Algae
Volume 18(3): 191-197, 2003

Putative Histone H2A Genes from a Red Alga, *Griffithsia japonica*

Yoo Kyung Lee* and Hong Kum Lee

National Research Lab for Marine Microbial Diversity, Korea Ocean Research and Development Institute, Ansan P.O. Box 29, Seoul 425-600, Korea

Histones are important proteins that interact with the DNA double helix to form nucleosome. Two putative histone genes, *GjH2A-1* and *GjH2A-2* were isolated from a red alga *Griffithsia japonica*. The putative open reading frame of *GjH2A-1* and *GjH2A-2* shared high similarity with the previously reported amino acid sequences of histone H2As. They have a motif consisting of seven amino acids A-G-L-Q-F-P-V, which matches the histone H2A motif [AC]-G-L-x-F-P-V. Phylogenetic trees were constructed from amino acid sequences of 38 histone H2As. The histone H2As were divided into two groups: major H2As and H2A.F/Z variants. The major histone H2A group consisted of animals, fungi, plants + green algae, and red algae H2A subgroups. The animal histone H2A subgroup was divided into vertebrates, echinoderms, nematodes, insects, and segmented worms H2As. The putative red algal histone genes, *GjH2A-1* and *GjH2A-2*, constituted an independent lineage. This is the first report on red algal histone genes.

Key Words: EST(expressed sequence tag), *Griffithsia japonica*, Histone

INTRODUCTION

Histones are small, basic proteins that interact with the DNA double helix to form nucleosomes (Mcghee and Felsenfeld 1980). There are five classes of histone proteins: H1, H2A, H2B, H3, and H4. The four core histones (H2A, H2B, H3 and H4) form an octameric assembly around which 146 base pairs of DNA wraps to form nucleosome (Kornberg and Thomas 1974). Histone H1 binds to the linker DNA found between these nucleosomes, playing a role in the stabilization and formation of higher-order chromatin structure (Noll and Kornberg 1977). The core histones are highly conserved across their entire sequences (Thatcher and Gorovsky 1994), therefore the histone sequences can be used as molecular markers for classification of higher hierarchy.

The total number of histone genes in a species is variable. For example, the genomic sequence of *Saccharomyces cerevisiae* includes two genes for each histone (Goffeau *et al.* 1996). *Aspergillus nidulans* has single genes encoding H2A, H2B, and H3 and two genes encoding slightly different H4 proteins (Ehinger *et al.* 1990). On the other hand, sea urchin has a few hundred copies of histone genes (Ingham and Davis 1988). The histone genes are organized into tandemly repeating quintets of the five histones (Melfi *et al.* 2000), formed into one or two large clusters (Wang *et al.* 1996), or dispersed throughout the genome in small groups of one to few copies (Roberts *et al.* 1987). For most histones are encoded by multigene families (Chaboute *et al.* 1993), classification of histone gene families is important for their taxonomic application as molecular markers.

In this study, we report two putative histone H2A genes of a red alga *Griffithsia japonica*. We also examine phylogenetic relationship of histone H2As and their potentiality as a molecular marker.

MATERIALS AND METHODS

Strains and culture conditions

Griffithsia japonica was collected from Dolsan Island on the southern coast of Korea on 3 April, 1992, and cultured as described in a previous paper (Lee *et al.* 1995). The samples were blotted with paper towel to remove extra moisture, quickly frozen in liquid nitrogen, and stored at –70°C until they were used.

Construction of cDNA library

Total RNA was extracted from *G. japonica* using the

*Corresponding author (leeyk@kordi.re.kr)

Table 1. A list of histone H2A protein sequences analyzed in this study.

Species	GI No.[a]	Group	Phylum	Common Name
Arabidopsis thaliana 1	7595337	Plants	Magnoliophyta	Thale cress
Arabidopsis thaliana 2	4585900	Plants	Magnoliophyta	Thale cress
Arabidopsis thaliana 3	2407800	Plants	Magnoliophyta	Thale cress
Aspergillus niger	2632155	Fungi	Ascomycota	
Bombyx mori	1842152	Animals	Arthropoda	Silkworm
Botryotinia fuckeliana	3242067	Fungi	Ascomycota	
Caenorhabditis elegans 1	2144740	Animals	Nematoda	The elegant worm
Caenorhabditis elegans 2	17541830	Animals	Nematoda	The elegant worm
Chaetopterus variopedatus	2564109	Animals	Annelida	Segmented worm
Chironomus thummi	2118978	Animals	Arthropoda	Midge
Chlamydomonas reinhardtii	2118984	Green algae	Chlorophyta	
Cricetulus longicaudatus	516303	Animals	Chordata	Long-taled hamster
Drosophila melanogaster 1	2144741	Animals	Arthropoda	Fruit fly
Drosophila melanogaster 2	84999	Animals	Arthropoda	Fruit fly
Gallus gallus 1	122000	Animals	Chordata	Chicken
Gallus gallus 2	121988	Animals	Chordata	Chicken
Griffithsia japonica 1	32401019	Red algae	Rhodophyta	
Griffithsia japonica 2	32401021	Red algae	Rhodophyta	
Homo sapiens 1	121986	Animals	Chordata	Human
Homo sapiens 2	11513399	Animals	Chordata	Human
Mus musculus 1	51325	Animals	Chordata	Mouse
Mus musculus 2	1575713	Animals	Chordata	Mouse
Oryctolagus cuniculus	3108211	Animals	Chordata	Rabbit
Oryza sativa	6319146	Plants	Magnoliophyta	Rice
Paracentrotus lividus	1654076	Animals	Echinodermata	Common urchin
Parechinus angulosus	70694	Animals	Echinodermata	Angulated urchin
Pinus taeda	2317760	Plants	Pinophyte	Loblolly pine
Platynereis dumerilii	102609	Animals	Annelida	Dumeril's clam worm
Psammechinus miliaris	85350	Animals	Echinodermata	Sand urchin
Rattus norvegicus 1	56347	Animals	Chordata	Norway rat
Rattus norvegicus 2	92380	Animals	Chordata	Norway rat
Saccharomyces cerevisiae	1582561	Fungi	Ascomycota	Baker's yeast
Schizosaccharomyces pombe 1	70703	Fungi	Ascomycota	Fission yeast
Schizosaccharomyces pombe 2	70704	Fungi	Ascomycota	Fission yeast
Triticum aestivum	536892	Plants	Magnoliophyta	Bread wheat
Urechis caupo	102645	Animals	Annelida	Spoonworm
Volvox carteri	99436	Green algae	Chlorophyta	
Xenopus laevis	1685280	Animals	Chordata	African clawed frog

[a] GenBank information number

TRI reagent (Molecular Research Center, Inc. USA). Poly(A)$^+$ RNA was isolated using the polyATtract mRNA isolation kit (Promega, USA) according to the manufacturer's manual. After electrophoresis of 50 μg mRNA, the fragments with size ≥ 500 bp were purified from the agarose gel. The cDNA library was constructed from the purified mRNA using ZAP-cDNA Gigapack III Gold Cloning Kit (Stratagene, USA) according to the manufacturer's manual.

Excision and sequencing

The cDNA library was excised using ExAssist helper phage (Stratagene, USA). Each monoclone containing *G. japonica* cDNA was cultured in 3 ml LB, and plasmid DNA containing *G. japonica* cDNA was extracted with AccuPrep plasmid Extraction Kit (Bioneer, Korea). The purified plasmid DNA was sequenced with T3 or T7 primers using an ABI PRISM 3100 Genetic Analyzer (Applied Biosystems, USA). Vector masking, contig assembly, and Blast search were performed using

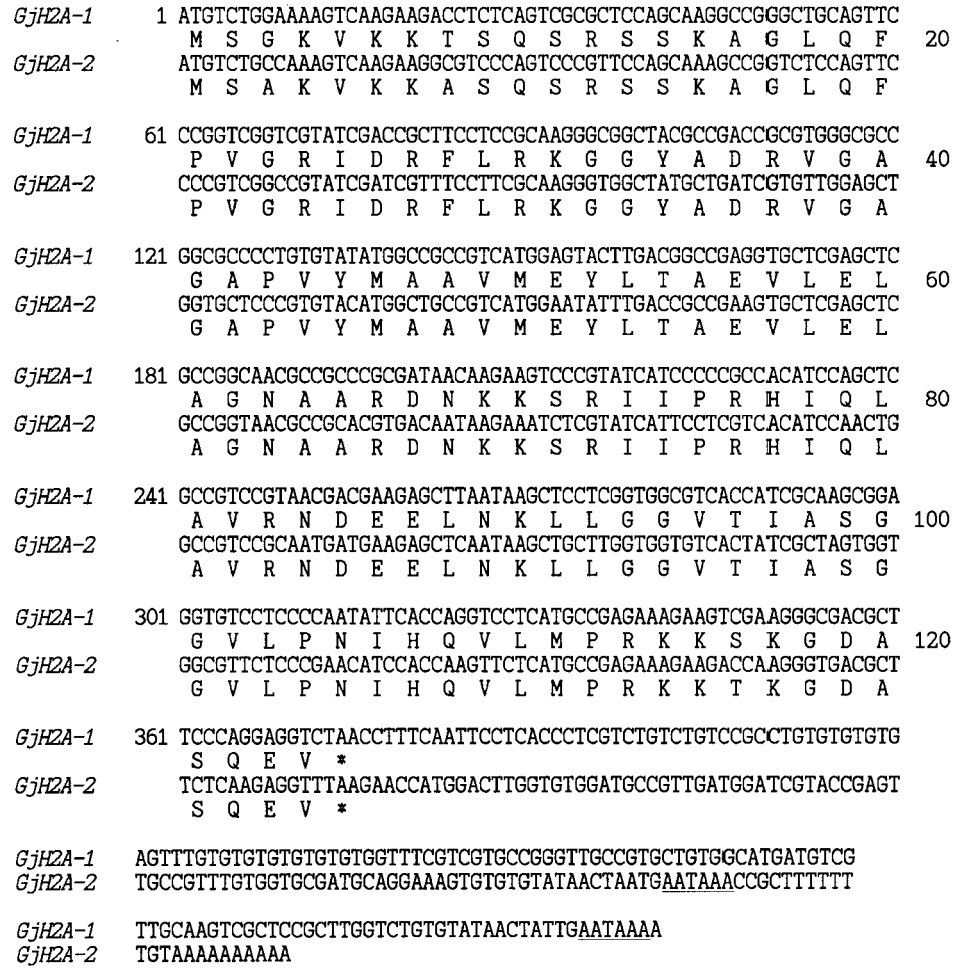

Fig. 1. Nucleotide and deduced amino acid sequences of *GjH2A-1* and *GjH2A-2*. Nucleotide numbers are on the left side and the amino acid numbers are on the right side. Putative termination codons are marked by asterisk(*). Putative polyadenylation signals are underlined.

GeneMaster version 2.0 (Ensoltek, Korea). Two cDNA clones for putative histone H2A genes were selected after random sequencing and similarity analysis.

Phylogenetic analysis

Histone H2A amino acid sequences were obtained from the histone sequence database (Sullivan *et al.* 2002; http://genome.nhgri.nih.gov/histones/). Partial sequences, overlapped sequences, or long sequences were eliminated from the database. Total 38 histone H2A sequences of major eukaryotic organisms were finally selected (Table 1). The selected histone H2A sequences were aligned using CLUSTAL-X (Thompson *et al.* 1997), and final alignment was checked by visual inspection. Phylogenetic analyses were conducted using MEGA version 2.1 (Kumar *et al.* 2001). Phylogenetic trees were reconstructed from the distances using the ML-distance (minimum evolution) method using Poisson correction model. Bootstrap consensus tree was obtained with 1,000 replicates.

GenBank accession numbers

Protein sequences of *G. japonica* histones were deposited in GenBank under the accession numbers AAP80715 (gi 32401019) and AAP80716 (gi 32401021).

RESULTS & DISCUSSION

Two putative histone genes of *G. japonica* were isolated, and designated as *GjH2A-1* and *GjH2A-2*. Nucleotide sequences of *GjH2A-1* and *GjH2A-2*, and their deduced amino acid sequences were determined (Fig. 1). Putative open reading frames for 124 amino acids started from the 5' end of the cDNA. Both *GjH2A-1* and *GjH2A-2* have

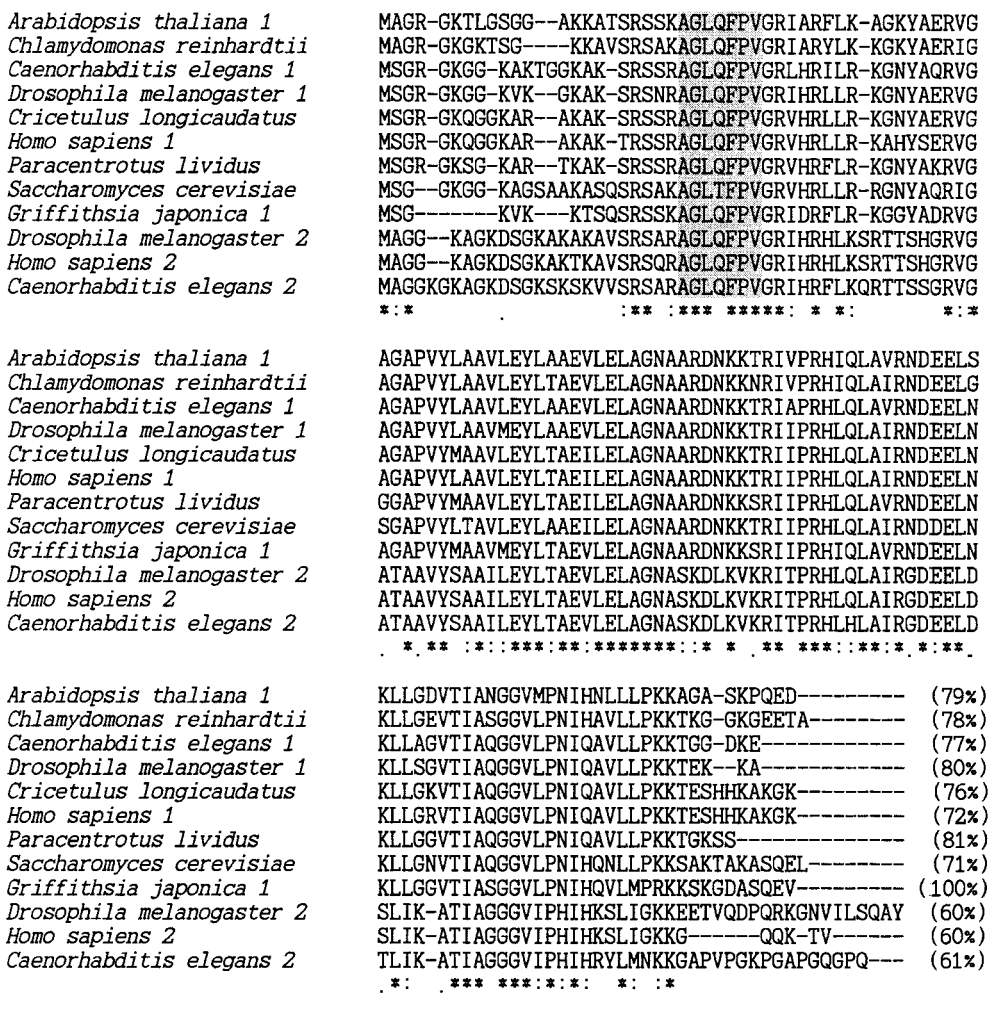

Fig. 2. Sequence alignment of the histone H2As from representative species. Numbers in parentheses are sequence identities with *Griffithsia japonica 1* (*GjH2A-1*). Residues identical to all the sequences are marked by asterisks (*). Residues reflecting conservative changes are indicated by colons (:) and dots (.). Residues indicating the histone H2A motif [AC]-G-L-x-F-P-V are marked with shadow. Gaps were introduced to optimize the alignments, and gaps are represented by dashes (-), which indicates one amino acid.

putative polyadenylation signals with the consensus sequence of AATAAA. They have short 3' UTR (untranslated region), which reported from other *G. japonica* genes (Lee *et al.* 1998, 2002)

G/C content of *GjH2A-1* and *GjH2A-2* was 60.0% and 51.9% in the coding region, which is similar with other *G. japonica* genes previously reported (Lee *et al.* 1998, 2002). In the codon usage, *GjH2A-1* was strongly biased to G and C in the third position: 81.6% of the third position in the codon was G or C, which is similar with other *G. japonica* genes previously reported (Lee *et al.* 1998, 2002). Strong bias toward C or G in the third base of codons was also reported in human and mouse histone genes (Marzluff *et al.* 2002).

The *GjH2A-1* and *GjH2A-2* shared high similarity with the previously reported amino acid sequences of histone H2As: the ORF shared 71-81% amino acid sequence identities to histones of representative animals, plants, fungi and green algae (Fig. 2). They also have a motif consisting of seven amino acids A-G-L-Q-F-P-V, which matches the histone H2A motif [AC]-G-L-x-F-P-V.

Phylogeny shows that the histone H2As are divided into two groups: major histone H2A and histone H2A.F/Z variants (Figs 3-4). These tree topologies agree with previous phylogenetic analysis of histone H2A (Thatcher and Gorovsky 1994). The histone H2A.F/Z variants are more closely related to each other than to the major H2As from the same species (Thatcher and Gorovsky 1994).

The tree topology of major histone H2A except of

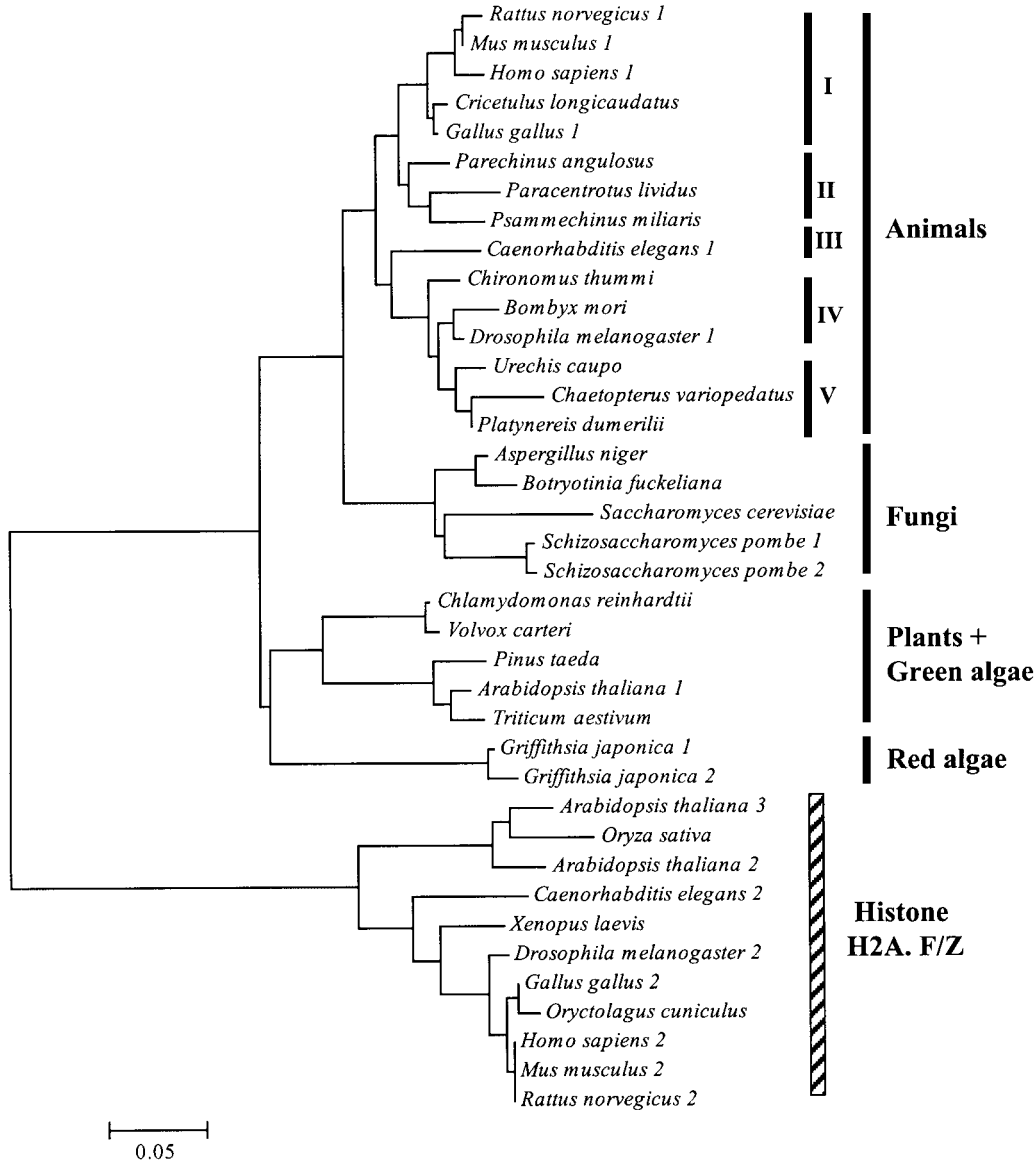

Fig. 3. Phylogeny of histone H2A amino acid sequences from representative species. Thirty-eight histone H2A sequences were selected from histone database, aligned and used to construct a minimum evolution tree (see Materials and Methods). The scale bar indicates the branch length that corresponds to 0.05 substitutions per position. The histone H2As are divided into two monophyletic groups: major H2As and H2A.F/Z variants. The major histone H2As consist of animals, fungal, plants + green algal, and red algal H2A subgroups. The animal histone H2A is divided into vertebrates (I), echinoderms (II), nematodes (III), insects (IV) and segmented worms (V) histone H2As.

H2A.F/Z variants are very similar to a eukaryotic phylogeny based on SSU rDNA sequences (Wainright *et al.* 1993). Animals, plants, fungal and red algal histone H2As form monophyletic group, respectively (Figs 3-4). The animal histone H2A has vertebrates (I), echinoderm (II), nematodes (III), insects (IV) and segmented worms (V) histone H2As (Figs 3-4). The closest relative *GjH2A-1* and *GjH2A-2* of is plants and green algal histone H2A (Fig. 3). But they can align with animal and fungal lineage in other phylogenies (data not shown). Bootstrap consensus tree shows that the *GjH2A-1* and *GjH2A-2* make an independent lineage (Fig. 4). More data on red algal histone H2A sequences will decipher the relationship of red algal H2A and other eukaryotic H2As.

Because the topology of major histone H2A phylogeny is similar to the eukaryotic phylogeny based on SSU rDNA sequences, histone H2A can be used as a molecular marker for classification of higher hierarchy. But for

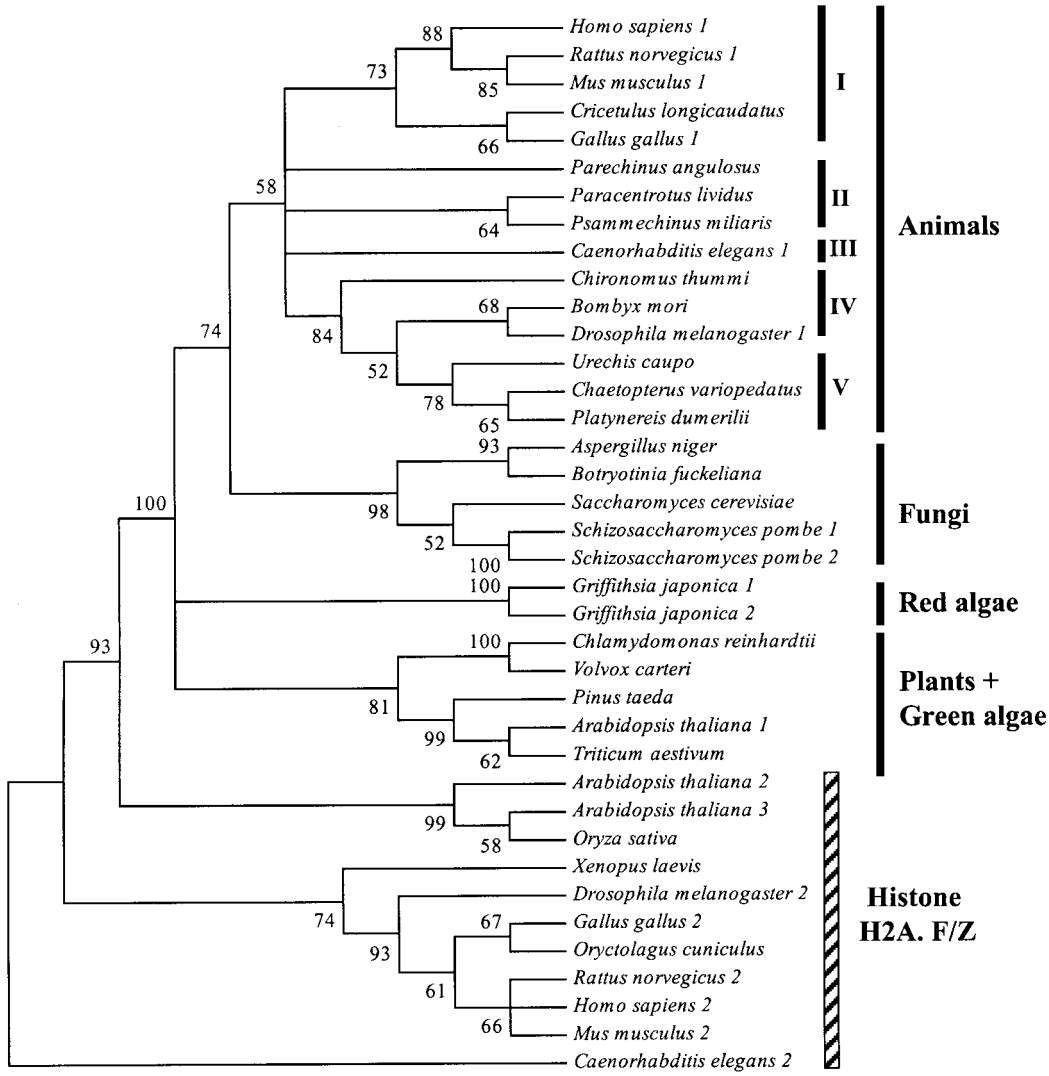

Fig. 4. Bootstrap consensus tree with bootstrap values (% of 1,000 replications) showing relationships between 38 histone H2As with 159 amino acid positions. The major histone H2As consist of animals, fungal, plants + green algal, and red algal H2A groups. The animal histone H2A is divided into vertebrates (I), echinoderms (II), nematodes (III), insects (IV), and segmented worms (V) histone H2As. The histone H2A.F/Z variants consist of animal, plant, and nematode H2A.F/Z. The *GjH2A-1* and *GjH2A-2* make an independent lineage.

histone H2A contains F/Z variants, pseudogenes and macroH2A genes (DeBry 1998; Jiang *et al.* 1998; Pehrson and Fuji 1998), we have to be careful to select genes in the same subfamily to use histone H2A as molecular marker.

ACKNOWLEDGEMENTS

This work was supported by a grant R03-2000-000-00009-0 from the Basic Research Program of the Korea Science and Engineering Foundation (KOSEF). We thank Ensoltek for providing GeneMaster.

REFERENCES

Chaboute M.E., Chaubet N., Gigot C. and Philipps G. 1993. Histones and histone genes in higher plants: structure and genomic organization. *Biochimie*. **75:** 523-531.

DeBry R.W. 1998. Comparative analysis of evolution in a rodent histone H2a. *J. Mol. Evol.* **46:** 355-360.

Ehinger A., Denison S.H. and May G.S. 1990. Sequence, organization and expression of the core histone genes of *Aspergillus nidulans*. *Mol. Gen. Genet.* **222:** 416-424.

Goffeau A., Barrell B.G., Bussey H., Davis R.W., Dujon B., Feldmann H., Galibert F., Hoheisel J.D., Jacq C., Johnston M., Louis E.J., Mewes H.W., Murakami Y., Philippsen P., Tettelin H. and Oliver S.G. 1996. Life with 6000 genes.

Science **274:** 563-567.

Ingham L.D. and Davis F.C. 1988 Cloning and characterization of a core histone gene tandem repeat in *Urechis caupo. Mol. Cell. Biol.* **8:** 4425-4432.

Jiang W., Guo X. and Bhavanandan V.P. 1998. Histone H2A.F/Z subfamily: the smallest member and the signature sequence. *Biochem. Biophys. Res. Commun.* **245:** 613-617.

Kornberg R. and Thomas J.O. 1974. Chromatin structure; oligomers of the histones. *Science* **184:** 865-868.

Kumar S., Tamura K., Jakobsen I.B. and Nei M. 2001. MEGA2: molecular evolutionary genetics analysis software. *Bioinformatics* **17:** 1244-1245.

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

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

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

Marzluff W.F., Gongidi P., Woods K.R., Jin J. and Maltais L.J. 2002 The human and mouse replication-dependent histone genes. *Genomics* **80:** 487-498.

Mcghee J.D. and Felsenfeld G. 1980. Nucleosome structure. *Annu. Rev. Biochem.* **49:** 1115-1156.

Melfi R., Palla F., Di Simone P., Alessandro C., Cali L., Anello L. and Spinelli G. 2000. Functional characterization of the enhancer blocking element of the sea urchin early histone gene cluster reveals insulator properties and three essential cis-acting sequences. *J. Mol. Biol.* **304:** 753-763.

Noll M. and Kornberg R.D. 1977. Action of micrococcal nuclease on chromatin and the location of histone H1. *J. Mol. Biol.* **109:** 393-404.

Pehrson J.R. and Fuji R.N. 1998. Evolutionary conservation of histone macroH2A subtypes and domains. *Nucleic Acids Res.* **26:** 2837-2842.

Roberts S.B., Sanicola M., Emmons S.W. and Childs G. 1987. Molecular characterization of the histone gene family of *Caenorhabditis elegans. J. Mol. Biol.* **196:** 27-38.

Sullivan S., Sink D.W., Trout K.L., Makalowska I., Taylor P.M., Baxevanis A.D. and Landsman D. 2002. The Histone Database. *Nucleic Acids Res.* **30:** 341-342.

Tatusov R.L., Altschul S.F. and Koonin E.V. 1994. Detection of conserved segments in proteins: iterative scanning of sequence databases with alignment blocks. *Proc. Natl. Acad. Sci. USA*, **91:** 12091-12095.

Thatcher T.H. and Gorovsky M.A. 1994. Phylogenetic analysis of the core histones H2A, H2B, H3, and H4. *Nucleic Acids Res.* **22:** 174-179.

Thompson, J.D., Gibson T.J. Plewniak F. Jeanmougin F. and Higgins. D.G. 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **24:** 4876-4882.

Wang Z.F., Krasikov T., Frey M.R., Wang J., Matera A.G. and Marzluff W.F. 1996. Characterization of the mouse histone gene cluster on chromosome 13: 45 histone genes in three patches spread over 1Mb. *Genome Res.* **6:** 688-701.

Wainright P.O., Hinkle G., Sogin M.L. and Stickel S.K. 1993. Monophyletic origins of the metazoa: an evolutionary link with fungi. *Science* **260:** 340-342.

Received 6 September 2003
Accepted 22 September 2003