



Identified Hybrid tRNA Structure Genes in Archaeal Genome

Uttam Roy Mandal ^{1,*}, Shib Sankar Das ², Brajadulal Chattopadhyay ³, Satyabrata Sahoo ⁴

¹ Department of Mathematics, Raidighi College, Raidighi, W.B., India

² Department of Mathematics, Uluberia College, Uluberia, Howrah, W.B, India

³ Department of Physics, Jadavpur University, Kolkata, W.B, India

⁴ Department of Physics, Dhruva Chand Halder College, Dakshin Barasat, W.B., India

* **Corresponding author:** Uttam Roy Mandal, Department of Mathematics, Raidighi College, Raidighi, W.B., India. Tel: +91-8335822418; E-mail: urmandal77@raidighicollege.in

Abstract

Background: In Archaea, previous studies have revealed the presence of multiple intron-containing tRNAs and split tRNAs. The full unexpurgated analysis of archaeal tRNA genes remains a challenging task in the field of bioinformatics, because of the presence of various types of hidden tRNA genes in archaea. Here, we suggested a computational method that searched for widely separated genes encoding tRNA halves to generate suppressive variants of missing tRNAs.

Objectives: The exploration of tRNA genes from a genome with varying hypotheses, among all three domain of life (eukaryotes, bacteria and archaea), has been rapidly identified in different ways in the field of bioinformatics. Like eukaryotic tRNA genes, it has been established that two separated regions of the coding sequence of a tRNA gene are essential and sufficient for promotion of transcription. Our objective is to find out the two essential regions in the genome sequence which comprises two halves of the hidden tRNAs.

Material and Methods: Considering the existence of split tRNA genes widely separated throughout the genome, we developed our tRNA search algorithm to predict such separated tRNA genes by searching both a conserved terminal 5'- and 3'-motif of tRNA in agreement with the split hypothesis on the basis of cloverleaf prediction and precise insilico determination of bulge-helix-bulge secondary structure at the splice sites.

Results: By a comprehensive search for all kinds of missing tRNA genes, we have constructed hybrid tRNA genes containing one essential region from tDNA (XYZ) and the other from tDNA (ABC), both from same species in the archaea. We have also found, this type of hybrid tRNA genes are identified in the different species of the archaea (XYZ: ASN, ARG and MET; ABC: ASP, SER, ARG and PRO). These hybrid split tRNA share a common structural motif called bulge-helix-bulge (BHB) a more relaxed bulge-helix loop (BHL), at the leader exon boundary and suggested to be evolutionary interrelated.

Conclusions: Analysis of the complete genome sequences of *Metallosphaera sedula* DSM 5348, *Desulfurococcus kamchatkensis* 1221n and *Ignicoccus hospitalis* KIN4/I in archaea by our algorithm revealed that a number of hybrid tRNAs are constructed from different tDNAs. Asymmetric combination of 5' and 3' tRNA halves may have generated the diversity of tRNA molecules. Our study of hybrid tRNA genes will provide a new molecular basis for upcoming tRNA studies.

Keywords: Helix (Snails); Open Reading Frames; RNA, Transfer

1. Background

Extremophiles, especially those in Archaea, have a myriad of adaptations that keep their cellular proteins stable and active under the extreme conditions in which they live. Rather than having one basic set of adaptations that works for all environments, Archaea have evolved separate protein features that are customized for each

environment (1, 2). The evolution of tRNA in archaea is an exciting topic widely discussed in the field of molecular evolution (3-7). tRNA is a short-coding RNA charged with a corresponding amino acid and delivered into ribosome during protein biosynthesis. In archaca, there have been four types of tRNA genes identified to date: intron-less,

intron-containing, split tRNA and permuted tRNA (8-10). Coding sequence for these tRNAs are separated with introns, fragmented, or permuted at the genome level. Split tRNA genes are encoded on two or three separate regions on genome and then these are processed into single tRNA gene (11, 12). Intron-containing and split tRNA gene share a common structural motif called bulge-helix-bulge (BHB), or more relaxed bulge-helix-loop (BHL), at the intron/ leader- exon boundary and these are suggested to be evolutionarily interrelated (12-14). Structural motif and location of tRNA introns in archaea have been affected by the change in the recognition and activity of the different unit compositions in the tRNA splicing endonucleases (15). But it is still unclear how tRNA molecules originated, evolutionary biologists continue to question how these chains of ribonucleotides became involved in the context of protein synthesis, and how they influenced the evolution of these biological system. When in the course of molecular evolution did tRNA molecule and its characteristic cloverleaf structure emerged is still an ongoing debate, however several evolutionary models representing the origin and convergence of proto-tRNA have been proposed (4, 16, 17).

It has been firmly established that the genes transcribed by RNA polymerase III contains their transcriptional signals (promoters) within the coding sequence. In 5S RNA genes, a 30-base pair (bp) region seems to contain all the information necessary for correct initiation of transcription (18, 19). In the VAI RNA gene of adenovirus, an internal control region, approximately 60 bp long, could be identified (20, 21).

In case of tRNA genes, a more precise analysis leads to the identification of two sequences of about 10 nucleotides each, located within the coding region, whose presence is essential for transcription (22, 23). The structure of tRNA is very similar in eukaryotes and prokaryotes (24). All tRNA genes are transcribed by RNA polymerase III (25) and it is reasonable to expect that the transcriptional signals must correspond to features common to all tRNA genes.

It is well known that tRNA plays a central role in the translation machinery and the molecular representation of the genetic code provides the mechanism to translate the genetic code from a three-base mRNA codon to the correct amino acid. Several experimental verifications led to the cloverleaf secondary structure model of tRNA known today. Throughout the translation process, it recognizes several protein and RNA molecules. What individual crystal structures are unable to provide is a sense of the embedded dynamic nature of tRNA. tRNA has an inherent flexibility that is critical for interactions with cellular proteins and its

ultimate role in protein biosynthesis. All characterized tRNA species bear numerous modifications of their bases. Whereas many of these modifications have significant effects on translation. As sequence and structure are evolutionarily conserved to preserve the function of biological molecules, residues critical for function can be determined computationally prior to simulation or experiment. Increasing numbers of research groups are now applying computational methods to understand tRNA-protein interactions pertaining to components of the protein biosynthetic machinery. While crystal structures clearly reveal distortions in tRNA structure, computational methods provide additional evidence for the inherent flexibility of tRNA structure under a variety of conditions and for differing tRNA species. Of particular interest are the effects of modified nucleosides in stabilizing tRNA structure. Compared to the ease of preparing unmodified tRNA by in transcription methods, obtaining modified nucleoside for tRNA dynamics is much difficult *in vitro*. Thus computational analysis of tRNA sequences is more challenging and remains active fields in bioinformatics research.

There are many hidden patterns around in genome sequences. Many sophisticated non-linear algorithms (26) exist for pattern formation and recognition (27, 28).

For instance, there are several computational approaches to detect tRNA genes from a genome (29). To identify these on the sequences, there are number of algorithms extensively used in this field. Notable amongst these are tRNAScan-SE (30) and ARAGORN (30).

Most of these tRNA- search programs key on primary sequence patterns and /or secondary structures specific to tRNAs. Some of the tRNA genes are either misidentified or missed by existing search algorithms. We developed our algorithm to circumvent some of the difficulties in searching missing tRNA sequences by existing programs. Our algorithm searched for two halves of cloverleaf, 1-37 and 38-72, from 5' and 3' end respectively, allowed maximum one non-Watson -Crick base pairing and the possibility of a single putative intron with the standard BHB structure at the exon intron boundary.

In eukaryotic tRNA genes it has been established that two separated regions with in the coding sequence of a tRNA gene are essential and sufficient for promotion of transcription (23, 24).

In search of missing tRNA genes in the genome of *Metallosphaera sedula* DSM 5348, *Desulfurococcus kamchatkensis* 1221n and *Ignicoccus hospitalis* KIN4/I, we inferred from our results that the two "essential regions" of the existing tRNAs can be combined to form a tRNA which we called as hybrid tRNA.

2. Objectives

In silico prediction of all tRNAs in a complete genome of an organism still remains an unsolved problem in the field of bioinformatics. The proper identification of tRNAs is critical for its essential role in the translation of mRNA into an amino acid sequence of a protein. So, the exploration of tRNA genes from a genome has been done with varying hypotheses in all three domain of life and has been rapidly identified in different ways in the field of bioinformatics. In eukaryotic tRNA genes, it has been established that two separated regions of the coding sequence of a tRNA gene are essential and sufficient for promotion of transcription. Our objective is to find out the two essential regions in the genome sequence which comprises two halves of the hidden tRNAs. In the present communication, we have developed a computational methodology towards genome analysis that searches for fragmented parts of a tRNA gene in widely separated coding sequences.

3. Materials and Methods

The entire genome sequences were obtained from National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/>). In the present communication, we focused our study on three archaeal genomes namely, *M. sedula* DSM 5348, (GenBank accession no NC_009440), *Desulfurococcus kamchatkensis* 1221n (GenBank accession no NC_011766) and *Ignicoccus hospitalis* KIN4/I (GenBank accession no NC_009776).

We developed a computational approach which was aimed at searching for any type of tRNA genes and was especially focused on intron containing tRNAs (and split tRNAs) not identified by tRNAscan-SE and ARAGORN. Here we have developed a procedure for generating and optimizing pattern descriptors that can be used to find structural motifs of tRNA. We have searched two halves of cloverleaf (split at 37/38 position), structure of tRNA from 5' end and 3' end respectively and then matched the base pairing of various stems with its required length. Up to one non-Watson-Crick base pairing in any of the four stem and a one in 3D-base pairs. We also adopted the standard cloverleaf model for studying the secondary structure of predicted tDNAs with salient features that i) T8, G18, R19, R53, Y55 and A58 were considered as conserved bases for archaeal tRNAs. ii) Base positions optionally occupied in D-loop were 17, 17a, 17b, 20a, 20b and 20c. iii) Extra arm or V-arm was taken into consideration. The constraint on length of V-arm was restricted to be less than 21 base. The exon-intron boundaries from a folded RNA structure, with 3nt or more nucleotide bulges on opposite strands on an RNA, separated by 4bp

central helix. The introns and split position constrained to harbor the Bulge-Helix-Bulge (BHB) secondary structure for splicing out during tRNA maturation. With these features in our algorithm, we extracted two different sets of tRNA pieces, one essential region from one tRNA and other from another tRNA. We determined the composition of a 5' half (one piece tRNA) deriving from one tDNA and the 3' half (other piece tRNA) from other tDNA from the above archaeal species. We constructed a new tRNA which we called a hybrid tRNA. Remarkably we have identified hybrid tRNA genes in the different species in the archaea.

4. Results

4.1. Construction of a Hybrid tRNA ASN(GTT)-tRNA ASP(GTC) Gene

The secondary structure of all tRNA molecules requires that there must be inverted repeats in all tRNA genes. In order to investigate the influence on transcription of the overall partial self-complementarity of a tRNA gene and in a comprehensive search of missing tRNA genes in the archaeal genome by our bioinformatics tool, we have identified a hybrid gene carrying one region from tDNA ASN(GTT) and other region from tDNA ASP(GTC) of *M. sedula* DSM 5348 i.e. the composition of a 5' half derived from tDNA ASN(GTT) and the 3' half from tDNA ASP(GTC) constructed a hybrid gene tRNA LEU(TAA). We now analyze in silico the secondary structure at the exon intron boundary – the bulge-helix-bulge (BHB). This is illustrated in (Fig. 1). It is the conformational structure most easily recognized and processed by archaeal splicing mechanism.

4.2. Construction of a Hybrid tRNA ASN(GTT) – tRNA SER(GGA) Gene

A comprehensive prediction of archaeal tRNA genes remains a challenging task in the field of bioinformatics, because various types of split tRNA genes are found in archaea. In search of missing tRNA genes in *Desulfurococcus kamchatkensis* 1221n in archaea, we have detected another hybrid tRNA gene containing one essential region from tDNA ASN(GTT) and the other essential region from tDNA SER(GGA) both from the same genome. Regardless of the overall regularities of the secondary structure of tRNA genes from the bioinformatics point of view, this hybrid tRNA gene has been identified as LEU(TAA). We also analyzed the secondary structure at the exon-intron boundary i.e. the bulge-helix-bulge (BHB). This is illustrated in (Fig. 2). It is the conformational structure most easily recognized and processed by archaeal splicing mechanism.

4.3. Construction of Hybrid tRNA ARG(TCT) – tRNA ARG(CCT) Gene and tRNA MET(CAT) – tRNA PRO(CGG) Gene.

Remarkably, our search algorithm identified hybrid tRNA genes GLU(TTC) and ARG(CCG) in *Ignicoccus hospitalis* KIN4/I. It has been observed that one essential region of the hybrid tRNA gene GLU(TTC) has been constructed from tDNA ARG(TCT) and the other region from

ARG(CCT) as illustrated in (Fig. 3) and the ARG(CCG) hybrid gene has been constructed from the composition of a 5' half derived from tDNA MET(CAT) and the 3' half from tDNA PRO(CGG) as illustrated in (Fig. 4). In both the cases we have analyzed the secondary structure of the bulge-helix –bulge (BHB) motif at the exon-intron boundary. It is the conformational structure most easily recognized and processed by archaeal splicing mechanism.

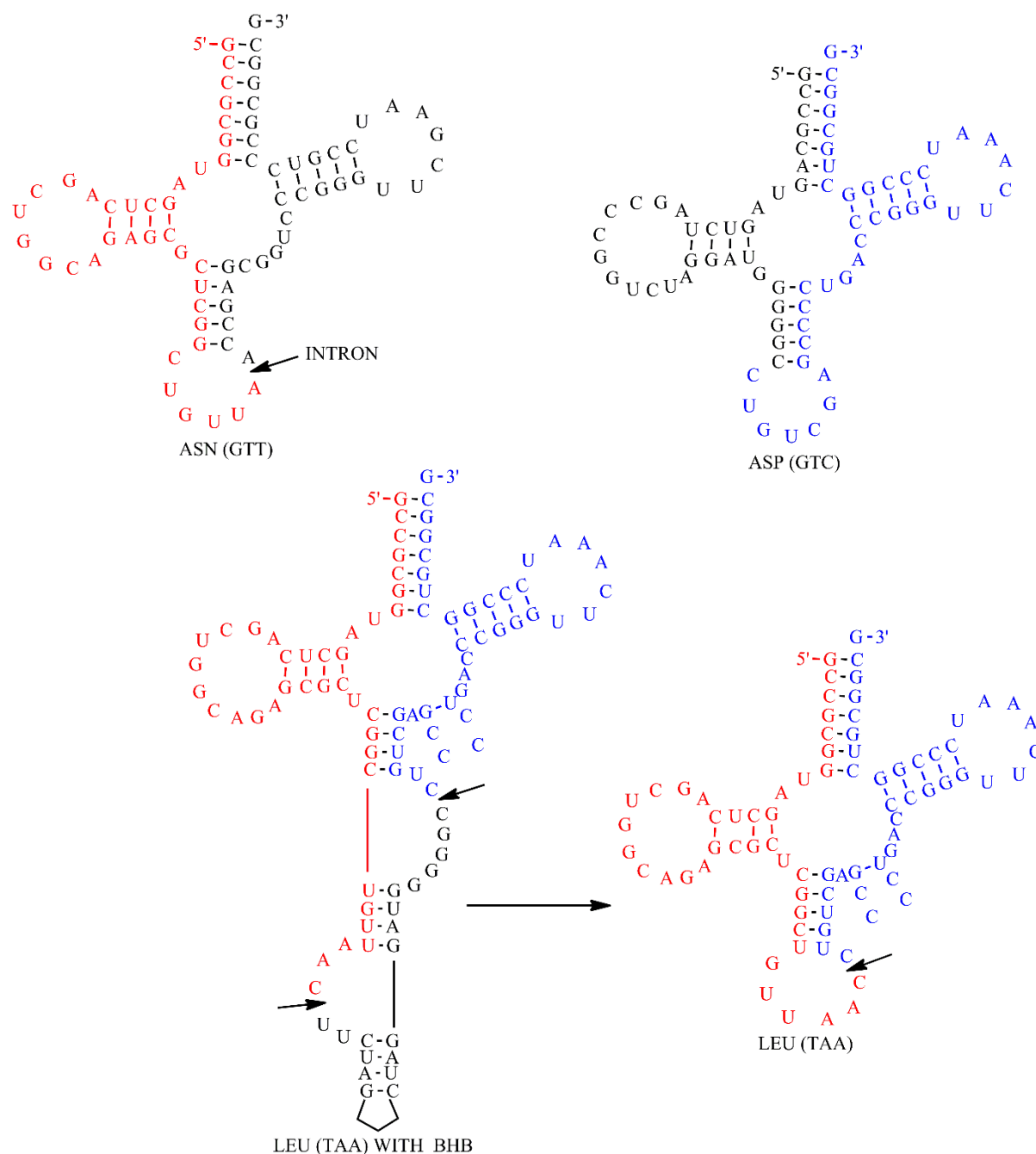


Figure 1. The locations and the gene sequences of the fragmented part with schematic diagrams of the tDNA ASN (GTT) and ASP (GTC) and predicted tRNA LEU (TAA) that split at 37/38 position along with the structures of the BHB motifs. Splice sites are indicated by arrows *Metallosphaera sedula* DSM 5348, ASN (GTT), C (2052586..2052621, 2052639..2052675), ASP (GTC), C (56542..56616), LEU (TAA), C (56543..56582, 2052637..2052675)

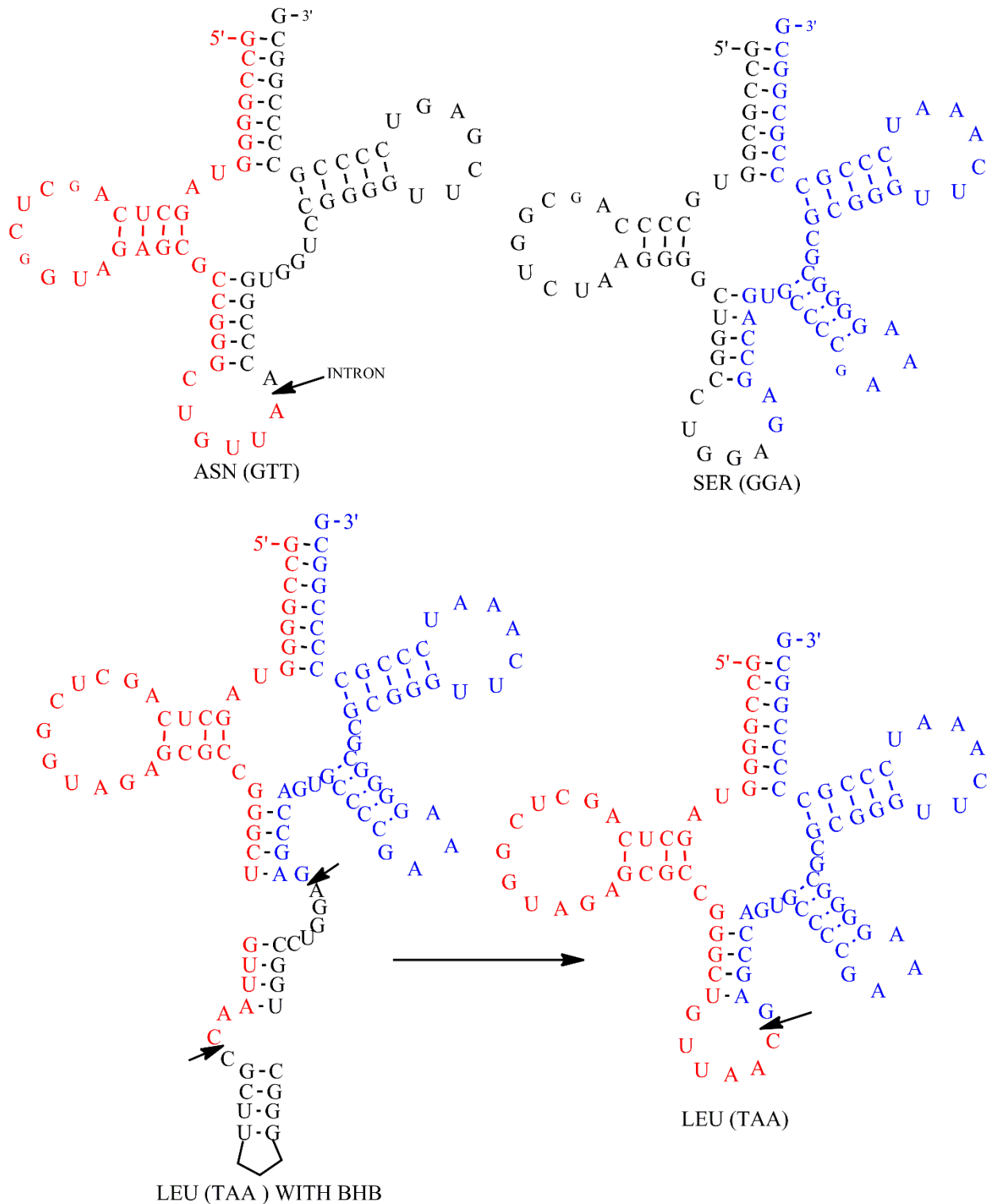


Figure 2. The locations and the gene sequences of the fragmented part with schematic diagrams of the tDNA ASN (GTT) and SER (GGA) and predicted tRNA LEU (TAA) that split at 37/38 position along with the structures of the BHB motifs. Splice sites are indicated by arrows *Desulfurococcus kamchatkensis* 1221n ASN (GTT), COB 1072..1231107,1231137..1231174) SER (GGA), C (1268007..1268092) LEU (TAA), C (1268008..1268052, 1231135..1231174)

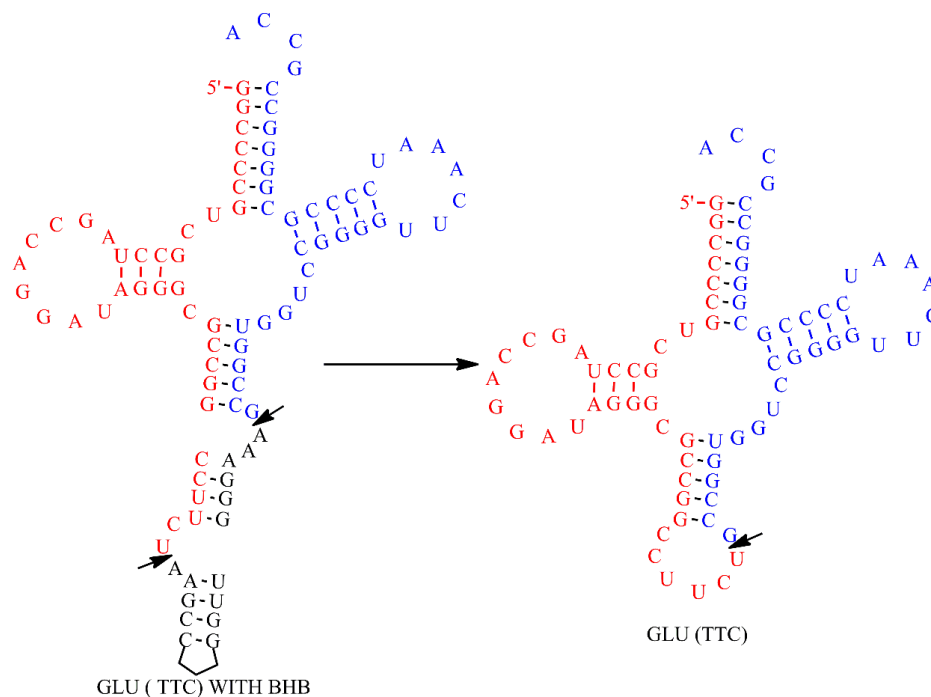


Figure 3. The locations of the tDNA ARG (TCT) and ARG (TCT) and the location and schematic diagrams of predicted tRNA GLU (TTC) that split at 37/38 position along with the structures of the BHB motifs. Splice sites are indicated by arrows
Ignicoccus hospitalis KIN4/1 ARG (TCT), (498947..499024) ARG (CCT), (1134528..1134624,1134567..1134585) GLU (TTC), (498947..498984,1134587-1134620)

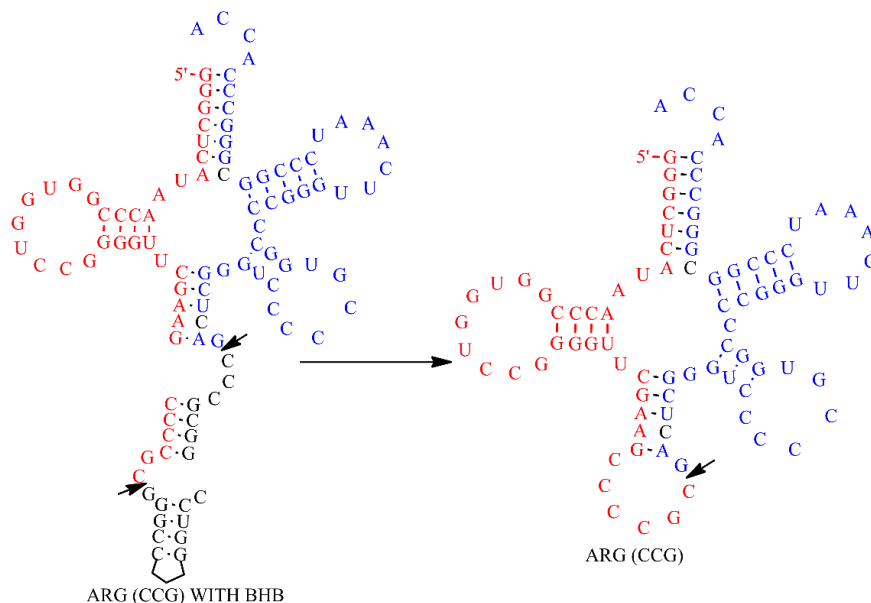


Figure 4. The locations of the tDNA MET (CAT) and PRO (CGG) and the location and schematic diagrams of predicted tRNA ARG (CCG) that split at 37/38 position along with the structures of the BHB motifs. Splice sites are indicated by arrows
 MET (CAT) (3936..4012) PRO (CGG) (539363-539477) ARG (CCG), (3965..4002, 539430..539473)

5. Discussion

The results of these analyses have shown that genetic information encoding functional RNAs is described in the genome cryptically. In spite of use of tRNAscan-SE along with the Split-tRNA-Search programme that led to the discovery of a variety of disrupted tRNA genes from the archaeal lineage, such as trans-spliced tRNAs (split tRNAs) joined at several positions in the cloverleaf structure (12-14) and cis-spliced tRNAs

containing one or multiple introns at non-canonical positions (31, 32), recently developed software has enabled the identification of additional distinct types of hybrid tRNA genes.

Although tRNAs are mostly transcribed from unfragmented genes, but occasionally also from split genes, with separated 5' and 3' halves. A reanalysis of the existing data on *Metallosphaera sedula* DSM 5348,

Desulfurococcus kamchatkensis 1221n and *Ignicoccus hospitalis* KIN4/I, hints of a novel hybrid gene that encodes both an un-fragmented and a 5'-split-half together in one. The corresponding 3'-complement-gene is located elsewhere in other gene on the genome. This hybrid of the split and the un-fragmented in one suggests a deeper synergy between the two, and hints of co-evolution. Furthermore, in a subtle contrast to the widely held idea of conservation of 3'-half, it is precisely the 3'-half that varies in these tRNAs; the 5'-half remains conserved. A further analysis also identified hybrid tRNA genes in an archaeal lineage highlighting the considerable diversity and wide distribution of tRNA gene disruption among organisms. While BHB motifs and the tRNA-intron splicing system must have been a prerequisite for the development of hybrid tRNA genes (33), their detailed mechanisms and physiological relevance remain unclear. Interestingly, split tRNAs and intron-containing tRNAs have common BHB consensus motifs at the boundaries, and hence require the same endonuclease. RNaseP (34) and tRNaseZ (35, 36) generally recognize the top half of the L-shaped tertiary structure of a tRNA corresponding to the acceptor-stem and the T-Ψ-C-arm, and do not require the mature body of the tRNA. Therefore, these enzymes may also perform endonucleolytic cleavage of the intervening sequence. Although some endoribonucleases require the linear ends of substrates to function (37), it is not known whether this condition holds for RNaseP and tRNaseZ.

As summarized in the model presented in the paper, maturation of hybrid pre-tRNAs probably starts with processing of the junction of the termini joined by the intervening sequence. The intervening sequence is then removed, possibly by RNaseP and tRNaseZ, which are universal endoribonucleases. Finally, the 3-terminal sequence is added to generate the functional acceptor-stem of the tRNA. Cleavage of the leader and trailer sequences at the junction of hybrid tRNAs is most likely performed by the tRNA-intron splicing machinery, because the sequences adjoining the processing sites potentially form a BHB motif, which is the dominant recognition element for archaeal tRNA-splicing endonucleases.

6. Conclusions

The comprehensive identification of the complete set of tRNAs in a given organism is of great interest in biology. In general, bioinformatics approaches for predicting tRNA loci from whole genome sequences are based on machine learning and pattern recognition. Although these methods correctly identified a number of tRNAs but also led to a high failure rate. Current method approaches this problem by identifying tRNA specific structures and sequence features, such as cloverleaf structures and base-pairing frequencies. Such features are then applied to whole genome scan windows to

predict the likelihood of a candidate locus being an authentic tRNA. The major drawback of this method is that, it is often too stringent to handle sequencing errors and natural variations in spliced products. Consequently it may produce high false negative rates and perform poorly on novel genomes. Secondly, many genomic sequences resemble tRNA cloverleaf structures, and additional information is required to eliminate such false positives. Considering these shortcomings in our present approach, we performed a comprehensive search of missing tRNA genes in the archaeal genomes of *M. sedula* DSM 5348, *D. kamchatkensis* 1221n and *I. hospitalis* KIN4/I by our bioinformatics tool. On the basis of the results presented in this paper, we concluded that the essential regions (5' half and 3' half) identified by our algorithm from different tDNAs constructed tRNA *de novo* which we called hybrid tRNA genes. These hybrid tRNA genes were actively involved to transport different amino acid to the ribosome for protein production. Our study of hybrid tRNA genes will provide a new molecular basis for upcoming tRNA studies. Asymmetric combination of 5' and 3' tRNA halves may have generated the diversity of tRNA molecules.

References

1. Cavicchioli R. General Characteristics and Important Model Organisms. In: Cavicchioli R, editor. *Archaea: Molecular and Cellular Biology*. Washington, DC, USA: ASM Press; 2007.
2. Schleper C, Jurgens G, Jonuscheit M. Genomic studies of uncultivated archaea. *Nat Rev Microbiol*. 2005;3(6):479-488. doi: 10.1038/nrmicro1159 pmid: 15931166
3. Maizels N, Weiner AM. Phylogeny from function: evidence from the molecular fossil record that tRNA originated in replication, not translation. *Proc Natl Acad Sci U S A*. 1994;91(15):6729-6734. doi: 10.1073/pnas.91.15.6729 pmid: 8041690
4. Weiner AM, Maizels N. The genomic tag hypothesis: modern viruses as molecular fossils of ancient strategies for genomic replication, and clues regarding the origin of protein synthesis. *Biol Bull*. 1999;196(3):327-328; discussion 329-330. doi: 10.2307/1542962 pmid: 10390830
5. Di Giulio M. The non-monophyletic origin of the tRNA molecule and the origin of genes only after the evolutionary stage of the last universal common ancestor (LUCA). *J Theor Biol*. 2006;240(3):343-352. doi: 10.1016/j.jtbi.2005.09.023 pmid: 16289209
6. Ertem G. Montmorillonite, oligonucleotides, RNA and origin of life. *Orig Life Evol Biosph*. 2004;34(6):549-570. doi: 10.1023/B:ORIG.0000043130.49790.a7 pmid: 15570708
7. Di Giulio M. The non-monophyletic origin of the tRNA molecule. *J Theor Biol*. 1999;197(3):403-414. doi: 10.1006/jtbi.1998.0882 pmid: 10089150
8. Tanaka T, Kikuchi Y. Origin of the cloverleaf shape of transfer RNA-the double-hairpin model: implication for the role of tRNA intron and the long extra loop. *Viva Origino*. 2001;29(134):134-142.

9. Marck C, Grosjean H. tRNomics: analysis of tRNA genes from 50 genomes of Eukarya, Archaea, and Bacteria reveals anticodon-sparing strategies and domain-specific features. *RNA*. 2002;**8**(10):1189-1232. doi: 10.1017/s1355838202022021 pmid: 12403461
10. Sugahara J, Yachie N, Sekine Y, Soma A, Matsui M, Tomita M, et al. SPLITS: a new program for predicting split and intron-containing tRNA genes at the genome level. *In Silico Biol*. 2006;**6**(5):411-418. pmid: 17274770
11. Sugahara J, Yachie N, Arakawa K, Tomita M. In silico screening of archaeal tRNA-encoding genes having multiple introns with bulge-helix-bulge splicing motifs. *RNA*. 2007;**13**(5):671-681. doi: 10.1261/rna.309507 pmid: 17369313
12. Randau L, Munch R, Hohn MJ, Jahn D, Soll D. Nanoarchaeum equitans creates functional tRNAs from separate genes for their 5'- and 3'-halves. *Nature*. 2005;**433**(7025):537-541. doi: 10.1038/nature03233 pmid: 15690044
13. Fujishima K, Sugahara J, Kikuta K, Hirano R, Sato A, Tomita M, et al. Tri-split tRNA is a transfer RNA made from 3 transcripts that provides insight into the evolution of fragmented tRNAs in archaea. *Proc Natl Acad Sci U S A*. 2009;**106**(8):2683-2687. doi: 10.1073/pnas.0808246106 pmid: 19190180
14. Randau L, Calvin K, Hall M, Yuan J, Podar M, Li H, et al. The heteromeric Nanoarchaeum equitans splicing endonuclease cleaves noncanonical bulge-helix-bulge motifs of joined tRNA halves. *Proc Natl Acad Sci U S A*. 2005;**102**(50):17934-17939. doi: 10.1073/pnas.0509197102 pmid: 16330750
15. Tocchini-Valentini GD, Fruscoloni P, Tocchini-Valentini GP. Coevolution of tRNA intron motifs and tRNA endonuclease architecture in Archaea. *Proc Natl Acad Sci U S A*. 2005;**102**(43):15418-15422. doi: 10.1073/pnas.0506750102 pmid: 16221764
16. Fujishima K, Sugahara J, Miller CS, Baker BJ, Di Giulio M, Takesue K, et al. A novel three-unit tRNA splicing endonuclease found in ultrasmall Archaea possesses broad substrate specificity. *Nucleic Acids Res*. 2011;**39**(22):9695-9704. doi: 10.1093/nar/gkr692 pmid: 21880595
17. Widmann J, Di Giulio M, Yarus M, Knight R. tRNA creation by hairpin duplication. *J Mol Evol*. 2005;**61**(4):524-530. doi: 10.1007/s00239-004-0315-1 pmid: 16155749
18. Sun FJ, Caetano-Anolles G. The origin and evolution of tRNA inferred from phylogenetic analysis of structure. *J Mol Evol*. 2008;**66**(1):21-35. doi: 10.1007/s00239-007-9050-8 pmid: 18058157
19. Sakonju S, Bogenhagen DF, Brown DD. A control region in the center of the 5S RNA gene directs specific initiation of transcription: I. The 5' border of the region. *Cell*. 1980;**19**(1):13-25. doi: 10.1016/0092-8674(80)90384-0 pmid: 7357599
20. Bogenhagen DF, Sakonju S, Brown DD. A control region in the center of the 5S RNA gene directs specific initiation of transcription: II. The 3' border of the region. *Cell*. 1980;**19**(1):27-35. doi: 10.1016/0092-8674(80)90385-2 pmid: 7357604
21. Fowlkes DM, Shenk T. Transcriptional control regions of the adenovirus VAI RNA gene. *Cell*. 1980;**22**(2 Pt 2):405-413. doi: 10.1016/0092-8674(80)90351-7 pmid: 7448868
22. Guilfoyle R, Weinmann R. Control region for adenovirus VA RNA transcription. *Proc Natl Acad Sci U S A*. 1981;**78**(6):3378-3382. doi: 10.1073/pnas.78.6.3378 pmid: 6943546
23. Hofstetter H, Kressman A, Birnstiel ML. A split promoter for a eucaryotic tRNA gene. *Cell*. 1981;**24**(2):573-585. doi: 10.1016/0092-8674(81)90348-2 pmid: 7237560
24. Ciliberto G, Castagnoli L, Melton DA, Cortese R. Promoter of a eukaryotic tRNAPro gene is composed of three noncontiguous regions. *Proc Natl Acad Sci U S A*. 1982;**79**(4):1195-1199. doi: 10.1073/pnas.79.4.1195 pmid: 6951168
25. Sprinzl M, Gauss DH. Compilation of tRNA sequences. *Nucleic Acids Res*. 1984;**12** Suppl(suppl):r1-57. doi: 10.1093/nar/12.suppl.r1 pmid: 6728685
26. Murray J. *Mathematical Biology: An Introduction*. Berlin: Springer 2002.
27. Cross MC, Hohenberg PC. Pattern formation outside of equilibrium. *Rev Mod Phys*. 1993;**65**(3):851-1112. doi: 10.1103/RevModPhys.65.851
28. el-Mabrouk N, Lisacek F. Very fast identification of RNA motifs in genomic DNA. Application to tRNA search in the yeast genome. *J Mol Biol*. 1996;**264**(1):46-55. doi: 10.1006/jmbi.1996.0622 pmid: 8950266
29. Lowe TM, Eddy SR. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res*. 1997;**25**(5):955-964. doi: 10.1093/nar/25.5.955 pmid: 9023104
30. Laslett D, Canback B. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Res*. 2004;**32**(1):11-16. doi: 10.1093/nar/gkh152 pmid: 14704338
31. Sugahara J, Fujishima K, Morita K, Tomita M, Kanai A. Disrupted tRNA gene diversity and possible evolutionary scenarios. *J Mol Evol*. 2009;**69**(5):497-504. doi: 10.1007/s00239-009-9294-6 pmid: 19826747
32. Chan PP, Lowe TM. GtRNAdb: a database of transfer RNA genes detected in genomic sequence. *Nucleic Acids Res*. 2009;**37**(Database issue):D93-97. doi: 10.1093/nar/gkn787 pmid: 18984615
33. Das S, Mitra S, Sahoo S, Chakrabarti J. Novel hybrid encodes both continuous and split tRNA genes? *J Biomol Struct Dyn*. 2011;**28**(5):827-831. doi: 10.1080/07391102.2011.10508610 pmid: 21294593
34. Altman S. Ribonuclease P. *Philos Trans R Soc Lond B Biol Sci*. 2011;**366**(1580):2936-2941. doi: 10.1098/rstb.2011.0142 pmid: 21930585
35. Spath B, Canino G, Marchfelder A. tRNase Z: the end is not in sight. *Cell Mol Life Sci*. 2007;**64**(18):2404-2412. doi: 10.1007/s00018-007-7160-5 pmid: 17599240
36. Minagawa A, Ishii R, Takaku H, Yokoyama S, Nashimoto M. The flexible arm of tRNase Z is not essential for pre-tRNA binding but affects cleavage site selection. *J Mol Biol*. 2008;**381**(2):289-299. doi: 10.1016/j.jmb.2008.05.016 pmid: 18602113
37. Suzuki H, Zuo Y, Wang J, Zhang MQ, Malhotra A, Mayeda A. Characterization of RNase R-digested cellular RNA source that consists of lariat and circular RNAs from pre-mRNA splicing. *Nucleic Acids Res*. 2006;**34**(8):e63. doi: 10.1093/nar/gkl151 pmid: 16682442