J Mol Evol (1996) 42:321–322

JOURNAL OF MOLECULAR EVOLUTION

Letter to the Editor

CUG Codons in *Candida* spp.

T.H. Jukes,[1] S. Osawa[2]

[1] Space Sciences Laboratory, University of California, Berkeley, 6701 San Pablo Avenue, Oakland, CA 94608 USA
[2] Biohistory Research Hall, 1-1 Murasaki-cho, Takatsuki, Osaka 569, Japan

Received: 16 October 1995 / Accepted: 21 October 1995

Abstract. Codon CUG is used for serine instead of for leucine, its usual assignment, in several yeasts of the genus *Candida.* We propose a series of steps for the reassignment, including disappearance of leucine CUG and its anticodon CAG, formation of a new serine tRNA, with anticodon CAG, from a duplication of the gene for serine tRNA (IGA), and then production of CUG codons by mutation at sites that are mostly "nonessential."

Key words: Genetic code — *Candida* yeasts — Leucine codons — Serine codons — Codon capture

Several species of yeast in the genus *Candida* use CUG as a codon for serine instead of leucine as in the "universal" code (Kawaguchi et al. 1989; Yokogawa et al. 1992). Thus, in *Candida cylindracea* (Kawaguchi et al. 1989; Ohama et al. 1993; Yokogawa et al. 1992), *parapsilosus, zeylanoides, albicans, ragosa,* and *melibiosica,* CUG codes for serine (Yokogawa et al. 1992) and in several others, including *magnolidae, azyma,* and *utilis,* CUG codes for leucine. In *C. cylindracea* the tRNA with anticodon CAG, charged with serine, translates codon CUG and has been identified and sequenced (Yokogawa et al. 1992). A comparison of this tRNA with its gene led to the conclusion that a single cytidine was inserted into the anticodon loop of the gene for tRNA (Ser) IGA during evolution so as to produce tRNA (Ser) CAG. (IGA is AGA in the gene) (Yokogawa et al. 1992). The translation of codon CUG into serine in *Candida albicans* has been commented on by Santos and Tuite (1995).

For the change from CUG coding Leu to Ser, the first prerequisite is that CUG must be removed as a leucine codon. This is easily accomplished by AT pressure (Osawa et al. 1992), which changes CUG codons to leucine codons CUA, UUG, and UUA. Therefore, no changes take place in amino acid sequences. The anticodon CAG for leucine must also disappear; otherwise, "new" CUG codons formed by mutation would be translated as leucine, and could not be codons for serine. If CUG coded both leucine and serine, this would lead to ambiguous translation of a gene into various polypeptide sequences rather than a single protein. Anticodons IAG, UAA, CAA, and UAG can persist for translation of other leucine codons. With the disappearance of CUG and its tRNA (Leu) CAG, CUG became an untranslatable nonsense codon that would stop the translation process, as in the case of codon CGG in *Mycoplasma capricolum* (Osawa et al. 1992). A unique event then took place, as described above—the formation of a strong tRNA with anticodon CAG. This tRNA and the gene for tRNA serine (IGA) were identified and sequenced by Yokogawa et al. (1992). Ohama et al. (1993) have also sequenced tRNA Ser (CAG) genes from five other *Candida* species.

Once present, serine tRNA (CAG) translated CUG codons when they are formed by mutation. The "regu-

Correspondence to: T.H. Jukes

lar'' serine codons UCN and ACY are translated in eukaryotes by anticodons IGA, UGA, CGA, and GCU (Osawa et al. 1992).

Santos and Tuite (1995) discuss the translation of codon CUG into serine in the yeast *Candida albicans.* They state that ''this major deviation to [*sic*] the pattern of genetic code changes found in eukaryotic cytoplasmic mRNAs cannot be explained by the ''codon reassignment theory'' (referring to the publication by Osawa et al. (1992) ''for two main reasons).

The ''main reasons'' stated by them are, first, that it is unlikely that codons disappear from the entire set of mRNAs due to GC or AT pressure, because the relative GC/AT content is not uniform in *S. cerevisiae* within chromosomes. However, in eubacteria, GC/AT content is also not uniform, as shown in highly and poorly expressed genes occurring in genomes with moderate GC contents. In bacteria with very high GC content, such as *Micrococcus luteus,* or high AT content, such as *Mycoplasma capricolum,* the GC/AT content becomes almost uniform throughout the genome, and certain codons have actually disappeared (Osawa et al. 1992). Therefore this ''main reason'' is not valid unless it is shown that GC/AT content is uneven and that CUG codons do not disappear in *Candida* species that are very high in AT. Second they state that ''an intermediary step is required in which the mutant anticodon decodes either the most frequently used leucine codon UUG or the proline-CCG codon.''

This first ''main reason''—that it is unlikely for CUG to disappear entirely—is wrong, for this disappearance is obligatory for a change in the meaning of a codon, as described above. Their second reason is also incorrect. The mutant anticodon CAG is produced in one step (not two steps) from the gene for tRNA Ser (IGA). Anticodon CAG pairs with codon CUG, and not with UUG (Leu) or CCG (Pro) as postulated by Santos and Tuite.

Our next proposal (see above) was that CUG leucine codons had disappeared from *C. cylindraceae* and tRNA Leu (CAG) had also disappeared by relaxation of functional constraints. It would not be necessary for UUG or CCG, in addition to CUG, also to disappear, as proposed by Santos and Tuite, because Leu anticodon CAG does not pair with either UUG or CCG. Their anticodons are, respectively, CAA and CGG. The ''new'' tRNA (Ser) CAG would pair with ''new'' CUG codons to translate them as Ser. These CUG codons were produced, according to our scheme, by mutations of various other codons, perhaps AUG (Met) and some leucine codons (but not serine codons UCN or AGY) to CUG under GC pressure, at sites that can accommodate neutral changes. At sites that do not tolerate such mutations, the result of their occurrence would have produced inferior genes, and several inferior lipase genes exist in *C. cylindracea* (Kawaguchi et al. 1989). Also, CUG is used for Ser in the catalytic center of a lipase gene of *C. cylindracea* (Kawaguchi et al. 1989), showing that a Ser codon at an important site can become CUG by a nearly neutral process. In this case, there is no indication that Ser replaced Leu for codon CUG, but the Ser codon CUG must have been formed from another Ser codon, UCN. This would pass through an intermediate codon such UUR, Leu, CCN Pro. A change from AGY to CUG is less likely, because this would entail three substitutions. The gene with the ''intermediate'' codon UUR or CCN in the catalytic center would become dysfunctional or inactive, but most eukaryotic genes exist in more than one copy and therefore the cell would continue to be viable. The inactive gene could be reactivated by a second mutation from UUR Leu or CCN Pro reverting to UCR, or by a mutation from UUG or CCG to CUG, preceded by formation of tRNA Ser (CAG). Note that high GC pressure increases the usage of CUG codons, regardless of whether they are assigned to Ser or to Leu. For example, CUG is a rare codon for Leu in *S. cerevisiae* (40% G + C) and for Leu in *C. albican* (34% G + C) but is a predominant Ser codon in *C. cylindracea* (63% G + C); see discussion in Ohama et al., (1993, p. 4044).

This series of events overcomes all the objections by Santos and Tuite. There is no need for ''the mutant anticodon to decode either Leu codon UUG or the proline CCG codon'' (Santos and Tuite 1995). There is no ''ambiguous decoding'' (Santos and Tuite 1995). There is no mutation of ''a serine codon UCN to the CUG codon'' (Santos and Tuite 1995). The requirements for codon capture are the disappearance of a codon and its reappearance with a new assignment (Osawa et al. 1992). These have been fulfilled. As stated previously (Ohama et al. 1993), the six *Candida* species examined in which CUG codes for serine belong to one distinct group of *Hemiascomycetes.*

References

Kawaguchi Y, Honda H, Taniguchi-Morimura J, Iwasaki S (1989) The codon CUG is read as serine in an asporogenic yeast, *Candida cylindracea.* Nature 341:164–166

Ohama T, Suzuki T, Mori M, Osawa S, Ueda T, Watanabe K, Nakase T (1993) Non-universal decoding of the leucine codon CUG in several *Candida* species. Nucleic Acids Res 21:4039–4045

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

Santos MAS, Tuite MF (1995) The CUG codon is decoded *in vivo* as serine and not as leucine in *Candida albicans.* Nucleic Acids Res 23:1481–1486

Yokogawa T, Suzuki T, Ueda T, Mori M, Ohama T, Kuchino Y, Yoshinari S, Moroki I, Nishikawa K, Osawa S, Watanabe K (1992) Serine tRNA complementary to the nonuniversal serine codon CUG in *Candida cylindracea*: Evolutionary implications. Proc Natl Acad Sci U.S.A. 89:7408–7411