International Journal of Engineering, Science and Mathematics
Vol. 7 Special Issue 4(2), April 2018,
ISSN: 2320-0294 Impact Factor: 6.765
Journal Homepage: http://www.ijmra.us, Email: editorijmie@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gage as well as in Cabell's Directories of Publishing Opportunities, U.S.A

Unidentified tRNA Structure Identified in the Methanosarcina Mezei Strain Go 1

Uttam Roy Mandal*

Abstract

The accumulation of archaeal tRNA genes was gradually becoming more important in the field of bioinformatics for its essential role in evaluating the origin and evolution of tRNA molecule. Hidden pattern of tRNA structure abound in genome sequence. As these investigations progressed I observed that several existing powerful tools do not indentified the hidden pattern. In this paper, I investigated in Methanosarcina Mazei strain Go1 of archaea genome and report that a rare and so far unidentified tRNAs and tRNAs structure are identified in split nature using my algorithm. We argued from bioinformatics evidence that two halves of clover leaf 1-37 and 38-72, from $5^{/}$ and $3^{/}$ ends respectively allow upto one-non-watson –crick base pairing and the possibility of a single putative intron with the BHB structure at the exon-intron boundary.

Keywords:

Unidentified tRNA;
Hidden pattern;
tRNA structure;
Bulge-helix-bulge (BHB);
Split-tRNA.

Author correspondence:

Uttam Roy Mandal,
Department of Mathematics
Raidighi College, Raidighi, South 24 Parganas, 743383, India
Email: urmandal@gmail.com

1. Introduction

Transfer RNA (tRNA) a central genetic element in the decoding of genome information for all of Earth's life forms [1]. Although tRNA is one of the most conserved RNA molecules in structure and functions.The tRNA gene is architecturally diverse.

In archaea, there have been four types of tRNA genes identified to date: intron-less, intron –containing, split tRNA and permuted tRNA [2]-[4]. Coding sequence for these tRNA are either 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 [5]-[6]. 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 [6]-[8]. Structural motif and location of tRNA introns in archaea have been affected by change in the recognition and activity of the different unit compositions in the tRNA splicing endonucleases [9]. 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 [10]-[12].

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 1a 30-base pair (bp) region seems to contain all the information necessary for correct initiation of transcription [13]-[14].

There are many hidden patterns around in the genome sequences. Many sophisticated non linear algorithms [15] exits for pattern formation and recognition [16]-[17]. For instance, there are several computational approaches to detect tRNA genes from a genome [18]. To identify these on the sequences,

* Department of Mathematics, Raidighi College, Raidighi, South 24 Parganas, 743383, India

there are number of algorithms extensively used in this field. Notable amongst these are tRNAScan-SE [19] and ARAGORN [20]. 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. I developing my algorithm to circumvent some of the difficulties in searching missing tRNA sequence by existing programs, my algorithm searched for two halves of cloverleaf, 1-37 and 38-72, from 5′ and 3′ end respectively, allowed upto one non-watson-crick base pairing and the possibility of a single putative intron with the standard BHB structure at the exon boundary.

The Archaeon Methanosarcina Mazei has great ecological importance as they are the only organism fermenting acetate, methylamines and methanol to methane, carbon dioxide and ammonia. It can be found in a multitude of environments including the rumen in cow, sheep, goats, deer and the large intestine in humans. The Methanosarcina Mazei strain Go1(MMSG) genome is 4096345bp long and GC content 41.48% . In Methanosarcina Mazei strain Go1(MMSG) of archaea genome I investigate from my results that a rear and so far unidentified tRNAs THR (AGU),ALA (AGC), SEC (UCA), LEU (AAG) ,PHE (AAA) and HIS (AUG) are identified , remarkably I have seen GLU (UUC) ,ARG (UCG) and THR (UGU), HIS (GUG) tRNAs are overlapping, more so a specific region of the nucleotide sequence (NTS) in this genome, LEU (GAG) tRNAs are looks like a twine in nature.

2. Research Method

The entire genome sequences were obtained from National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/).In the present communication, I focused my study on Methanosarcina Mazei strain Go1 MMSG) (GenBank accession no NC_00390).

I developed a computational approach which was aimed at searching for any type of tRNA genes and was especially focused on intron containing tRNA (and split tRNA) not identified by tRNAscan-SE and ARAGORN. Here we have developed a procedure for generation 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 stands 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, I extracted many secondary structures of tRNAs which were not indentified so far in the Methanosarcina Mazei strain Go1 (MMSG) genome.

3. Results and Analysis

3.1. tRNA THR (AGU), ALA (AGC), SEC (UCA) , PHE (AAA),LEU (AAG) and HIS (AUG) are identified genes of MMSG.

I investigated the genome of the archaea MMSG. In this archaea, the rare threonine, alanine and selenocysteine tRNAs gene with anticododon AGU, AGC and UCA respectively in Figure 1 were not annalist earlier. I indentified its between (192111---192150, 1422566---1422613), (3343146---3343183, 3773243---3773287) and (3451162---3451198, 1619345---1619391) on the genome and analyse in silico the secondary structure at the exon –intron boundary –the bulge-helix-bulge (BHB) . It is the conformation structure most easily recognized and processed by archaeal splicing mechanism. I elucidate this identification now. Note that intron splicing occurs during tRNA maturation following the transcription of the gene. Splicing in archaea is enzyme catalyzed, initiated by an endonuclease that excises the intron to yield half-molecules with ends containing a 2′-3′ cyclic phosphate and a 5′ -OH. The entire sequence of this tRNAs gene and silico secondary structure at exon –intron boundary –the BHB is shown in Figure 1.

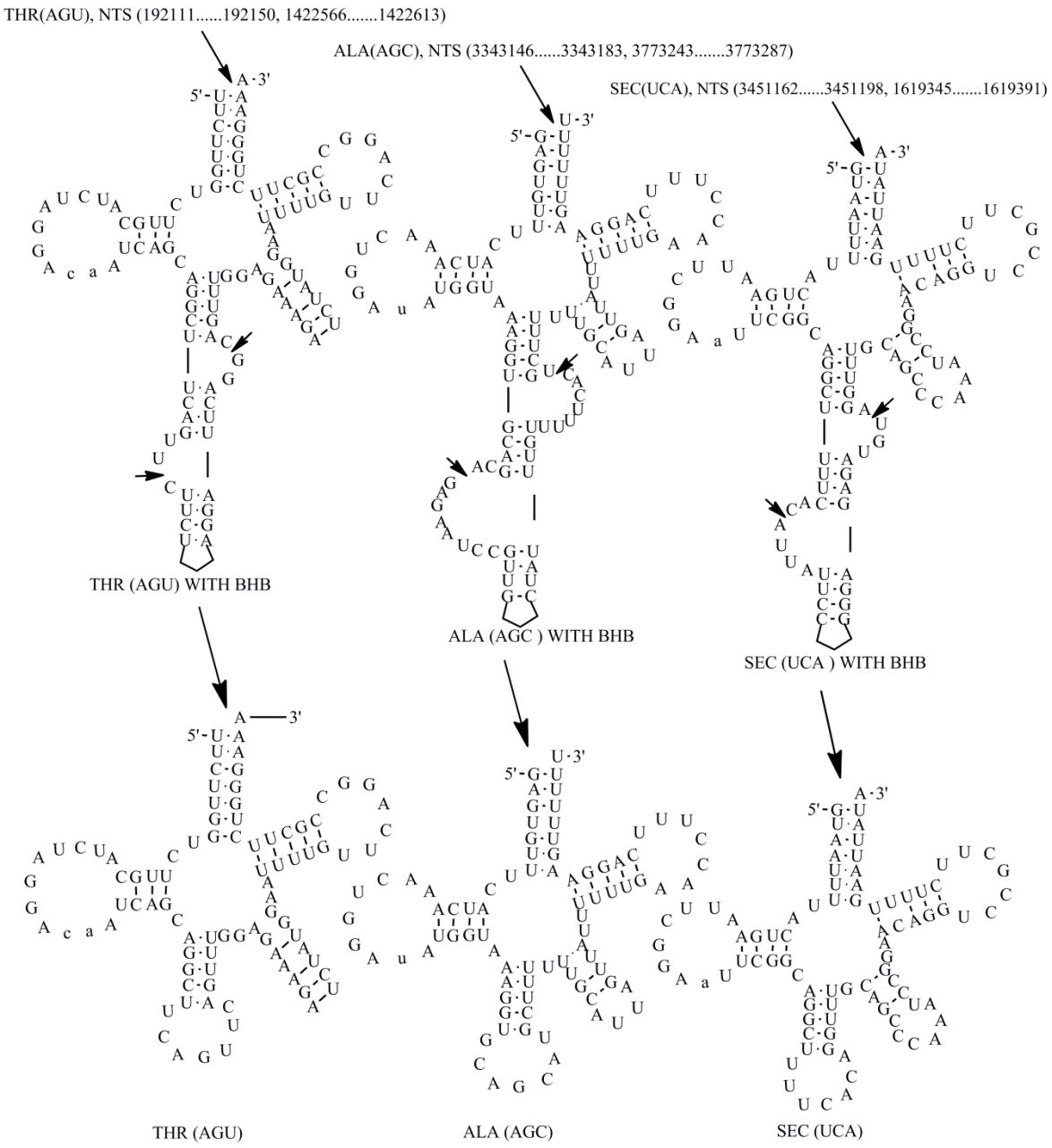

Figure-1

Figure 1.Structure of tRNA genes THR (AGU), ALA (AGC) and SEC (UCA). The schematic diagrams of the typical tRNAs that split at 37 or 38 position along with the structures of the BHB motifs. Splice sites are indicated by arrows.

Also leucine, phenylalanine and histidine tRNAs gene with anticodon AAA, AAG and AUG respectively indetified between (3363264---3363300, 278664---2786711), (733643---733682, 3773242---3773287) and complement (2951595---2951637, 336766---336801) on the genome and silico secondary structure at exon – intron boundary –the BHB is shown in Figure 2.

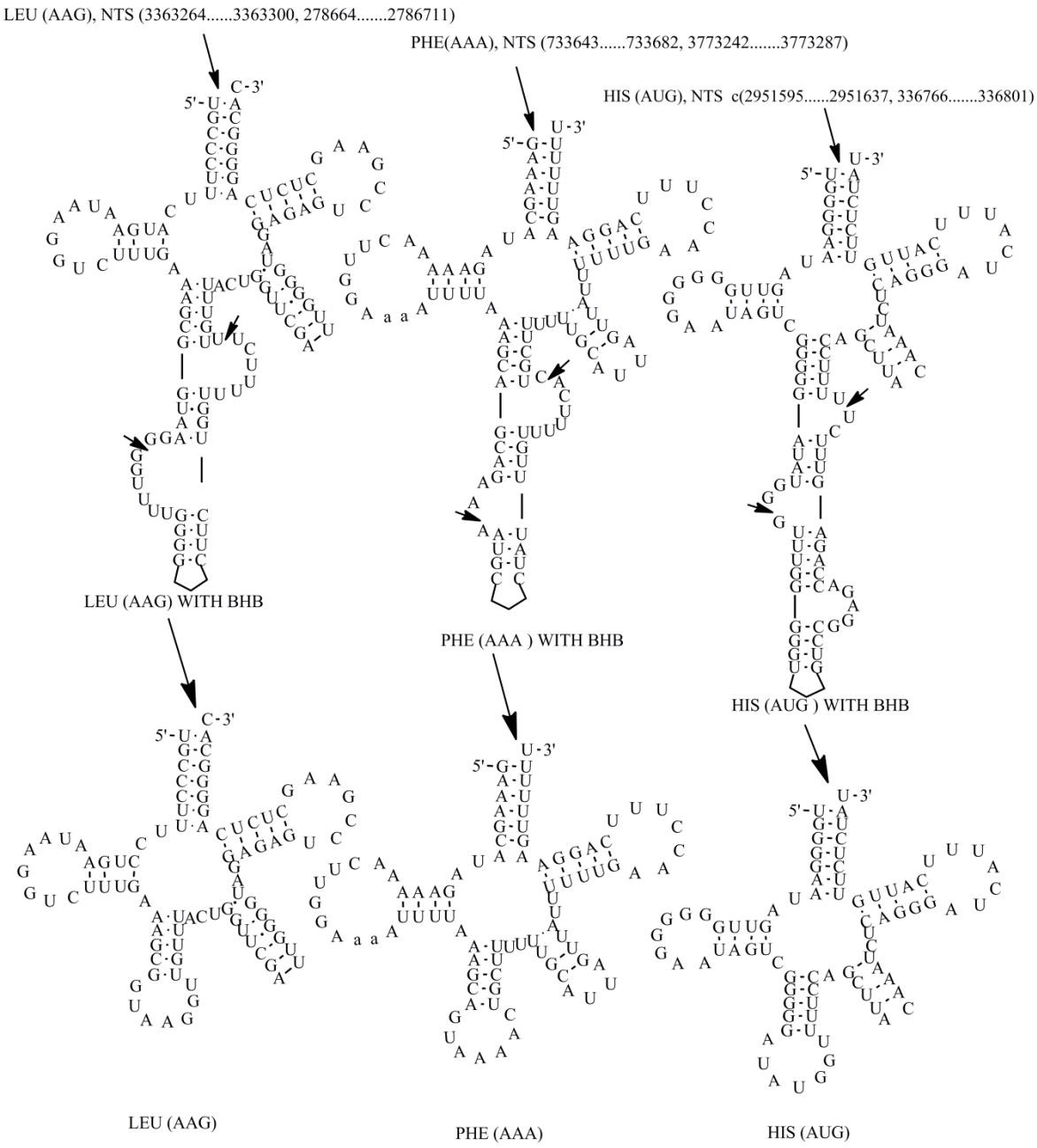

Figure-2

Figure 2.Structure of tRNA genes LEU (AAG), PHE (AAA) and HIS (AUG).
The schematic diagrams of the typical tRNAs that split at 37 or 38 position along with the structures of the BHB motifs. Splice sites are indicated by arrows.

Interestingly , I analyse and observe the tRNA gene summary from the tRNA Database of MMSG the unidentified shaded tRNAs mentioned in the Table 1 are indentified by own investigation. From the split nature of the tRNA we can find the missing tRNAs in the archaea genome.

Table 1. tRNA Gene Summary

Isotype	tRNA count by anticodon						Total/Database	Total/Result
THR	AGT 1	GGT 1	CGT 1	TGT 1			3	4
ALA	AGC 1	GGC 1	CGC 1	TGC 1			3	4
SEC	TCA 1						0	1
PHE	AAA 1	GAA 2					2	3
LEU	AAG 1	GAG 2	CAG 1	TAG 1	CAA 1	TAA 1	6	7
HIS	ATG 1	GTG 1					1	2

3.2. tRNAs THR (UGU)/ HIS (GUG) and GLU (UUC)/ ARG (UCG) Overlapping genes

From my result the tRNA THR (UGU) and GLU (UUC) are located in the genome sequence between complement (c) (302540---302572, 302575---302611) and complent (c) (3631184---3631216, 3631239-3631276) respectively. Also from the bioinformatics point of view its has the conserved bases and base-pairs of other archaeal threonine and glycine tRNAs and analyse silico the secondary structure at the exon –intron boundary- the bulge-helix-bulge (BHB) or relaxed BHB. The entire sequence and BHB of these tRNAs gene is shown in the Figure 3.

From the existing tRNAs database the tRNA HIS (GUG) and ARG (UCG) were identified in between c(302539---302611) and c(3631183---3631217, 3631238---3631276). Remarkably, the tRNA THR (UGU) / HIS (GUG) and GLU (UUC)/ ARG (UCG) are located in the same region in the genome sequence i.c the tRNAs are overlapping eacg other. So, from the same location in the genome tRNA carries the different amino acids. In the different environments the archaea species can produce amino acid from the hidden pattern gene in the genome.

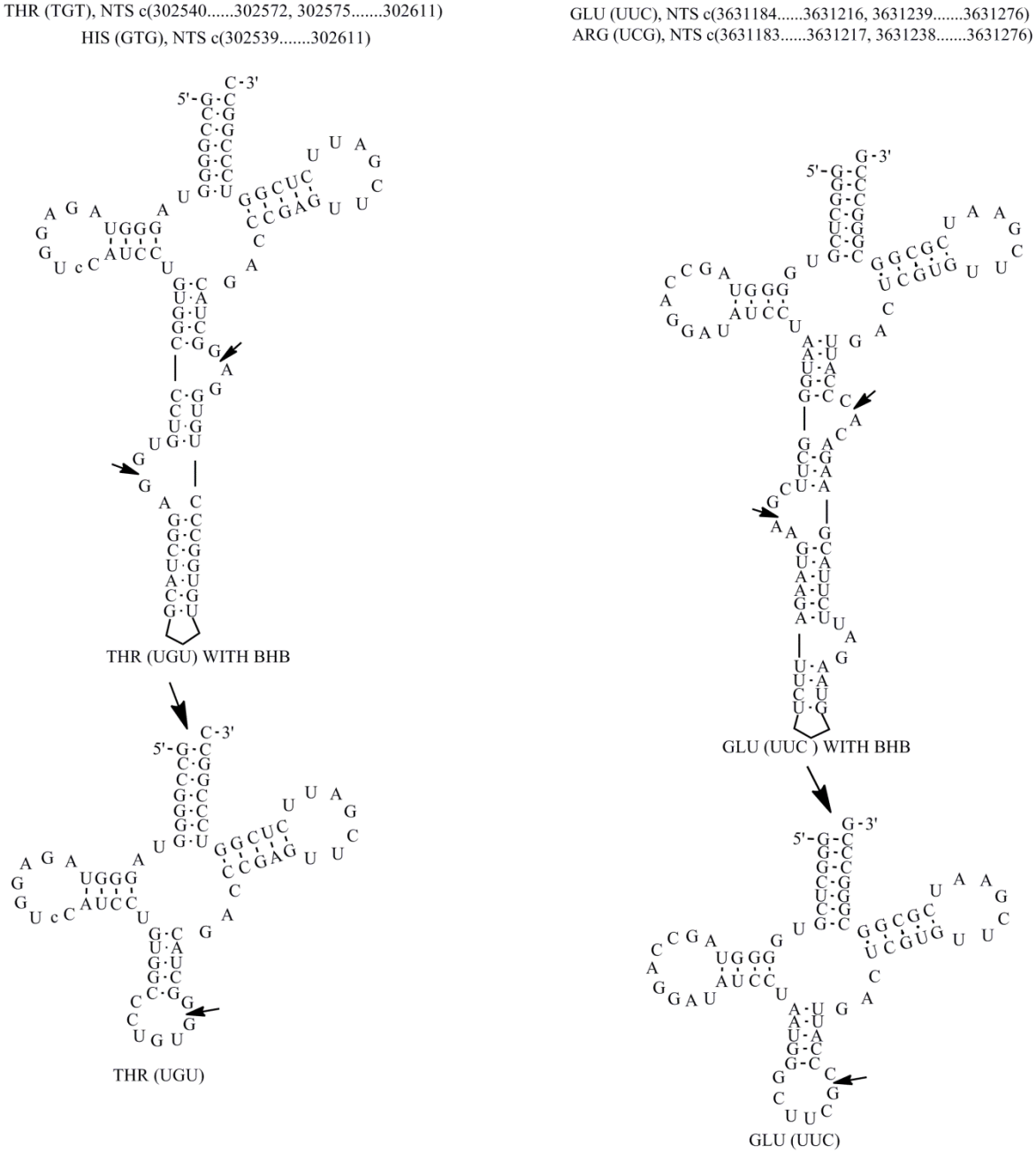

Figure-3

Figure 3. tRNA THR (UGU)/ HIS (GUG) and GLU (UUC)/ ARG (UCG) are overlapping tRNA

The schematic diagrams of the typical tRNAs that split at 37 or 38 position along with the structures of the BHB motifs. Splice sites are indicated by arrows.

3.3. Structure of tRNA leucine gene of MMSG

In MMSG all the tRNA genes could not be located using tRNAscan-SE and other existing software. Some of the missing pattern of the gene may be active in the extreme conditions for survival. Therfore through out the results I investigate such hidden patterns of the tRNA genes which are not indentified so far.

The tRNAs LEU (GAG) were indentified in between (2560527---2560614) and (2560706---2560790) using tRNAscan-SE shown in the Figure 4. I identified the tRNAs LEU (GAG) between (2560527---2560794) on the genome. The entire method for findind tRNA leucine gene is shown in the Figure 4. My algorithm located LEU(GAG) between (2560527---2560566, 2560746---2560794) and (2560706---2560745, 2560567---2560614) and splicing mechanism i.c silico BHB for conformational structure are satisfied. The details of the tRNAs sequence are shown in the Figure 5.

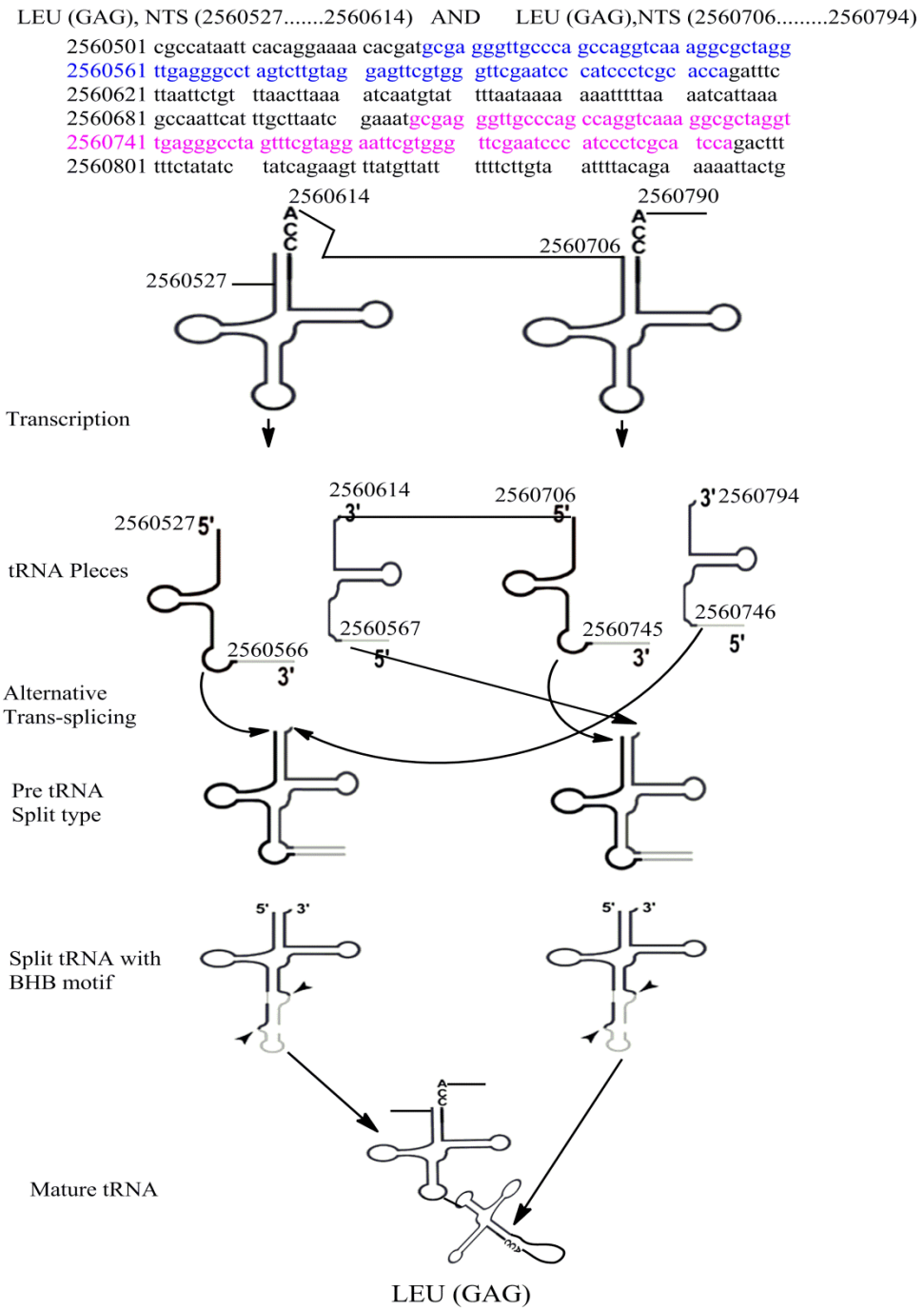

Figure-4

Figure 4. Formation of two tRNAs LEU (GAG)

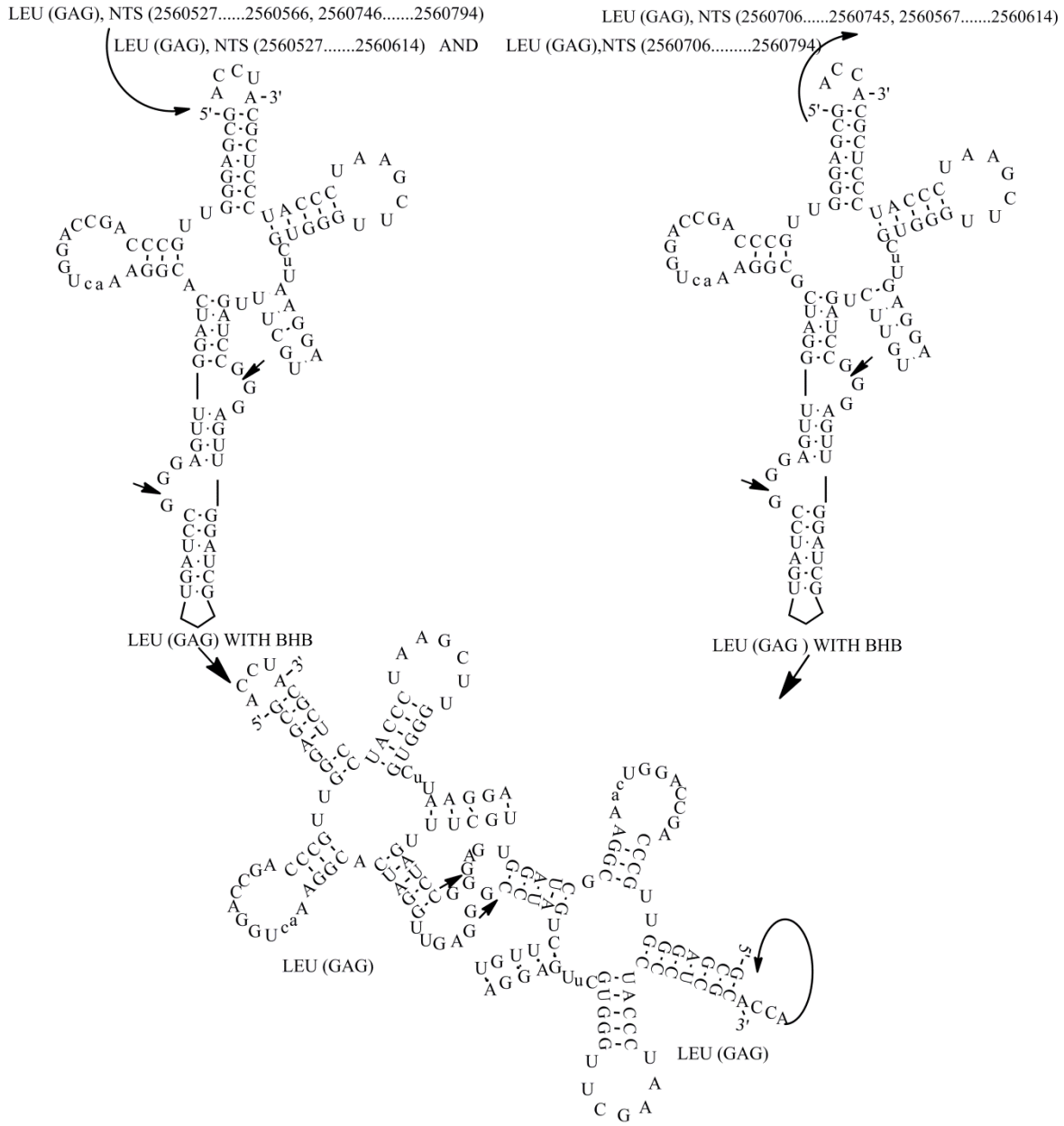

Figure-5

Figure 5. Structure of Split LEU (GAG) tRNA genes
The schematic diagrams of the typical tRNAs that split at 37 or 38 position along with the structures of the BHB motifs. Splice sites are indicated by arrows.

4. Conclusion

In archaea, I studied the availabity of all tRNAs over the entire set of sequenced genomes. Extending my investigation to the newly sequenced i.c the hidden pattern tRNA genes and report that a rare and so far unidentified tRNAs, structural variation of the tRNA in a specific region.Interestingly, the different types of overlapping tRNA has been identified in the different archaea genome. In some of the primary transcripts of mitochondrial tRNA of animals, tRNA genes are known to overlap by one several baser. We belive that the hidden pattern of the tRNA gene carry amino acid for building the protine in extreme environments.So that they are adapted to them and in fact, have a hard time acclimating to less extreme conditions.

I am investigation features to find out tRNA gene which are unidentified so far in the different archaeal species.

References

[1] Maruyama, S., Sugahara J., Kanai A. and Nozaki H., "Permuted tRNA Genes in the Nuclear and Nucleomorph Genomes of Photosynthetic Eukaryotes," Mol Biol Evol, vol. 27(5), pp. 1070-1076, 2010.

[2] Marck, C. and Grosjean H.,"tRNomics: analysis of tRNA genes from 50 genome of Eukarya, Archaea, and Bacteria relveals anticodon-sparing strategies and domain-specific features," Rna, vol. 8, pp.1189-1232, 2002.

[3] Sugahara, J., Yachie N., Sekine Y., Soma A., Matsui M.,et al. "SPLITS: a new program for predicting split and intron-containing tRNA genes at the genome level," In Silico Biol, vol. 6, pp. 411-418, 2006.

[4] 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, vol.13, pp. 671-681, 2007.

[5] Randau, L., Munch R., Hohn M.J., Jahn D., Soll D., "Nanoarchaeum equitans creates functional tRNAs from separate genes for their 5′- and 3′-halves", Nature, vol. 433(7025), pp. 537–541, 2005.

[6] Fujishimaa, K., Sugaharaa J., Kikuta K., 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 USA, vol. 106(8) ,pp. 268—2687, 2009.

[7] Randau, L., et al. "The heteromeric Nanoarchaeum equitans splicing endonuclease cleaves noncanonical bulge-helix-bulge motifs of joined tRNA halves", Proc Natl Acad Sci USA, vol. 102(50), pp.17934–17939, 2005.

[8] Tocchini-Valentini, G.D., Fruscoloni P., "Tocchini-Valentini GP. Coevolution of tRNA intron motifs and tRNA endonuclease architecture in Archaea", Proc Natl Acad Sci USA, vol. 102(43), pp. 15418–15422, 2005.

[9] Fujishima, K., Sugahara J., Miller C.S., Baker B.J., Di Giulio M.,"Takesue K, et al. A novel three-unit Trna splicing endonuclese found in ultrasmall Archaea possesses broad substrate specificity", Nucleic Acid Res, vol. 39, pp. 9695-9704, 2011.

[10] Weiner, A.M., 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", Bio Bull, vol. 196, pp. 327, 1999.

[11] Widmann, J., Giulio M.D., Yarus M., Knight R., "tRNA creation by hairpin duplication", J Mol Evol, vol. 61, pp. 524-530, 2005.

[12] Sun, F.J., Caetano-Anolles G., "The origin and evolution of tRNA inferred from phylogenetic analysis of structure", J Mol Evol, vol. 66, pp. 21-35, 2007.

[13] Sakonju, S., Bogenhagen D.F., Brown D.D., "A control region in the center of the 5S RNA gene directs specific initiation of transcription: I. The 5' border of the region", Cell, vol. 19(1), pp. 13-25, 1980.

[14] Bogenhagen, D.F., Sakonju S., Brown D.D., "A control region in the center of the 5S RNA gene directs specific initiation of transcription: II. The 3' border of the region", Cell, vol. 19(1), pp. 27-35, 1980.

[15] Murray, J.D., "An Introduction (3rd) (Berlin:Springer Verlag) ", Mathematical Biology , 2002.

[16] Cross, M.C. and Hohenberg., "Pattern formation of outside equilibrium", Rev. Mod. Phys, vol. 65, pp. 851, 1993.

[17] el-Mabrouk, N. and Lisacek F., "Very fast identification of RNA motifs in genomic DNA: application to tRNA search in the yeast genome", J. Mol. Biol, vol. 264, pp. 46–55, 1996.

[18] Lowe, T.M. and Eddy S.R., "tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence", Nucleic Acids Res., vol.25, pp. 955–964, 1997.

[19] Lowe, T.M. and Eddy S.R., "tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence", Nucleic Acids Res., vol.25, pp. 955–964, 1997.

[20] Laslett, D. and Canback, B., "ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences", Nucleic Acids Res., vol. 32, pp. 11–16, 2004.