Permuted tRNA Genes in the Nuclear and Nucleomorph Genomes of Photosynthetic Eukaryotes

Shinichiro Maruyama,*1 Junichi Sugahara,2,3 Akio Kanai,2,3,4 and Hisayoshi Nozaki1

1Department of Biological Sciences, Graduate School of Science, University of Tokyo, 7-3-1 Hongo, Bunkyo, Tokyo 113-0033, Japan
2Institute for Advanced Biosciences, Keio University, Tsuruoka 997-0017, Japan
3Systems Biology Program, Graduate School of Media and Governance, Keio University, Fujisawa 252-8520, Japan
4Faculty of Environment and Information Studies, Keio University, Fujisawa 252-8520, Japan

*Corresponding author: E-mail: maruyama@biol.s.u-tokyo.ac.jp.

Abstract

Transfer RNA (tRNA) is a central genetic element in the decoding of genome information for all of Earth’s life forms. Nevertheless, there are a great number of missing tRNAs that have been left without examination, especially in microbial genomes. Two tRNA gene families remarkable in their structure and expression mechanism have been reported: split and permuted tRNAs. Split tRNAs in archaea are encoded on the genome as two or three fragmented genes and then processed into single tRNA molecules. Permuted tRNAs are organized with the 5′-3′ halves of the gene positioned in reverse on the genome and hitherto have been found only in one tiny red alga. Here we reveal a wide-ranging distribution of permuted tRNA genes in the genomes of photosynthetic eukaryotes. This includes in the smallest eukaryotic genome known to date, the nucleomorph genome of the chlorarachniophyte alga Bigelowiella natans. Comparison between cDNA and genomic DNA sequences of two nucleomorph-encoded tRNAser genes confirms that precursors are circularized and processed into mature tRNA molecules in vivo. In the tRNAser(AGA), adenine at the wobble position of the codon is likely modified to inosine to expand capacity of the codon recognition. We also show the presence of permuted tRNAs in the ultrasmall free-living green algae Ostreococcus and Micromonas, which are closely related to the B. natans nucleomorph. Conserved intron/leader sequence structures in the intron-containing and permuted tRNAs suggest the ancient origin of the splicing machinery in the common ancestor of eukaryotes and archaea. Meanwhile, a wide but patchy distribution of permuted tRNA genes in the photosynthetic eukaryotes implies that extant permuted tRNAs might have emerged multiple times. Taken together, our data demonstrate that permuted tRNA is an evolutionarily conserved and fundamental element in tiny eukaryotic genomes.

Key words: permuted tRNA, intron, SPLITS, nucleomorph, green algae, endosymbiosis.

Introduction

Among all three domains of life (eukaryotes, bacteria, and archaea), transfer RNA (tRNA) plays a key role in protein biosynthesis. Although tRNA is one of the most conserved RNA molecules in structure and function, the tRNA gene is architecturally diverse. In archaea, there have been three types of tRNA genes identified to date: intron-lacking, intron-containing, and split tRNA (Randau and Söll 2008; Sugahara et al. 2008, 2009). Split tRNA genes are encoded on two or three separated regions on the genome and then processed into single tRNA molecules (Randau, Münch, et al. 2005; Fujishima et al. 2009). Intron-containing and split tRNAs share a common structural motif called bulge-helix-bulge (BHB), or more relaxed bulge-helix-loop (BHL), at the intron/leader–exon boundary and suggested to be evolutionarily interrelated (Randau, Calvin, et al. 2005; Tocchini-Valentini et al. 2005; Fujishima et al. 2009). In eukaryotic genomes, intron-lacking, intron-containing tRNAs, as well as permuted tRNAs have been found. Permuted tRNA genes, in which the 3′ and 5′ regions of the gene are encoded at a single locus in reverse order and processed into circular and then canonical linear cloverleaf forms, have been found only in the ultrasmall red alga Cyanidioschyzon merolae nuclear genome (Soma et al. 2007).

Although the finding of permuted tRNA revealed that the integrity of the genome information in C. merolae was maintained by these odd tRNAs (Soma et al. 2007), the evolutionary history and distribution of the permuted tRNA species in the eukaryotic genomes is totally unknown. This is partly due to the lack of genome sequence data from red algae and the difficulty in searching for the counterparts of the C. merolae permuted tRNA genes in other red algae. Comparatively, genomic analysis is more advanced in green plants (green algae and land plants). A growing amount of the genome sequence data showed that the large genomes of land plants are abundant in canonical tRNA genes (Chan and Lowe 2009), which is generally considered to be sufficient for protein biosynthesis. On the other hand, although a number of genome sequences from green algae have been published, many “missing” tRNAs in the small genomes of green algae and the related tiny eukaryotic genomes remain to be identified.

To date, only a few algal genomes such as the nuclear genomes of C. merolae (Nozaki et al. 2007) and the green alga Ostreococcus “lucimarinus” CCMP2972 (Palenik et al. 2005; Tocchini-Valentini et al. 2005; Fujishima et al. 2009).
Permuted tRNAs in Photosynthetic Eukaryotes

Materials and Methods

Computational Prediction of tRNA Genes in the B. natans Nucleomorph Genome

Genome-wide prediction of permuted tRNA genes encoded in the B. natans nucleomorph genome (Gilson et al. 2006) and the nuclear genomes of Chlamydomonas reinhardtii (Merchant et al. 2007), O. tauri, O. lucimarinus (Palenik et al. 2007), Micromonas sp. RCC299 and M. pusilla CCMP1545 (Worden et al. 2009) was performed using SPLITS software (Sugahara et al. 2006). SPLITS parameters used in this analysis were “-c -p 0.55 -F -3 -h 3” (Soma et al. 2007) or “-c -p 0.6 -F -1.” Prediction of tandem-duplicated tRNA genes was conducted using SPLITS or tRNAscan-SE (Lowe and Eddy 1997). For proper memory utilization, scaffolds larger than 2.5 Mb in the C. reinhardtii genome data were fragmented into a number of 2-Mb subsequences with 10-kb overlapping regions.

Cell Culture and cDNA Preparation

Bigelowiella natans strain CCMP621 was cultured in artificial seawater supplemented with 1× Daigo IMK medium (Nippon Pharmaceutical, Tokyo, Japan) at 25 °C under a 16:8 h light:dark cycle. Total RNA was extracted using RNAiso Plus (Takara Bio, Shiga, Japan) and treated with RQ1 RNase-Free DNase (Promega, Madison, WI) and aliquots were self-ligated into circular molecules using T4 RNA ligase (Takara Bio) according to manufacturer’s instructions. Preparation of cDNA was conducted with SuperScript II reverse transcriptase (Invitrogen, Carlsbad, CA) using primers Bnpmt1_R1 (5′-TCGAGTCATCCTCTTAAACA-3′) or Bnpmt1_R2 (5′-GATTGCAACCTGTGCGTAA-3′) for tRNA Ser (CGA) and Bnpmt2_R1 (5′-TCAGTCCTCTTAAACCATCCTAG-3′) or Bnpmt2_R2 (5′-AGATTGCAACC-TATGCCGAT-3′) for tRNA Ser (AGA). Polymerase chain reaction (PCR) products were amplified with ExTaq (Takara Bio) using primer pairs Bnpmt1_F1 (5′-ATGGTTTT-TACCCCAACAGG-3′) and Bnpmt1_R1, Bnpmt1_F2 (5′-CGGTTCTGTGCGATGTTG-3′) and Bnpmt1_R2 for tRNA Ser (CGA) and Bnpmt2_F1 (5′-TGAGTTACTCGCA-TAGGTTCG-3′) and Bnpmt2_R1, Bnpmt2_F2 (5′-GACCATGCCCTGAGTTAAG-3′) and Bnpmt2_R2 for tRNA Ser (AGA). The PCR products were visualized in ethidium bromide–stained 12% polyacrylamide–tris-borate-ethylenediaminetetraacetic acid gels. The B. natans permuted tRNA gene sequences were deposited in DDBJ/EMBL/GenBank under the accession numbers AB537483 and AB537484.

Results

Using SPLITS, we identified two candidate permuted tRNA genes in the B. natans nucleomorph genome. Both were 1) genes encoding tRNA Ser, 2) homologs of each other, and 3) likely members of the tRNA family spliced at position 37/38 (fig. 1). Both genes have distinct 5-bp intervening sequences, flanked by upstream 3′ and downstream 5′ halves. One gene, tRNA Ser (CGA), is located just 6 bp away from tRNA Ile (TAT); however, it does not overlap flanking genes or other known tRNA genes. The other gene is tRNA Ser (AGA), located 5 bp from tRNA Tyr (GTA), and it overlaps no known genes. We also surveyed the nucleomorph genomes of two cryptophytes, Guillardia theta (Douglas et al. 2001) and Hemiselmis andersenii (Lane et al. 2007). Regardless of a common ancestry between the cryptophyte nucleomorph and C. merolae nucleus, we failed to identify any candidate permuted tRNA genes in the nucleomorph genomes of G. theta and H. andersenii. Introns on tRNA genes as previously found in the G. theta nucleomorph (Douglas et al. 2001; Kawach et al. 2005) were not identified in the B. natans nucleomorph in this study.

To verify whether these B. natans permuted tRNA Ser genes are active genes that are actually transcribed and processed into a mature form, we sequenced reverse transcription RT–PCR products prepared from total RNA (Soma et al. 2007) of B. natans CCMP621. We detected RT–PCR products corresponding to the precursor form (3′ half, intervening region, 5′ half), the circular form (detected as two rounds of RT around the circular RNA molecules; data not shown), as well as the mature form with canonical cloverleaf secondary structure (5′ half to 3′ half), suggesting that the nucleomorph tRNA genes are actively transcribed and processed (fig. 2A). Furthermore, we sequenced RT–PCR products using total RNA circularized
FIG. 1. A model for the maturation process of the *Bigelowiella natans* permuted tRNA genes. (A) Models of permuted tRNA genes predicted by SPLITS software. Magenta, 3' half of tRNA; green, 5' half; blue, intervening spacer sequence; black, flanking sequence. Arrowheads indicate positions of splicing sites. (B) Schematic representation of permutation, processing, and addition of the CCA sequence at the 3' terminus.
Permuted tRNAs in Photosynthetic Eukaryotes · doi:10.1093/molbev/msp313

by T4 RNA ligase and confirmed that the CCA sequence was added at the 3′ terminus of the mature tRNA molecule (data not shown), indicative of the maturation process conserved among all eukaryotes (Weiner 2004; Soma et al. 2007).

It is of note that a comparison between the genomic and cDNA sequences of one B. natans permuted tRNA gene, tRNA\textsuperscript{Ser}(AGA), revealed that the anticodon sequence of the cDNA was GGA. This suggested that the A residue at the wobble position of the anticodon was modified by deamination into a nucleotide capable of pairing with C (fig. 28). Given the pairing affinity of inosine (I) with C during RT–PCR, our data suggest that the GGA anticodon likely represents the I modification of the permuted tRNA anticodon (Gerber and Keller 1999; Wolf et al. 2002). Generally, I modification is important for extending the codon recognition capacity of tRNA (Gerber and Keller 1999; Wolf et al. 2002). In the B. natans nucleomorph, tRNA\textsuperscript{Ser}(CGA) and the presumed tRNA\textsuperscript{Ser}(IGA) are capable of recognizing all four TCN codons for serine (supplementary fig. 1, Supplementary Material online). Our results indicate that I modification might play a role in maximizing the codon recognition capacity by the limited number of tRNA genes in a highly reduced genome.

To investigate the evolutionary origin of permuted tRNA genes in the B. natans nucleomorph, we surveyed publicly available nuclear genome sequences of green algae (cholorophytes): the chlorophycean C. reinhardtii (Palenik et al. 2007), the prasinophytes O. lucimarinus and O. tauri (Palenik et al. 2007), and Micromonas sp. RCC299 and M. pusilla CCMP1545 (Worden et al. 2009) (fig. 3). First, we searched for canonical tRNA genes and then examined their location on the genomes. As a result, we found tandem-duplicated tRNA genes, composed of a single tRNA species, in the genomes of C. reinhardtii and O. lucimarinus (supplementary fig. 2, Supplementary Material online). These tRNA genes are essentially identical to the proposed ancestral form of permuted tRNA genes (Soma et al. 2007; Di Giulio 2008). Then we examined if permuted tRNA genes were found in these green algal genomes using the SPLITS program. Prediction by SPLITS identified gene models for novel permuted tRNA genes in prasinophyte genomes (Ostreococcus and Micromonas) (supplementary fig. 3, Supplementary Material online), but not in C. reinhardtii. Permuted tRNA genes found in prasinophytes are likely essential because their anticodons recognize the codons that are not covered by canonical tRNA genes (e.g., the codon table of the O. lucimarinus nuclear genome in supplementary fig. 1, Supplementary Material online). Permuted tRNA\textsuperscript{Ser} genes are found in all four prasinophyte genomes, as in the B. natans nucleomorph. The tRNA\textsuperscript{Ser}(UGA) gene in O. lucimarinus is predicted to possess an additional intron outside the anticodon loop. In the O. lucimarinus genome, large chromosomal regions (about 150 kb) that include tRNA\textsuperscript{Ser}(GCA) genes are duplicated on chromosome 14, and downstream of these two identical tRNA\textsuperscript{Ser}(GCA) genes, the 5′ halves were additionally duplicated (supplementary fig. 3A, Supplementary Material online).

**Discussion**

We found novel permuted tRNA genes in the nucleomorph genome of the chlorarachniophyte B. natans, which are similar in structure and processing manner (circularization and cleaving the intervening sequence and terminal CCA sequence addition; see fig. 1) to the permuted genes first reported in C. merolae (Soma et al. 2007). The intervening sequences of both B. natans genes are 5 bp, the shortest examined so far, possibly reflecting the compact nature of the nucleomorph genome. Our data showed that the sequence structures and the possible mechanism for expressing mature tRNA from permuted genes were similar in both organisms. One can suppose that, in the chlorarachniophyte nucleomorph, a continuous trend of genome reduction might pressure the redundantly duplicated tRNA genes to invent a novel permuted tRNA gene structure, potentially able to be processed into the mature form using the eukaryotic RNA processing system. Actually, we identified tandem repeats of duplicated tRNA genes composed of single tRNA species in the nuclear genomes of the green algae C. reinhardtii and O. lucimarinus (supplementary fig. 2, Supplementary Material online), which are closely related to the chlorarachniophyte
nucleomorph (Rogers et al. 2007; Takahashi et al. 2007). These duplicated tRNA genes are structurally identical to the plausible intermediate stage of permuted tRNA evolution in previously proposed models (Soma et al. 2007; Di Giulio 2008). In addition, a previous study identified an unusually large number of tRNA genes and a number of recent tRNA duplications within the tRNA gene clusters in the C. reinhardtii nuclear genome (Merchant et al. 2007). These data suggest that the green algal ancestor of the chlorarachniophyte nucleomorph had a potential to generate permuted tRNAs through tRNA gene duplication.

In addition, the presence of permuted tRNA genes in the nuclear genomes of the prasinophyte algae Ostreococcus and Micromonas (fig. 3; supplementary fig. 3, Supplementary Material online) was predicted by SPLITS. Almost all the permuted tRNAs identified in this study are members of the tRNA type spliced at position 37/38, except for a tRNA$^{\text{Ser}}$(CGA) in O. tauri, which has the junction at the D loop. One tRNA$^{\text{Ser}}$(UGA) in O. lucimarinus has a unique intron at position 27/28 on the anticodon stem, rarely found in eukaryotes and archaea (Kawach et al. 2005; Sugahara et al. 2008). The secondary structures of these permuted tRNAs possess BHB or BHL consensus splicing motifs, which are conserved in intron-containing tRNAs found in archaea (Randau, Calvin, et al. 2005; Tocchini-Valentini et al. 2005) and most permuted tRNAs in C. merolae (Soma et al. 2007). The eukaryotic splicing endonuclease has been shown to cleave the intron at a fixed distance from the mature tRNA domain (Reyes and Abelson 1988), although it is also able to recognize and cleave the BHB motif like archaeal enzymes (Fabbri et al. 1998; Di Segni et al. 2005). Secondary structure analyses demonstrated that BHB and BHL motifs are formed at the intron–exon boundary in the red alga–type tRNA genes (G. theta nucleomorph, Zauner et al. 2000 and C. merolae nuclear tRNAs, Soma et al. 2007), as well as the green plant lineage (supplementary fig. 4, Supplementary Material online), suggesting that BHB motif-based splicing is an evolutionarily conserved system to process the permuted and intron-containing tRNA genes in plant nuclear genomes. Although it is possible that the B. natans nucleomorph and the prasinophytes independently acquired permuted tRNA genes via the reduction of duplicated tRNA genes, it is more parsimonious and likely that the green algal ancestor of the chlorarachniophyte nucleomorph already possessed permuted tRNA genes prior to the secondary endosymbiosis (fig. 3). It is not known why permuted tRNA$^{\text{Ser}}$ genes were commonly found in green algae and if there might be functional constraints on this tRNA species.

Is there any common background or selective pressure at the birth of these peculiar tRNA molecules in red and green algae? Or did they simply evolve in a parallel way? Prior to this analysis, one hypothesis proposed that permuted tRNA structures had been of help to prevent the integration of viruses and other mobile genetic elements by disrupting contiguous tRNA gene structures known to be highly conserved and typical target sites for viral integration, which assumed relatively “recent” origins of permuted tRNA (Randau and Söll 2008). Another hypothesis suggested that permuted tRNA genes might be a plesiomorphic form of all tRNAs with “ancient” origin (Di Giulio 2008). Based on the patchy distribution of the permuted tRNA species and its apparently small contribution to the large genomes of green plants, such as land plants and C. reinhardtii (fig. 3), one plausible scenario is that the common ancestor of eukaryotes (or at least plants)
had the potential to express and process RNA genes with permuted structure and, in each lineage of the offspring, permuted genes have been repeatedly lost and gained depending on the selective pressure that the genome has experienced. This explanation assumes that the extant permuted tRNA genes might be subject to relatively "neutral" selective pressure and that permuted tRNA genes could duplicate and regenerate into the canonical form of tRNA genes within genomes and vice versa. We postulate that permutation provides an alternative way to efficiently regenerate a singlet tRNA gene from duplicated genes along with the loss of flanking chromosomal regions while maintaining the essential tRNA gene set required for genome integrity (fig. 4).

Our results show that permuted tRNAs, which may be thought of as a peculiar and isolated phenomenon in a single red algal genome, are evolutionarily conserved and potentially fundamental to maintain genome function in photosynthetic eukaryotes. Given the evolutionary relationship between intron-containing tRNA and split tRNA (Fujishima et al. 2009), it is likely that permuted tRNA in eukaryotes, split tRNA in archaea, and intron-containing tRNA in both domains are evolutionarily related, and all three tRNA species share a similar molecular background (e.g., splicing machinery). In some eukaryotic lineages such as yeasts, another splicing mechanism for the tRNA intron might have secondarily evolved (Reyes and Abelson 1988). Our data also indicate that the permuted genes were retained within the nucleomorph and algal nuclear genomes through multiple rounds of genome reconstruction processes, such as endosymbiosis (Gilson et al. 2006), speciation (Palenik et al. 2007), and ecological adaptation (Worden et al. 2009). In this study, not all the "missing" tRNAs were identified in the B. natans nucleomorph (supplementary fig. 1, Supplementary Material online), and these tRNAs are probably imported from the nuclear genome into the periplastid space or encoded by unknown genes with unusual structures, which are not recognized by the currently available prediction programs. Further development of bioinformatic programs for tRNA prediction (Sugahara et al. 2007, 2008), as well as other algal genome sequences, will uncover the evolutionary history of permuted tRNA and help to understand the mechanism maintaining the integrity of eukaryotic genomes.

**Supplementary Material**

Supplementary figures 1–4 are available at Molecular Biology and Evolution online (http://mbe.oxfordjournals.org/).

**Acknowledgments**

We thank Dr Akiko Soma for helpful suggestions on tRNA extraction and RT–PCR experiments and Dr John M. Archibald for critical reading of the manuscript. Computation time was provided by the Super Computer System, Human Genome Center, Institute of Medical Science, University of Tokyo. Grants-in-Aid for Research Fellowships for Young Scientists (No. 20-9894 to S.M.) from the Japan Society for the Promotion of Science; Scientific Research (Nos 21657024 and 20247032 to H.N.) from The Ministry of Education, Culture, Sports, Science and Technology, Japan.

**References**


Gilson PR, Su V, Slamovits CH, Reith ME, Keeling PJ, McFadden GL. 2006. Complete nucleotide sequence of the chlorarachniophyte...


*Maruyama et al.* · doi:10.1093/molbev/msp313