Current Biology, Vol. 13, 442–447, March 4, 2003, ©2003 Elsevier Science Ltd. All rights reserved. PII S0960-9822(03)00126-X

A New Noncanonical Nuclear Genetic Code: Translation of UAA into Glutamate

Rocío Sánchez-Silva,[1] Eduardo Villalobo,[1,*] Loïc Morin,[2] and Antonio Torres[1]
[1]Departamento de Microbiología
Facultad de Biología
Universidad de Sevilla
41012 Sevilla
Spain
[2]Laboratoire de Biologie Cellulaire 4
Université Paris-Sud
91405 Orsay
France

Summary

Deviant genetic codes reported in ciliates share the same feature: one (UGA) or two (UAR) of the three canonical stop codons are translated into one particular amino acid. In many genera, such as *Oxytricha*, *Paramecium*, and *Tetrahymena*, UAR codons are translated into glutamine. UGA is translated into cysteine in *Euplotes* or into tryptophan in *Colpoda inflata* and *Blepharisma americanum*. Here, we show that three peritrich species (*Vorticella microstoma*, *Opisthonecta henneguyi*, and *Opisthonecta matiensis*) translate UAA into glutamate and that at least UAA in *O. matiensis* is decoded through a mutant suppressor-like tRNA. This kind of genetic code has never been reported for any living organism. Phylogenetic analysis with α-tubulin sequences corroborates that peritrichs, peniculines (*Paramecium*), and hymenostomates (*Tetrahymena*) form a monophyletic group (class Oligohymenophorea). The differential translation (glu/gln) of UAR codons, the monophyly of the Oligohymenophorea, and the common evolutionary origin of glutamate and glutamine suggest that deviant genetic codes of present-day oligohymenophoreans could have the same origin.

Results and Discussion

When analyzing the "phylogeny" of genetic code deviations [1], it becomes evident that reassignment of stop codons to sense codons is a relatively frequent phenomenon. This is particularly apparent in ciliates, in which two (UAR) or one (UGA) of the three canonical stop codons are translated into one particular amino acid. In many genera, such as *Oxytricha*, *Paramecium*, and *Tetrahymena*, UAR codons are translated into glutamine. On the other hand, UGA is translated into cysteine in *Euplotes* [2] or into tryptophan in *Colpoda* and *Blepharisma* [3]. We have investigated the peritrich ciliates *Vorticella microstoma*, *Opisthonecta henneguyi*, and *Opisthonecta matiensis* for the occurrence of deviant genetic codes. For this purpose, we used actins and tubulins as informational molecules because the alignment of their coding sequences is straightforward. We amplified, cloned, and sequenced several DNA fragments representing about 95% and 85% of typical actin and α-tubulin genes, respectively (see the Supplementary Material available with this article online). The deduced amino acid sequences indicate that *Vorticella* and *Opisthonecta* actin and α-tubulin coding regions contain UAR codons in frame. Preliminary alignment of the deduced amino acid sequences of peritrich actins shows that in-frame UAA and UAG codons (see Figure 1, amino acid positions 59, 127, 197, 216, 228, 272, and 313) are at positions that normally encode a glutamate residue (also see Figure S1 in the Supplementary Material). Likewise, in a preliminary α-tubulin alignment, in-frame UAA codons of peritrichs (see Figure 2, amino acid positions 77, 297, and 411) occur at glutamate residues. We then decided to extend the analysis to a wider number of actin- and α-tubulin-deduced amino acid sequences and to calculate the degree of conservation of glutamate residues at the abovementioned amino acid positions. This analysis was carried out with 306 actin and 136 α-tubulin sequences, which are representatives of most of the eukaryotic kingdoms. Glutamate shows (see Table 1) a high degree of conservation (above 94%) in all of the analyzed positions, except at position 313 of actins, where an aspartic residue (equivalent to glutamate) is the predominant amino acid (above 78%). Strikingly, E411 of tubulins is invariant, i.e., glutamate is at this position in all of the cases analyzed. In addition, a glutamine to glutamate substitution (see Table 1) is rarely observed, and it occurs in less than 1.5% of sequences at only four (amino acids positions 59, 197, 216 of actins and 77 of tubulins) out of ten analyzed positions. In summary, our data (six independent in-frame UAA codons in *O. henneguyi*, five in *V. microstoma*, and two in *O. matiensis*) strongly support that UAA is translated into glutamate. The support for UAG to glutamate is less strong because only two in-frame UAG codons have been found in *O. matiensis* actin, and one of them (position 313) is at a nonconserved site. Nonetheless, alignment of another available cDNA protein-coding sequence in the database [4] also displays a UAG at a position of a moderately conserved glutamate residue (86% E, 2% Q). The extended analysis (see Table 1 and Figure 1) also shows that some glutamine residues, which are invariant (position 133 of tubulins) or more than 93% conserved (positions 91, 176, 233, 256, and 342 of tubulins; 61, 123, 355 of actins), are encoded by UAR codons in *Paramecium* and/or *Tetrahymena*, but by canonical codons (CAR) in *Vorticella and Opisthonecta*. Since UAA and UAG are sense codons, the remainder of the canonical stop codons (UGA) must be used to mark the ends of translation. We thus amplified, cloned, and sequenced the remaining 3′ and 5′ regions of actin genes in *V. microstoma* and *O. henneguyi* by performing inverse PCR with sequence-specific, outward-facing primers and circularized DNA as template

*Correspondence: evpolo@us.es

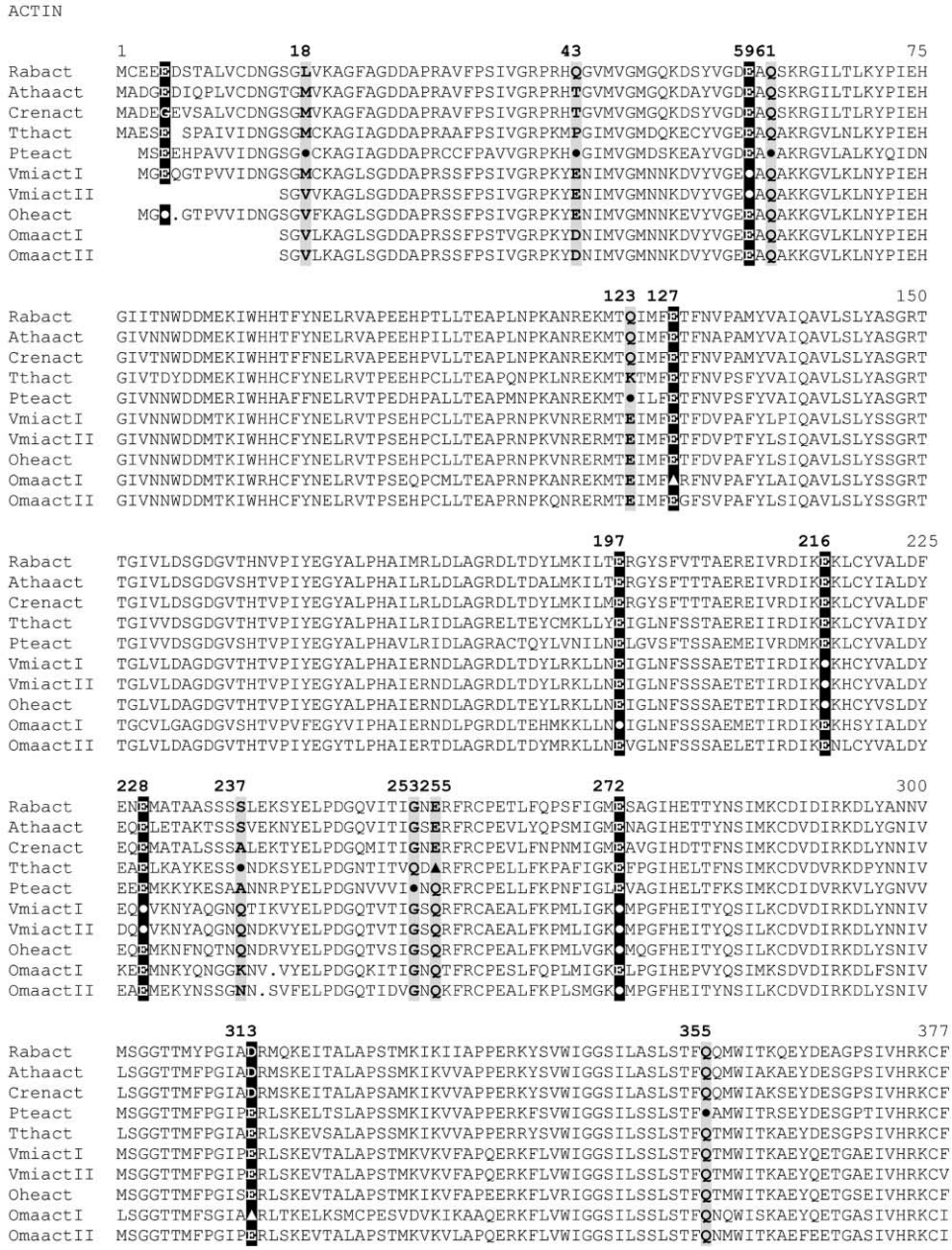

Figure 1. Alignment of Deduced Amino Acid Sequences of Actin

Vorticella microstoma, *Opisthonecta heneguyii*, and *Opisthonecta matiensis* actin are aligned to actin from rabbit (accession number NM_007392), *Arabidopsis thaliana* (accession number M20016), *Chlamydomonas reinhardtii* (accession number D50838.1), *Tetrahymena thermophila* (accession number M13939), and *Paramecium tetraurelia* (our unpublished data). Filled circles and filled triangles denote UAA and UAG codons, respectively, at the nucleotide sequence (see Figure S1). Gaps, introduced to improve alignments, are indicated by dots. The black and gray shadowed amino acids represent, respectively, glutamate and glutamine residues encoded by UAR in ciliates. The numbers on top of the alignments refer to the sequence in the first row. Rab, rabbit; Atha, *Arabidopsis thaliana*; Cre, *Chlamydomonas reinhardtii*; Tth, *Tetrahymena thermophila*; Pte, *Paramecium tetraurelia*; Vmi, *Vorticella microstoma*; Ohe, *Opisthonecta henneguyi*; Oma, *Opisthonecta matiensis*.

(see the Experimental Procedures in the Supplementary Material). Sequence analyses of these actin 5′ regions show that both species use UGA as the true stop codon. Therefore, all of this evidence strongly supports that UAA codons are translated into glutamate in peritrichs. Whether UAG is translated into glutamate must be further verified. These genetic code deviations, namely translation of canonical stop codons into glutamate, had never been described in ciliates or in any other living organism.

According to wobble pairing rules, canonical glutamate tRNAs, with YUC anticodons, would not be able

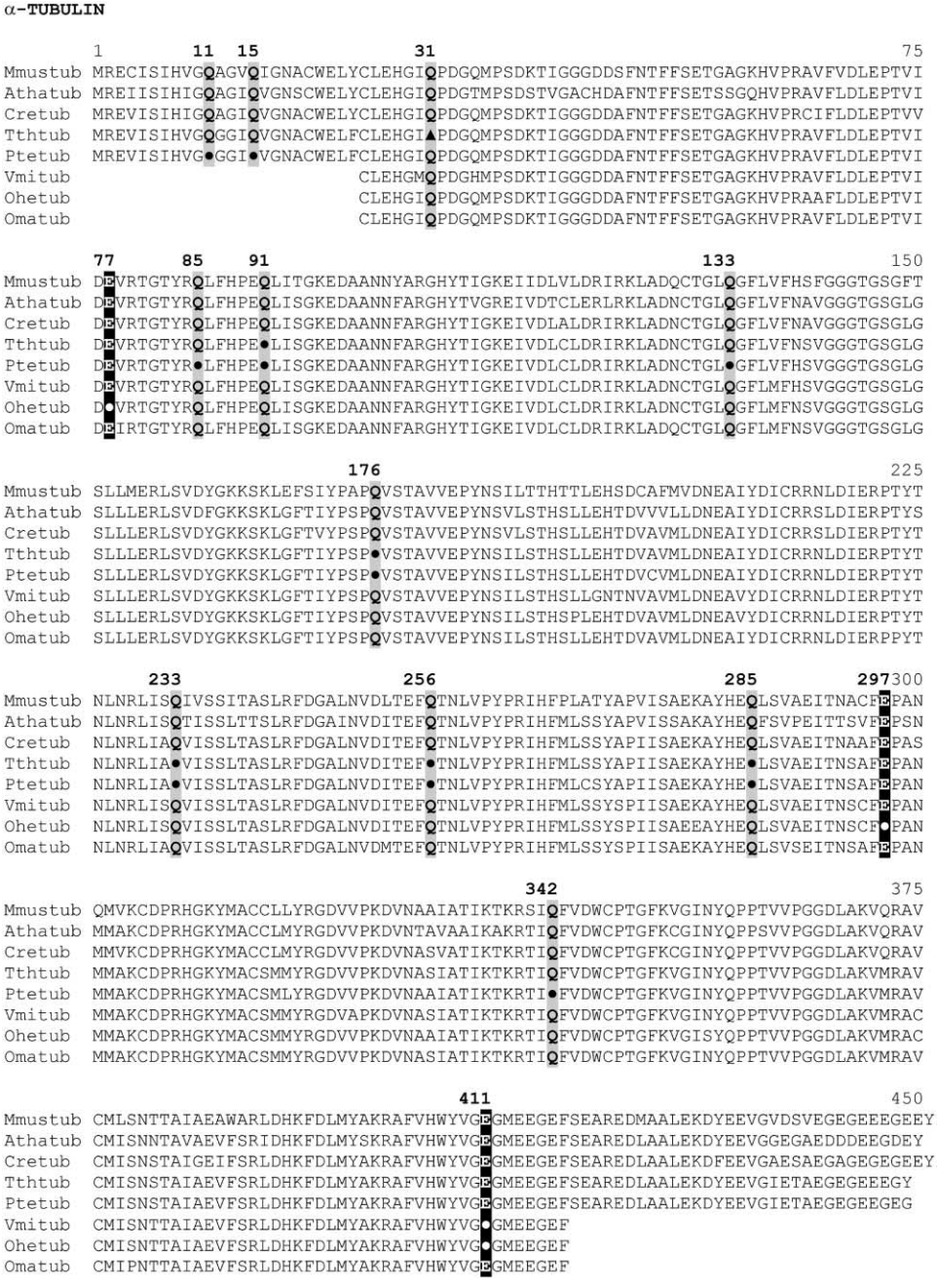

Figure 2. Alignment of Deduced Amino Acid Sequences of α-Tubulin

V. microstoma, *Opisthonecta heneguyii*, and *Opisthonecta matiensis* α-tubulin sequences are aligned to α-tubulin from *Mus musculus* (accession number BC008117.1), *Arabidopsis thaliana* (accession number AY091372.1), *Chlamydomonas reinhardtii* (accession number M11447.1), *Tetrahymena thermophila* (accession number M86723), and *Paramecium tetraurelia* (accession number X99489). Filled circles and filled triangles denote UAA and UAG codons, respetively, at the nucleotide sequence (see Figure S1). The black and gray shadowed amino acids represent, respectively, glutamate and glutamine residues encoded by UAR in ciliates. The numbers on top of the alignments refer to the sequence in the first row. Mmus, *Mus musculus*; Atha, *Arabidopsis thaliana*; Cre, *Chlamydomonas reinhardtii*; Tth, *Tetrahymena thermophila*; Pte, *Paramecium tetraurelia*; Vmi, *Vorticella microstoma*; Ohe, *Opisthonecta henneguyi*; Oma, *Opisthonecta matiensis*.

to pair UAR codons. Since two suppressor-like tRNAs ($tRNA^{Gln}_{UUA}$ and $tRNA^{Gln}_{CUA}$), which decode UAR into glutamine, have been reported in the hymenostomate *Tetrahymena thermophila* [5], we hypothesized the existence of YUA-bearing tRNAs in peritrichs. Searching for these tRNAs, we cloned and sequenced (see Figure S2) two DNA fragments in *O. matiensis*, which were recognized as tRNAs when they were analyzed with the tRNAscan-SE software [6]. This software also displays that one of these nucleotide sequences bears a UUC anticodon while the other bears a UUA anticodon. Thus, the first sequence most likely codes for a putative canonical

Table 1. Percentage of Glutamate and Glutamine Residues Encoded by UAR Codons in Actin and α-Tubulin of Oligohymenophoreans

Protein	Amino Acid Position	% E	% Q	% Other Amino Acids
Actin				
	59	98.69 (100)	0.32 (0)	0.99
	127	99 (100)	0	1
	197	95.8 (100)	0.65 (0)	3.55
	216	98.64 (95)	0.32 (0)	0.99
	228	94.4 (40)	0	5.6
	272	96 (100)	0	4
	313	18.3 (100)	0	81.7
	18	-	0.32 (5)	99.68
	43	-	59.4 (10)	40.6
	61	-	98.7 (100)	1.3
	123	-	98.7 (90)	1.3
	253	-	0.98 (15)	99.02
	255	-	3.6 (40)	94.4
	355	-	93.1 (35)	6.9
α-Tubulin				
	77	95.6 (100)	1.47 (0)	4.4
	297	96.32 (100)	0	3.68
	411	100 (100)	0	0
	11	-	100 (100)	0
	15	-	100 (100)	0
	31	-	82 (100)	18
	85	-	83 (100)	17
	91	-	93.4 (100)	6.6
	133	-	100 (100)	0
	176	-	95.6 (100)	4.4
	233	-	99.26 (100)	0.74
	256	-	99.26 (100)	0.74
	285	-	85.3 (100)	14.7
	342	-	95.6 (100)	4.4

These values were calculated only in the positions encoded by the UAR codon in peritrichs (see the text and Figure 1). The percentage of glutamine was calculated only in the positions encoded by the UAR codon in *Tetrahymena* and/or *Paramecium* (see the text and Figure 1). Actin (306 sequences)- and α-tubulin (136 sequences)-deduced amino acid sequences were retrieved from the Interpro database (at the EBI) and have accession numbers PS00406 and PS00227, respectively. Amino acid positions are referred to as rabbit actin (GenBank accession number NM_007392) and *Mus musculus* α-tubulin (GenBank accession number BC008117.1). The numbers in parentheses correspond to values obtained when considering only ciliate sequences.

glu-tRNA, and the second one most likely codes for a putative suppressor-like tRNA.

Phylogenetic studies based on rRNA sequences [7] reveal a solid and consistent association between peritrichs (*Vorticella* and *Opisthonecta*), hymenostomates (*Tetrahymena*), and peniculines (*Paramecium*). In order to validate the monophyletic character of the class Oligohymenophorea, we reexamined the phylogenetic position of the peritrich ciliates by using a UAR-containing marker. We used α-tubulin, which allows one to infer phylogenetic relationships at a low taxonomic level within ciliates [8]. The general pattern of the α-tubulin tree (see Figure 3) is rather similar to a multifurcation leading to different subgroups. *Opisthonecta* and *Vorticella* species branch close to the hymenostomates (*Tetrahymena*) and peniculines (*Paramecium*) with high bootstrap values. These species constitute a solid monophyletic unit, the class Oligohymenophorea. Consequently, UAR codons are differentially translated within a solid monophyletic group into glutamine in *Tetrahymena* and *Paramecium* and into glutamate in *Vorticella* and *Opisthonecta*. A similar situation was already described in pseudohypotrichs since *Diophrys* translates UAR into glutamine [9], whereas *Euplotes* translates UGA into cysteine and UAR are stop codons [2]. This scattered pattern of UAR reassignment within and between ciliate classes led to the hypothesis, first proposed by Baroin-Tourancheau et al. [10], that these nuclear genetic code deviations have arisen independently several times within the phylum. Nevertheless, we suggest that UAR reassignment could have a single origin at least in oligohymenophoreans. This is based on the solid monophyly of the Oligohymenophorea class, the translation of UAR codons with different specificity (glu or gln) within this class, and the accepted common evolutionary origin of the glutamate and glutamine translational pathways [11]. In modern organisms, glutamine translation follows two different pathways. According to the indirect pathway, $tRNA^{YUG}$ (canonical gln-tRNAs) are first glutamylated by a glutamyl-tRNA synthetase (gluRS) and are then converted into glutaminyl-$tRNA^{YUG}$ by an amidotransferase. Alternatively, in the direct pathway, $tRNA^{YUG}$ are glutaminylated by a glutaminyl-tRNA synthetase (glnRS). Both pathways are found in modern prokaryotes, although it is assumed that glnRS evolved in eukaryotes. A horizontal gene transfer probably accounted for the direct pathway in prokaryotes [11]. It has been suggested that the glutamine direct pathway probably evolved from that of glutamate since gluRS and glnRS synthetases are paralogous genes [11]. An

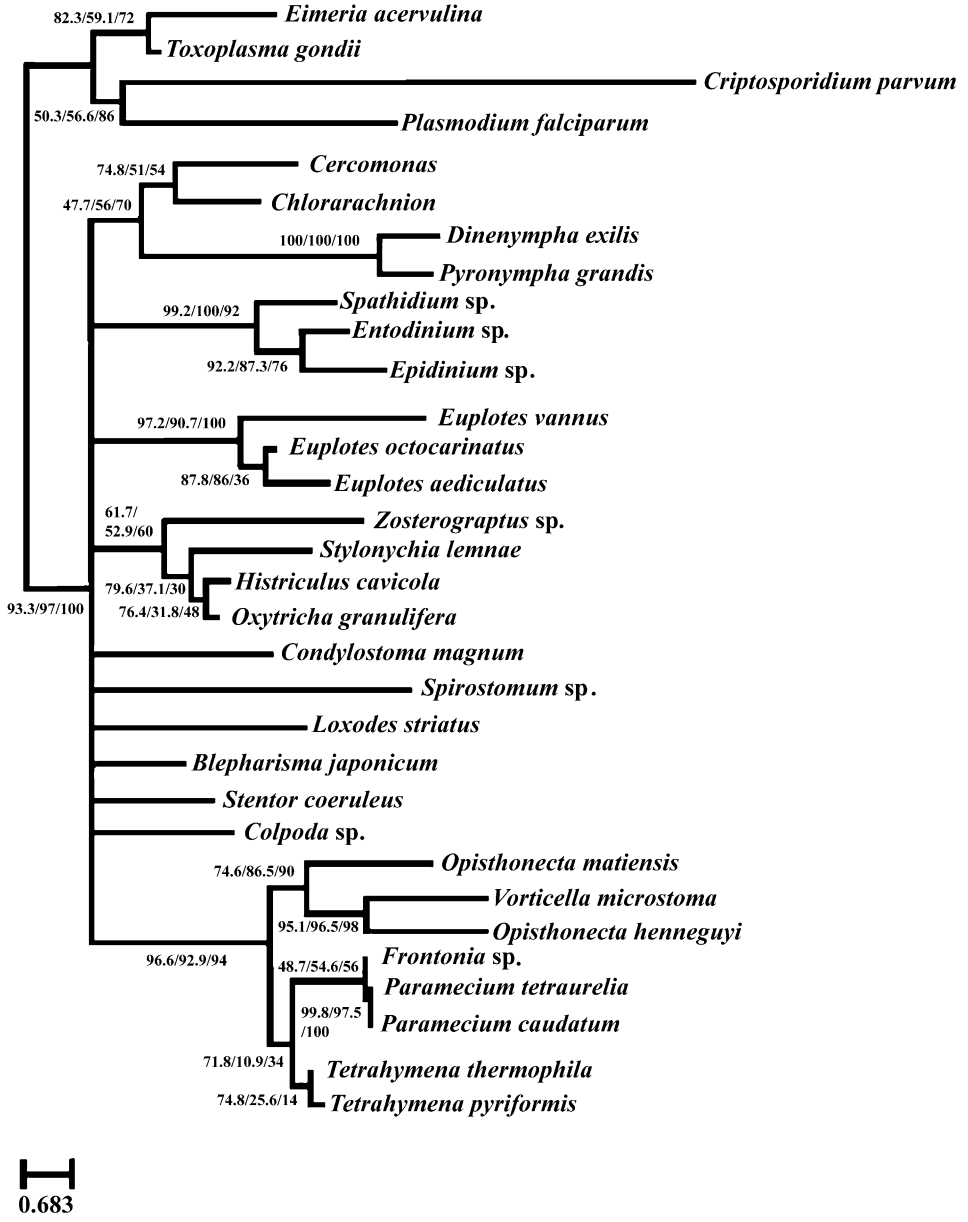

Figure 3. Consensus Tree Based on α-Tubulin-Deduced Amino Acid Sequences

The resulting bootstrap values are displayed at each node; neighbor-joining, parsimony, and maximum likelihood are listed from left to right. The scale bar represents 0.6083 expected amino acid replacements per position.

ancestral glxRS might have glutamylated $tRNA^{YUS}$ (canonical glu- and gln-tRNAs); however, the amidotransferase selected only the $tRNA^{YUG}$ (canonical gln-tRNAs) in order to convert glutamyl-tRNAs to glutaminyl-tRNAs [12]. Later on, the duplication of the glxRS and the divergence of one copy made possible the specialization of two related enzyme activities, glu-tRNA and gln-tRNA [12, 13]. This implied not only changes in the acylation activity, but also changes in the preference for the last position of the tRNA anticodons. Meaningfully, it has been suggested that the evolving glnRS could change its anticodon preference through the use of intermediate suppressor-like $tRNA^{YUA}$ [12]. All of this is not surprising if one bears in mind that, in modern organisms, incorporation of nonstandard amino acids, such as the recently discovered pyrrolysine [14] or selenocysteine [15], is made by taking over the tRNA synthetase system of standard amino acids (lysine and serine, respectively) and suppressor-like tRNAs. In summary, there may have been an oligohymenophorean ancestor with a deviant genetic code. Its code diverged later, giving different specificity to UAR codons. This divergence was possible through modifications on the aminoacylation apparatus that we suggest could be similar to those that gave rise to the glutamine translational pathway.

Supplementary Material

Supplementary Material including the Experimental Procedures, nucleotide sequence alignments, and a table with sequence features is available at http://images.cellpress.com/supmat/supmatin.htm.

Acknowledgments

We would like to thank A. Villalobo for his fruitful comments on this paper and Josh Raizman and Carmen Velasco for their help with the English.

Received: September 24, 2002
Revised: November 11, 2002
Accepted: December 19, 2002
Published: March 4, 2003

References

1. Knight, R.D., Freeland, S.J., and Landweber, L.F. (2001). Rewiring the keyboard: evolvability of the genetic code. Nat. Rev. Genet. *2*, 49–58.
2. Meyer, F., Schmidt, H.J., Plumper, E., Hasilik, A., Mersmann, G., Meyer, H.E., Engstrom, A., and Heckmann, K. (1991). UGA is translated as cysteine in pheromone 3 of *Euplotes octocarinatus*. Proc. Natl. Acad. Sci. USA *88*, 3758–3761.
3. Lozupone, C.A., Knight, R.D., and Landweber, L.F. (2001). The molecular basis of nuclear genetic code change in ciliates. Curr. Biol. *11*, 65–74.
4. Maciejewski, J.J., Vacchiano, E.J., McCutcheon, S.M., and Buhse, H.E. (1999). Cloning and expression of a cDNA encoding a *Vorticella convallaria* spasmin: an EF-hand calcium-binding protein. J. Eukaryot. Microbiol. *46*, 165–173.
5. Hanyu, N., Kuchino, Y., and Nishimura, S. (1986). Dramatic events in ciliate evolution: alterations of UAA and UAG termination codons to glutamine codons due to anticodon mutations in two *Tetrahymena* $tRNAs^{Gln}$. EMBO J. *5*, 1307–1311.
6. Lowe, T.M., and Eddy, S.R. (1997). tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. *25*, 955–964.
7. Miao, W., Yu, Y.-H., and Shen, Y.-F. (2001). Phylogenetic relationships of the Subclass Peritrichia (Olygohymenophorea, Ciliophora) with emphasis on the genus *Epistylis*, inferred from small subunit rRNA gene sequences. J. Eukaryot. Microbiol. *48*, 583–587.
8. Baroin Tourancheau, A., Villalobo, E., Tsao, N., Torres, A., and Pearlmann, E. (1998). Protein coding trees in ciliates: comparison with rRNA-based phylogenies. Mol. Phylogenet. Evol. *10*, 299–309.
9. Perez-Romero, P., Villalobo, E., and Torres, A. (2001). Different stop codon usage in two pseudohypotrich ciliates. FEMS Microbiol. Lett. *205*, 259–263.
10. Baroin Tourancheau, A., Tsao, N., Klobutcher, L.A., Pearlman, R.E., and Adoutte, A. (1995). Genetic code deviation in the ciliates: evidence for multiple and independent events. EMBO J. *14*, 3262–3267.
11. Gagnon, Y., Lacoste, L., Champagne, N., and Lapointe, J. (1996). Widespread use of the glu-$tRNA^{Gln}$ transamidation pathway among bacteria. A member of the alpha purple bacteria lacks glutaminyl-tRNA synthetase. J. Biol. Chem. *271*, 14856–14863.
12. Siatecka, M., Rozek, M., Barciszewski, J., and Mirande, M. (1998). Modular evolution of the Glx-tRNA synthetase family. Rooting of the evolutionary tree between the bacteria and archaea/eukarya branches. Eur. J. Biochem. *256*, 80–87.
13. Rogers, K.C., and Söll, D. (1995). Divergence of glutamate and glutamine aminoacylation pathways: providing the evolutionary rationale for mischarging. J. Mol. Evol. *40*, 476–481.
14. Srinivasan, G., James, C.M., and Krzycki, J.A. (2002). Pyrrolysine encoded by UAG in Archaea: charging of a UAG-decoding specialized tRNA. Science *24*, 1459–1462.
15. Ibba, M., and Söll, D. (2000). Aminoacyl-tRNA synthesis. Annu. Rev. Biochem. *69*, 617–650.