11076.

Stekel, D.J., Y. Git, and F. Falciani. 2000. The comparison of gene expression from multiple cDNA libraries. *Genome Res*. 10, 2055-2061.

Sun, J., T. Liu, B. Guo, D. Jin, M. Weng, Y. Feng, P. Xu, D. Duan, and B. Wang. 2006. Development of SSR primers from EST sequences and their application in germplasm identification of *Porphyra* lines (Rhodophyta). *Eur. J. Phycol*. 41, 329-336.

Susko, E. and A.J. Roger. 2004. Estimating and comparing the rates of gene discovery and expressed sequence tag (EST) frequencies in EST surveys. *Bioinformatics* 20, 2279-2287.

Teo, S.-S., C.-L. Ho, S. Teoh, W.-W. Lee, J.-M. Tee, R.A. Rahim, and S.-M. Phang. 2007. Analyses of expressed sequence tags from an agarophyte, *Gracilaria changii* (Gracilariales, Rhodophyta). *Eur. J. Phycol*. 42, 41-46.

Waaland, S.D. 1978. Parasexually produces hybrids between male and female plants of *Griffithsia tenuis* C. Agardh, a red alga. *Planta (Berl)* 138, 65-68.

Zhou, Y.H. and M.A. Ragan. 1995. Characterization of the polyubiquitin gene in the marine red alga *Gracilaria verrucosa. Biochim. Biophys. Acta*. 1261, 215-222.