J Mol Evol (1998) 47:378–384

JOURNAL OF MOLECULAR EVOLUTION

Directionally Evolving Genetic Code: The UGA Codon from Stop to Tryptophan in Mitochondria

Yuji Inagaki,[1,*] Megumi Ehara,[2] Kazuo I. Watanabe,[2] Yasuko Hayashi-Ishimaru,[1,] Takeshi Ohama[1]**

[1] JT Biohistory Research Hall, 1-1 Murasaki-cho, Takatsuki, Osaka 569-1125, Japan
[2] Department of Biology, Faculty of Science, Osaka University, Machikaneyama-cho, Toyonaka, Osaka 560-0043, Japan

Received: 26 January 1998 / Accepted: 24 April 1998

Abstract. For the comprehensive analyses of deviant codes in protistan mitochondria (mt), we sequenced about a 1.1-kb region of a mitochondrial (mt) gene, the cytochrome *c* oxidase subunit I (*coxI*) in two chlorarachniophytes, the filose amoeba *Euglypha rotunda*, the cryptomonad *Cryptomonas ovata*, the prymnesiophyte (haptophyte) *Diacronema vlkianum* (Pavlovales), and the diatom Melosira ambigua. As a result of this analysis, we noticed that the UGA codon is assigned to tryptophan (Trp) instead of being a signal for translational termination in two chlorarachniophytes and in *E. rotunda*. The same type of deviant code was reported previously in animals, fungi, ciliates, kinetoplastids, *Chondrus crispus* (a red alga), *Acanthamoeba castellanii* (an amoeboid protozoon), and three of the four prymnesiophyte orders with the exception of the Pavlovales. A phylogenetic analysis based on the COXI sequences of 56 eukaryotes indicated that the organisms bearing the modified code, UGA for Trp, are not monophyletic. Based on these studies, we propose that the ancestral mitochondrion was bearing the universal genetic code and subsequently reassigned the codon to Trp independently, at least in the lineage of ciliates, kinetoplastids, rhodophytes, prymnesiophytes, and fungi. We also discuss how this codon was directionally captured by Trp tRNA.

Key words: COXI phylogenetic tree — UGA codon — Deviant genetic code — Directional codon reassignment

Introduction

A wide variety of metazoan mitochondrial (mt) genomes has been completely sequenced and many types of deviant genetic codes have been reported (for a review, see Wolstenholme 1992). Of especial interest is the codon UGA, which usually serves as a translational termination signal (stop), but which is used for Trp in the mt genomes of animals, fungi, ciliates, and kinetoplastids (reviewed by Osawa et al. 1992). Recent studies on protistan mitochondria (mt) revealed the same kind of reassignment for the UGA codon in *Acanthamoeba* (*Ac.*) *castellanii* (an amoeboid protozoon) (Burger et al. 1995), *Chondrus* (*Cho.*) *crispus* (Rhodophyta/red alga) (Boyen et al. 1994), and prymnesiophytes (haptophytes) (Hayashi-Ishimaru et al. 1997). It has been reported that the codon is utilized as a stop codon in land plants (e.g., Oda et al. 1992), a phaeophyte (brown alga) (Fontaine et al. 1995), a rhodophyte (Viehmann et al. 1996), an oomycete (see the public database site, http://megasun.bch.umontreal.ca/People/lang/FMGP/), and a freshwater heterotrophic flagellate, *Reclinomonas americana* (Lang et al. 1997). Curiously, the UGA codon was not detected throughout the mt genomes of apicomplexans (Feagin et al. 1992; Kairo et al. 1994), chlorophytes (green algae) (Boer et al. 1985; Feagin et al. 1992; Wolff et al. 1994),

* *Present address:* Department of Biochemistry, Dalhause University, Halifax NS, B3H 4H7, Canada
** *Present address:* Department of Cell Biology, National Institute for Basic Biology, Okazaki, Aichi 444-0867, Japan
Correspondence to: Takeshi Ohama; *e-mail:* Takeshi.Ohama@ims.brh.co.jp

and an ancestral fungus (Paquin et al. 1996). In these genomes, no tRNA to translate the UGA codon is encoded.

To change the assignment of a codon from a stop to a sense, it is obvious that peptide chain release factor (RF) must be altered so as not to recognize the stop codon, and a tRNA must appear to recognize the codon. This process can be well explained by the codon capture theory, which is composed of three subsequent steps (Osawa et al. 1992). The initial step is the disappearance of a specific stop codon through its replacement with synonymous codons throughout the genome. Secondarily, the corresponding RF loses its ability to recognize the codon. This does not have any unfavorable effect, as the codon is absent from the genome. In this context, the codon is referred to as ''unassigned,'' as it is recognized neither by RFs nor by tRNAs. Finally, a mutated tRNA that can decode the unassigned codon appears and converts the unassigned codon into a sense codon. This theory does not allow an ambiguity on codon assignment, while recently Schultz and Yarus (1994, 1996) proposed another another theory which includes an intermediate state where a codon is read more than one way (ambiguous intermediate theory).

In this study, we detected the modified genetic code, UGA for Trp, in the sequenced region of the cytochrome *c* oxidase subunit I (COXI) gene of two chlorarachniophytes and a filose amoeba. Based on the COXI phylogenetic tree, we propose the independent occurrence of the same codon reassignment in several mt lineages and discuss why the UGA stop codon is modified specifically for Trp.

Materials and Methods

Biological Materials

Protists were purchased from the following culture stock centers: National Institute for Environmental Studies (NIES; Japan), Center for Culture of Marine Phytoplankton (CCMP; USA), American Type Culture Collection (ATCC; USA), and Culture Collection of Algae and Protozoa (CCAP; UK). The strains used in this study are as follows: *Crypthecodinium* (*Crypth.*) *cohnii* sibling species I (ATCC 30541), *Prorocentrum micans* (NIES 12), *Chlorella* (*Chlore.*) *reisiglii* (CCAP 211/59), *Chlore. vulgaris* (NIES 227), *Hydrodictyon reticulatum* (Kyoto University, Kyoto, Japan), *Chlorarachnion* (*Chlora.*) sp. (CCMP 238), *Chlora. reptans* (CCMP 240), *Euglypha rotunda* (CCAP 1520/1), *Cryptomonas* (*Crypto.*) *ovata* (NIES 274), *Diacronema vlkianum* (CCAP 914/1), and *Melosira ambigua* (NIES 20).

Polymerase Chain Reaction (PCR) and Sequence Analysis

Total DNA was extracted by SDS/proteinase K treatment followed by phenol/chloroform extraction and ethanol precipitation as described previously (Hayashi-Ishimaru et al. 1996). The obtained crude DNA was used as a template for PCR.

To amplify by PCR the 1.1-kb region of the COXI gene of *Chlora.* sp., *Chlora. repentance, Crypto. ovata, D. vlkianum,* and *M. ambigua,* the following set of primers was used: p1C (5′-TGGTTNTTYTCNACNAAYCAYAARGAYAT-3′; N—A, C, G, or T; Y—T or C; R—A or G) and CX1AS1 (5′-AARTGIGCIACIACRTARTAIGTRTCRTG-3′; I, inosine). For the extension of sequenced regions up to 1.1 kb in *Chlore. vulgaris, Chlore. reisiglii,* and *H. reticulatum,* we used p1C and each of the specific primers whose sequences are based on the reported DNA sequence (Hayashi-Ishimaru et al. 1996). The reaction was performed in 30 μl of a reaction mixture containing a 2.5 m*M* concentration of each deoxyribonucleoside, a 1 μ*M* concentration of the set of primers, and 2 units of Ex Taq DNA polymerase (Takara shuzo, Japan) by applying 30 cycles of 15 s at 94°C, 1 min at 40°C, and 1 min at 72°C. To expand the sequenced region of the COXI gene for *Crypth. cohnii* and *P. micans,* the adapter-mediated PCR walking method was adopted using the LA PCR in vitro cloning kit (Takara shuzo, Japan). The PCR product was cloned into the pT7 Blue T-vector (Novagen, USA). DNA sequencing was carried out on at least five clones by the dye-terminator cycle sequencing method using a DNA sequencer, Model 377 (Applied Biosystems, USA).

Phylogenetic Analysis

Fifty-seven amino acid sequences of COXI were aligned using CLUSTAL W (Higgins et al. 1988) and subsequently edited by eye. Floating extra amino acid sequences peculiar to the COXI gene of two ciliates (total, 107 residues) were omitted because it was difficult to evaluate their information. Phylogenetic trees based on this alignment (365 sites) were inferred using maximum-likelihood (ML), distance, and parsimony methods. The ML analysis was conducted by the star decomposition search using the JTT transition probability matrix in the PROTML program from the MOLPHY Version 2.2 package (Adachi et al. 1992). Evolutionary distances were calculated by referring to the Dayhoff (1978) PAM250 substitution matrix. A phylogenetic tree based on the distance matrix was constructed by the neighbor-joining method (Saitou and Nei 1987) included in the computer program SINCA (Version 3.0) (Fujitsu System Engineering, Japan). Unweighted parsimony trees were found using the PROTPARS program from the PHYLIP Version 3.5c package (Felsenstein 1993), and the consensus branching pattern among those trees was extracted. Bootstrap resampling (100 times) (Felsenstein 1985) was carried out to quantify the relative support for branches of the inferred trees.

Codon Usage of UGA

For the 52 mt genomes that have been sequenced completely and found to bear two Trp codons (UGA and UGG), we analyzed the correlation between the preference for the UGA Trp codon and the GC content of the third nucleoside of codons (codon 3rd GC%).

For 72 bacteria and 18 chloroplasts, we investigated the correlation between the usage of stop codons and the third codon GC%. For these analyses, the Codon Usage Tabulate from Genbank was utilized (Nakamura et al. 1997).

Results

Mitochondrial Genetic Code

In this study, no deviant genetic code was detected in the coding region of two chlorophytes (*Chlore. reisiglii* and *Chlore. vulgaris*) and two dinoflagellates (*Cryphe. coh-*

Table 1. The UGA codon at conserved Trp sites in chlorarachniophytes and *E. rotunda*

	Position No.[a]									
Species	9	69	91	105	114	175	225	276	311	316
Planta/Chlorophyceae										
Chlore. vulgaris	A	W[b]	W	L	W	W	W	W	W	W
Chlore. reisiglii	A	W	W	L	W	W	W	W	W	W
Dinoflagellata										
Crypth. cohnii	F	Y	L	I	W	W	W	W	W	L
Pror. micans	F	Y	L	I	W	W	W	W	W	L
Cryptophyta										
Crypto. ovata	I	W	W	L	W	W	W	W	W	W
Heterokonta/Bacillariophyceae										
M. ambigua	I	W	W	L	W	W	W	W	W	W
Chlorarachniophyta										
Chlorarachnion sp.	A	TGA	TGA	L	TGA	TGA	TGA	TGA	TGA	TGA
Chlora. reptans	A	W	W	L	TGA	W	W	W	TGA	TGA
A filose amoeba										
E. rotunda	TGA	L	W	TGA	W	W	TGA	TGA	TGA	W

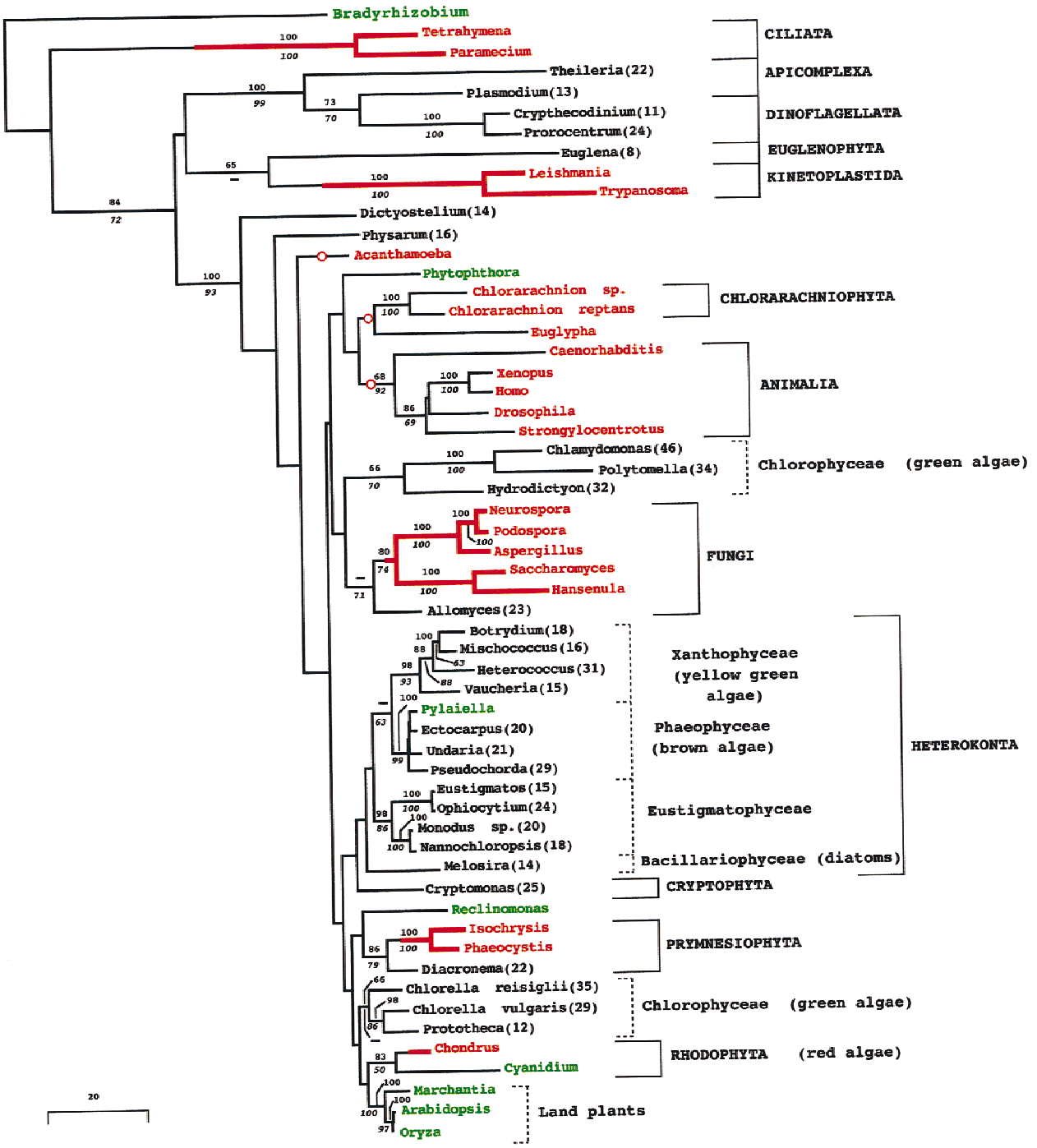
[a] The deduced amino acid sequence was numbered based on that of *Chlore. vulgaris.*
[b] ''W'' is coded by UGG.

nii and *P. micans*). In the coding frame for a chlorophyte *H. reticulatum,* the UAG codon appeared at conserved alanine sites (data not shown). These were in agreement with the results presented previously (Hayashi-Ishimaru et al. 1996; Inagaki et al. 1997). We newly determined an approximately 1.1-kb region of the COXI gene of the cryptophyte *Crypto. ovata,* the prymnesiophyte *D. vlkianum,* the bacillariophyte *M. ambigua,* two chlorarachniophytes (*Chlorarachnion* sp. and *Chlora. reptans*), and the filose amoeba *E. rotunda.* No deviant code was observed in *Crypto. ovata, D. vlkianum,* or *M. ambigua,* whereas we detected UGA codons at conserved Trp sites in the coding frame of two chlorarachniophytes and *E. rotunda* (Table 1). In the two chlorarachniophytes, one UGA codon was also detected in the partially sequenced region of the COXII gene (data not shown). From *Chlora. reptans,* we obtained a PCR product showing a high homology to prokaryotic Trp tRNA ($tRNA^{Trp}$). The anticodon sequence of this $tRNA^{Trp}$ was UCA, which allows it to translate the UGA codon (data not shown).

Phylogenetic Trees Based on the COXI Sequences

A phylogenetic tree based on the deduced amino acid sequences of the COXI gene from 56 eukaryotes was constructed by the ML method (Fig. 1). The phyloge-

Fig. 1. Phylogenetic tree based on the deduced amino acid sequence of COXI (365 positions). The tree was constructed by the maximum-likelihood method (Adachi et al. 1992). The cytochrome oxidase subunit I sequence of the bacterium *Bradyrhizobium japonicum* (Genbank accession No. X68547) was used as the outgroup. Bootstrap values above 50% are indicated. The values indicated in *roman* characters were based on the distance method and those indicated in *italics* were based on the parsimony method. Species shown in *red* and *green* use the UGA codon for stop and Trp, respectively. Species in which no UGA codon was detected so far are shown in *black.* Their GC contents of the third nucleoside of codons are indicated in *parentheses.* The five lineages which are strongly suggested to have experienced a change in the assignment of the UGA codon, from stop to Trp, are indicated by *thick red lines. Red circles* indicate the points where reassignments of the UGA codon were supposed to have occurred. The species and Genbank accession numbers used in this study are as follows. Ciliata: *Tetrahymena pyriformis* (Accession No. X06133), *Paramecium tetraurelia* (M15281). Kinetoplastida: *Trypanosoma brucei* (X01094), *Leishmania tarentolae* (M10126). Euglenophyta: *Euglena gracilis* (U49052). Apicomplexa: *Plasmodium falciparum* (M99416), *Theileria parva* (Z23263). Dinoflagellata: *Crypthecodinium cohnii* (AB000122), *Prorocentrum micans* (AB000133). Chlorophyceae: *Chlorella reisiglii* (AB009363), *Chlorella vulgaris* (D63763), *Hydrodictyon reticulatum* (AB009364), *Chlamydomonas reinhardtii* (U03843), *Polytomella* sp. (U31972), *Prototheca wickerhamii* (U02970). Eustigmatophyceae: *Eustigmatos magnus* (AB000205), *Nannochloropsis oculata* (AB000209), *Mondodus* sp. (AB000207), *Ophiocytium majus* (AB000210). Xanthophyceae: *Botrydium granulatum* (AB000204), *Heterococcus caespitosus* (AB000206), *Mischococcus sphaerocephalus* (AB0002208), *Vaucheria sessilis* (AB000212). Prymnesiophyceae: *Isochrysis galbana* (AB000119), *Phaeocystis pouchetii* (AB000120), *Diacronema vlkianum* (AB009420). Rhodophyta: *Chondrus crispus* (Z47547), *Cyanidium caldarium* (Z48930). Phaeophyceae: *Pylaiella littoralis* (Z72500), *Pseudochorda nagaii* (AF037992), *Ectocarpus* sp. (AF037994), *Undaria pinnatifida* (AF037993). Cryptophyta: *Cryptomonas ovata* (AB009419). Bacillariophyceae: *Melosira ambigua* (AB009418). Fungi: *Allomyces macrogynus* (U41288), *Saccharomyces cerevisiae* (J01481), *Hansenula wingei* (D31785), *Aspergillus nidulans* (X00790), *Podospora anserina* (X55026), *Neurospora crassa* (X01850). Animalia: *Caenorhabditis elegans* (X54252), *Strongylocentrotus purpuratus* (X12631), *Drosophila melanogaster* (U37541), *Xenopus laevis* (M10217), *Homo sapiens* (J01415). Land plants: *Marchantia polymorpha* (M68929), *Arabidopsis thaliana* (Y08502), *Oryza sativa* (M57903). Chlorarachniophyta: *Chlorarachnion* sp. (AB009396), *Chlorarachnion reptans* (AB009416). Other protists: *Acanthamoeba castellanii* (U12386), *Dictyostelium discoideum* (D50297), *Physarum polycephalum* (L14769), *Euglypha rotunda* (AB009417), *Reclinomonas americana* (AF007261). The amino acid sequence for the oomycete *Phytophthora infestans* was obtained from the public database site of the Fungal Mitochondrial Genome Project at Université de Montréal (http://megasun.bch.umontreal.ca/People/lang/FMGP/).



netic analyses were also conducted using the distance and parsimony methods. The results from the three methods agreed well (data not shown). Each of the robust clades in the COXI tree basically corresponded to a taxon, phylum, or class. In our COXI tree, the cluster of two ciliates was separated from the rest of eukaryotes with a high statistical probability. *Euglena gracilis* and the cluster of kinetoplastids were connected. This evolutionary connection of *E. gracilis* and kinetoplastids has been supported with sufficient reliance in the small ribosomal RNA (SSU rRNA) tree (e.g., Cavalier-Smith 1993) and the hsp60 tree (e.g., Yasuhira and Simpson 1997). *Ac. castellanii* showed no close relationship to any other organism. A clade of the two chlorarachniophytes and a filose amoeba *E. rotunda* were connected (Fig. 1). This relationship has been confirmed by the SSU rRNA tree (Bhattacharya et al. 1995). Animals were linked to the cluster of chlorarachniophytes/*E. rotunda* in our COXI tree, however, animals show a close affinity to fungi in recent phylogenetic trees based on the SSU

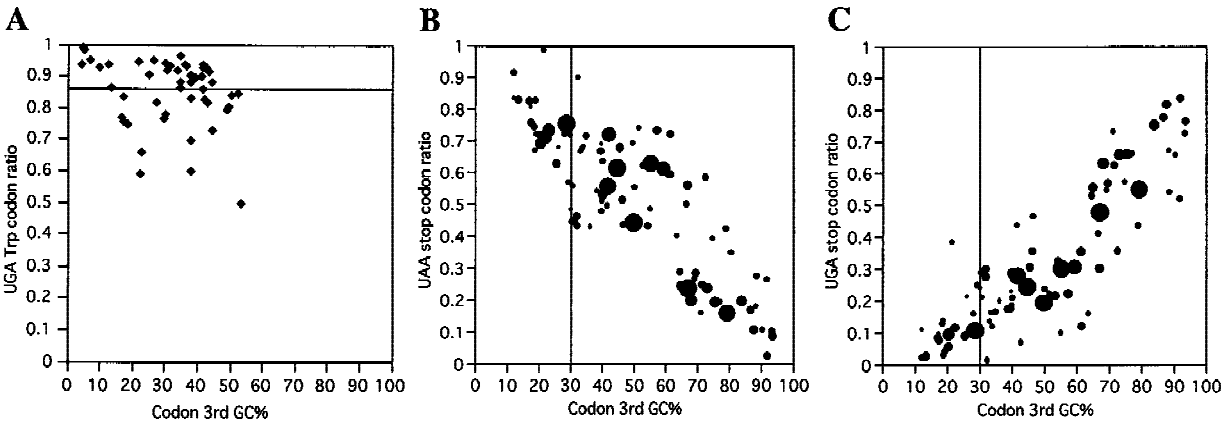


Fig. 2. **A** Correlation between the UGA Trp codon ratio [UGA/(UGA+UGG)] and the GC content of the third nucleoside of codons (3rd GC%) in the complete sequences of 51 mitochondrial (mt) genomes. The *line* shows the average UGA Trp codon ratio (ca. 5/1) among these mitochondria. **B** Correlation between the UAA stop codon ratio [UAA/(UAA+UAG+UGA)] and the codon 3rd GC%. **C** Correlation between the UGA stop codon ratio [UGA/(UAA+UAG+UGA)] and the codon 3rd GC%. In B and C, each ratio was calculated for 72 bacteria and 18 chloroplasts, in which more than 50 genes have been reported, using the Codon Usage Tabulate from Genbank (CUTG) (Nakamura et al. 1997). Each *dot* indicates the ratio of a bacterium or chloroplast. The diameter of the dot is proportional to the number of analyzed three-stop codons. The *vertical line* indicates the average codon 3rd GC% (30%) of 89 mitochondria listed in the CUTG (Nakamura et al. 1997).

rRNA (e.g., Cavalier-Smith 1993) and multiple proteins (e.g., Baldauf et al. 1993). The prymnesiophyte, *D. vlkianum* was strongly tied to other prymnesiophytes (Fig. 1). Two rhodophytes, *Cho. crispus* and *Cyanidium caldarium*, made a monophyletic cluster. In the fungal clade, *Al. macrogynus* was judged to be the organism that diverged earlier than other fungi (Fig. 1). This has been confirmed by the phylogenetic study based on the NADH dehydrogenase subunit 5 sequences (Paquin et al. 1995).

The names of organisms shown in red and green (Fig. 1) correspond to species bearing the UGA codon for Trp and stop, respectively. For the names of species shown in black (Fig. 1), no UGA codon has so far been detected in the sequenced region.

Codon Usage of UGA as a Trp or Stop Codon

In the mt that use the UGA codon for Trp, the codon is utilized more frequently than the cognate UGG Trp codon without substantial effects of the third codon GC%. Among these, the average codon ratio (UGA/UGG) was ca. 5/1 (Fig. 2A).

The usage analyses for UAA and UGA stop codons in bacteria and chloroplasts clearly indicated that the choice of these stop codons is directly affected by the third codon GC%. When the third codon GC% is low, the UAA stop codon is dominant, while the UGA stop codon is rare (Figs. 2B and C). The choice pattern for the stop codons, UAA or UGA, is completely reversed when the third codon GC% is high. The frequency of the UAG stop codon was also investigated but no obvious relationship with the third codon GC% was found (data not shown).

Discussion

UGA Assignment in mt: Trp, Stop, or ''Unassigned''

In this study, we found UGA codons at conserved Trp sites in the sequenced region of *Chlora.* sp., *Chlora. reptans*, and *E. rotunda* (Table 1). Moreover, evidence for the existence of $tRNA^{Trp}$ with the anticodon UCA ($tRNA^{Trp}_{UCA}$) was found in *Chlora. reptans* (data not shown). Considering these data, we conclude that the UGA codon is assigned for Trp in those mt genomes. We suppose that $tRNA^{Trp}_{UCA}$ appeared through duplication of a $tRNA^{Trp}_{CCA}$ gene and a following mutation which replaced the anticodon CCA to UCA or through a series of mutations; i.e., a mutation that facilitated the irregular base pairing between UGA codon and CCA anticodon (see below) was followed by another mutation which replaced the anticodon CCA to UCA.

No UGA codon was detected in the five analyzed Trp sites of *Crypth. cohnii* and *P. micans* and the eight sits of *Chlore. reisiglii*, *Chlore. vulgaris*, *Crypto. ovata*, *D. vlkianum*, and *M. ambigua* (Table 1). Through the usage analysis of cognate Trp codons, UGA and UGG, in 52 mt genomes, the ratio of Trp codon (UGA/UGG) was expected to be ca. 5/1 (Fig. 2A). This suggests that if the codon is assigned for Trp, at least one UGA codon appears in five Trp sites with the remarkably high probability of $1 - (1/6)^5$. Thus, it is highly unlikely that no UGA Trp codon is accidentally used in the analyzed region. Hence, we concluded that UGA is not assigned for Trp in these organisms. Applying the same reasoning to the 56 eukaryotes shown in Fig. 1, UGA would not be assigned for Trp in 28 lineages in which no UGA has been detected so far (shown in black in Fig. 1).

When the UGA codon is not assigned for Trp, what is the codon used for? The average third codon GC% in 89 mt listed in the Codon Usage Tabulate from Genbank (Nakamura et al. 1997) was 30%. Under these conditions of GC content, the UAA stop codon is utilized much more frequently, and UGA is a rare stop codon (Figs. 2B and C). Considering the limited number of protein-coding genes in the mt genome [the known maximum number is 67 in *R. americana* (Lang et al. 1997)] and the deduced usage of the UGA stop codon (Fig. 2C), one possibility is that it is an active stop codon but rarely (or not) used in these genomes. Another possibility is that it is an unassigned codon. To determine whether UGA is a latent stop codon or unassigned, the assessment of the recognition ability of RFs will be essential.

Evolution of the UGA Assignment

In our COXI tree, the lineages bearing the deviant code, UGA for Trp (shown in red in Fig. 1), were not monophyletic and were located among those using the UGA codon for stop or unassigned (shown in green or black in Fig. 1). Supposing that the mt of ancestral eukaryotes used the devalant code, UGA for Trp, the assignment must be changed from Trp to stop or to unassigned in multiple lineages (species shown in green or black in Fig. 1). As long as the GC content of mt genome is kept rather low, as in the present mt, UGA Trp codon is strongly preferred (Fig. 2A). In this context, sudden conversion of the UGA assignment from Trp to stop or to unassigned obviously leads to the synthesis of immature peptides. To avoid such a lethal event, it seems essential that the UGA Trp codon is completely replaced by the UGG codon or, at least, the number is markedly reduced. However, so far there is no proved or examined driving force that reduces the UGA Trp codon (but not in the case of the UGA stop; see below) in the AT rich mt genome. Hence, we do not discuss the probability of this scenario further.

Here we propose that the ancestral mt used the universal code, and the assignment changed independently in various mt lineages. The UGA stop codon is rarely used as long as the third codon GC% is rather low, as is the case in the present mt (the average is 30%) (Fig. 2C) and the gene content is extremely limited as mentioned above. In this context, it seems probable that the UGA stop codon could disappear from the whole genome even by chance or through an ambiguous intermediate step in which the remaining UGA stop codons have two assignments simultaneously, not only as a stop but also as a sense; i.e., the UGA codon is recognized by a relase factor and a tRNA (Schultz and Yarus 1994, 1996).

This scenario brings up another question. Why was this codon always captured by $tRNA^{Trp}$? Usually the tRNA that recognizes the deviant code possesses a mutation in the anticodon triplet to permit a regular base-pairing with the codon. $tRNA^{Trp}_{UCA}$, which can translate the UGA codon, has been detected in many mt bearing the deviant code, UGA for Trp (reviewed by Osawa et al. 1992). However, only in $tRNA^{Trp}_{CCA}$, a mutation(s) outside the anticodon most probably leads to the structural changes that facilitate irregular base-pairing between the third nucleoside of the codon, adenosine (A), and the first nucleoside of the anticodon, cytosine, as exemplified below. In the mt of *Ac. castellanii*, it is supposed that UGA is regularly translated by the bizarre $tRNA^{Trp}_{CCA}$ including a U–U mismatch in the anticodon stem (Burger et al., 1995). Another example of UGA recognition by mt $tRNA^{Trp}_{CCA}$ has been reported in a fungus, *Schizophyllum commune* (Paquin et al. 1997). In *Escherichia coli*, the opal (UGA) suppresser $tRNA^{Trp}_{CCA}$, $Su^{+}9$, has one point mutation, guanosine to A, in the D stem (Hirsh and Gold 1971). Therefore, in $tRNA^{Trp}_{CCA}$, the number of potential mutation sites that can result in the ability to decode the UGA codon seems much larger compared to other tRNAs, and this feature seems the most likely reason that UGA is reassigned for Trp specifically.

Here we propose that the assignment change of the UGA codon, from stop to Trp, occurred independently in the lineages of ciliates, kinetoplastids, prymnesiophytes, rhodophytes, and fungi (indicated by thick red lines, Fig. 1). With reference to recent SSU rRNA (e.g., Cavalier-Smith 1993; De Peer et al. 1996) and multiple protein trees (e.g., Baldauf et al. 1993), *Ac. castellanii*, animals, and chlorarachniophytes/*E. rotunda* also seem to have independently experienced the same assignment change (the expected points of reassignment are shown by red circles in Fig. 1).

Acknowledgments. Y.I. and M.E. were supported by Research Fellowships of the Japanese Society for the Promotion of Science for Young Scientists. A visiting associate professor position was arranged for T.O. at Osaka University in order to permit collaboration with postgraduates and use of the various facilities.

References

Adachi J, Hasegawa M (1992) MOLPHY, program for molecular phylogenetics. I. PROTML, maximum likelihood inference of protein phylogeny (Inst. Stat. Math., Tokyo), Version 2.2, Computer Science Monographs No. 27

Baldauf SL, Palmer JD (1993) Animals and fungi are each other's closest relatives: Congruent evidence from multiple proteins. Proc Natl Acad Sci USA 90:11558–11562

Bhattacharya D, Helmchen T, Melkonian M (1995) Molecular evolutionary analyses of nuclear-encoded small subunit ribosomal RNA identify an independent rhizopod lineage containing the Euglyphina and the Chlorarachniophyta. J Euk Moicrobiol 42:65–69

Boer PH, Bonen L, Lee RW, Gray NW (1985) Genes for respiratory chain proteins and ribosomal RNAs are present on a 16-kilobase-pair DNA species from *Chlamydomonas reinhardtii* mitochondria. Proc Natl Acad Sci USA 82:3340–3344

Boyen C, Leblanc C, Bonnard G, Grienenberger JM, Kloareg B (1994) Nucleotide sequence of the cox3 gene from *Chondrus crispus*: Evidence that UGA encodes tryptophan and evolutionary implications. Nucleic Acids Res 22:1400–1403

Burger G, Plante I, Lonergan KM, Gray MW (1995) The mitochondrial DNA of the amoeboid protozoon, *Acanthamoeba castellanii*: complete sequence, gene content and genome organization. J Mol Biol 245:522–537

Cavalier-Smith T (1993) Kingdom protozoa and its 18 phyla. Microbiol Rev 57:953–994

Dayhoff MO (1978) A model of evolutionary change in proteins. Matrices for detecting distant relationships. In: Dayhoff MO (ed) Atlas of protein sequence and structure. National Biomedical Research Foundation, Washington DC, Vol 5, Suppl 3, pp 345–358

De Peer V, Rensing SA, Maier UG, De Wachter R (1996) Substitution rate calibration of small subunit ribosomal RNA identifies chlorarachniophyte endosymbionts as remnants of green algae. Proc Natl Acad Sci USA 93:7732–7736

Feagin JE, Werner E, Gardner MJ, Williamson DH, Wilson RJ (1992) Homologies between the contiguous and fragmented rRNAs of the two *Plasmodium falciparum* extrachromosomal DNAs are limited to core sequences. Nucleic Acids Res 20:879–887

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39:783–791

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

Fontaine JM, Rousvoal S, Delaroque N, Loiseaux-de Goer S (1995) Characterization of the cox3, nad7 and atp6 genes from the mitochondrial genome of the brown alga *Pylaiella littoralis.* Plant Physiol Biochem 33:605–609

Hayashi-Ishimaru Y, Ohama T, Kawatsu Y, Nakamura K, Osawa S (1996) UAG is a sense codon in several chlorophycean mitochondria. Curr Genet 30:29–33

Hayashi-Ishimaru Y, Ehara M, Inagaki Y, Ohama T (1997) A deviant genetic code in prymnesiophytes (yellow-algae): UGA codon for tryptophan. Curr Genet 32:296–299

Higgins DG, Sharp PM (1988) CLUSTAL: A package for performing multiple sequence alignment on microcomputer. Gene 73:237–244

Hirsh D, Gold L (1971) Translation of the UGA triplet *in vitro* by tryptophan transfer RNA's. J Mol Biol 58:459–68

Inagaki Y, Ehara M, Hayashi-Ishimaru Y, Igarashi I, Ohama T (1997) Algae or protozoa: Phylogenetic position of euglenophytes and dinoflagellates as inferred from mitochondrial sequences. J Mol Evol 45:295–300

Kairo A, Fairlamb AH, Gobright E, Neve V (1994) A 7.1 kb linear DNA molecule of *Theileria parva* has scarmbled rDNA sequences and open reading frames for mitochondrially encoded proteins. EMBO J 13:898–905

Lang BF, Burger G, O'Kelly CJ, Cedergren R, Golding GB, Lemieux C, Sankoff D, Turmel M, Gray MW (1997) An ancestral mitochondrial DNA resembling a eubacterial genome in miniature. Nature 387:493–497

Nakamura Y, Gojobori T, Ikemura T (1997) Codon usage tabulated from the international DNA sequence databases. Nucleic Acids Res 25:244–245

Oda K, Yamato K, Ohta E, Nakamura Y, Takemura M, Nozato N, Kohchi T, Ogura Y, Kanegae T, Akashi K, Ohyama K (1992) Gene organization deduced from the complete sequence of liverwort *Marchantia polymorpha* mitochondrial DNA. A primitive form of plant mitochondrial genome. J Mol Biol 223:1–7

Osawa S, Jukes TH, Watanabe K, Muto A (1992) Recent evidence for evolution of the genetic code. Microbiol Rev 56:229–264

Paquin B, Lang BF (1996) The mitochondrial DNA of *Allomyces macrogynus*: The complete genomic sequence from an ancestral fungus. J Mol Biol 225:688–701

Paquin B, Roewer I, Wang Z, Lang BF (1995) A robust fungal phylogeny using the mitochondrially encoded NAD5 protein sequence. Can J Bot 73:S180–S185

Paquin B, Laforest MJ, Forget L, Roewer I, Wang Z, Longcore J, Lang BF (1997) The fungal mitochondria genome project: Evolution of fungal mitochondrial genome and their gene expression. Curr Genet 31:380–395

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

Schultz DW, Yarus M (1994) Transfer RNA mutation and the malleability of the genetic code. J Mol Biol 235:1377–1380

Schultz DW, Yarus M (1996) On malleability in the genetic code. J Mol Evol 42:597–601

Viehmann S, Richard O, Boyen C, Zetsche K (1996) Genes for two subunits of succinate dehydrogenase form a cluster on the mitochondrial genome of Rhodophyta. Curr Genet 29:199–201

Wolff G, Plante I, Lang BF, Kück U, Burger G (1994) Complete sequence of the mitochondrial DNA of the chlorophyte alga *Prototheca wickerhamii.* Gene content and genome organization. J Mol Biol 237:75–86

Wolstenholme DR (1992) Animal mitochondria DNA: Structure and evolution. Int Rev Cytol 141:173–216

Yasuhira S, Simpson L (1997) Phylogenetic affinity of mitochondria of *Euglena gracilis* and kinetoplastids using cytochrome oxidase I and hsp60. J Mol Evol 44:341–347