Origin of the 16S Ribosomal Molecule from Ancestor tRNAs

Savio Torres de Farias 1,2,*, Thais Gaudêncio Régo 3 and Marco V. José 4

1 Laboratório de Genética Evolutiva Paulo Leminsk, Departamento de Biologia Molecular, Universidade Federal da Paraíba, João Pessoa 58051-900, Brazil
2 Departamento de Filosofia, Universidade Federal de Santa Catarina, Florianópolis 88040-900, SC, Brazil
3 Departamento de Informática, Universidade Federal da Paraíba, João Pessoa 58051-900, Brazil
4 Theoretical Biology Group, Instituto de Investigaciones Biomédicas, Universidad Nacional Autónoma de México, México D.F. C.P. 04510, Mexico
* Correspondence: stfarias@yahoo.com.br

Abstract: We tested the hypothesis that concatemers of ancestral tRNAs gave rise to the 16S ribosomal RNA. We built an ancestral sequence of proto-tRNAs that showed a significant identity of 51.69% and a percentage of structural identity of 0.941 with the 16S ribosomal molecule. We also propose a hypothesis for the emergence of translation.

Keywords: 16S rRNA; proto-tRNAs; evolution; translation system; origins

1. Introduction

The origin and evolution of the ribosome represent a challenge for modern biology. The centrality of this complex in the transition between the RNA world to the ribonucleoprotein world, put its origin at the center of the discussion around the origin of translation. Many researchers have focused their studies in the comprehension of the origin of the large ribosomal subunit, mainly the catalytic center of the ribosome, the peptidyl transferase center (PTC) [1–9]. Farias et al. [7] reconstructed the ancestor sequences of tRNAs, and they showed a similarity of 50.52% between concatemers of proto-tRNAs with the modern PTC. Root-Bernstein and Root-Bernstein [9] suggested, based in comparative analysis between tRNA and the PTC, that the catalytic center of the large ribosomal subunit originated from tRNAs and that could have operated as a primitive genome. Farias et al. [10] reconstructed from ancestral sequences of tRNA, a structural model of the ancestral PTC, and by molecular docking experiments they also found that the ancestral PTC worked as an attraction center for proto-tRNAs, which made possible the initial synthesis of the first peptides, but without codification. Although most studies have focused in the origin of the PTC, the comprehension of the origin and evolution of the small ribosomal subunit is essential for understanding the origin of the translation system. Currently, the small subunit of the ribosome is an important molecule in phylogenetic studies, and it was crucial for determining the three kingdoms of life (Eubacteria, Archaea, and Eukarya) and the existence of the last universal common ancestor (LUCA) [11]. Petrov et al. [12] proposed that the small ribosomal subunit at the beginnings acted as a cofactor, positioning the activated ends of tRNAs within the PTC such that the large ribosomal subunit and the small ribosomal subunit were organized via a coevolution process. Based on phylogenetic methods, and using the decomposition of the substructures, Harish and Caetano-Anolles [1] found that the smaller and larger ribosomal subunits evolved independently as ribonucleoprotein complexes and proposed a sequence of structural events that led to the formation of modern molecules. Bloch et al. [13,14] analyzed the similarities between
tRNAs and 16S ribosomal molecules, and they were the first to point out that both molecules may have a common origin. Based on the evidence that the tRNAs could have participated in the origin of the catalytic center of the large ribosomal subunit (PTC) and that they also could be involved in the origin of the 16S ribosomal molecules, we analyzed the likelihoods of similarity between ancestor tRNAs and the modern 16S molecules.

2. Results and Discussion

2.1. Sequence Analysis

The ancestral sequences for each of the tRNA types (Supplementary Material) were aligned individually with the *Thermus Thermophilus* 16S ribosomal sequence. This initial analysis allowed us to identify whether there was a similarity between individual tRNAs and different parts of this *T. Thermophilus* ribosomal molecule. In Table 1, the alignment between the individual canonical ancestor tRNAs and the 16S molecule from *T. thermophilus* with the initial position and the final position of similarity, as well as the percentage of similarity between them, are shown. The result shows that different types of tRNAs' ancestral sequences have similarities with a sequential portion in the modern molecule of the 16S ribosomal molecule, which can indicate that the primitive 16S molecules were formed by junctions of primitive or ancestral tRNAs/proto-tRNAs.

<table>
<thead>
<tr>
<th>tRNA</th>
<th>Initial Position</th>
<th>Final Position</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>InMet</td>
<td>180</td>
<td>240</td>
<td>40%</td>
</tr>
<tr>
<td>Asp</td>
<td>195</td>
<td>246</td>
<td>46.1%</td>
</tr>
<tr>
<td>Trp</td>
<td>202</td>
<td>246</td>
<td>45.5%</td>
</tr>
<tr>
<td>Glu</td>
<td>242</td>
<td>291</td>
<td>50%</td>
</tr>
<tr>
<td>Ser</td>
<td>283</td>
<td>307</td>
<td>60%</td>
</tr>
<tr>
<td>Sel</td>
<td>296</td>
<td>340</td>
<td>50%</td>
</tr>
<tr>
<td>Val</td>
<td>382</td>
<td>429</td>
<td>46.8%</td>
</tr>
<tr>
<td>Ile</td>
<td>394</td>
<td>432</td>
<td>48.7%</td>
</tr>
<tr>
<td>Gly</td>
<td>505</td>
<td>552</td>
<td>48.9%</td>
</tr>
<tr>
<td>His</td>
<td>507</td>
<td>561</td>
<td>31.03%</td>
</tr>
<tr>
<td>Thr</td>
<td>670</td>
<td>710</td>
<td>58.9%</td>
</tr>
<tr>
<td>Cys</td>
<td>736</td>
<td>789</td>
<td>44.2%</td>
</tr>
<tr>
<td>Gln</td>
<td>831</td>
<td>877</td>
<td>40%</td>
</tr>
<tr>
<td>Tyr</td>
<td>881</td>
<td>916</td>
<td>48.5%</td>
</tr>
<tr>
<td>Pro</td>
<td>893</td>
<td>950</td>
<td>47.36%</td>
</tr>
<tr>
<td>Ala</td>
<td>1003</td>
<td>1044</td>
<td>56.4%</td>
</tr>
<tr>
<td>Lys</td>
<td>1013</td>
<td>1048</td>
<td>57.14%</td>
</tr>
<tr>
<td>Asn</td>
<td>1028</td>
<td>1085</td>
<td>50.1%</td>
</tr>
<tr>
<td>Arg</td>
<td>1155</td>
<td>1188</td>
<td>47%</td>
</tr>
<tr>
<td>Met</td>
<td>1270</td>
<td>1319</td>
<td>42.8%</td>
</tr>
<tr>
<td>Phe</td>
<td>1323</td>
<td>1371</td>
<td>50%</td>
</tr>
<tr>
<td>Leu</td>
<td>1461</td>
<td>1499</td>
<td>56.4%</td>
</tr>
</tbody>
</table>

From the similarities observed in the alignment between the individual ancestral tRNAs and the 16S ribosomal molecule, we observed that the ancestral tRNAs had similarities with different portions. It was also observed that the positions were sequential for a group of ancestral tRNA sequences in the *T. thermophilus* 16S ribosomal molecule. The ancestral tRNAs that showed to align in sequential positions in 3’ terminal was used to construct a concatemer. The concatamer formed by joining the sequences of Gln/Tyr-Pro-Ala/lys-Asn-Arg-Met-Phe-Leu, ancestral tRNAs were aligned with the complete 16S ribosomal molecule. Figure 1 shows the alignments between concatemers formed by the ancestor tRNAs sequence Gln-Pro-Ala-Asn-Arg-Met-Phe-Leu and the part of the 16S
molecule that had similarity. In the alignment, there were 325 nucleotides in the ancestor sequence with 168 nucleotides with exact base matches, corresponding to 51.69% of similarities without gaps. The concatemers had a match with the 16S molecule from *T. thermophilus* in the upper 3′ domain of positions 943 to 1286. Petrov et al. [12] suggested that this portion of the small subunit had origins at the beginning of the organization of this subunit, and thereby, it is an early part of this molecule [12].

The alternative sequences with Lys at the Ala position (Figure 1—Supplementary Material) and Tyr at the Gln position (Figure 2—Supplementary Material) had lower similarity than the sequence used in Figure 1. A concatemer was also constructed with the ancestral tRNA sequences that showed similarity to the 5′ portion of the *T. thermophilus* 16S ribosomal sequence. The ancestral tRNAs used in the construction of the concatemer were sequentially: Asp-Glu-Ser-Sel-Val-Ile-Gly. The alignment yielded 39.68% of similarity with the sequence between the positions 191-468 of *T. thermophilus* 16S ribosome (Figure 3—Supplementary Material). The sequence obtained by the concatemer construct was aligned with the ancestral sequence obtained from modern 16S ribosomal sequences. The alignment of these sequences showed a similarity of 39.07% from the first position to the final position that had match (Figure 4—Supplementary Material). The similarity difference observed between the alignments of the 16S ribosomal ancestor (39.07%) concatemer sequence and the *T. thermophilus* 16S ribosomal modern sequence concatamer sequence (51.69%) can be explained by a bias in the methodology, since the method can be influenced by the regions of greater conservation, as well as by the chosen evolutionary model.

![Figure 1. Alignment between the concatemers of ancestral tRNAs sequences and the modern 16S ribosomal sequence from *T. thermophilus*. The consensus is the indication that the same nucleotide is in the same position, however, in the position with gap, only one position is accept to construct the consensus sequence. The occupancy indicates whether there is no or there is gap in that position.](image-url)

Similar results were obtained by Farias et al. [7] and Root-Bernstein and Root-Bernstein [9] for the 16s and 23S ribosomal sequence. An important point to note is that Farias et al [6] used the method of reconstructing ancestral sequences as a query, while Root-Bernstein and Root-Bernstein [8] used modern tRNA sequences. Notwithstanding the use of different methodologies, they produced the same results. Root-Bernstein and Root-Bernstein [9] observed that between positions 943 and 1286 there were similarities with 8 tRNAs, 4 of which are also observed in this work (Gln, Lys, Phe, and Met). One noteworthy observation is the similar position of leucine tRNA in helix-44 in the work by Root-Bernstein and Root-Bernstein [9] and in the present work. The same helix was also described...
by Harish and Caetano-Anolles [1] in a study on ribosome evolution as the oldest portion of the 16S ribosomal molecule. Caetano-Anolles and Caetano-Anolles [15] analyzed the antiquity of ribosomal RNA, as well as tRNAs, in which they demonstrated that there was a coevolution process, older tRNAs showed homology to older regions of ribosomal RNA, with Ser and leu being the most abundant in these regions.

Farias et al. [7] suggested that the PTC was built by concatemers of ancestor tRNAs, and the PTC in modern organisms was found at the 3’ terminal. Here, our data suggest that the small ribosomal subunit was also built by ancestor tRNAs and was also located at the 3’ terminal. These similarities suggest a common pattern in the evolution of the ribosomal RNA, always with the 5’ portion being more recent than the 3’ portion. The 3’ major domain of the small subunit was part of the P and A sites, which suggests that this was important for stabilizing the P site [12]. The upper 3’ domain has been described as important for interacting with the tRNA anticodon in the A site [2]. It is also suggested that the initial organization and evolution of the small and large ribosomal subunit had an independent evolution [12].

2.2. Structural Analysis

From the ancestor tRNAs concatemers, we reconstructed the tri-dimensional (3D) structure of the 16S molecule via homology. In Figure 2, we show the model generated for the ancestor 16S molecule and the structural alignment superimposition between the 3D model for the ancestor 16S molecule and the 16S molecule from T. thermophilus, Escherichia coli, and Mycobacterium smegmatis. The structural alignment with T. thermophilus had a value of root-mean-square deviation (RMSD) = 0.595 and a percentage of structural identity (PSI) = 0.941, E. coli had the values of RMSD = 1.159 and PSI = 0.906, and Mycobacterium smegmatis had the values of RMSD = 1.201 and PSI = 0.861.

![Figure 2](image_url)

**Figure 2.** (A) Structural alignment generated between ancestral concatemers of tRNAs (blue), and modern 16S ribosomal molecule from T. thermophilus (red) with RMSD = 0.602 and PSI = 0.940, E. coli (yellow) with RMSD = 1.159 and PSI = 0.906, and Mycobacterium smegmatis (green) with RMSD = 1.201 and PSI = 0.861. (B) Consensual tridimensional model derivate from multiple structural alignment.

The results suggest that during the evolutionary process, the structure of the 16S molecule was undergoing strong selective pressure to maintain their structure, as we can observe in the high similarity between the structural ancestor predicted structure and modern. Despite a moderate level of similarity in sequence, we can infer that the selection process was to maintain the structure, which is directly related to the function of this molecule. In Figure 3, the modern small subunit (green) complexed with mRNA (orange) and tRNAs (blue), and the red section highlights the portion of the modern small subunit that had similarities with the concatemers of ancestral tRNAs. The reconstructed ancestral structure indicated that the first portion was involved in the tRNA anticodon interaction, and together
with the central domain and the bottom 3’ domain, formed the decoding site. In spite of participating structurally the decoding site, the upper 3’ domain does not interact directly with the mRNA, with this function being developed by the bottom 3’ domain [16,17].

Figure 2. (A) Structural alignment generated between ancestral concatemers of tRNAs (blue), and modern 16S ribosomal molecule from *T. thermophilus* (red) with RMSD = 0.602 and PSI = 0.940, *E. coli* (yellow) with RMSD = 1.159 and PSI = 0.906, and *Mycobacterium smegmatis* (green) with RMSD = 1.201 and PSI = 0.861. (B) Consensual tridimensional model derivate from multiple structural alignment.

Figure 3. The *T. thermophilus* small ribosomal subunit without proteins and complexed with tRNAs (blue) and mRNA (orange). In red, the portion of the 16S rRNA that has similarities with the concatemers of ancestral tRNAs is shown.

This result suggests that initially, the small ribosomal subunit worked just as a binding domain. Farias et al. [10] proposed that the initial peptide synthesis occurred without codification and only with the emergence of the first mRNA, and the anticodon stem loop was co-opted to the decoding system. Thus, the results presented in this work are in line with the proposition that the first peptides were synthesized without codification and that the small and large subunit had independent evolutionary histories.

2.3. A Hypothesis of the Emergence of Translation

In Figure 4, we propose a model for the emergence of the primitive translation system. The primitive small and large subunits were formed via junctions of the proto-tRNAs. The first part of the small ribosomal subunit worked as a point to bind RNAs in an open structure configuration and the primitive large subunit (PTC) worked as a ribozyme to carry out peptide bonds randomly without decoding information [10]. Farias et al. [10] demonstrated the possibility that the tRNA stem loop interacted with the PTC via the corresponding primitive acceptor site. Farias et al. [18] demonstrated that this same portion of the tRNA could interact with the primitive aminoacyl-tRNA synthetase, thus serving as an amino acid delivery system to the emergent PTC. At this stage, the primitive small and large subunits worked independently (Figure 4B). The present model is in line with the model presented...
by Harish and Caetano-Anolles [1], in which the ribosome subunits, in their origin, had independent evolutionary processes and that the subunit junction, with consequent primitive ribosome formation, was mediated by interactions with tRNA molecules.

With the accretion of the central domain and the bottom 3’ domain in the small subunit, the decoding site appeared, and then it was possible the interaction between this portion and the nascent primitive mRNA. The accretion of new parts in the small subunit made possible an increase of the affinity between the small and large subunit, and it was possible for the interaction between the anticodon stem loop and the PTC with the mRNA connecting the decoding site in the primitive small ribosomal subunit (Figure 4C). With the interaction between the subunits, the tRNA anticodon stem loop, that initially worked as a delivery system to the PTC [10], was co-opted to interact with

**Figure 4.** Hypothesis for the emergence of the primitive translation system. (A) The primitive structure of the primitive tRNA anticodon stem loop and small and large subunit as obtained from junctions of ancestral tRNAs. (B) Interactions between the primitive small subunit with the primitive mRNA and the PTC, and with the tRNA anticodon stem loop. (C) Accretion of new parts and increased stability of the small and large subunit made possible the interaction between small and large subunit and emergence of the primitive translation system.
the mRNA bound in the small subunit, and precisely at this moment, the first correlation between codon-anticodon was established and may have been the first step towards the organization of the genetic code as we know it. Caetano-Anolles and Caetano-Anolles [15] also suggested a similar scenario for the organization of the primitive translation system, with proto-tRNAs assisting in the formation of ribosomal subunits and consequently in the formation of the primitive genome, as well as playing an important role in the formation and establishment of the primitive genetic code, which enabled the formation of the first peptides. In their model the whole process occurs in coevolution until the assembly of the primitive translation system, thus being, in its origin, a ribonucloproteic system [15].

3. Material and Methods

3.1. Ribosomal Ancestral Sequence Reconstruction and Alignment

The tRNA ancestral sequence were obtained from Farias [19] (Supplementary material). We analyzed the individual similarity between each type of ancestor tRNAs and the 16S ribosomal molecule of *Thermus thermophilus*. From the ancestor tRNAs that had similarities in sequential position in the 16S ribosomal sequence of *T. thermophilus*, we constructed a concatemer sequence to construct a global alignment with the 16S ribosomal molecule. From the individual alignment, we constructed the concatemers with the cognate tRNAs to: Gln/Tyr-Pro-Ala/Lys-Asn-Arg-Met-Phe-Leu. For the reconstruction of the 16S ancestral ribosomal sequence were used ([19]) (Supplementary material), 93 Archaea sequences, 110 Bacteria sequences, and 100 Eukarya sequences. The best evolutionary model was TN93+I+G, gamma 0.81 and invariable site 0.11. To the alignment was used the Maff webserver [20]. To the best evolutionary model and the ancestral reconstruction was used MEGA software Version 6 [21].

The concatemers’ ancestral sequence of tRNAs and the current *T. thermophilus*’ 16S sequence, as well as the ancestral sequence of tRNA and the 16S ancestral sequence that had similarities were aligned using the Maff webserver [20].

3.2. Structural Reconstruction

From the concatemers’ ancestral sequence, the 3D-structure was modeled using the ModeRNA Server [22]. After reconstruction of the 3D-model of the ancestral 16S molecule, we aligned the structure of the model with the template from *T. thermophilus* (ID4V8X), *E. coli* (4ADV), and *Mycobacterium smegmatis* (5XYU) using the Setter program [23].

Supplementary Materials: The following are available online at http://www.mdpi.com/2413-4155/1/2/46/s1.


Funding: MVJ was financially supported by PAPIIT-IN201019; UNAM; México.

Conflicts of Interest: The authors declare no conflict of interest.

References

10. Farias, S.T.; Rêgo, T.G.; José, M.V. Peptidyl Transferase Center and the Emergence of the Translation System. Life 2017, 7, 21. [CrossRef]

© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).