International Journal of Computational Biology
ISSN: 2229-6700 & E-ISSN:2229-6719, Volume 5, Issue 1, 2014, pp.-068-075.
Available online at http://www.bioinfopublication.org/jouarchive.php?opt=&jouid=BPJ0000220

MUTATION PRESSURE AND NATURAL SELECTION AFFECT PATTERN OF CODON USAGE BIAS IN MITOCHONDRIAL ATPASE 6 AND ATPASE 8 GENES IN DIFFERENT MAMMALIAN SPECIES

UDDIN A. AND CHAKRABORTY S.*

Department of Biotechnology, Assam University, Silchar - 788 011, Assam, India.
*Corresponding Author: Email- supriyoch_2008@rediffmail.com

Received: May 03, 2014; Accepted: November 06, 2014

Abstract- *Background*: The non uniform usage of synonymous codons during translation makes codon usage bias *i.e* some codons are used more frequently than others. However the synonymous codon usage pattern differs among the genes of an organism and also among different organisms. Mitochondrial genome is small, highly conserved and maternally inherited with high mutation rate. The study of codon usage bias pattern is useful for better understanding the process of evolution in terms of molecular biology.

Methods and Result: In this study we analyzed codon usage pattern of genes encoding mitochondrial ATPase 6 and ATPase 8 in different species of mammals. From our study it was evident that leucine is the most frequent amino acid in both genes in the mammalian species. ENC value ranged from 57 to 60 with a mean of 59.25 in ATPase 6. In ATPase 8, ENC value ranged from 42 to 60 with a mean of 54.1. In ATPase 6, a significant positive correlation was observed between overall A and A3% (r =0.9065, p<0.001), C and C3% (r = 0.9238, p< 0.001), GC and GC3% (r = 0.9244, p<0.001). ENC showed a highly significant positive correlation with overall GC content (r =0.9479, p<0.001) and GC3 (r = 0.9387, p<0.001) in ATPase 6. But in ATPase 8, significant positive correlations were observed between A% and A3% (r=0.6502, p<0.05), C% and C3% (r =0.9804, p< 0.001), GC% and GC3% (r =0.9732, p<0.001). In addition, a highly positive correlation was found between ENC and GC (r = 0.9333 p<0.001), ENC and GC3 (r = 0.9811, p<0.001) in ATPase 8.

Conclusion: In our findings, we observed the codon usage bias in ATPase 6 and ATPase 8 in mammalian species is not remarkable. However in both the genes, the frequent codons favour A or C at the 3rd codon position. Our results supported that mutation pressure plays a major role in the codon usage pattern in mitochondrial ATPase 6 and ATPase 8 in addition to natural selection in some mammalian species.

Keywords- Codon, Codon usage, Synonymous codon, Mitochondrial genome, Mammals

Citation: Uddin A. and Chakraborty S. (2014) Mutation Pressure and Natural Selection Affect Pattern of Codon Usage Bias in Mitochondrial ATPase 6 and ATPase 8 Genes in Different Mammalian Species. International Journal of Computational Biology, ISSN: 2229-6700 & E-ISSN:2229-6719, Volume 5, Issue 1, pp.-068-075.

Introduction

The genetic code is degenerate. Except methionine and tryptophan all other amino acids are encoded by more than one codon which usually ranges from 2 to 6. The frequency of synonymous codons coding same amino acid are not same in protein coding genes [1,2]. Bias in codon usage is found in highly expressed genes [3]. Codon usage bias exists in diverse organisms ranging from unicellular organism to multicellular organism [4-6]. Each organism has unique pattern of codon usage and it differs from species to species [7]. Several factors which influence the codon bias are gene expression [8-11], gene length [10], base compositional mutational bias [12] and natural selection [13-17]. Due to mutation pressure, the codon usage varies with extremely high A+T or G+C contents in some prokaryotes [18] and mammals [19]. However, in Drosophila and some plants codon usage bias is due to translational selection [20-23].

For understanding the evolution of living organism at molecular level and its environmental adaptation it is imperative to study codon usage [24]. Several studies have been done on codon usage bias in nuclear genomes, but very few literature is available on codon usage bias in mitochondrial genomes. Mitochondrial genomes are small in size and evolutionarily conserved. The standard genetic code and the mitochondrial genetic codes are not same [25]. The ATPase is an important enzyme in mitochondria and it is encoded by two genes namely ATPase 6 and ATPase 8 gene. Analysis of codon usage bias pattern in ATPase gene would help us understand the mechanism of variation in codon usage, its distribution and the evolutionary forces involved in shaping the codon usage bias.

Materials and Methodology

Sequences Data

The coding sequences (cds) of ATPase 6 and ATPase 8 gene were retrieved from National Center for Biotechnology Information, USA (http://www.ncbi.nlm.nih.gov/) for several mammalian species. The detailed information about different species is given in [Table-1].

Table 1- Accession no, gene length, ENC, CAI and Overall GC% in some mammalian species for ATPase 6 and ATPase 8 genes

Species	Accession No	Gene length	ENC	CAI	Overall GC%
ATPase 6					
H.sapiens	V00662	681	60	0.4926	44.2
G.gorilla	D38114	681	60	0.4428	44.6
P.paniscus	D38113	681	59	0.1723	43.3
P.troglogytes	D38116	681	59	0.3772	43.8
P.pygmaeus	D38115	681	60	0.4626	45.5
H.lar	X99256	681	60	0.142	45.5
E.caballus	X79547	681	59	0.242	43.3
E.asinus	X97337	681	57	0.3563	41
ATPase 8					
G.gorilla	D38114	207	58	0.3719	38.6
P.paniscus	D38113	207	57	0.2602	39.6
P.troglodyle	D38116	207	59	0.3955	40.1
P.pygmaeus	D38115	207	60	0.2378	46.4
H.lar	X99256	207	60	0.3584	42
E.caballus	X79547	204	54	0.1771	35.8
E.asinus	X97337	204	49	0.2356	33.8
R.unicornis	X97336	204	53	0.3737	35.8
R.norvegicus	X14848	204	49	0.2206	35.3
A.jamaicensis	AF061340	204	42	0.2207	27.5

Compositional Properties

Base compositions (A, T, G & C) of ATPase 6 and ATPase 8 genes, their nucleotide composition at the codon 3rd position, overall GC contents along with GC1, GC2, and GC3 in percentage are calculated for each coding sequence (cds) using a perl programme developed by SC.

Relative Synonymous Codon Usage

The relative synonymous codon usage (RSCU) value is calculated according to the formula of Sharp *et.al* [26]. The codon is said to be frequently used than expected, if its RSCU value is >1 and if RSCU<1, codon is less frequently used than expected. If RSCU= 1 it means the codon is used randomly and equally [27]. If the RSCU value of a codon is >1.6, the codon is over-represented [28] and if the RSCU value is <0.6, the codon is under represented.

Effective Number of Codon (ENC)

ENC is used to measure the codon usage bias and depends upon the nucleotide composition of the gene. Its value ranges from 20 to 61. Higher value means low codon bias [29]. ENC value of a gene lower than 35 is considered as significant codon usage bias [30].

Codon Adaptation Index (CAI)

Codon adaptation index is the most widely used measure for gene expression. Its value ranges from 0-1. High CAI value indicates that gene expression is high. It was described by Sharp and Li [31]. All the bioinformatics analysis was done using a perl programme, developed by Supriyo Chakraborty.

Statistical Analysis

Correlation analysis was used to assess the relationship between nucleotide composition and each base in 3rd codon position. Correlation of ENC with CAI, ENC with GC, ENC with GC3, and GC12 with GC3 were also estimated. All the statistical analysis was done using SPSS software.

Result

Compositional Properties

The value of overall nucleotide composition and its composition at codon's 3rd position for ATPase 6 and ATPase 8 are shown in [Table-2]. The distributions of overall A, T, G and C% and A3, T3, G3 and C3 in mitochondrial ATPase 6 and ATPase 8 in some mammalian species are presented in [Fig-1] and [Fig-2]. From [Fig-1] & [Fig-2] it is evident that A and C occurred more frequently than G and T in both genes. We have found that C3 occurred more frequently but G3 occurred less frequently in ATPase 6. But in ATPase 8, A3 occurred more frequently and G3 occurred less frequently. If the mutation pressure alone is responsible for codon usage, then the% of A and T should have been equal to G and C at the wobble position [32] but in case of ATPase 6 and ATPase 8 the nucleotide composition variation indicates that in addition to mutation pressure, natural selection might have influenced the pattern of codon usage bias in the mammals.

Table 2- Overall nucleotide composition and its composition at codon's 3rd position for ATPase 6 and ATPase 8 gene

Species	% of A	A3%	% of T	T3%	% of G	G3%	% of C	C3%
ATPase 6								
H.sapiens	30.24	35.68	25.55	18.06	10.42	6.16	33.77	40.08
G.gorilla	30.1	37	25.25	17.62	11.16	6.6	33.48	38.76
P.paniscus	29.95	37.44	26.72	21.58	10.71	5.28	32.59	35.68
P.troglogytes	29.95	37.44	26.28	20.7	10.86	5.28	32.89	36.56
P.pygmaeus	29.36	34.8	25.11	17.18	10.27	5.72	35.24	42.29
H.lar	30.39	37.44	24.08	17.18	10.71	6.16	34.8	39.2
E.caballus	30.83	40.08	25.84	17.18	11.89	6.16	31.42	36.56
E.asinus	31.86	41.4	27.16	21.14	10.86	4.84	30.1	32.59
ATPase 8								
G.gorilla	37.19	43.47	24.15	17.39	7.24	10.14	31.4	28.98
P.paniscus	38.16	44.92	22.22	17.39	6.28	8.69	33.33	28.98
P.troglodyle	38.16	44.92	21.73	14.49	6.76	10.14	33.33	30.43
P.pygmaeus	34.78	42.02	18.84	10.14	6.76	10.14	39.61	37.68
H.lar	35.74	39.13	22.22	15.94	7.24	11.59	34.78	33.33
E.caballus	38.23	47.05	25.98	20.58	9.8	14.7	25.98	17.64
E.asinus	38.23	48.52	27.94	26.47	8.82	10.29	25	14.7
R.unicornis	38.72	48.52	25.49	20.58	6.86	11.76	28.92	19.11
R.norvegicus	36.76	50	27.94	23.52	6.37	5.88	28.92	20.58
A.jamaicensis	38.72	47.05	33.82	36.76	5.88	5.88	21.56	10.29

The overall GC contents and the GC contents at codon's 3rd position in different species of mammals varied from 41-45.5 with a mean of 43.9 and 37.4 -46.3 with a mean of 43.51 respectively. The distributions of GC, GC1, GC2 and GC3% are shown in [Fig-3] for ATPase 6 and in [Fig-4] for ATPase 8. Overall GC content varies from 27.5% to 46.4% with a mean of 37.49±5.12 whereas the GC content at 3rd codon position varies from 16.2 to 47.8 with a mean of 34.11±9.76.

In ATPase 6, the ENC of different species of mammals ranges from 57 to 60 with a mean of 59.25. Higher ENC means lower bias. The distribution of ENC across species is shown in [Fig-3]. In ATPase 8, the ENC values were much lower and ranged from 42 to 60 with a mean of 54.1. Distributions of ENC are shown in [Fig-4] for ATPase 8. This indicated that codon usage bias is not remarkable for ATPase 6 and ATPase 8 in these species.

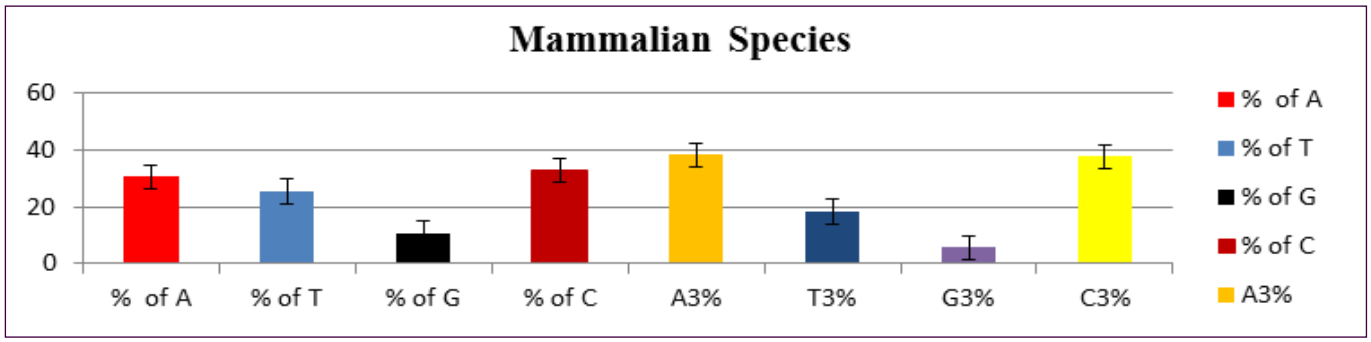

Fig. 1- Overall nucleotide composition and its 3rd codon position in different species in mammals for ATPase 6 gene.

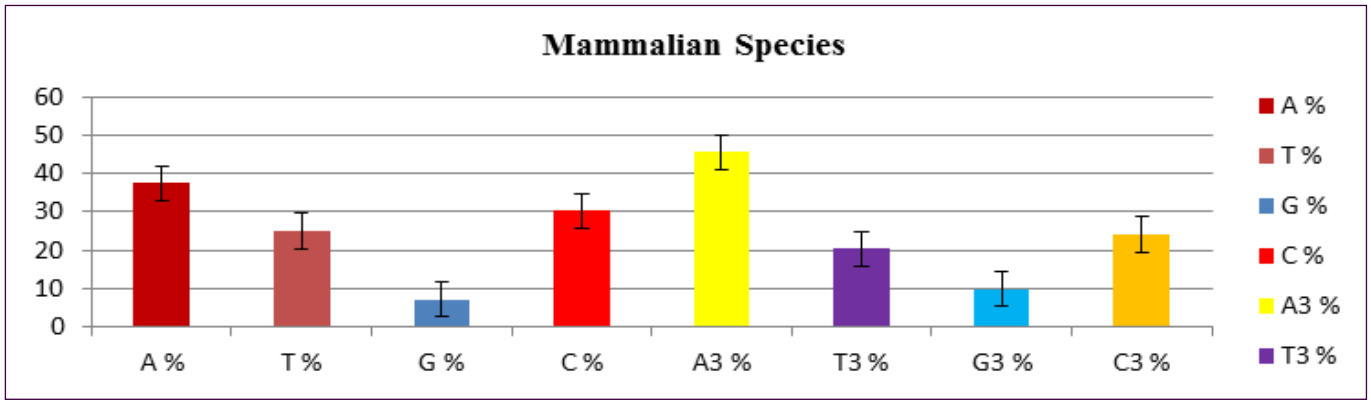

Fig. 2- Distribution of overall nucleotide composition and its 3rd codon position for ATPase 8 gene

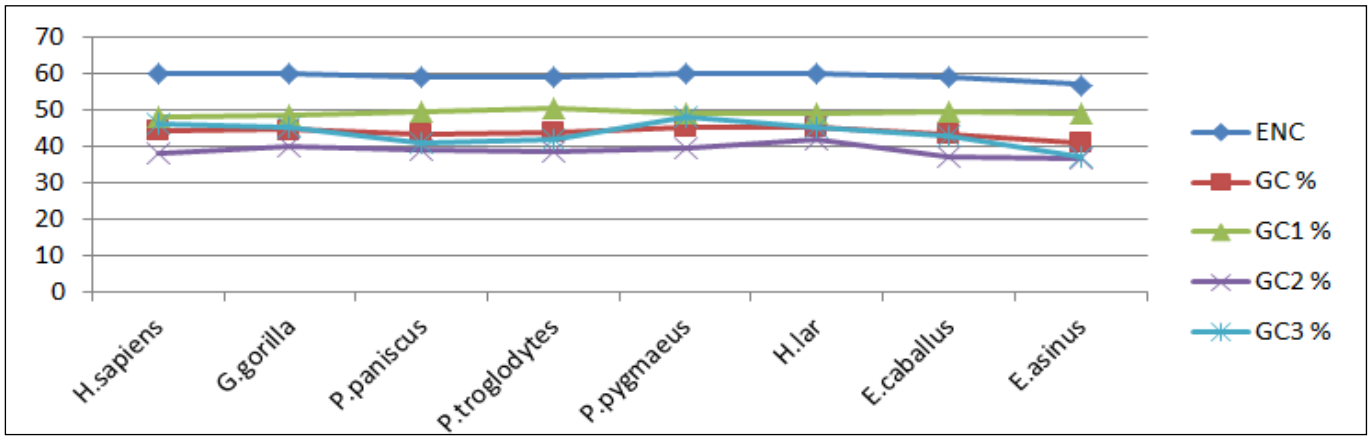

Fig. 3- Distribution of ENC, GC contents, GC1, GC2, GC3 contents for ATPase 6 gene in different species of Mammals

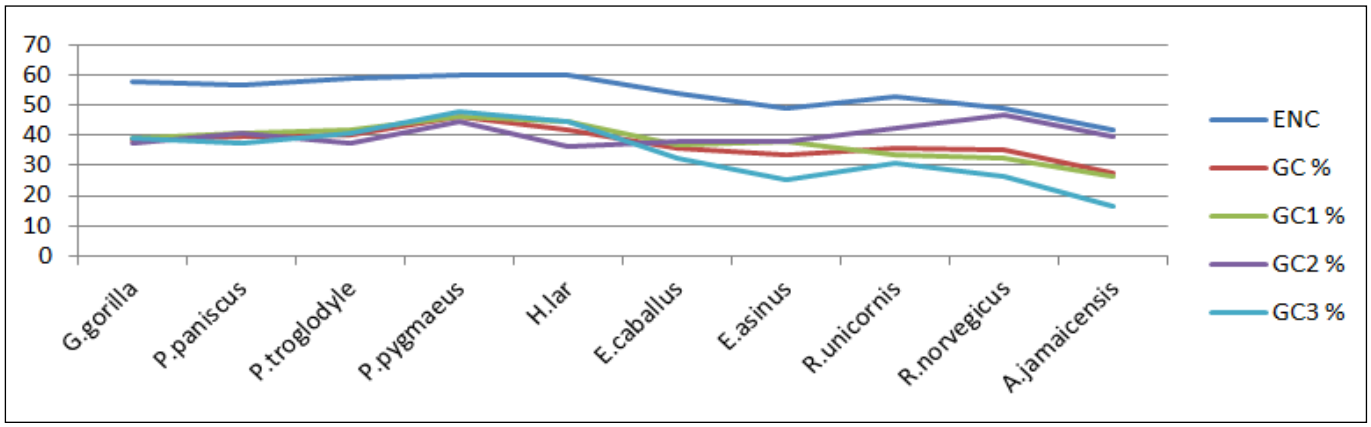

Fig. 4- Distribution of ENC, overall GC%, GC1%, GC2% and GC3% for ATPase 8 gene

To determine the differences between nucleotide composition and codon selection in each species correlation was estimated between ENC and CAI. CAI is a directional measure of codon usage bias similar to relative codon bias score. The distribution of CAI in different mammalian species is shown in [Fig-5] and [Fig-6] for ATPase 6 and ATPase 8 respectively.

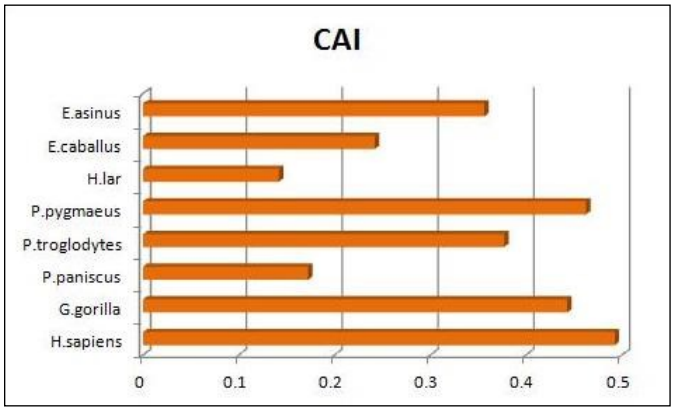

Fig. 5- Distribution of CAI for ATPase 6 gene in different species of mammals

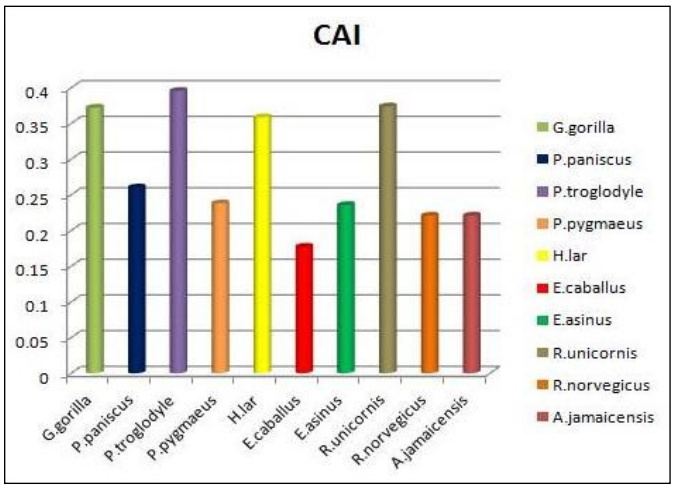

Fig. 6- Distribution of CAI for ATPase 8 gene in different mammalian species

The correlation between ENC and CAI provides a good qualitative assessment between the nucleotide composition and codon bias imposed by selection. For both the genes, similar trend of positive correlation (Pearson correlation) was found between two parameters in mammalian species as in [Table-3].

Table 3- Correlation coefficient between codon usage bias parameters for ATPase 6 and ATPase 8 genes

Correlation between	Correlation Coefficient
ATPase 6	
ENC and CAI	0.1591
ENC and GC	0.9479***
ENC and GC3	0.9387***
GC12 with GC3	0.3611
ATPase 8	
ENC and CAI	0.5202
ENC and GC	0.9333***
ENC and GC3	0.9811***
GC12 and GC3	0.8270**

If $p<0.05$*, $p<0.01$**, $p<0.001$***

Codon Usage in Mammalian Species

For ATPase 6, the frequently used codons as revealed from RSCU value were TCA, TCC, AGC, TTC, CTA, CTC, TAC, CCC, CAC, CAA, CGA, CGC, ATC, ATT, ACA, ACC, AAC, GAC, GTA, GCC, AAA, GAA, GGA & GGC. Codons favoring the base A or C at the 3rd codon position are frequently used. Among these, the codons TCA, AGC, CTA, TAC, CCC, CAC, GAA, CGA, GTA, GCC, AAA, GAA & GGC were over represented. Less frequently used codons were TCT, TCG, TTT, TTG, TAT, TGT, TGC,AGT, AGA, AGG, ACT, ATA, ACG, AAT, AAG, GAT, GTG, GCG, GGT, GAG, GTT, GCA,GTC, GCT, GGC, GGT, GGG, CTG, CTT, CCA, CCG, CAT, CAG & CGG. Amongst these the under represented codons were TCG, TGT, TGC, TTG, TAT, AGT, ACG, AAT, AAG,AGA, AGG, GAT, GTG, GCG, GAG, GTT, GGT, CTG, CCG, CAT, CAG, CGG & GGG. The comparisons of RSCU values of 59 codons in different species are shown in [Fig-7]. The distribution of over and under represented codons for ATPase 6 gene are shown in [Fig-8] & [Fig-9] respectively. For ATPase 8 gene, the codons which occurred more frequently as revealed from RSCU values were TCA, TCC, TCG, TTT, CTA, CTT, TAT, CCA, CCC, CAC, CAA, ATA, ATC, ACA, ACC, AAT, GTA, GCC, AAA & GAA. This indicates that most of the frequent codons end with A or C in its 3rd codon position. Amongst these, the over represented codons were TCA, CTA, CCC, GCC, AAA & GAA. Less frequently used codons were TCT,TTC, TTA, TTG, CTC, CTG, TAC, TGT, TGC, CCG, CCT, CAT, CAG, CGA, CGC, CGG, CGT, AGC,AGT, AGA, AGG, ATT, ACG, ACT, AAC, GAT, GAC, GTC, GTG,GTT, GCA,GCG, GCT, AAG, GAG, GGA, GGC, GGG & GGT. Among the less frequently used codons, the under represented ones were TCT, TTG, CTG, TGT, TGC, CCG, CAG, CGG, CGT, AGA, AGG, ACT, GAC, GTG, GTT, GCA, GCG, GCT, AAG, GAG, GGA, GGC, GGG & GGT.

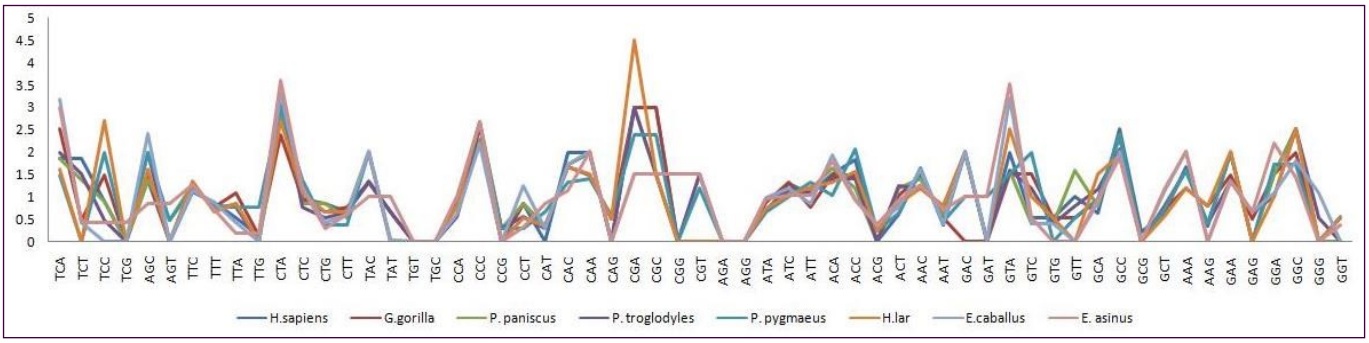

Fig. 7- Comparison of RSCU value for 59 codons for ATPase 6 gene in different species of Mammals

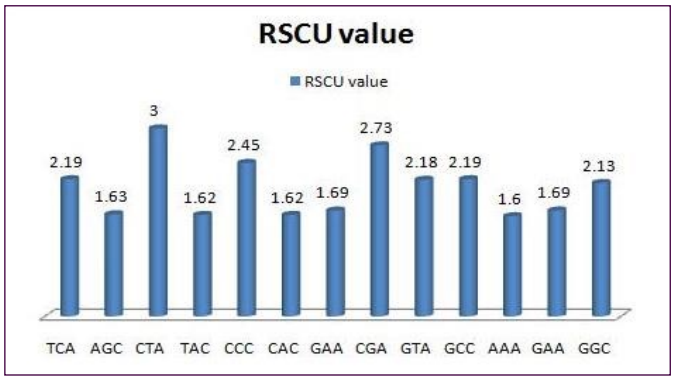

Fig. 8- Distribution of over represented codons for ATPase 6 gene

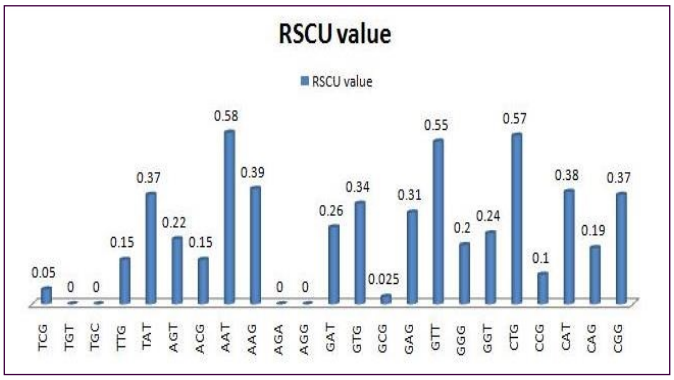

Fig. 9- Distribution of under represented codons for ATPase 6 gene

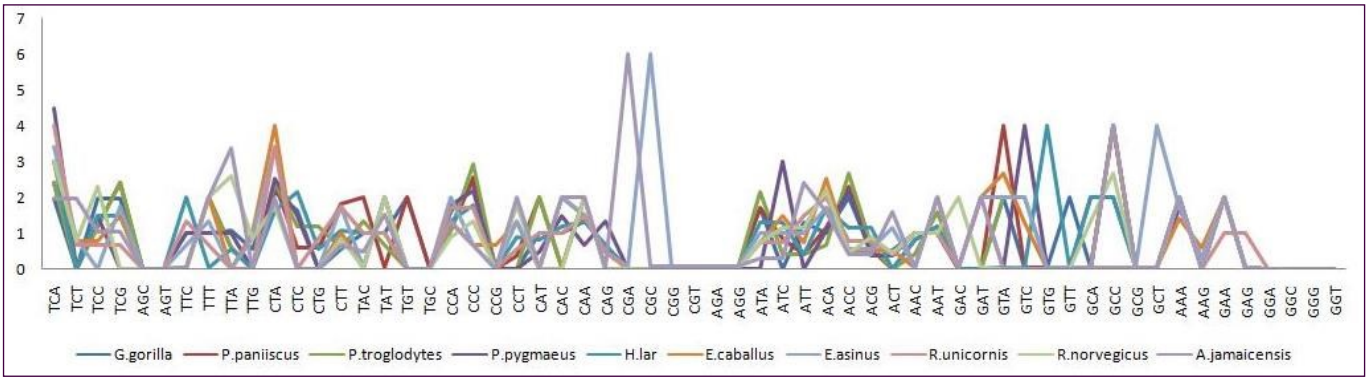

Fig. 10- Comparison of RSCU value for 59 codons for ATPase 8 gene in different species of Mammals

These results provide evidence that compositional constraint plays an important role in codon usage in mitochondrial ATPase 8 gene in mammals. The comparisons of RSCU value for 59 codons in different species of mammals are shown in [Fig-10]. The distribution of the over represented and the under represented codons are shown in [Fig-11] and [Fig-12], respectively. These results suggest that a base composition is a major contributing factor in codon usage pattern in the mammalian species.

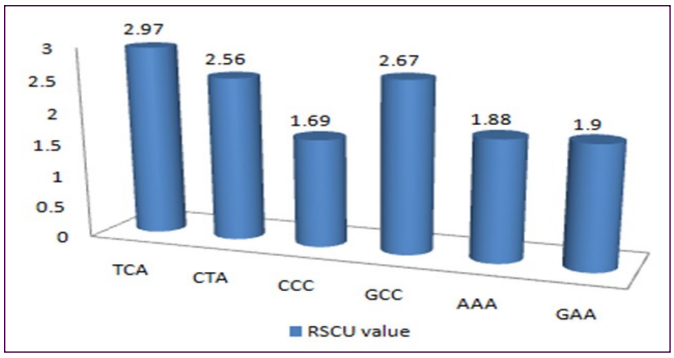

Fig. 11- Distribution of over represented codons in ATPase 8 gene

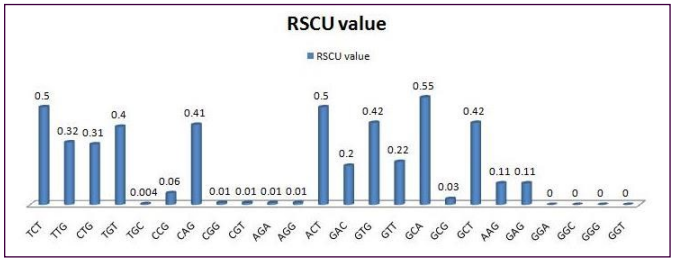

Fig. 12- Distribution of under represented codons in ATPase 8 gene

Amino Acids Contribute Differently to a Gene's Codon Usage Bias (CUB)

For ATPase 6 gene, except methionine and tryptophan, the number of other amino acids used is different. The amino acid, leu is clearly the single amino acid that accounts for the greatest usage in a gene and the amino acids cys and asp account for the least use as shown in [Fig-13]. In case of ATPase 8, leu is distinctly the single amino acid that accounts for the greatest usage and gly accounts for the least usage [Fig-14].

Effect of Mutational Bias on Codon Usage Variation

To ascertain whether the evolution of codon usage bias for ATPase 6 and ATPase 8 genes in the mammalian species had been driven by mutation pressure alone or by translation selection, we compared the correlation between the overall nucleotide composition (A, T, G, C and GC%) and the nucleotide composition at the 3rd codon position (A3, T3, G3, C3 and GC3%) using Pearson correlation analysis [Table-4]. In ATPase 6, a significant positive correlation was observed between A% and A3% (r =0.9065, p<0.001), C% and C3% (r = 0.9238, p<0.001), GC% and GC3% (r = 0.9244, p< 0.001) but negative correlation was observed between most of other nucleotide pairs.

For ATPase 8 gene, significant positive correlations was observed between A% and A3% (r=0.6502, p<0.05), C% and C3% (r =0.9804, p<0.001), GC% and GC3% (r =0.9732, p<0.001) but significant negative correlation was found for most of other nucleotide comparisons [Table-4]. These results further suggest that compositional constraint under mutation pressure is a major factor for the pattern of codon usage bias for ATPase 6 gene in mammalian species.

That the pattern of codon usage bias is primarily governed by muta-

tion pressure was further supported by examining the correlation between ENC and GC, ENC and GC3. High significant positive correlation was found between ENC and GC content (r =0.9479, p<0.001) and also highly significant positive correlation between (r = 0.9387, p<0.001) ENC with GC3 for ATPase 6 gene [Table-3]. For ATPase 8 gene, highly significant positive correlation was found between ENC and GC (r = 0.9333 p<0.001), ENC and GC3 (r = 0.9811, p<0.001) [Table-3].

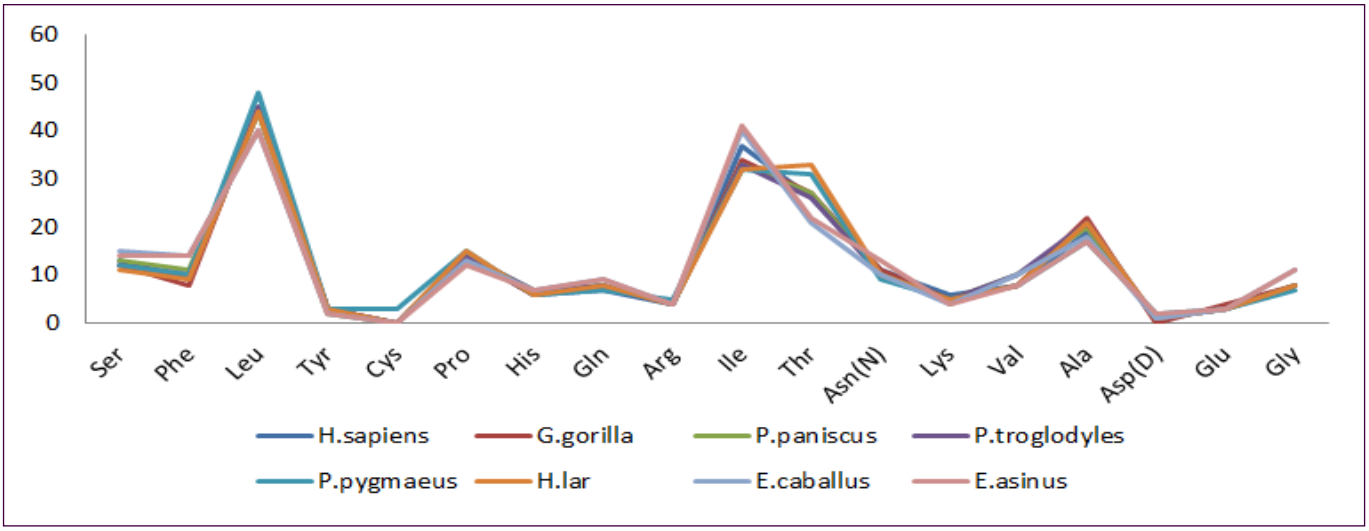

Fig. 13- Distribution of amino acid in ATPase 6 gene product in different mammalian species

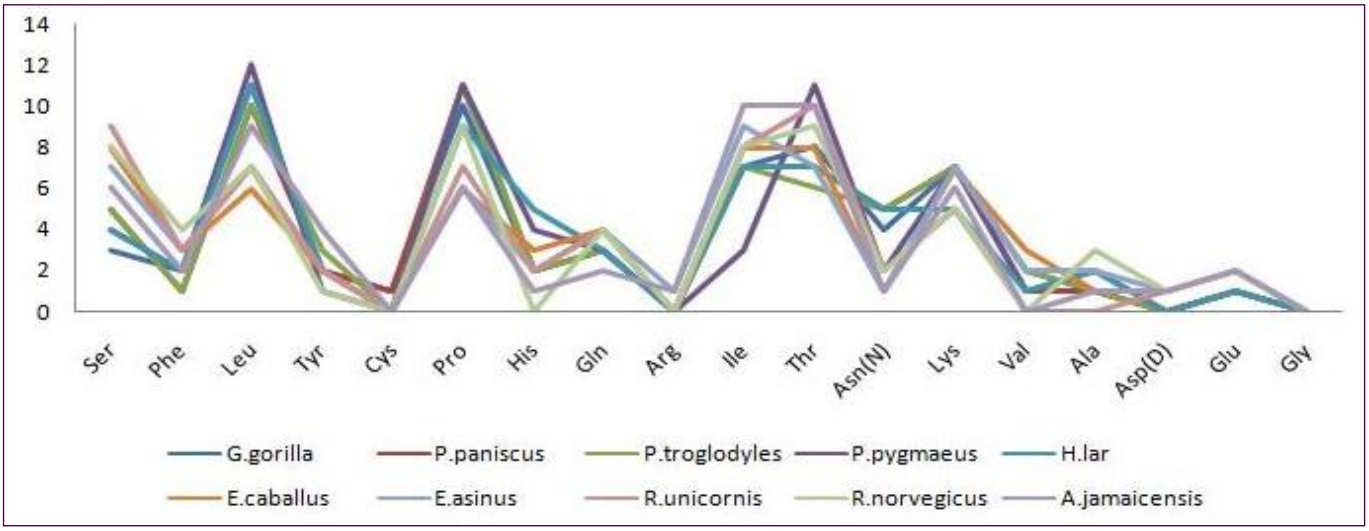

Fig. 14- Distribution of Amino acid in ATPase 8 gene product

Table 4- Correlation between overall nucleotide composition and composition at 3rd position for ATPase 6 and ATPase 8 genes

Nucleotide	A3%	T3%	G3%	C3%	GC3%
ATPase 6					
A%	r =0.9065***	r =0.2756	r =-0.3245	r =-0.7682	r =-0.7350*
T%	r =0.5474	r == 0.8485**	r =-0.7639*	r =-0.7946**	r =-0.8321
G%	r =0.6544	r =-0.1346	r =0.2516	r =-0.4337	r= -0.3394
C%	r = -0.9060***	r =-0.5726	r == 0.5111	r =0.9238***	r =0.9034**
GC%	r = -0.8251**	r =-0.6996	r =0.6656	r =0.9174***	r =0.9244***
ATPase 8					
A%	0.6502*	0.5987	-0.006	-0.7391**	-0.6744*
T%	0.6526*	0.9851***	-0.4384	-0.9318***	-0.9684***
G%	0.1125	-0.0541	0.7645**	-0.2253	0.0027
C%	-0.7093*	-0.9241***	0.1771	0.9804***	0.9413***
GC%	-0.7063*	-0.9675***	0.3624	0.9600***	0.9732***

These results further revealed that mutation pressure is a major factor for codon usage bias in mammals.

Furthermore, the GC content at the first and second codon positions (GC1% and GC2%) was compared with the GC content at the third codon position (GC3%). Positive correlation was found between GC12% with GC3% (r=0.3611) for ATPase 6 gene. But for ATPase 8, a highly positive correlation (r = 0.8270 p<0.01) was observed between GC12 with GC3. This result also supported the hypothesis that mutation pressure on nucleotide was a major determinant for the pattern of codon usage in mammalian species.

Discussion

The present work is consonant with the previous work on mitochondrial genomes. Codon usage bias in mitochondrial genomes is mainly due to mutation pressure because in mitochondria, rate of mutation is high. Huge numbers of free radicals are formed inside

the mitochondria during respiration process near the mitochondrial DNA which could lead to its high mutation rate. As a consequence of high mutation rate in mitochondrial genome, the evolution of mitochondrial genes is much faster than nuclear genes. Mutation rate is different in the two strands of DNA duplex [33]. It was also found that mutation pressure is the major force which contributes to the pattern of codon usage bias for two mitochondrial genes ATPase 6 and ATPase 8 in mammalian species. Analysis of the codon usage bias for ATPase 6 and ATPase 8 genes in mitochondrial genomes is expected to contribute to the understanding of the characteristics and the molecular evolution of these genes.

We have analyzed the synonymous codon usage bias in ATPase 6 and ATPase 8 genes in different species of mammals. In this study, we found that the most frequent codons end with A or C at the 3rd codon position. The ENC values were high, so the codon usage bias was not remarkable. This finding may be the result of compositional constraint and the active process of mutation that occurred in codon usage pattern in ATPase 6 and ATPase 8 genes in some mammalian species.

Mutational pressure has been shown to influence the codon usage bias in these species. Highly significant positive correlation was observed between A% and A3%, C% and C3%, GC% and GC3% but negative correlation between G% and C3% (r = -0.4337), GC% and T3% (r = -0.6996) for ATPase 6. Similarly, significant positive correlations were observed between A% and A3% (r=0.6502, $p<0.05$), C% and C3% (r =0.9804, $p<0.001$), GC% and GC3% (r =0.9732, $p<0.001$) but significant negative correlation was found for most of other nucleotide comparisons for ATPase 8. These results suggested that mutation pressure is a major factor for the pattern of codon usage bias in ATPase 6 and ATPase 8 in these species. Earlier workers also found that codon usage is known to be influenced by nucleotide composition [34]. Previous evidence further suggested that mutation is the main force responsible for codon usage bias in mitochondrial genomes because in mitochondria, the mutation rate is higher [35]. The positive correlation between ENC and CAI in both ATPase 6 and ATPase 8 indicates that the codon usage bias across the species has very distinct relationships with nucleotide composition of the coding sequences. Earlier work on mitochondrial genomes did not find any translational selection in the mitochondrial genomes [35].

That the pattern of codon usage bias is primarily governed by mutation pressure was further supported by examining the correlation between ENC and GC, ENC and GC3. High significant positive correlation was found between ENC and GC content (r =0.9479, $p<0.001$) between (r = 0.9387, $p<0.001$) ENC and GC3 for ATPase 6. In ATPase 8 gene, highly significant positive correlation was found between ENC and GC3 (r = 0.9811, $p<0.001$), ENC and GC (r = 0.9333 $p<0.001$), respectively.

Furthermore, GC content at the first and second codon positions (GC1% and GC2%) was compared with the GC content at the third codon position (GC3%). Positive correlation between GC12% and GC3in ATPase 6 (r=0.3611, p>0.05) and in ATPase 8 (r = 0.8270 $p<0.01$), suggests that mutation pressure was an important factor in shaping the codon usage bias in mitochondrial and ATPase 6 and ATPase 8 genes in mammalian species. In addition, all these results support the hypothesis that compositional constraint under mutation pressure is responsible for codon usage bias in mammalian mitochondria.

Conclusion

This is the first work on the pattern of codon usage in mitochondrial ATPase 6 and ATPase 8 genes in mammalian species. This work is initiated for understanding the pattern of evolution of codon usage in these species. Codon usage bias is low for mitochondrial genes in mammals. Nucleotide constraint and compositional constraint are significant factors that affect codon usage pattern in these species. We have found that mutation pressure is the main factor for pattern of codon usage bias coupled with natural selection. However further analysis is needed for understanding if there is any other factor which is responsible for codon usage bias in mitochondrial genes

Acknowledgement: We are thankful to the Department of Biotechnology, Assam University, Silchar for providing lab facility. AU is also thankful to the University Grant Commission, India for SRF-MANF (Maulana Azad National Fellowship).

Conflict of Interest: Authors declare that there is no conflict of interest in this research work.

References

[1] Ikemura T. (1981) *Journal of Molecular Biology*, 151(3), 389-409.

[2] Ikemura T. (1985) *Molecular Biology and Evolution*, 2(1), 13-34.

[3] Behura S.K. & Severson D.W. (2012) *PloS one*, 7(8), e43111.

[4] Akashi H. & Eyre-Walker A. (1998) *Current Opinion in Genetics & Development*, 8(6), 688-693.

[5] Akashi H. (2001) *Current Opinion in Genetics & Development*, 11(6), 660-666.

[6] Duret L. (2002) *Current Opinion in Genetics & Development*, 12 (6), 640-649.

[7] Grantham R., Gautier C. & Gouy M. (1980) *Nucleic Acids Research*, 8(9), 1893-1912.

[8] Akashi H. (1997) *Gene*, 205(1), 269-278.

[9] Powell J.R. & Moriyama E.N. (1997) *Proceedings of the National Academy of Sciences*, 94(15), 7784-7790.

[10] Moriyama E.N. & Powell J.R. (1998) *Nucleic Acids Research*, 26(13), 3188-3193.

[11] Powell J.R., Sezzi E., Moriyama E.N., Gleason J.M., Caccone A. (2003) *Journal of Molecular Evolution*, 57(1), S214-S225.

[12] Hou Z.C. & Yang N. (2002) *Yi chuan xue bao= Acta genetica Sinica*, 29(8), 747-752.

[13] Sharp P.M. & Li W.H. (1986) *Journal of Molecular Evolution*, 24 (1-2), 28-38.

[14] Sharp P.M. & Cowe E. (1991) *Yeast*, 7(7), 657-678.

[15] Stenico M., Lloyd A.T. & Sharp P.M. (1994) *Nucleic Acids Research*, 22(13), 2437-2446.

[16] Duret L. & Mouchiroud D. (1999) *Proceedings of the National Academy of Sciences*, 96(8), 4482-4487.

[17] Miyasaka H. (2002) *Journal of Molecular Evolution*, 55(1), 52-64.

[18] Sharp P.M., Stenico M., Peden J.F. & Lloyd A.T. (1993) *Biochemical Society Transactions*, 21(4), 835-841.

[19] Francino M.P. & Ochman H. (1999) *Nature*, 400(6739), 30-31.

[20] Powell J.R. & Moriyama E.N. (1997) *Proceedings of the Nation-*

al Academy of Sciences, 94(15), 7784-7790.

[21] Chiapello H., Lisacek F., Caboche M. & Hénaut A. (1998) *Gene*, 209(1), GC1-GC38.

[22] Wang H.C. & Hickey D.A. (2007) *BMC Evolutionary Biology*, 7 (1), S6.

[23] Ingvarsson P.K. (2007) *Molecular Biology and Evolution*, 24(3), 836-844.

[24] Angellotti M.C., Bhuiyan S.B., Chen G. & Wan X.F. (2007) *Nucleic Acids Research*, 35(2), W132-W136.

[25] Swire J., Judson O.P. & Burt A. (2005) *Journal of Molecular Evolution*, 60(1), 128-139.

[26] Sharp P.M., Tuohy T.M. & Mosurski K.R. (1986) *Nucleic Acids Research*, 14(13), 5125-5143.

[27] Gupta S.K. & Ghosh T.C. (2001) *Gene*, 273(1), 63-70.

[28] Ma J.J., Zhao F., Zhang J., Zhou J.H., Ma N.L., Ding Z.Y., Chen T.H., Gu X.Y. & Liu S.Y. (2013) *Journal of Animal and Veterinary Advances,* 12(1), 88-98.

[29] Wright F. (1990) *Gene*, 87(1), 23-29.

[30] Zhang Z., Dai W. & Dai D. (2013) *PloS one*, 8(11), e81469.

[31] Sharp P.M. & Li W.H. (1987) *Nucleic Acids Research*, 15(3), 1281-1295.

[32] Zhang Z., Dai W., Wang Y., Lu C. & Fan H. (2013) *Archives of Virology*, 158(1), 145-154.

[33] Plotkin J.B. & Kudla G. (2010) *Nature Reviews Genetics*, 12(1), 32-42.

[34] Jenkins G.M. & Holmes E.C. (2003) *Virus Research*, 92(1), 1-7.

[35] Jia W. & Higgs P.G. (2008) *Molecular Biology and Evolution*, 25 (2), 339-351.