©JLS

Genetic Study of the Class Dinophyceae Including Red Tide Microalgae Based on a Partial Sequence of SSU Region: Molecular Position of Korean Isolates of *Cochlodinium polykrikoides* Margalef and *Gyrodinium aureolum* Hulburt

Eun Seob Cho*

South Sea Fisheries Research Institute, NFRDI, Yeosu 556-823

The nucleotide sequence for a nuclear-encoded small subunit rDNA (SSU rDNA) was determined for 43 species of the class Dinophyceae, including harmful algae *Cochlodinium polykrikoides* and *Gyrodinium aureolum*. These sequences and data analyses were performed by parsimony, distances and maximum likelihood methods in PHYLIP (Phylogenetic Inference Package) version 3.573c. The species *Noctiluca scintillans*, *Gonyaulax spinifera* and *Crypthecodinium cohnii* occupied a basal position within the Dinophyceae in our analyses. The genera *Alexandrium* and *Symbiodinium* were monophyletic (supported by a bootstrap value of >70%), whereas the genera *Gymnodinium* and *Gyrodinium* formed polyphyletic nodes, for which bootstrap support was strong (>70%) in the neighbor-joining and maximum likelihood methods except for the PHYLIP parsimony analysis (=59%). The sequence divergence between *G. aureolum* and *G. dorsum*/*G. galathenum* was the largest at 7.4% (45 bp), whereas *G. aureolum* and *G. mikimotoi* showed an extremely low value of genetic divergence of 0.9% (5 bp). The genetic divergence between *C. polykrikoides* and *G. aureolum* was a low value of 5.2% (31 bp). In the phylogenetic analysis, the placement of *G. aureolum* and *C. polykrikoides* was closer to the genus *Gymnodinium* than to the genus *Gyrodinium*, which was supported by a moderate bootstrap value.

Key words – *Cochodinium polykrikoides*, Dinophyceae, Genetic study, *Gyrodinium aureolum*, morphology, phylogeny, red tide, SSU

Harmful algal blooms (HABs) are increasing in frequency or severity in many coastal environments and worldwide. The blooms of unicellular marine algae can cause mass mortality in a variety of marine organisms and can cause illness and even death in humans who consume contaminated seafood. At present, about 57 HAB species are capable of producing potent toxins and causing serious problems globally [7]. Most HAB species are of the class Dinophyceae, which can be divided into five types (Gymnodinioids, Gonyaulacoids, Dinophysoids, Prorocentroids, and Peridinoids) according to the morphological features described by Taylor [29]. Furthermore, the identification of dinoflagellates species is sometimes painstaking and the discrimination of each is sometimes confusing because of closely related morphology in a few species and of possible variability in morphological features depending on environmental conditions [1,22]. For example, the genera *Alexandrium*, *Gymnodinium* and *Prorocentrum* have ongoing considerable studies and discussions on the morphological characteristics, which are subjects of interest to taxonomists.

Recent advances in DNA amplification and sequencing provide powerful tool to examine phylogenetic relationships of both closely and distantly related organisms. Nuclear ribosomal RNA genes are attractive candidates for molecular studies at taxonomic levels. The sequence analysis of the small subunit (SSU) is one method that can be used to classify populations and to identify specific genetic markers of *Alexandrium* [9,24-26]. In addition, the sequence data from SSU region provide information on heterogeneity within the genus *Prorocentrum* [12]. The dinoflagellate *Cochlodinium polykrikoides* occurred for the first time in 1982 in Korean coastal waters, and was associated with massive fish kills [18]. Impacts of *C. polykrikoides* appear to have increased in frequency, intensity and geographic distribution. Previously, we have studied the molecular structures and systematics of *C. polykrikoides* using only internal transcribed spacers (ITS) regions [3-6]. However, little has been investigated of the sequence analysis of *C. polykrikoides* targeted to the SSU region.

Recently, Dr. K. J. Cho, Inje University, Korea, isolated *Gyrodinium aureolum* from Chinhae Bay, Korea, which is a species closely related to *Gymnodinium mikimotoi* [2]. Despite the fact that the morphological characterstics of *G. aureolum*

*Corresponding author
Tel : +82-61-690-8959, Fax : +82-61-686-1588
E-mail : eun-5657@hanmail.net

have been much studied and debated, some genetic characteristics remain to be determined and understood. As mentioned above, the sequence analysis of SSU has been proved to be useful for the study of intra and interspecific genetic variation in Dinophyceae. This study sequenced the SSU to investigate the genetic divergence of *C. polykrikoides* and *G. aureolum* from related species, and also obtained other sequences through a GenBank search to form a combined data set for the analysis of genetic divergence and phylogenetic relationships.

Materials and Methods

Cultures

Cochlodinium polykrikoides was isolated from the coast of Tongyong, Korea, in 1997 and has been grown in f/2-Si medium [13] containing an antibiotic mixture [15]. *Gyrodinium aureolum* was provided by Dr. K. J. Cho, Inje University, Kimhae, Korea. These species have maintained at 20 ℃ under a photon flux from cool white fluorescent tubes of 50 μmol m^{-2} s^{-1} in a 14: 10 h L:D cycle.

Isolation of genomic DNA

Cultures in exponential phase were harvested by centrifugation. Pelleted cells were frozen immediately and stored at -20 ℃ until required. Cells were thawed in 500 μL of extraction buffer (100 mM Tris-HCl, pH 8.0; 40 mM EDTA) and 150 μL of 10% sodium dodecyl sulphate (SDS), and incubated at 55 ℃ for 30 min. The supernatant was extracted twice with phenol : chloroform : isoamylalchol (25 : 24 : 1, v/ v/v) and RNase (1 mg mL^{-1}) was added. The algal cells were pelleted and the supernatant mixed with 40 μL of 3 M sodium acetate (pH 4.0) and incubated on ice for 1 h. The DNA was precipitated from the tube at room temperature for 10 min by adding 2.5 volumes of 100% ice-cold ethanol. The DNA pellet was washed with 2 volumes of 70% ethanol and was then dried by desiccator. The DNA was dissolved in 20 μL TE buffer (10 mM Tris, 0.1 mM EDTA, pH 7.5) and was kept at -20 ℃. Gel electrophoresis was performed at voltage of 100 V for 20 min at room temperature. After electrophoresis, the 1.5% agarose gel was visualized by embedding in ethidium bromide (2 μL ml^{-1}) and irradiating by ultraviolet light (312 nm).

DNA amplification and sequencing

The partial region of nuclear SSU rDNA was amplified from genomic DNA using the polymerase chain reaction

(PCR) with the two primers [17], NS1 (5′-GTAGTCATA-TGCTTGTCTC-3′) and NS2 (5′-GGCTGCTGGCACCAGAC-TTGC-3′). PCR was performed in a 20 uL reaction (Korea Biotech. Inc., Ltd., Taejon, Korea) mixture using Perkin-Elmer 2400 Thermocycler under the following cycle conditions; pre-denaturating at 95 ℃ for 30 sec, followed by 35 cycles of denaturating at 95 ℃ for 1 min, primer annealing at 50 ℃ for 1 min, and 5 min chain extenstion at 72 ℃ for 5 min. The PCR product after the amplification was subjected to preparative electrophoresis in a 1.5% agarose gel in TBE buffer. All PCR products yielded only a single visible band. Target fragments were cut off from the gel and withdrawn using a QIAGEN gel elution kit (Qiagen, Wartworth CA) according to the instructions by the manufacture. Purified rDNA region was suspended in distilled water and stored at -20 ℃ until use. Amplified PCR products were sequenced directly on the Perkin-Elmer Applied Biosystems (ABI) 377A DNA sequencer using a ABI PRISM Big Dye™ Terminator Ready Reaction Kit (Perkin Elmer) following the manufacture's protocol.

Sequence analysis

Sequence data were aligned using the multiple alignment program Clustal W [30] and were determined by parsimony, distances and maximum likelihood (ML) methods (http:// bioweb.pasteur.fr/intro-uk.html#phylo). To understand the possible genetic relationships, PHYLIP (Phylogenetic Inference Package) version 3.573c [10] was used in this study. PHYLIP was used dinoflagellates such as *Symbiodinium corculorum* L13717, *S. pilosum* X62650, *S. meandrinae* L13718, *S. microadriatic* M88521, *Gymnodinium beii* U41087, *G. simplex* U41086, *Symbiodinium* sp. M88509, *Gyrodinium impudicum* AF022197, *Lepidodinium viride* AF022199, *G. catenatum* AF022193, *Gymnodinium* sp. AF022196, *G. fuscum* AF022194, *Heterocapsa triquetra* AF022198, *G. sanguineum* U41085, *Scrippsiella nutricula* U52357, *Pentapharsodinium tyrrhenicum* AF022201, *Prorocentrum micans* M14649, *G. breve* AF172714, *G. mikimotoi* AF022195, *Gloeodinium viscum* L13716, *Peridinium* sp. AF022202, *Gyrodinium dorsum* AF274261, *G. uncatenum* AF274263, *Ceratococys horrida* AF022154, *Gonyaulax cochlea* AF274258, *Pyrodinium bahamense* AF274275, *Gyrodinium galatheanum* AF274262, *Scrippsiella trochoidea* AF274277, *Prorocentrum minimum* Y16238, *P. lima* Y16235, *Noctiluca scintillans* AF022200, *Alexandrium fundyense* U09048, *A. tamarense* X54946, *A. margaelefii* U27498, *A. minutum* U27499, *A. ostenfeldii* U27500, *Ceratium tenue* AF022192, *Pyrocystis*

noctiluca AF022156, *Amphidinium belauense* L13719, *Gonyaulax spinifera* AF022155 and *Crypthecodinium cohnii* M64245. This search for parsimony analysis was repeated several times from different random staring points using the stepwise addition option to make certain the most parsimonious tree was found. For distance analysis, subprogram DNADIST in PHYLIP was used to obtain a matrix of Kimura's two-parameter distance[19]. Distance matrix was analyzed by subprogram NEIGHBOR in PHYLIP with algorithms based on Saitou and Nei's neighbor-joining (NJ) method [21]. All nucelotide substituions were equally weighted and unordered alignment gaps were treated as missing information. As an indication of confidence in the branching order, a bootstrap analysis (100 replications) was completed for both distance and parsimony methods. Reliability of the tree was constructed using subprogram CONSENSE in PHYLIP after pairwise sequence distances were estimated by Kimura's two-parameter method, which attempts to correct observed dissimilarities for multiple substitutions in sequences evolving with a transition bias.

Results

The SSU rRNA gene sequence was determined for 41 species except for Korean *C. polykrikoides* and *G. aureolum*, which were retrieved from the GenBank database (Fig. 1). A Kimura two-parameter NJ analysis and a substitution rate calibration analysis each produced phylogenetic trees, and the substitution rate calibration tree was presented in Fig. 2. Three groups were inferred. Group I contained three different taxa (*Noctiluca*, *Gymnodinium*, and *Symbiodinium*). *Notiluca scintillans* played the role of a sister to Group I. Several species of *Symbiodinium* had a monophyletic relationship. Group II consisted of various genera and was trisected with Group II-1 (*Prorocentrum*, *Gyrodinium*, *Scrippsiella*, *Pyrodinium*, *Gonyaulax*, and *Ceratocorys*), Group II-2 (*Peridinium*, *Gloeodinium*, and *Gymnodinium*), and Group II-3 (*Heterocapsa*, *Cochlodinium*, *Lepidodinium*, and *Pentapharsodinium*). Group II was fragmented with *P. minimum*/*P. lima* as a sister to Group II-1 (bootstrap 74%), *P. micans* as a sister to Group II-2 (bootstrap 92%) and *P. tyrrhemicum* as a sister to Group II-3 (bootstrap 58%). In particular, *S. nutricula* and *G. sanguineum* had an independent subgroup, which was supported by a weak bootstrap value of 51%. The phylogenetic placement of Korean *G. aureolum* was closer to *G. mikimotoi* and *G. breve* than that of *C. polykrikoides*, which joined *Gymnodinium* sp. The position of *C. polykrikoides* located the branch

with *G. impudicum*, and *G. catenatum*. Group III contained six taxa (*Ceratim*, *Pyrocystis*, *Alexandrium*, *Ampidinium*, *Gonyaulax*, and *Crypthecodinium*). *Alexandrium* formed the most basal clade, a sister to the remaining taxa and had a monophyletic relationship. However, Group I and II were supported by a moderate bootstrap value of 72%, while Group III was supported by a very weak bootstrap value of 51%. The PHYLIP NJ and ML analyses produced a tree that was similar to the substitution rate calibration tree, and all groups were recovered (Figs. 3, 4). The essential aspects of topology were retained. For example, the tree members of each group according to parsimony and NJ analyses formed their own groups, exactly distinctive monophyletic groups. The unrooted phylogenetic tree obtained from ML PHYLIP was slightly changed which tree members were inserted or not. The essential aspects of topology were that *G. aureolum* formed an extremely strong monophyletic group with *G. mikimotoi* and *G. breve* (bootstrap >90%) and *C. polykrikoides* appeared to be affected according to the topology of the trees. The DNA similarity between taxa was shown in Table 1. Among the species used in this study, the DNA similarity was high, over 70%. *Cochlodinium polykrikoides* had a high similarity (88-90%) to the genus *Gymnodinium* including *G. aureolum* and *G. impudicum*. However, *G. aureolum* was higher value of similarity between taxa than did *C. polykrikoides*. *Gonyaulax cochllea* had the lowest value of similarity among all tested species. To further illustrate the genetic relationships between the genera *Gymnodinium* and *Gyrodinium* including *C. polykrikoides*, we constructed a matrix of Kimura genetic distances between pairs of taxa, as shown in Table 2. The sequence divergence in pairwise comparisons ranged from 0.0% to 9.5% (0 to 58 bp) and the largest sequence divergence occurred in a comparison of *G. fuscum* and *C. polykrikoides*. The sequence divergence between *G. aureolum* and *G. dorsum*/*G. galathenum* was the largest at 7.4% (45 bp), while *G. aureolum* and *G. mikimotoi* showed an extremely low value of genetic divergence of 0.9% (5 bp). The genetic divergence between *C. polykrikoides* and *G. aureolum* was a low value of 5.2% (31 bp), compared with other taxa. The genetic relationships between the genera *Gymnodinium* and *Gyrodinium* including *C. polykrikoides* sequences were shown in PHYLIP and were in Fig. 5. In the pairwise analysis, *G. aureolum* and *C. polykrikoides* consisted of a member of Group I and of Group II, respectively; the two groups were supported by a moderate bootstrap value of 84% (Fig. 5). Group III was supported by a weak bootstrap value of 62%.

```
1   -AACCTGGTTGATCCTGCCAGTAGTCATAT-GCTTGTCTCAAAGATTAAGCCATGCATGT
21  ------GGTTGATCCTGCCAGTAGTCATAT-GCTTGTCTCAAAGATTAAGCCATGCATGT
7   -AACCTGGTTGATCCTGCCAGTAGTCATAT-GCTTGTCTCAAAGATTAAGCCATGCATGT
9   CAACCTGGT-GATCCTGCCAGTAGTCATAT-GCTTGTCTCAAAGATTAAGCCATGCATGT
10  -AACCTGGTTGATCCTGCCAGTAGTCATAT-GCTTGTCTCAAAGATTAAGCCATGCATGT
12  -AACCTGGTTGATCCTGCCAGTAGTCATAT-GCTTGTCTCAAAGATTAAGCCATGCATGT
13  -AACCTGGTTGATCCTGCCAGTAGTCATAT-GCTTGTCTCAAAGATTAAGCCATGCATGT
11  -AACCTGGTTGATCCTGCCAGTAGTCATAT-GCTTGTCTCAAAGATTAAGCCATGCATGT
19  -CACCTGGTTGATCCTGCCAGTAGTCATAT-GCTTGTCTCAAAGATTAAGCCATGCATGT
33  ------GGTTGATCCTGCCAGTAGTCATATTGCTTGTCTCAAAGATTAAGCCATGCATGT
14  --ACCTGGTTGATCCTGCCAGTAGTCATAT-GCTTGTCTCAAAGATTAAGCCATGCATGT
18  --ACCTGGTTGATCCTGCCAGTAGTCATAT-GCTTGTCTCAAAGATTAAGCCATGCATGT
32  -CACCTGGTTGATCCTGCCAGTAGTCATAT-GCTTGTCTCAAAGATTAAGCCATGCATGT
22  -CACCTGGTTGATCCTGCCAGTAGTCATAT-GCTTGTCTCAAAGATTAAGCCATGCATGT
29  -CACCTGGTTGATCCTGCCAGTAGTCATAT-GCTTGTCTCAAAGATTAAGCCATGCATGT
27  -CACCTGGTTGATCCTGCCAGTAGTCATAT-GCTTGTCTCAAAGATTAAGCCATGCATGT
26  -CACCTGGTTGATCCTGCCAGTAGTCATAT-GCTTGTCTCAAAGATTAAGCCATGCATGT
24  -CACCTGGTTGATCCTGCCAGTAGTCATAT-GCTTGTCTCAAAGATTAAGCCATGCATGT
28  -CACCTGGTTGATCCTGCCAGTAGTCATAT-GCTTGTCTCAAAGATTAAGCCATGCATGT
35  --ATCTGGTTGATCCTGCCAGTAGTCATAT-GCTTGTCTCAAAGATTAAGCCATGCATGT
23  -CACCTGGTTGATCCTGCCAGTAGTCATAT-GCTTGTCTCAAAGATTAAGCCATGCATGT
25  -CACCTGGTTGATCCTGCCAGTAGTCATAT-GCTTGTCTCAAAGATTAAGCCATGCATGT
34  -AACCTGGTTGATCCTGCCAGTAGTCATAT-GCTTGTCTCAAAGATTAAGCCATGCATGT
4   -AACCTGGTTGATCCTGCCAGTAGTCATAT-GCTTGTCTCAAAGATTAAGCCATGCATGT
6   -AACCTGGTTGATCCTGCCAGTAGTCATAT-GCTTGTCTCAAAGATTAAGCCATGCATGT
8   --ACCTGGTTGATCCTGCCAGTAGTCATAT-GCTTGTCTCAAAGATTAAGCCATGCATGT
5   -AACCTGGTTGATCCTGCCAGTAGTCATAT-GCTTGTCTCAAAGATTAAGCCATGCATGT
3   ------GGTTGATCCTGCCAGTAGTCATAT-GCTTGTCTCAAAGATTAAGCCATGCATGT
17  --ACCTGGTTGATCCTGCCAGTAGTCATAT-GCTTGTCTCAAAGATTAAGCCATGCATGT
16  --ACCTGGTTGATCCTGCCAGTAGTCATAT-GCTTGTCTCAAAGATTAAGCCATGCATGT
2   ------GGTTGATCCTGCCAGTAGTCATAT-GCTTGTCTCAAAGATTAAGCCATGCATGT
30  -CACCTGGTTGATCCTGCCAGTAGTCATAT-GCTTGTCTCAAAGATTAAGCCATGCATGT
20  ----------------------------------------AAGATTAAGCC-TGCATGT
43  ------------------AGTCATAT-GCTTGTCTCAAAGATTAAGCCATGCATGT
39  ------------------AGTCATAT-GCTTGTCTCAAAGATTAAGCCATGCATGT
40  ------------------AGTCATAT-GCTTGTCTCAAAGATTAAGCCATGCATGT
42  -------------------AGTCATAT-GCTTGTCTCAAAGATTAAGCCATGCATGT
38  --------------------AGTCATAT-GCTTGTCTCAAAGATTAAGCCATGCATGT
41  ---------------------AGTCATAT-GCTTGTCTCAAAGATTAAGCCATGCATGT
36  ------------TCCTGCCAGTAGTCATAT-GCTTGTCTCAAAGATTAAGCCATGCATGT
37  ------------TCCTGCCAGTAGTCATAT-GCTTGTCTCAAAGATTAAGCCATGCATGT
31  -CACCTGGTTGATCCTGCCAGTAGTCATAT-GCTTGCCTCAAAGATTAAGCCATGCATGT
15  -CACCTGGTTGATCCTGCCAGTAGTCATAT-GCTTGTCTCAAAGATTAAGCCATGCGCT
                              **********  *******

1   CTGAGTCTATGCCT-TTACACGGTTAAGCTGCGAATGGCTCATT-ATAACAGTAATGATC
21  CTCAGAATAAGCTT-TT-CATGGCAAGTCTGCGAATGGCTCATT-AAACCAGTTATAGTT
7   CTCAGCATAAGCGT-TT-CACAGCAAGGCTGCAAATGGCTCATT-AAAACAGCAATGATT
9   CTCAGTATAAGCTT-TCATTTTGTGAAACTGCAAATGGCTCATT-AAAACAGTTATAATG
10  CTTAGTATAAGCTT-CAATTTTGTGAAACTGCAAATGGCTCATT-AAAACAGTCACAAYG
12  CTCAGTATAAGCAT-TCATTTGGCGAAACTGCAAATGGCTCATT-AAAACAGTTATAATG
13  CTCAGTATAAGCAT-TCATTTGGCAAAACTGCAAATGGCTCATT-AAAACAGTTATAATG
11  CTCAGTATAAGCAT-TTATTTGGCAAAACTGCAAATGGCTCATT-AAAACAGTTATAATG
19  CTGAGTTTAAGCTA-ATTTTTGGTGAAGCTGCGAATGGCTCATT-ATATCAGTTATAATA
33  CTCAGCTTAAGCTT-ACCTACAGTGAAGCTGCAAATGGCTCATT-AAAACAGTTATAAAT
14  CTCAGTATAAGCTT-CTATATGGCAAAACTGCGAATGGCTCATT-AAAACAGTTATAGTT
18  CTCAGTATAAGCTT-CTACACGGCAAAACTGCGAATGGCTCATT-AAAACAGTTATAGTT
32  CTCAGTATAAGCTT-CTATACGGCAAAACTGCGAATGGCTCATT-AAAACAGTTATAGTT
22  CTCAGTATAAGCTT-CTACACGGCAAAACTGCGAATGGCTCATT-AAAACAGTTATCGTT
29  CTCAGTATAAGCTT-CTACACGGCGAAACTGCGAATGGCTCATT-AAAACAGTTATCGTT
27  CTCAGTATAAGCTT-TCACACGGCGAAACTGCGAATGGCTCATT-AAAACAGTTATCGTT
26  CTCAGTATAAGCTT-CTACACGGCGAAACTGCGAATGGCTCATT-AAAACAGTTATCGTT
24  CTCAGTATAAGCTT-CTACACGGCGAAACTGCGAATGGCTCATT-AAAACAGTTATCGTT
28  CTCAGTATAAGCTT-TTATACGGCGAAACTGCGAATGGCTCATT-AAAACAGTTATAGTT
35  CTCAGTATAAGCTTTCTATACGGCGAAACTGCGAATGGCTCATT-AAAACAGTTATAGTT
23  CTCAGTATAAGCTT-CTATACGGCGAAACTGCGAATGGCTCATT-AAAACAGTTATAGTT
25  CTCAGTATAAGCTT-CTATACGGCGAAACTGCGAATGGCTCATT-AAAACAGTTATAGTT
34  CTCAGTATAAGCTT-CTATACGGCGAAACTGCGAATGGCTCATT-AAAACAGTTATAGTT
4   CTCAGTATAAGCTT-CTGTACGGCAAAACTGCGAATGGCTCATT-AAAACAGTTACCGTT
6   CTCAGCATAAGCTT-CTACACGGCGAAACTGCGAATGGCTCATT-AAAGCAGTTATAATT
8   CTCAGTATAAGCTT-CTACACGGCGAAACTGCGAATGGCTCATT-AAAGCAGTTATAATT

5   CTCAGTATAAGCTT-CTACACGGCGAAACTGCGAATGGCTCATT-AAAGCAGTTATAATT
3   CTCAGTATAAGCTT-TTACACGGCGAAACTGCGAATGGCTCATT-AAAGCAGTTATAATT
17  CTCAGTATAAGCTT-CTACATGGTGAAACTGCGAATGGCTCATT-AAAGCAGTTATAGTT
16  CTCAGTATAAGCTT-TCATACGGTGAAACTGCGAATGGCTCATT-AAAGCAGTTATAGTT
2   CTCAGTATAAGCAG-TCAAATGGTGAAACTGCGAATGGCTCATT-AAAGCAGTTATAGTT
30  CTCAGTATCAACTT-TTACACGGTAACACTGCGAATGGCTCATT-AAAACAGTTATAGTT
20  CTCAGTATAAAT-T-CAATATGGTTAAACTGCGAATGGCTCATTCAA-ACAGTTATAGTT
43  CTCAGAATATGCTC-TGATG-AACTAATCTGCAAATGGCTCATT-ATAACAGTT----TT
39  CTCAGTATAAGTGT-TTGGATAGTGAAACTGCGAATGGCTCATT-AAAACAGTTATAATT
40  CTCAGTATAAGCCT-CTACATGACGAAACTGCGAATGGCTCATT-AAAACAGTTATAGTT
42  CTCAGTATAAGCCT-CTACATGACGAAACTGCGAATGGCTCATT-AAAACAGTTATAGTT
38  CTCAGTATAAGCTT-CTATACGGTGAAACTGCGAATGGCTCATT-AAAACAGTTACAGTT
41  CTCAGTATAAGCTT-CTATACGGCGAAACTGCGAATGGCTCATT-AAAACAGTTATAGTT
36  CTCAGCATAAGCTT-CCATCCGGCGAAACTGCGAATGGCTCATT-AAAACAGTTACAATT
37  CTCAGTATAAGCTT-CTATACGGCGAAACTGCGAATGGCTCATT-AAAACAGTTATAGTT
31  CTCAGTACAAACTT-GTATATGGTGAAGCTGCGAATGGCTCATT-AAATCAGTTACAGCT
15  CTCAGTATAAGCTT-CTACACGGCGAAACTGCGAATGGCTCATT-AAAACAGTTATCGTT
     ** **          *  ****  *********** *   ***

1   TATTTGATGATTATCG--TTACATGG-ATAACTACAGTAATTCTGGAGCCAATACATGTT
21  TCATTGGTTGTTGGTGC-TTACATGG-ATACCTGTGGTAATTCTATAGCTAATACATGAT
7   TTCGCAGTGCTTCATC--ACACATGG-ATAACTGTGGAAAATCTAGGGCTAATACATG-C
9   TATTTGACAATCAACT--CTATATGG-ATATCTGTGGAAATTCTATAGTTAATACATGCA
10  CATTTGGCGATCAATT--CTAAATGG-ATATCTGTGGRAATTCTATAGTTAATACATGCA
12  CACTTGAYGATCGATTGCTTACATGG-ATAACTGTGGTAATTCTATAGCTAATACATGCA
13  CACTTGATGGTCGATTGCT-ACATGG-ATAACTGTGGTAATTCTATAGCTAATACATGCA
11  TGTTTGATGGTTGATTGCTTRCATGG-ATAACTTTGGTAATTCTACAGCTAATGCATGCA
19  TGCTCAGTCACTT-TTGCTTACATGG-ATAACTGTGGTAATTCTAGAGCTAATACATGCA
33  CATATGTTGGTTG-CAGTCTACATGG-ATAACCTCGGTAATTCTGCAGCTAATTCATGCC
14  CATTTGATGGT-CTCTCACTACATGG-ATAACTGTGGTAATTCTAGAGCTAATACATGCG
18  TATTTGATGGT-CAATCCTTACATGG-ATAACTGTGGTAATTCTAGAGCTAATACATGCG
32  TATTTGATGGT-CATTCTTTACATGG-ATAACCGTGGTAATTCTAGAGCTAATACATGCG
22  TATTTGGTGGT-CATTCAGTACATGG-ATAACCATGGTAATTCTAGAGCTAATACATGCG
29  TATTTGGTGGT-CATTCATTACATGG-ATAACCGTGGTAATTCTAGAGCTAATACATGCG
27  TATTTGGTGGT-CATTCATTACATGG-ATAACCATGGTAATTCTAGAGCTAATACATGCG
26  TATTTGGTGGT-CATTCATTACATGG-ATAACCATGGTAATTCTAGAGCTAATACATGCG
24  CATTTGGTGGT-CATTTCCTACATGG-ATAACCGTGGTAATTCTAGAGCTAATACATGCG
28  TATTTGATGGT-CATTC-TTACATGG-ATAACCGTGGTAATTCTAGAGCTAATACATGCG
35  TATTTGATGGT-CACTCTTTACATGG-ATAACTGTGCTAATTGTAGAGCTAATACATGCG
23  TATTTGATGGT-CATTCTTTACATGG-ATAACTTTGGAAATTCTAGAGCTAATACATGCG
25  TATTTGATGGT-CATTCTTTACATGG-ATAACTTTGGAAATTCTAGAGCTAATACATGCG
34  TATTTGATGGT-CATTCATTACATGG-ATAACTTTGGAAATTCTAGAGCTAATACATGCG
4   TATTTGATGGT-CATTC-TGACACGG-ATAACCGTGGTAATTCTAGAGCTAATACGCGCT
6   TATTTGATGGT--CACTGCTACATGG-ATAACTGTGGTAATTCTAGAGCTAATACATGCA
8   TATTTGATGGT--CACTGCTACATGG-ATAACTGTGGTAATTCTAGAGCTAATACATGCA
5   TATT-GATGGT--CACTGTCACATGG-ATAACTGTGGTAATTCTAGAGCTAATACATGCA
3   TATTTGATGGT--CACTGCTACATGG-ATAACTGTGGTAATTCTAGAGCTAATACATGCA
17  TATTTGATGGT--CACTGTTATATGG-ATAACTGTGGTAATTCTAGAGCTAATACATGCA
16  TATTTGATGGT--CACTGTTACATGG-ATAACTGTGGTAATTCTAGAGCTAATACATGCA
2   TATTTGATGGT--TGCTGCTACATGG-ATAACTGTGGTAATTCTAGAGCTAATACATGCA
30  TATTTGATGGC--ATTCATTATATGG-ATACCTTTGGTAATTCTAGAGCTAATACATGCA
20  CACTTGATGGTGTGTGTTCAGATCC--ATAACTGTGGTAATTCTAGCAGTAATACATGCT
43  GATCTTGTGATTCGTTTTCTGAATTGGATATCTTGGGTAATTCTAAAGCTAATACGTGCT
39  TATTTGGTGGT-CATTCTCTACATGG-ATAACTGTGGTAATTCTATAGCTAATACATGCT
40  TATTTGATGGT-CGCTTAC-ACATGG-ATAACTGTGGTAATTCTAGAGCTAATACATGCA
42  TATTTGATGGT-CGCTTACTACATGG-ATAACTGTGGTAATTCTAGAGCTAATACATGCA
38  TATTTGATGGT-CATTCCTTACATGG-ATAACTGTGGTAATTCTAGAGCTAATACATGCG
41  TATTTGATGGT-CATTCATTACATGG-ATAACTGTGGTAATTCTAGAGCTAATACATGCG
36  TATTTGGTGGTTCACTCTT-ACATGG-ATACCCGTGGAAATGCTAGAGCTAATACATGCG
37  TATTTGATGGTCATTCTTT-ACATGG-ATAACTGTGCTAATTGTAGAGCTAATACATGCG
31  TTTTTGATGAT-GATTCTTTACATGG-ATAGCCATGGTAATTCTGCTGCTAATACATGTG
15  TATTTGGTGGT-CATTCATTACATGG-ATAACCATGGTAATTCTAGAGCTAATACATGCG
                  *** *     ** *      *** * *

1   TCAACACCCAACTTATTGGA---G-GGTGGTGCTTACCAA-ATTCAGAACCAAT-CCAAG
21  TGCATACCCGACTTTCCGGAA--G-GGTTGTGCTTATCAG-CTCCAGAACCATT-CCTGG
7   GACATACCTGACTTTTGGGGA--ATGGTTGCTATCATGAG-CCACAGAATCGGC-GCAGG
9   TTGAAACCGATCTTTGGGGA---AAGGTTGTGGTCGTTAG-TTACAGAACCAAT-TCAAG
10  CTAAAACTTATCTTTGGGGA---AAGGTTGTGGTCGTTAG-TTACAGAACCAAT-TCAAG
12  TCTAAACCTGACTTTTGG-A---AGGGTTGTGGTCATTAG-TTACAGAACCAAC-CCAAG
13  TCCAAACCTGACTATTGGGA---AGGGTTGTGGTCATTAG-TTACAGAACCAAC-CCAGG
```

```
11  ACCAAGCCKGACTTGTGGGA---AGGGCTGAAGTCATTAG-TTACAGAACCAAC-CCAAG
19  CACAAACCTGACCTCAAAGT---ATGGTTGTGTTTATTGT-TTGCAAAACCAAT-TCAAG
33  ACCAAACTTGACTTGTGGAAGATAAAGTTGTGGTTCTTAGCTTTTAGAATCAAC-CCAAG
14  CCCTAACCCGACTTCGTAGA---AGGGTTGTGCTTATTAG-ACACAGAACCAAC-CCAGG
18  CCCAAACCCGACTCCGTGGA---AGGGTTGTGCTTATTAG-TTACAGAACCAAC-CCAGG
32  CCCAAACCCGACTCCGTGGA---AGGGTTGTGTTTATTAG-WTACAGAACCAAC-CCAGG
22  CCCAGACCCAACTTTGTGGA---AGGGTTGTGTTTATTAG-GTACAGAACCAAC-CCAGG
29  TCCAGACCCAACTTCGTGGA---AGGGTTGTGTTTATTAG-GTACAGAACCAAC-CCAGG
27  CCCAGACCCAACTTCGTGGA---AGGGTTGTGTTTATTAG-GTACAGAACCAAC-CCAGG
26  CCCAAACCCAACTTCTTGGA---CGGGTTGTGTTTATTAG-GTACAGAACCAAC-CCAGG
24  CCCATACCCAACTTCTTGGA---AGGGTTGTGTTTATTAG-GTACAGAACCAAC-CCAGG
28  CCAAAACCCGACTTCTTGGA---AGGGTTGTGTTTATTAG-TTACAAAACCAAC-CCAGG
35  CCCAAACCCGACTTCGAGGA---AGGGTTGTGTTTATTAG-TTACAGAACCAAC-CCAGG
23  CCCAAACC-GACTTCTTGGA---AGGGTTGTGTTTATTAG-ATACAGAACCAAC-CCAGG
25  CCCAAACC-GACTTCTTGGA---AGGGTTGTGTTTATTAG-ATACAGAACCAAC-CCAGG
34  CCCAAACCCGACTTCTTGGA---AGGGTTGTGTTTATTAG-ATACAGAACCAAC-CCAGG
4   GCTAAACCCAACTTCGAGGA---AGGGTTGTATTTATTAG-GTACAGAACCAAC-GCAGG
6   CCAAAACCCAACTTCGCAGA---AGGGTTGTATTTATTAG-ATACAGAACCATC-GCAGG
8   CCAAAACCCAACTTCGCAGA---AGGGTTGTATTTATTAG-ATACAGAACCATC-GCAGG
5   CCAAAACCCAACTTCGCAGA---AGGGTTGTATTTATTAG-ATACAGAACCATC-GCAGG
3   CCAAAACCCAACTTCGCAGA---AGGGTTGTATTTATTAG-ATACAGAACCAAC-GCAGG
17  CCAAAACCCAACTTCGCAGA---AGGGTTGTGTTTATTAG-ATACAGAACCAAC-CCAGG
16  CCAAAACCCGACTTCG--GA---AGGGTTGTGTTTATTAG-ATACAGAACCATC-ACAG-
2   TCCAAGCCCGACTTTGCAGA---AGGGTTGTGTTTATTAG-ATCCAGAACCAAC-CCAGG
30  TACAGACCCGACTCAT--GA---AGGGTTGTGTTTATTAG-ATTCCAAACCTTC-CAA--
20  -TCTTGACCCAACTTTGTGG----GGGTGTT--TTATTAG-TTACAAAACTAAC-CCAGA
43  GTAATGCCTCACTTCTGGGA---CGGGCATTGCTTGTTGG-CTTCAGAACCAATACCTGG
39  TACCAACCCAACTTCATGGA---AGGGTTGTGGTTATTAG-TTACAGAACCAAC-CCAAG
40  CAAATTCCCAACTTCGCGGA---AGGGATGTGTTTATTAG-ATACAGAACCAAC-ACAGG
42  CAAATTCCCAACTTCGCGGA---AGGGATGTGTTTATTAG-ATACAGAACCAAC-ACAGG
38  CCCAAACCCGACTCCGTGGA---AGGGTTGTGTTTATTAG-TTACAGAACCAAC-CCAGG
41  CCCAAACCCGACTTCGTGGA---AGGGTTGTGTTTATTAG-ATACAGAACCAAC-CCAGC
36  CTCCTACCCAACTTAACAGA---AGGGTTGTGGTTATTAG-TTACAGAACTAGC-CCAGG
37  CCCAAACCCGACTTATAGGA---AGGGTTGTGTTTATTAG-TTACAGAACCAGC-CCAGG
31  CACAAACCTGACTCTGTGGA---AGGGTTGTGTTTGTTAG-GAACAGAACCAAC-CTGGG
15  CCCAAACCCAACTTCTTGGA---CGGGTTGTGTTTATTAG-GTACAGAACCAAC-CCAGG
                          *        *        **
```

```
1   C-TTGCTTGGTT-TCATAATGAAGTATGGTAAGA-GGATGAATCACATGGTCTTT-CTGG
21  CATTGCCAGGTT-CTTTGGTGATCTATGGTAAGT-GTGCGGATCACATGGCTGCTGCTGG
7   CTCTGCCTTGCTGCTCGGGTGAGTTATGATGATT-TAGCTGATCGCACAGCACTTGCTGG
9   CATTGCTTGAATACTTGATTGATTCATAATGACA-AAATGAATTACATGGCAACAGCTGG
10  CTATGCTTGGACACTTGAATGATTCACAATGACA-AA-TGAATTACATGGCAACAGCTGG
12  CTCTGCTTGGA-ATTTTGGTGATTCATGATGACT-GAATGAATTACATGGCAACTGCTGG
13  CTCTGCTTGGA-ACCTTGGTGATTCATAATGACT-GAATGAATTACATTGCAACCGCTGG
11  CTCTGCTTGGA-ATTTTGGTGATTCACATTGACT-GAATGAATTACATGGCATCTGCTGG
19  CCCTGCTTGATATGTTGCATGACTGATTATAAAT-GAATGAATTGCATGGCGCACGCTGG
33  TTCTGCTTGGTTACGTG-GTGATGCATAGGGACC-CAATGAATTGCATAGCATCTGCTGG
14  CTCTGCCTGGTCTTGTG-GTGAGTCATAATAACT-GAACGAATCGCATG-CATC-GCCGG
18  CTCTGCCTGGTTTTGTG-GTGATTCATAATAACT-CAACGAATCGCATGGCATCCGCCGG
32  CTCTGCCTGGTCTTGTG-GTGATTCATAATAACC-AAACGAATCGCATGGCATCAGCTGG
22  CTCTGCCTGGTCTTTTG-GTGATTCATGATAACC-GAACGAATCGCATGGCATCAGCTGG
29  CTCTGCCTGGTCTTCTG-GTGATTCATGATAACC-GAACGAATCGCATGGCATCAGCTGG
27  CTCTGCCTGGTCTTTTG-GTGATTCATAATAACC-GAACGAATCGCATGGCCTCGCCTAG
26  CTCCGCCTGGCTTTTTG-GTGATTCATGATAACC-GAACGAATCGCATGGCATTCGCTGG
24  CTCTGCCTGGTCTTCTG-GTGATTCATGATAACC-GAACGAATCGCATGGCATCCGCTGG
28  CTCCGCCTGGTCTTCTG-GTGATTCATAATAACC-GAACGAATCGCATGGCATCCGCTGG
35  CTCTGC-TGTCTTTGTG-GTGATTCATAATAACCAGAACGAATCGCATGGCATCCGCTGG
23  CTCTGCCTGGTCTCGTG-GTGATTCATAATAACC-GAACGAATCGCATTGCATCAGCTGG
25  CTCTGCCTGGTCTCGTG-GTGATTCATAATAACC-GAACGAATCGCATTGCATCAGCTGG
34  CCCTGCCGTGTCTTGTG-GTGATTCRTAATAACC-GAACGAATCGCATTGCATCAGCTGG
4   TTC-GCCTGATCTTGTG-GTGAGTCATGATAACT-CAACGAGTCGCATTTCATCAGTTGG
6   CTCTGCCTGGTT--GTG-GTGATTCATGATAACT-CGATGAATCGTGTGGC-TTGGCCGA
8   CTCTGCCTGGTT--GTG-GTGATTCATGATAACT-CGATGAATCGTGTGGC-TTGGCCGA
5   CTCTGCCTGGTT--GTG-GTGATTCATGATAACT-CGATGAATCGTGTGGC-TTGGCCGA
3   CTCCGCCTGGTT--GTG-GTGATTCATGATAACT-CGATGAATCGTGTGGC-TTGGCCGA
17  TTC-GCCTGGTTTTGTG-GTGACTCATGATAACT-CGATGAATCGTGTAGC-TTCGCTGA
16  CTCCGCTTGGTCTTGTG-GTGACTCATGATAACT-CGATGAATCGTGTGGC-TTCGCTGA
2   CTCCGCCTGGTCATGTG-GTGATTCATGATAACT-TGACGAATCGTGCGGC-CTTGCCGA
30  CTTCGGTTGTGTTCTTG-GTGATTCATGATAAAA-TACTGA-TCGC----C-TTTGGCGA
20  CTCTGTCTGGTATTGTG-TTGATTCATGATGTTT-GGATGGATTGCATGGCAGCTGCTGG
```

```
43  CTTTGTCAGGAATGGT--TTGAATTAAAACAA---GGATG--TGATCTAACA-TTGGTGG
39  CTTTGCTTGGTCTTGCG-GTGATTCATGATGACT-GAATGAATTGCATGGCA-TTGCTGG
40  CTCTGCCTGTTCTCATG-GTGACTCATAGTAACC-GAACGAATCGCATGGCATCTGCTGG
42  CTCTGCCTGTTCTCATG-GTGACTCATAGTAACC-GAACGAATCGCATGGCATCTGCTGG
38  CTCTGCCTGGTCTTGTG-GTGATTCATAATAACC-GAATGAATCGCATTGCATCAGCTGG
41  TTCGTGCTGGTATTGTG-GTGATTCATAATAACC-GAACGAATCGCATAGCTTCCGCTGG
36  CTT-GCCTGGTTACTCG-GTGACTCATGATAATG-GAATTAGTCGTATGGCGTCTGCTGA
37  CTCCGCCTGGTCCTTTG-GTGATTCATAATAACC-GAACGAATCGCATAGC-TTCGCTGG
31  CTGTGCCTGGCCATGAG-GTGATTCATAATACCT-GAACGAATCGTGCAGTGTCAGCTGG
15  CTCCGCCTGGCCGTGGT-TTGCCACGGGATTGTT--TTTTCCTCACGGGGGAGCTGTCCC
                          **                  *
```

```
1   TGATGCATCTCATGAGTTTCTGACCTATC--AGCT-TTGGATGGTAGGGTATTGGCCTAC
21  TGATTTGTCTTTTAAGTTTCTGACCTATC--AGCT-TTCGACGGGAAGGTATTGGCTTCC
7   TGATGAACTGCACAAGTTTCTGACCTATC--AGCT-TCCGACGCTAGGGTATTGGCCTAC
9   TAATAATTCATTCCAGTTTCTGACCTATC--AGCT-TTCGACGGTAAGGTATTGGCTTAC
10  TGATAATTCATTCCAGTTTCTGACCTATC--AGCT-TTCGACGGTAAGGTATTGGCTTAC
12  TAATAAATCATTCCAGTTTCTGACCTATC--AGCT-TCCGACGGTAAGGTATTGGCTTAC
13  TAATAAATCATTCCAGTTTCTGACCTATC--AGCT-TCCGACGGTAGGGTATTGGCTTAC
11  TAATATATCATTTCAGTTTCTGACCTATC--AGCT-TCCGACGGTAAGGTATTGGCTTAC
19  CAATGTATTGTTTTAGTTTCTGACCTATC--AGCT-TCCGACGGTAAGGTATTGGCTTAC
33  TGATGCAACATTCACGTTTCTGACCTATC--AGCT-TCCGACGGTAGGGTATTGGCTTAC
14  CGATAAATCATTCAAGTTTCTGACCTATC--AGCT-TCCGACGGTAGGGTATTGGCCTAC
18  CGATGAATCATTCAAGTTTCTGACCTATC--AGCT-TCCGACGGTAGGGTATTGGCCTAC
32  CGATGAATCATTCAAGTTTCTGACCTATC--AGCT-TCCGACGGTAGGGTATTGGCCTAC
22  CGATGAATCATTCAAGTTTCTGACCTATC--AGCT-TCCGACGGTAGGGTATTGGCCTAC
29  CGATGAATCATTCAAGTTTCTGACCTATC--AGCT-TCCGACGGTAGGGTATTGGCCTAC
27  CGATGGACTATTCAAGTTTCTGACCTATC--AGCT-TCCGACGGTAGGGTATTGGCCTAC
26  CGATGAACCATTCA-GTTTCTGACCTATC--AGCT-TCCGACGGTAGGGTATTGGCCTAC
24  CGATGTATCATTCAAGTTTCTGACCTATC--AGCT-TCCGACGGTAGGGTATTGGCCTAC
28  CGATGAATCATTCAAGTTTCTGACCTATC--AGCT-TCCGACGGTAGGGTATTGGCCTAC
35  CGATGAATCATTCAAGTTTCTGACCTATC--AGCT-TCCGACGGTAGGGTATTGGCCTAC
23  CGATAGATCATTCAAGTTTCTGACCTATC--AGTT-TCCGACGGTAGGGTATTGGCCTAC
25  CGATAGATCATTCAAGTTTCTGACCTATC--AGTT-TCCGACGGTAGGGTATTGGCCTAC
34  CGATAGATCATTCAAGTTTCTGACCTATC--AGTT-TCCGACGGTAGGGTATTGGCCTAC
4   CGATGAATGATTCAAGTTTCTGACCTATC--AGCT-TCCGACGGTAGGGTATTGGCCTAC
6   CGATGCATCTTTCAAGTTTCTGACCTATC--AGCT-TCCGACGGTAGGGTATGGGCCTAC
8   CGATGCATCTTTCAAGTTTCTGACCTATC--AGCT-TCCGACGGTAGGGTATGGGCCTAC
5   CGATGCATCTTTCAAGTTTCTGACCTATC--AGCT-TCCGACGGTAGGGTATGGGCCTAC
3   CGATGCATCTTTCAAGTTTCTGACCTATC--AGCT-TCCGACGGTAGGGTATGGGCCTAC
17  CGATGCATCATTCAAGTTTCTGACCTATC--AGCT-TCCGACGGTAGGGTATTGGCCTAC
16  CGATGCATCATTCAAGTTTCTGACCTATC--AGCT-TCCGACGGTAGGGTATTGGCCTAC
2   CGATGCGTCATTCAAGTTTCTGACCTATC--AGCT-TCCGACGGTAGGGTATTGGCCTAC
30  CGA---GTCTTTCAAGTTTCTGACCTATC--AGCT-TTCGATGGTAGGGTATTGGCCTAC
20  TAACACATCATTCAAGTTTCTCACCTATC--AGCT-TCCGACGGTAAG-TATTGGCATAC
43  --ACTTGTTTTTTGAGTTTCTGACCTATC--AGCT-GTGGACAGTAAGATATTTGCTTCC
39  CGATATATCATTCAAGTTTCTGACCTATC--AGCT-TCCGACGGTAGGGTATTGGCTTAC
40  CGATGAATCATCCGAGCTTCTGACCTATC--AGCT-TCCGACGGTAGGGTATTGGCCTAC
42  CGATGAATCATCCGAGCTTCTGACCTATC--AGCT-TCCGACGGTAGGGTATTGGCCTAC
38  CGATGAATCATTCAAGTTTCTGACCTATC--AGCT-TCCGACGGTAGTGTATTGGACTAC
41  CGATAAATCATTCAAGTTTCTGACCTATC--AGCT-TCCGACGGTAGGGTATTGGCCTAC
36  CGATAAACCATTCAAGCTTCTGACCTATC--AGCT-TCCGACGGTAGGGTATTGGCCTAC
37  CGATGAATCATTCAAGTTTCTGACCTATC--AGCT-TCCGACGGTAGGGTATTGGCCTAC
31  CGATGGGTCATTCAGGTTTCTGACCTATC--AGCT-TCCGTCGGTAGGGTATTGGCCTAC
15  GGCAGAAACTTGAAAGTCCATCCCTTGGCCAAGCCATCCATCCCCCCCGGGTTTATGAAT
        *    *    *    *   **                    *
```

```
1   CATGGCAATGACGGGT-AACGGAGAATTAGGGTTCGATTCCGGAGA----GGGAGCCTGA
21  CGTGGCAATGACGGGT-AACGGAGAATTAGGGTTTGATTCCGGAGA----GGGAGCCTGA
7   TGTGGCGTTGACGGGT-GACGGAGAATTAGGGTTCGATTCCGGAGA----GGGAGCCTGA
9   CGTGGCAATGACAGGT-GACGGAGGATTAGGGTTTGATTCCGGAGA----GGGAGCTTGA
10  CGTGGCAATGACAGGT-GACGGAGGATTAGGGTTTGATTCCGGAGA----GGGAGCTTGA
12  CGTGGCAATGACGGGT-GACGGAGGATTAGGGTTTGATTCCGGAGA----GGGAGCTTGA
13  CGTGGCAATGACGGGT-GACGGAGGATTAGGGTTTGATTCCGGAGA----GGGAGCTTGA
11  CGTGGCAATGACGGGT-GACGGAGGATTAGGGTTTGATTCCGGAGA----GGGAGCTTGA
19  CGTGGCAATGACGGGT-GACGGAGAATTAGGGTTTGATTCCGGAGA----GGGAGCCTGA
33  CGTGGCATTGGCGGGT-AACGGGGGGATTAGGGTTTGATTCCGGAGA----GGGAGCCCGG
14  CGTGGCAATGACGGGT-AACGGAGAATTAGGGTTCGATTCCGGAGA----GGGAGCCTGA
18  CGTGGCAATGACGGGT-AACGGAGAATTAGGGTTCGATTCCGGAGA----GGGAGCCTGA
32  CGTGGCAATGACGGGT-AACGGAGAATTAGGGTTCGATTCCGGAGA----GGGAGCCTGA
22  CGTGGCAATGACGGGT-AACGGAGAATTAGGGTTCGATTCCGGAGA----GGGAGCCTGA
```

29	CGTGGCAATGACGGGT-AACGGAGAATTAGGGTTCGATTCCGGAGA----GGGAGCCTGA
27	CGTGGCAATGACGGGT-AACAGAGAATTAGGGTTCGATTCTGGAGA----GGGAGCCTGA
26	CGTGGCAATGACGGGTTAACGGAGAATTAGGGTTCGATTCCGGAGA----GGGAGCCTGA
24	CGTGGCAATGACGGGT-AACGGAGAATTAGGGTTCGATTCGGAGA----GGGAGCCTGA
28	CGTGGCAATGACGGGT-AACGGAGAATTAGGGTTCGATTCCGGAGA----GGGAGCCTGA
35	CGTGGCAATGACGGGT-AACGGAGAATTAGGGTTCGATTCCGGAGA----GGGAGCCTGA
23	CGTGGCATTGACGGGT-AACGGAGAATTAGGGTTCGATTCCGGAGA----GGGAGCCTGA
25	CGTGGCATTGACGGGT-AACGGAGAATTAGGGTTCGATTCCGGAGA----GGGAGCCTGA
34	CGTGGCATTGACGGGT-AACGGAGAATTAGGGTTCGATTCCGGAGA----GGGAGCCTGA
4	CGTGGCAATGACGGGT-AACGGAGAATTAGGGTTTGATTCCGGAGA----GGGAGCCTGA
6	CGTGGCAATGACGGGT-AACGGAGAATTAGGGTTTGATTCCGGAGA----GGGAGCCTGA
8	CGTGGCAATGACGGGT-AACGGAGAATTAGGGTTTGATTCCGGAGA----GGGAGCCTGA
5	CGTGGCAATGACGGGT-AACGGAGAATTAGGGTTTGATTCCGGAGA----GGGAGCCTGA
3	CGTGGCAATGACGGGT-AACGGAGAATTAGGGTTTGATTCCGGAGA----GGGAGCCTGA
17	CGTGGCAATGACGGGT-AACGGAGAATTAGGGTTTGATTCCGGAGA----GGGAGCCTGA
16	CGTGGCAATGACGGGT-AACGGAGAATTAGGGTTTGATTCCGGAGA----GGGAGCCTGA
2	CGTGGCAATGACGGGT-AACGGAGAATTAGGGTTTGATTCCGGAGA----GGGAGCCTGA
30	CATGGCAGTGACGGGT-AACAGAGAATTAGGGTTCGATTCTGGAGA----GGGAGCCTGA
20	CGTGGCAATGACGGGTTAACGGAGAATTAGGGTTCGATTCCGGAGA----GGGAGCCTGA
43	TGTGGCTATGACGGGT-GACGGGGGATTAGGGTTTGATTCCGGAGA----GGGAGCTTGA
39	CGTGGCAATGACGGGT-AACGGAGGATTAGGGTTTGATTCCGGAGA----GGGAGCCTGA
40	CGTGGCAATGACGGGT-AACGGAGAATTAGGGTTTGATTCCGGAGA----GGGAGCCTGA
42	CGTGGCAATGACGGGT-AACGGAGAATTAGGGTTTGATTCCGGAGA----GGGAGCCTGA
38	CGTGGCAATGACGGGT-AACGGAGAATTAGGGTTCGATTCCGGAGA----GGGAGCCTGA
41	CGTGGCAATGACGGGT-AACGGAGAATTAGGGTTCGATTCCGGAGA----GGGAGCCTGA
36	CGCGGCTTTGACGGGT-AACGGAGAATTAGGGTTTGATTCCGGAGA----GGGAGCCTGA
37	CGTGGCAATGACGGGT-AACGGAGAATTAGGGTTCGATTCCGGAGA----GGGAGCCTGA
31	TGTGGCAATGACGGGT-AACGGAGAATTAGGGTTCGATTCCGGAGA----GGGCGCCTGA
15	CCCAAAAACACGCAGC--ACCCTGAATTGGGCTCCCTATAAAAAAACCTCAACAGCCCCC
	* ** * *** ** * * * * **

1	GAAATGGCTACCACATCTAAGGAAGGCAGCAGGCGCGCAAATTACCCAATC-TTGACACA
21	GGAATGGCTACCACATCTAAGGAAGGCAGCAGGCGCGCAAATTACCCAATC-TTGATACG
7	CACATGGCTACCACATCTAAGGAAGGCAGCAGGCGCGCAAATTACCCAATC-CTGACACG
9	GAAATGGCTACCACATCTAAGGAAGGCAGCAGGCGCGCAAATTACCCAATC-CTGGCACA
10	GAAATGGCTACCACATCTAAGGAAGGCAGCAGGCGCGCAAATTACCCAATC-CTGGCACA
12	GAAATGGCTACCACATCTAAGGAAGGCAGCAGGCGCGCAAATTACCCAATC-CTGACACA
13	GAAATGGCTACCACATCTAAGGAAGGCAGCAGGCGCGCAAATTACCCAATC-CTGACACA
11	GAAATGGCTACCACATCTAAGGAAGGCAGCAGGCGCGCAAATTACCCAATC-CTGACACA
19	GAAATGGCTACCACATCTAAGGAAGGCAGCAGGCGCGCAAATTACCCAATC-TTGGCATA
33	GCAACGGCTACCACATCTAAGGAAGGCAGCAGGCGCGCAAATTACCCAATC-CTGACACA
14	GAAACGGCTACCACATCTAAGGAAGGCAGCAGGCGCGCAAATTACCCAATC-C-GACACG
18	GAAACGGCTACCACATCTAAGGAAGGCAGCAGGCGCGCAAATTACCCAATC-CTGACACA
32	GAAACGGCTACCACATCTAAGGAAGGCAGCAGGCGCGCAAATTACCCAATC-CTGACACA
22	GAAACGGCTACCACATCTAAGGAAGGCAGCAGGCGCGCAAATTACCCAATC-CTGACACA
29	GAAACGGCTACCACATCTAAGGAAGGCAGCAGGCGCGCAAATTACCCAATC-CTGACACA
27	GAAACGGCTACCACATCTAAGGAAGGCAGCAGGCGCGCAAATTACCCAATC-CTGACACA
26	GAAACGGCTACCACATCTAAGGAAGGCAGCAGGCGCGCAAATTACCCAATC-CTGACACA
24	GAGATGGCTACCACATCTAAGGAAGGCAGCAGGCGCGCAAATTACCCAATC-CTGACACA
28	GAAACGGCTACCACATCTAAGGAAGGCAGCAGGCGCGCAAATTACCCAATC-CTGACACA
35	GAAACGGCTACCACATCTAAGGAAGGCAGCAGGCGCGCAAATTACCCAATC-CTGACACA
23	GAAACGGCTACCACATCTAAGGAAGGCAGCAGGCGCGCAAATTACCCAATC-CTGACACA
25	GAAACGGCTACCACATCTAAGGAAGGCAGCAGGCGCGCAAATTACCCAATC-CTGACACA
34	GAAACGGCTACCACATCTAAGGAAGGCAGCAGGCGCGCAAATTACCCAATC-CTGACACA
4	GAAACGGCTACCACATCTAAGGAAGGCAGCAGGCGCGCAAATTACCCAATC-CTGGCACA
6	GAAACGGCTACCACATCTAAGGAAGGCAGCAGGCCCGCAAATTACCCAATC-CTGACACA
8	GAAACGGCTACCACATCTAAGGAAGGCAGCAGGCGCGCAAATTACCCAATC-CTGACACA
5	GAAACGGCTACCACATCTAAGGAAGGCAGCAGGCGCGCAAATTACCCAATC-CTGACACA
3	GAAACGGCTACCACATCTAAGGAAGGCAGCAGGCGCGCAAATTACCCAATC-CTGACACA
17	GAAACGGCTACCACATCTAAGGAAGGCAGCAGGCGCGCAAATTACCCAATC-CTGACACA
16	GAAACGGCTACCACATCTAAGGAAGGCAGCAGGCGCGCAAATTACCCAATC-CTGACACA
2	GAAACGGCTACCACATCTAAGGAAGGCAGCAGGCGCGCAAATTACCCAATC-CTGACACA
30	GAAACGGCTACCACTTCTAAGGAAGGCAGCAGGCGCGTAAATTACCCAATC-CTGACACA
20	GAAACGGCTACCACATCTAAGGAAGGCAGCAGGCGCGCAAATTACCCAATC-TTGACACA
43	GGATCGGCTACCACATCTAAGGAAGGCAGCAGGCGCGCAAATTACCCATTC-TTGACACA
39	GAAACGGCTACCACATCTAAGGAAGGCAGCAGGCGCGCAAATTACCCAATC-CTGACACT
40	GAAACGGCTACCACATCTAAGGAAGGCAGCAGGCGCGCAAATTACCCAATC-CTGACACA
42	GAAACGGCTACCACATCTAAGGAAGGCAGCAGGCGCGCAAATTACCCAATC-CTGACACA
38	GAAACGGCTACCACATCTAAGGAAGGCAGCAGGCGCGCAAATTACCCAATC-CTGACACA
41	GAAACGGCTACCACATCTAAGGAAGGCAGCAGGCGCGCAAATTACCCAATC-CTGACACA
36	GAAATAGCTACCACATCTAAGGAAGGCAGCAGGCGCGCAAATTACCCAATC-CTGACACA

37	GAAACGGCTACCACATCTAAGGAAGGCAGCAGGCGCGCAAATTACCCAATC-CTGACACA
31	GAAACGGCTACCACATCTAAGGAAGGCAGCAGGCGCGCAAATTACCCAATC-TTGAAACG
15	CACACAACCTCCACCCTTCCCATAAGTCCCAC-CTCCCACCTTCCCAAATCGCCCGCATC
	* **** * * * ** * * * ** ** * ** *

1	AGGAGGTAGTGACAAGAAATAACAATACAGGGCATCCATGTCTTGTAATTGGAATGAGCA
21	AGGAGGTAGTGACAAGAAATAACAATACAGGGCATCCATGTCTTGTAATTGGAATGGGCT
7	GGGAGGTAGTGACAAGCAATAACAATATGGGGCATCCATGTCTTGTAATTGGA-TGAACT
9	GGGAGGTAGTGACAAGAAATAACAATACAAGGCATCCATGTCTTGTAATTGGAATGAATG
10	GGGAGGTAGTGACAAGAAATAACAATACAAGGCATCCATGTCTTGTACTTGGAATGAATG
12	GGGAGGTAGTGACAAGAAATAACAATACAAGGCATCCATGTCTTGTACTTGGAATGAATG
13	GGGAGGTAGTGACAAGAAATAACAATACAAGGCATCCATGTCTTGTACTTGGAATGAATG
11	TGGAGGTAGTGACAAGAAATAACAATACAAGGCATCCATGTCTTGTACTTGGAATGAATG
19	GGGAGGTAGTGACAAGAAATAACAATACAGGGCATCCATGTTTTGTAATTGGAATGGACA
33	GGGAGGTAGTGACAAGATATAACAATACAAGGCATCCATGTTTTGTAATTGGAATGAATG
14	GGGAGGTAGTGACAAGAAATAACAATACAGGGCATCCATGTCTTGTAATTGGAATGAGTA
18	GGGAGGTAGTGACAAGAAATAACAATACAGGGCATCCATGTCTTGTAATTGGAATGAATA
32	GGGAGGTAGTGACAAGAAATAACAATACAGGGCATCCATGTCTTGTAATTGGAATGAATA
22	GGGAGGTAGTGACAAGAAATAACAATACAGGGCATCCATGTCTTGTAATTGGAATGAATA
29	GGGAGGTAGTGACAAGAAATAACAATACAGGGCATCCATGTCTTGTAATTGGAATGAATA
27	GGGAGGTAGTGACAAGAAATAACAATACAGGGCATCCATGTCGTGTAATTGGAATGGATA
26	GGGAGGTAGTGACAAGAAATAACAATACAGGGCATCCATGTCTTGTAATTGGAATGAATA
24	GGGAGGTAGTGACAAGAAATAACAATACAGGGCATCCATGTCTTGTAATTGGAATGAGTA
28	GGGAGGTAGTGACAAGAAATAACAATACAGGGCATCCATGTCTTGTAATTGGAATGAGTA
35	GGGAGGTAGTGACAAGAAATAACAATACAGGGCATATCTGTCTTGTAATTGGAATGAGTA
23	GGGAGGTAGTGACAAGAAATAACAATACAGGGCATCCATGTCTTGTAATTGGAATGAGTA
25	GGGAGGTAGTGACAAGAAATAACAATACAGGGCATCCATGTCTTGTAATTGGAATGAGTA
34	GGGAGGTAGTGACAAGAAATAACAATACAGGGCATCCATGTCTTGTAATTGGAATGAGTA
4	GGGAGGTAGTGACAAGAAATAACAATACAGGGCATCCATGTCTTGTAATTGGAATGAGTA
6	GGGAGGTAGTGACAAGAAATAACAATACAGGGCATCCATGTCTTGTAATTGGAATGAGTA
8	GGGAGGTAGTGACAAGAAATAACAATACAGGGCATCCATGTCTTGTAATTGGAATGAGTA
5	GGGAGGTAGTGACAAGAAATAACAATACAGGGCATCCATGTCTTGTAATTGGAATGAGTA
3	GGGAGGTAGTGACAAGAAATAACAATACAGGGCATCCATGTCTTGTAATTGGAATGAGTA
17	GGGAGGTAGTGACAAGAAATAACAATACAGGGCATCCATGTCTTGTAATTGGAATGAGTA
16	GGGAGGTAGTGACAAGAAATAACAATACAGGGCATTTCTGTCTTGTAATTGGAATGAGTA
2	GGGAGGTAGTGACAAGAAATAACAATACAGGGCATCCATGTCTTGTAATTGGAATGAACA
30	GGGAGGTAGTGACAAGAAATAACAATACAGGGCATCCATGTCTTGTAATTGGAATGAGTA
20	AGGAGGTAGTGACAAGAAATAACAATACAGGGCATCCATGTCTTGTAATTGGAATGAATA
43	AGGAGGTAGTGACAAGAAATAACAATACATGGCATGCATGTCTTGTAATTGGAATGTGCT
39	GGGAGGTAGTGACAAGAAATAACAATACAGGGCATCCATGTCTTGTAATTGGAATGAATG
40	GGGAGGTAGTGACAAGAAATAACAATACAGGGCATCCATGTCTTGTAATTGGAATGAGTA
42	GGGAGGTAGTGACAAGAAATAACAATACAGGGCATCCATGTCTTGTAATTGGAATGAGTA
38	GGGAGGTAGTGACAAGAAATAACAATACAGGGCATCCATGTCTTGTAATTGGAATGAGTA
41	GGGAGGTAGTGACAAGAAATAACAATACAGGGCATCCATGTCTTGTAATTGGAATGAGTA
36	GGGAGGTAGTGACAAGAAATAACAATACAGGGCATCCATGTCTTGTAATTGGAATGAGTA
37	GGGAGGTAGTGACAAGAAATAACAATACAGGGCATAATTGTCTTGTAATTGGAATGAGTA
31	GTGAGGTAGTGACAAGAAATAACAATACAGGGCATCCATGTCTTGTAATTGGAATGAGTA
15	AAATCCCCTTCCCATCCCCCCACCTTTTAACGCCCCCAC----TCTAAC--GGAGAATTA
	* ** ** * ** * ** *

1	GAATTTAAAACACTTTGCAAGTATCAATTGGAGGGCAAGTCT---GGTGCCAGCAGCCGC
21	GAATTTAAATATTTCTGCGAGTATCAATTGGAGGGCAAGTCT---GGTGCCAGCAGCCGC
7	AGCTCCAAACATGTTGGTGAGTATCAATTGGAGGGCAAGTCT---GGTGCCAGCAGCCGC
9	GATTTTAAACCTTTCTGTGAAGTATCAATTGGAGGGCAAGTCT---GGTGCCAGCAGCCGC
10	GATTTTAAACCTTTCTATAAGTATCAATTGGAGGGCAAGTCT---GGTGCCAGCAGCCGC
12	GATATTAAACCTTTCTATGAGTATCAATTGGAGGGCAAGTCT---GGTGCCAGCAGCCGC
13	GATATTAAACCTTTCTATGAGTATCAATTGGAGGGCAAGTCT---GGTGCCAGCAGCCGC
11	GATATTAAACCTTTCTATGAGTATCAATTGGAGGGCAAGTCT---GGTGCCAGCAGCCGC
19	GATGTTAACCTTCTCTGTGAGTATCAATTGGAGGGCAAGTCT---GGTGCCAGCAGCCGC
33	GAATTTAAATCTCTTTCTGAGTATCGATTGGAGGGCAAGTCT---GGTGCCAGCAGCCGC
14	GAATTTAAATCCCTTTACGAGTATCGATTGGAGGGCAAGTCT---GGTGCCAGCAGCCGC
18	GAATTTAAATCCCTTTACGAGTATCGATTGGAGGGCAAGTCT---GGTGCCAGCAGCCGC
32	GAATTTAAATCCCTTTACGAGTATCGATTGGAGGGCAAGTCT---GGTGCCAGCAGCCGC
22	GAATTTAAATCCCTTTATGAGTATCAATTGGAGGGCAAGTCT---GGTGCCAGCAGCCGC
29	GAATTTAAATCCCTTTATGAGTATCGATTGGAGGGCAAGTCT---GGTGCCAGCAGCCGC
27	GAATTTAAAYCCCTTTATGAGTATCGATTGGAGGGCAAGTCT---GGTGCCAGCAGCCGC
26	GAATTTAAATCCCTTTATGAGTATCGATTGGAGGGCAAGTCT---GGTGCCAGCAGCCGC
24	GAATTTAAATCCCTTTATGAGTATCGATTGGAGGGCAAGTCT---GGTGCCAGCAGCCGC
28	GAATTTAAATCCCTTTACGAGTATCGATTGGAGGGCAAGTCT---GGTGCCAGCAGCCGC
35	GAATTTAAATCCCTTTACGAGTACCAATTGGAGGGCAAGTCT---GGTGCCAGCAGCCGC
23	GAATTTAAATCCCTTTGCGAGTATCAATTGGAGGGCAAGTCT---GGTGCCAGCAGCCGC

```
25  GAATTTAAATCCCTTTGCGAGTATCAATTGGAGGGCAAGTCT---GGTGCCAGCAGCCGC        20  GGTAATT--CCAGCTCCAATAGCGTATATTAAAGTTGTTGCGGTTAAAAAGCTCGTAGTT
34  GAATTTAAATCCCTTTGCGAGTATCAATTGGAGGGCAAGTCT---GGTGCCAGCAGCCGC        43  GGTAATT--CCAGCTCCAATAGCGTATATTAAAGTTGTTGCGGTTAAAAAGCTCGTAGTT
4   GAATTCAAATCCCTTTGCAAGTACCAATTGGAGGGCAAGTCT---GGTGCCAGCAGCCGC        39  GGTAATT--CCAGCTCCAATAGCGTATATTAAAGTTGTTGCGGTTAAAAAGCTCGTAGT-
6   GAATTTAAACCCCTTTATGAGTATCAATTGGAGGGCAAGTCT---GGTGCCAGCAGCCGC        40  GGTAATT--CCAGCTCCAATAGCGTATATTAAAGTTGTTGCGGTTAAAAAGCTCGTAGT-
8   GAATTTAAACCCCTTTATGAGTATCAATTGTAGGGCAAGTCT---GGTGCCAGCAGCCGC        42  GGTAATT--CCAGCTCCAATAGCGTATATTAAAGTTGTTGCGGTTAAAAAGCTCGTAG--
5   GAATTTAAACCCCTTTATGAGTATCAATTGGAGGGCAAGTCT---GGTGCCAGCAGCCGC        38  GGTAATT--CCAGCTCCAATAGCGTATATTAAAGTTGTTGCGGTTAAAAAGCTCGTAG--
3   GAATTTAAACCCCTTTATGAGTATCAATTGGAGGGCAAGTCT---GGTGCCAGCAGCCGC        41  GGTAATT--CCAGCTCCAATAGCGTATATTAAAGTTGTTGCGGTTAAAAAGCTCGTAG--
17  GAATTTAAATCCCTTTATGAGTATCAATTGGAGGGCAAGTCT---GGTGCCAGCAGCCGC        36  GGTAATT--CCAGCTCCAATAGCATATATTAAAGTTGTTGCGGTTAAAA----------
16  GAATTTAAATCTCTTTATGAGTATCAATTGGAGGGCAAGTCT---GGTGCCAGCAGCCGC        37  GGTAATT--CCAGCTCCAATAGCGTATATTAAAGTTGTTGCGGTTAAAA----------
2   GAATTTAAATCTCTTTATGAGTATCAATTGGAGGGCAAGTCT---GGTGCCAGCAGCCGC        31  GGTAATT--CCAGCTCCAATAGCGTATATTAAAGTTG---------------------
30  AAATTTAAATCTCTCTACGAGTATCAATTGGAGGGCAAGTCT---GGTGCCAGCAGCCGC        15  GCAAATTACCCAATCCTGACACAGGGAGGTATATC----------------------
20  GAATTCAAACCCTTTTGTTAGTATCGATTGGAGGGCAAGTCT---GGTGCCAGCAGCCGC            *
43  AATCCTAAACTATTTTACGAACTTCAATTGGAGGGCAAGTCT---GGTGCCAGCAGCCGC
39  GAAATTAAACCTCTTTGTAAGTATCGATTGGAGGGCAAGTCT---GGTGCCAGCAGCCGC         1  ------------------
40  GACTTTAAATCACTTTGCAAGTATCGATTGGAGGGCAAGTCT---GGTGCCAGCAGCCGC        21  ------------------
42  GACTTTAAATCACTTTGCAAGTATCGATTGGAGGGCAAGTCT---GGTGCCAGCAGCCGC         7  ------------------
38  GAATTTAAATCCCTTTATGAGTATCGATTGGAGGGCAAGTCT---GGTGCCAGCAGCCGC         9  ------------------
41  GAATTTAAATCCCTTTTCGAGTATCAATTGGAGGGCAAGTCT---GGTGCCAGCAGCCGC        10  ------------------
36  GAACTTAAATCTCTTTGTGAGTACCAATTGGAGGGCAAGTCT---GGTGCCAGCAGCCGC        12  ------------------
37  GAATTTAAATCCCTTTACGAGTACCAATTGGAGGGCAAGTCT---GGTGCCAGCAGCCGC        13  ------------------
31  GAAATTAAAAGCCTTTGCGAGTACCAATTGGAGGGCAAGTCT---GGTGCCAGCAGCCGC        11  ------------------
15  GGGTTCGA--TTCCGGAGAGGGAGCAAACGGCTACCACATCTAAGGAAGGCAGCAGGCGC        19  ------------------
         *          * *  *     **  ***   *  * ******  ***       33  ------------------
                                                                    14  ------------------
1   GGTAATT--CCAGCTCCAATAGCGTATATTAAAGTTGTTGC-----------------         18  ------------------
21  GGTAATT--CCAGCTCCAACAGCGTATATTAAAGTTGTTGCGG---------------         32  ------------------
7   G-TAATT--CCAGCTCCAATAGCGTATATTAAAGCTGTTG-----------------         22  ------------------
9   G-TAATT--CCAGCTCCAATAGCGTATATTAAAGTTGT-------------------         29  ------------------
10  GGTAATT--CCAGCTCCAATAGCGTATATTAAAGTTGT-------------------         27  ------------------
12  GGTAATT--CCAGCTCCAATAGCGTATATTAAAGTTG--------------------         26  ------------------
13  GGTAATT--CCAGCTCCAATAGCGTATATTAAAGTTG--------------------         24  ------------------
11  GGTAATT--CCAGCTCCAATAGCGTATATTAAAGTT---------------------         28  ------------------
19  GGTAATT--CCAGCTCCAATAGCGTATATTAAAGTT---------------------         35  ------------------
33  GGTAATT--CCAGCTCCAAAAGCGTATATTAAAGTTG--------------------         23  ------------------
14  GGTAATT--CCAGCTCCAATAGCGTATATTAAAGTTGTTGC----------------         25  ------------------
18  GGTAATT--CCAGCTCCAATAGCGTATATTAAAGTTGT-------------------         34  ------------------
32  GGTAATT--CCAGCTCCAATAGCGTATATTAAAGTTG--------------------         4   ------------------
22  GGTAATT--CCAGCTCCAATAGCGTATATTAAAGTTG--------------------         6   ------------------
29  GGTAATT--CCAGCTCCAATAGCGTATATTAAAGTTG--------------------         8   ------------------
27  GGTAATT--CCAGCTCCAATAGCGTATATTAAAGTTG--------------------         5   ------------------
26  GGTAATT--CCAGCTCCAATAGCGTATATTAAAGTTG--------------------         3   ------------------
24  GGTAATT--CCAGCTCCAATAGCGTATATTAAAGTTG--------------------         17  ------------------
28  GGTAATT--CCAGCTCCAATAGCGTATATTAAAGTTGT-------------------         16  ------------------
35  GGTAATT--CCAGCTCCAATAGCGTATATTAAAGTTG--------------------         2   ------------------
23  GGT-----------------------------------------------------         30  ------------------
25  GGTAATT--CCAGCTCCAATAGCGTATATTAAAGTTG--------------------         20  GGATTTATGCTGAAGACAA
34  G-TAATT--CCAGCTCCAATAGCGTATATTAAAGTTGT-------------------         43  TGATGTGT----------
4   GGTAATT--CCAGCTCCAATAGCGTATATTAAAGTTGTT------------------         39  ------------------
6   GGTAATT--CCA-CTCCAATAGCGTATATTAAAGTTGTTGCG---------------         40  ------------------
8   GGTAATT--CCAGCTCCAATAGCGTATATTAAAGTTGTTGCG---------------         42  ------------------
5   GGTAATT--CCAGCTCCAATAGCGTATATTAAAGTTGTTGC----------------         38  ------------------
3   GGTAATT--CCAGCTCCAATAGCGTATATTAAAGTTGTTGCGGTTA-----------         41  ------------------
17  GGTAATT--CCAGCTCCAATAGCGTATATTAAAGTTGTTGC----------------         36  ------------------
16  GGTAATT--CCAGCTCCAATAGCGTATATTAAAGTTGTTGCGG--------------         37  ------------------
2   GGTAATT--CCAGCTCCAATAGCGTATATTAAAGTTGTTGCGGT-------------         31  ------------------
30  GGTAATT--CCAGCTCCAATAGCGTATATTAAAGTTGTTGCGGTTAAAAAG------         15  ------------------
```

Fig. 1. Sequence alignments of the partial SSU region for 43 species of dinoflagellates. CLUSTAL W generated the alignment. A hypen represents a gap and a period represents a base identical to that of the top sequence.

An asterisk represents an identical sequence on vertical lines. 1, *C. cohnii* M64225; 21, *G. spinifera* AF022155; 7, *A. belauense* L13719; 9, *A. tamarense* X62650; 10, *A. fundyense* U09048; 12, *A. minutum* U27499; 13, *A. ostenfeldii* U27500; 11, *A. margaelefii* U27498; 19, *C. tenue* AF022192; 33, *P. noctiluca* AF022156; 14, *G. sanguineum* U41085; 18, *S. nutricula* U52357; 32, *P. tyrrhenicum* AF022201; 22, *G. catenatum* AF022193; 29, *L. viride* AF022199; 27, *G. impudicum* AF022197; 26, *Gymnodinium* sp. AF022196; 24, *G. fuscum* AF022194; 28, *H. triquetra* AF022198; 35, *P. micans* M14649; 23, *G. aureolum*; 25, *G. mikimotoi* AF022195; 34, *G. breve* AF172714; 4, *G. viscum* L13716; 6, *S. meandrinae* L13718; 8, *S. pilosum* X62650; 5, *S. corculorum* L13717; 3, *S. microadraticum* M88521; 17, *G. beii* U41087; 16, *G. simplex* U41086; 2, *Symbiodinium* M88509; 30, *N. scintillans* AF022200; 20, *C. horrida* AF022154; 43, *G. cochlea* AF274258; 39, *P. bahamense* AF274275; 40, *G. uncatenum* AF274263; 42, *G. dorsum* AF274261; 38, *S. trocidea* AF274277; 41, *G. galatheanum* AF274262; 36, *P. lima* Y16235; 37, *P. minimum* Y16238; 31, *Peridinium* sp. AF022202; 15, *C. polykrikoides.*
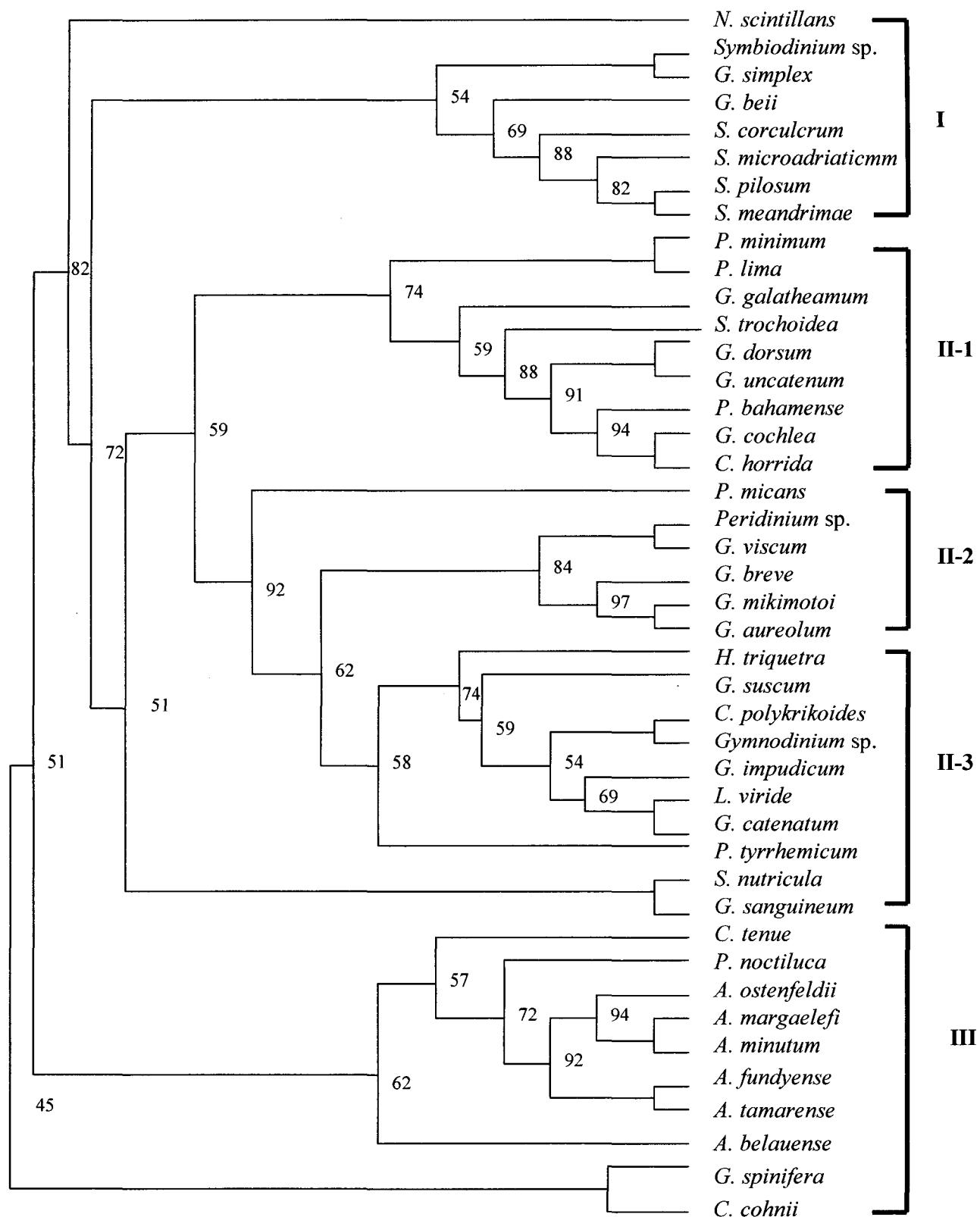
Fig. 2. PHYLIP analysis of 41 species obtained from GenBank database including *C. polykrikoides* and *G. aureolum*. The tree was obtained using subprogram DNAPARS in PHYLIP. The topology represents the consensus tree from an heuristic search yielding two equally most parsimonious tree. Bootstrap values (100 replications) are given above the internal nodes. This is an unrooted tree.
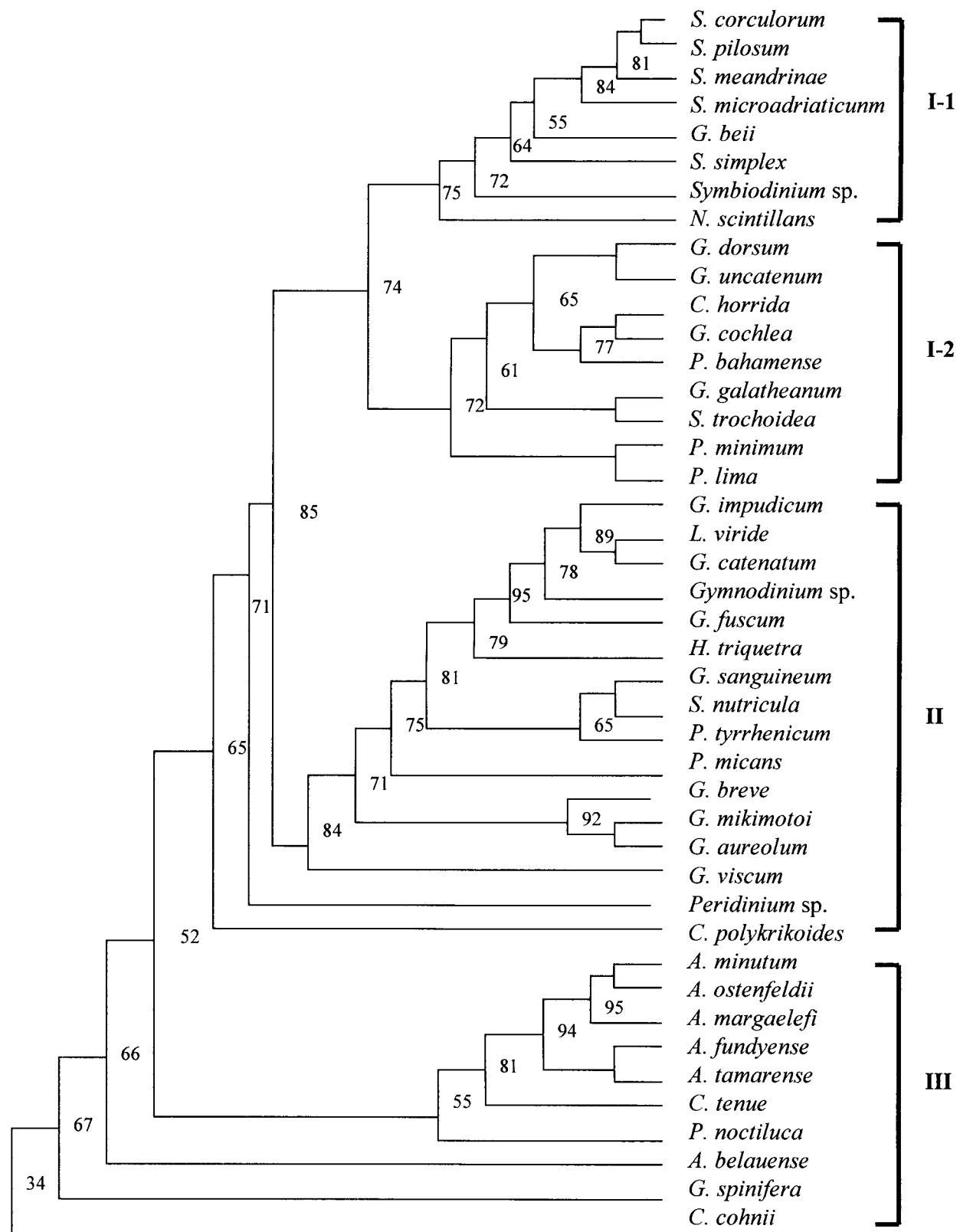
Fig. 3. PHYLIP analysis of 41 species obtained from GenBank database including *C. polykrikoides* and *G. aureolum*. The tree was obtained using subprogram NEIGBOR in PHYLIP with the option of Kimura's two-parameter method. Bootstrap values (100 replications) are given above the internal nodes. This is an unrooted tree.

Fig. 4. PHYLIP analysis of 41 species obtained from GenBank database including *C. polykrikoides* and *G. aureolum*. The tree was obtained using subprogram DNAMLK in PHYLIP with the option of Kimura's two-parameter method. Bootstrap values (100 replications) are given above the internal nodes. This is an unrooted tree.

Table 1. Percentage of similarity of nucleotide sequences of Dinophyceae

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43
1	-																																										
2	84	-																																									
3	85	94	-																																								
4	84	88	90	-																																							
5	86	93	98	91	-																																						
6	85	93	98	81	99	-																																					
7	79	81	81	81	82	81	-																																				
8	85	93	98	91	99	99	80	-																																			
9	81	83	84	84	82	84	80	84	-																																		
10	81	83	86	83	85	86	80	85	95	-																																	
11	82	86	87	85	87	88	82	86	89	89	-																																
12	83	88	87	86	88	88	83	88	92	92	95	-																															
13	82	88	92	86	88	87	82	87	84	91	95	97	-																														
14	85	92	85	92	93	86	83	93	80	85	87	88	89	-																													
15	81	84	81	85	86	88	80	86	85	85	82	83	84	86	-																												
16	86	94	95	91	95	95	81	95	85	86	87	89	88	93	96	-																											
17	86	94	93	92	93	96	83	96	85	85	87	89	89	93	85	96	-																										
18	82	86	84	85	86	86	79	85	84	86	88	89	90	94	86	93	94	-																									
19	76	82	81	91	81	80	72	80	77	83	85	86	85	85	85	81	85	84	-																								
20	84	86	85	83	85	85	80	85	81	81	88	80	79	82	80	86	81	81	75	-																							
21	86	91	91	93	92	92	82	92	85	81	87	85	84	85	81	93	85	86	82	85	-																						
22	85	85	92	93	91	91	82	92	85	86	87	89	89	94	86	92	93	89	84	84	95	-																					
23	85	90	92	92	92	92	83	92	84	85	88	89	89	95	81	92	93	95	80	80	99	95	-																					
24	86	91	92	87	92	92	82	92	86	85	87	84	84	93	86	91	93	95	81	84	94	96	94	-																				
25	86	88	91	90	92	92	83	91	84	84	88	87	87	93	82	94	93	95	80	85	97	95	96	95	-																			
26	85	91	92	92	92	92	80	92	85	85	85	86	86	94	85	88	93	95	81	84	88	88	97	88	94	-																	
27	86	90	91	92	91	92	79	91	82	84	84	84	84	88	82	91	92	94	85	81	89	90	87	94	88	96	-																
28	86	90	91	92	92	92	79	92	81	85	86	87	87	93	85	94	94	97	81	84	97	98	93	96	90	88	95	-															
29	85	90	91	92	91	92	82	91	83	86	88	88	89	93	82	91	93	88	82	81	96	98	94	96	88	89	93	96	-														
30	86	91	92	93	93	87	83	93	85	84	87	83	84	94	85	89	88	89	85	85	99	95	97	94	90	93	93	98	88	-													
31	84	88	88	86	85	86	80	87	82	80	82	83	83	87	82	92	88	81	76	84	88	86	86	89	97	95	88	88	89	84	-												
32	82	87	86	86	86	86	79	86	84	81	82	89	84	89	84	88	88	83	81	81	89	89	89	88	88	97	89	98	88	88	86	-											
33	85	91	92	93	92	90	82	93	81	86	88	85	89	96	87	86	94	90	81	85	97	97	91	96	90	87	89	97	89	96	85	85	-										
34	81	85	84	87	84	89	82	84	86	83	83	88	88	85	85	93	86	88	76	82	86	86	93	85	86	88	93	96	87	88	83	83	88	-									
35	83	90	91	90	92	90	82	91	85	85	85	89	89	94	82	85	93	90	80	85	95	95	93	94	85	87	88	95	85	89	84	84	93	95	-								
36	82	91	92	90	92	88	81	92	82	83	87	84	84	88	85	92	89	88	85	82	89	88	87	88	82	90	90	92	88	87	88	82	89	93	92	-							
37	82	85	89	88	88	90	81	88	85	85	83	87	87	93	82	88	91	93	85	85	93	93	89	93	85	87	87	95	87	88	80	83	89	89	88	90	-						
38	83	89	91	90	91	87	81	91	83	83	85	86	84	92	85	92	88	93	81	82	92	92	92	92	82	90	91	93	93	90	93	84	93	93	96	87	92	-					
39	83	89	91	88	89	88	79	90	82	85	84	87	87	90	82	91	90	90	82	83	88	87	88	88	85	84	87	89	87	84	82	82	89	89	90	90	93	90	-				
40	83	89	93	90	90	90	79	90	82	82	86	84	85	92	83	89	92	93	83	84	90	89	90	89	90	90	88	90	90	88	86	90	93	93	93	90	93	96	92	-			
41	83	89	90	88	89	89	79	89	82	82	83	86	84	90	81	89	92	90	85	83	92	93	90	90	93	88	88	90	84	84	93	88	82	89	89	87	88	90	92	93	-		
42	83	89	90	90	89	89	79	89	82	82	84	86	85	90	81	81	90	90	85	85	89	89	89	90	90	89	90	90	90	82	90	82	82	74	74	76	78	80	79	99	93	-	
43	76	76	77	74	74	76	73	76	75	76	76	76	76	75	77	76	75	74	75	78	74	74	74	75	75	74	73	75	74	75	73	75	74	74	74	75	78	74	80	79	79	79	-

Note: 1. C. cohnii M64245; 2. Symbiodinium sp. M88521; 3. S. microadriatic M88509; 4. G. viscum L13716; 5. S. corculorum L13717; 6. S. meandrinae L13718; 7. A. belauense L13719; 8. S. pilosum X62650; 9. A. tamarense X54946; 10. A. fundyense U09048; 11. A. margaelefii U27498; 12. A. minutum U27499; 13. A. ostenfeldii U27500; 14. G. sanguineum U41085; 15. C. polykrikoides; 16. G. simplex U41086; 17. G. beii U41087; 18. S. nutricula U52357; 19. C. tenue AF022192; 20. C. horrida AF022154; 21. G. spinifera AF022155; 22. G. catenatum AF022193; 23. G. aureolum; 24. G. fuscum AF022194; 25. G. mikimotoi AF022195; 26. Gymnodinium sp. AF022196; 27. G. impudicum AF022197; 28. H. triquetra AF022198; 29. L. viride AF022199; 30. N. scintillans AF022200; 31. Peridinium sp. AF022202; 32. P. tyrrhenicum AF022201; 33. P. noctiluca AF022156; 34. G. breve AF172714; 35. P. micans M14649; 36. P. lima Y16235; 37. P. minimum Y16238; 38. S. trochoidea AF274277; 39. P. bahamense AF274275; 40. G. uncatenum AF274263; 41. G. galathenum AF274262; 42. G. dorsum AF274261; 43. G. cochlea AF274258

Table 2. Pairwise distances between fourteen species including *C. polykrikoides* of the genus *Gymnodinium* and *Gyrodinium* from the aligned SSU rDNA gene sequences on the basis of Kimura two-parameter estimates

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	-	19	50	50	35	42	42	49	46	45	41	41	43	49
2	0.0319	-	42	42	31	34	36	45	38	39	39	39	38	52
3	0.0828	0.0695	-	0	43	40	42	53	48	42	46	42	45	50
4	0.0826	0.0694	0.0000	-	43	40	42	53	48	42	46	42	45	50
5	0.0583	0.0510	0.0711	0.0710	-	24	26	36	29	32	20	19	20	53
6	0.0693	0.0563	0.0659	0.0658	0.0403	-	30	40	34	31	25	24	26	54
7	0.0695	0.0601	0.0698	0.0697	0.0440	0.0494	-	17	12	16	27	24	26	43
8	0.0805	0.0745	0.0870	0.0869	0.0604	0.0656	0.0285	-	20	26	36	34	36	42
9	0.0770	0.0638	0.0793	0.0792	0.0477	0.0566	0.0200	0.0337	-	18	31	29	12	42
10	0.0751	0.0655	0.0698	0.0697	0.0532	0.0514	0.0268	0.0441	0.0302	-	32	29	30	58
11	0.0681	0.0640	0.0759	0.0757	0.0334	0.0426	0.0443	0.0603	0.0514	0.0532	-	6	6	39
12	0.0681	0.0640	0.0702	0.0701	0.0316	0.0409	0.0407	0.0566	0.0477	0.0476	0.0100	-	5	36
13	0.0721	0.0639	0.0745	0.0743	0.0335	0.0432	0.0430	0.0598	0.0504	0.0504	0.0105	0.0094	-	31
14	0.0812	0.0864	0.0825	0.0823	0.0882	0.0897	0.0721	0.0703	0.0695	0.0956	0.0654	0.0598	0.0521	-

Note: 1. *G. simplex* U41086; 2. *G. beii* U41087; 3. *G. uncatenum* AF274263; 4. *G. dorsum* AF274261; 5. *G. galathenum* AF274262; 6. *G. sanguineum* U41085; 7. *G. catenatum* AF022193; 8. *G. impudicum* AF022197; 9. *Gymnodinium* sp. AF022196; 10. *G. fuscum* AF022194; 11. *G. breve* AF172714; 12. *G. mikimotoi* AF022195; 13. *G. aureolum*; 14. *C. polykrikoides* Numbers above the diagonal are absolute distance values; numbers below the diagonal are mean distance values.
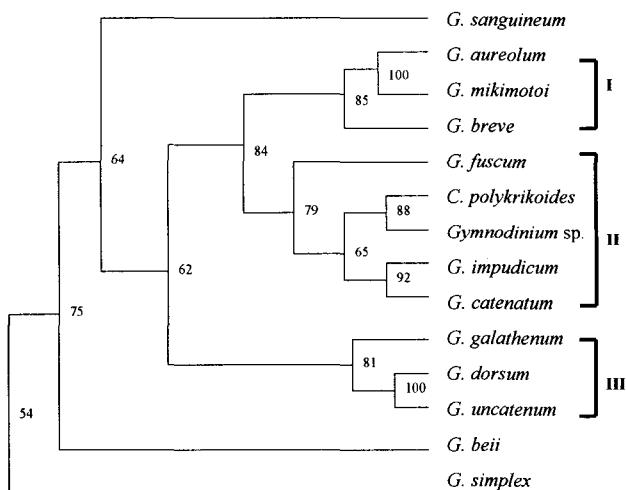
Fig. 5. PHYLIP analysis 12 species within the genus *Gymnodinium* and *Gyrodinium* obtained from GenBank database including *C. polykrikoides* and *G. aureolum*.
The tree was obtained using subprogram DNAPARS in PHYLIP. The topology represents the consensus tree from an heuristic search yielding two equally most parsimonous tree. Bootstrap values (1000 replications) are given above the internal nodes. This is an unrooted tree.

Fig. 6. PHYLIP analysis 12 species within the genus *Gymnodinium* and *Gyrodinium* obtained from GenBank database including *C. polykrikoides* and *G. aureolum*.
The tree was obtained using subprogram NEIGBOR in PHYLIP with the option of Kimura's two-parameter method. Bootstrap values (1000 replications) are given above the internal nodes. This is an unrooted tree.

Gymnodinium beii and *G. simplex* were sister groups to Group I and to Group II, respectively. The unrooted phylogenetic trees obtained from the NJ and ML methods were similar to the result obtained from the parsimony analysis (Figs. 6, 7). Likewise, *C. polykrikoides* and *G. aureolum* were closer to the genus *Gymnodinium* than to the genus *Gyrodinium*, and
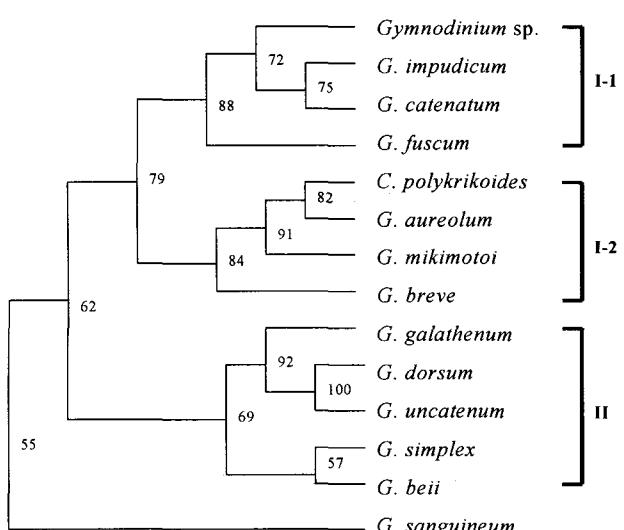
were supported by a moderate bootstrap value of 79%.

Discussion

Dinoflagellates

Fensome *et al.* [11] separated *N. scintillans* from the class
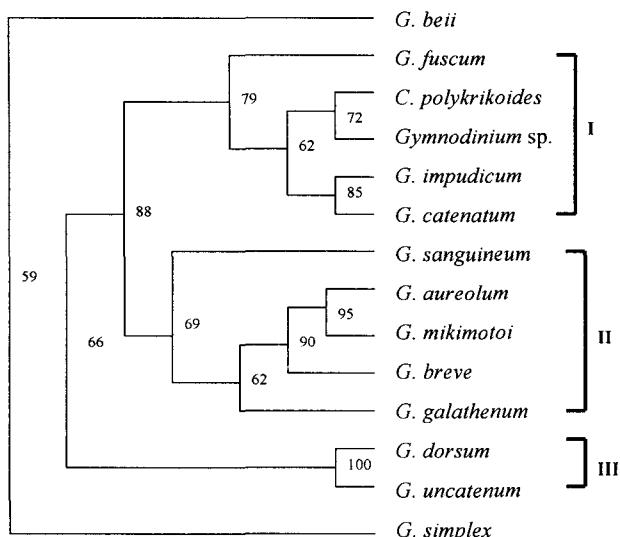
Fig. 7. PHYLIP analysis 12 species within the genus *Gymno-dinium* and *Gyrodinium* obtained from GenBank database including *C. polykrikoides* and *G. aureolum*.
The tree was obtained using subprogram DNAMLK in PHYLIP with the option of Kimura's two-parameter method. Bootstrap values (1000 replications) are given above the internal nodes. This is an unrooted tree.

Dinophyceae and placed it in the class Noctiluciphyceae. In the present study, *N. scintillans* played a role as a primitive dinoflagellate, but also *G. spinifera* and *C. cohnii*, which were a member of the gonyaulacoid dinoflagellates reflected many genetic divergences of other taxa. Since *C. cohnii* appears intermediately between Noctiluciphycidae and Balstodini-phycidae, some studies have found that this species diverged early in dinoflagellate evolution [12,23]. This study showed that *N. scintillans*, *C. cohnii* and *G. spinifera* did not enough to support the dinoflagellates. Fensome *et al.* [11] suggested *C. cohnii* was able to contribute to its own family within the Gonyaulacales. The phylogenetic position of *C. cohnii* in our study was possibly responsible for its joining to the order Gonyaulacales. Fensome *et al.* [11] reported that *G. viscum* was not assigned to a subclass, which has been done uncertainly, but that gave strong bootstrap support to *Gym-ndinium* (Figs. 2-4). Thus, the three species of *N. scintillans*, *G. spinifera* and *C. cohnii* represent a separate lineage outside the Dinophyceae according to this SSU phylogeny. According to the taxonomic scheme presented in Fensome *et al.* [11], the genus *Prorocentrum* was considered to belong to the independent subclass Prorocentrophycidae, but was closely related and diverged relatively to the subclass Gymnodiniphycidae in our analysis, with strong bootstrap support (Figs. 2-4).

Gymnodinium/Gyrodinium

The classification of naked microalgae bearing a resemblance to the *Gymnodinium* or *Gyrodinium* genera is difficult, as the unique distinction between these two genera is the degree of displacement of the cingulum separating the epicone and hypocone. Cingular displacement must be less than one fifth of the cell's body length for the genus *Gymnodinium* and greater than one fifth of the body length for the genus *Gyrodinium* [16]. Such an arbitrary criterion for determining genera does not account for the variations between individual cells of a unialgal isolate. Important attention has been paid to *G. mikimotoi* and *G. aureolum*, whose circumscription s has been confused for a long time. Since two very similar dinoflagellates have been cultured, both of these species were called *G. nagasakiense* [28], prior to the name change to *G. mikimotoi* [27]. Some researchers have considered these to synonymous species, while others have proposed to remain uncertain [20]. However, the name *G. aureolum* is still widely used in Europe [14].

This study indicated that Korean *G. aureolum* and *G. mikimotoi* were genetically different targeted to the SSU, although the genetic divergence was extremely low (Table 2). The result obtained from the large subunit (LSU) gene sequence was that the genetic divergence between them was a higher value of approximately 18.7% compared with that of 0.9% when targeted to the SSU. This indicates that it is likely that the LSU is expected to be desirable as a genetic marker for discriminating *G. aureolum* from *G. mikimotoi* and for determining the clear genetic positions of the two species. Thus, study based on only molecular data has supported the distinction of *G. mikimotoi* from Korean *G. aureolum*. Although *G. aureolum* and *G. impudicum* belong to the genus *Gyrodinium*, the two species are closer to the genus *Gymno-dinium* than to the genus *Gyrodinium* based on the analysis of morphological characters and gene sequences in the present work. Thus, they did not move to the taxa of *G. galathenum*, *G. dorsum* and *G. uncatenatum* with the trees produced by the parsimony, distance, and ML methods (Figs. 5-7). We suggest that *G. aureolum* and *G. impudicum* are limited to the genus *Gyrodinium*, but transferred to the genus *Gymnodinium*. More recent study has defined the need for improving the classification of gymnodinioid dinofla-gellates by the analysis of morphological and molecular data [8,31]. Although Korean *G. aureolum* probably belongs to the genus *Karenia*, it did not play a role as a member of the genus *Gyrodinium*.

Cochlodinium polykrikoides

One example of the typically different character between *C. polykrikoides* and the genus *Gymnodinium* of morphological features is that a girdle in *C. polykrikoides* makes 1.8-1.9 turns around the cell [4]. Because of this problem, three species, *C. polykrikoides*, *G. impudicum*, and *G. catenaum*, are difficult to discriminate and often are misidentified as each other under a light microscope [3]. Previously, we analyzed gene sequences targeted to internal transcribed spacer (ITS) regions and the phylogenetic placement of them indicated that *C. polykrikoides* was distant from *G. impudicum* and *G. catenatum* [3,4]. At present, the genetic distance between *C. polykrikoides* and *G. impudicum/G. catenatum* based on the SSU is also similar to the result obtained from the ITS region, that *G. impudicum* was close to *G. catenaum* than to *C. polykrikoides* (Table 2). However, as can be seen in Figs. 5-7, *C. polykrikoides* did not transfer more to the genus *Gymnodinium* than to the genus *Gyrodinium*, under any type of topology, and also was supported by a moderate bootstrap. Although cigulum plays an important role in the potential taxonomic identification of *C. polykrikoides* and the genus *Gymnodinium*, the genetic studies on sequence divergence and phylogenetic analyses give us assurance that *C. polykrikoides* may be closer to the genus *Gymnodinium* than to the genus *Gyrodinium*.

Acknowledgements

References

1. Adachi, M., Y. Sako and Y. Ishida. 1993. The identification of conspecific dinoflagellates *Alexandrium tamarense* from Japan and Thailand by monoclonal antibodies. *Nipppon Suisan Gakkaishi* **59**, 327-332.

2. Chang, F.H. 1996. A review of knowledge of a group of closely related economically important toxic *Gymnodinium/Gyrodinium* (Dinophyceae) species in New Zealand. *J. Roy. Soc. NZ* **26**, 381-394.

3. Cho, E. S., G. Y. Kim and Y. C. Cho. 2001b. Molecular analysis of morphologically similar dinoflagellates *Cochlodinium polykrikoides, Gyrodinium impudicum* and *Gymnodinium caten-*

4. Cho, E. S., G. Y. Kim, B. D. Choi, L. L. Rhodes, T. J. Kim, G. H. Kim and J. D. Lee. 2001a. A comparative study of the harmful dinoflagellates *Cochlodinium polykrikoides* and *Gyrodinium impudicum* using transmission electron microscopy, fatty acid composition, carotenoid content, DNA quantification and gene sequences. *Bot. Mar.* **44**, 57-66

5. Cho, E. S., G. Y. Kim, T. H. Lee, Y. C. Cho and J. D. Lee. 2001c. Phylogenetic relationship among several Korean coastal red tide dinoflagellates based on their rDNA internal transcribed spacer sequences. *J. Kor. Lif. Sci.* **11**, 74-80.

6. Cho, E. S., H. G. Kim and Y. C. Cho. 2000. Sequence analysis of *Cochlodinium polykrikoides* isolated from Korean coastal waters using sequences of Internal Transcribed Spacers and 5.8S rDNA. *J. Kor. Soc. Oceano.* **35**, 158-160.

7. Costas, E., R. Zardoya, J. Bautista, A. Garrido, C. Rojo and V.L. Rodas. 1995. Morphospecies vs. genospecies in toxic marine dinoflagellates: an analysis of *Gymnodinium catenatum/Gyrodinium impudicum* and *Alexndrium minutum/A. lusitanicum* using antibodies, lectins and gene sequences. *J. Phycol.* **31**, 801-807.

8. Daugbjerg, N., G. Hansen, J. Larsen and O. Moestrup. 2000. Phylogeny of some of the major genera of dinoflagellates based on ultrastructure and partial LSU rDNA sequence data, including the erecton of three new genera of unarmoured dinoflagellates. *Phycologia* **39**, 302-317.

9. Destombe, C., A.D. Cembella, C.A. Murphy and M.A. Ragan. 1992. Nucleotide sequence of the 18S ribosomal RNA genes from the marine dinoflagellate *Alexandrium tamarense* (Gonyaulacales, Dinophyta). *Phycologia* **31**, 121-124.

10. Felsenstein, J. 1993. PHYLIP (Phylogeny Inference Package) Version 3.5c. Department of Genetics, University of Washington, Seattle.

11. Fensome, R. A., F. J. R. Taylor, G. Norris, W. A. S. Sarjeant, D. I. Wharton and G. L. Williams. 1993. *A classification of living and fossil dinoflagellates.* Micropaleontological Species Publication. 351 pp.

12. Grzebyk, D., Y. Sako and B. Berland. 1998. Phylogenetic analysis of nine species of *Prorocentrum* (Dinophyceae) inferred from 18S ribosomal DNA sequences, morphological comparisons, and description of *Prorocentrum panamensis*, sp. nov. *J. Phycol.* **34**, 1055-1068.

13. Guillard, R. R. L. and J. H. Ryther. 1962. Studies of marine planktonic diatoms 1. *Cyclotella nana* Hustedt and *Detonula confervacea* (Cleve) Gran. *Can. J. Microbiol.* **8**, 229-239.

14. Hansen, P.J. 1995. Grwoth and grazing response of a ciliate feeding on the red tide dinoflagellate *Gyrodinium aureolum* in monoculture and in mixture with a non-toxic alga. *Mar. Ecol. Prog. Ser.* **121**, 65-72.

15. Hasui, M., M. Matsuda, S. Yoshimatsu and K. Okutani. 1995. Production of a lactate-associated galactan sulfate by a dinoflagellate *Gymnodinium A₃*. *Fisheries Science* **61**, 321-326.

16. Hulburt, E. M. 1957. The taxonomy of unarmored Dinophyceae of shallow embayments of Cape Cod, Massachusetts. *Biol. Bull.* **112**, 196-219.

17. Innis, M. A., D. H. Gelfamd, J. J. Sninsky and T. J. White. 1990. *PCR protocols a guide to methods and applications.* Academic Press. 322 pp.

18. Kim, H. G., S. G. Lee, K. H. An, S. H. Youn, P. Y. Lee, C. Y. Lee, E. S. Cho, J. B. Kim, H. G. Choi and P. Y. Kim. 1997. *Recent red tides in Korean coastal waters.* Kudeok Publishing, Pusan, 280 pp. (in Korean)

19. Kimura, M. 1980. A simple method for estimating evolutionary rate of base substitution through comparative studies of nucleotide sequences. *J. Mol. Evol.* **116**, 111-120.

20. Partensky, F., D. Vaulot, A. Coute and A. Sournia. 1988. Morphological and nuclear analysis of the bloom-forming dinoflagellates *Gyrodinium* cf. *aureolum* and *Gymnodinium nagasakiense.* *J. Phycol.* **24**, 408-415.

21. Saitou, N. and M. Nei. 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**, 406-425.

22. Sako, Y., C. H. Kim, H. Ninomiya, M. Adachi and Y. Ishida. 1990. Isozyme and cross analysis of mating populations in the *Alexandrium catenella/tamarense* species complex. In Graneli, E., B. Sundstrom, L. Edler and D.M. Anderson (eds.). *Toxic Marine Phytoplankton.* Elsevier, New York, pp. 320-323.

23. Saunders, G. W., D. R. A. Hill, J. P. Sexton and R. A. Anderson. 1997. Small-subunit ribosomal RNA sequences from selected dinoflagellates: testing classical evolutionary hypotheses with molecular systematic methods. *Pant Syst. Evol.* **11**, 237-259.

24. Scholin, C. A. and D.M. Anderson. 1993. Population analysis of toxic and non-toxic *Alexandrium* species using ribosomal RNA signature sequences. In Graneli, E., B. Sundstrom, L. Edler and D.M. Anderson (eds.). *Toxic Marine Phytoplankton.* Elsevier, New York, pp. 95-102.

25. Scholin, C. A. and D. M. Anderson. 1994. Identification of group and strain-specific genetic markers for globally distributed *Alexandrium* (Dinophyceae). I. RFLP analysis of SSU rRNA genes. *J. Phycol.* **30**, 744-754.

26. Scholin, C. A., D. M. Anderson and M. L. Sogin. 1993. Two distinct small-subunit ribosomal RNA genes in the north American toxic dinoflagellate *Alexandrium fundyense* (Dinophyceae). *J. Phycol.* **29**, 209-216.

27. Takayama, H. and K. Matsuoka. 1991. A reassessment of the specific characters of *Gymnodinium mikimotoi* Miyake et Kominami et Oda and *Gymnodinium nagasakiense* Takayama et Adachi. Bull. *Plankton Soc. Jpn.* **38**, 53-68.

28. Takayama, H. and R. Adachi. 1984. *Gymnodinium nagasakiense* sp. nov., a red-tide forming dinophyte in the adjacent waters of Japan. Bull. *Plankton Soc. Jpn.* **31**, 7-14.

29. Taylor, F. J. R. 1985. The taxonomy and relationships of red tide dinoflagellates.In Anderson, D. M., A. W. White and D. C. Baden (eds.). *Toxic Dinoflagellates.* Elsevier, New York, pp. 11-26.

30. Thomson, J. D., D. G. Higgins and T. J. Gibson. 1994. Clustal W: improving the sensitivity of progressive multiple sequences alignment through sequence weighting, position-specific fap penalties and weight matrix choice. Nucl. Acid Res. **22**, 4673-4680.

31. Yang, Z. B., H. Takayama, K. Matsuoka and I. J. Hodgkiss. 2000. *Karenia digitata* sp. nov. (Gymnodiniales, Dinophyceae), a new harmful algal bloom species from the coastal waters of west Japan and HongKong. *Phycologia* **39**, 463-470.

초록 : SSU 부위의 유전자 염기서열 분석에 의한 한국연안에서 분리한 *Cochlodinium polykrikoides* Margalef와 *Gyrodinium aurelum* Hulburt 적조생물의 분자생물학적 연구

조 은 섭*

(국립수산과학원 남해수산연구소)

유해성 적조생물 *Cochlodinium polykrikoides*/*Gyrodinium aurelum*을 포함한 43 종류의 와편조류를 대상으로 SSU 부위 유전자를 분석했다. 유전자 염기서열에 의거한 상호 계통수는 parisomny, distance, maxium 방법으로 실행 했다. *Noctiluca scintillans*, *Gonyaulax spinifera*와 *Crythecodinium cohnii* 종 들은 와편모조류 중 가장 유전적으로 먼 것으로 보였다. *Alexandrium*과 *Symbiodinium* 종 간의 bootstrap는 70% 이상의 상호 단일 계통도를 보인 반면 에, *Gymnodinium*과 *Gyrodinium*은 근립절약계수와 최대 유사도 방법에서 다형 계통도를 나타내었다. *Gyroddinium aurelum*과 *G. dorsum*/*G. galathenum*의 유전적 분화율은 7.4% (45 bp) 였고, *G. aurelum*과 *G. mikimotoi* 상호간에는 0.9% (5bp) 밖에 나타나지 않았다. 또한 *C. polykrikoides*와 *G. aurelum*도 5.2% (31bp)로 낮은 유전적 분화율을 보였 다. 계통도를 분석한 결과 *G. aureolum*과 *C. polykrikoides*는 *Gyrodinium* 보다 *Gymnodinim* 속에 훨씬 더 근접하게 나타났다.