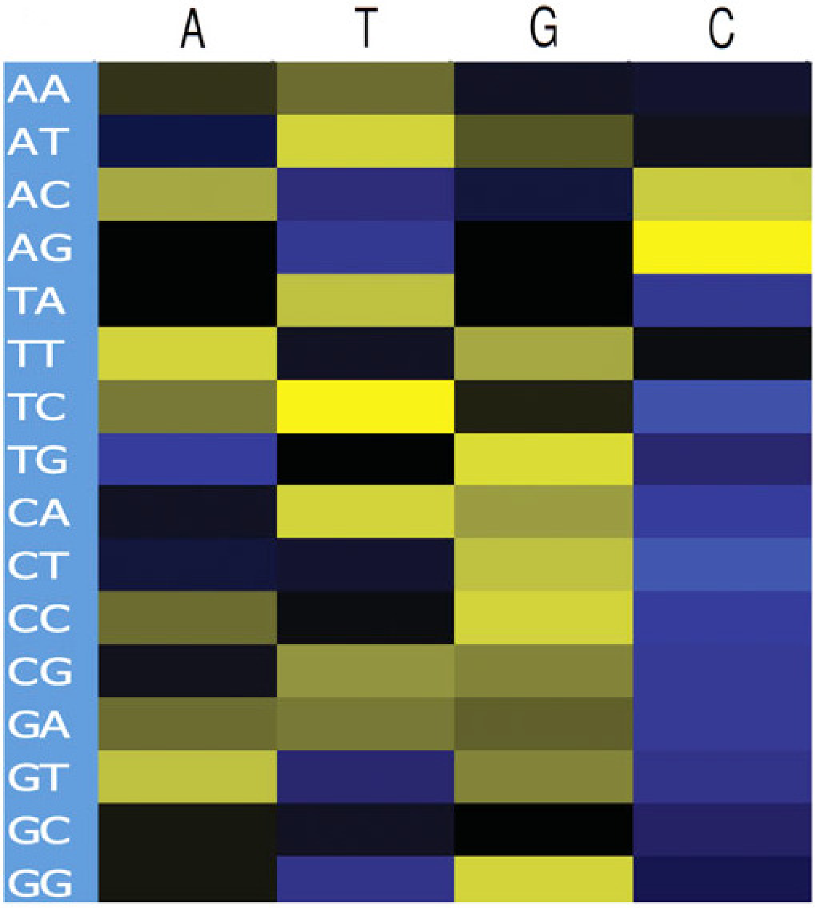
CHAPTER-4

RESULTS



4. Results

4.1. Analysis of codon usage bias in MT-ND gene among Pisces, Aves and Mammals

4.1.1 Codon usage pattern in MT-ND gene among pisces, aves and mammals

To understand the pattern of random usage of synonymous codons in pisces, aves and mammals for MT-ND gene, relative synonymous codon usage (RSCU) of individual codons was calculated and compared.

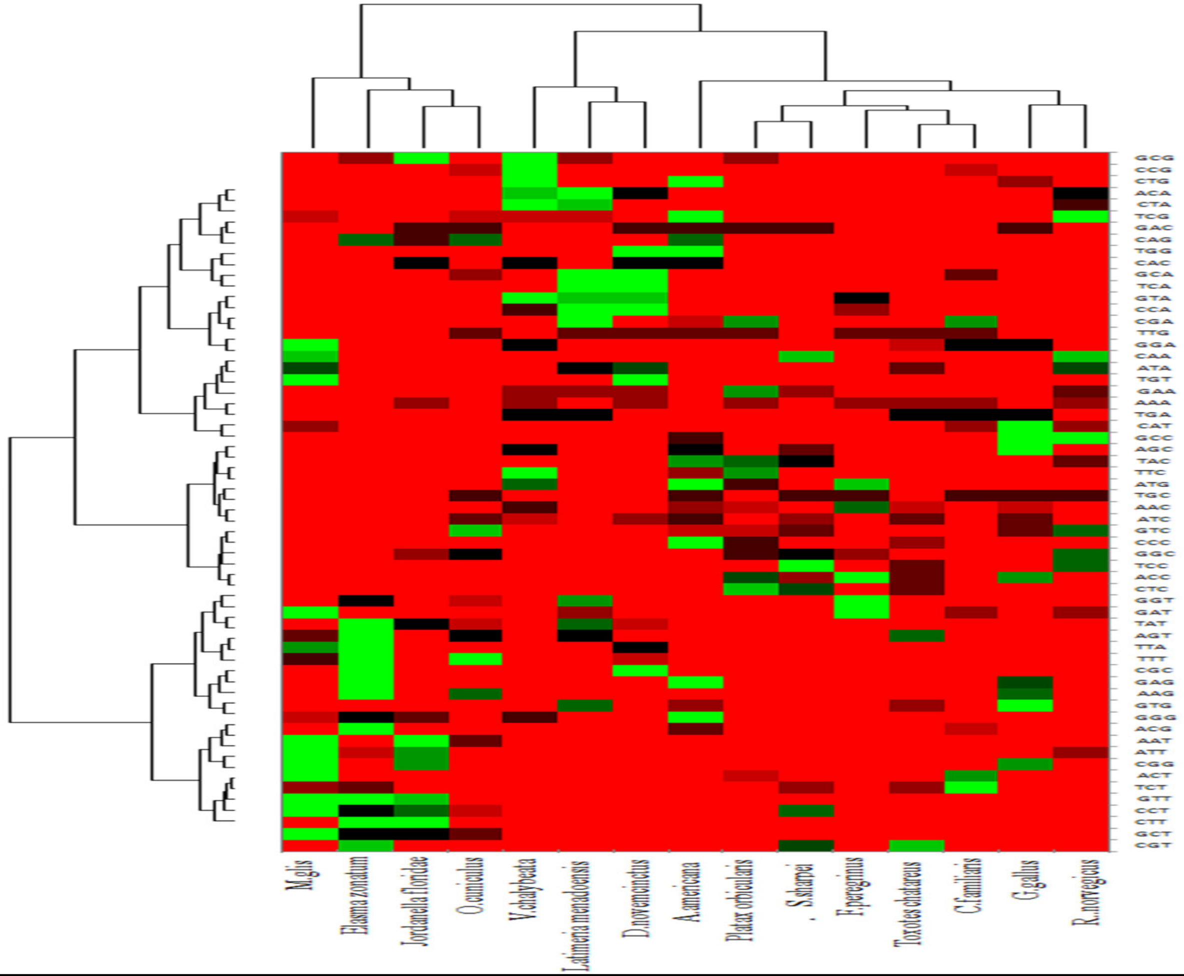


Fig. 4.1.1 Hierarchial Clustering of RSCU values of each codon among pisces, aves and mammals. Each rectangular box on the map represents the RSCU value of a codon (shown in rows) corresponding to MT-ND1 gene among pisces, aves and mammals (shown in columns). The color coding indicates different RSCU values: green indicates RSCU>1.6, black indicates RSCU>1 and dark red indicates RSCU<1 and red indicates RSCU <0.6.

In ND1 gene, out of 60 codons, GCG, CTG, CTA, TCG, TGG, TCA, GTA, CCA, CAA, CCT encoding amino acids ala, leu, leu, ser, trp, ser, val, pro, gln, and pro respectively were used as over-represented codons in some species but as under-presented codons in most of the species as shown in **Figure 4.1.1.** In ND2 gene, ACG, CGT, AGC, CCG, TCA, AAG, and GAT encoding amino acids thr, arg, ser, pro, ser, lys, and asp respectively were used as over-represented codons in some species but as under-presented codons in most of the species as shown in **Figure 4.1.2.**

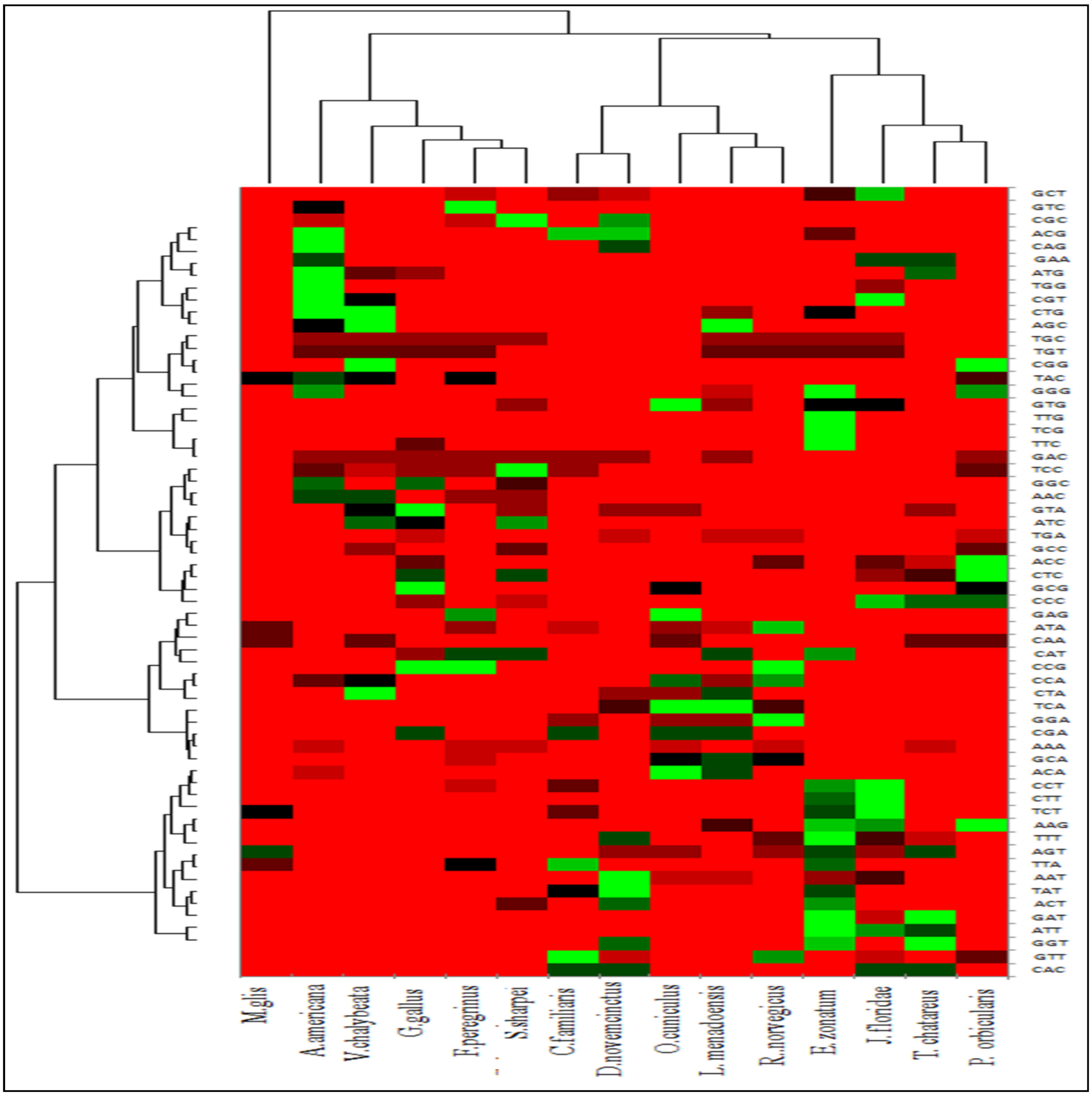


Figure 4.1.2 Hierarchial Clustering of RSCU values of each codon among pisces, aves and mammals for ND2 gene

In ND3 gene, the codons GAT, CAC, CCG, CGC, TAT and TGG encoding amino acids asp, his, pro, arg, tyr and trp respectively were used as over-represented codons in some species but as under-presented codons in most of the species as shown in **Figure 4.1.3.**

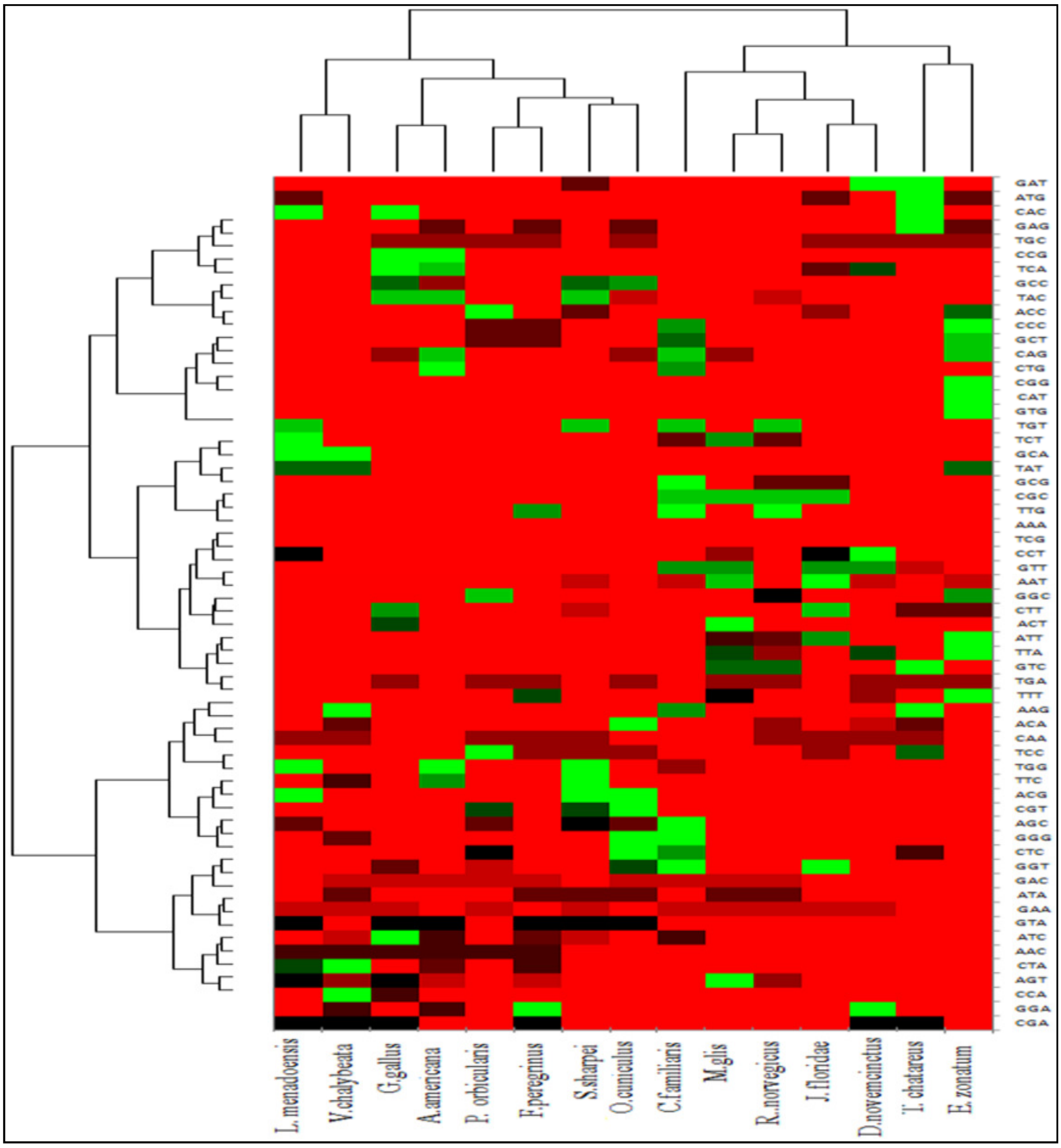


Figure 4.1.3 Hierarchial Clustering of RSCU values of each codon among pisces, aves and mammals for ND3 gene

In ND4 gene, the codons GCA, ACA, ACG, TTT, AAG, ACT, TTA and TTG encoding amino acids ala, thr, thr, phe, lys, thr, leu and leu respectively were used as over-represented codons in some species but as under-presented codons in most of the species as shown in **Figure 4.1.4.**

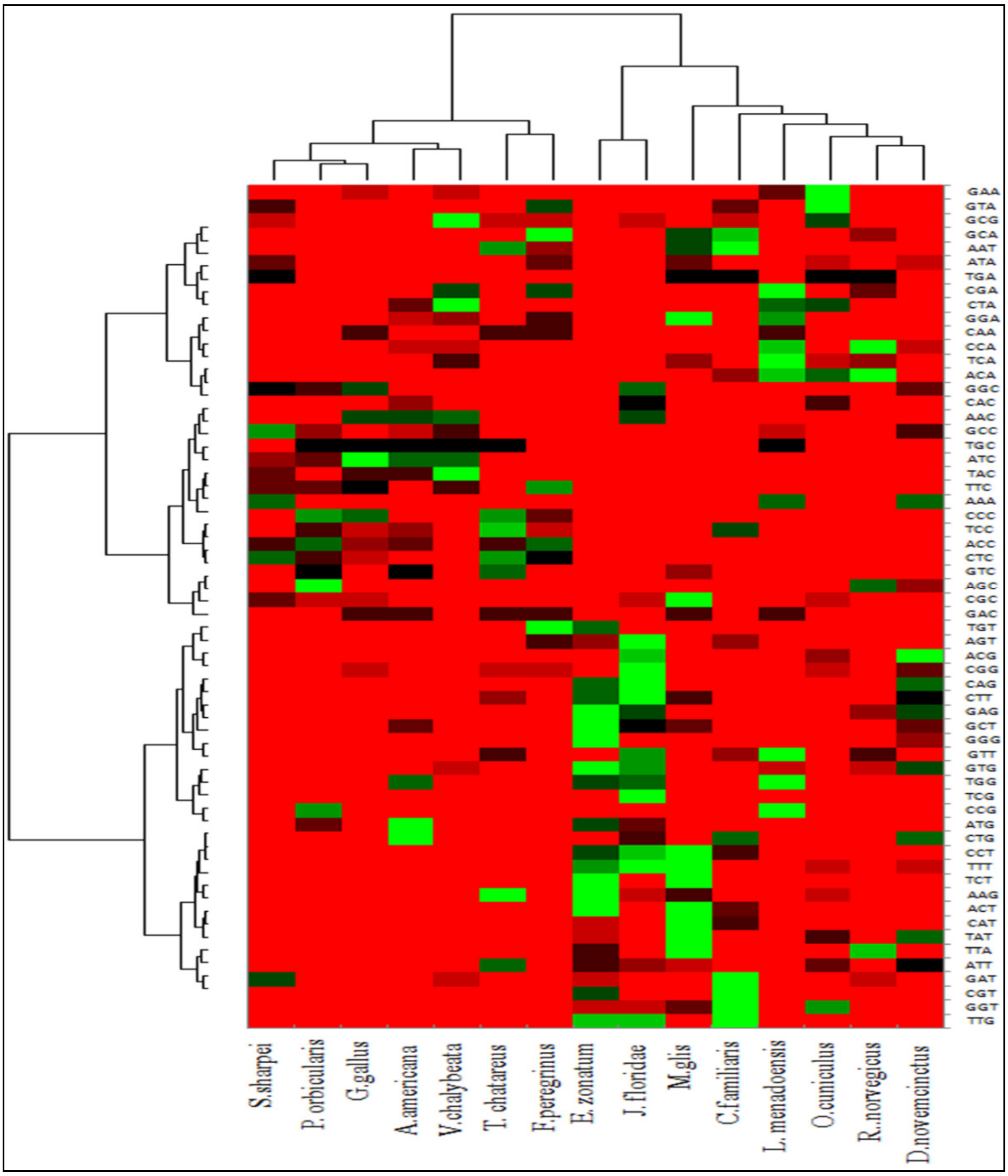


Figure 4.1.4 Hierarchial Clustering of RSCU values of each codon among pisces, aves and mammals for ND4 gene

In ND4L gene, the codons CTC, GCT, AAG, ACT, CTT, TTG, GTT, TAT, GCC and GTG encoding amino acids leu, ala, lys, thr, leu, leu, val, tyr, ala and val respectively were used as over-represented codons in some species but as under-presented codons in most of the species as shown in **Figure 4.1.5.**

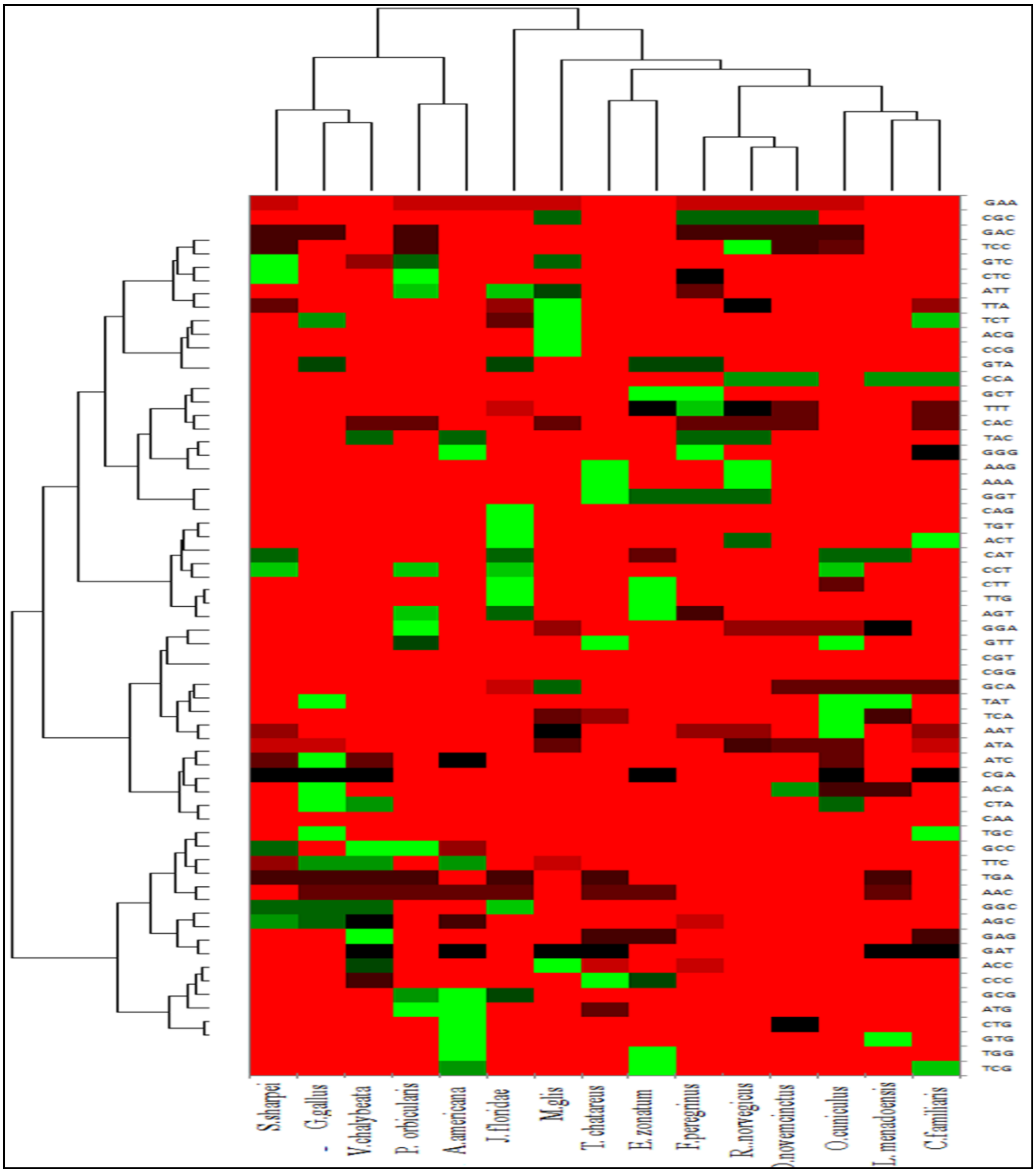


Figure 4.1.5 Hierarchial Clustering of RSCU values of each codon among pisces, aves and mammals for ND4l gene

In ND5 gene, the codons CTA, TCA, TGT, TCC, CCG, GTG, CAG, CGG, ACG and GGG encoding amino acids leu, ser, cys, ser, pro, val, gln, arg, thr and gly respectively, were used as over-represented codons in some species but as under-presented codons in most of the species as shown in **Figure 4.1.6**.

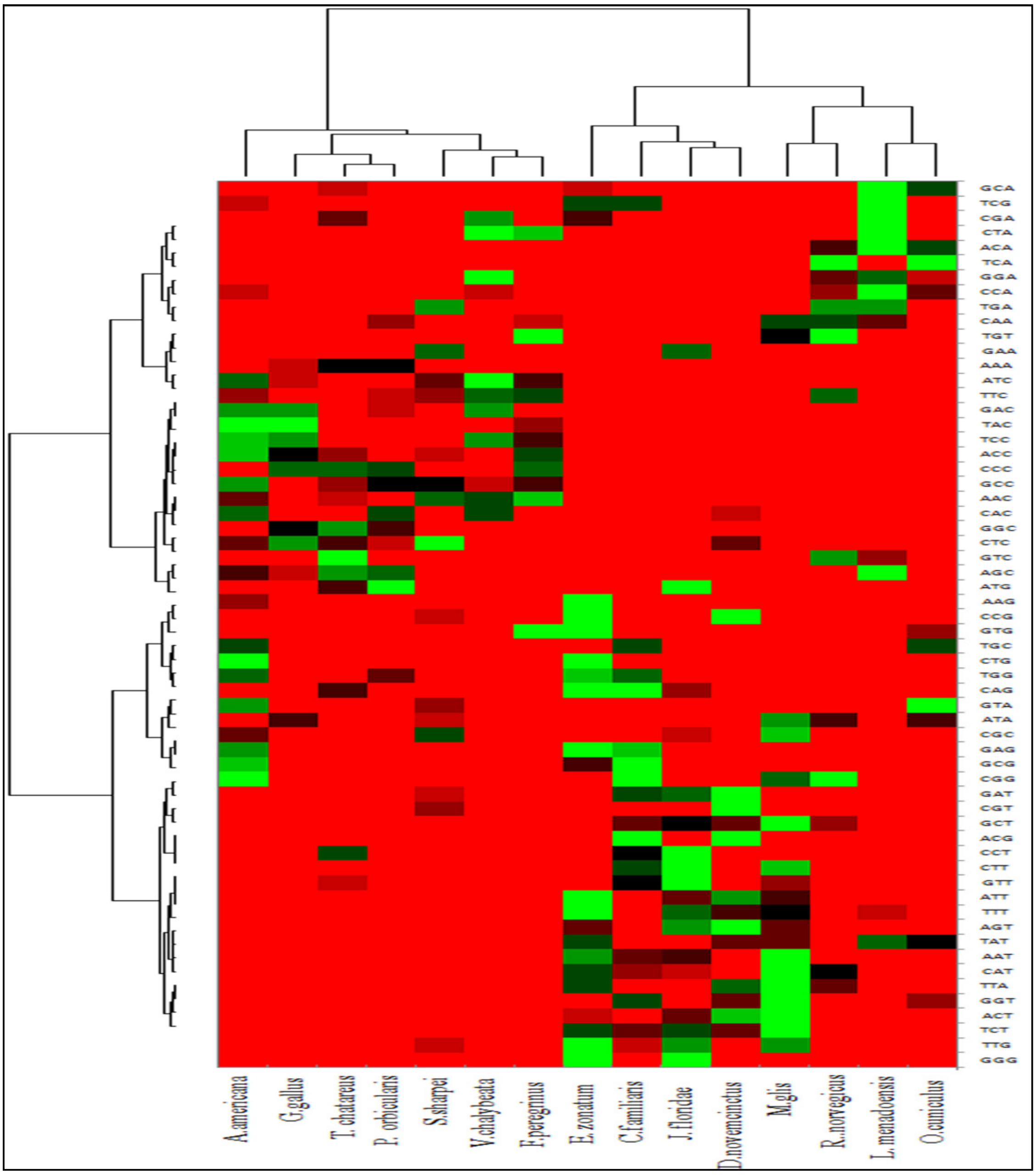


Figure 4.1.6 Hierarchial Clustering of RSCU values of each codon among pisces, aves and mammals for ND5 gene

In ND6 gene, the codons TAC, CCT, GGA, GAG, TTT, TAT and GAT encoding amino acids tyr, pro, gly, glu, phe, tyr and asp respectively were used as over-represented codons in some species but as under-presented codons in most of the species as shown in **Figure 4.1.7.**

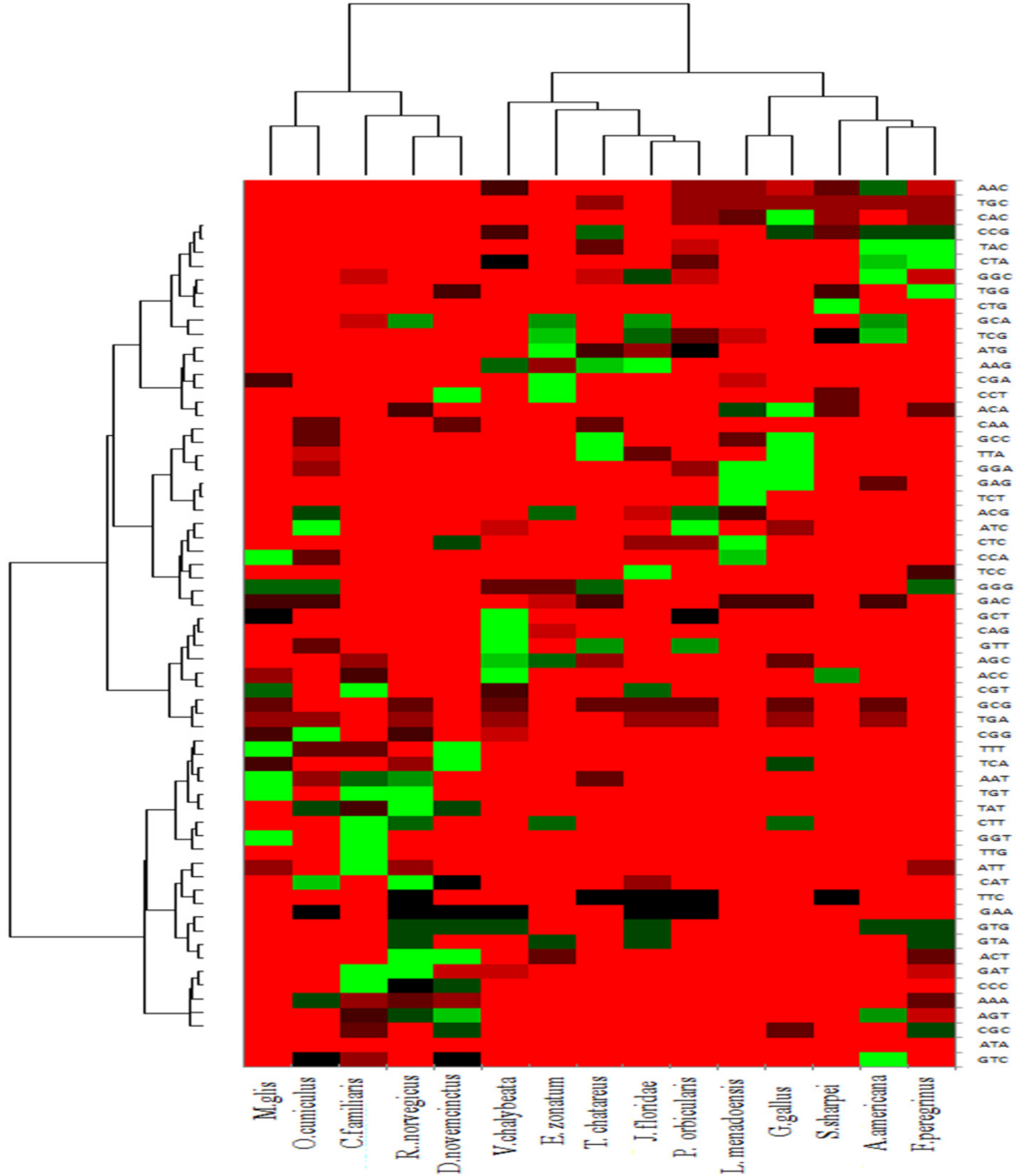


Figure 4.1.7 Hierarchial Clustering of RSCU values of each codon among pisces, aves and mammals for ND6 gene

Based on hierarchical clustering of relative synonymous codon usage, it was found that over- represented and under-represented codons were species specific in pisces, aves and mammals. In all ND genes, some codons were used either as over-represented codons in one species or as rare codons in other species and vice versa (Fig 3). Further, it was observed that some codons were more frequently used in some species but less frequently used in other species. Based on RSCU analysis and nucleotide composition we deduced that the existence of preferred codons in coding sequences has been mostly influenced by compositional constraints, which account for the presence of mutation pressure.

4.1.2 Analysis of codon usage bias in mitochondrial ND gene

In ND1 gene, the ENC values in different species of pisces, aves and mammals were (Mean ±SD) 57.40±3.20, 59.80±0.44, 56±2.64 respectively. The mean ENC values in mammals were lower than aves and pisces which suggest that codon usage bias was high in different species of mammals in comparison to different species of aves and mammals for ND1 gene. From **Figure 4.1.8**, it was found that *M.glis* showed low ENC value in comparison to other species for ND1 gene which indicates high codon usage bias in *M.glis* as compared to other species.

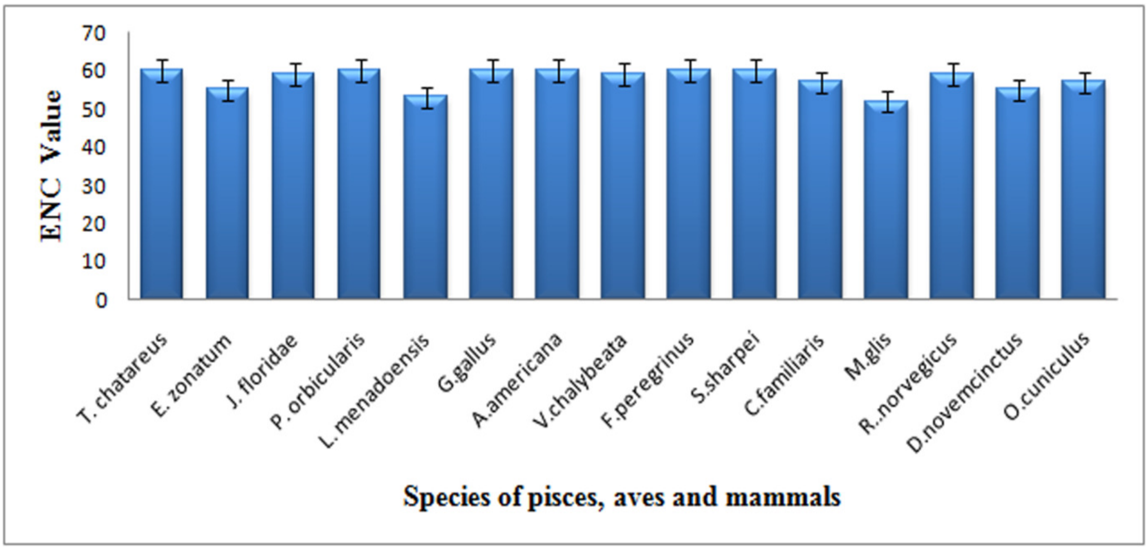


Figure 4.1.8 Distribution of ENC in different species of pisces, aves and mammals for MT-ND1 gene

In ND2 gene, the ENC values in different species of pisces, aves and mammals were (Mean ±SD) 57±2.91, 59.80±0.44, 55±1.58 respectively. The mean ENC values in mammals were lower than aves and pisces which suggest that codon usage bias was high in different species of mammals in comparison to different species of aves and mammals for ND2 gene. From **Figure 4.1.9**, it was found that *L. menadoensis* showed low ENC value in comparison to other species for ND2 gene which indicates high codon usage bias in *L. menadoensis* as compared to other species.

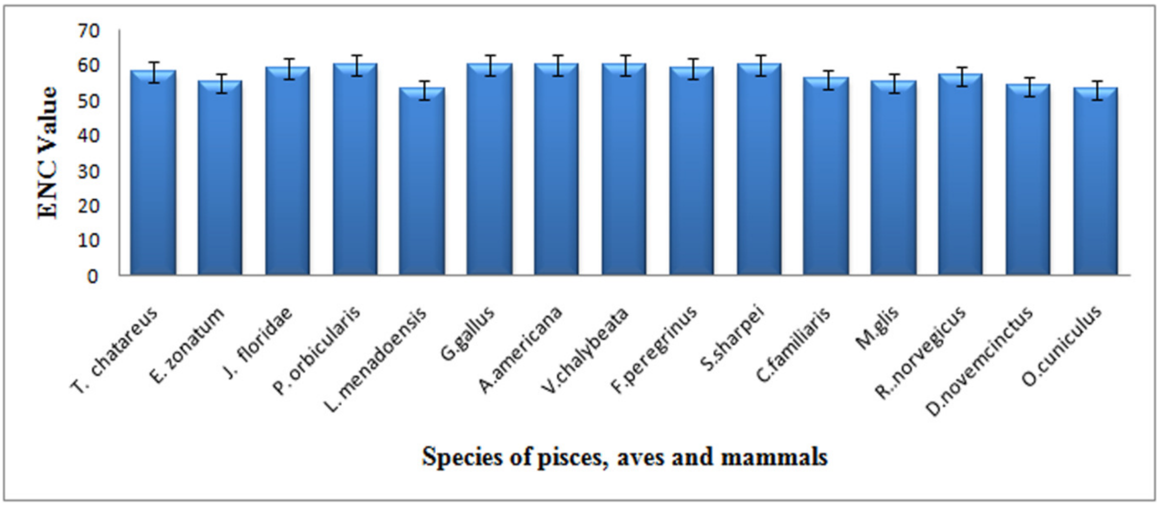


Figure 4.1.9 Distribution of ENC in different species of pisces, aves and mammals for MT-ND2 gene

In ND3 gene, the ENC values in different species of pisces, aves and mammals were (Mean ±SD) 57.8±2.48, 59±1.41, 53.2±3.63 respectively. The mean ENC values in mammals were lower than aves and pisces which suggest that codon usage bias was high in different species of mammals in comparison to different species of aves and mammals for ND3 gene. From **Figure 4.1.10**, it was found that *M.glis* showed low ENC value in comparison to other species in ND3 gene which indicates high codon usage bias in *M.glis* as compared to other species.

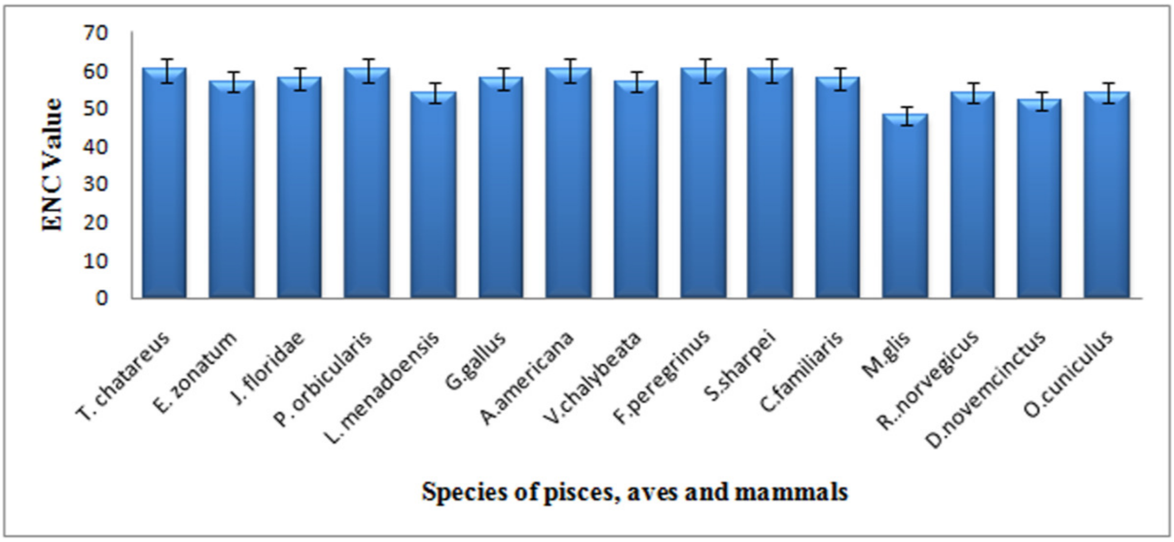


Figure 4.1.10 Distribution of ENC in different species of pisces, aves and mammals for MT-ND3 gene

In ND4 gene, the ENC values in different species of pisces, aves and mammals were (Mean ±SD) 58±2.82, 59.8±0.44, 54.4±3.20 respectively. The mean ENC values in mammals were lower than aves and pisces which suggest that codon usage bias was high in different species of mammals in comparison to different species of aves and mammals for ND4 gene. From **Figure 4.1.11**, it was found that *M.glis* showed low ENC value in comparison to other species in ND4 gene which indicates high codon usage bias in *M.glis* as compared to other species.

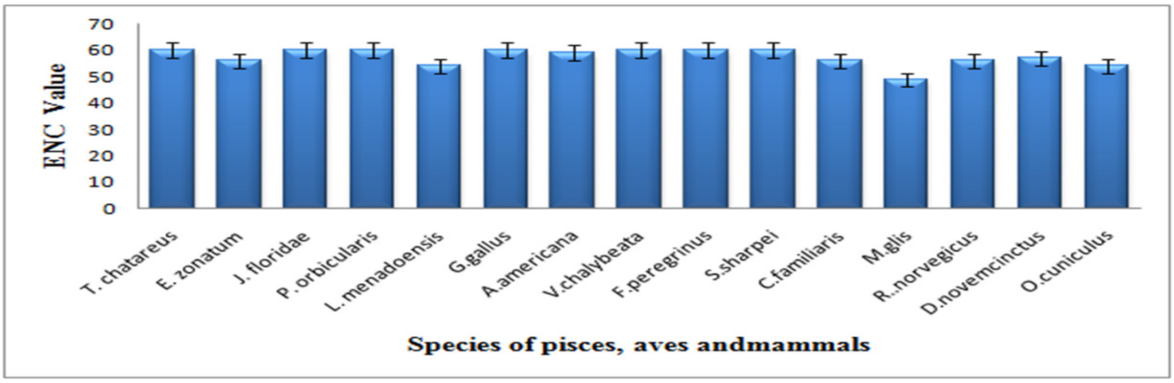


Figure 4.1.11 Distribution of ENC in different species of pisces, aves and mammals for MT-ND4 gene

In ND4L gene, the ENC values in different species of pisces, aves and mammals were (Mean ±SD) 56.6±2.79, 59.2±0.83, 52.6±3.04 respectively. The mean ENC values in mammals were lower than aves and pisces which suggest that codon usage bias was high in different species of mammals in comparison to different species of aves and mammals for ND4L gene. From **Figure 4.1.12**, it was found that *D. nevemcinctus* and *C. familiaris* showed low ENC value in comparison to other species for ND4L gene which indicates high codon usage bias in *D. nevemcinctus* and *C. familiaris* as compared to other species.

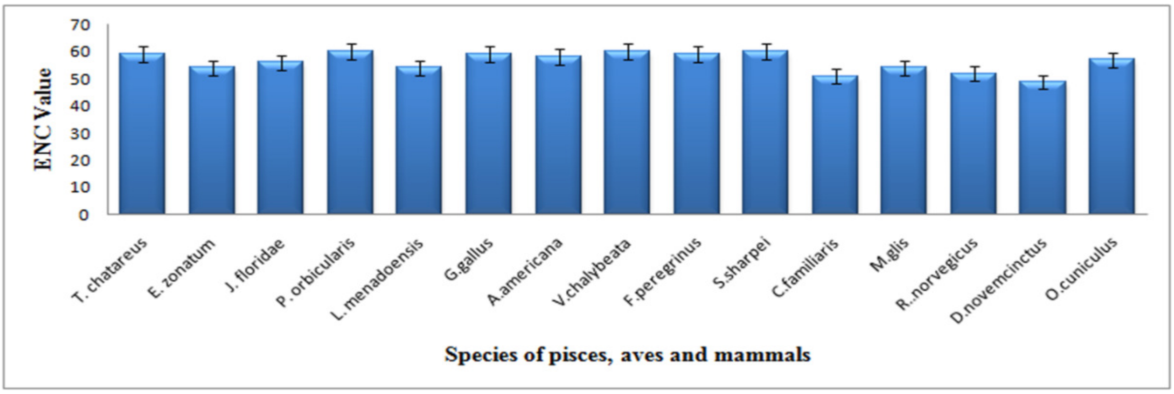


Figure 4.1.12 Distribution of ENC in different species of pisces, aves and mammals for MT-ND4L gene

In ND5 gene, the ENC values in different species of pisces, aves and mammals were (Mean ±SD) 58.2±1.64, 59.80±0.44, 56.2±2.58 respectively. The mean ENC values in mammals were lower than aves and pisces which suggest that codon usage bias was high in different species of mammals in comparison to different species of aves and mammals for ND5 gene. From **Figure 4.1.13**, it was found that *M. glis* showed low ENC value in comparison to other species for ND5 gene which indicates high codon usage bias in *M. glis* as compared to other species.

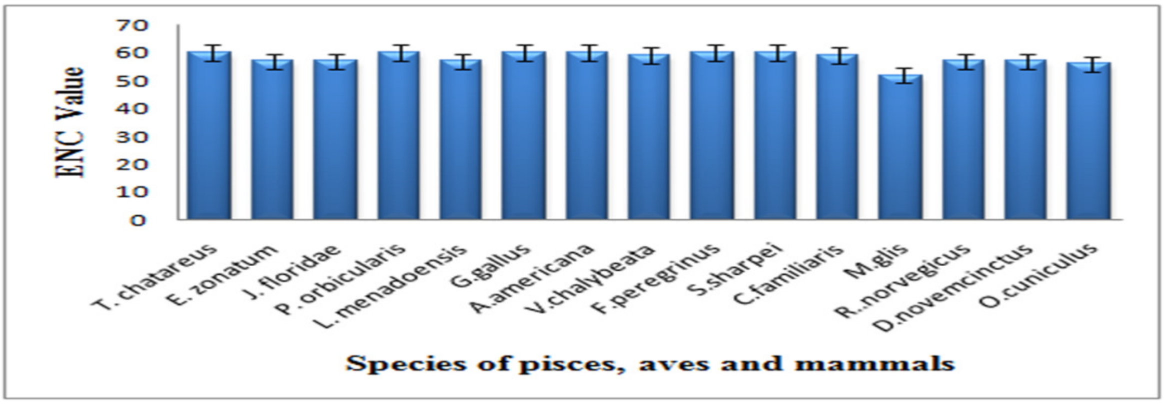


Figure 4.1.13 Distribution of ENC in different species of pisces, aves and mammals for MT-ND5 gene

In ND6 gene, the ENC values in different species of pisces, aves and mammals were (Mean ±SD) 59.6±0.89, 58.4±2.07, 59.4±0.89 respectively. The mean ENC values in aves were lower than pisces and mammals which suggest that codon usage bias was high in different species of aves in comparison to different species of pisces and mammals for ND6 gene.

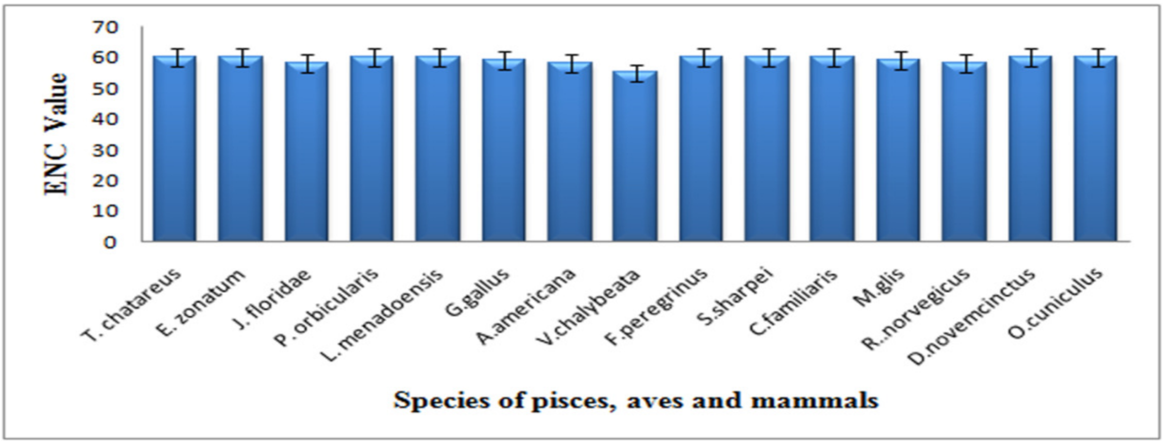


Figure 4.1.14 Distribution of ENC in different species of pisces, aves and mammals for MT-ND6 gene

From **Figure 4.1.14**, it was found that *V. chalybeata* showed low ENC value in comparison to other species in ND6 gene which indicates high codon usage bias in *V. chalybeata* as compared to other species.

4.1.3 Prediction of expression level for mitochondrial ND gene

The CAI is a directional measure of codon usage bias and its higher value means higher gene expression. In many bacteria and small eukaryotes, highly expressed gene revealed stronger bias. In ND1 gene, the mean±SD of different species of pisces, aves and mammals were 0.78±0.053, 0.79±0.066, 0.76±0.03 respectively, which suggest that the expression level of ND1 in different species of pisces, aves and mammals was high. The expression level of *A. americana* was high in comparison to others as shown in **Figure 4.1.15.**

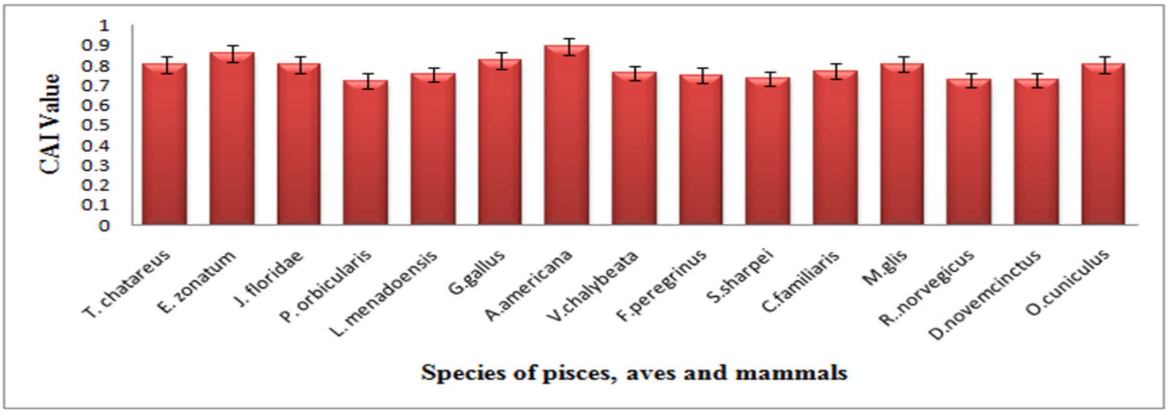


Figure 4.1.15 Expression level in different species of pisces, aves and mammals for MT-ND1 gene

In ND2 gene**,** the mean±SD of different species of pisces, aves and mammals were 0.78±0.05, 0.76±0.05, 0.76±0.02 respectively, which suggest that the expression level of ND2 in different species of pisces, aves and mammals was high. The expression level of *E. zonatum* was high in comparison to others as shown in **Figure 4.1.16.**

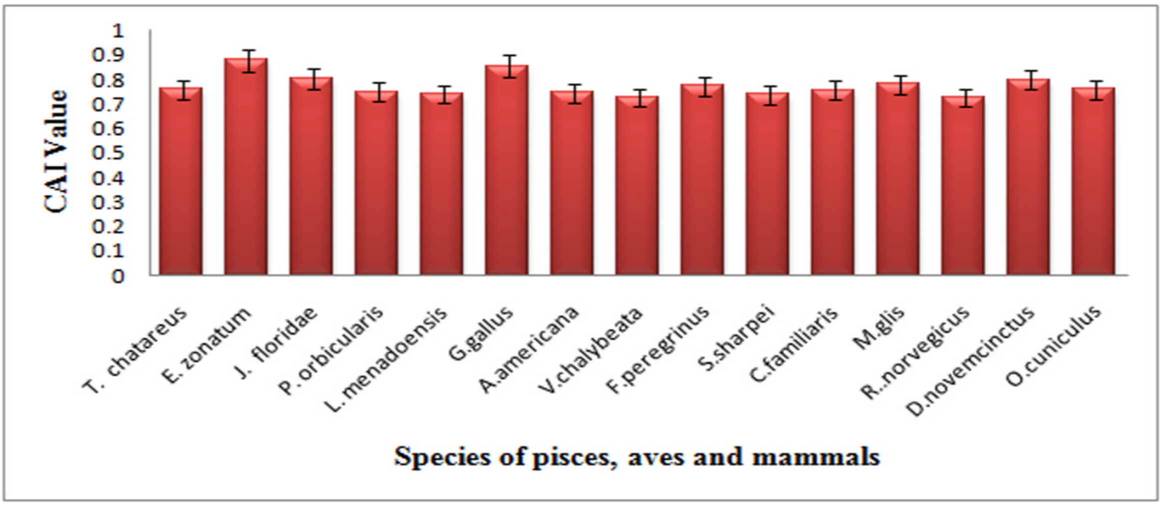


Figure 4.1.16 Expression level in different species of pisces, aves and mammals for MT-ND2 gene

In ND3 gene, the mean±SD of different species of pisces, aves and mammals were 0.44±0.11, 0.30±0.056, 0.46±0.014, respectively, which suggest that the expression level of ND3 in different species of pisces, aves and mammals was moderate. The expression level of *E. zonatum and C. familiaris was* high in comparison to others as shown in **Figure 4.1.17.**

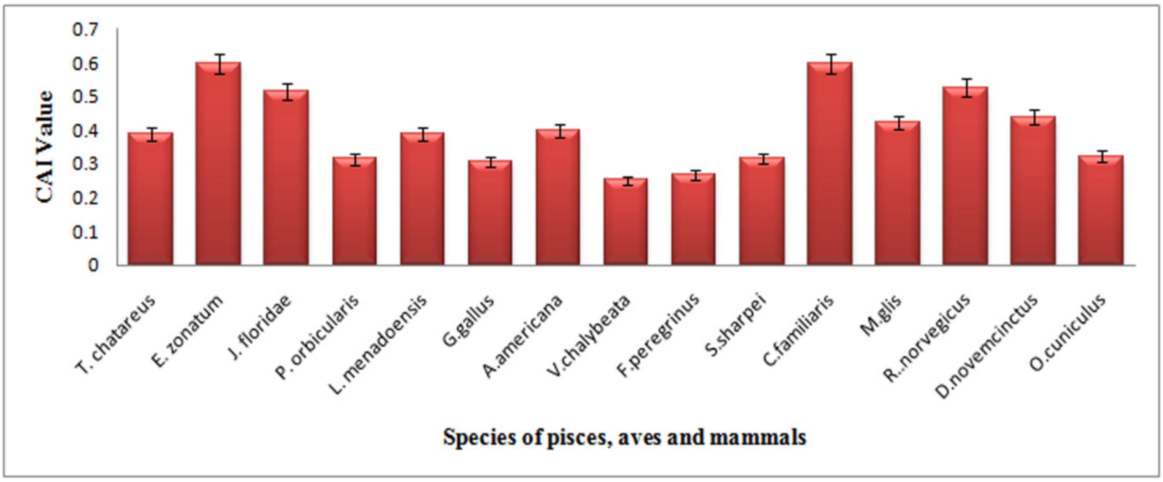


Figure 4.1.17 Expression level in different species of pisces, aves and mammals for MT-ND3 gene

In ND4 gene, the mean±SD of different species of pisces, aves and mammals were 0.85±0.05, 0.81±0.02, 0.84±0.015, respectively, which suggest that the expression level of ND4 in different species of pisces, aves and mammals was high. The expression level of *J. floridae* was high in comparison to others as shown in **Figure 4.1.18.**

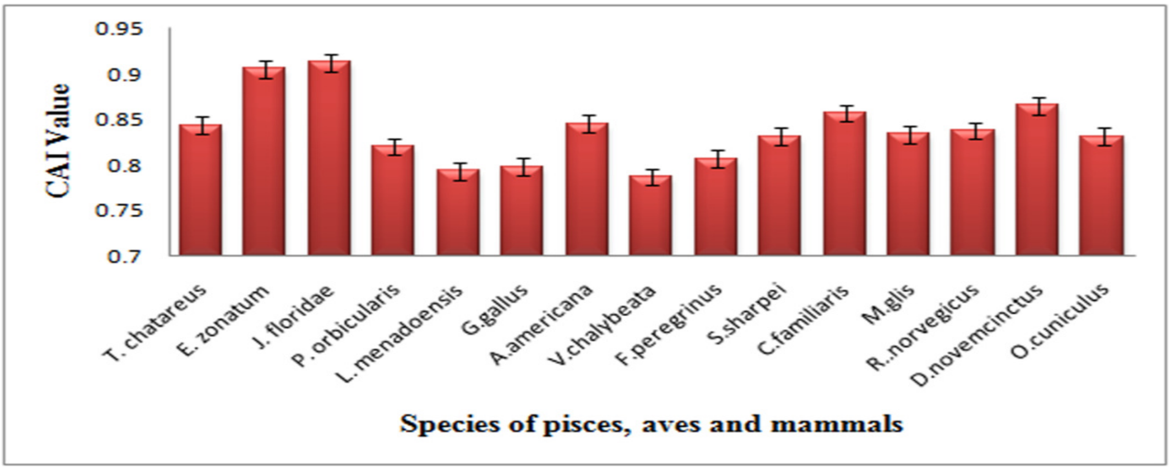


Figure 4.1.18 Expression level in different species of pisces, aves and mammals for MT-ND4 gene

In ND4l gene, the mean±SD of different species of pisces, aves and mammals were 0.35±0.036, 0.33±0.077, 0.36±0.104 respectively, which suggest that the expression level of ND4L in different species of pisces, aves and mammals was moderate. The expression level of *M.glis* was high in comparison to others as shown in **Figure 4.1.19.**

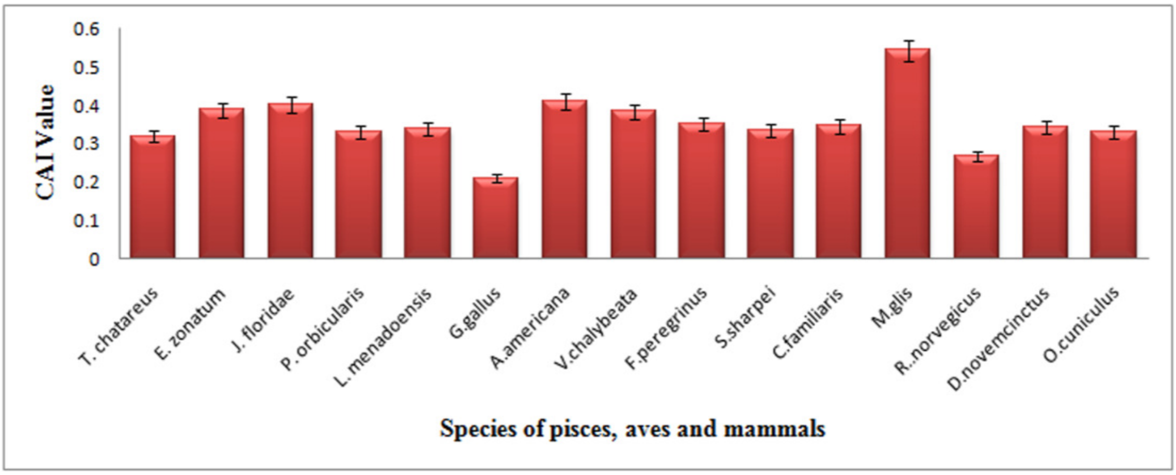


Figure 4.1.19 Expression level in different species of pisces, aves and mammals for MT-ND4l gene

In ND5 gene, the mean±SD of different species of pisces, aves and mammals were 0.88±0.034, 0.86±0.021, 0.89±0.023 respectively, which suggest that the expression level of ND6 in different species of pisces, aves and mammals was high. The expression level of *E. zonatum* was high in comparison to others as shown in **Figure 4.1.20.**

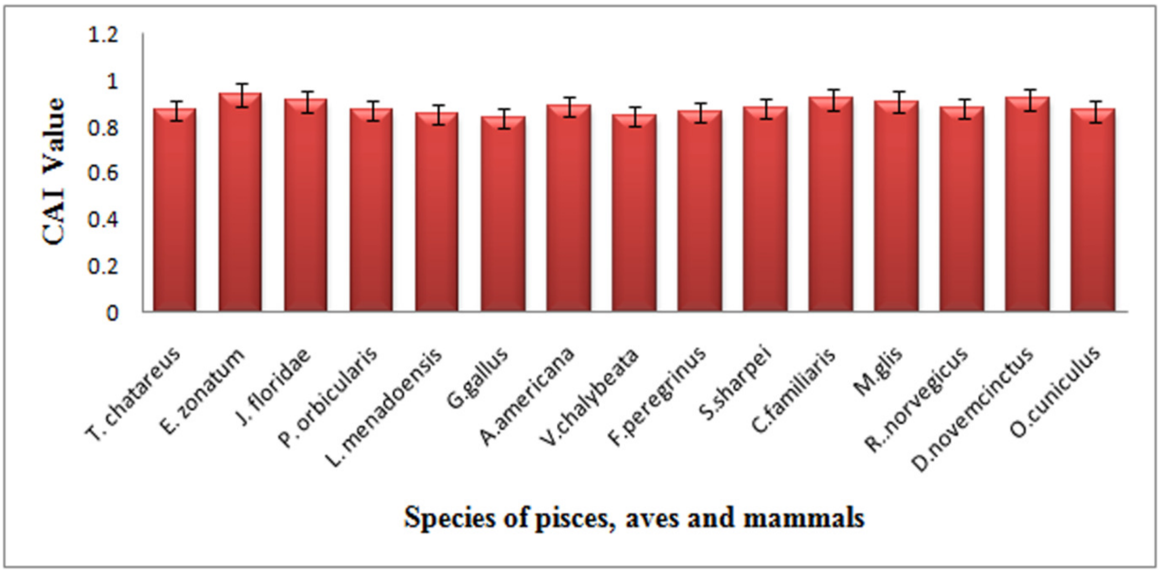


Figure 4.1.20 Expression level in different species of pisces, aves and mammals for MT-ND5 gene

In ND6 gene, the mean±SD of different species of pisces, aves and mammals were 0.53±0.076, 0.53±0.065, 0.51±0.048 respectively, which indicate that the expression level of ND6 in different species of pisces, aves and mammals was high. The expression level of *E. zonatum* was high in comparison to others as shown in **Figure 4.1.21.**

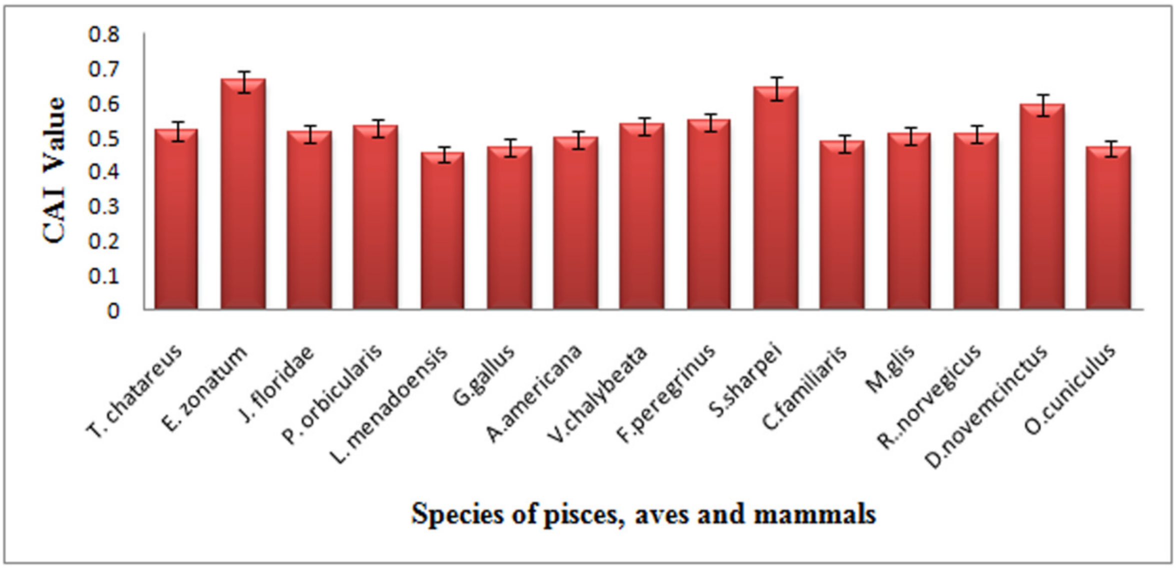


Figure 4.1.21 Expression level in different species of pisces, aves and mammals for MT-ND6 gene

4.1.4. Relationships of gene expression (CAI) with codon usage bias (ENC) for MT-ND gene among pisces, aves and mammals

Comparison of ENC and CAI is used to judge the nucleotide composition and codon selection. We correlated ENC and CAI to know the nucleotide composition variation and codon selection among different species of pisces, aves and mammals for MT-ND gene.

In ND1 gene, we found negative correlation between ENC and CAI in pisces and mammals while positive correlation was found between ENC and CAI in aves. Pearson correlation coefficient was -0.186 ($p>0.05$) in pisces, 0.266 ($p>0.05$) in aves and -0.483 ($p>0.05$) in mammals and these suggest codon usage bias has a very weak relationship with nucleotide composition.

In ND2 gene, we found negative correlation between ENC and CAI in pisces, aves and mammals. Pearson correlation coefficient was -0.156 ($p>0.05$) in pisces, -0.0706 ($p>0.05$) in aves and -0.647 ($p>0.05$) in mammals and these suggest codon usage bias shows a very weak relationship with nucleotide composition.

In ND3 gene, negative correlation was observed between ENC and CAI in pisces, while positive correlation was found between ENC and CAI in aves and mammals. Pearson correlation coefficient was -0.260 ($p>0.05$) in pisces, 0.543 ($p>0.05$) in aves and 0.537 ($p>0.05$) in mammals and these reveal that codon usage bias has a very weak relationship with nucleotide composition.

In ND4 gene, positive correlation was found between ENC and CAI in pisces and mammals while negative correlation was found between ENC and CAI in aves. Pearson correlation coefficient was 0.282 ($p>0.05$) in pisces, -0.728 ($p>0.05$) in aves and 0.658 ($p>0.05$) in mammals and these suggest codon usage bias has a very weak relationship with nucleotide composition.

In ND4l gene, negative correlation was found between ENC and CAI in pisces and aves but positive correlation between ENC and CAI in mammals. Pearson correlation coefficient was -0.605 ($p>0.05$) in pisces, -0.117 ($p>0.05$) in aves and 0.216 ($p>0.05$) in mammals and these indicate that codon usage bias has a very weak relationship with nucleotide composition.

In ND5 gene, negative correlation was observed between ENC and CAI in pisces while positive correlation was found between ENC and CAI in aves and mammals. Pearson correlation coefficient was -0.452 ($p>0.05$) in pisces, 0.396 ($p>0.05$) in aves and 0.089 ($p>0.05$) in mammals and these suggest codon usage bias has a very weak relationship with nucleotide composition.

In ND6 gene, positive correlation was found between ENC and CAI in pisces, aves and mammals. Pearson correlation coefficient was 0.164 ($p>0.05$) in pisces, 0.307 ($p>0.05$) in aves and 0.067 ($p>0.05$) in mammals and these suggest codon usage bias has a very weak relationship with nucleotide composition.

4.1.5 Compositional features of MT-ND gene among pisces, aves and mammals

The non uniform usage of codons *i.e.* some codons are more preferred than others (codon usage bias) can be influenced by compositional constraints of the genome**s** (Jenkins and Holmes 2003). We therefore analysed compositional features of different species of pisces, aves and mammals for MT-ND gene.

In ND1 gene, mean±SD of nucleotide composition and its composition at 3rd position were shown in **Table 4.1**.1. **From the Figure 4.1.22**, the nucleobase C was the highest in pisces and aves but nucleobase A in mammals, whereas G was the lowest in pisces, aves and mammals respectively. This suggests unequal distribution of A, T, G and C bases and codons ending with C/A are more preferred over codons ending with G. However, the analysis of nucleotide composition at the 3rd position of codons provides a lucid depiction that could influence the preference of codon usage in different species of pisces, aves and mammals. The nucleobase A/C at the 3rd codon position was the highest in pisces and mammals but in aves C/A was found to be the highest, whereas G was the lowest in pisces, aves and mammals.

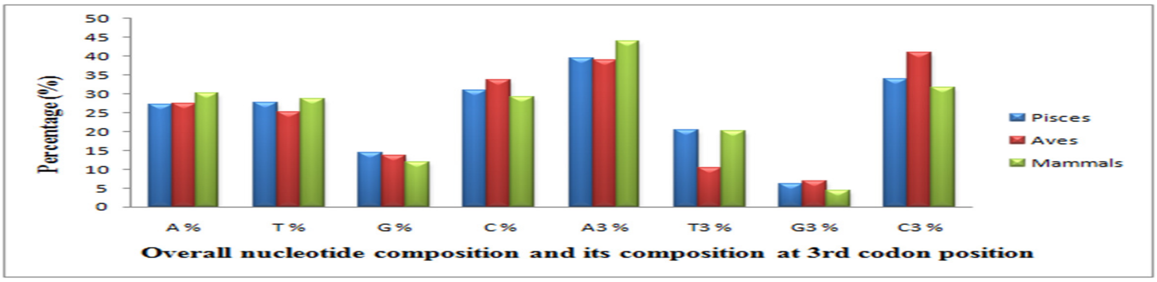


Figure 4.1.22 Nucleotide composition and its composition at 3rd codon position for MT-ND1 gene

The overall GC% was the highest in aves followed by pisces, mammals respectively and GC content is lower than 50% which indicates that MT-ND1 gene is AT rich. Further, the greatest difference of GC content was found between 1st and 2nd codon position in aves while between 1st and 3rd codon position in pisces and mammals, where most neutral mutations occur (Zhang Wen-Juan et al. 2007) (**Figure 4.1.23**).

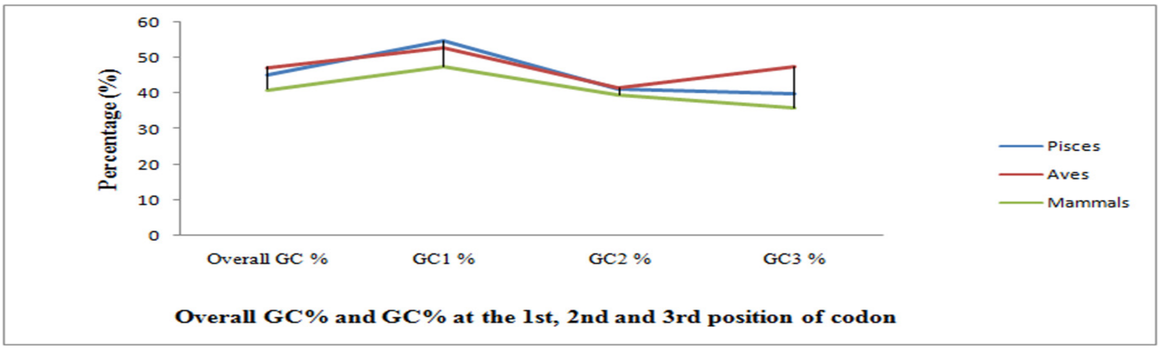


Figure 4.1.23 Overall GC content and GC content at 1st, 2nd and 3rd position in pisces, aves and mammals for MT-ND1 gene

Table 4.1.1 Nucleotide compositions among pisces, aves and mammals for ND1 gene

Species	**A %**	**T %**	**G %**	**C %**	**A3 %**	**T3 %**	**G3 %**	**C3 %**
T. chatareus	26.97	25.64	14.05	33.33	38.15	16.92	4.3	40.61
E. zonatum	26.25	32.2	15.69	25.84	37.84	28.61	9.53	24
J. floridae	23.28	30.05	15.28	31.38	31.07	27.07	7.38	34.46
P. orbicularis	26.46	24.82	13.12	35.58	37.23	14.15	3.69	44.92
L. menadoensis	32.61	25.1	13.58	28.7	53.39	15.74	5.55	25.3
Mean±SD	27.11±3.4	27.56±3.4	14.34±1.1	30.96±3.8	39.53±8.2	20.49±6.8	6.09±2.4	33.85±9.2
G.gallus	27.17	25.33	12.71	34.46	37.23	13.53	6.15	43.07
A.americana	25.25	24.74	15.74	34.25	33.74	10.73	11.96	43.55
V.chalybeata	30.16	24.13	15.03	30.67	47.23	10.12	8.89	33.74
F.peregrinus	27.48	26.15	12.41	33.94	39.38	0.16	4	40.61
S.sharpei	26.78	25.97	12.26	34.96	36.5	17.17	2.76	43.55
Mean±SD	27.36±1.8	25.26±0.8	13.63±1.6	33.65±1.7	38.81±5.1	10.34±6.3	6.75±3.7	40.90±4.2
C.familiaris	29.67	28.94	11.91	29.46	42.63	19.12	4.07	34.16
M.glis	29.66	32.18	12.05	26.1	40.56	30.5	5.03	23.89
R..norvegicus	30.82	26.64	11.18	31.34	44.51	14.73	2.82	37.93
D.novemcinctus	32.28	26.64	12.33	28.73	51.09	14.42	4.38	30.09
O.cuniculus	28.84	28.84	12.22	30.09	40.75	21.31	5.32	32.6
Mean±SD	30.25±1.3	28.64±2.3	11.93±0.5	29.14±1.9	43.9±4.3	20.01±6.5	4.32±0.9	31.73±5.2

In ND2 gene, mean±SD of nucleotide composition and its composition at 3rd position were shown in **Table 4.1.2**. From the **Figure 4.1.24**, the nucleobase C/A was the highest in pisces and aves but nucleobase A/C was the highest in mammals, whereas G was the lowest in pisces, aves and mammals respectively. This suggests unequal distribution of A, T, G and C bases and codons ending with C/A are more preferred over codons ending with G. The nucleobase A/C at the 3rd codon position was the highest in pisces, aves and mammals, whereas G was the lowest in pisces, aves and mammals.

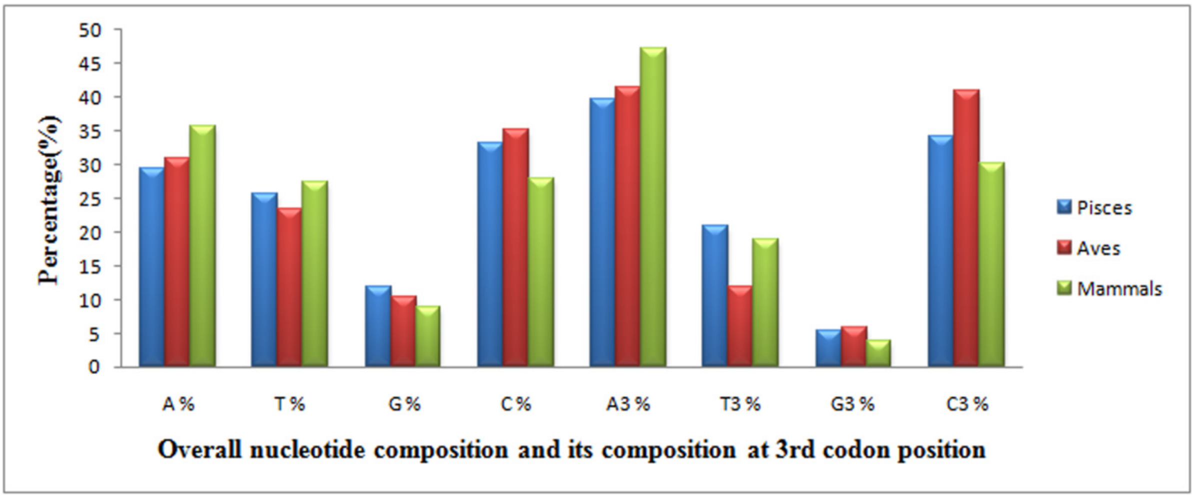


Figure 4.1.24 Nucleotide composition and its composition at 3rd codon position for MT-ND2 gene

The overall GC content was highest in aves followed by pisces and mammals for ND2 gene. The GC content was lower than AT content *i.e.* gene is AT rich. Further, the greatest difference of GC content was found between 1st and 2nd codon position in pisces while in aves and mammals, the greatest difference of GC content was found between 2nd position and 3rd position (**Figure 4.1.25**).

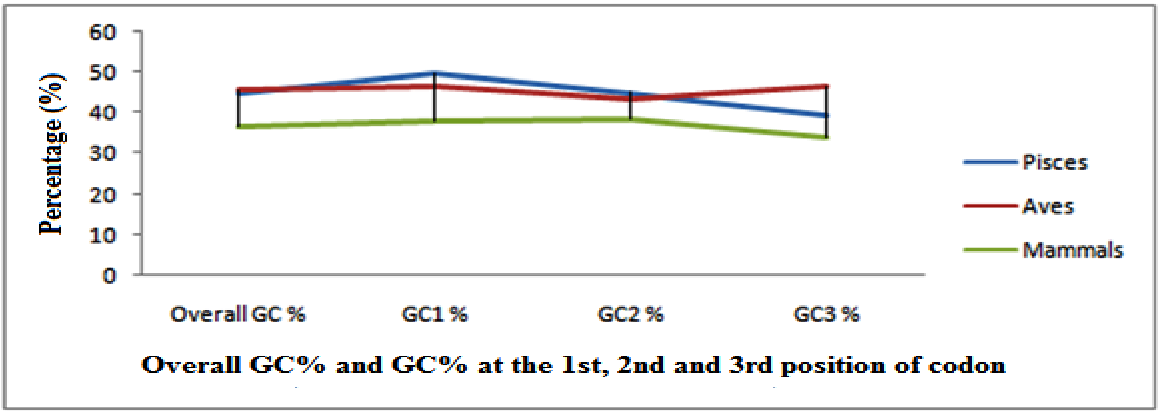


Figure 4.1.25 Overall GC content and GC content at 1st, 2nd and 3rd position in pisces, aves and mammals for MT-ND2 gene

Table 4.1.2 Nucleotide compositions among pisces, aves and mammals for ND2 gene

Species	**A %**	**T %**	**G %**	**C %**	**A3 %**	**T3 %**	**G3 %**	**C3 %**
T. chatareus	29.2	24.9	11.4	34.5	41	19.2	3.7	36.1
E. zonatum	28.7	28.7	13.8	28.8	37.8	27.8	8.9	25.5
J. floridae	25.2	28.9	11.6	34.3	29.8	28.1	4.9	37.2
P. orbicularis	26.3	22.9	11.5	39.3	33.8	15.2	4.3	46.7
L. menadoensis	38.2	22.6	10.8	28.4	55.9	13.5	5.4	25.2
Mean±SD	29.52±5.1	25.60±3.0	11.82±1.1	33.06±4.5	39.66±10	20.76±6.8	5.44±2.0	34.14±9.0
G.gallus	32.6	23	8.6	35.8	42.4	8.9	4.3	44.4
A.americana	29.1	21.8	13.1	36	38	9.2	10.7	42.1
V.chalybeata	31	23	11.9	34.1	45.5	11	6.1	37.4
F.peregrinus	32.2	25.1	9.4	33.3	43.5	14.1	4.6	37.8
S.sharpei	29.7	24.5	8.7	37.1	37.8	15.6	3.2	43.4
Mean±SD	30.92±1.5	23.48±1.3	10.34±2.0	35.26±1.5	41.44±3.4	11.76±2.9	5.78±2.9	41.02±3.2
C.familiaris	35.4	28	9.2	27.4	46	18.1	5.2	30.7
M.glis	33.2	30.1	9.3	27.4	42.7	23.1	3.5	30.7
R..norvegicus	36	25	7.9	31.1	47.1	15.9	2	35
D.novemcinctus	39.1	26.1	8.3	26.5	45.7	22.7	4	27.6
O.cuniculus	35	28.3	9.1	27.6	54.6	14.4	4.6	26.4
Mean±SD	35.74±2.1	27.50±±1.9	8.76±0.6	28±1.7	47.22±4.4	18.84±3.9	3.86±1.2	30.08±3.3

In ND3 gene, mean±SD of nucleotide composition and its composition at 3rd position were shown in **Table 4.1.3**. From the **Figure 4.1.26**, the nucleobase C/T was the highest in pisces but in aves, C/A was the highest while in mammals, the nucleobase A/T was the highest. The base G was the lowest in pisces, aves and mammals respectively. However, the analysis of nucleotide composition at the 3rd position of codons revealed that the nucleobase A/C at the 3rd codon position was the highest in pisces, aves and mammals, whereas G was the lowest in pisces, aves and mammals.

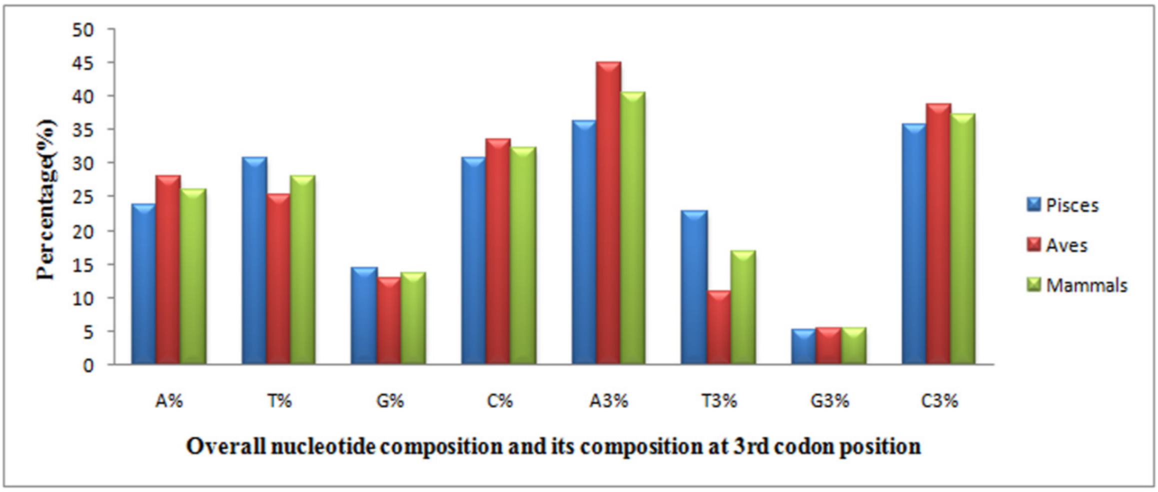


Figure 4.1.26 Nucleotide composition and its composition at 3rd codon position for MT-ND3 gene

The overall GC content was the highest in aves followed by pisces and mammals. The GC% was lower than 50% *i.e.* the ND3 gene is AT rich. Further, the greatest difference of GC content was found between 1st and 2nd codon position in pisces and aves but between 1st and 3rd codon position in mammals (**Figure 4.1.27**).

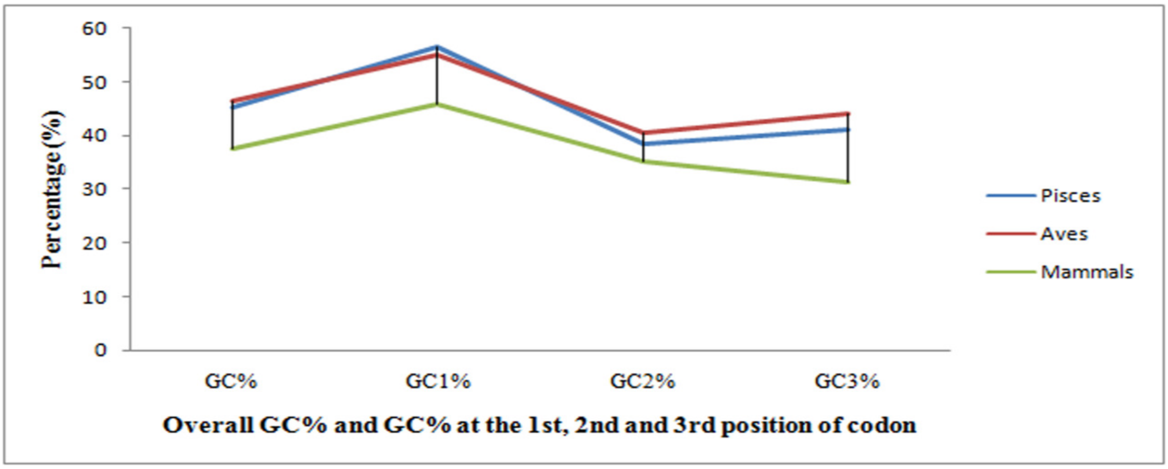


Figure 4.1.27 Overall GC content and GC content at 1st, 2nd and 3rd position in pisces, aves and mammals for MT-ND3 gene

Table 4.1.3 Nucleotide compositions among pisces, aves and mammals for ND3 gene

Species	**A%**	**T%**	**G%**	**C%**	**A3%**	**T3%**	**G3%**	**C3%**
T. chatareus	21.7	31.6	14.2	32.5	32.5	22.2	5.1	40.2
E. zonatum	23.9	34.2	15.1	26.8	35	27.4	7.7	29.9
J. floridae	21.4	32.8	14	31.8	31.6	29.1	5.1	34.2
P. orbicularis	21.7	28.5	14.2	35.6	31.6	17.9	2.6	47.9
L. menadoensis	30.5	27.4	14.5	27.6	50.4	17.1	6	26.5
Mean±SD	23.84±3.8	30.9±2.87	14.4±0.43	30.86±3.64	36.22±8.04	22.74±5.42	5.3±1.84	35.74±8.50
G.gallus	27.9	27.1	13.1	31.9	43.6	17.1	4.3	35
A.americana	25.1	23.9	16	35	41.9	7.7	11.1	39.3
V.chalybeata	31.6	24.8	12.5	31.1	53.8	8.5	3.4	34.3
F.peregrinus	29.6	24.5	11.1	34.8	47.9	7.7	3.4	41
S.sharpei	26.2	26.5	12	35.3	37.6	13.7	5.1	43.6
Mean±SD	28.08±2.6	25.35±1.37	12.94±1.86	33.62±1.96	44.96±6.16	10.94±4.25	5.46±3.23	38.64±3.95
C.familiaris	29.9	28.2	14.1	27.8	40.5	20.7	10.3	28.5
M.glis	29.6	36.5	11	22.9	40.9	34.8	1.7	22.6
R..norvegicus	32.8	30.7	10.1	26.4	44	24.1	2.6	29.3
D.novemcinctus	31.3	32.8	11.5	24.4	47.4	23.3	1.7	27.6
O.cuniculus	33.3	27.2	12.2	27.3	53	13.9	5.2	27.9
Mean±SD	31.38±1.6	31.08±3.73	11.78±1.50	25.76±2.05	45.16±5.18	23.36±7.54	4.3±3.64	27.18±2.64

In ND4 gene, mean±SD of nucleotide composition and its composition at 3rd position were shown in **Table 4.1.4**. From the **Figure 4.1.28**, the nucleobase C/A was the highest in pisces and aves but nucleobase A/T in mammals, whereas G was the lowest in pisces, aves and mammals respectively. However, the analysis of nucleotide composition at the 3rd position of codons suggests that the nucleobase A/C at the 3rd codon position was the highest in pisces and mammals but in aves C/A was found to be the highest, whereas G was the lowest in pisces, aves and mammals.

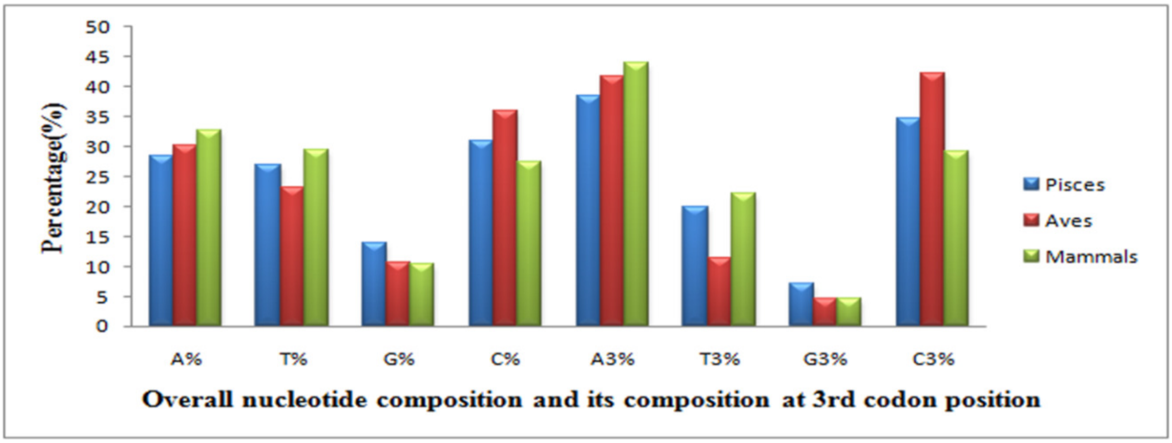


Figure 4.1.28 Overall nucleotide composition and its composition at 3rd codon position for MT-ND4 gene

The overall GC content was the highest in aves followed by pisces and mammals. The overall GC content was lower than 50% *i.e.* the ND4 gene is AT rich**.** Further, the greatest difference of GC content was found between 1st and 3rd codon position in pisces and mammals, while between 1st and 2nd codon position in aves (**Figure 4.1.29**).

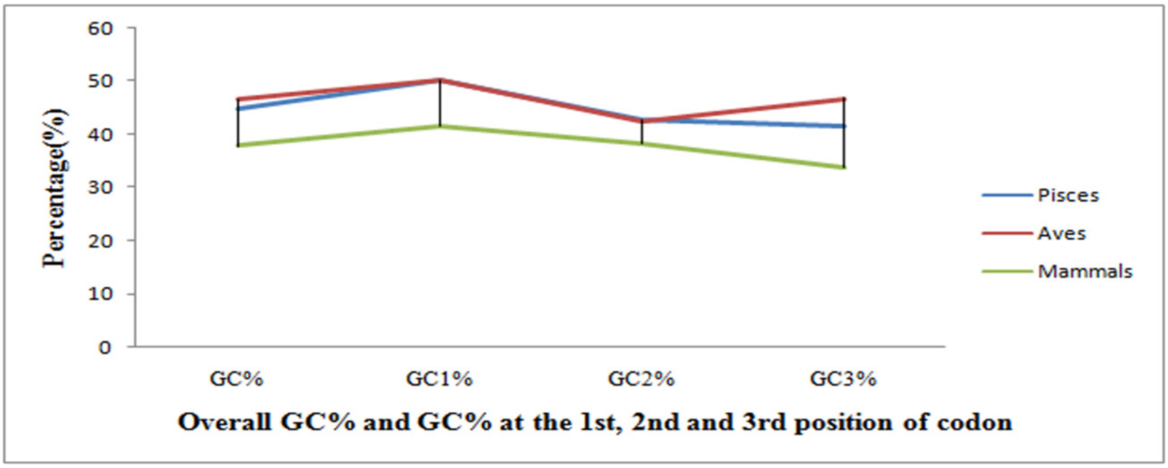


Figure 4.1.29 Overall GC content and GC content at 1st, 2nd and 3rd position in pisces, aves and mammals for MT-ND4 gene

Table 4.1.4 Nucleotide compositions among pisces, aves and mammals for ND4 gene

Species	**A%**	**T%**	**G%**	**C%**	**A3%**	**T3%**	**G3%**	**C3%**
T. chatareus	27.2	26.2	13.6	33	36.2	19.7	4.8	39.3
E. zonatum	27.8	30.2	15.2	26.8	37.1	27.3	10.2	25.4
J. floridae	24.9	29.1	14.8	31.2	30.8	24.3	9.5	35.4
P. orbicularis	26.9	24.1	13.8	35.2	36.4	13.2	5.6	44.8
L. menadoensis	34.7	24.8	12.1	28.4	51.8	15	4.8	28.4
Mean±SD	28.3±3.7	26.88±2.6	13.9±1.2	30.92±3.4	38.46±7.9	19.9±5.98	6.98±2.7	34.66±7.9
G.gallus	30.1	23.7	10.1	36.1	40.7	11.3	3.3	44.7
A.americana	29.2	23	12.1	35.7	40.2	9.6	8.3	41.9
V.chalybeata	32.2	21.7	10.9	35.2	46.5	8.5	4.6	40.4
F.peregrinus	31	22.8	9.9	36.3	43.9	12	3.7	40.4
S.sharpei	28.8	24.9	9.8	36.5	37.6	15.7	3	43.7
Mean±SD	30.26±1.4	23.22±1.2	10.56±0.9	35.96±0.5	41.78±3.5	11.42±2.8	4.58±2.2	42.22±1.9
C.familiaris	31.3	30.2	10.9	27.6	40.4	23.9	6.5	29.2
M.glis	32.4	33.1	9.7	24.8	44.4	29.2	2.6	23.8
R..norvegicus	34.6	27.6	9.9	27.9	47.2	16.5	3.5	32.8
D.novemcinctus	30.2	29.3	11.4	29.1	38.5	23.3	6.3	31.9
O.cuniculus	35.1	27	9.9	28	49.7	17.6	4.1	28.6
Mean±SD	32.72±2.1	29.44±2.4	10.36±0.7	27.48±1.6	44.04±4.6	22.1±5.2	4.6±1.7	29.26±3.5

In ND4l gene, mean±SD of nucleotide composition and its composition at 3rd position were shown in **Table 4.1.5.** From the **Figure 4.1.30**, the base C/T was the highest in pisces while nucleobase C/A was the highest in aves but T/A was the highest in mammals. The nucleobase G was the lowest in pisces, aves and mammals respectively. However, the analysis of nucleotide composition at the 3rd position of codons showed that the nucleobase A/C at the 3rd codon position was the highest in pisces and mammals but in aves C/A was found to be the highest, whereas G was the lowest in pisces, aves and mammals.

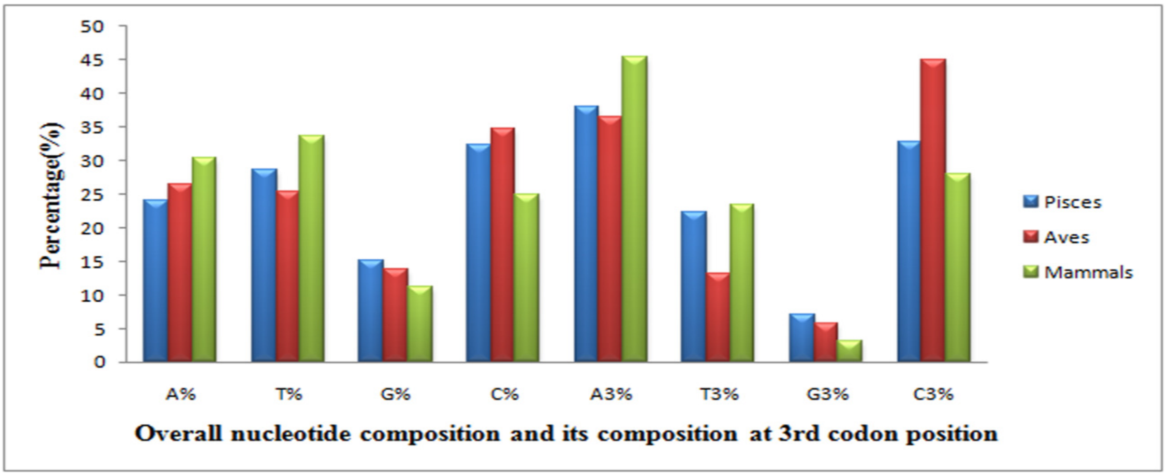


Figure 4.1.30 Overall nucleotide composition and its composition at 3rd codon position for MT-ND4L gene

The overall GC content was the highest in aves followed by pisces and mammals. The overall GC content was lower than 50% *i.e.* the ND4l gene is AT rich. Further, the greatest difference of GC content was found between 1st and 3rd codon position in pisces and mammals but between 1st and 2nd codon position in aves (**Figure 4.1.31).**

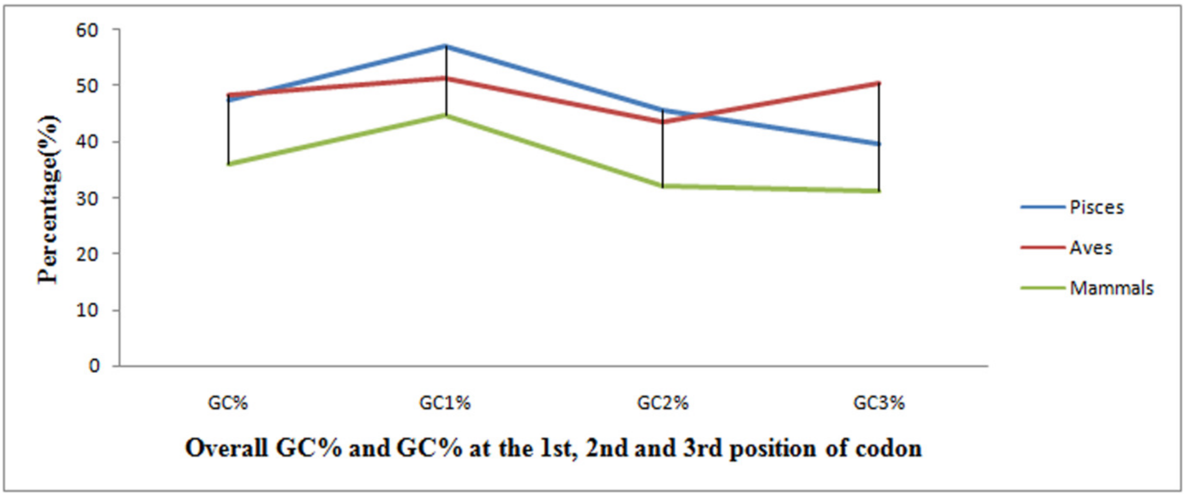


Figure 4.1.31 Overall GC content and GC content at 1st, 2nd and 3rd position in pisces, aves and mammals for MT-ND4L gene

Table 4.1.5 Nucleotide compositions among pisces, aves and mammals for ND4L gene

Species	**A%**	**T%**	**G%**	**C%**	**A3%**	**T3%**	**G3%**	**C3%**
T. chatareus	22.6	28.6	14.8	34	34.3	22.2	6.1	37.4
E. zonatum	21.2	32.7	15.8	30.3	35.4	32.3	9.1	23.2
J. floridae	22.6	31.6	15.5	30.3	34.3	30.3	7.1	28.3
P. orbicularis	22.6	25.3	13.8	38.3	32.3	13.1	6.1	48.5
L. menadoensis	31	24.6	15.5	28.9	53.5	13.1	7.1	26.3
Mean±SD	24±3.95	28.56±3.62	15.08±0.80	32.36±3.82	37.96±8.75	22.2±9.12	7.1±1.22	32.74±10.27
G.gallus	27.9	24.9	12.5	34.7	44.4	12.1	2	41.5
A.americana	23.9	23.6	18.2	34.3	29.3	10.1	17.2	43.4
V.chalybeata	26.6	24.6	12.8	36	36.4	9.1	5.1	49.4
F.peregrinus	27.9	26.3	12.5	33.3	39.4	18.2	2	40.4
S.sharpei	25.3	26.9	12.5	35.3	32.3	16.2	2	49.5
Mean±SD	26.32±1.72	25.26±1.33	13.7±2.51	34.72±1.02	36.36±5.92	13.14±3.92	5.66±6.58	44.84±4.34
C.familiaris	31.3	34.3	12.5	21.9	47.5	24.2	6.1	22.2
M.glis	26.9	38.4	10.4	24.3	39.4	27.3	2	31.3
R..norvegicus	32	32	9.4	26.6	46.5	23.2	1	29.3
D.novemcinctus	29.6	34.3	10.4	25.7	46.5	27.3	1	25.2
O.cuniculus	32	29.6	12.8	25.6	47.5	15.2	5.1	32.2
Mean±SD	30.36±2.16	33.72±3.25	11.1±1.47	24.82±1.82	45.48±3.43	23.44±4.95	3.04±2.39	28.04±4.23

In ND5 gene, mean±SD of nucleotide composition and its composition at 3rd position were shown in **Table 4.1.6.** From the **Figure 4.1.32**, the base A/C was the highest in pisces while nucleobase C/A was the highest in aves but A/T was the highest in mammals. The nucleobase G was the lowest in pisces, aves and mammals respectively. However, the analysis of nucleotide composition at the 3rd position of codons revealed that the nucleobase A/C at the 3rd codon position was the highest in pisces and mammals but in aves C/A was found to be the highest, whereas G was the lowest in pisces, aves and mammals.

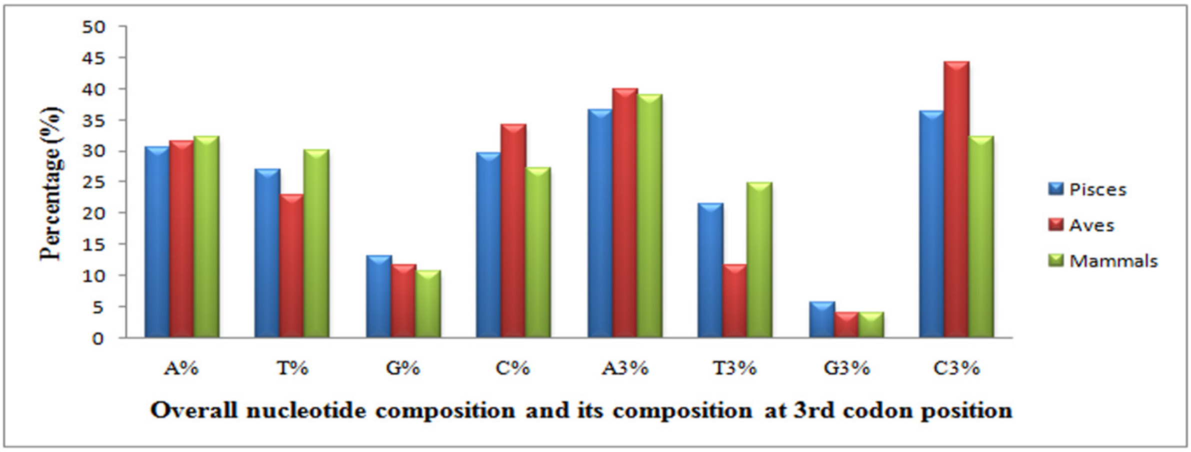


Figure 4.1.32 Overall nucleotide composition and its composition at 3rd codon position for MT-ND5 gene

The overall GC content was the highest in aves followed by pisces and mammals. The overall GC content was lower than 50% *i.e.* the ND5 gene is AT rich. Further, the greatest difference of GC content was found between 1st and 2nd codon position in pisces while in aves greatest difference of GC content was found between 2nd and 3rd position but in mammals between 1st and 3rd codon position (**Figure 4.1.33**).

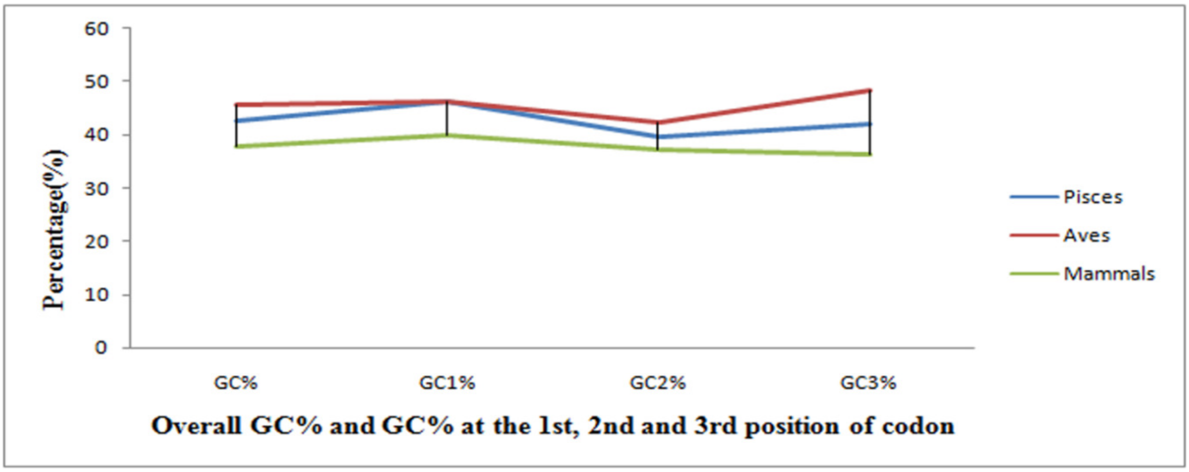


Figure 4.1.33 Overall GC content and GC content at 1st, 2nd and 3rd position in pisces, aves and mammals for MT-ND5 gene

Table 4.1.6 Nucleotide compositions among pisces, aves and mammals for ND5 gene

Species	A%	T%	G%	C%	A3%	T3%	G3%	C3%
T. chatareus	28.8	25.8	12.5	32.9	33.1	19.1	4.1	43.7
E. zonatum	28.7	30.1	15.2	26	34.6	27.1	10.4	27.9
J. floridae	27.4	30.8	13.3	28.5	31	30.8	5.9	32.3
P. orbicularis	30.1	24.7	12.5	32.7	36.1	16.5	3.9	43.5
L. menadoensis	37.3	23.6	11.4	27.7	48	14.2	4.2	33.6
Mean±SD	30.46±3.9	27±3.3	12.98±1.4	29.56±3.0	36.56±6.7	21.54±7.1	5.7±2.7	36.2±7.07
G.gallus	31	23.4	10.6	35	38.3	14.2	2.5	45
A.americana	29.8	21.7	13.7	34.8	37.8	8.9	7.4	45.9
V.chalybeata	31.7	22.6	12.4	33.3	43.4	9.6	4.1	42.9
F.peregrinus	33.1	21.8	11.3	33.8	42.6	10.4	3.6	43.4
S.sharpei	31.4	24.9	10.4	33.3	37.4	15.6	2.8	44.2
Mean±SD	31.4±1.19	22.88±1.32	11.68±1.37	34.04±0.81	39.9±2.86	11.74±2.97	4.08±1.96	44.28±1.20
C.familiaris	30.7	29.7	11.9	27.7	34.1	25.4	6.4	34.1
M.glis	31.4	33.6	10.2	24.8	37.9	32	2.8	27.3
R..norvegicus	34.5	28.2	9.2	28.1	42.7	20	2.3	35
D.novemcinctus	30.8	30.6	11.5	27.1	35.4	27.3	4.6	32.7
O.cuniculus	33.7	28	10.3	28	44.7	19.1	3.8	32.4
Mean±SD	32.22±1.7	30.02±2.3	10.62±1.08	27.14±1.36	38.96±4.59	24.76±5.33	3.98±1.61	32.3±2.98

In ND6 gene, mean±SD of nucleotide composition and its composition at 3rd position were shown in **Table 4.1.7.** From the **Figure 4.1.34**, the base A/C was the highest in pisces, aves and mammals while G was the lowest in pisces, aves and mammals respectively. However, the analysis of nucleotide composition at the 3rd position of codons indicated that the nucleobase C/A at the 3rd codon position was the highest in pisces, aves and mammals whereas T was the lowest in pisces, aves but nucleobase G was the lowest in mammals.

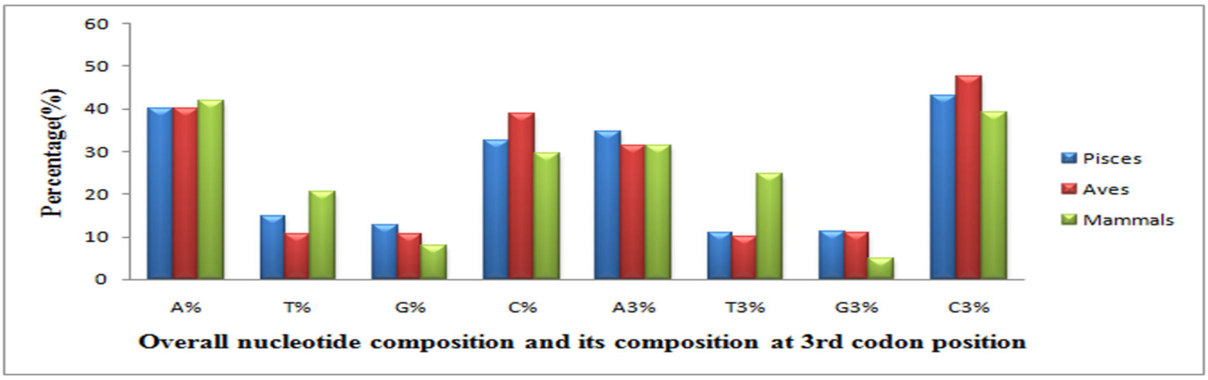


Figure 4.1.34 Overall nucleotide composition and its composition at 3rd codon position for MT-ND6 gene

The overall GC content was the highest in aves followed by pisces and mammals. The overall GC content was lower than 50% *i.e.* the ND6 gene is AT rich. Further, the greatest difference of GC content was found between 1st and 3rd codon position in pisces and mammals while in aves the greatest difference of GC content was found between the 2nd and 3rd position (**Figure 4.1.35**).

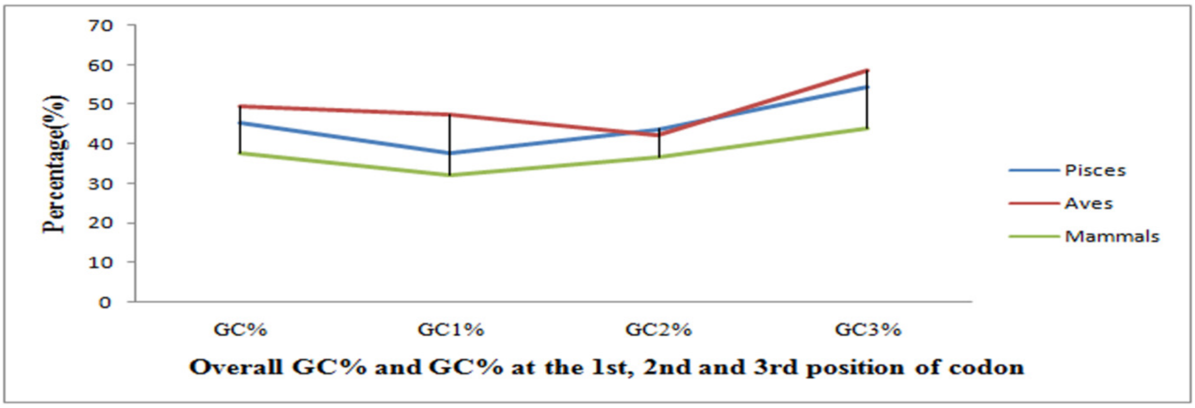


Figure 4.1.35 Overall GC content and GC content at 1st, 2nd and 3rd position in pisces, aves and mammals for MT-ND6 gene

Table 4.1.7 Nucleotide compositions among pisces, aves and mammals for ND6 gene

Species	A%	T%	G%	C%	A3%	T3%	G3%	C3%
T. chatareus	41	16.3	11.9	30.8	35.1	13.2	11.5	40.2
E. zonatum	39.8	16.1	15.3	28.8	31.6	13.2	13.8	41.4
J. floridae	33.9	16.9	13.8	35.4	28.7	11.5	12.6	47.2
P. orbicularis	38.5	12.5	12.5	36.5	35.1	9.2	9.8	45.9
L. menadoensis	46.9	11.7	10.7	30.7	43.1	8	8	40.9
Mean±SD	40.02±4.6	14.7±2.4	12.84±1.8	32.44±3.3	34.72±5.4	11.02±2.4	11.14±2.3	43.12±3.2
G.gallus	41.6	10.2	9.4	38.8	34.5	8	9.2	48.3
A.americana	39.7	9.8	11.9	38.6	30.5	8.6	11.5	49.4
V.chalybeata	36.4	11.9	14.1	37.6	23.1	11	16.8	49.1
F.peregrinus	42.6	10.6	9.1	37.7	34.1	12.5	8	45.4
S.sharpei	39.5	10.9	8	41.6	35.1	9.8	9.2	45.9
Mean±SD	39.96±2.4	10.68±0.8	10.5±2.5	38.86±1.7	31.46±5.0	9.98±1.8	10.94±3.5	47.62±1.8
C.familiaris	40	21.8	10.2	28	28.4	27.3	5.7	38.6
M.glis	42.7	22.7	7.6	27	32.6	24.6	6.9	35.9
R..norvegicus	43.4	20.2	6.9	29.5	34.7	26	4.6	34.7
D.novemcinctus	40.4	19.8	6.5	33.3	28.6	25.1	3.4	42.9
O.cuniculus	42.6	18.8	8.1	30.5	31.8	21	3.4	43.8
Mean±SD	41.82±1.5	20.66±1.6	7.86±1.4	29.66±2.4	31.22±2.7	24.8±2.3	4.8±1.5	39.18±4.1

4.1.6 Interrelationships among different compositional features of MT-ND gene among pisces, aves and mammals

Two major evolutionary forces namely mutation pressure and natural selection are considered to shape the codon usage pattern (TATARINOVA *et al.* 2010). The mutational pressure, which affects the whole mitochondrial genome, accounts for codon usage bias in ribbon worm (Chen et al. 2014). Therefore to determine the extent of the influence of mutation pressure in codon usage pattern among pisces, aves and mammals for MT-ND gene, we performed correlation analysis between general nucleotide composition and nucleotide composition at 3rd codon position to determine whether evolutionary process is driven by mutation pressure alone or by both mutation pressure and natural selection.

In ND1 gene, highly significant positive correlation was found between A and A3, T and T3, T and G3, G and T3, G and G3, C and C3, C and GC3, GC and C3, GC and GC3 in pisces. In aves, significant positive correlation was found between A and A3, G and G3, C and C3, GC and GC3 while significant negative correlation was found between A and C3, C and A3. In mammals, significant positive correlation was found between A and A3, T and T3, C and C3 and C and GC3, GC and C3, GC and GC3 as shown in **Table 4.1.8.** These results suggest that the compositional constraint arising from mutation pressure and natural selection determines the pattern of codon usage in ND1 gene.

Table 4.1.8 Correlation between overall nucleotide composition (%) and its composition at 3rd codon position in pisces, aves and mammals in ND1 gene

Species	Nucleotide	A3%	T3%	G3%	C3%	GC3%
Pisces	A %	**0.995**	-0.582	-0.280	-0.392	-0.564
	T %	-0.487	**0.987**	**0.954**	-0.538	-0.359
	G %	-0.482	**0.986**	**0.924**	-0.535	-0.364
	C %	-0.317	-0.633	-0.855	**0.975**	**0.921**
	GC %	-0.542	-0.417	-0.703	**0.978**	**0.972**
Aves	A %	**0.987**	-0.160	-0.109	**-0.927**	-0.856
	T %	-0.470	-0.184	-0.831	0.601	-0.115
	G %	0.150	0.016	**0.962**	-0.356	0.417
	C %	**-0.922**	0.218	-0.383	**0.980**	0.535
	GC %	-0.829	0.254	0.550	0.695	**0.990**
Mammals	A %	**0.958**	-0.625	-0.464	0.079	-0.003
	T %	-0.724	**0.988**	0.621	-0.757	-0.736
	G %	0.144	0.284	0.863	-0.638	-0.547
	C %	0.152	-0.787	-0.604	**0.974**	**0.984**
	GC %	0.214	-0.808	-0.450	**0.920**	**0.956**

Note: One asterisk (*) and two asterisks (**) indicate correlations at a level of 0.05 and 0.01, respectively

Furthermore, correlation between ENC and various GC contents as shown in **Table 4.1.9** suggests that the nucleotide composition under from mutation pressure and natural selection affect the synonymous codon usage in MT-ND1 gene for pisces, aves and mammals.

Table 4.1.9 Correlation coefficients among ENC, CAI, GC contents for MT-ND1 gene

		Pisces		**Aves**		**Mammals**	
SL No	Correlation between	Correlation coefficient	p value	Correlation coefficient	p value	Correlation coefficient	p value
1	ENC and GC %	**0.941****	**0.017**	0.568	0.318	**0.948****	**0.014**
2	ENC and GC3 %	**0.970****	**0.006**	0.562	0.324	**0.997****	**0.000**
3	GC and GC3 %	**0.972****	**0.006**	**0.990****	**0.001**	**0.956****	**0.011**
4	GC1 and GC3 %	0.672	0.214	-0.173	0.781	0.484	0.409
5	GC2 and GC3%	0.106	0.865	0.039	0.951	-0.461	0.434
6	GC12 and GC3 %	0.642	0.243	-0.205	0.741	0.269	0.662

Note: One asterisk (*) and two asterisks (**) indicate correlations at a level of 0.05 and 0.01, respectively

If the pattern of synonymous codon usage is solely governed by mutation pressure then the frequency of nucleotide A and T should be equal to that of G and C at synonymous 3rd codon position. In ND1 gene, the frequencies of those nucleotides *i.e.* A and T are not equal to G and C in pisces, aves and mammals indicating that other factor such as natural selection might have played a role in codon usage pattern.

To quantify the extent of natural selection and mutation pressure in the codon usage pattern of MT-ND1, we drew neutrality plot between GC12 and GC3. GC12 is the average of GC1 and GC2. The regression coefficient of GC12 on GC3 is 0.121 indicating relative neutrality of 12.1 % and relative constraint of 87.9 % for GC3 % in pisces as shown in **Figure 4.1.36 (a).** The GC12 was influenced by mutation pressure and natural selection with a ratio of 0.121/0.879= 0.137. In aves, the regression coefficient of GC12 on GC3 was 0.016 indicating relative neutrality 1.6% (absolute) and relative constraint 98.4% as shown in **Figure 4.1.36 (b)**. The GC12 was influenced by mutation pressure and natural selection with a ratio of 0.016/0.984= 0.016. In mammals, the regression coefficient of GC12 on GC3 was 0.046 indicating relative neutrality of 4.6 % and relative constraint of 95.4 % for GC3 % as shown in **Figure 4.1.36 (c).** The GC12 was influenced by mutation pressure and natural selection with a ratio of 0.046/0.954= 0.048. These results suggest that natural selection played a major role while mutation pressure played a minor role in shaping codon usage pattern in ND1 gene among pisces, aves and mammals.

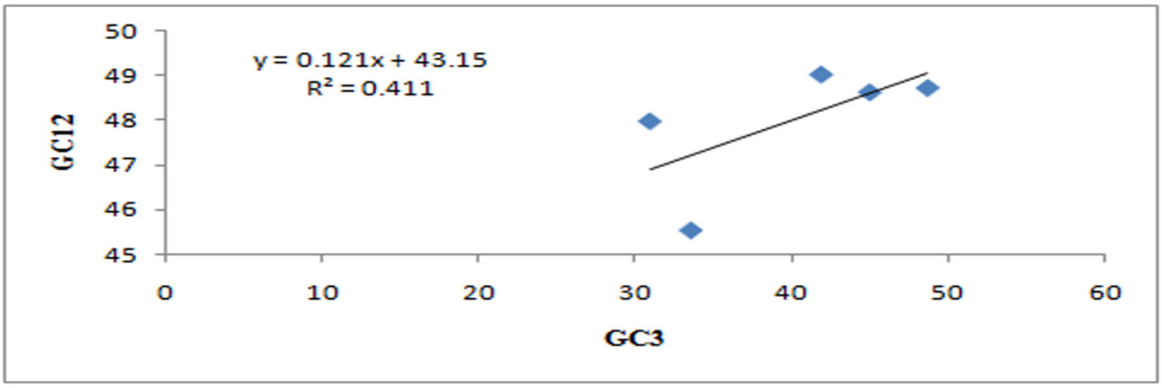


Figure 4.1.36 (a) Neutrality plot of GC12 with GC3 in different species of pisces for ND1 gene

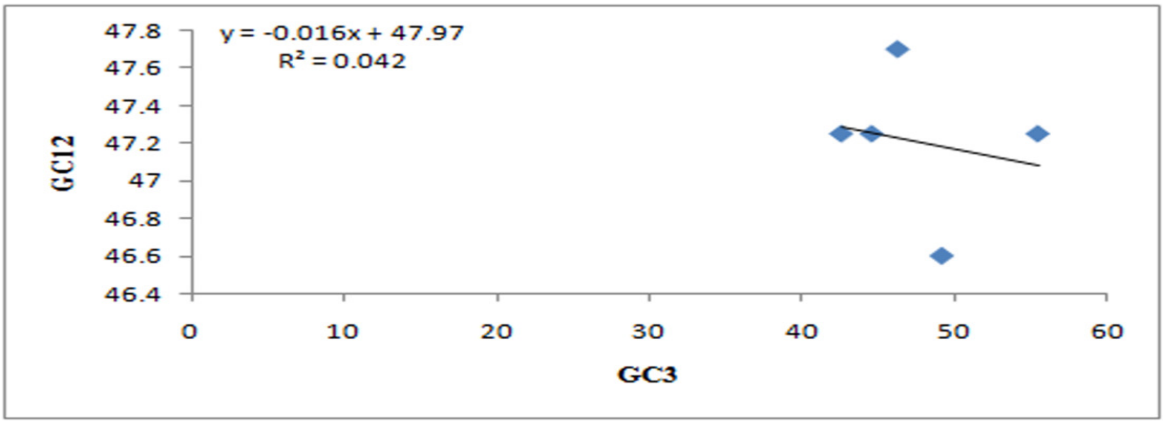


Figure 4.1.36 (b) Neutrality plot of GC12 with GC3 in different species of aves for ND1 gene

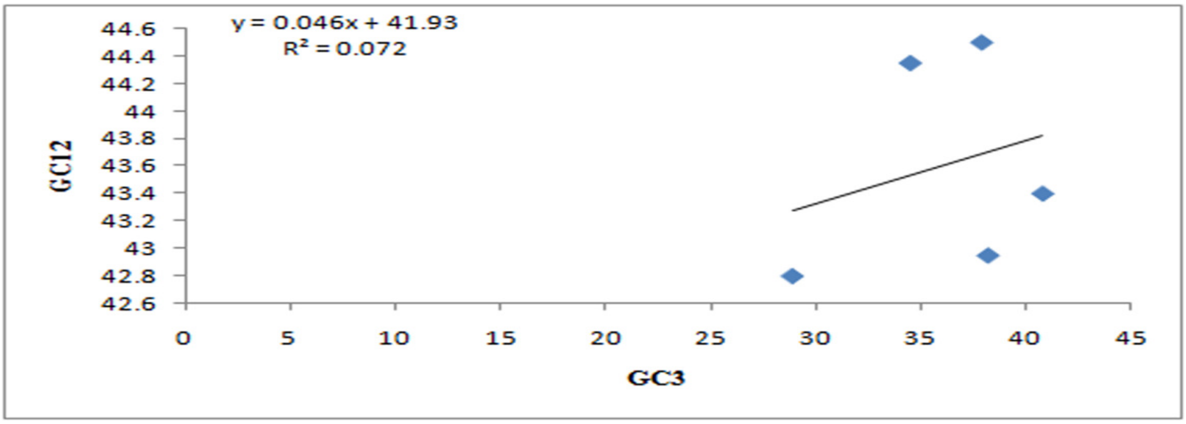


Figure 4.1.36 (c) Neutrality plot of GC12 with GC3 in different species of mammals for ND1 gene

In ND2 gene, highly significant positive correlation was found between A and A3, T and T3, C and C3, C and GC3, GC and C3, GC and GC3 in pisces. In aves, significant positive correlation was found between G and G3, C and C3. In mammals, significant positive correlation was found between GC and C3 as shown in **Table 4.1.10.** These results suggest that the compositional constraint arising from mutation pressure and natural selection determine the pattern of codon usage in ND2 gene.

Table 4.1.10 Correlation between overall nucleotide composition (%) and its composition at 3rd codon position in pisces, aves and mammals in ND2 gene

	Nucleotide	**A3**	**T3**	**G3**	**C3**	**GC3**
Pisces	A %	**0.992****	-0.582	0.081	-0.675	-0.755
	T %	-0.590	**0.999****	0.537	-0.228	-0.127
	G %	-0.364	0.700	0.865	-0.325	-0.154
	C %	-0.631	-0.192	-0.672	**0.998****	**0.977****
	GC %	-0.761	-0.014	-0.462	**0.959****	**0.986****
Aves	A %	0.734	-0.104	-0.546	-0.183	-0.528
	T %	0.166	0.875	-0.791	-0.264	-0.754
	G %	-0.029	-0.484	**0.917****	-0.358	0.359
	C %	-0.833	-0.004	0.002	**0.884***	0.695
	GC %	-0.566	-0.433	0.792	0.278	0.770
Mammals	A %	0.034	0.145	-0.010	-0.213	-0.287
	T %	-0.114	0.290	0.492	-0.370	-0.231
	G %	0.030	0.124	0.718	-0.448	-0.229
	C %	0.076	-0.542	-0.788	0.826	0.683
	GC %	-0.542	0.091	-0.798	**0.903****	0.780

Note: One asterisk (*) and two asterisks (**) indicate correlations at a level of 0.05 and 0.01, respectively

Furthermore, correlation between ENC and various GC contents as shown in **Table 4.1.11**, suggests that the nucleotide composition under from mutation pressure and natural selection affect the synonymous codon usage in MT-ND2 gene for pisces, aves and mammals.

Table 4.1.11 Correlation coefficients among ENC, CAI, GC contents for MT-ND2 gene

		Pisces		**Aves**		**Mammals**	
SL No	Correlation between	Correlation coefficient	p value	Correlation coefficient	p value	Correlation coefficient	p value
1	ENC and CAI	-0.157	0.801	-0.071	0.910	-0.648	0.237
2	ENC and GC %	**0.951****	**0.013**	0.695	0.193	0.870	0.055
3	ENC and GC3 %	**0.941****	**0.017**	0.594	0.291	**0.984****	**0.002**
4	GC and GC3 %	**0.986****	**0.002**	0.770	0.128	0.780	0.120
5	GC1 and GC3 %	**0.945****	**0.015**	0.319	0.601	0.180	0.773
6	GC2 and GC3%	0.631	0.254	-0.530	0.358	0.517	0.372
7	GC12 and GC3 %	**0.914****	**0.030**	0.264	0.668	0.195	0.753

Note: One asterisk (*) and two asterisks (**) indicate correlations at a level of 0.05 and 0.01, respectively

In ND2 gene, the frequencies of those nucleotides *i.e.* A and T are not equal to G and C in pisces, aves and mammals indicating that other factor such as natural selection might have played a role in codon usage pattern.

To quantify the extent of natural selection and mutation pressure in the codon usage pattern of MT-ND2, we drew the neutrality plot between GC12 and GC3. GC12 is the average of GC1 and GC2. The regression coefficient of GC12 on GC3 in pisces is 0.318 indicating relative neutrality of 31.8 % and relative constraint of 68.2 % for GC3 % in pisces as shown in **Figure 4.1.37 (a).** The GC12 was influenced by mutation pressure and natural selection with a ratio of 0.318/0.682= 0.466. In aves, the regression coefficient of GC12 on GC3 was 0.144 indicating relative neutrality 14.4% and relative constraint 85.6% as shown in **Figure 4.1.37 (b)**. The GC12 was influenced by mutation pressure and natural selection with a ratio of 0.144/0.856= 0.168. In mammals, the regression coefficient of GC12 on GC3 was 0.182 indicating relative neutrality of 18.2 % and relative constraint of 81.8 % for GC3 % as shown in **Figure 4.1.37 (c).** The GC12 was influenced by mutation pressure and natural selection with a ratio of 0.182/0.818= 0222. These results suggest that natural selection played a major role while mutation pressure played a minor role in shaping codon usage pattern in ND2 gene among pisces, aves and mammals.

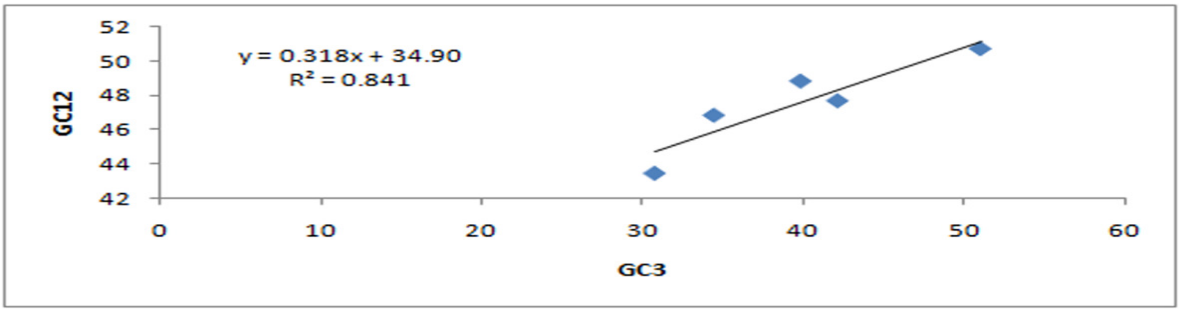


Figure 4.1.37 (a) Neutrality plot of GC12 with GC3 in different species of pisces for ND2 gene

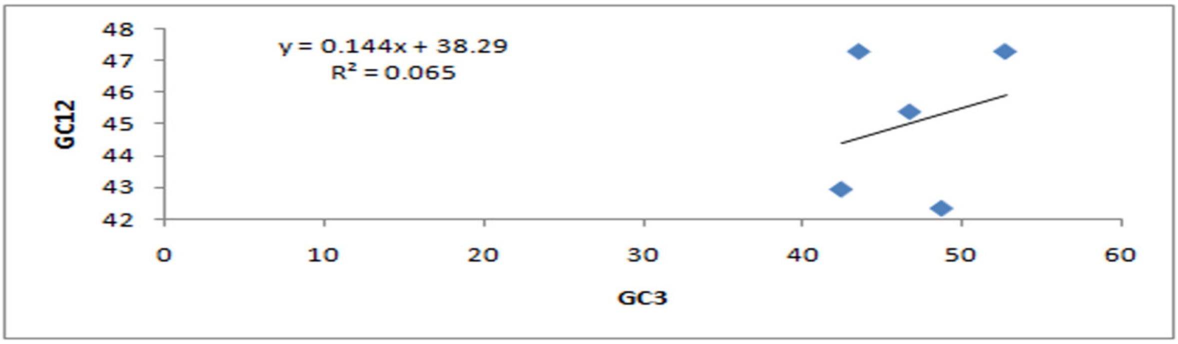


Figure 4.1.37 (b) Neutrality plot of GC12 with GC3 in different species of aves for ND2 gene

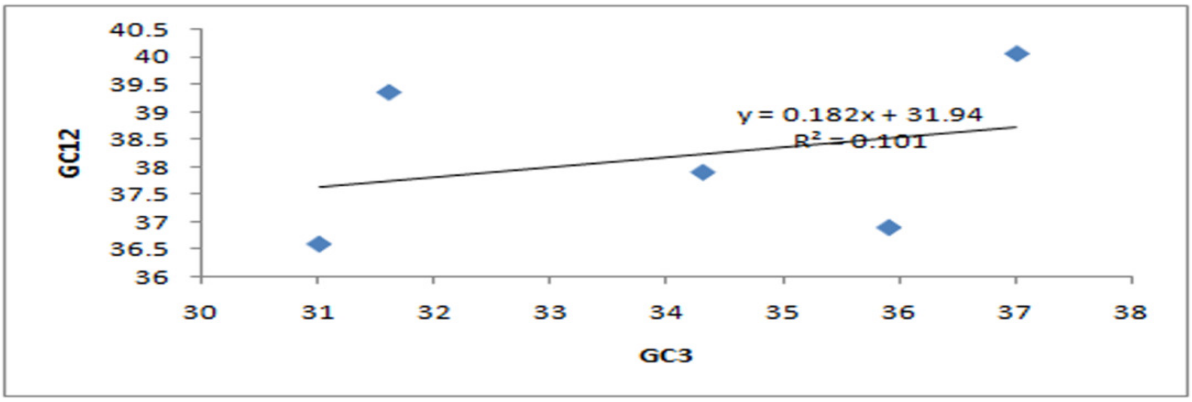


Figure 4.1.37 (c) Neutrality plot of GC12 with GC3 in different species of mammals for ND2 gene

In ND3 gene, highly significant positive correlation was found between A and A3, T and T3, C and C3, C and GC3, GC and C3, GC and GC3 in pisces while negative correlation was found between C and G3, GC and G3. In aves, significant positive correlation was found between A and A3, T and T3, G and G3, C and C3, C and GC3, GC and G3, GC and GC3. In mammals, significant positive correlation was found between T and T3, G and G3, C and GC3, GC and G3, GC and GC3 as shown in **Table 4.1.12.** These results suggest that the compositional constraint arising from mutation pressure and natural selection determine the pattern of codon usage in ND3 gene.

Table 4.1.12 Correlation between overall nucleotide composition (%) and its composition at 3rd codon position in pisces, aves and mammals in ND3 gene

	Nucleotide	**A3 %**	**T3 %**	**G3 %**	**C3 %**	**GC3%**
Pisces	A %	**0.995****	-0.494	0.404	-0.714	-0.764
	T %	-0.583	**0.938****	0.539	-0.163	-0.053
	G %	0.295	0.132	0.759	-0.527	-0.442
	C %	-0.627	-0.232	**-0.942****	**0.946****	**0.902****
	GC %	-0.642	-0.242	**-0.939****	**0.965**	**0.927****
Aves	A %	**0.915****	-0.229	-0.781	-0.541	-0.837
	T %	-0.407	**0.975****	-0.437	-0.057	-0.290
	G %	-0.261	-0.142	**0.919****	-0.192	0.394
	C %	-0.682	-0.242	0.470	**0.940****	**0.938****
	GC %	-0.646	-0.262	**0.933****	0.527	**0.913****
Mammals	A %	0.803	-0.668	-0.195	0.601	0.160
	T %	-0.462	**0.939****	-0.710	-0.798	**-0.881***
	G %	-0.092	-0.446	**0.906****	0.205	0.724
	C %	0.256	-0.838	0.782	0.813	**0.940****
	GC %	0.104	-0.748	**0.948****	0.626	**0.961****

Note: One asterisk (*) and two asterisks (**) indicate correlations at a level of 0.05 and 0.01, respectively

Furthermore, correlation between ENC and various GC contents, as shown in **Table 4.1.13**, suggests that the nucleotide composition under from mutation pressure and natural selection affect the synonymous codon usage in MT-ND3 gene for pisces, aves and mammals.

Table 4.1.13 Correlation coefficients among ENC, CAI, GC contents for MT-ND3 gene

		Pisces		**Aves**		**Mammals**	
SL No	Correlation between	Correlation coefficient	p value	Correlation coefficient	p value	Correlation coefficient	p value
1	ENC and CAI	-0.260	0.672	0.543	0.344	0.537	0.350
2	ENC and GC %	0.836	0.078	0.751	0.143	**0.930****	**0.022**
3	ENC and GC3 %	**0.936****	**0.019**	**0.913****	**0.031**	**0.995****	**0.000**
4	GC and GC3 %	**0.927****	**0.024**	**0.913****	**0.031**	**0.961****	**0.009**
5	GC1 and GC3 %	0.575	0.310	0.412	0.490	0.417	0.485
6	GC2 and GC3%	-0.457	0.439	0.405	0.499	0.626	0.258
7	GC12 and GC3 %	0.534	0.354	0.513	0.376	0.823	0.087

Note: One asterisk (*) and two asterisks (**) indicate correlations at a level of 0.05 and 0.01, respectively

In ND3 gene, the frequencies of two nucleotides *i.e.* A and T are not equal to G and C in pisces, aves and mammals indicating that other factor such as natural selection might have played a role in codon usage pattern.

To quantify the extent of natural selection and mutation pressure in the codon usage pattern of MT-ND3, we drew neutrality plot between GC12 and GC3. GC12 is the average of GC1 and GC2. The regression coefficient of GC12 on GC3 in pisces is 0.169 indicating relative neutrality of 16.9 % and relative constraint of 83.1 % for GC3 % in pisces as shown in **Figure 4.1.38 (a).** The GC12 was influenced by mutation pressure and natural selection with a ratio of 0.169/0.831= 0.203. In aves, the regression coefficient of GC12 on GC3 was 0.186 indicating relative neutrality of 18.6% and relative constraint of 81.4% as shown in **Figure 4.1.38 (b)**. The GC12 was influenced by mutation pressure and natural selection with a ratio of 0.186/0.814= 0.228. In mammals, the regression coefficient of GC12 on GC3 was 0.357 indicating relative neutrality of 35.7 % and relative constraint of 64.3 % for GC3 % as shown in **Figure 4.1.38 (c).** The GC12 was influenced by mutation pressure and natural selection with a ratio of 0.357/0.643= 0.555. These results suggest that natural selection played a major role while mutation pressure played a minor role in shaping codon usage pattern in ND3 gene among pisces, aves and mammals.

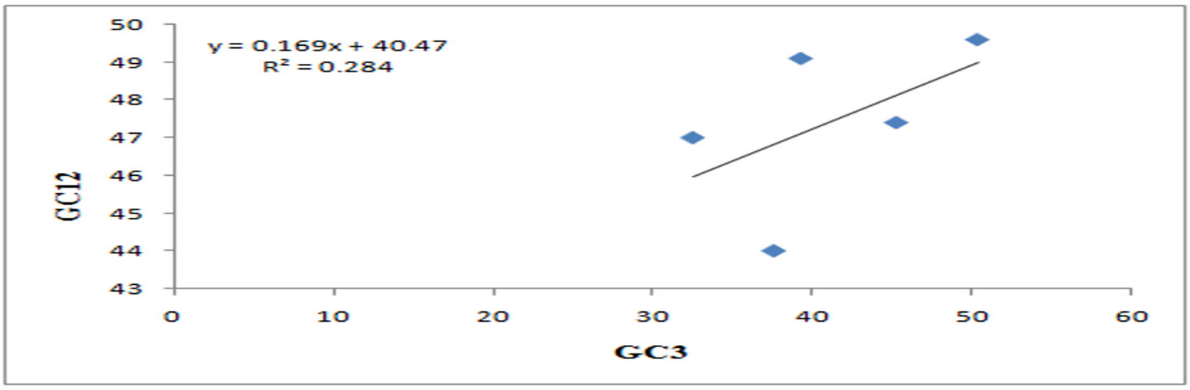


Figure 4.1.38(a) Neutrality plot of GC12 with GC3 in different species of pisces for ND3 gene

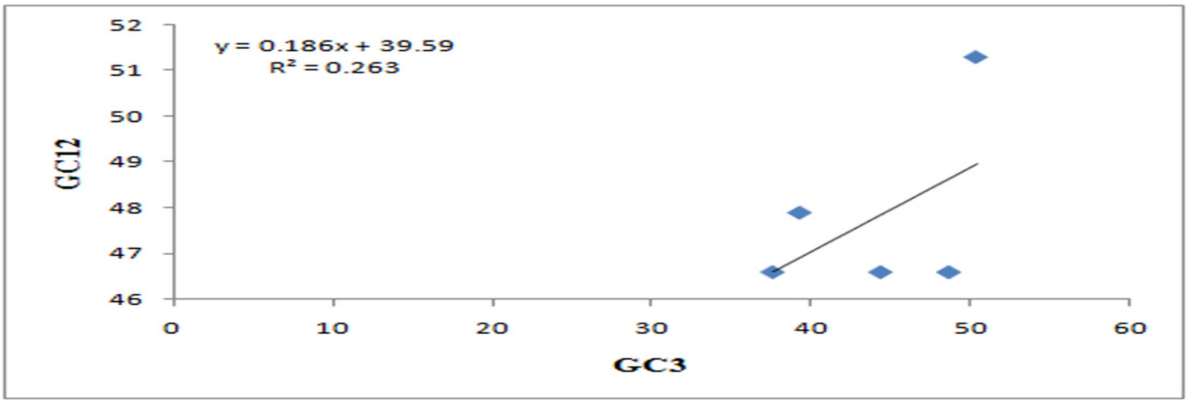


Figure 4.1.38(b) Neutrality plot of GC12 with GC3 in different species of aves for ND3 gene

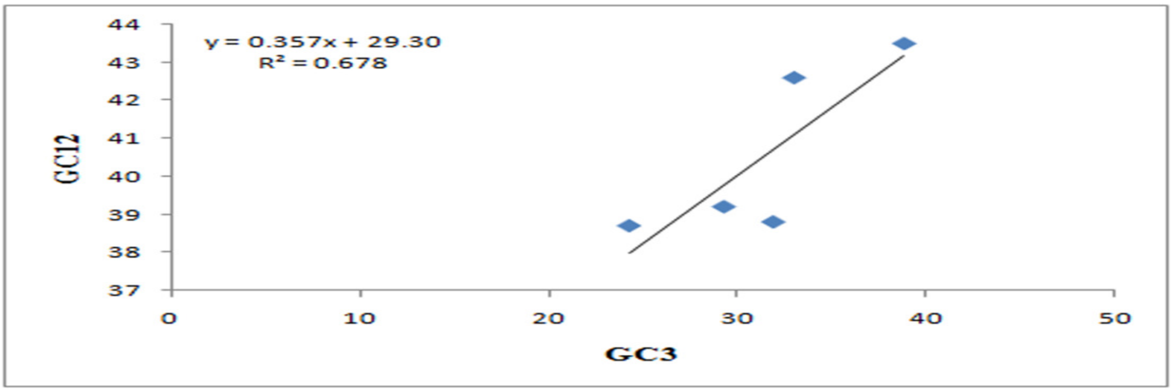


Figure 4.1.38 (c) Neutrality plot of GC12 with GC3 in different species of mammals for ND3 gene

In ND4, highly significant positive correlation was found between A and A3, T and T3, T and G3, C and C3, C and GC3, GC and C3, GC and GC3 in pisces. In aves, significant positive correlation was found between A and A3, T and T3, G and G3, C and T3, GC and G3, while significant negative correlation was found between T and A3. In mammals, significant positive correlation was found between A and A3, T and T3, G and G3, C and C3, C and GC3, GC and GC3 as shown in **Table 4.1.14.** These results suggest that the compositional constraint arising from mutation pressure and natural selection determine the pattern of codon usage in ND4 gene.

Table 4.1.14 Correlation between overall nucleotide composition (%) and its composition at 3rd codon position in pisces, aves and mammals in ND4 gene

	Nucleotide	**A3 %**	**T3 %**	**G3 %**	**C3 %**	**GC3%**
Pisces	A %	**0.998****	-0.452	-0.499	-0.484	-0.728
	T %	-0.500	**0.992****	**0.924****	-0.562	-0.278
	G %	-0.836	0.800	0.872	-0.066	0.255
	C %	-0.409	-0.567	-0.486	**0.999****	**0.930****
	GC %	-0.687	-0.274	-0.171	**0.949****	**0.994****
Aves	A %	**0.979****	-0.627	-0.171	-0.663	-0.684
	T %	**-0.933****	**0.898***	-0.366	0.793	0.308
	G %	0.108	-0.693	**0.972****	-0.291	0.634
	C %	-0.679	**0.906****	-0.516	0.496	-0.064
	GC %	-0.348	-0.266	**0.932****	-0.040	0.800
Mammals	A %	**0.985****	-0.670	-0.660	0.009	-0.246
	T %	-0.453	**0.971****	-0.172	-0.743	-0.629
	G %	-0.867	0.141	**0.939****	0.474	0.718
	C %	-0.205	-0.651	0.688	**0.887***	**0.936****
	GC %	-0.445	-0.443	0.835	0.824	**0.945****

Note: One asterisk (*) and two asterisks (**) indicate correlations at a level of 0.05 and 0.01, respectively

Furthermore, correlation between ENC and various GC contents, as shown in **Table 4.1.15**, suggests that the nucleotide composition under from mutation pressure and natural selection affect the synonymous codon usage of MT-ND4 gene in pisces, aves and mammals.

Table 4.1.15 Correlation coefficients among ENC, CAI, GC contents for MT-ND4 gene

		Pisces		**Aves**		**Mammals**	
SL No	Correlation between	Correlation coefficient	p value	Correlation coefficient	p value	Correlation coefficient	p value
1	ENC and CAI	0.282	0.646	-0.729	0.162	0.658	0.227
2	ENC and GC %	**0.943****	**0.016**	**-0.992****	**0.001**	**0.938****	**0.018**
3	ENC and GC3 %	**0.933****	**0.021**	-0.783	0.117	**0.993****	**0.001**
4	GC and GC3 %	**0.994****	**0.001**	0.800	0.104	**0.945****	**0.015**
5	GC1 and GC3 %	**0.968****	**0.007**	-0.367	0.543	0.698	0.190
6	GC2 and GC3%	0.507	0.383	-0.147	0.814	-0.102	0.871
7	GC12 and GC3 %	**0.947****	**0.014**	-0.535	0.353	0.557	0.329

Note: One asterisk (*) and two asterisks (**) indicate correlations at a level of 0.05 and 0.01, respectively

To quantify the extent of natural selection and mutation pressure in the codon usage pattern of MT-ND4, we drew the neutrality plot between GC12 and GC3. GC12 is the average of GC1 and GC2. The regression coefficient of GC12 on GC3 in pisces is 0.237 indicating relative neutrality of 23.7% and relative constraint of 76.3% for GC3% in pisces as shown in **Figure 4.1.39 (a).** The GC12 was influenced by mutation pressure and natural selection with a ratio of 0.237/0.763= 0.310. In aves, the regression coefficient of GC12 on GC3 was 0.146 indicating relative neutrality of 14.6% (absolute) and relative constraint of 85.4% as shown in **Figure 4.1.39 (b)**. The GC12 was influenced by mutation pressure and natural selection with a ratio of 0.146/0.854=0.170. In mammals, the regression coefficient of GC12 on GC3 was 0.153 indicating relative neutrality of 15.3% and relative constraint of 84.7% for GC3 % as shown in **Figure 4.1.39 (c).** The GC12 was influenced by mutation pressure and natural selection with a ratio of 0.153/0.847= 0.180. These results suggest that natural selection played a major role while mutation pressure played a minor role in shaping the codon usage pattern in ND4 gene among pisces, aves and mammals

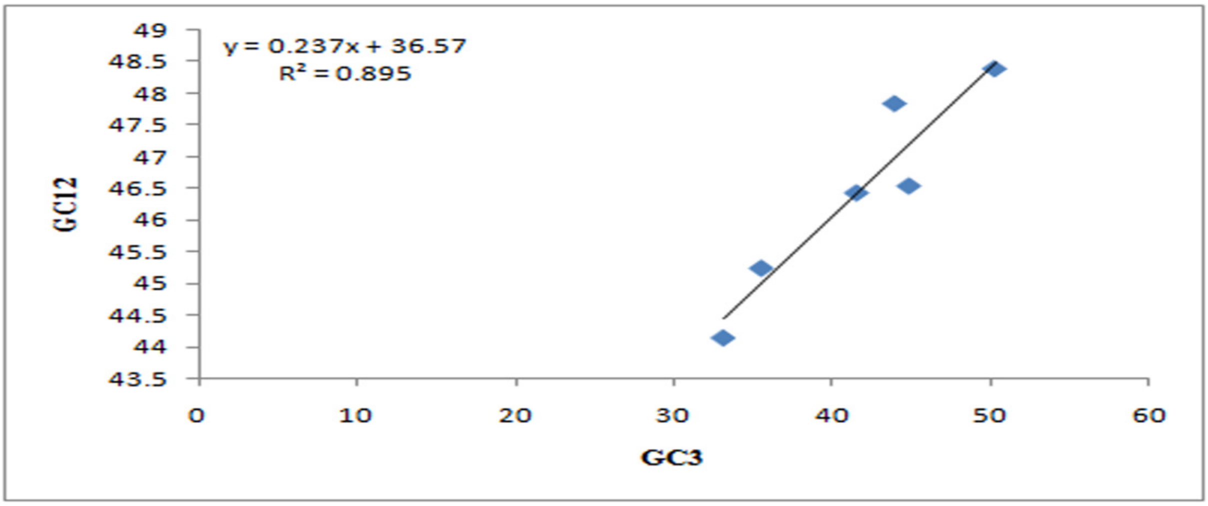


Figure 4.1.39(a) Neutrality plot of GC12 with GC3 in different species of pisces for ND4 gene

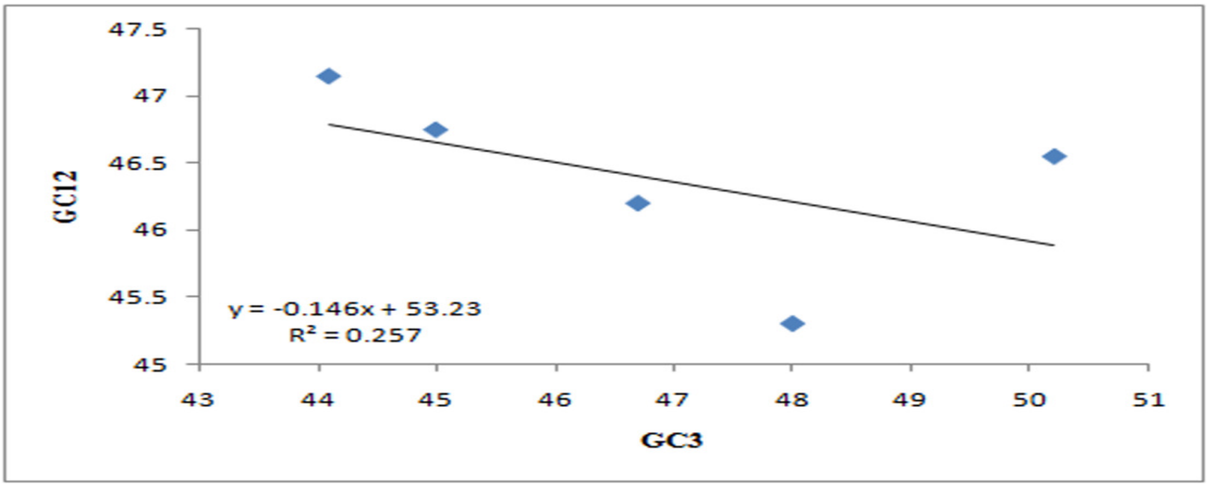


Figure 4.1.39(b) Neutrality plot of GC12 with GC3 in different species of aves for ND4 gene

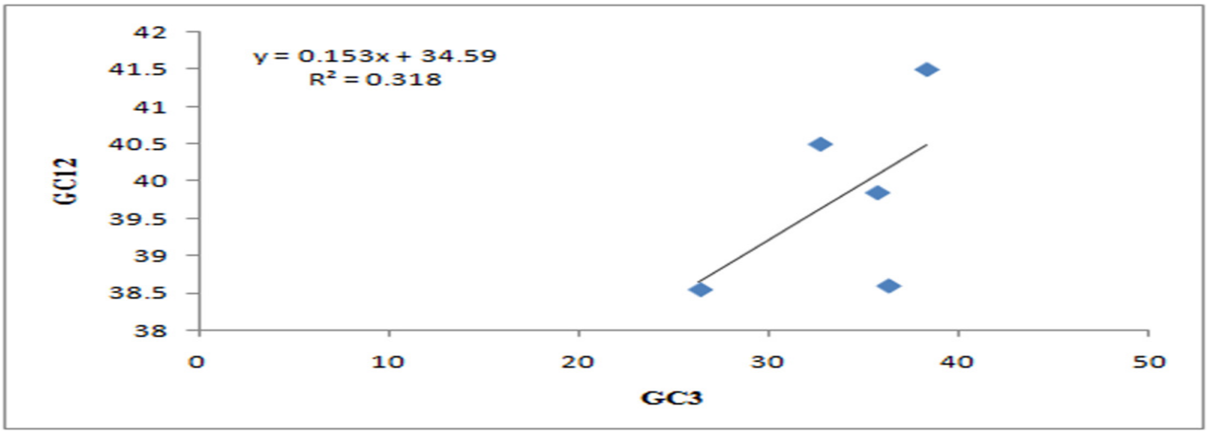


Figure 4.1.39(c) Neutrality plot of GC12 with GC3 in different species of mammals for ND4 gene

In ND4L gene, highly significant positive correlation was found between A and A3, T and T3, C and C3, C and GC3, GC and C3, GC and GC3 in pisces while negative correlation was found between G and C3, G and GC3. In aves, significant positive correlation was found between A and A3, T and T3, G and G3, GC and G3, GC and GC3 while negative correlation was found between A and GC3. In mammals, significant positive correlation was found between A and A3, G and G3, as shown in **Table 4.1.16.** These results suggest that the compositional constraint arising from mutation pressure and natural selection determine the pattern of codon usage in ND4L gene.

Table 4.1.16 Correlation between overall nucleotide composition (%) and its composition at 3rd codon position in pisces, aves and mammals in ND4L gene

	Nucleotide	**A3 %**	**T3 %**	**G3 %**	**C3 %**	**GC3%**
Pisces	A %	__0.967**__	-0.627	-0.144	-0.250	-0.298
	T %	-0.511	__0.997**__	0.647	-0.527	-0.492
	G %	0.405	0.630	0.761	__-0.996**__	__-0.995**__
	C %	-0.602	-0.430	-0.625	__0.970**__	__0.985**__
	GC %	-0.644	-0.376	-0.581	__0.952**__	__0.972**__
Aves	A %	__0.946**__	0.348	-0.788	-0.409	__-0.912**__
	T %	0.168	__0.879*__	-0.777	0.156	-0.579
	G %	-0.674	-0.469	__0.988**__	-0.156	0.762
	C %	-0.158	-0.579	-0.090	0.875	0.419
	GC %	-0.746	-0.696	__0.950**__	0.204	__0.934**__
Mammals	A %	__0.911**__	-0.685	0.387	-0.156	0.064
	T %	-0.821	0.852	-0.296	-0.165	-0.329
	G %	0.398	-0.613	__0.947**__	-0.142	0.387
	C %	0.062	-0.212	-0.697	0.593	0.198
	GC %	0.423	-0.821	0.071	0.578	0.610

Note: One asterisk (*) and two asterisks (**) indicate correlations at a level of 0.05 and 0.01, respectively

Furthermore, correlation between ENC and various GC contents, as shown in **Table 4.1.17**, suggests that the nucleotide composition under from mutation pressure and natural selection affect the synonymous codon usage of MT-ND4L gene in pisces, aves and mammals.

Table 4.1.17 Correlation coefficients among ENC, CAI, GC contents for MT-ND4L gene

		Pisces		**Aves**		**Mammals**	
SL No	Correlation between	Correlation coefficient	p value	Correlation coefficient	p value	Correlation coefficient	p value
1	ENC and CAI	-0.605	0.280	-0.117	0.851	0.216	**0.026**
2	ENC and GC %	__0.921**__	**0.026**	-0.506	0.384	0.586	0.299
3	ENC and GC3 %	__0.939**__	**0.018**	-0.198	0.750	__0.997**__	**0.000**
4	GC and GC3 %	__0.972**__	**0.000**	**0.934**	**0.020**	0.610	0.275
5	GC1 and GC3 %	0.372	**0.006**	0.124	0.843	-0.234	0.705
6	GC2 and GC3%	-0.477	0.538	-0.724	0.166	-0.063	0.919
7	GC12 and GC3 %	-0.163	0.793	-0.189	0.760	-0.366	0.545

Note: One asterisk (*) and two asterisks (**) indicate correlations at a level of 0.05 and 0.01, respectively

To quantify the extent of natural selection and mutation pressure in the codon usage pattern of MT-ND4L, we drew the neutrality plot between GC12 and GC3. GC12 is the average of GC1 and GC2. The regression coefficient of GC12 on GC3 in pisces is 0.019 indicating relative neutrality of 1.9 (absolute) % and relative constraint of 98.1 % for GC3 % in pisces as shown in **Figure 4.1.40 (a).** The GC12 was influenced by mutation pressure and natural selection with a ratio of 0.019/0.981= 0.019. In aves, the regression coefficient of GC12 on GC3 was 0.034 indicating relative neutrality 3.4% (absolute) and relative constraint is 96.6% as shown in **Figure 4.1.40 (b)**. The GC12 was influenced by mutation pressure and natural selection with a ratio of 0.034/0.966= 0.035. In mammals, the regression coefficient of GC12 on GC3 was 0.171 indicating relative neutrality of 17.1% (absolute) and relative constraint of 82.9 % for GC3 % as shown in **Figure 4.1.40 (c).** The GC12 was influenced by mutation pressure and natural selection with a ratio of 0.171/0.829= 0.206. These results suggest that natural selection played a major role while mutation pressure played a minor role in shaping the codon usage pattern in ND4L gene among pisces, aves and mammals.

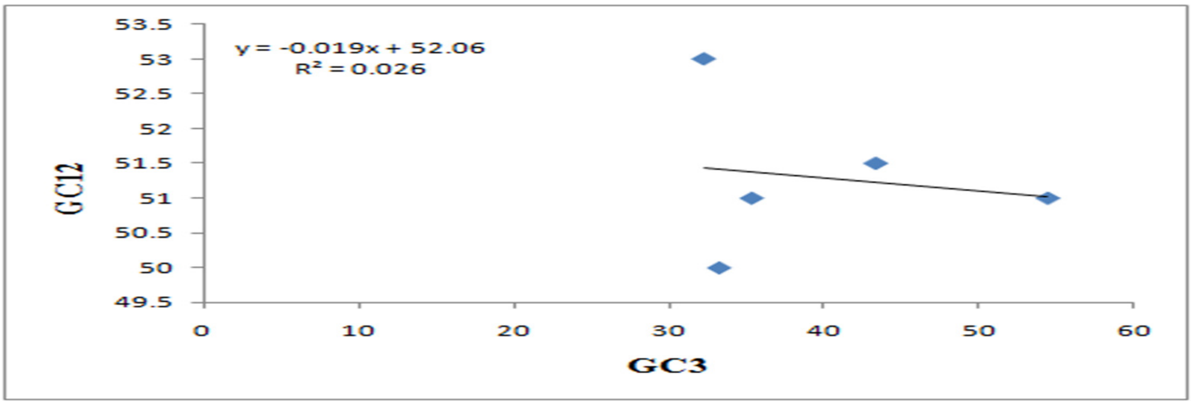


Figure 4.1.40(a) Neutrality plot of GC12 with GC3 in different species of pisces for ND4L gene

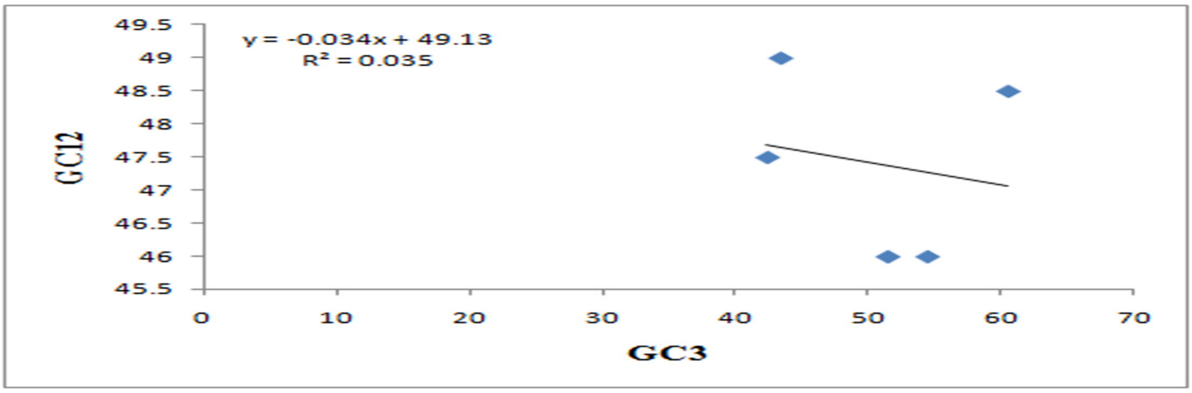


Figure 4.1.40(b) Neutrality plot of GC12 with GC3 in different species of aves for ND4L gene

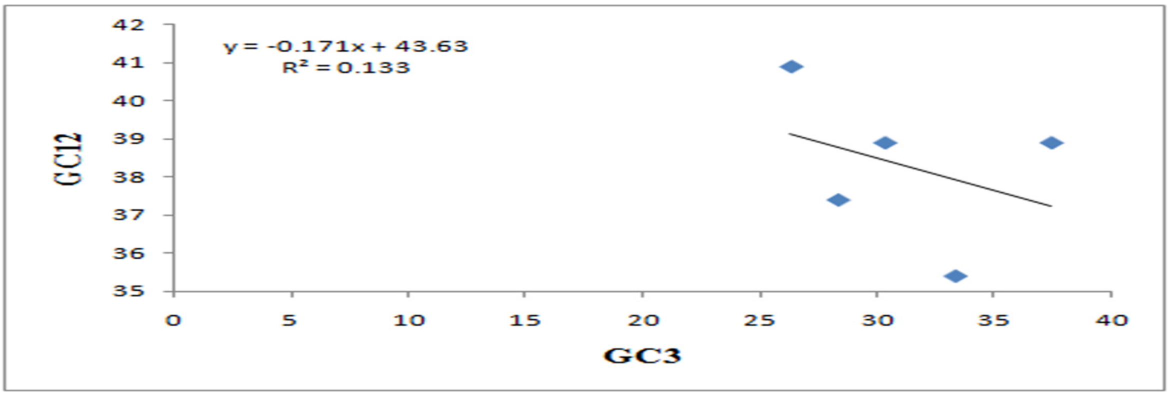


Figure 4.1.40 (c) Neutrality plot of GC12 with GC3 in different species of mammals for ND4L gene

In ND5 gene, highly significant positive correlation was found between A and A3, T and T3, G and G3, C and C3, C and GC3, GC and GC3 in pisces. In aves, significant positive correlation was found between T and T3, G and G3, while negative correlation was found between G and T3. In mammals, significant positive correlation was found between A and A3, T and T3, G and G3, C and C3, GC and GC3 while negative correlation was found between C and T3 as shown in **Table 4.1.18.** These results suggest that the compositional constraint, arising from mutation pressure and natural selection, determine the pattern of codon usage in ND5 gene.

Table 4.1.18 Correlation between overall nucleotide composition (%) and its composition at 3rd codon position in pisces, aves and mammals for ND5 gene

	Nucleotide	**A3 %**	**T3 %**	**G3 %**	**C3 %**	**GC3%**
Pisces	A %	**0.995****	-0.736	-0.366	-0.057	-0.276
	T %	-0.702	**0.994****	0.741	-0.624	-0.452
	G %	-0.592	0.759	**0.934****	-0.567	-0.271
	C %	-0.260	-0.454	-0.740	**0.988****	**0.951****
	GC %	-0.614	-0.111	-0.358	0.828	**0.939****
Aves	A %	0.705	0.045	-0.586	-0.831	-0.777
	T %	-0.482	**0.920****	-0.657	-0.057	-0.490
	G %	0.139	**-0.879***	**0.951****	0.292	0.801
	C %	-0.485	-0.077	0.322	0.818	0.584
	GC %	0.380	-0.144	-0.276	-0.099	-0.238
Mammals	A %	**0.938****	-0.792	-0.674	0.338	-0.022
	T %	-0.544	**0.975****	-0.083	-0.860	-0.710
	G %	-0.806	0.371	**0.937****	0.069	0.452
	C %	0.336	**-0.896***	0.263	**0.942****	0.850
	GC %	-0.228	-0.480	0.787	0.782	**0.947****

Note: One asterisk (*) and two asterisks (**) indicate correlations at a level of 0.05 and 0.01, respectively

Furthermore, correlations between ENC and various GC contents in pisces, aves and mammals, as shown in **Table 4.1.19**, which suggests that the nucleotide composition under mutation pressure and natural selection affect the synonymous codon usage of MT-ND5 gene in pisces, aves and mammals.

Table 4.1.19 Correlation coefficients among ENC, CAI, GC contents for MT-ND5 gene

		Pisces		**Aves**		**Mammals**	
SL No	Correlation between	Correlation coefficient	p value	Correlation coefficient	p value	Correlation coefficient	p value
1	ENC and CAI	-0.452	0.444	0.016	0.979	0.089	0.887
2	ENC and GC %	**0.924****	**0.025**	-0.545	0.342	**0.938****	**0.018**
3	ENC and GC3 %	**0.999****	**< 0.0001**	0.274	0.655	**0.999****	**< 0.0001**
4	GC and GC3 %	**0.939****	**0.018**	-0.238	0.699	**0.947****	**0.014**
5	GC1 and GC3 %	0.757	0.139	-0.580	0.305	0.654	0.232
6	GC2 and GC3%	0.411	0.492	-0.357	0.555	0.207	0.738
7	GC12 and GC3 %	0.654	0.231	0.651	0.234	0.571	0.315

Note: One asterisk (*) and two asterisks (**) indicate correlations at a level of 0.05 and 0.01, respectively

To quantify the extent of natural selection and mutation pressure in the codon usage pattern of MT-ND5, we drew the neutrality plot between GC12 and GC3. GC12 is the average of GC1 and GC2. The regression coefficient of GC12 on GC3 in pisces is 0.227 indicating relative neutrality of 22.7 % and relative constraint of 77.3 % for GC3 % in pisces as shown in **Figure 4.1.41 (a).** The GC12 was influenced by mutation pressure and natural selection with a ratio of 0.227/0.773= 0.293. In aves, the regression coefficient of GC12 on GC3 was 0.349 indicating relative neutrality 34.9% and relative constraint 65.1% as shown in **Figure 4.1.41 (b)**. The GC12 was influenced by mutation pressure and natural selection with a ratio of 0.349/0.651= 0.536. In mammals, the regression coefficient of GC12 on GC3 was 0.157 indicating relative neutrality of 15.7 % and relative constraint of 84.3 % for GC3 % as shown in **Figure 4.1.41 (c).** The GC12 was influenced by mutation pressure and natural selection with a ratio of 0.157/0.843= 0.186. These results suggest that natural selection played a major role while mutation pressure played a minor role in shaping the codon usage pattern in ND5 gene among pisces, aves and mammals.

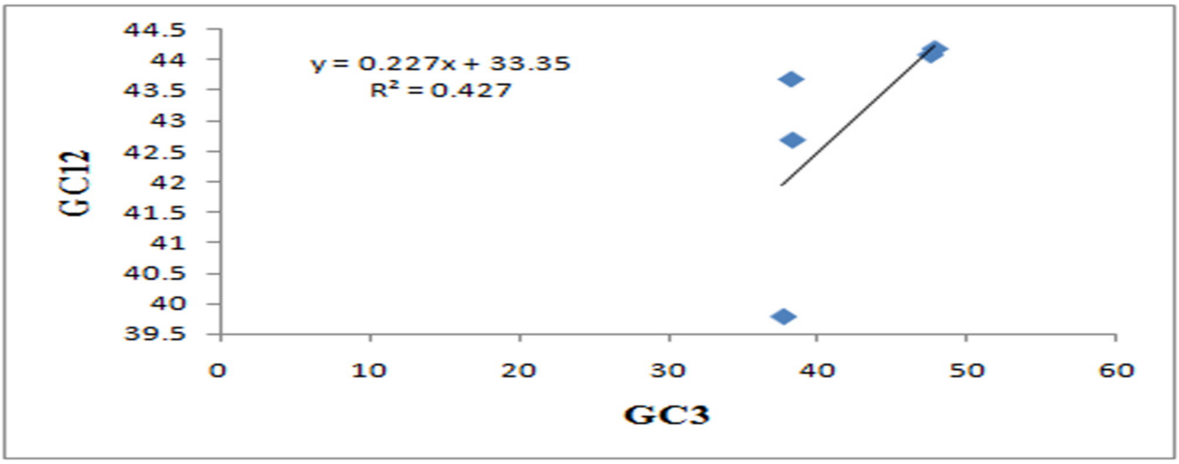


Figure 4.1.41(a) Neutrality plot of GC12 with GC3 in different species of pisces for ND5 gene

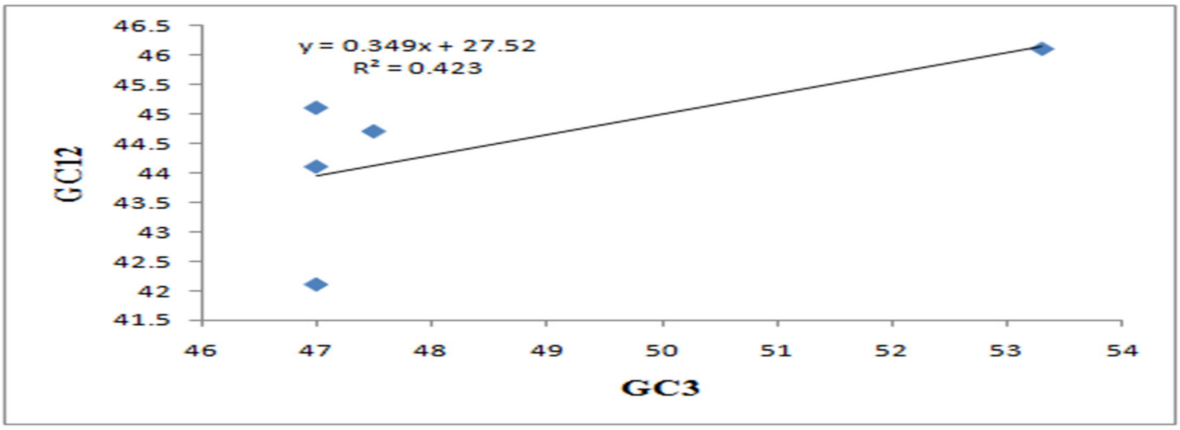


Figure 4.1.41(b) Neutrality plot of GC12 with GC3 in different species of aves for ND5 gene

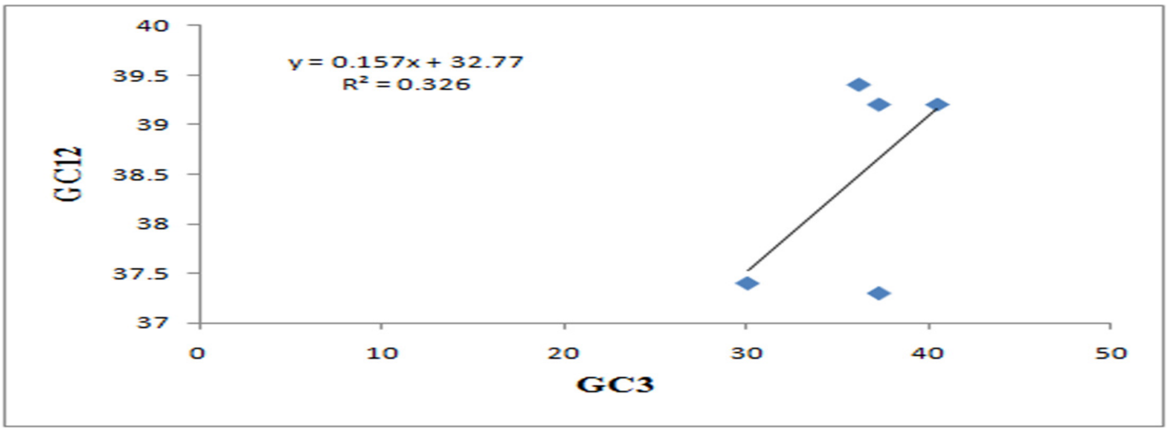


Figure 4.1.41(c) Neutrality plot of GC12 with GC3 in different species of mammals for ND5 gene

In ND6 gene, highly significant positive correlation was found between A and A3, T and T3, G and G3, C and C3, GC and C3, GC and GC3 in pisces. In aves, significant positive correlation was found between G and G3, G and GC3, while negative correlation was found between G and A3. In mammals, significant positive correlation was found between A and A3, T and G3, while negative correlation was found between C and G3 as shown in **Table 4.1.20.** These results suggest that the compositional constraint, arising from mutation pressure and natural selection, determine the pattern of codon usage in ND6 gene.

Table 4.1.20 Correlation between overall nucleotide composition (%) and its composition at 3rd codon position in pisces, aves and mammals in ND6 gene

	Nucleotide	A3 %	T3 %	G3 %	C3 %	GC3%
Pisces	A %	0.938**	-0.437	-0.668	-0.783	-0.971
	T %	-0.827	0.911**	0.901**	0.079	0.558
	G %	-0.834	0.651	0.911**	0.276	0.715
	C %	-0.280	-0.390	-0.195	0.901**	0.586
	GC %	-0.666	-0.040	0.267	0.964**	0.889*
Aves	A %	0.846	0.012	-0.923**	-0.550	-0.860
	T %	-0.640	0.545	0.655	-0.049	0.447
	G %	-0.963**	0.047	0.931**	0.793	0.950**
	C %	0.539	-0.358	-0.386	-0.375	-0.405
	GC %	-0.767	-0.278	0.858	0.720	0.877
Mammals	A %	0.977**	-0.434	0.086	-0.428	-0.513
	T %	-0.049	0.587	0.967**	-0.668	-0.391
	G %	-0.422	0.251	0.379	-0.007	0.174
	C %	-0.326	-0.257	-0.901**	0.700	0.468
	GC %	-0.706	-0.119	-0.796	0.833	0.690

Note: One asterisk (*) and two asterisks (**) indicate correlations at a level of 0.05 and 0.01, respectively

Furthermore, correlation was found between ENC and various GC contents as shown in **Table 4.1.21**, which suggests that the nucleotide composition under mutation pressure and natural selection affect the synonymous codon usage in MT-ND6 gene for pisces, aves and mammals.

Table 4.1.21 Correlation coefficients among ENC, CAI, GC contents for MT-ND6 gene

		Pisces		Aves		Mammals	
SL No	Correlation between	Correlation coefficient	p value	Correlation coefficient	p value	Correlation coefficient	p value
1	ENC and CAI	0.165	0.791	0.307	0.615	0.067	0.914
2	ENC and GC %	-0.606	0.279	-0.795	0.108	0.720	0.170
3	ENC and GC3 %	-0.746	0.148	-0.976**	0.004	0.941**	0.017
4	GC and GC3 %	0.889*	0.044	0.877	0.051	0.690	0.198
5	GC1 and GC3 %	0.558	0.329	-0.069	0.913	0.038	0.951
6	GC2 and GC3%	0.947**	0.015	0.262	0.671	0.748	0.146
7	GC12 and GC3 %	0.744	0.150	0.013	0.983	0.237	0.701

Note: One asterisk (*) and two asterisks (**) indicate correlations at a level of 0.05 and 0.01, respectively

To quantify the extent of natural selection and mutation pressure in the codon usage pattern of MT-ND6 gene, we drew the neutrality plot between GC12 and GC3. GC12 is the average of GC1 and GC2. The regression coefficient of GC12 on GC3 in pisces is 0.673 indicating relative neutrality of 67.3% and relative constraint of 32.7% for GC3 % in pisces as shown in **Figure 4.1.42 (a).** The GC12 was influenced by mutation pressure and natural selection with a ratio of 0.673/0.327= 2.05. In aves, the regression coefficient of GC12 on GC3 was 0.003 indicating relative neutrality of 0.3% and relative constraint of 99.7% as shown in **Figure 4.1.42 (b)**. The GC12 was influenced by mutation pressure and natural selection with a ratio of 0.003/0.997= 0.003. In mammals, the regression coefficient of GC12 on GC3 was 0.171 indicating relative neutrality of 17.1 % and relative constraint of 82.9% for GC3% as shown in **Figure 4.1.42 (c).** The GC12 was influenced by mutation pressure and natural selection with a ratio of 0.171/0.829= 0.206. These results suggest that natural selection played the major role while mutation pressure played the minor role in shaping the codon usage pattern in ND6 gene between aves and mammals while in pisces, mutation pressure played the major role while natural selection played the minor role in shaping the codon usage pattern in ND6 gene.

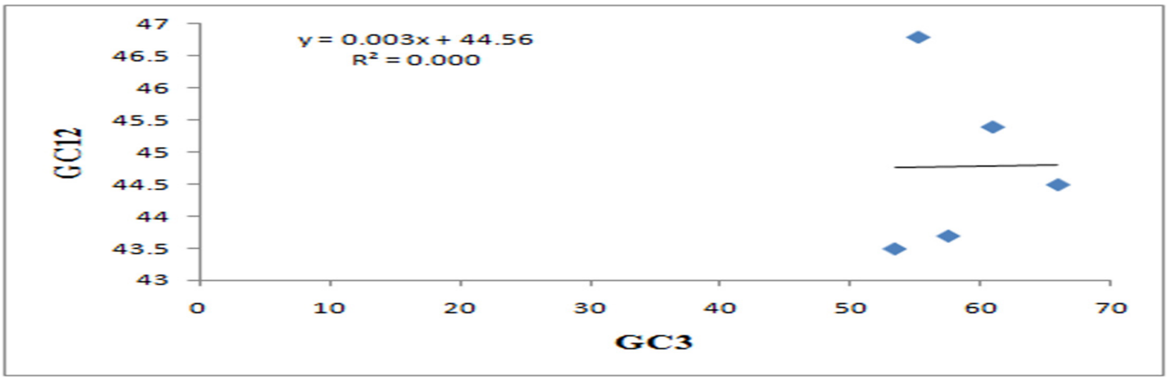


Figure 4.1.42(a) Neutrality plot of GC12 with GC3 in different species of pisces for ND6 gene

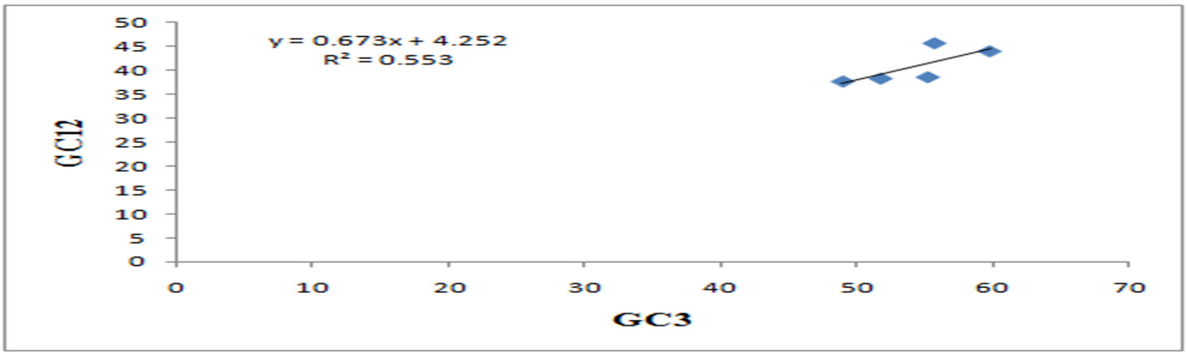


Figure 4.1.42(b) Neutrality plot of GC12 with GC3 in different species of aves for ND6 gene

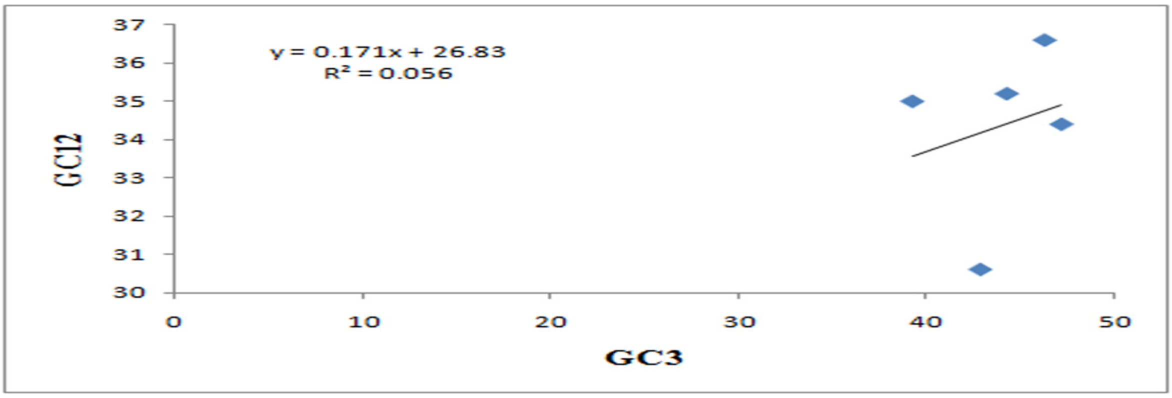


Figure 4.1.42 (c) Neutrality plot of GC12 with GC3 in different species of mammals for ND6 gene

4.1.7 Correspondence analysis (COA) of MT-ND gene among pisces, aves and mammals

To determine the trends in codon usage variation among different species of pisces, aves and mammals for MT-ND gene, we performed COA using RSCU values of codons.

In ND1 gene, the four main contributors *i.e.* axis1, axis2, axis3 and axis4 for pisces, aves and mammals were shown in **Figure 4.1.43.**

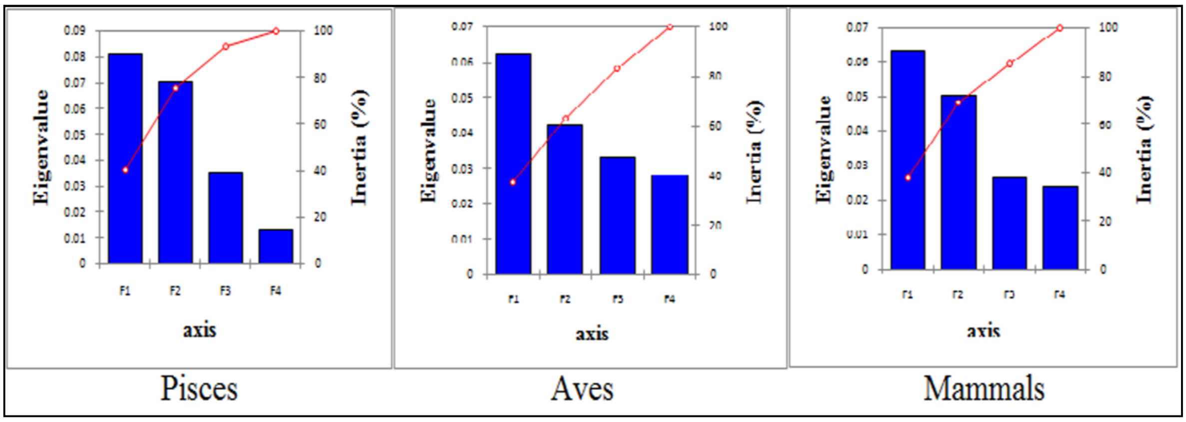


Figure 4.1.43 Contributions of the axes for MT-ND1 gene in pisces, aves and mammals

In ND1 gene, for pisces, the axis 1 accounted for 40.62% of the total variation; axis 2 accounted for 35.16%, of the total variation. In aves, axis 1 accounted for 37.55% of the total variation and axis 2 accounted for 25.49% of the total variation and in mammals, axis 1 accounted for 38.40% of the total variation and axis 2 accounted for 30.65% of the total variation as shown in **Figure 4.1.44**.

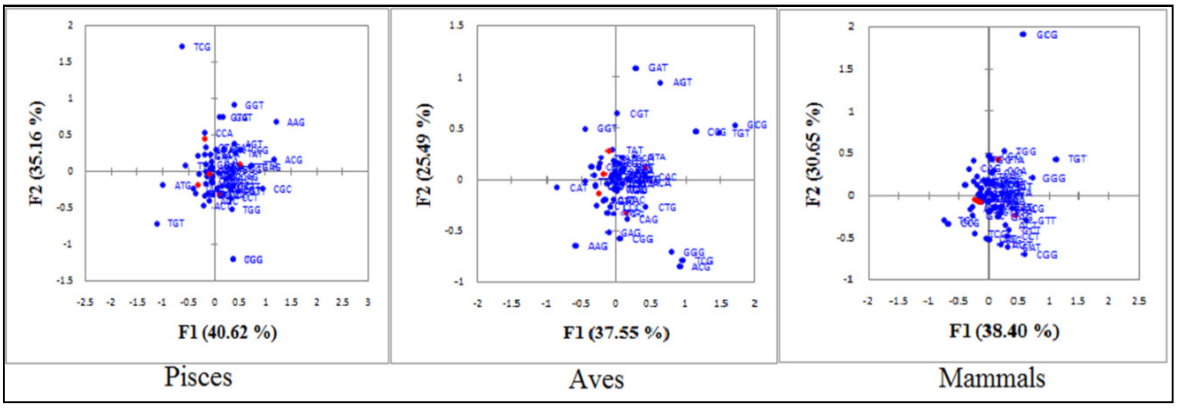


Figure 4.1.44 Correspondence analysis of the synonymous codon usage towards the codons in ND1 gene for pisces, aves and mammals

In ND2 gene, the four main contributors namely axis1, axis2, axis3 and axis4 for pisces, aves and mammals were shown in **Figure 4.1.45.**

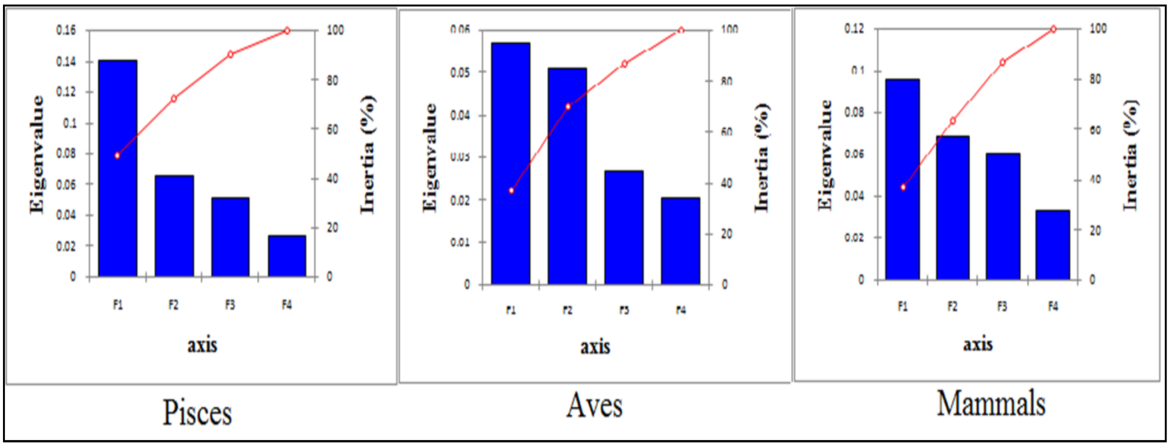
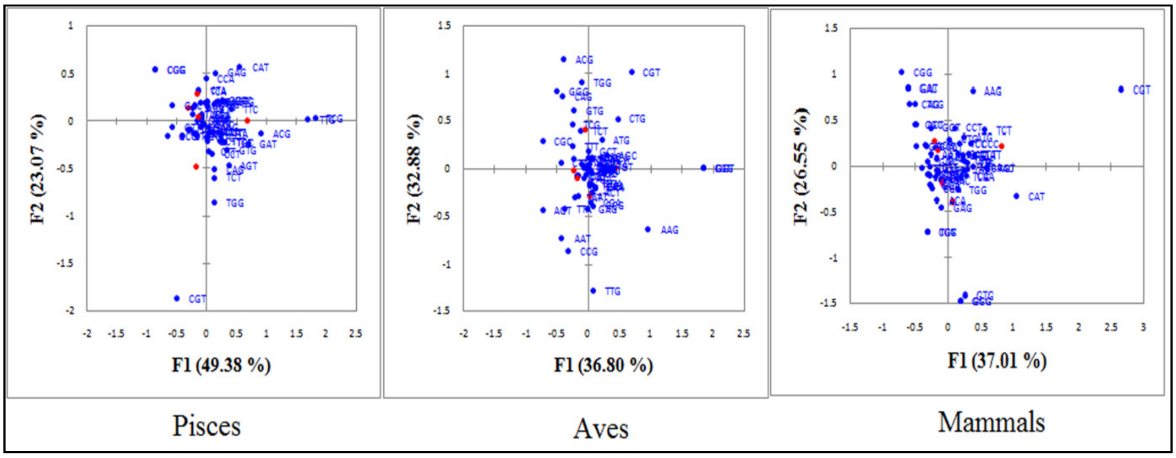


Figure 4.1.45 Contributions of the axes for MT-ND2 gene in pisces, aves and mammals

In ND2 gene, for pisces, the axis 1 accounted for 49.38% of the total variation; axis 2 accounted for 23.07%, of the total variation. In aves, axis 1 accounted for 36.80% of the total variation and axis 2 accounted for 32.88% of the total variation and in mammals, axis 1 accounted for 37.01% of the total variation and axis 2 accounted for 26.55% of the total variation as shown in **Figure 4.1.46**.



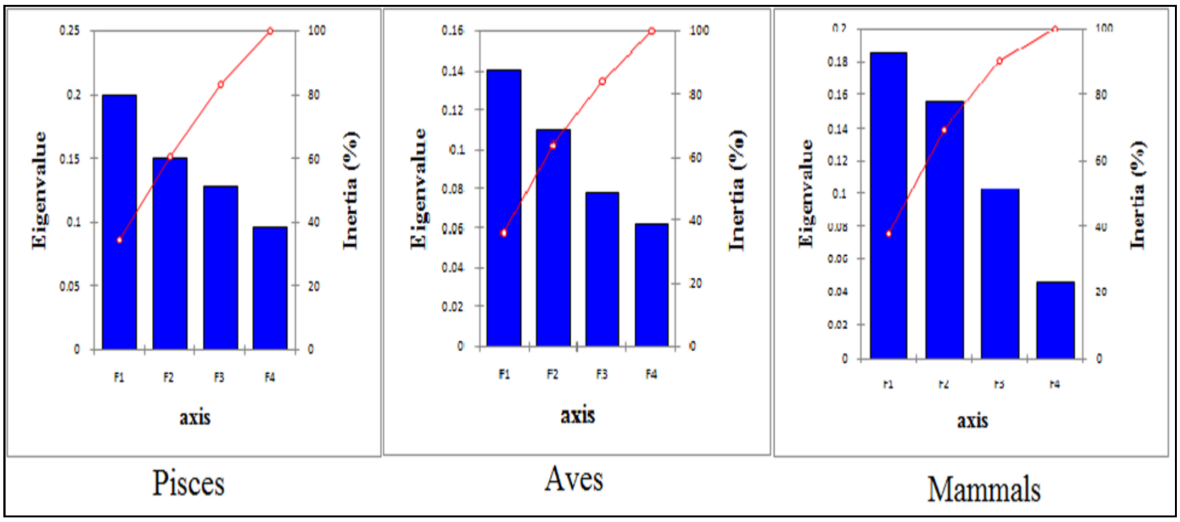
Pisces

Aves

Mammals

Figure 4.1.46 Correspondence analysis of synonymous codon usage towards the codons in ND2 gene for pisces, aves and mammals

In ND3 gene, the four main contributors of variation namely axis1, axis2, axis3 and axis4 for pisces, aves and mammals were shown in **Figure 4.1.47.**



Pisces

Aves

Mammals

Figure.4.1.47 Contributions of the axes for MT-ND3 gene in pisces, aves and mammals

In ND3 gene, for pisces, the axis 1 accounted for 34.63% of the total variation; axis 2 accounted for 26.24%, of the total variation. In aves, axis 1 accounted for 36% of the total variation and axis 2 accounted for 28.09% of the total variation and in mammals, axis 1 accounted for 37.77% of the total variation and axis 2 accounted for 31.80% of the total variation as shown in **Figure 4.1.48.**

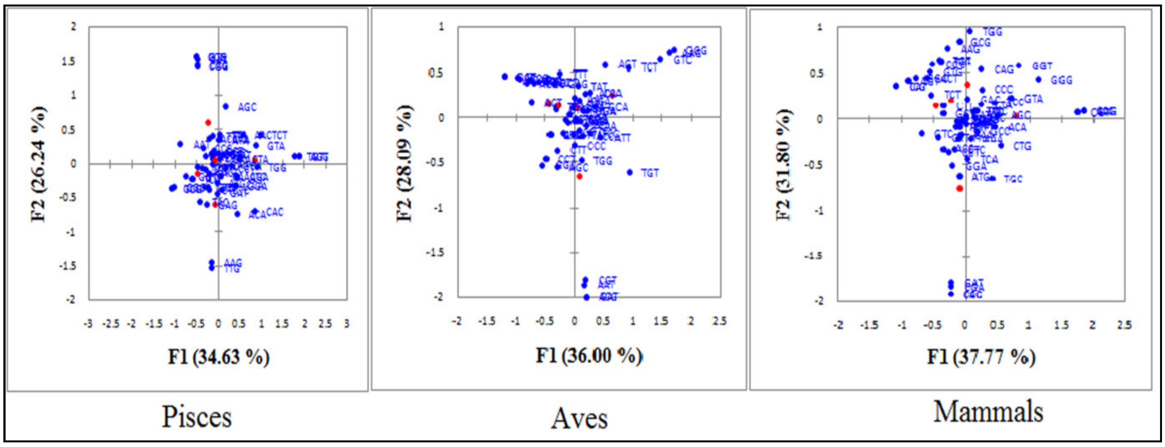


Figure 4.1.48 Correspondence analysis of synonymous codon usage towards the codons in ND3 gene for pisces, aves and mammals

In ND4 gene, the four main contributors of variation *i.e.* axis1, axis2, axis3 and axis4 for pisces, aves and mammals were shown in **Figure 4.1.49.**

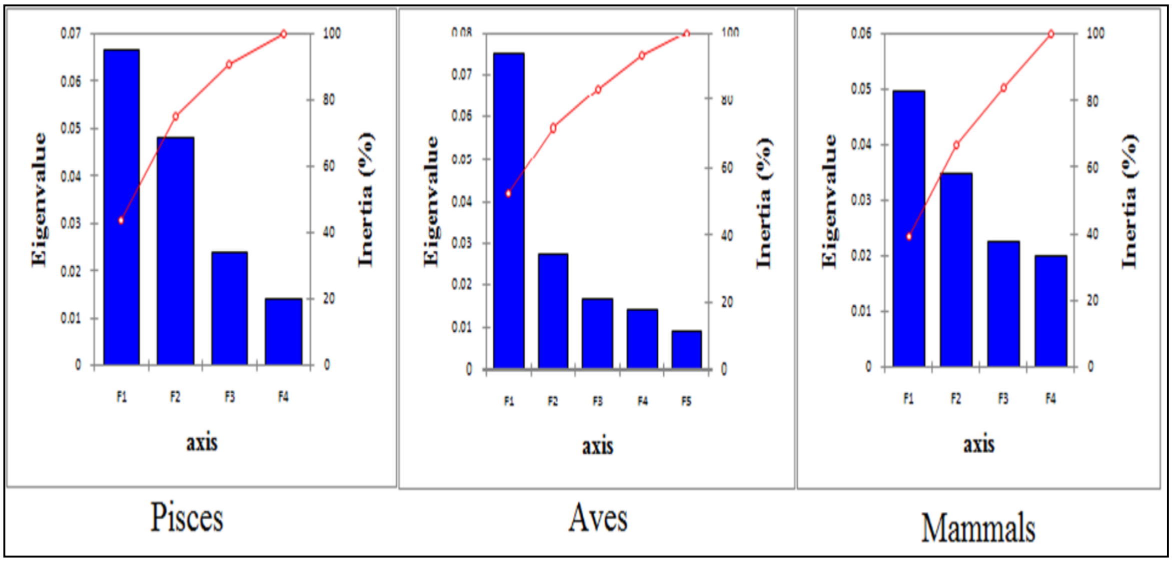


Figure 4.1.49 Contributions of the axes for MT-ND4 gene in pisces, aves and mammals

In ND4 gene, for pisces, the axis 1 accounted for 43.66% of the total variation; axis 2 accounted for 31.44%, of the total variation. In aves, axis 1 accounted for 52.44% of the total variation and axis 2 accounted for 19.33% of the total variation and in mammals, axis 1 accounted for 39.17% of the total variation and axis 2 accounted for 27.37% of the total variation as shown in **Figure 4.1.50.**

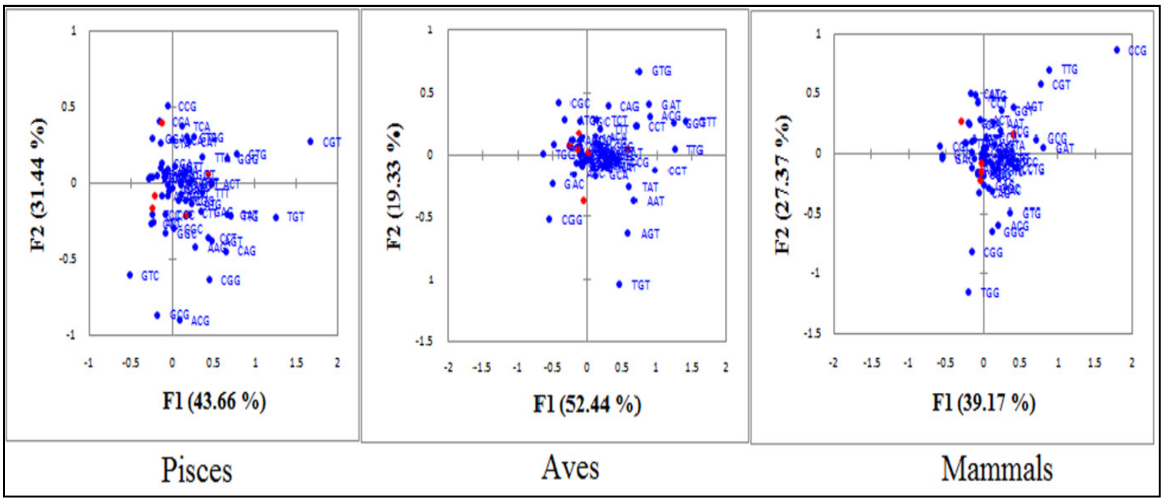


Figure 4.1.50 Correspondence analysis of synonymous codon usage towards the codons in ND4 gene for pisces, aves and mammals

In ND4L gene, the four main contributors namely axis1, axis2, axis3 and axis4 for pisces, aves and mammals were shown in **Figure 4.1.51.**

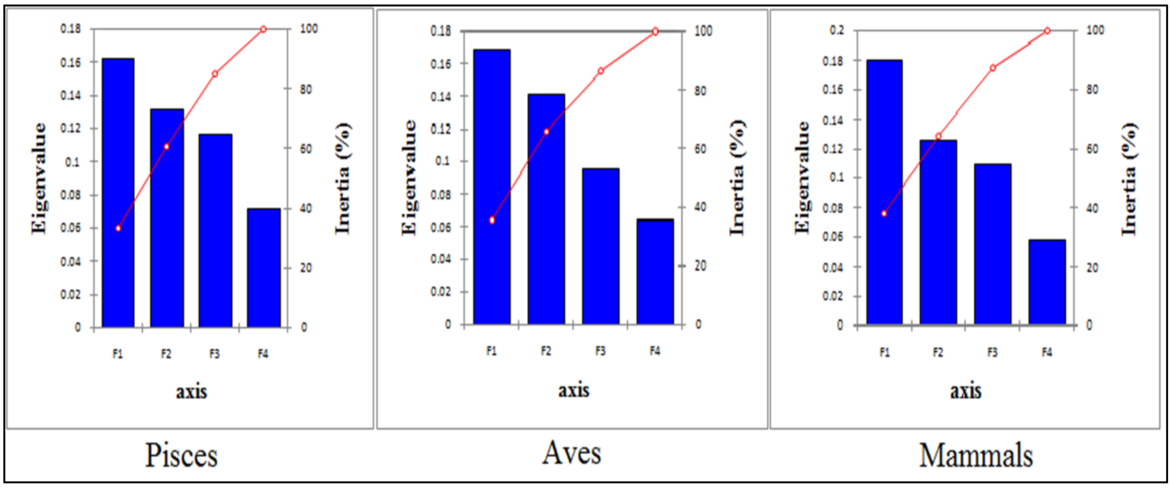


Figure 4.1.51 Contributions of the axes for MT-ND4L gene in pisces, aves and mammals

In ND4L gene, for pisces, the axis 1 accounted for 33.59% of the total variation; axis 2 accounted for 27.36% of the total variation. In aves, axis 1 accounted for 35.91% of the total variation and axis 2 accounted for 30.13% of the total variation and in mammals, axis 1 accounted for 38.12% of the total variation and axis 2 accounted for 26.64% of the total variation as shown in **Figure 4.1.52.**

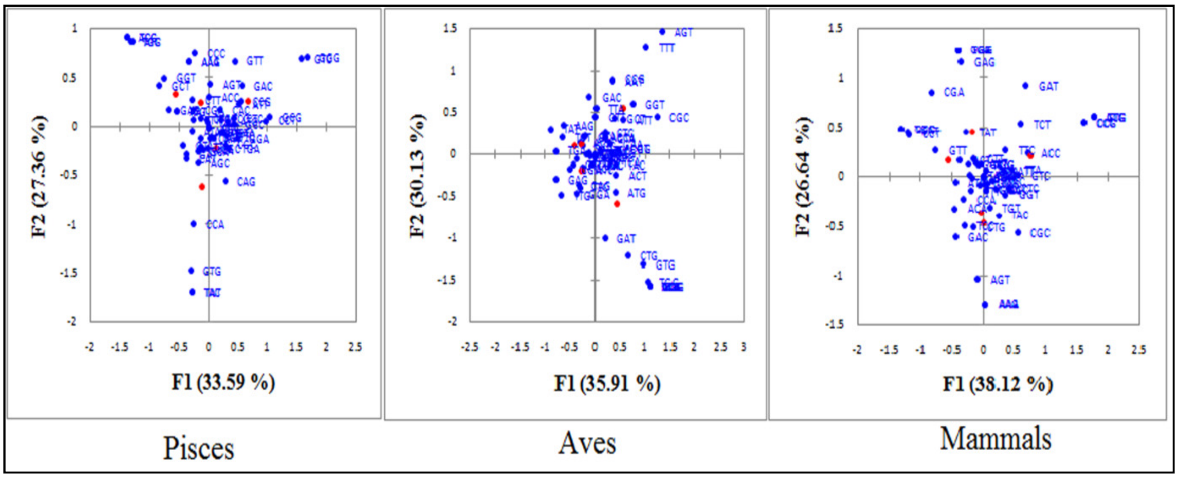


Figure 4.1.52 Correspondence analysis of synonymous codon usage towards the codons in ND4L gene for pisces, aves and mammals

In ND5 gene, the four main contributors of variation *i.e.* axis1, axis2, axis3 and axis4 for pisces, aves and mammals were shown in **Figure 4.1.53.**

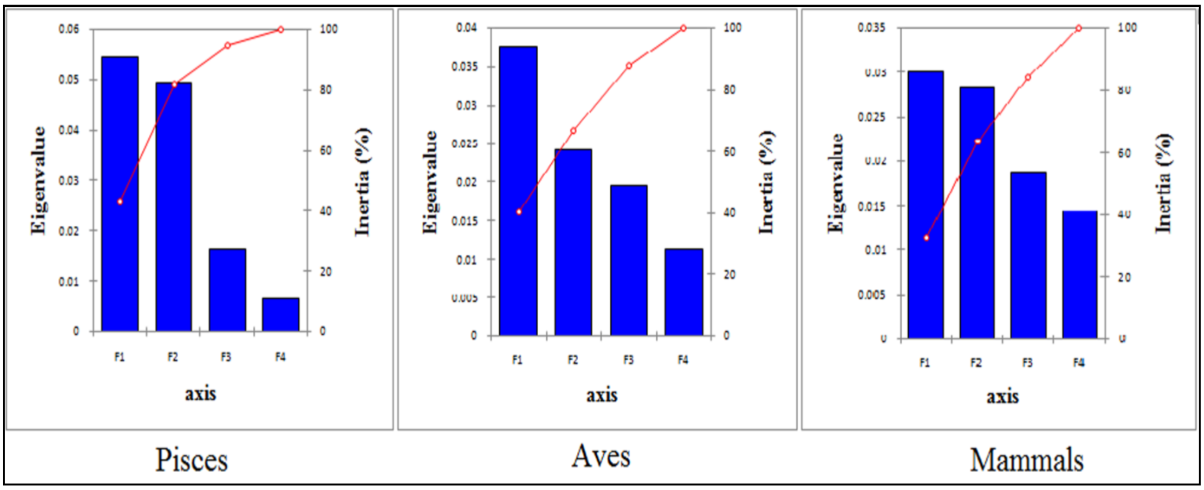
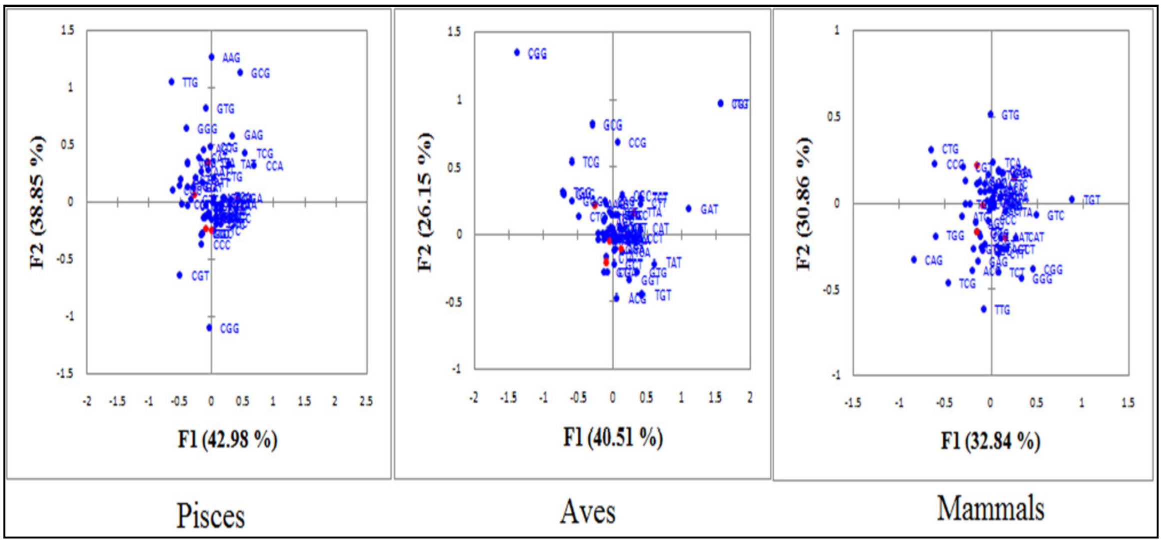


Figure 4.1.53 Contributions of the axes for MT-ND5 gene in pisces, aves and mammals

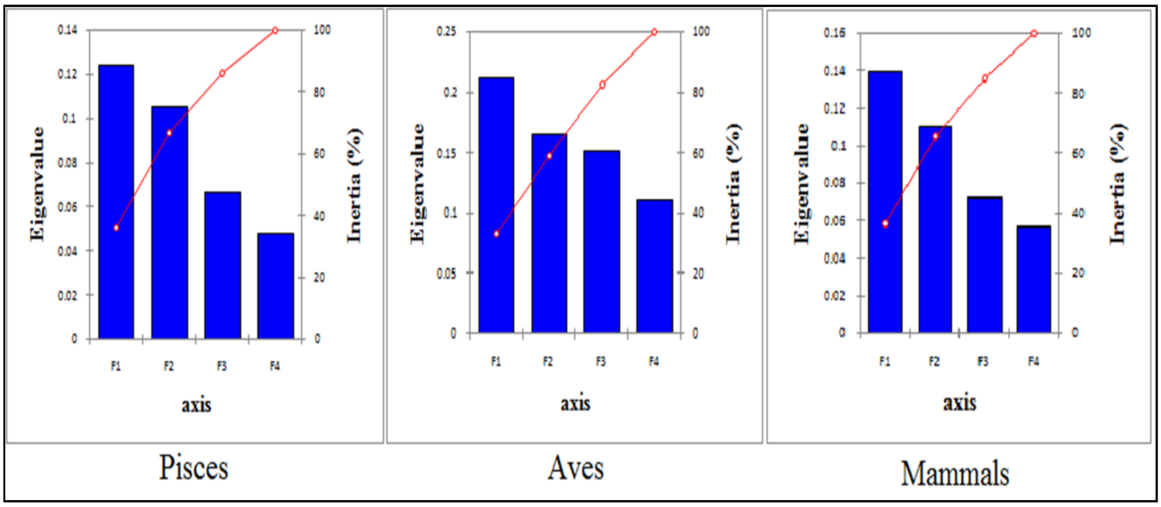
In ND5 gene, for pisces, the axis 1 accounted for 42.98% of the total variation; axis 2 accounted for 38.85% of the total variation. In aves, axis 1 accounted for 40.51% of the total variation and axis 2 accounted for 26.15% of the total variation and in mammals, axis 1 accounted for 32.84% of the total variation and axis 2 accounted for 30.86% of the total variation as shown in **Figure 4.1.54.**



Pisces Aves Mammals

Fig. 4.1.54 Correspondence analysis of synonymous codon usage towards the codons in ND5 gene for pisces, aves and mammals

In ND6 gene, the four main contributors *i.e.* axis1, axis2, axis3 and axis4 for pisces, aves and mammals were shown in **Figure 4.1.55.**



Pisces Aves Mammals

Figure 4.1.55 Contributions of the axes for MT-ND6 gene in pisces, aves and mammals

In ND6 gene, for pisces, the axis 1 accounted for 36.09% of the total variation; axis 2 accounted for 30.70% of the total variation. In aves, axis 1 accounted for 33.14% of the total variation and axis 2 accounted for 25.89% of the total variation and in mammals, axis 1 accounted for 36.83% of the total variation and axis 2 accounted for 29.04% of the total variation as shown in **Figure 4.1.56.**

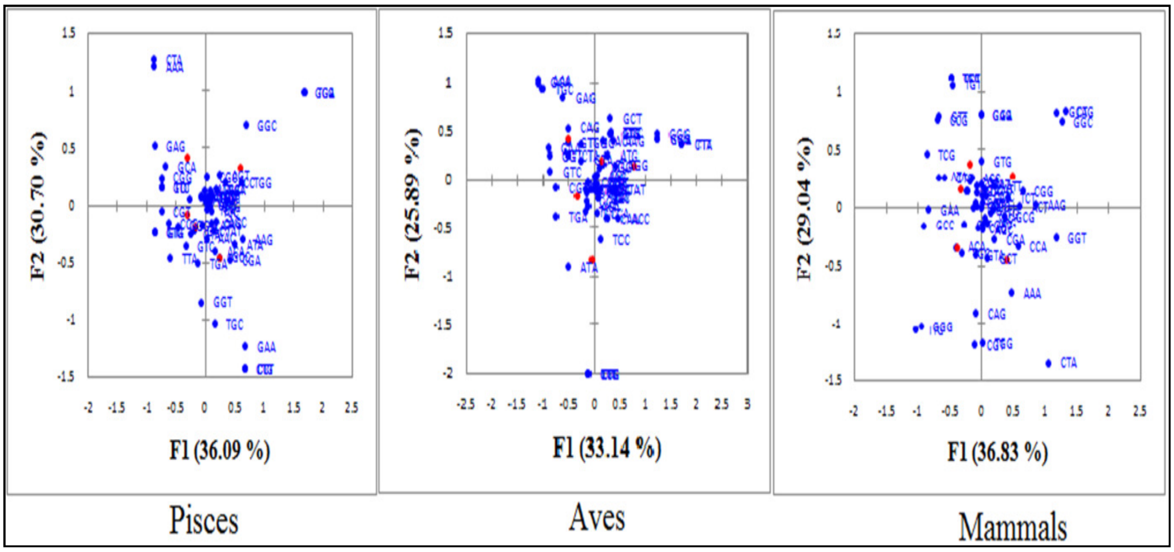


Figure 4.1.56 Correspondence analysis of synonymous codon usage towards the codons in ND6 gene for pisces, aves and mammals

In ND1, ND2, ND3, ND4, ND4l, ND5 and ND6 genes, the points in different plots were different, which indicate that codon usage patterns different in these genes. The positions of most codons are more close to axes with a concentrate distribution among pisces, aves and mammals, indicating that the base composition for mutation bias might correlate to the codon usage bias. A few codons are in a discrete distribution, indicating that other factors like natural selection played a role in codon usage of ND1, ND2, ND3, ND4, ND4L, ND5 and ND6 genes.

4.1.8 Amino acid composition of ND protein in different species of pisces, aves and mammals

In ND1 protein, the amino acid frequency of the encoded proteins in different species of pisces aves and mammals was estimated. The frequency of leucine residue was the highest in the amino acid composition of pisces, aves and mammals while cysteine, histidine, aspartate residue were lower in the proteins as shown in **Figure 4.1.57.**

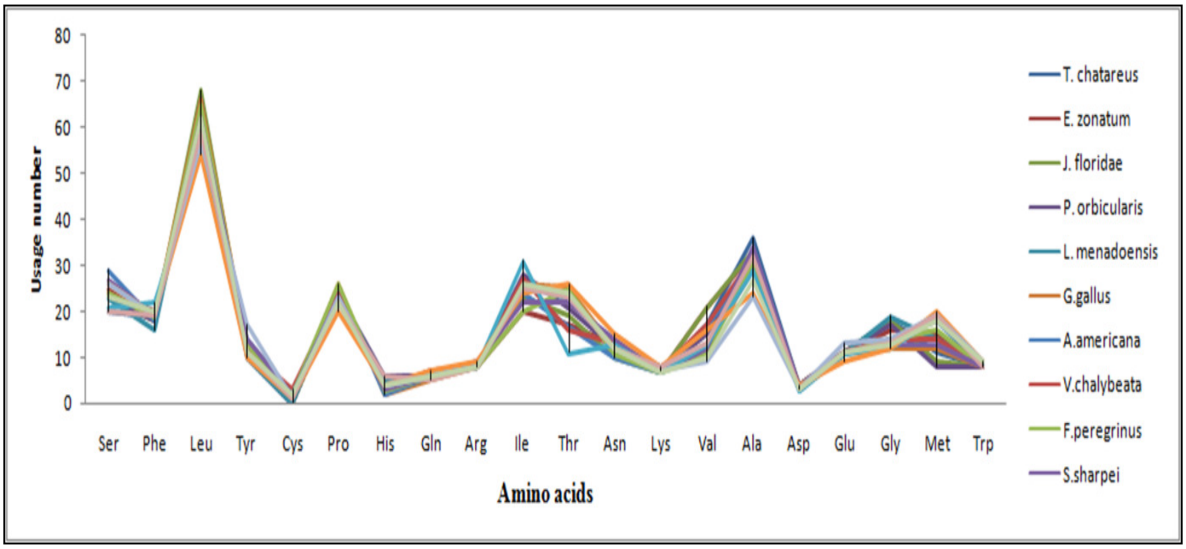


Figure 4.1.57 Amino acid distribution in different species of pisces, aves and mammals for ND1 protein

In ND2 protein, the usage of leucine in different species of pisces, aves and mammals was the highest. The usages of tyrosine, cysteine, histidine, glutamine, arginine, asparagine, valine, aspartate and glutamate were lower in the amino acid sequence as shown in **Figure 4.1.58.**

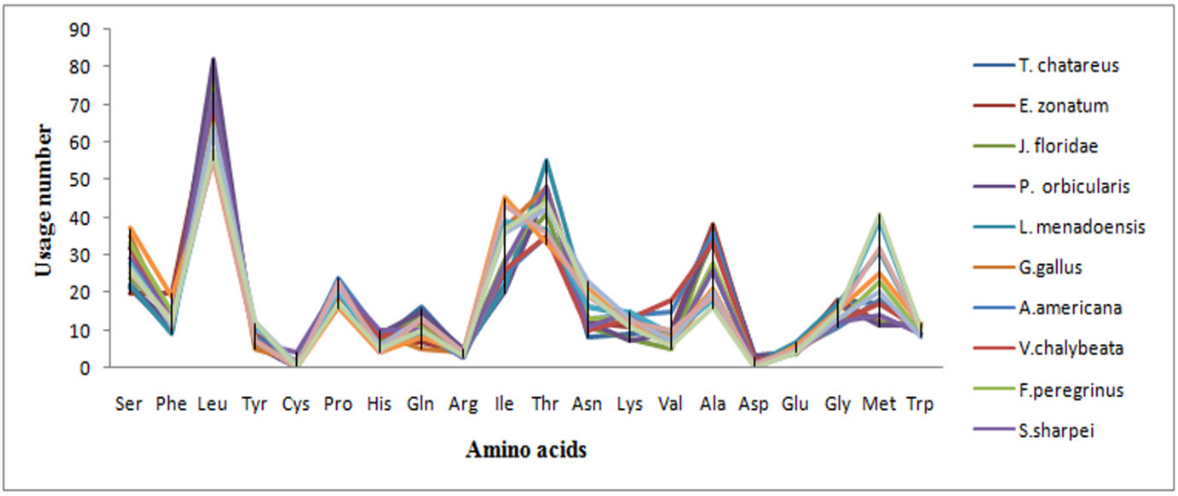


Figure 4.1.58 Amino acid distribution in different species of pisces, aves and mammals for ND2 protein

In ND3 protein, the usage of amino acid leucine was the highest in the amino acid sequence of ND3 protein. The usages of amino acids such as arginine, asparagine, aspartate, cysteine, glutamine, histidine, lysine, methionine, tyrosine and valine were the lowest in ND3 protein (**Figure 4.1.59**).

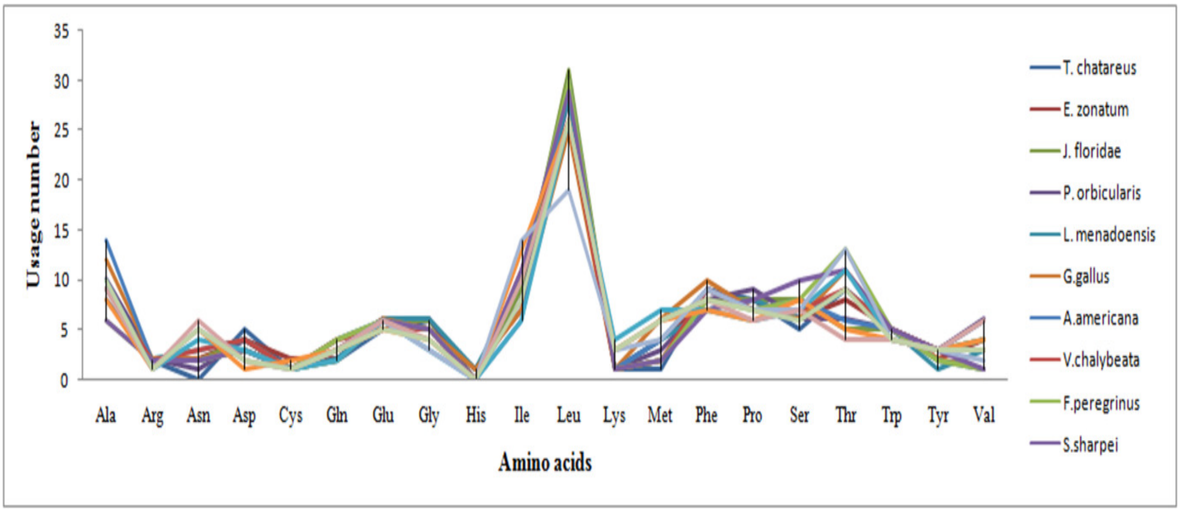


Figure 4.1.59 Amino acid distribution in different species of pisces, aves and mammals for ND3 protein

In ND4 protein, the usage of amino acid leucine was the highest while amino acids namely arginine, aspartate, cysteine, glutamine, glutamate, histidine, lysine, tyrosine and valine were lower in the amino acid sequence of protein (**Figure 4.1.60**).

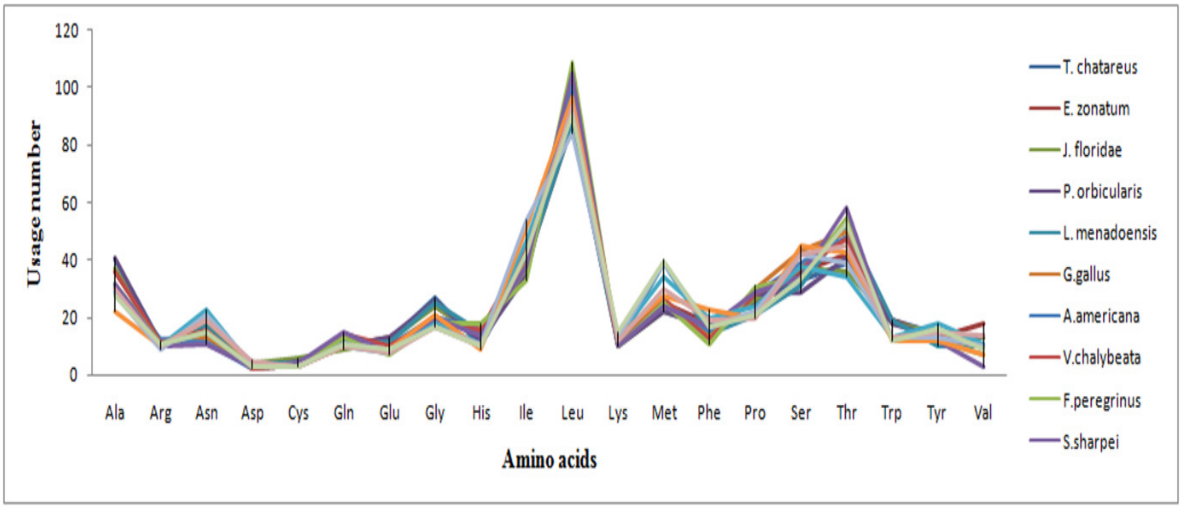


Figure 4.1.60 Amino acid distribution in different species of pisces, aves and mammals for ND4 protein

In ND4L protein, the usage of amino acid leucine was the highest while arginine, asparagine, asparatate, glutamine, histidine, isoleucine, lysine, proline, theonine, tryptophan, tyrosine and valine were lower in usage in the amino acid sequence of protein (**Figure 4.1.61**).

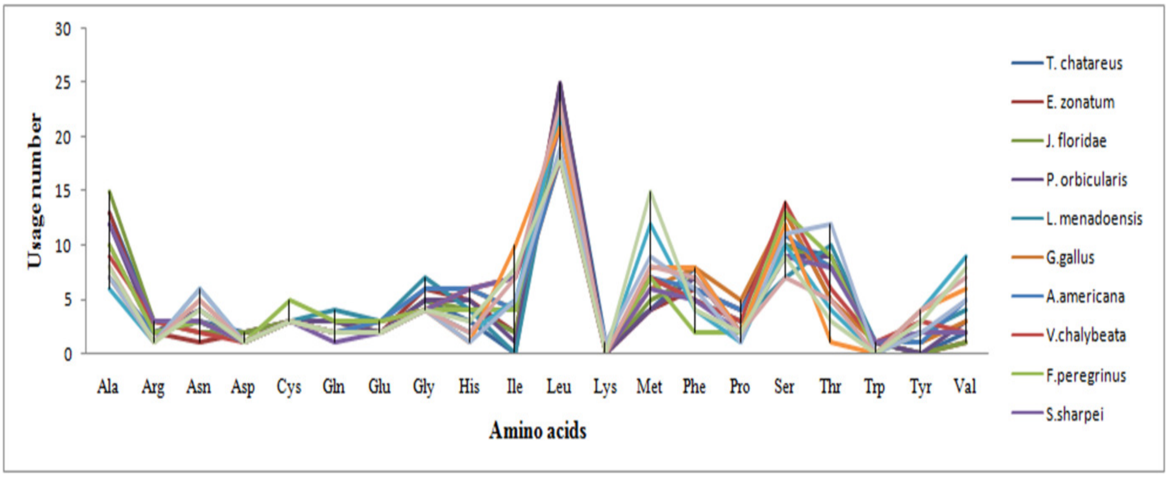


Figure 4.1.61 Amino acid distribution in different species of pisces, aves and mammals for ND4L protein

In ND5 protein, the usage of amino acid leucine was the highest while arginine, aspartate, cysteine, glutamate, histidine and valine were lower in the amino acid sequence of protein (**Figure 4.1.62**).

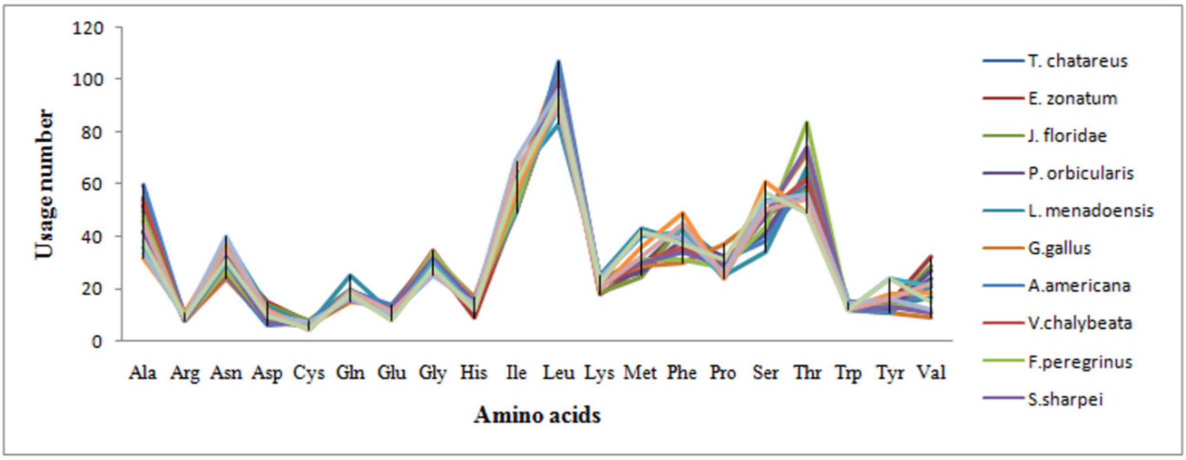


Figure 4.1.62 Amino acid distribution in different species of pisces, aves and mammals for ND5 protein

In ND6 protein, the amino acids such as asparagine, glutamine, histidine, lysine and proline were higher while alanine, aspartate, cysteine, glutamate, glycine, phenylalanine, tryptophan, tyrosine and valine were lower in usage in the amino acid sequence of protein (**Figure 4.1.63**).

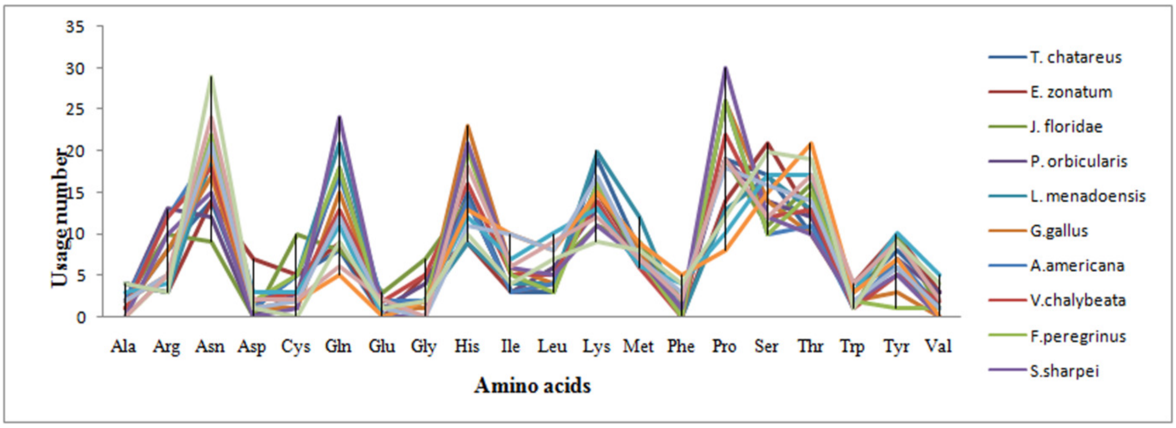


Figure 4.1.63 Amino acid distribution in different species of pisces, aves and mammals for ND6 protein

4.2 Analysis of codon usage bias in MT-CO gene among Pisces, Aves and Mammals

4.2.1 Codon usage pattern of CO gene in pisces, aves and mammals

To understand the pattern of random usage of synonymous codons in pisces, aves and mammals for MT-CO gene, RSCU of individual codons was calculated and compared. In COI gene, out of 60 codons, ACG, ATC, GTT, TTA and TGT encoding the amino acids thr, ile, val, leu, and cys respectively were used as over-represented codons in some species but as under-presented codons in most of the species as shown in **Figure 4.2.1**.

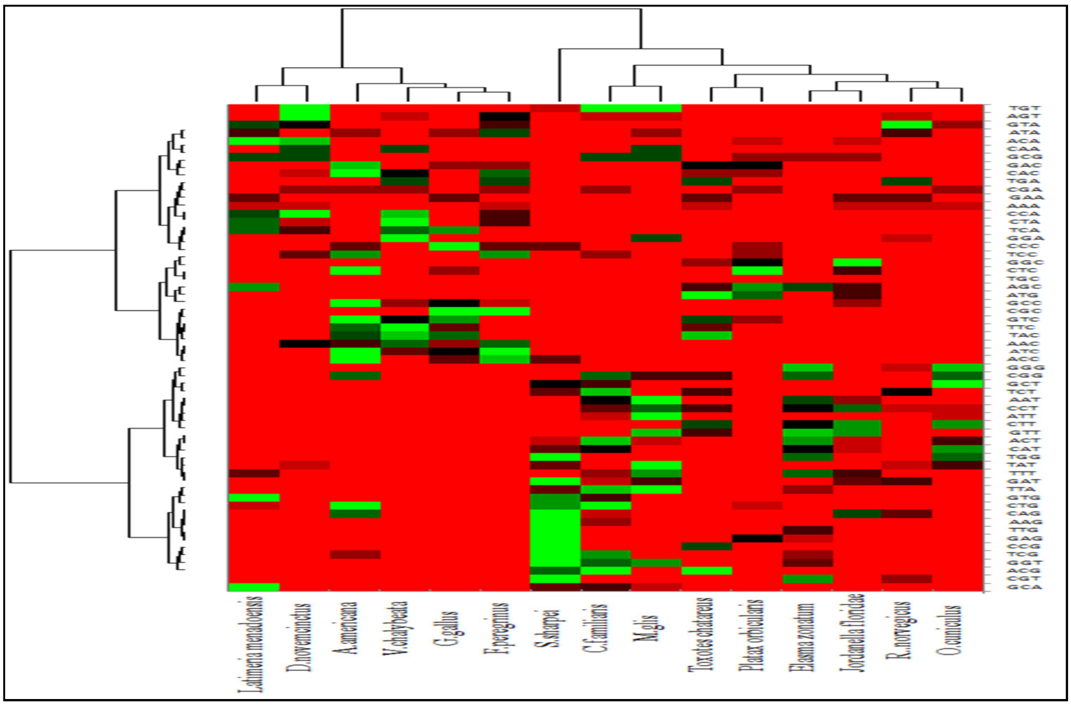


Figure 4.2.1 Hierarchial Clustering of RSCU values of each codon among pisces, aves and mammals for COI

In COII gene, AGT, GCA, GGG, TCC, TCG, TGG, TTG, CAT, CGT and CGG encoding the amino acids ser, ala, gly, ser, ser, trp, leu, his, arg, and arg respectively were used as over-represented codons in some species but as under-presented codons in most of the species as shown in **Figure 4.2.2**.

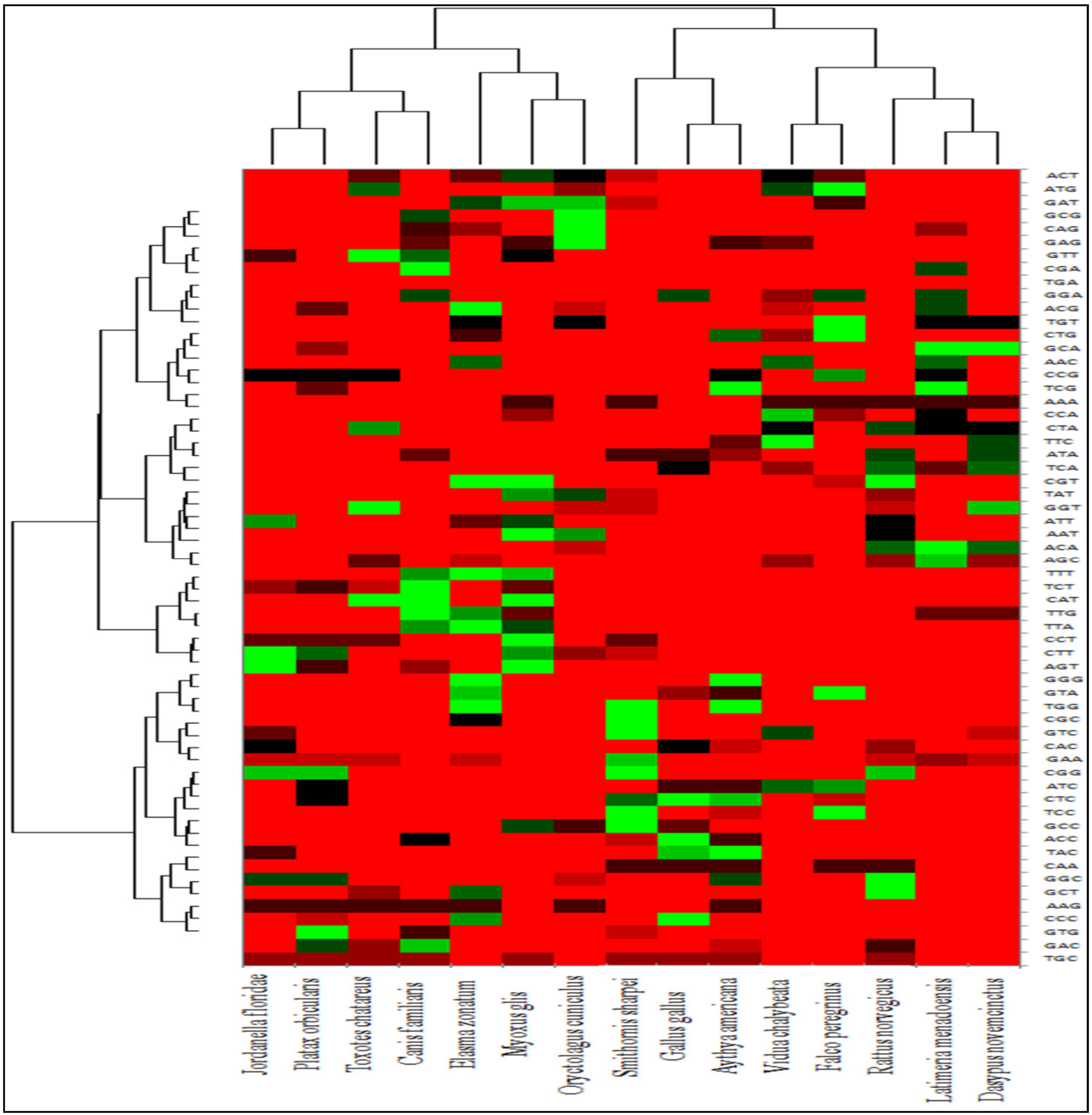


Figure 4.2.2 Hierarchial Clustering of RSCU values of each codon among pisces, aves and mammals for COII gene

In COIII gene, AAT, AGC, TGT, GCG, CAG and CCG encoding the amino acids asn, ser, cys, ala, gln and pro respectively were used as over-represented codons in some species but as under-presented codons in most of the species as shown in **Figure 4.2.3**.

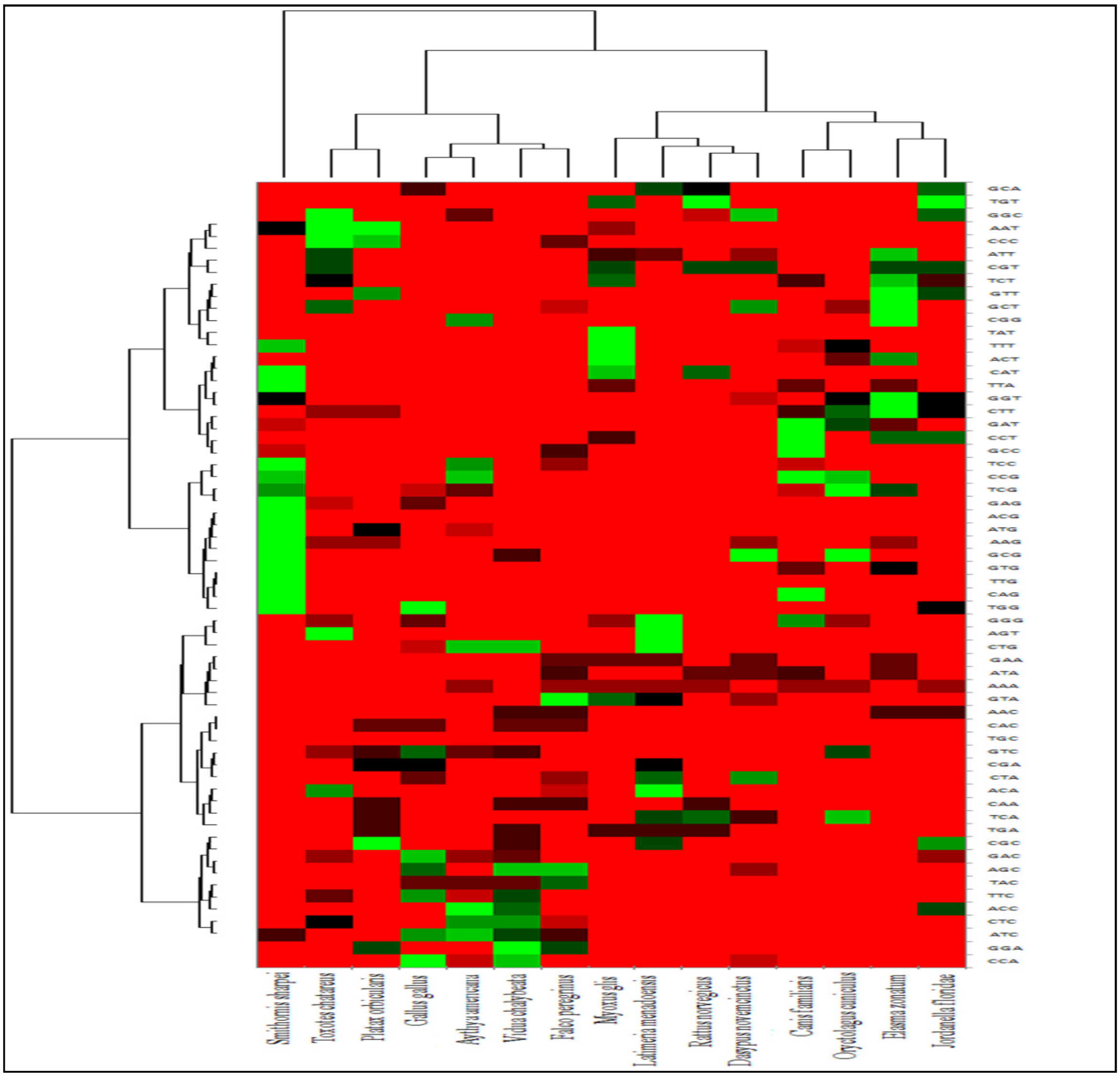


Figure 4.2.3 Hierarchial Clustering of RSCU values of each codon among pisces, aves and mammals for COIII gene

From hierarchial clustering, the over represented and the under represented codons were clearly evident. Further, it was observed that some codons were more frequently used in some species but less frequently used in other species. Moreover, some codons were under-represented in most of the species whereas less frequently used in other species. Based on RSCU analysis and nucleotide composition we deduced that the existence of preferred codons in coding sequences has been mostly influenced by compositional constraints, which account for the presence of mutation pressure.

4.2.2 Analysis of codon usage bias in mitochondrial CO

In COI gene, the ENC values in different species of pisces, aves and mammals were (Mean ±SD) 58.40±1.81, 59.80±0.44, 53.80±3.49 respectively. The mean ENC values in mammals were lower than aves and pisces which suggest that codon usage bias was high in different species of mammals in comparison to different species of aves and mammals for COI gene. From **Figure 4.2.4**, it was found that *M. glis* showed low ENC value in comparison to other species in COI gene which indicates high codon usage bias in *M. glis* in comparison to other species.

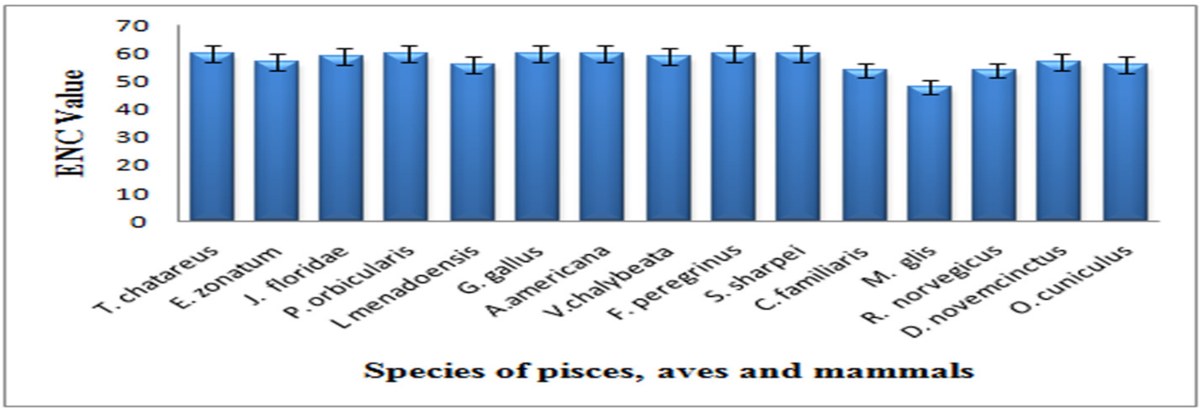


Figure 4.2.4 Distribution of ENC in different species in pisces, aves and mammals for MT-COI gene

In COII gene, the ENC values in different species of pisces, aves and mammals were (Mean ±SD) 56.8±1.92, 59.8±0.44, 53.6±3.13 respectively. The mammalian species showed low (average) ENC value than aves and pisces, which suggests that codon usage bias was high in different species of mammals for COII gene in comparison to different species of aves and mammals. From **Figure 4.2.5**, it was found that *M. glis* showed low ENC value in comparison to other species in COII gene which indicates high codon usage bias in *M. glis.*

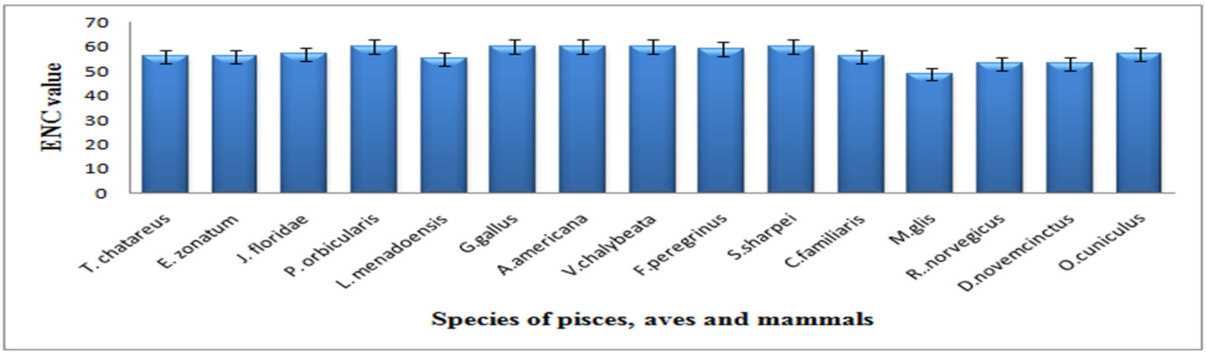


Figure 4.2.5 Distribution of ENC in different species in pisces, aves and mammals for MT-COII gene

In COIII gene, the ENC values in different species of pisces, aves and mammals were (Mean ±SD) 57.8±2.58, 59.8±0.44, 55.2±4.65 respectively. The mammalian species showed low (average) ENC value than aves and pisces, which suggests that codon usage bias was high in different species of mammals for COIII gene in comparison to different species of aves and pisces. From **Figure 4.2.6**, it was found that *M. glis* showed low ENC value as compared to other species in COIII gene which indicates high codon usage bias.

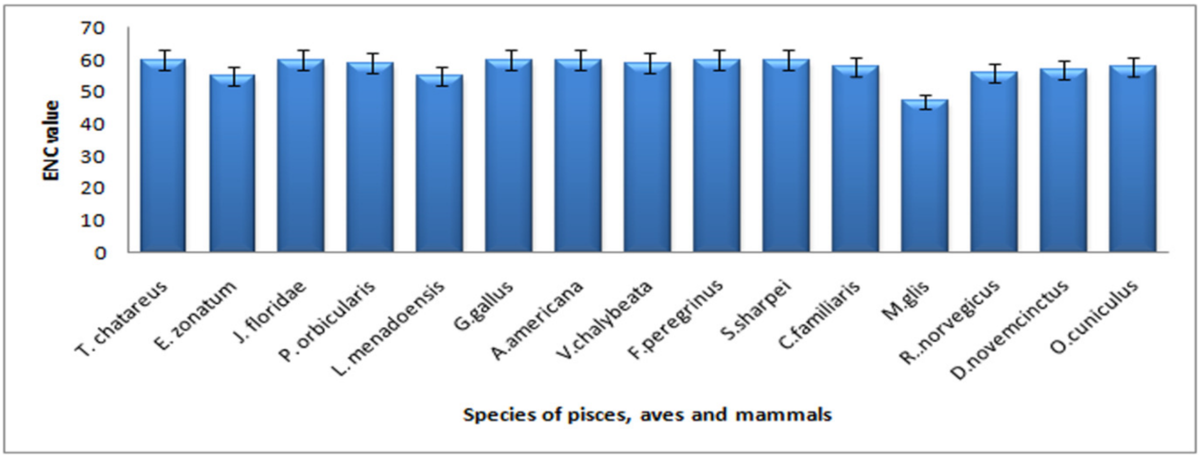


Figure 4.2.6 Distribution of ENC in different species in pisces, aves and mammals for MT-COIII gene

In COI, COII and COIII genes, the ENC value > 50 in different species of pisces, and aves indicates that codon usage bias was low and is apparently maintained at a stable level.

4.2.3 Prediction of expression level for mitochondrial CO gene

The CAI is a directional measure of codon usage bias and its higher value means higher gene expression. In many bacteria and small eukaryotes, highly expressed gene revealed stronger bias. CAI measures the gene expression level with respect to a reference set of genes [36]. In COI gene, the mean±SD of different species of pisces, aves and mammals were 0.87±0.02, 0.84±0.03, 0.86±0.03, which suggest that the expression level of COI in different species of pisces, aves and mammals was high. The expression level of *E. zonatum* and *C. familiaris* was high in comparison to others as shown in **Figure 4.2.7.**

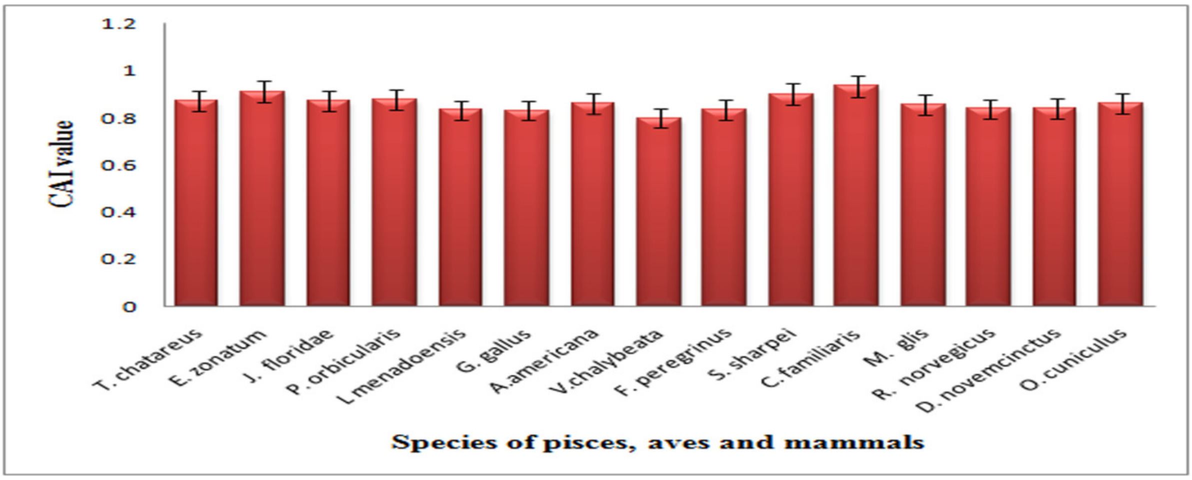


Figure 4.2.7 Distribution of CAI in different species of pisces, aves and mammals for COI gene

In COII gene, the mean±SD of different species, aves and mammals were 0.74±0.09, 0.61±0.01, 0.66±0.02, which suggest that the expression level of COII gene in different species of pisces, aves and mammals was high. The expression level of *E. zonatum* and *P. orbicularis* was high in comparison to others as shown in **Figure 4.2.8.**

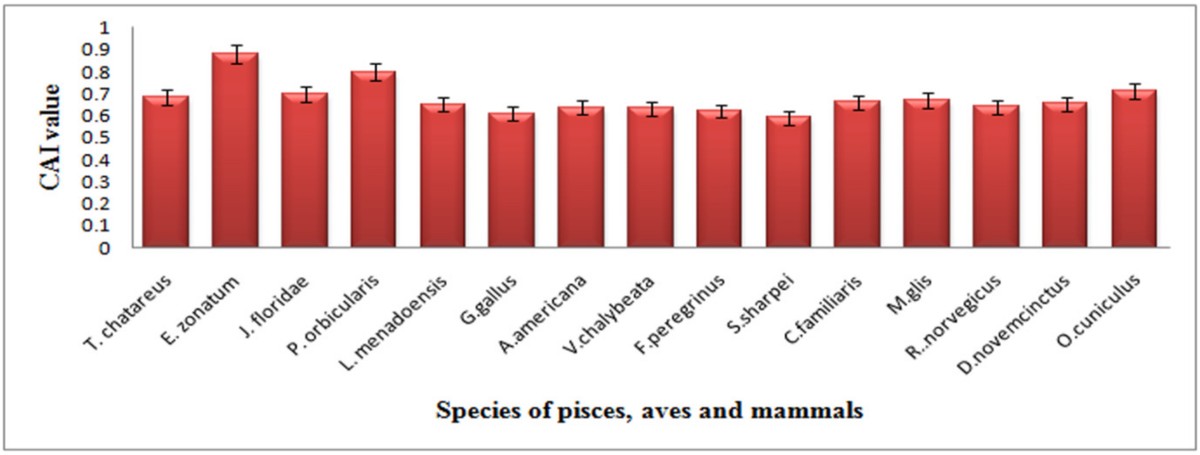


Figure 4.2.8 Distribution of CAI in different species of pisces, aves and mammals for COII gene

In COIII gene, the mean±SD of different species, aves and mammals were 0.73±0.10, 0.71±0.12, 0.73±0.06, which suggest that the expression level of COIII gene in different species of pisces, aves and mammals was high. The expression level of *J. floridae* and *A .americana* was high in comparison to others as shown in **Figure 4.2.9.**

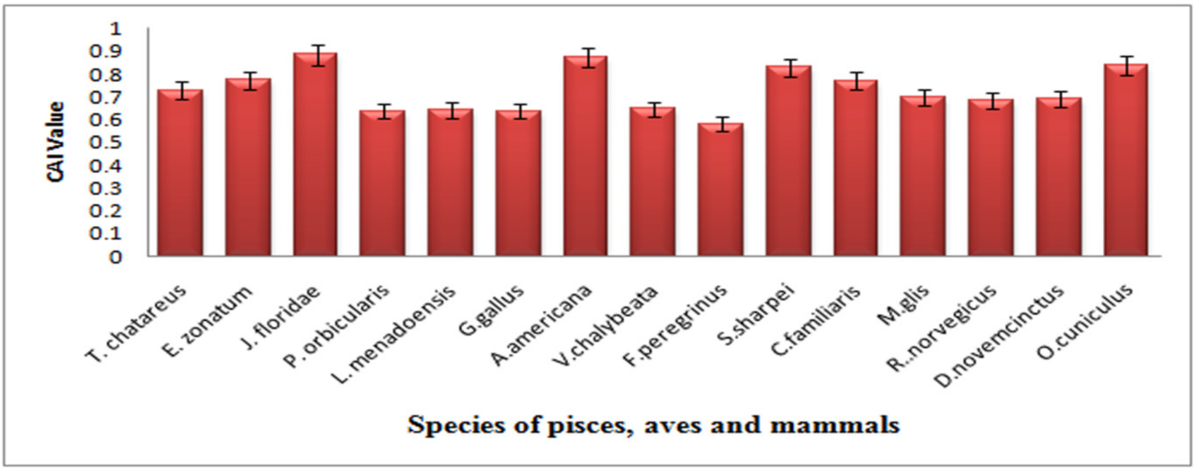


Figure 4.2.9 Distribution of CAI in different species of pisces, aves and mammals for COIII gene

4.2.4 Relationships of gene expression (CAI) with codon usage bias (ENC) for mitochondrial CO gene among pisces, aves and mammals

Comparison of ENC and CAI is used to judge the nucleotide composition and codon selection (Behura and Severson 2013)**.** We correlated ENC and CAI to understand the nucleotide composition variation and codon selection among different species of pisces, aves and mammals for CO gene. In COI gene, positive correlation was found between ENC and CAI in pisces, aves while negative correlation in mammals. In pisces, Pearson correlation coefficient was 0.235 (p>0.05), in aves 0.683 (p>0.05) and in mammals -0.041 (p>0.05) and these suggest codon usage bias had a very weak relationship with nucleotide composition in mammals in comparison to pisces and aves. In COII gene, positive correlation was found in all species pisces (0.369, p>0.05), aves (0.271, p>0.05) and mammals (0.429, p>0.05) which suggest that codon usage bias has a weak relationship with gene expression. ENC positively influenced the expression level. In COIII gene, positive correlation was found between ENC and CAI in pisces (0.328, p>0.05), aves (0.327, p>0.05) and mammals (0.437, p>0.05) which suggest that codon usage bias had a weak relationship with CAI and that ENC positively influenced the expression level of COIII gene.

4.2.5 Compositional features of mitochondrial CO gene among pisces, aves and mammals

Overall nucleotide composition may influence the codon usage bias of a genome (Jenkins and Holmes 2003). We therefore analyzed nucleotide composition of coding sequences of different species of pisces, aves and mammals.

In COI gene, the overall nucleotide composition and the composition at 3rd codon position were shown in **Table 4.2.1.** The nucleobase T and C occurred more frequently A and G in pisces, but in aves, C and A occurred more frequently than T and G while in mammals A and T occurred more frequently than C and G as shown in **Figure 4.2.10**. However, analysis of nucleotide composition at the 3rd position of codons (A3 %, T3 %, G3 % and C3 %) and GC1 %, GC2 %, GC 3 % provides a clear picture about the preference of codon usage in different species of pisces, aves and mammals. The mean±SD of A3 was the highest, followed by C3 and T3 in pisces and in aves, C3 was the highest, followed by A3 and T3 while in mammals A3 was the highest followed by T3 and C3. The G3 was the lowest in mammals followed by aves and pisces as shown in **Figure 4.2.10**.

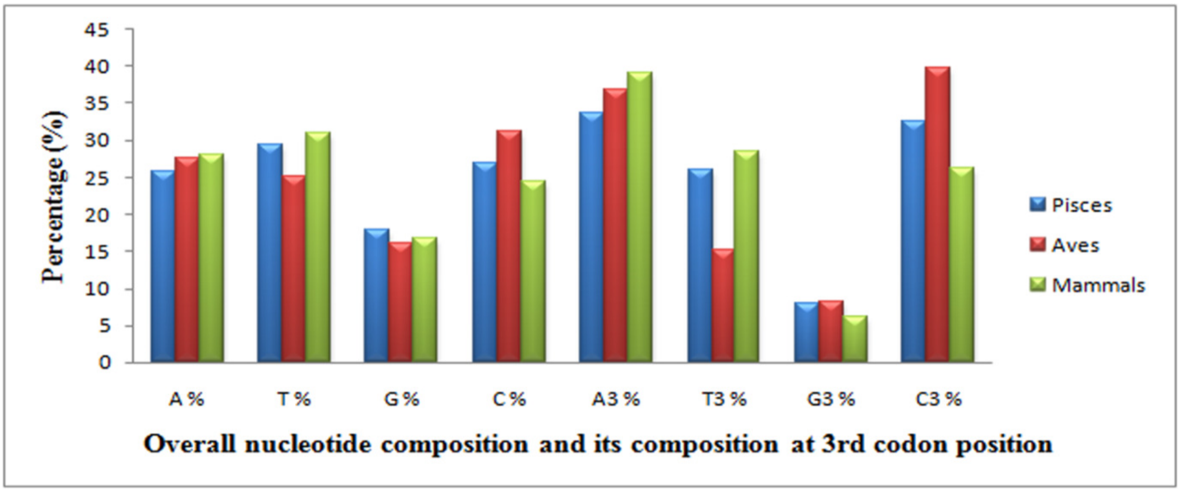


Figure 4.2.10 Nucleotide composition and its composition at 3rd codon position for MT-COI gene

The overall GC% was the highest aves, followed by mammals and then pisces and the GC content was lower than AT content *i.e.* the COI gene is AT rich. In pisces and mammals, the greatest difference of GC content was found for GC content between the 1st position and the 3rd position while in aves the greatest difference of GC content was found for GC content between the 1st position and the 2nd position as shown in **Figure 4.2.11.**

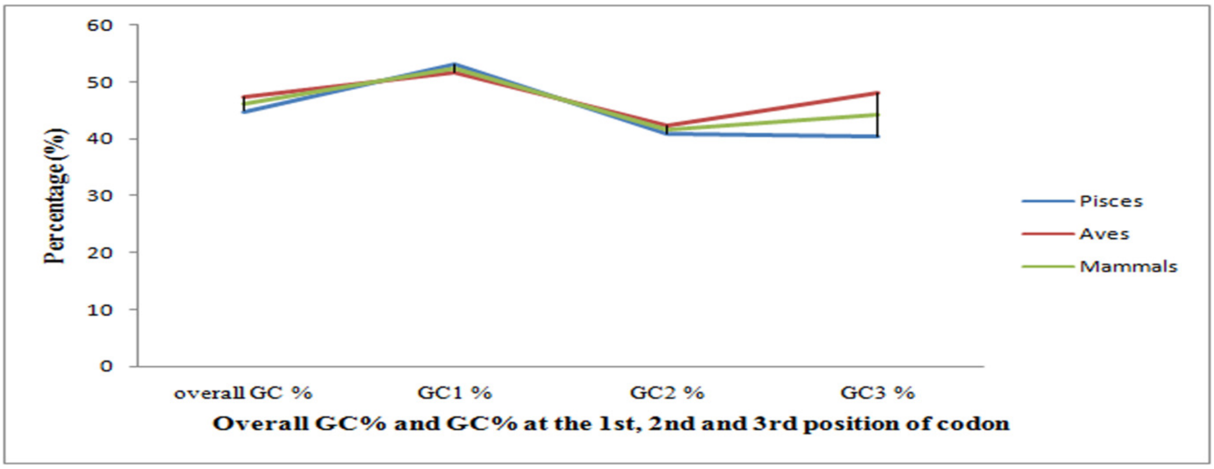


Figure 4.2.11 GC content and its content at codon's 1st, 2nd and 3rd position in pisces, aves and mammals for MT-COI gene

Table 4.2.1 Nucleotide compositions among pisces, aves and mammals for COI gene

Species	**A %**	**T %**	**G %**	**C %**	**A3 %**	**T3 %**	**G3 %**	**C3 %**
Toxotes chatareus	24.3	29.2	18.18	28.3	29.98	24.56	8.12	37.33
Elasma zonatum	24.37	32.68	18.24	24.69	29.78	33.65	9.67	26.88
Jordanella floridae	23.93	31.08	17.56	27.41	28.76	30.69	6.75	33.78
Platax orbicularis	25.14	27.85	18.05	28.94	32.1	22.82	8.12	36.94
Latimeria menadoensis	30.8	26.21	17.34	25.63	47.37	17.86	6.79	27.96
Mean± SD	25.70±2.8	29.40±2.5	17.87±0.40	26.99±1.79	33.59±7.79	25.91±6.30	7.89±1.20	32.57±4.92
Gallus gallus	27.32	25.58	15.89	31.2	37.59	16.08	3.68	42.63
Aythya americana	25.72	23.79	17.27	33.2	32.88	10.63	7.54	48.93
Vidua chalybeata	29.72	24.04	16.31	29.91	45.26	12.18	4.25	38.29
Falco peregrinus	28.49	25.08	15.4	31.01	40.81	14.89	2.9	41.39
Smithornis sharpei	26.61	27.13	15.05	31.2	27.51	22.28	22.86	27.32
Mean± SD	27.57±1.5	25.12±1.3	15.98±0.86	31.30±1.18	36.81±6.89	15.21±4.50	8.24±8.35	39.71±7.93
Canis familiaris	26.86	32.36	17.86	22.91	34.75	32.42	9.9	22.91
Myoxus glis	28.33	33.52	16.21	21.91	40.66	35.6	4.08	19.64
Rattus norvegicus	28.54	30.03	16.11	25.3	41.35	25.43	4.27	28.93
Dasypus novemcinctus	29.63	27.75	16.27	26.32	44.35	18.67	5.05	31.9
Oryctolagus cuniculus	26.392	30.99	17.38	25.22	34.82	29.96	7.39	27.82
Mean± SD	27.95±1.3	27.95±1.3	16.76±0.7	24.33±1.8	39.18±4.2	28.41±6.5	6.13±2.4	26.24±4.9

In COII gene, the overall nucleotide composition and its composition at 3rd codon position were shown in **Table 4.2.2.** The nucleobase A was the highest, followed by C and T, with G being the lowest in pisces. In aves, C % was the highest, followed by A and T , with G being the lowest while in mammals, A was the highest, followed by T and C, with G being the lowest as shown in **Figure 4.2.12**.

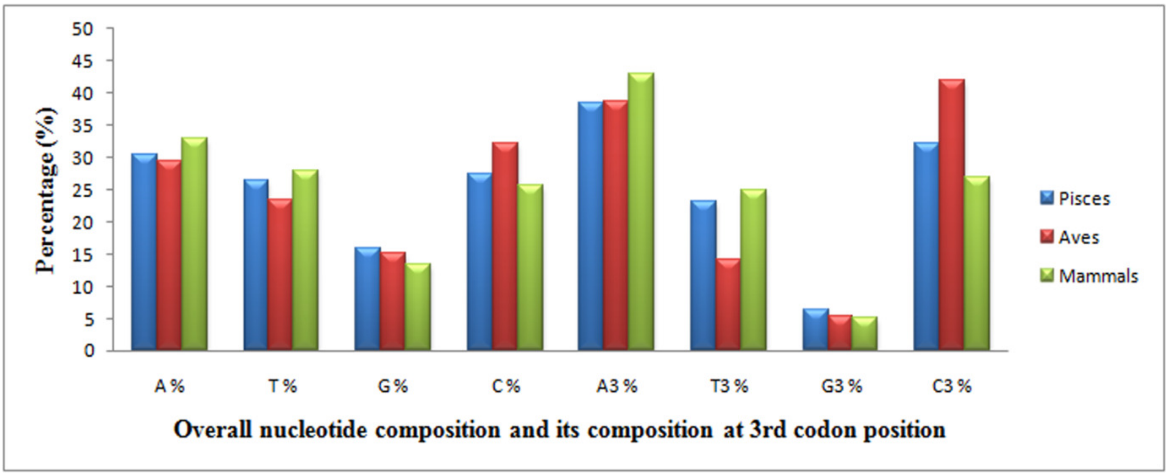


Figure 4.2.12 Nucleotide composition and its composition at 3rd codon position for COII gene

However, analysis of nucleotide composition at the 3rd position of codons (A3 %, T3 %, G3 % and C3 %) and GC1 %, GC2 %, GC 3 % provides a clear picture about the preference of codon usage in different species of pisces, aves and mammals. The nucleobase A3 was the highest, followed by C3 and T3 in pisces. The nucleobase C3 was the highest, followed by A3 and T3 in aves while in mammals, A3 was the highest followed by C3 and T3 (**Figure 4.2.12**). The nucleobase G3 was the lowest in mammals followed by aves and pisces as shown in **Figure 4.2.12.** The overall GC content was the highest in aves followed by pisces and then in mammals. In pisces, and aves, the greatest difference in GC content was found between the 1st and 2nd position of codon while in mammals, greatest difference was found between the 1st and 3rd position of codon as shown in **Figure 4.2.13.**

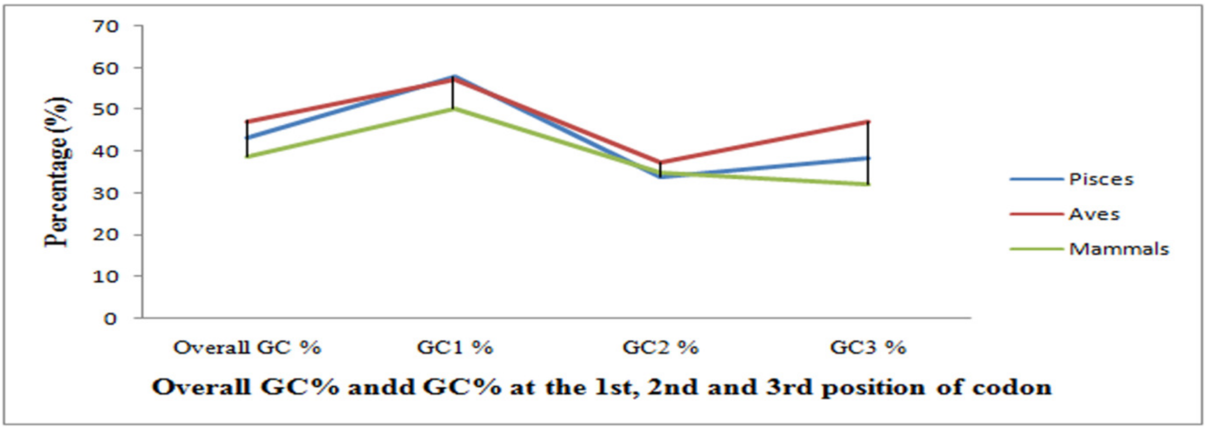


Figure 4.2.13 GC content and its content at codon's 1st, 2nd and 3rd position in pisces, aves and mammals for MT-COII gene

Table 4.2.2 Nucleotide compositions among pisces, aves and mammals for COII gene

Species	A %	T %	G %	C %	A3 %	T3 %	G3 %	C3 %
T. chatareus	29.87	27.12	15.87	27.12	36.79	27.27	5.62	30.3
E. zonatum	29.72	28.42	16.3	25.54	36.36	27.27	7.35	29
J. floridae	29.14	27.7	15.15	27.99	35.06	26.83	4.32	33.76
P. orbicularis	29	23.95	16.45	30.59	34.19	18.61	7.35	39.82
L. menadoensis	34.19	24.24	15.72	25.82	50.21	16.01	6.49	27.27
Mean±SD	30.38±2.2	26.28±2.0	15.89±0.5	27.41±2.0	38.52±6.6	23.19±5.4	6.22±1.3	32.03±4.9
G.gallus	29.38	22.8	14.47	33.33	37.71	13.15	3.94	45.17
A.americana	28.67	22.12	16.59	32.6	37.55	10.91	7.86	43.66
V.chalybeata	29.67	22.66	16.22	31.43	40.78	13.59	5.7	39.91
F.peregrinus	30.99	23.39	14.91	30.7	41.22	14.47	6.57	37.71
S.sharpei	28.5	25.43	13.74	32.3	35.96	18.42	2.63	42.98
Mean±SD	29.44±0.9	23.28±1.3	15.18±1.2	32.07±1.0	38.64±2.3	14.10±2.7	5.34±2.1	41.88±3.0
C.familiaris	32.89	27.63	13.74	25.73	42.54	22.36	6.14	28.94
M.glis	31.72	31.28	13.88	23.09	40.35	34.64	4.38	20.61
R..norvegicus	34.21	26.9	12.71	26.16	45.17	23.68	2.63	28.5
D.novemcinctus	35.23	25.43	12.71	26.6	50.87	17.98	3.07	28.07
O.cuniculus	30.4	28.65	14.32	26.6	35.52	26.31	9.21	28.94
Mean±SD	32.89±1.9	27.97±2.1	13.47±0.7	25.63±1.4	42.89±5.6	24.99±6.1	5.08±2.6	27.01±3.5

In COIII gene, the nucleotide composition and its composition at 3rd codon position were shown in **Table 4.2.3**. In pisces, nucleobase C was the highest, followed by T and A, with G being the lowest. In aves, C was the highest, followed by A and T, and G was the lowest while in mammals T was the highest followed by A, C, and G as shown in **Figure 4.2.14**. However nucleobase at the 3rd codon position revealed that nucleobase A was the highest followed by C and T in pisces. In aves, nucleobase C was the highest followed by A and T while in mammals, A was the highest followed by C and T. The nucleobase G was the lowest in mammals followed by pisces, and aves as shown in **Figure 4.2.14.**

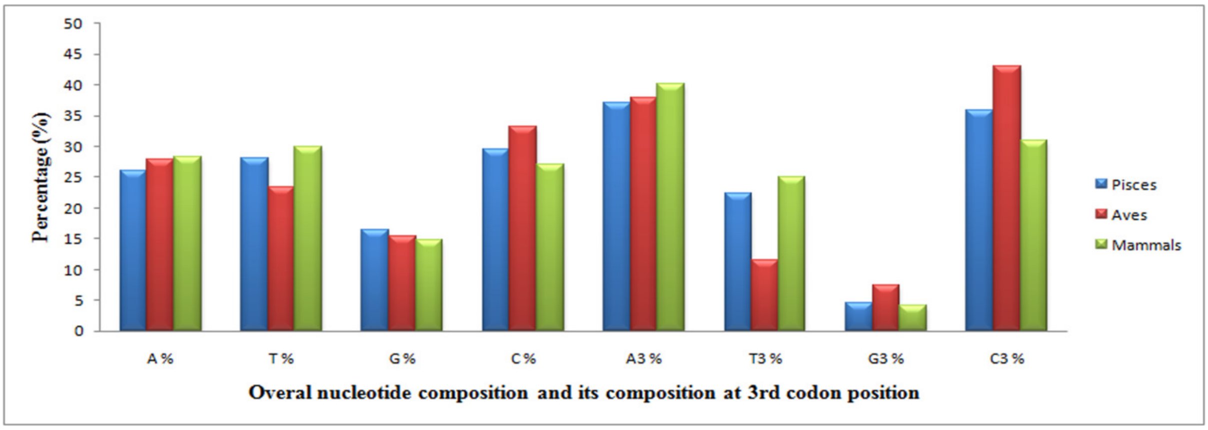


Figure 4.2. 14. Nucleotide composition and its composition at 3rd codon position for COIII gene

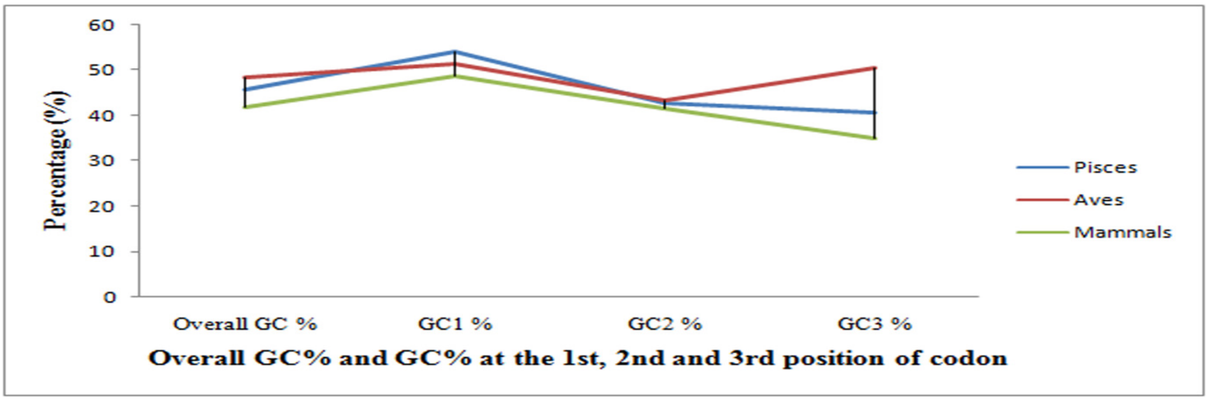


Figure 4.2.15 GC content and its content at codon's 1st, 2nd and 3rd position in pisces, aves and mammals for MT-COIII gene

The overall GC content was the lowest in mammals and pisces and aves. In pisces and mammals, the greatest difference of GC content was found between the 1st and 3rd position of codon while in aves, greatest difference of GC content was found between the 1st and 2nd position of codon as shown in **Figure 4.2.15.**

Table 4.2.3 Nucleotide compositions among pisces, aves and mammals for COIII gene

Species	A %	T %	G %	C %	A3 %	T3 %	G3 %	C3 %
T. chatareus	24.93	26.43	16.79	31.8	33.96	19.46	4.96	41.6
E. zonatum	24.55	32.31	16.92	26.2	32.06	33.58	6.48	27.86
J. floridae	24.17	28.88	16.41	30.53	31.67	23.66	4.58	40.07
P. orbicularis	26.97	26.08	16.15	30.78	40.07	17.17	3.05	39.69
L. menadoensis	30.02	26.71	15.52	27.73	47.7	17.93	4.19	30.15
Mean±SD	**26.12±2.42**	**28.08±2.60**	**16.35±0.55**	**29.40±2.34**	**37.09±6.81**	**22.36±6.75**	**4.65±1.24**	**35.87±6.36**
G.gallus	27.98	22.77	15.9	33.33	40.45	8.39	5.72	45.41
A.americana	27.09	23.02	16.15	33.71	37.4	8.39	6.1	48.09
V.chalybeata	27.22	22.51	16.28	33.96	40.83	8.39	4.19	46.56
F.peregrinus	28.38	23.53	14.37	32.69	43.89	10.3	1.14	44.65
S.sharpei	28.48	25.41	13.79	32.31	26.81	22.22	20.3	30.65
Mean±SD	**27.83±0.64**	**23.44±1.15**	**15.29±1.13**	**33.2±0.69**	**37.87±6.59**	**11.53±6.02**	**7.49±7.42**	**43.07±7.06**
C.familiaris	26.33	30.78	16.03	26.84	35.11	25.95	6.87	32.06
M.glis	29.24	33.46	14.04	23.24	41.76	34.86	3.06	20.3
R..norvegicus	30.02	28.24	13.74	27.98	44.27	19.46	1.9	34.35
D.novemcinctus	28.73	28.35	14.81	28.09	42.52	20.3	3.44	33.71
O.cuniculus	27.35	28.75	15.26	28.62	36.64	24.04	5.34	33.96
Mean±SD	**28.33±1.48**	**29.91±2.23**	**14.77±0.92**	**26.95±2.17**	**40.6±3.96**	**24.92±6.16**	**4.12±1.97**	**30.87±5.97**

4.2.6 Interrelationships among different compositional features of MT-CO gene among pisces, aves and mammals

Two major evolutionary factors namely mutation pressure and natural selection shape the codon usage pattern in a species. Mutation pressure affects the whole genome, which accounts for the majority of codon usage among different RNA viruses. We performed correlation analysis between general nucleotide composition and nucleotide composition at 3rd codon position to determine whether evolutionary process is driven by mutation pressure alone or by both mutation pressure and natural selection.

In COI gene, highly significant positive correlation was found between A and A3, T and T3, G and G3, C and C3, C and GC3, GC and GC3 in pisces. In aves, significant positive correlation was found between T and T3, C and GC3 while significant negative correlation was found between A and GC3. In mammals, significant positive correlation was found between A and A3, T and T3 and C and GC3 while significant negative correlation was found between A and GC3 as shown **in Table 4.2.4.** These results suggest that the compositional constraint, arising from mutation pressure and natural selection, determines the pattern of codon usage in COI gene.

Table 4.2.4 Correlation between overall nucleotide composition (%) and its composition at 3rd codon position in pisces, aves and mammals for COI gene

	Nucleotide	**A3 %**	**T3 %**	**G3 %**	**C3 %**	**GC3%**
Pisces	A %	1.000**	-0.774	-0.463	-.0480	-0.600
	T %	-0.771	0.994**	0.571	-0.192	-0.055
	G %	-0.689	0.463	0.886*	0.282	0.509
	C %	-.0352	-0.280	-0.270	0.982**	0.929*
	GC %	-0.469	-0.174	-0.080	0.986**	0.981**
Aves	A %	0.874	-0.189	-0.478	-0.148	-0.931*
	T %	-0.606	0.994**	0.731	-0.807	-0.057
	G %	0.167	-0.861	-0.426	0.792	0.493
	C %	-0.596	-0.243	0.120	0.529	0.937*
	GC %	-0.325	-0.557	-0.111	0.715	0.856
Mammals	A %	0.990**	-0.639	-0.747	0.377	-0.931*
	T %	-0.606	0.994**	0.731	-0.807	-0.057
	G %	0.167	-0.861	-0.426	0.792	0.493
	C %	-0.596	-0.243	0.120	0.529	0.937*
	GC %	-0.325	-0.557	-0.111	0.715	0.856

Note: One asterisk (*) and two asterisks (**) indicate correlations at a level of 0.05 and 0.01, respectively.

Furthermore, significant correlations among ENC and various GC contents as shown in **Table 4.2.5**, suggest that the nucleotide composition under mutation pressure and natural selection affect the synonymous codon usage in MT-COI gene for pisces, aves and mammals.

Table 4.2.5 Correlation coefficients among ENC, CAI, GC contents for COI gene

		Pisces		**Aves**		**Mammals**	
SL No	Correlation between	Correlation coefficient	p value	Correlation coefficient	p value	Correlation coefficient	p value
1	ENC and GC %	0.961**	0.009	0.337	0.579	0.987**	0.002
2	ENC and GC3 %	0.980**	0.003	0.543	0.345	0.998**	0
3	GC and GC3 %	0.981**	0.003	0.856	0.064	0.982**	0.003
4	GC1 and GC3 %	0.707	0.182	-0.250	0.685	0.639	0.246
5	GC2 and GC3%	-0.506	0.385	-0.042	0.946	-0.902*	0.036
6	GC12 and GC3 %	0.573	0.313	-0.301	0.616	0.305	0.618

Note: One asterisk (*) and two asterisks (**) indicate correlations at a level of 0.05 and 0.01, respectively

If the pattern of synonymous codon usage is solely governed by mutation pressure then the frequency of nucleotide A and T should be equal to that of G and C at synonymous 3rd codon position (Zhang Zhicheng *et al.* 2013). In COI gene, the frequencies of those nucleotides were not same in pisces, aves and mammals indicating that other factor such as natural selection might have played a role in codon usage pattern.

To quantify the extent of natural selection and mutation pressure in the codon usage pattern of MT-COI gene, we drew the neutrality plot between GC12 and GC3. GC12 is the average of GC1 and GC2. The regression coefficient of GC12 on GC3 in pisces is 0.075 indicating relative neutrality of 7.5 % and relative constraint of 92.5 % for GC3 % in pisces as shown in **Figure 4.2.16 (a).** The GC12 was influenced by mutation pressure and natural selection with a ratio of 0.075/0.925= 0.081. In aves, the regression coefficient of GC12 on GC3 was 0.084 indicating relative neutrality 8.4% (absolute) and relative constraint is 91.6% as shown in **Figure 4.2.16 (b)**. The GC12 was influenced by mutation pressure and natural selection with a ratio of 0.084/0.916= 0.0917. In mammals, the regression coefficient of GC12 on GC3 was 0.033 indicating relative neutrality of 3.3 % and relative constraint of 96.7 % for GC3 % as shown in **Figure 4.2.16 (c).** The GC12 was influenced by mutation pressure and natural selection with a ratio of 0.033/0.967= 0.034. These results suggest that natural selection played a major role while mutation pressure played a minor role in shaping codon usage pattern in COI gene among pisces, aves and mammals.

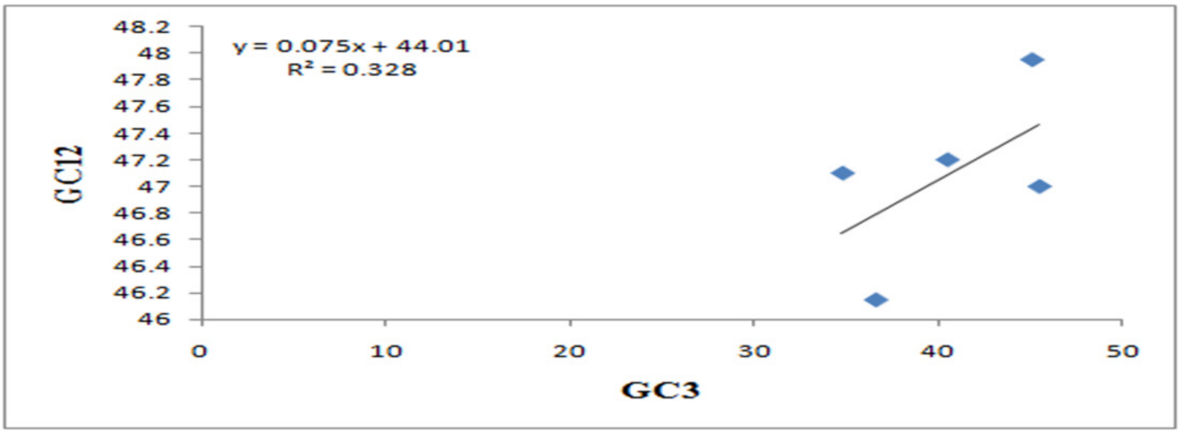


Figure 4.2.16 Neutrality plot of GC12 with GC3 in different species of pisces for COI gene

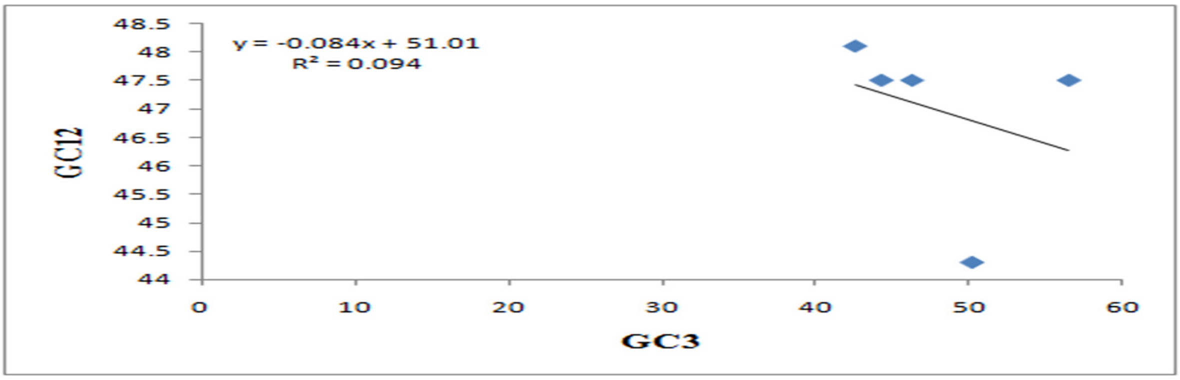


Figure 4.2.16 (b) Neutrality plot of GC12 with GC3 in different species of aves for COI gene

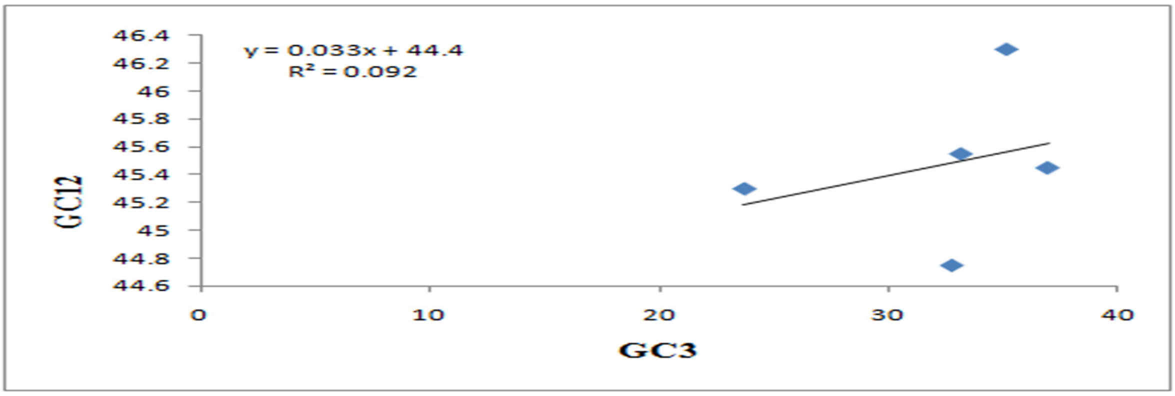


Figure 4.2.16 (c) Neutrality plot of GC12 with GC3 in different species of mammals for COI gene

In COII gene, highly significant positive correlation was found between A and A3, T and T3, G and G3, C and C3, C and GC3, GC and GC3 in pisces. In aves, significant positive correlation was found between A and A3, T and T3, C and C3, GC and GC3 while in mammals, significant positive correlation was found between A and A3, T and T3, G and G3, C and C3 and GC and C3 as shown in **Table 4.2.6.** These results suggest that the compositional constraint, arising from mutation pressure and natural selection, determine the pattern of codon usage in COII gene.

Table 4.2.6 Correlation between overall nucleotide composition (%) and its composition at 3rd codon position in pisces, aves and mammals for COII gene

	Nucleotide	**A3 %**	**T3 %**	**G3 %**	**C3 %**	**GC3%**
Pisces	A %	1.000**	-0.643	0.125	-0.657	-0.592
	T %	-0.455	0.951*	-0.376	-0.341	-0.418
	G %	-0.199	-0.218	0.934*	0.263	0.481
	C %	-0.554	-.0222	0.012	0.978**	0.932*
	GC %	-0.546	-0.270	0.231	0.963**	0.973**
Aves	A %	0.881*	-.0100	0.325	-0.795	-0.573
	T %	-0.438	0.979**	-0.765	-0.033	-0.565
	G %	0.417	-0.813	0.854	-0.162	0.430
	C %	-0.783	-0.184	-0.348	0.994**	0.760
	GC %	-0.218	-0.791	0.449	0.574	0.889*
Mammals	A %	0.982**	-0.689	-0.845	0.258	-0.276
	T %	-0.723	0.980**	0.335	-0.789	-0.386
	G %	-0.917*	0.595	0.885*	-0.229	0.319
	C %	0.246	-0.854	0.168	0.953*	0.780
	GC %	0.207	-0.872	0.255	0.980**	0.847

Note: One asterisk (*) and two asterisks (**) indicate correlations at a level of 0.05 and 0.01, respectively

In addition, highly significant correlation was found between ENC and various GC contents as shown in **Table 4.2.7**, which suggests that the nucleotide composition under mutation pressure and natural selection affect the synonymous codon usage in MT-COII gene for pisces, aves and mammals.

Table 4.2.7 Correlation coefficients among ENC, CAI, GC contents for COII gene

		Pisces		Aves		Mammals	
SL No	Correlation between	Correlation coefficient	p value	Correlation coefficient	p value	Correlation coefficient	p value
1	ENC and GC %	0.973^{**}	0.005	-0.201	0.746	0.877	0.051
2	ENC and GC3 %	0.996^{**}	0	-0.352	0.561	0.994^{**}	0.001
3	GC and GC3 %	0.973^{**}	0.005	0.889^{*}	0.044	0.847	0.070
4	GC1 and GC3 %	0.311	0.610	0.975^{**}	0.005	-0.260	0.673
5	GC2 and GC3%	0.955^{*}	0.011	-0.191	0.759	-.0862	0.060
6	GC12 and GC3 %	0.66	0.226	0.415	0.487	-0.633	0.251

Note: One asterisk (*) and two asterisks (**) indicate correlations at a level of 0.05 and 0.01, respectively

In COII gene, the frequencies of those nucleotides were not same in each of pisces, aves and mammals indicating that other factor such as natural selection might have played a role in codon usage pattern.

To quantify the extent of natural selection and mutation pressure in the codon usage pattern of MT-COII, we drew neutrality plot between GC12 and GC3. GC12 is the average of GC1 and GC2. The regression coefficient of GC12 on GC3 is 0.118 indicating relative neutrality of 11.8 % and relative constraint of 88.2 % for GC3 % in pisces as shown in **Figure 4.2.17 (a).** The GC12 was influenced by mutation pressure and natural selection with a ratio of 0.118/0.882= 0.133. In aves, the regression coefficient of GC12 on GC3 was 0.150 indicating relative neutrality 15% and relative constraint 85% in shown in **Figure 4.2.17 (b)**. The GC12 was influenced by mutation pressure and natural selection with a ratio of 0.150/0.85= 0.176. In mammals, the regression coefficient of GC12 on GC3 was 0.085 indicating relative neutrality of 8.5% (absolute) and relative constraint of 91.5 % for GC3 % as shown in **Figure 4.2.17 (c).** The GC12 was influenced by mutation pressure and natural selection with a ratio of 0.085/0.915= 0.092. These results suggest that natural selection played a major role while mutation pressure played a minor role in shaping codon usage pattern in COII gene among pisces, aves and mammals.

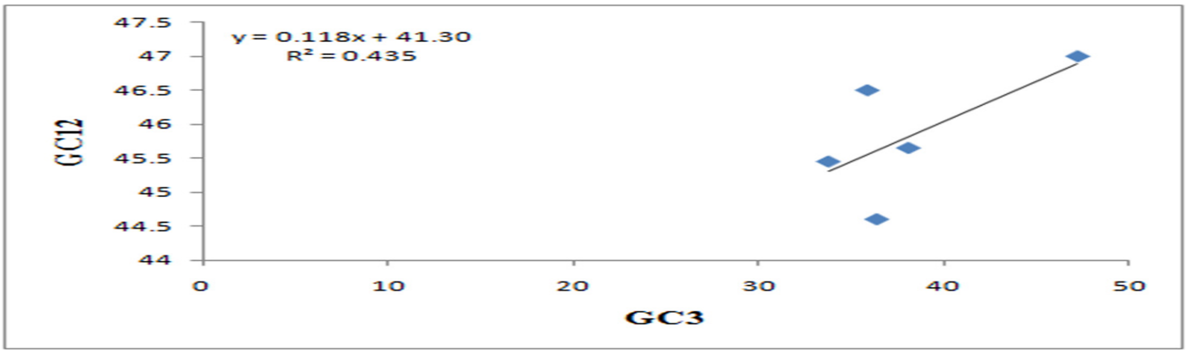


Figure 4.2.17(a) Neutrality plot of GC12 with GC3 in different species of pisces for COII gene

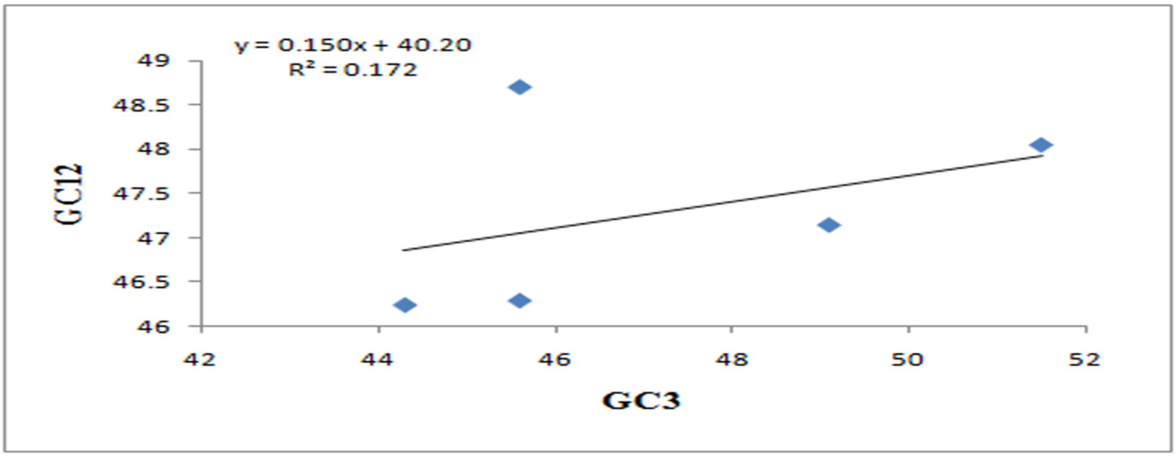


Figure 4.2.17(b) Neutrality plot of GC12 with GC3 in different species of aves for COII gene

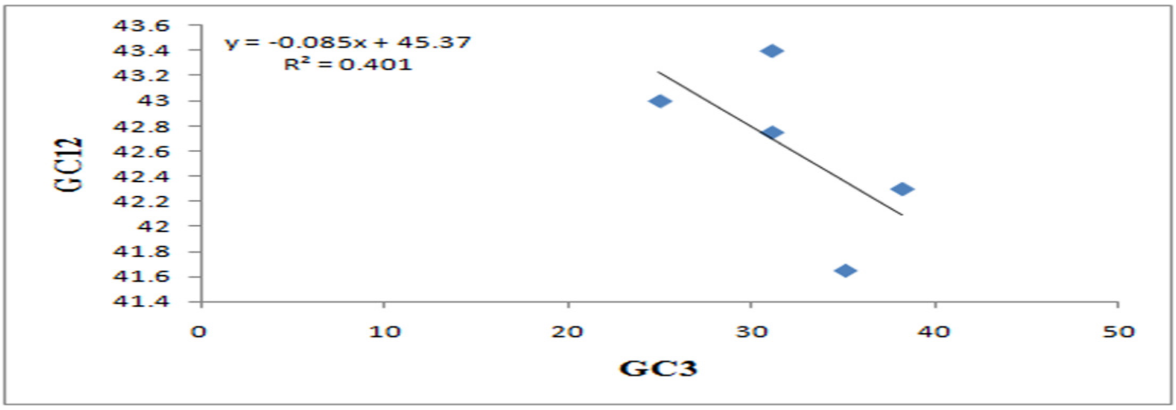


Figure 4.2.17(c) Neutrality plot of GC12 with GC3 in different species of mammals for COII gene

In COIII gene**,** highly significant positive correlation was found between A and A3, T and T3, C and C3, C and GC3, GC and C3, GC and GC3 in pisces while negative correlation was found between G and A3. In aves, significant positive correlation was found between T and T3, while significant negative correlation was found between T and C3. In mammals, significant positive correlation was found between A and A3, T and T3, G and G3, C and C3, C and GC3, GC and C3, GC and GC3 as shown in **Table 4.2.8.** These results suggest that the compositional constraint, arising from mutation pressure and natural selection, determines the pattern of codon usage in COIII gene.

Table 4.2.8 Correlation between overall nucleotide composition (%) and its composition at 3rd codon position in pisces, aves and mammals for COIII gene

	Nucleotide	A3 %	T3 %	G3 %	C3 %	GC3%
Pisces	A %	0.998^{**}	-0.585	-0.516	-.0348	-0.494
	T %	-0.576	0.991^{**}	0.847	-0.600	-0.473
	G %	-0.916^{*}	0.656	0.673	0.154	0.315
	C %	-0.174	-0.650	-0.567	0.988^{**}	0.960^{**}
	GC %	-0.367	-0.480	-0.391	0.979^{**}	0.989^{**}
Aves	A %	-0.290	0.645	0.397	-0.696	-0.653
	T %	-0.834	0.976^{**}	0.857	$-.0954^{*}$	-0.120
	G %	0.504	-0.824	-0.559	0.820	0.539
	C %	0.500	-0.798	-0.566	0.808	0.497
	GC %	0.509	-0.820	-0.570	0.822	0.521
Mammals	A %	0.973^{**}	-0.088	-0.998^{**}	-0.226	-0.497
	T %	-0.146	0.973^{**}	0.113	$-.0943^{*}$	-0.820
	G %	-0.913^{*}	-0.042	0.972^{**}	0.328	0.582
	C %	-0.126	-0.921^{*}	0.152	0.982^{**}	0.933^{*}
	GC %	-0.436	-0.780	0.479	0.935^{*}	0.986^{**}

Note: One asterisk (*) and two asterisks (**) indicate correlations at a level of 0.05 and 0.01, respectively

In addition, highly significant correlation was found between ENC and various GC contents in pisces, aves and mammals as shown in **Table 4.2.9**, which suggests that the nucleotide composition under mutation pressure and natural selection affects the synonymous codon usage in MT-COIII gene for pisces, aves and mammals.

Table 4.2.9 Correlation coefficients among ENC, CAI, GC contents for COIII gene

		Pisces		Aves		Mammals	
SL No	Correlation between	Correlation coefficient	p value	Correlation coefficient	p value	Correlation coefficient	p value
1	ENC and GC %	0.969^{**}	0.007	-0.421	0.480	0.985^{**}	0.002
2	ENC and GC3 %	0.992^{**}	0.001	-0.670	0.216	0.999^{**}	0
3	GC and GC3 %	0.989^{**}	.0001	0.521	0.368	0.986^{**}	0.002
4	GC1 and GC3 %	0.867	0.057	-0.018	0.977	0.830	0.082
5	GC2 and GC3%	0.106	0.865	-0.066	0.916	-0.258	0.675
6	GC12 and GC3 %	0.829	0.082	-0.32	0.959	0.645	0.24

Note: One asterisk (*) and two asterisks (**) indicate correlations at a level of 0.05 and 0.01, respectively.

In COIII gene, the frequencies of those nucleotides *i.e.* A and T were not equal to G and C at the 3rd position of codon in each of pisces, aves and mammals indicating that other factor such as natural selection could play a role in codon usage pattern.

To quantify the extent of natural selection and mutation pressure in the codon usage pattern of MT-COIII, we drew the neutrality plot between GC12 and GC3. GC12 is the average of GC1 and GC2. The regression coefficient of GC12 on GC3 is 0.127 indicating relative neutrality of 12.7 % and relative constraint of 87.3 % for GC3 % in pisces as shown in **Figure 4.2.18 (a).** The GC12 was influenced by mutation pressure and natural selection with a ratio of 0.127/0.873= 0.145. In aves, the regression coefficient of GC12 on GC3 was 0.024 indicating relative neutrality 2.4% (absolute) and relative constraint 97.6 % as shown in **Figure 4.2.18 (b)**. The GC12 was influenced by mutation pressure and natural selection with a ratio of 0.024/0.976= 0.024. In mammals, the regression coefficient of GC12 on GC3 was 0.082 indicating relative neutrality of 8.2% and relative constraint of 91.8 % for GC3 % as shown in **Figure 4.2.18 (c).** The GC12 was influenced by mutation pressure and natural selection with a ratio of 0.082/0.918= 0.089. These results suggest that natural selection played major role while mutation pressure played minor role in shaping codon usage pattern in COII gene among pisces, aves and mammals.

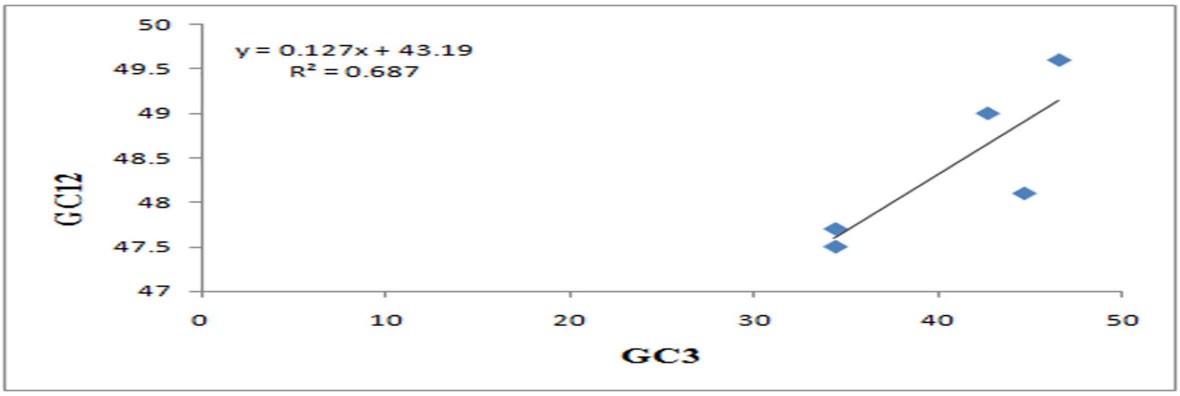


Figure 4.2.18 (a) Neutrality plot of GC12 with GC3 in different species of pisces for COIII gene

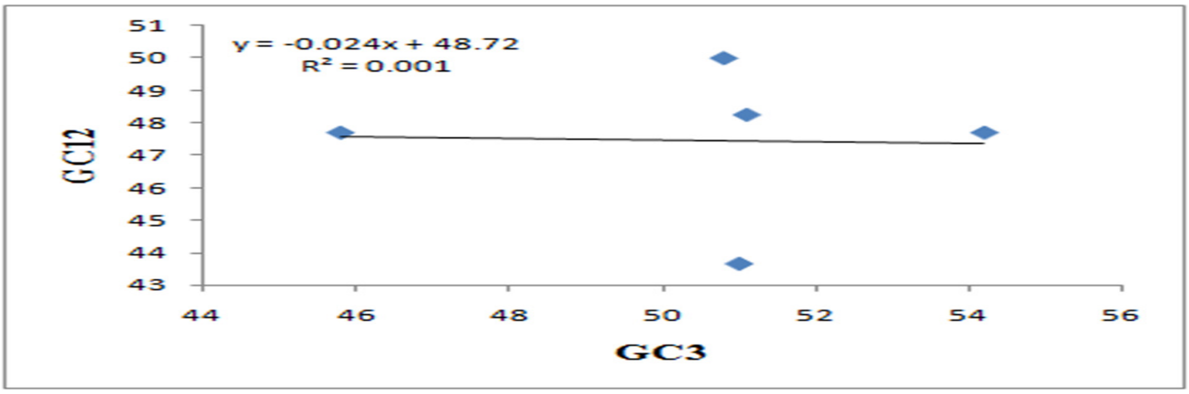


Figure 4.2.18 (b) Neutrality plot of GC12 with GC3 in different species of aves for COIII gene

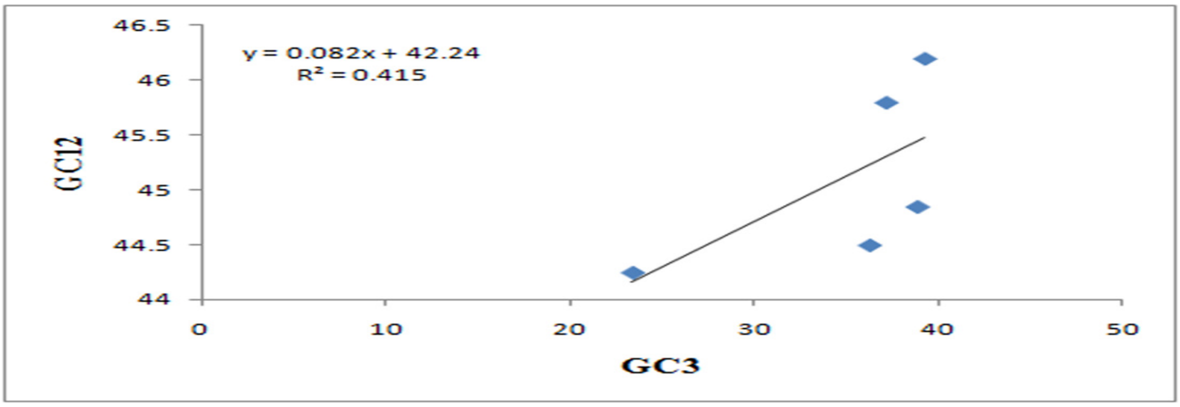
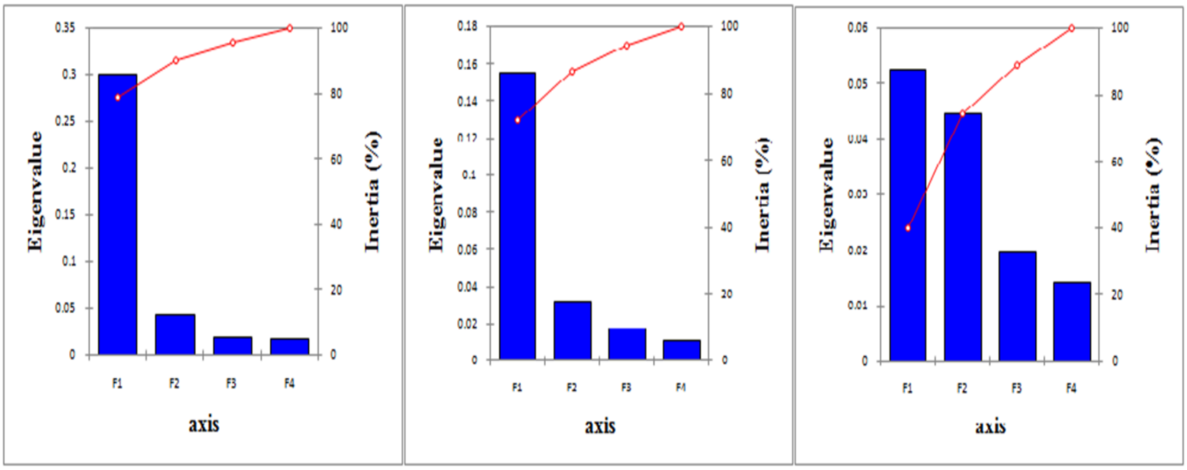


Figure 4.2.18(c) Neutrality plot of GC12 with GC3 in different species of mammals for COIII gene

4.2.7 Correspondence analysis (COA) of mitochondrial CO gene among pisces, aves and mammals

To determine the trends in codon usage variation among different species of pisces, aves and mammals for MT-CO gene, we performed COA using RSCU values.

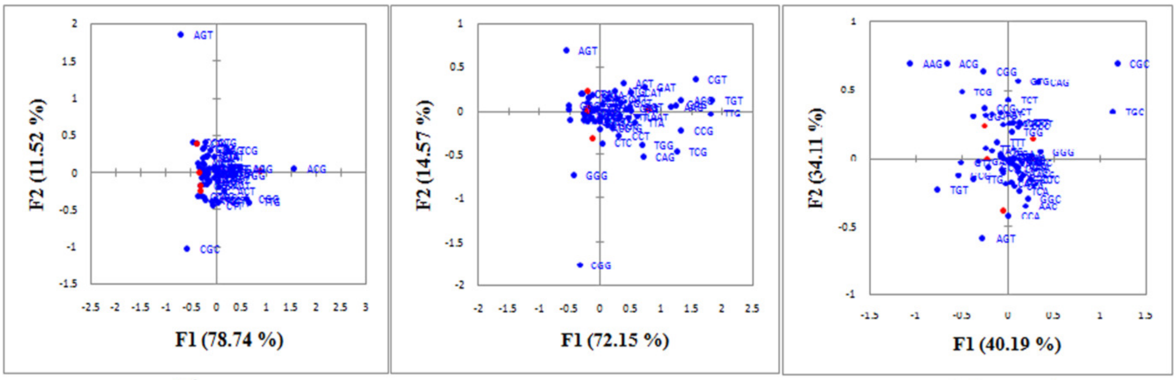
In pisces, aves and mammals, we arrived at four main contributors which were axis1, axis2, axis3 and axis4 for COI and contributions of the axes were shown in **Figure 4.2.19.**



Pisces Aves Mammals

Figure 4.2.19 Contributions of the axes for MT-COI gene in pisces, aves and mammals

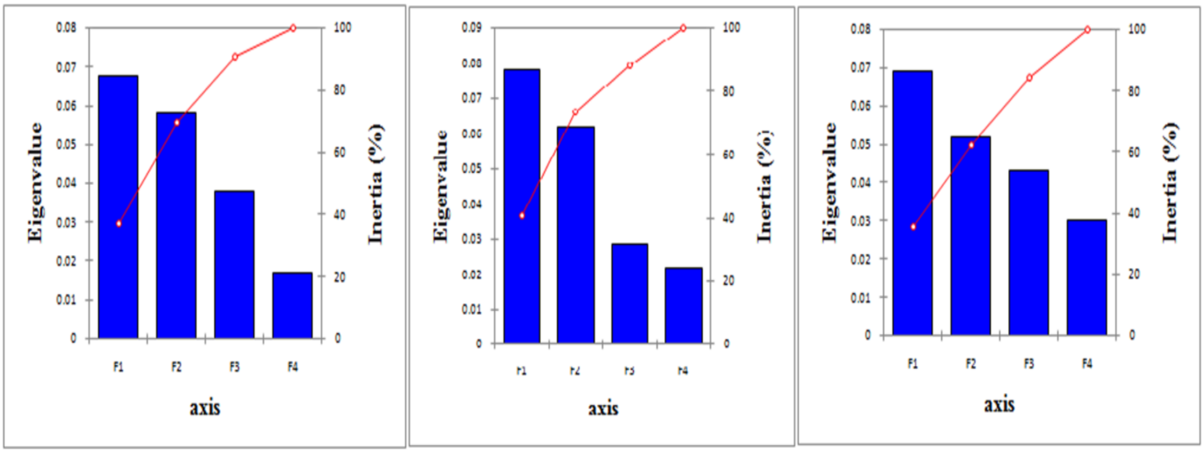
In COI gene for pisces, axis 1 accounted for 78.74% of the total variation; axis 2 accounted for 11.52% of the total variation. In aves, axis 1 accounted for 72.15% of the total variation and axis 2 accounted for 14.57% of the total variation and in mammals, axis 1 accounted for 40.19 of the total variation and axis 2 accounted for 34.11% of the total variation as shown in **Figure 4.2.20**.



Pisces Aves Mammals

Figure 4.2.20 Correspondence analysis of the synonymous codon usage for COI gene in pisces, aves and mammals

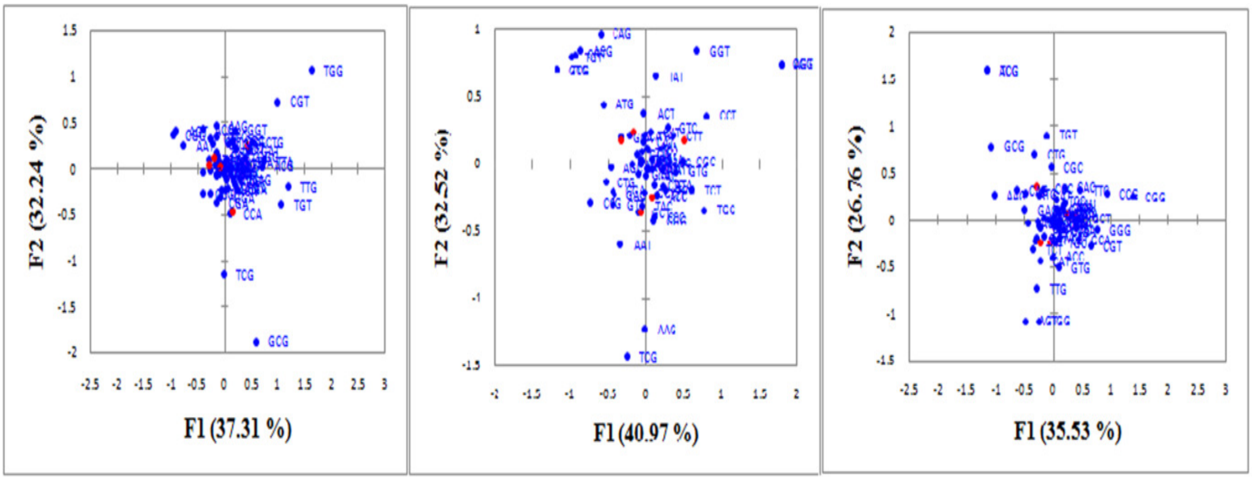
In COII gene, the four main contributors *i.e.* axis1, axis2, axis3 and axis4 for pisces, aves and mammals were shown in **Figure 4.2.21.**



Pisces Aves Mammals

Figure 4.2.21 Contributions of the axes for MT-COII gene in pisces, aves and mammals

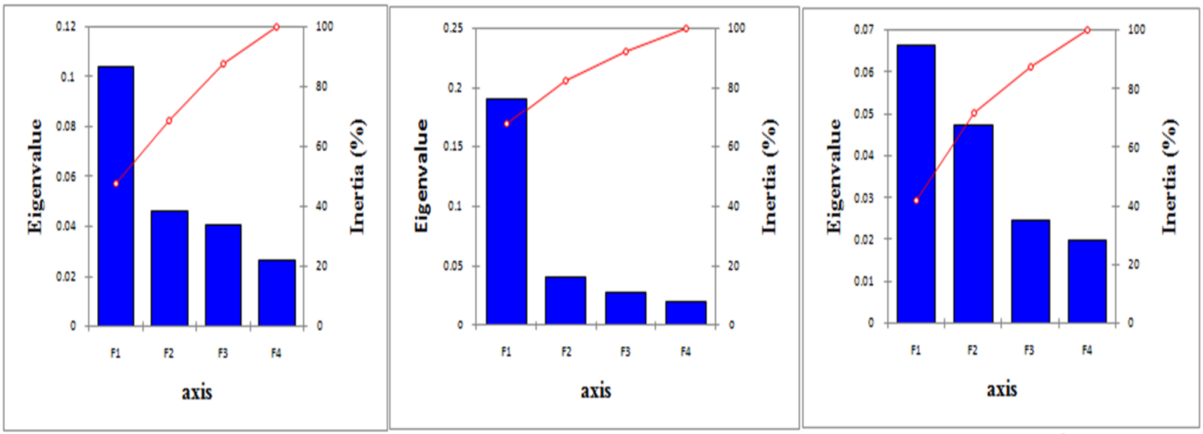
In COII gene, for pisces, axis 1 accounted for 37.31% of the total variation; axis 2 accounted for 32.24%, of the total variation. In aves, axis 1 accounted for 40.97% of the total variation and axis 2 accounted for 32.52% of the total variation and in mammals, axis 1 accounted for 35.53 of the total variation and axis 2 accounted for 26.76% of the total variation as shown in **Figure 4.2.22**.



Pisces Aves Mammals

Figure 4.2.22 Correspondence analysis of the synonymous codon usage for COII gene in pisces, aves and mammals

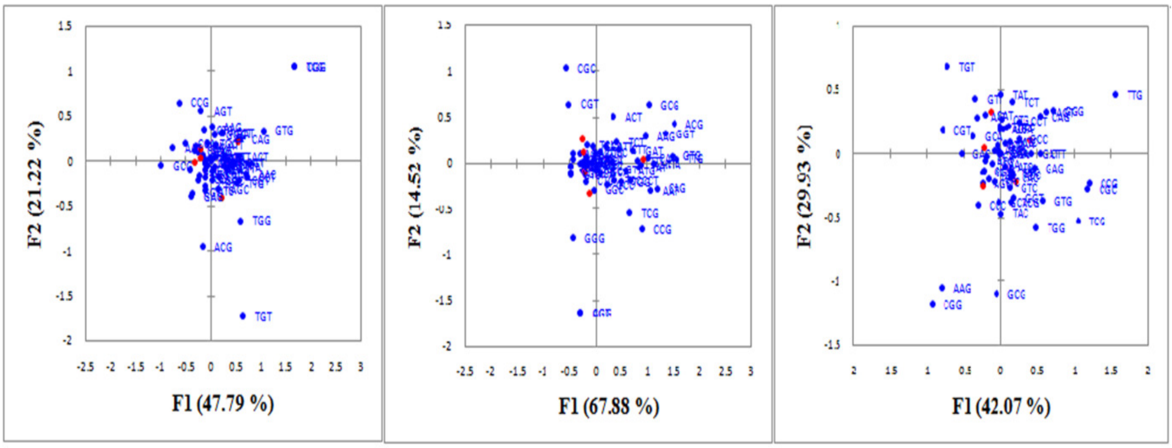
In COIII gene, the four main contributors *i.e.* axis1, axis2, axis3 and axis4 for pisces, aves and mammals were shown in **Figure 4.2.23.**



Pisces Aves Mammals

Figure 4.2.23 Contributions of the axes for MT-COIII gene in pisces, aves and mammals

In COIII gene for pisces, axis 1 accounted for 47.79% of the total variation; axis 2 accounted for 21.22%, of the total variation. In aves, axis 1 accounted for 67.88% of the total variation and axis 2 accounted for 14.52% of the total variation and in mammals, axis 1 accounted for 42.07% of the total variation and axis 2 accounted for 29.93% of the total variation as shown in **Figure 4.2.24**.



Pisces Aves Mammals

Figure 4.2.25 Correspondence analysis of the synonymous codon usage for COIII gene in pisces, aves and mammals

In COI, COII and COIII genes, the positions of most codons are more close to axes with a concentrate distribution among pisces, aves and mammals, indicating that the base composition for mutation bias might correlate to the codon usage bias. But, a few codons were in a discrete distribution, indicating that other factors, like natural selection, affect codon usage of COI, COII and COIII genes.

4.2.8 Amino acid composition of CO protein among pisces, aves and mammals

In COI protein, the amino acid frequency of the encoded proteins in different species of pisces aves and mammals was estimated. The frequency of leucine residue was the highest in the amino acid composition of pisces, aves and mammals while cysteine residue was the least in the proteins as shown in **Figure 4.2.26.** The usage of amino acids such as glutamine, arginine, lysine, glutamate and tryptophan was lower in COI protein.

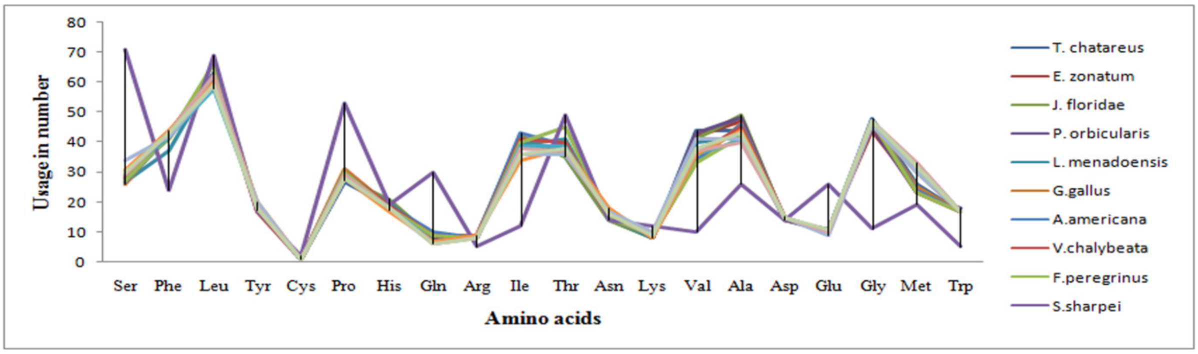


Figure 4.2.26 Comparison of amino acids in different species of pisces, aves and mammals for COI protein

In COII protein**,** the frequency of leucine residue was the highest in all species of pisces, aves and mammals while cysteine residue was the least in the proteins as shown in **Figure 4.2.27.** The usage of amino acids such as asparagine and lysine were lower in COII protein.

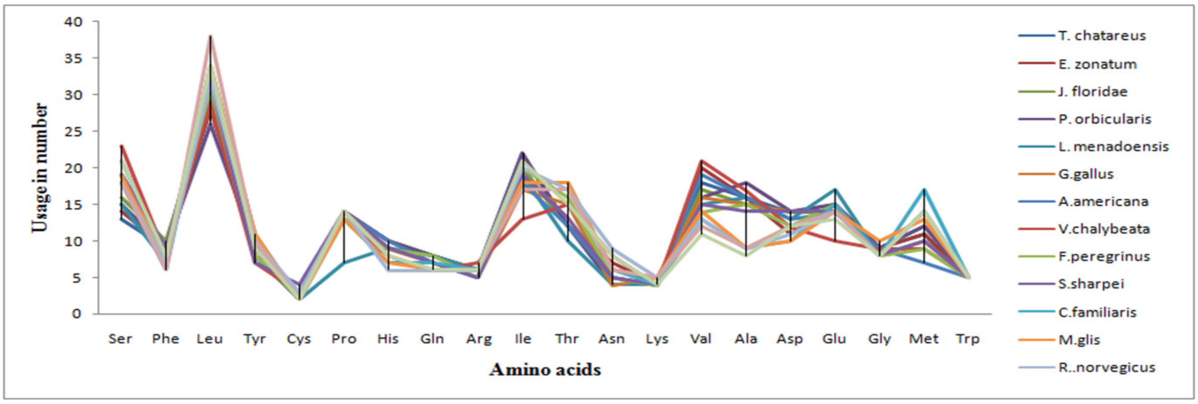


Figure 4.2.27 Comparison of amino acids in different species of pisces, aves and mammals for COII protein

In COIII protein, the amino acid frequency of the encoded proteins in different species of pisces aves and mammals was estimated. The frequency of leucine residue was the highest in the amino acid composition of most of the species of pisces, aves and mammals while cysteine, arginine, asparagine, lysine, aspartate residue were low in the proteins as shown in **Figure 4.2.28.** The usage of amino acids such as serine was the highest in only *S. sharpei* for COIII protein.

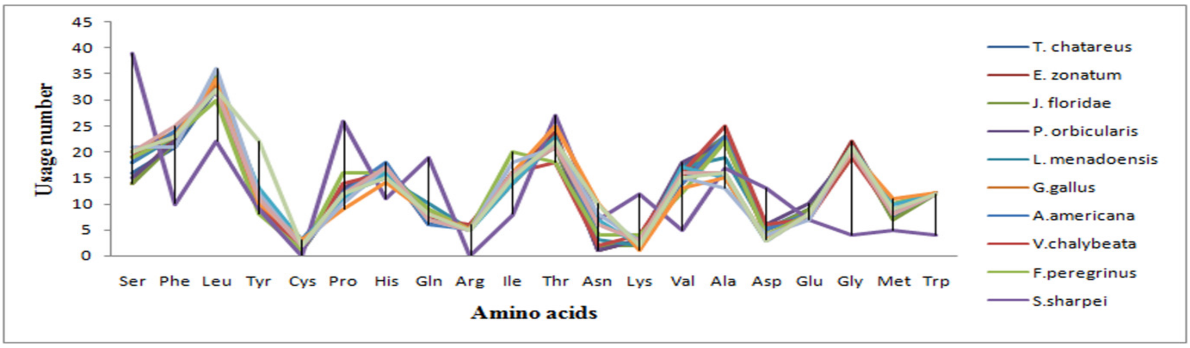


Figure 4.2.28 Comparison of amino acids in different species of pisces, aves and mammals for COIII protein

4.3. Analysis of codon usage bias in MT-CYB gene among Pisces, Aves and Mammals

4.3.1 Codon usage pattern of MT-CYB gene in pisces, aves and mammals

We performed RSCU analysis to determine the pattern of synonymous codon usage of MT-CYB gene and the extent of G/C ending codons in coding sequences. Furthermore, we divided the RSCU data into two groups (a) RSCU value >1.6: overrepresented (b) RSCU value<0.6: underrepresented. From the heat map, the over represented and the under represented codons are clearly evident as shown in **Figure 4.3.1.**

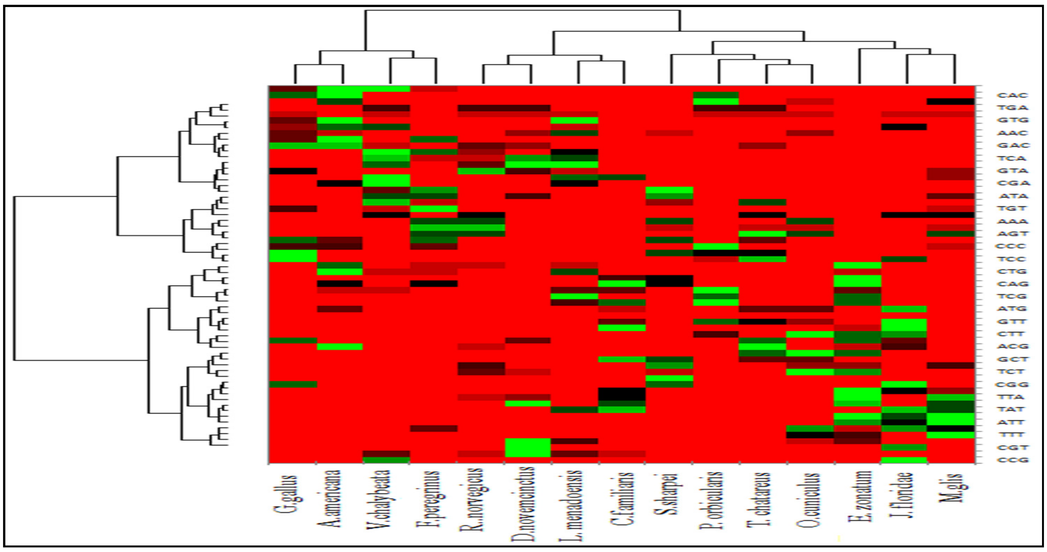


Figure 4.3.1 Hierarchial Clustering of RSCU values of each codon among pisces, aves and mammals. Each rectangular box on the map represents the RSCU value of a codon (shown in rows) corresponding to MT-CYB gene among pisces, aves and mammals (shown in columns). The color coding indicates different RSCU values: green indicates RSCU>1.6, black indicates RSCU>1 and dark red indicates RSCU>1 and red indicates RSCU <0.6.

In CYB gene, out of 60 codons, AGT, CTT, GTA, AGT, CAG encoding amino acids ser, leu, val, ser, gln respectively were used as over-represented codons in some species but under-presented codons in most of the species as shown in **Figure 4.3.1**. Further, it was observed that some codons were more frequently used in some species but less frequently used in other species. Based on RSCU analysis and nucleotide composition we deduced that the existence of preferred codons in coding sequences has been mostly influenced by compositional constraints, which account for the presence of mutation pressure.

4.3.2 Analysis of codon usage bias in MT-CYB gene among pisces, aves and mammals

The effective number of codon (ENC) was calculated to quantify the degree of codon usage bias among different species of pisces, aves and mammals. In MT-CYB gene, the ENC value ranges from 59.80±0.44 in pisces and aves and 58.20±1.30 in mammals in the present study indicating low codon usage bias. The ENC values of different species were shown in **Figure 4.3.2**. The high ENC value indicates conserved genomic composition among pisces, aves and mammals.

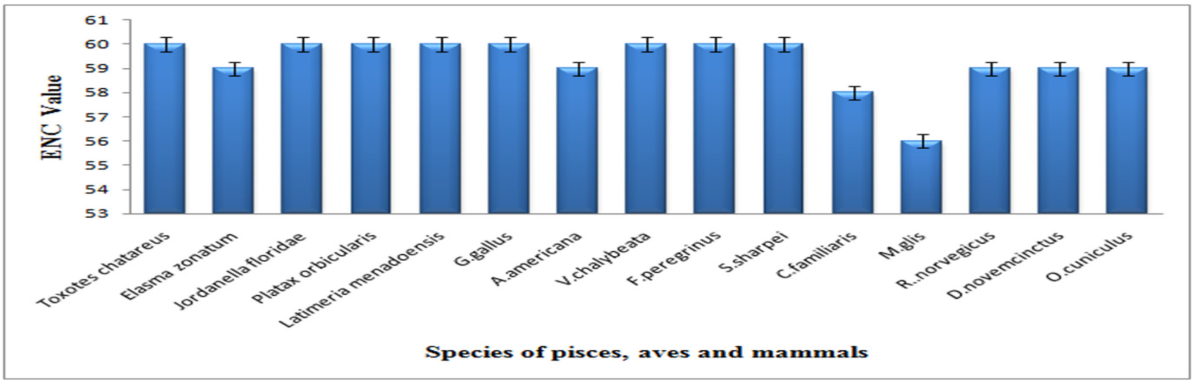


Figure 4.3.2 Distribution of ENC in different species of pisces, aves and mammals for MT-CYB gene

4.3.3 Prediction of gene expression level in MT-CYB gene among pisces, aves and mammals

The CAI is a directional measure of codon usage bias and its higher value indicates higher gene expression level and also higher codon usage bias. In CYB gene, the mean±SD of different species, aves and mammals were 0.82±0.05, 0.76±0.05, 0.78±0.01, which suggest that the expression level of CYB gene in different species of pisces, aves and mammals was high. The expression level of *E. zonatum* and *P. orbicularis* was high in comparison to others as shown in **Figure 4.3.3.**

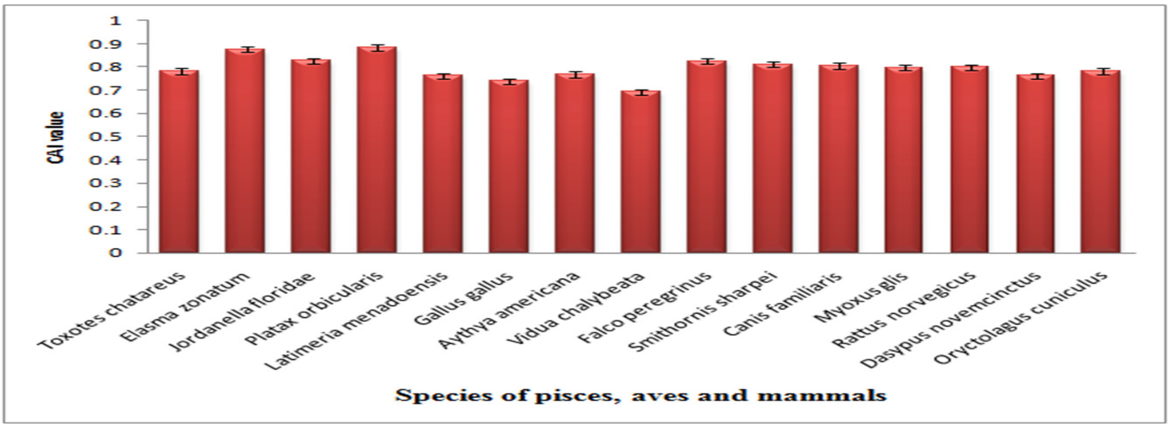


Figure 4.3.3 Expression level in different species of pisces, aves and Mammals for MT-CYB gene

4.3.4 Relationship of gene expression (CAI) with codon usage bias (ENC) for MT-CYB gene among pisces, aves and mammals

In CYB, we found negative correlation between ENC and CAI in pisces, aves and in mammals. Pearson correlation coefficient was -0.511, (p>0.05) in pisces, –0.014, (p>0.05) in aves and -0.459, (p>0.05) in mammals and these suggest that codon usage bias has a very weak relationship with nucleotide composition.

4.3.5 Compositional features of MT-CYB gene among pisces, aves and mammals

Overall nucleotide composition may influence the codon usage bias of a genome (Jenkins and Holmes 2003). We therefore analyzed nucleotide composition of coding sequences of CYB gene in different species of pisces, aves and mammals as shown in **Table 4.3.1**. In pisces, mean±SD value of C% was the highest, followed by T%, A% and G%. In aves, mean ±SD value of C% was the highest, followed by A%, T%, and G%. In mammals, mean ±SD value of A% was the highest, followed by T%, C%, and G%.

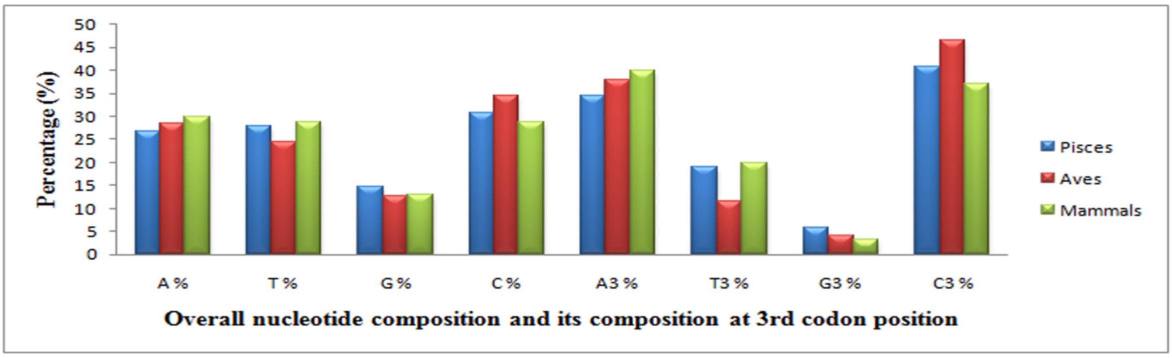


Figure 4.3.4 Nucleotide composition and its 3rd codon position for MT-CYB gene

This suggests unequal distribution of A, T, G, C % among the codons in different species of pisces, aves and mammals with more preference for C ending codons in pisces and aves followed by T/A ending codons in pisces, but A/T ending codons in aves. In mammals, the increased preference of A ending codons was observed, which was followed by T/C ending codons. The preference of G ending codons was less in aves followed by mammals and then pisces. However, analysis of nucleotide composition at the 3rd position of codons (A3%, T3%, G3% and C3%) and GC1%, GC2%, GC3% provides a clear picture about the preference of codon usage in different species of pisces, aves and mammals as shown in **Table 4.3.1**. The mean±SD value of C3% in pisces and in aves was the highest, followed by A3% and T3% respectively. The mean±SD value of A3% was the highest followed by C3% and T3% in mammals. The G3% was the lowest in mammals followed by aves and pisces as shown in **Figure 4.3.4.** Therefore, from the initial nucleotide composition analysis, it was expected that nulceobase C/A might be more preferred in coding sequences.

The mean ±SD values of overall GC% were 45.34±1.88, 47.12±1.86 and 41.66±1.45 in pisces, aves and mammals respectively. The overall GC% was the lowest in mammals followed by pisces and aves.

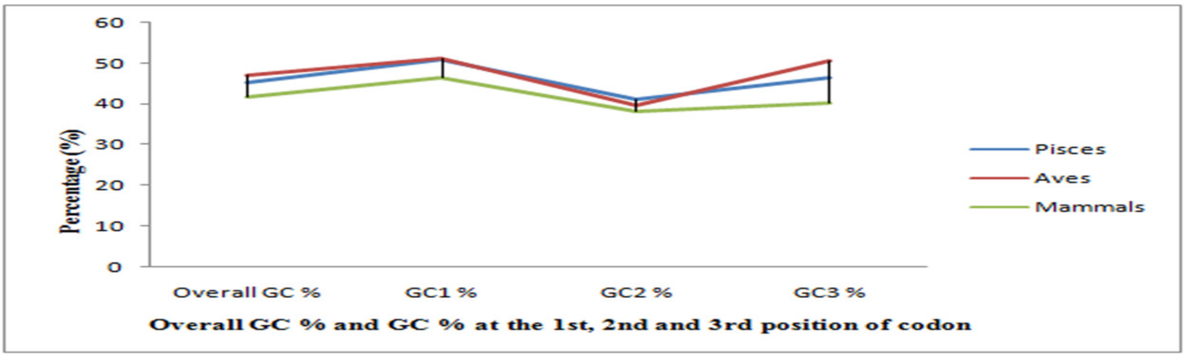


Figure 4.3.5 GC content and its content at codon's 1st, 2nd and 3rd position in pisces, aves and mammals for MT-CYB gene

In pisces and aves and mammals, GC1 % was the highest followed by GC3% and GC2% and the greatest difference of GC content were found between the 1st position and the 2nd codon position as shown in **Figure 4.3.5.** From the GC3 content analysis, it was evident that in pisces and mammals AT ending codons might be more preferred to GC ending codons whereas in aves both AT and GC ending codons might be equally preferred at the 3rd codon position. From the overall GC content analysis it was found that AT ending codons were more preferred to GC ending codons. These results suggest that there might be compositional constraint in presence of mutation pressure which affects the mitochondrial CYB gene.

Table 4.3.1 Nucleotide compositions among pisces, aves and mammals for CYB gene

Species	**A %**	**T %**	**G %**	**C %**	**A3 %**	**T3 %**	**G3 %**	**C3 %**
T.chatareus	25.63	26.94	14.34	33.07	32.8	17.32	3.93	45.93
E. zonatum	26.4	30.35	15.39	27.73	33.33	23.35	8.92	34.38
J. floridae	24.47	30	15.43	30.08	30.78	23.94	5.78	39.47
P.orbicularis	25.45	27.2	14.26	33.07	32.28	17.58	4.19	45.93
L. menadoensis	31.32	24.84	14.26	29.57	42.78	13.12	6.29	37.79
Mean±SD	26.65±2.69	27.86±2.30	14.73±0.61	30.70±2.33	34.39±4.74	19.06±4.54	5.82±2.00	40.70±5.11
G.gallus	27.47	24.14	12.07	36.3	34.9	10.49	3.14	51.44
A.americana	26.85	23.44	14.17	35.52	34.64	7.87	7.34	50.13
V.chalybeata	30.18	23.44	13.38	32.98	44.88	8.92	3.14	43.04
F.peregrinus	29.74	24.2	12.42	33.42	40.41	12.33	4.19	43.04
S.sharpei	27.99	26.68	11.63	33.68	35.17	17.58	2.88	44.35
Mean±SD	28.44±1.44	24.38± 1.33	12.73±1.02	34.38±1.44	38.00±4.52	11.43±3.82	4.13±1.85	46.40±4.06
C.familiaris	29.03	29.12	13.94	27.89	39.21	20.78	5	35
M.glis	29.94	31.84	12.36	26.84	38.68	26.31	2.36	32.63
R. norvegicus	30.35	27.2	12.59	29.83	42.25	14.96	2.62	40.15
D.novemcinctus	31.49	26.57	13.15	28.77	43.94	14.73	3.38	37.63
O.cuniculus	28.07	28.94	12.63	30.35	35	22.89	2.36	39.73
Mean±SD	29.77±1.30	28.73±2.05	12.93±0.63	28.73±1.42	39.81±3.45	19.93±5.04	3.14±1.11	37.02±3.19

4.3.6 Interrelationships among different compositional features of MT-CYB gene among pisces, aves and mammals

We performed correlation analysis between general nucleotide composition and nucleotide composition at 3rd codon position to determine whether evolutionary process is driven by mutation pressure alone or by both mutation pressure and natural selection. In pisces, aves and mammals, highly significant positive correlation was found between A and A3% , T and T3%, C and C3% , GC and GC3% respectively and negative correlation was observed for most of various nucleotide comparisons as shown in **Table 4.3.2.** These results suggest that the compositional constraint, arising from mutation pressure and natural selection might have played a significant role in codon usage of MT-CYB gene.

Table 4.3.2 Correlation between overall nucleotide composition (%) and its composition at 3rd codon position in pisces, aves and mammals

	Nucleotide	A3 %	T3 %	G3 %	C3 %	GC3%
Pisces	A %	0.997**	-0.732	0.249	-0.379	-0.446
	T %	-0.746	0.995**	-0.379	0.489	-0.295
	G %	-0.471	0.929*	0.668	-0.647	-0.607
	C %	-0.289	-0.394	-0.967**	0.999**	0.979**
	GC %	-0.499	-0.162	-0.934*	0.977**	0.965**
Aves	A %	0.942*	0.027	-0.485	-0.854	-0.851
	T %	-0.393	0.974**	-0.476	-0.261	-0.382
	G %	0.217	-0.837	0.782	0.189	0.435
	C %	-0.754	-0.34	0.369	0.991**	0.919*
	GC %	-0.463	-0.736	0.717	0.881*	0.959*
Mammals	A %	0.943*	-0.625	-0.012	-0.06	-0.035
	T %	-0.613	0.956*	-0.173	-0.761	-0.869
	G %	0.178	-0.244	0.985**	-0.158	0.197
	C %	-0.117	-0.475	-0.3	0.980**	0.902*
	GC %	-0.047	-0.563	0.124	0.893*	0.967**

Note: One asterisk (*) and two asterisks (**) indicate correlations at a level of 0.05 and 0.01, respectively.

Furthermore, the correlations between ENC and various GC contents as shown in **Table 4.3.3**, suggest that the nucleotide composition under from mutation pressure and natural selection affect the synonymous codon usage of MT-CYB gene in pisces, aves and mammals.

Table 4.3.3 Correlation coefficients among ENC, CAI, GC contents for MT-CYB gene

		Pisces		**Aves**		**Mammals**	
SL No	Correlation between	Correlation coefficient	p value	Correlation coefficient	p value	Correlation coefficient	p value
1	ENC and CAI	-0.511	0.379	-0.014	0.983	-0.459	0.437
2	ENC and GC %	0.668	0.218	-0.775	0.124	0.954*	0.012
3	ENC and GC3 %	0.559	0.327	-0.756	0.14	0.984**	0.002
4	GC and GC3 %	0.965**	0.008	0.959*	0.01	0.967**	0.007
5	GC1 and GC3 %	0.811	0.096	-0.09	0.886	0.893*	0.042
6	GC2 and GC3%	-0.542	0.345	0.704	0.185	-0.467	0.428
7	GC12 and GC3 %	-0.064	0.919	0.135	0.829	0.73	0.161

Note: One asterisk (*) and two asterisks (**) indicate correlations at a level of 0.05 and 0.01, respectively.

If the pattern of synonymous codon usage is solely governed by mutation pressure then the frequency of nucleotide A and T should be equal to that of G and C at synonymous 3rd codon position (Zhang Zhicheng et al. 2013). In case of pisces, aves and mammals, the frequencies of those nucleotides were not same indicating that other factors such as natural selection might have played a role in codon usage pattern.

To determine the extent of mutation pressure against natural selection in the codon usage pattern in mitochondrial MT-CYB gene, we drew the neutrality plot. Neutrality plot is the regression of GC12 (average of GC1 and GC2) on GC3. In pisces, the regression coefficient of GC12 on GC3 for MT-CYB gene is 0.033 which reveals the relative neutrality of 3.3% (absolute) and relative constraint of 96.7% for GC3. The GC12 was influenced by mutation pressure and natural selection with a ratio of 0.033/0.967= 0.034. The regression coefficient of GC12 on GC3 in aves is 0.020 which indicates the relative neutrality of 2% and relative constraint of 98% for GC3. The GC12 was affected by mutation pressure and natural selection with a ratio of 0.020/0.98=0.0204 in aves. In mammals, the regression coefficient of GC12 on GC3 for MT-CYB gene is 0.181 which suggests that the relative neutrality is 18.1% and relative constraint is 81.9% for GC3. The GC12 was influenced by mutation pressure and natural selection with a ratio of 0.181/ 0.819=0.221 (SUEOKA 1988). As shown in **Figure 4.3.6 (a) (b) (c)** the points are not in diagonal distribution rather the points are in a narrow range of distribution suggesting that GC12 versus GC3 are not due to mutational bias (Hebert et al. 2003). These results suggest that natural selection played a major role while mutation pressure played a minor role in shaping codon usage pattern of CYB gene among pisces, aves and mammals.

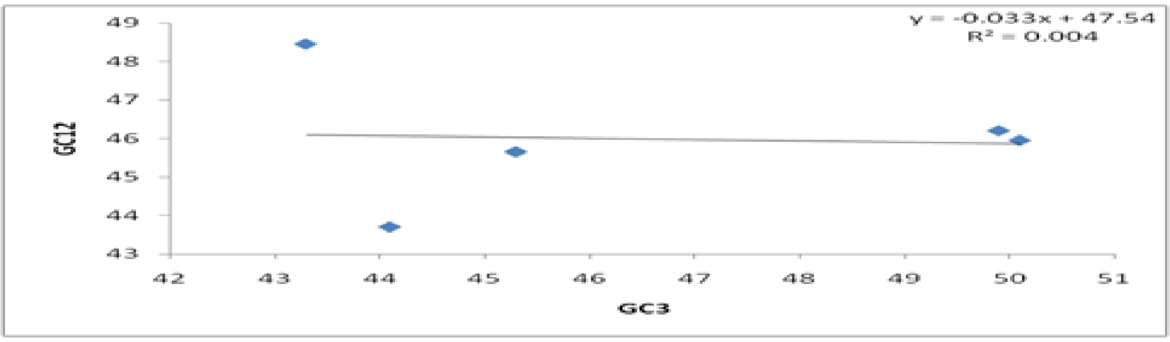


Figure 4.3.6 (a) Neutrality plot of GC12 with GC3 in different species of pisces for CYB gene

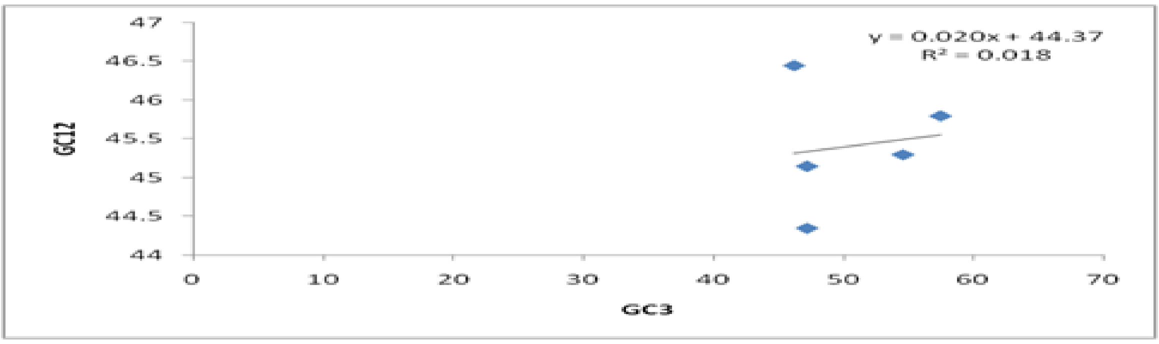


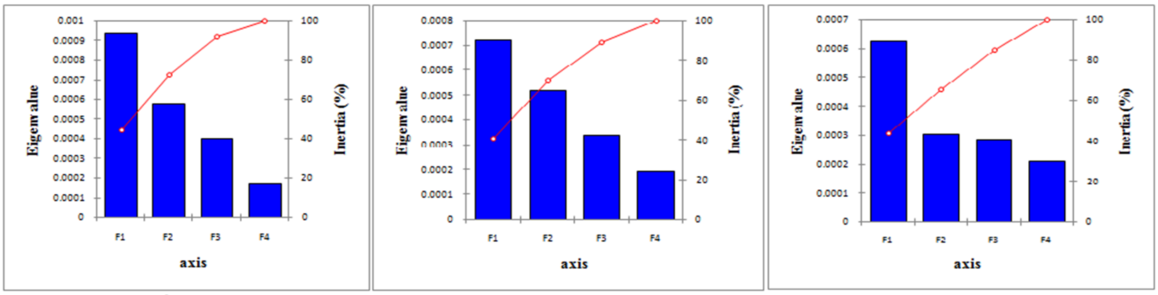
Figure 4.3.6 (b) Neutrality plot of GC12 with GC3 in different species of aves for CYB gene

Figure 4.3.6 (c) Neutrality plot of GC12 with GC3 in different species of mammals for CYB gene

4.3.7 Correspondence analysis (COA) of MT-CYB gene among pisces, aves and mammals

COA is a multivariate statistical technique which is used to determine the trends in codon usage variation among different species for MT-CYB gene. Therefore, we performed COA using RSCU values of codons. Based on the 60 synonymous codons, MT-CYB gene for different species was represented by the points in super dimensional axes and carried on to the correspondence analysis.

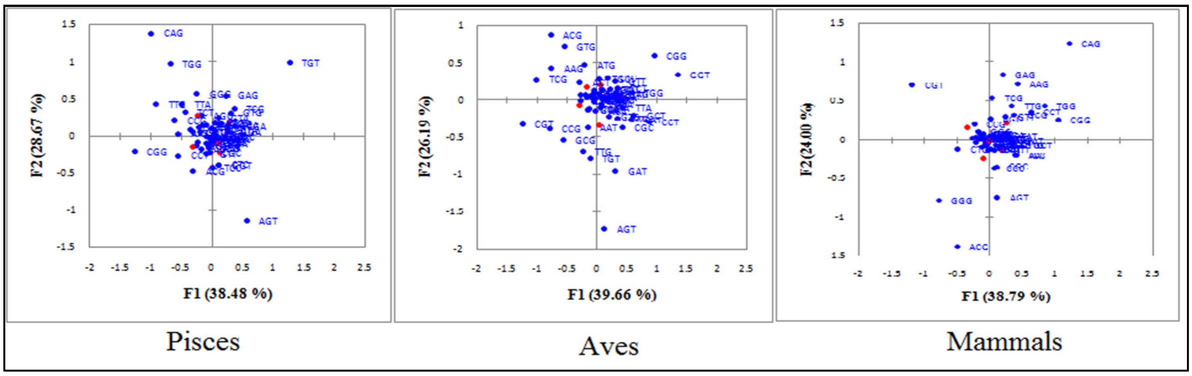
In pisces, aves and mammals, the contributions of the axes are shown in **Figure 4.3.7** and we found four main contributors *i.e.* axis1, axis2, axis3 and axis4 for CYB gene.



Pisces Aves Mammals

Figure 4.3.7 Contributions of the axes for MT-CYB in pisces, aves and mammals

In CYB gene, the first two-main-dimensional coordinates, axis1 and axis2 can explain 38.48% and 28.67% respectively of the total variation, which led to the first axis as the major contributor to the codon bias in pisces (**Figure 4.3.8**). In aves, axis1 contributes 39.66% of the total variation while axis2 contributes 26.19% respectively of the total variation, and indicate the first axis as the major contributor to the codon bias (**Figure 4.3.8**). In mammals, axis1 and axis2 can explain 38.79% and 24.00% respectively of the total variation, and as a result, the first axis was the major contributor to the codon bias (**Figure 4.3.8**). The positions of most codons are more close to axes with a concentrate distribution (**Figure 4.3.8**), indicating that the base composition for mutation bias might correlate to the codon usage bias. But a few genes are in a discrete distribution, which indicates that there are many other factors which could affect the codon usage.



Pisces Aves Mammals

Figure 4.3.8 Correspondence analysis of the synonymous codon usage for CYB gene in pisces, aves and mammals

4.3.8 Amino acid composition in different species of pisces, aves and mammals for CYB protein

The frequencies of 20 amino acids in the encoded proteins in different species of pisces, aves and mammals were estimated. The frequency of leucine was the highest in the amino acid composition of pisces, aves and mammals while cysteine residue was the least in the proteins as shown in **Figure 4.3.9.** The other amino acids such as glutamine, arginine, aspartate, glutamate were low in usage in the amino acid sequence of CYB protein.

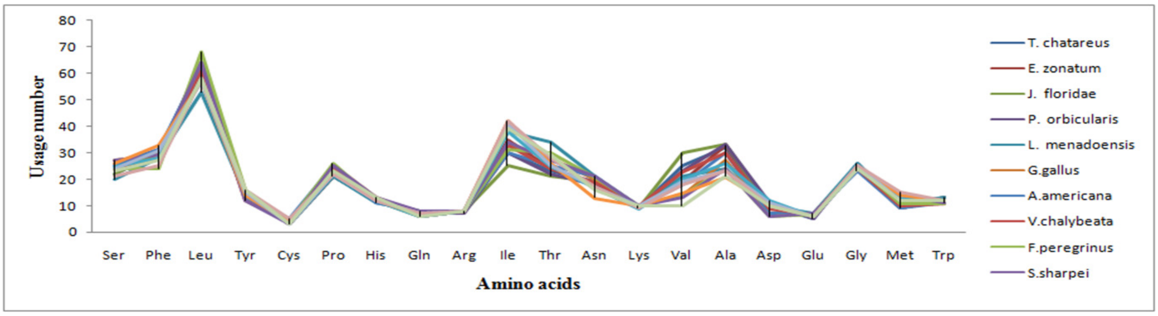


Figure 4.3.9 Comparison of amino acids in different species of pisces, aves and mammals in CYB protein

4.4 Analysis of codon usage bias in MT-ATP gene among Pisces, Aves and Mammals

4.4.1 Codon usage pattern in MT-ATP gene among pisces, aves and mammals

To understand the pattern of unequal usage of synonymous codons in ATP gene in different species of pisces, aves and mammals, we estimated relative synonymous codon usage (RSCU) of individual codons and compared them among the different species of pisces, aves and mammals.

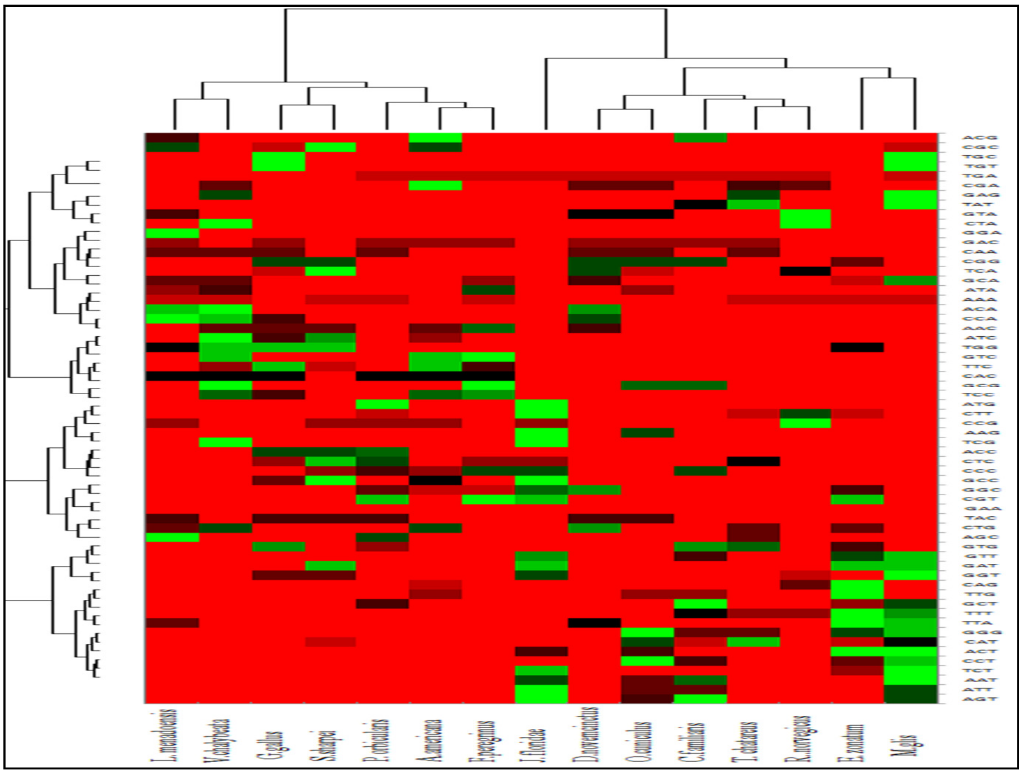


Figure 4.4.1 Hierarchial Clustering of RSCU values of each codon among pisces, aves and mammals. Each rectangular box on the map represents the RSCU value of a codon (shown in rows) corresponding to MT-ATP6 gene among pisces, aves and mammals (shown in

columns). The color coding indicates different RSCU values: green indicates RSCU>1.6, black indicates RSCU>1 and dark red indicates RSCU>1 and red indicates RSCU <0.6.

In ATP6 gene, of these 60 codons, GAT, GGT, GTC, TGG encoding the amino acids asp, gly, val, trp respectively were used as over-represented in some species but under-presented in most of the species as shown in **Figure 4.4.1**.

In ATP8 gene, codons such as GAT, TAT, TCT, and CTA encoding the amino acids asp, tyr, ser and leu respectively were used as over-represented in some species while under-represented in other species as shown in **Figure 4.4.2**.

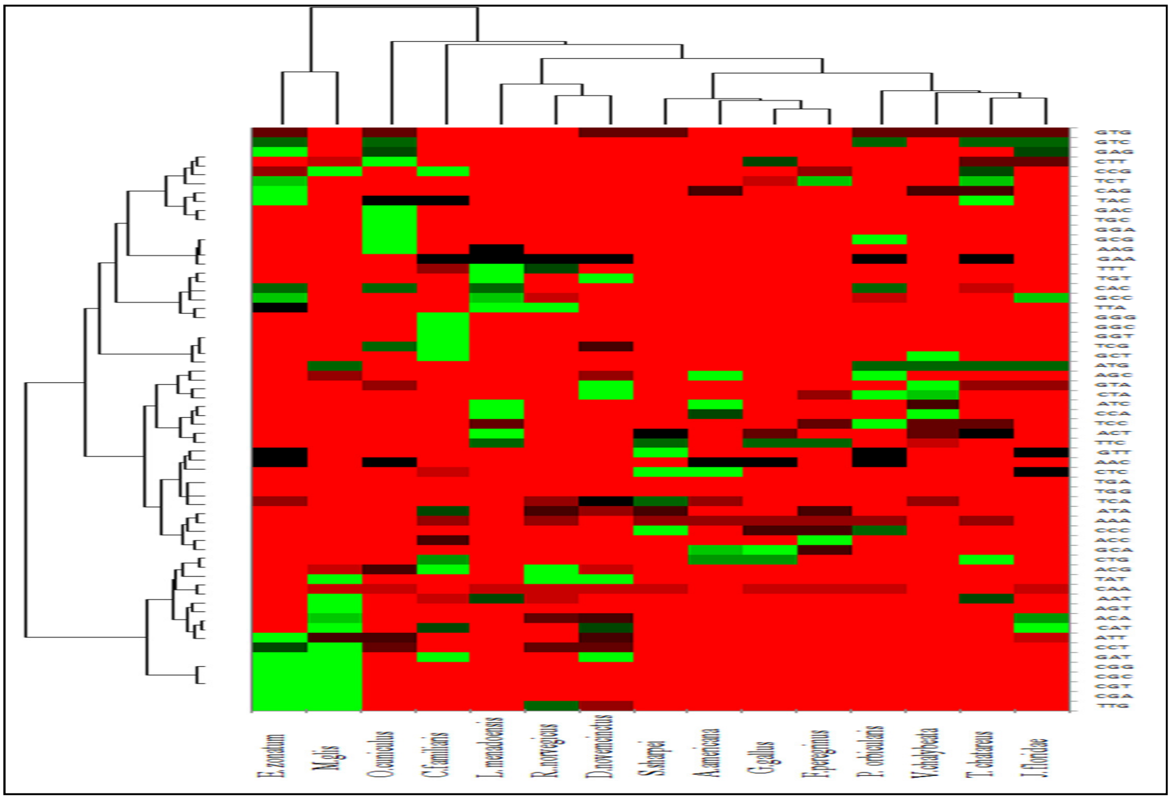


Figure 4.4.2 Hierarchial Clustering of RSCU values of each codon among pisces, aves and mammals for MT-ATP8

Furthermore, it was also found that in both genes, specific codons are frequently used in each species. Based on RSCU analysis it was found that the preferred codons in these two genes of different species have been mostly influenced by compositional constraints under mutation pressure.

4.4.2 Analysis of codon usage bias in MT-ATP gene among pisces, aves and mammals

The effective number of codon (ENC) was calculated to quantify the degree of codon usage bias among different species of pisces, aves and mammals. In ATP6 gene, the ENC values in different species of pisces, aves and mammals were (Mean ±SD) 57.8±2.16, 59.80±0.44, 53.20±4.20 respectively. The mammalian species showed low (average) ENC value than pisces and aves, which suggest that codon usage bias was high in mammals for ATP6 in comparison to different species of aves and mammals. From **Figure 4.4.3**, it was found that *M. glis* showed low ENC value as compared to other species in ATP 6 gene which indicates low codon usage bias. The high ENC value in different species of pisces, and aves indicates that codon usage bias was low.

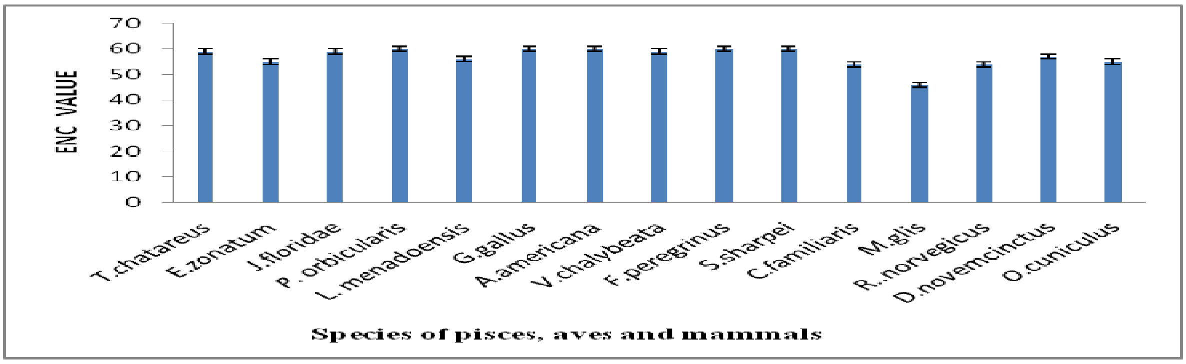


Figure 4.4.3 Distribution of ENC in different species of pisces, aves and mammals for MT-ATP6

In ATP8 gene, the ENC values in different species of pisces, aves and mammals were (Mean ±SD) 56.8±2.88, 59.2±0.83 and 50.2±3.70 respectively. The different species of mammals showed low ENC value (average) in comparison to different species of aves and mammals which indicates that mammalian ATP8 contains high codon usage bias than aves and mammals. The high ENC value in pisces and aves indicates that codon usage bias was low and was conserved. From **Figure 4.4.4**, *M. glis* also showed low ENC value than others which suggest that *M. glis* showed high codon usage bias than others for ATP8.

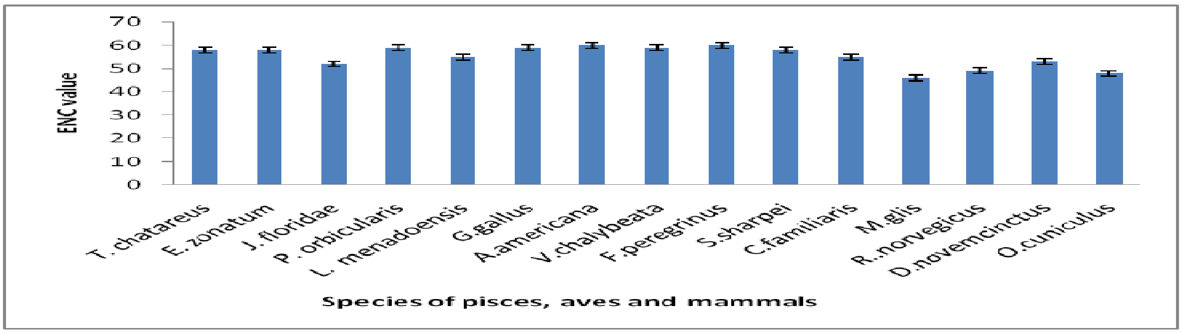


Figure 4.4.4 ENC distribution in different species of pisces, aves and mammals for ATP 8

4.4.3 Prediction of gene expression level in MT-ATP gene among pisces, aves and mammals

In ATP6 gene, the mean±SD of different species, aves and mammals were 0.6624±0.056, 0.6807±0.065, 0.6814±0.092, which suggest that the expression level of ATP6 in different species of pisces, aves and mammals was high. The expression level of *F. peregrinus* and *C. familiaris* were high in comparison to others as shown in **Figure 4.4.5.**

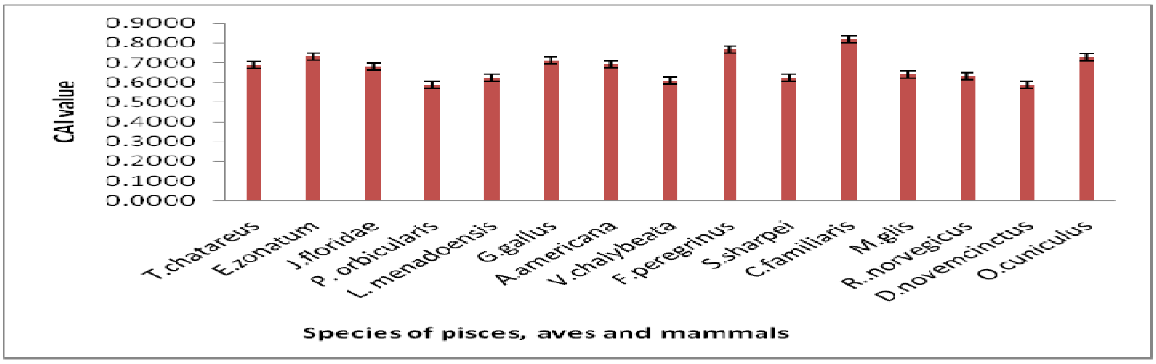


Figure 4.4.5 Expression level in different species of pisces, aves and Mammals for MT-ATP6

In ATP8 gene**,** the mean±SD of different species, aves and mammals were 0.2839±0.20, 0.2882±0.087, 0.3632±0.171, which suggest that the expression level of ATP8 in different species of pisces, aves and mammals were moderate. The expression level of *T. chatareus* and *D. norvemcinctus* were high in comparison to others as shown in **Figure 4.4.6.**

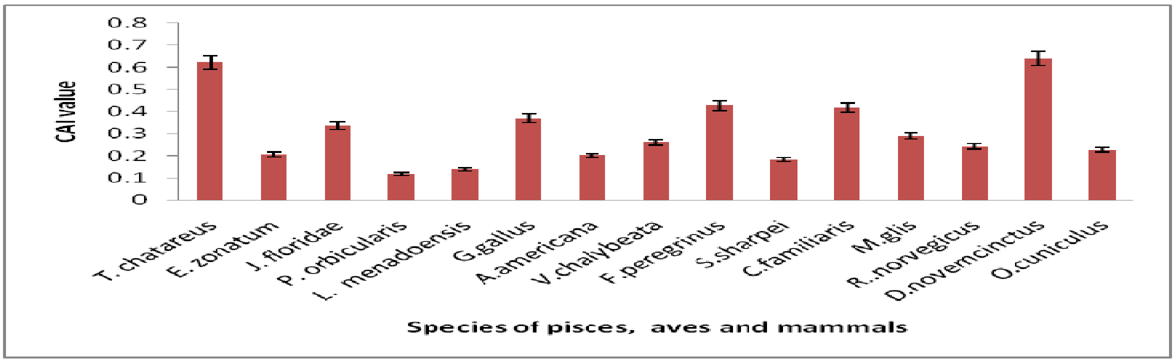


Figure 4.4.6 Expression level in different species of pisces, aves and mammals for MT-ATP8

4.4.4 Relationships of gene expression (CAI) with codon usage bias (ENC) for MT-ATP among pisces, aves and mammals

We correlated ENC and CAI to understand the nucleotide composition variation and codon selection among different species of pisces, aves and mammals in MT-ATP6 and MT-ATP8 gene. In MT-ATP6 gene, we found negative correlation between ENC and CAI in pisces, positive correlation in aves and mammals. In pisces, Pearson correlation coefficient was -0.483, ($p>0.05$), in aves, 0.621, ($p>0.05$) and in mammals, 0.06, ($p>0.05$). In MT-ATP8 gene, we found negative correlation between ENC and CAI in pisces and positive correlation in aves and mammals. In pisces, Pearson correlation coefficient was -0.019, ($p>0.05$), 0.437, ($p>0.05$) and in mammals 0.701, ($p>0.05$), which suggest codon usage bias has a very weak relationship with nucleotide composition.

4.4.5 Compositional features of MT-ATP gene among pisces, aves and mammals

In ATP6, overall nucleotide composition may influence the codon usage bias of a genome (Jenkins and Holmes 2003). We therefore analyzed nucleotide composition of coding sequences of different species of pisces, aves and mammals. From the **Table 4.4.1**, in pisces, mean±SD value of C% was the highest, followed by T% and A%, with G% being the lowest. In aves, mean ±SD value of C% was the highest, followed by A% and T%, with G% being the lowest. The mean ±SD value of A% was the highest, followed by T% and C%, with G% being the lowest in mammals.

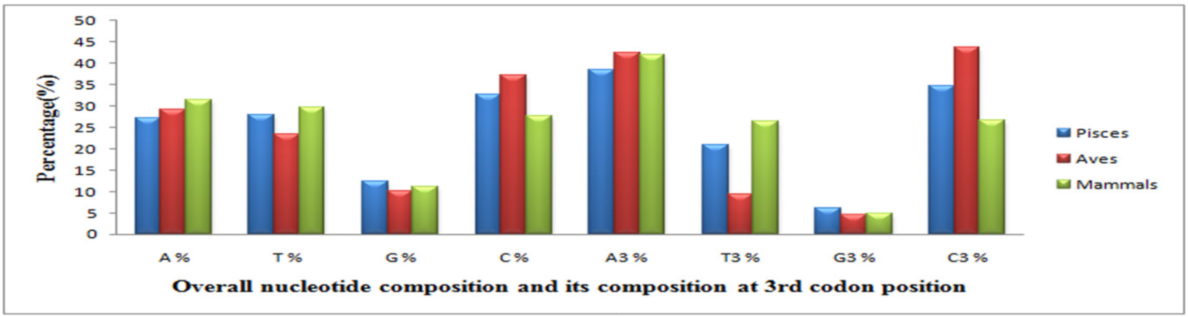


Figure 4.4.7 Nucleotide composition and its 3rd codon position for MT-ATP6

This suggests unequal distribution of A, T, G, C% among the codons in different species of pisces, aves and mammals with more preference of C ending codons in pisces and aves but in mammals A ending codons. The preference of G ending codons was less in aves followed by mammals and then pisces (**Figure 4.4.7**). However, analysis of nucleotide composition at the 3rd position of codons (A3%, T3%, G3% and C3%) provides a clear picture about the preference of codon usage in different species of pisces, aves and mammals (**Table 4.4.1**). The mean±SD value of A% was the highest, followed by C3%, T3% and G3% in pisces. The mean±SD value of % of C3 was the highest followed by A3, T3 and G3% in aves. The mean±SD of A3% was the highest followed by C3%, T3% and G3% in mammals. Therefore from the initial nucleotide composition analysis it was expected that nulceobase C/A might be more preferred. The values of mean ±SD of overall

GC% in pisces, aves and mammals were 44.92±3.88, 47.42±1.47 and 38.78±2.76 respectively which suggest that MT-ATP6 gene is AT-rich in pisces, aves and mammals. In pisces and aves, GC1% was the highest followed by GC3% and GC2% but in mammals GC1% was the highest followed by GC2% and GC3% as shown in **Figure 4.4.8**.

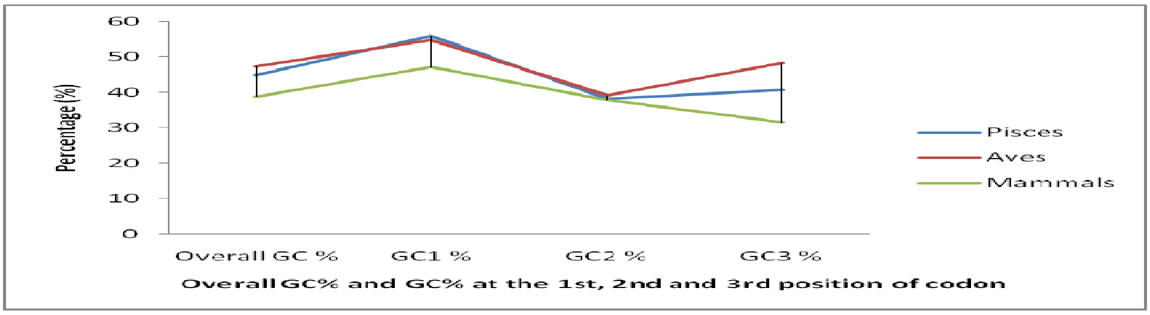


Figure 4.4.8 Overall GC content and its content at codon's 1st, 2nd and 3rd position in pisces, aves and mammals

Table 4.4.1 Nucleotide compositions among pisces, aves and mammals for ATP6 gene

Species	A %	T %	G %	C %	A3 %	T3 %	G3 %	C3 %
T.chatareus	26.02	26.9	13.3	33.77	38.15	20.61	6.14	35.08
E.zonatum	27.63	31.72	12.86	27.77	38.59	27.19	8.33	25.87
J.floridae	23.97	30.55	11.4	34.06	28.07	29.82	5.7	36.4
P. orbicularis	25.14	25	12.28	37.57	33.77	15.78	4.82	45.61
L. menadoensis	33.33	25.14	11.84	29.67	53.94	10.52	5.26	30.26
Mean±SD	27.21±3.66	27.863.10	12.33±0.76	32.56±3.87	38.5±9.61	20.78±7.95	6.05±1.36	34.64±7.41
G.gallus	28.5	22.51	10.38	38.59	41.66	8.33	4.82	45.17
A.americana	28.36	23.24	11.4	36.98	40.35	8.33	7.45	43.85
V.chalybeata	31.14	22.95	10.67	35.23	49.56	7.45	5.26	37.71
F.peregrinus	30.11	24.12	9.21	36.54	42.98	11.4	2.19	43.42
S.sharpei	27.77	24.26	9.06	38.88	37.71	10.96	3.07	48.24
Mean±SD	29.17±1.39	23.41±0.75	10.14±0.99	37.24±1.50	42.45±4.42	9.29±1.76	4.55±2.04	43.67±3.83
C.familiaris	29.36	30.83	12.04	27.75	37.88	29.51	6.16	26.43
M.glis	31.56	34.51	10.76	23.15	40.7	37.71	4.38	16.66
R..norvegicus	32.15	28.19	11.01	28.63	44.49	22.9	4.4	28.19
D.novemcinctus	34.5	25.69	10.71	29.07	46.69	16.29	4.4	32.59
O.cuniculus	29.51	29.66	11.3	29.51	40.52	25.55	4.84	29.07
Mean±SD	31.41±2.11	29.77±3.26	11.16±0.54	27.62±2.58	42.05±3.50	26.39±7.94	4.83±0.76	26.58±5.98

In ATP8 gene of pisces, the nucleobases A and C occurred more frequently whereas in aves C and A occurred more frequently than T and G respectively. In mammals, A and T occurred more frequently than C and G (**Figure 4.4.9.**).

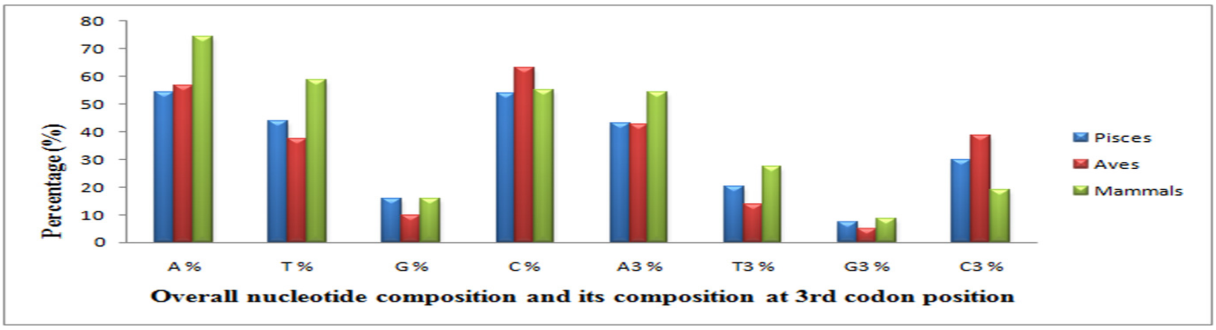


Figure 4.4.9 Nucleotide composition and its composition at 3rd codon position for ATP8 gene

The nucleotide A/C occurred most frequently at the third codon position (average) than T/G in pisces and aves respectively while in mammals, A/T occurred more frequently than C/A (**Table 4.4.2**). The mean±SD values of GC content were 41.56±2.48, 43.6±2.72 and 34.7±1.91 respectively in pisces, aves and mammals. In pisces, aves and mammals, the greatest difference of GC contents was found for the GC content at the codon 2nd and 3rd position as shown in **Figure 4.4.10**.

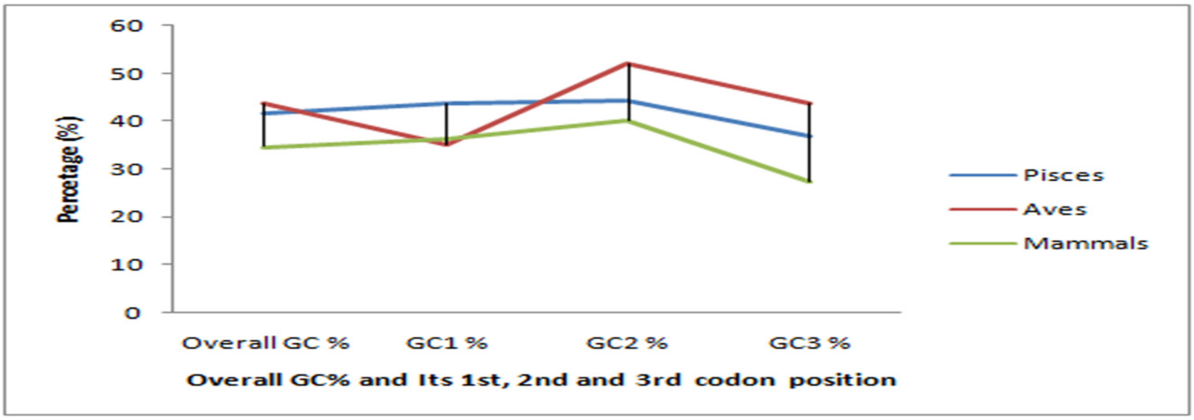


Figure 4.4.10 Overall GC% and its content at 1st, 2nd and 3rd codon position for ATP8 gene

Table 4.4.2 Nucleotide compositions among pisces, aves and mammals for ATP8 gene

Species	**A %**	**T %**	**G %**	**C %**	**A3 %**	**T3 %**	**G3 %**	**C3 %**
T. chatareus	55	43	16	54	37.5	23.2	8.9	30.4
E. zonatum	49	48	20	51	33.9	26.8	12.5	26.8
J. floridae	53	48	15	52	44.6	25	5.4	25
P. orbicularis	55	37	16	60	42.9	16.1	3.6	37.4
L. menadoensis	60	43	12	53	55.4	10.7	5.4	28.5
Mean±SD	54.4±3.9	43.8±4.5	15.8±2.8	54±3.5	42.86±8.2	20.36±6.7	7.16±3.5	**29.62±4.7**
G.gallus	57	40	8	60	41.8	16.4	3.6	38.2
A.americana	57	31	13	67	48.2	7.1	7.1	37.6
V.chalybeata	56	42	12	58	41.1	16.1	8.9	33.9
F.peregrinus	58	34	9	67	39.3	10.7	3.6	46.4
S.sharpei	56	41	7	64	42.9	17.9	1.8	37.4
Mean±SD	56.8±0.8	37.6±4.8	9.8±2.5	63.2±4.0	42.66±3.3	13.64±4.5	5±2.90	**38.7±4.62**
C.familiaris	71	61	17	55	36.8	29.4	10.3	23.5
M.glis	76	61	17	50	38.2	39.7	10.3	11.8
R..norvegicus	75	57	13	59	50	23.5	5.9	20.6
D.novemcinctus	70	58	16	60	44.1	25	7.4	23.5
O.cuniculus	80	57	15	52	57.4	19.1	8.8	14.7
Mean±SD	74.4±4.0	58.8±2.0	15.6±1.6	55.2±4.3	54.3±8.5	27.34±7.8	8.54±1.9	**18.82±5.3**

The overall nucleotide composition and the composition at the third codon position in different species for ATP6 and ATP8 genes suggest that compositional constraint under mutation pressure might be influencing the codon usage pattern of these genes in these species.

4.4.6 Interrelationships among different compositional features of MT-ATP gene among pisces, aves and mammals

We performed correlation analysis between general nucleotide composition and nucleotide composition at 3rd codon position to determine whether evolutionary process is driven by mutation pressure alone or by both mutation pressure and natural selection.

In ATP6 gene, highly significant positive correlation was found between A and A3 %, T% and T3%, G% and G3% (except pisces), C and C3 %, GC and GC3 % in pisces, aves and in mammals and negative correlation was observed for most of various nucleotide comparisons as shown in **Table 4.4.3.** These results suggest that the compositional constraint arising from mutation pressure and natural selection might determine the pattern of codon usage in MT-ATP6 gene.

Table 4.4.3 Correlation between overall nucleotide composition (%) and its composition at 3rd codon position in pisces, aves and mammals for ATP6 gene

	Nucleotide	**A3%**	**T3%**	**G3%**	**C3%**	**GC3%**
Pisces	A %	0.982**	-0.675	-0.015	-0.546	-0.633
	T %	-0.467	0.925*	0.791	-0.532	-0.443
	G %	0.071	0.004	0.496	-0.188	-0.112
	C %	-0.569	-0.103	-0.718	0.981**	0.977**
	GC %	-0.56	-0.102	-0.626	0.951*	0.962**
Aves	A %	0.944*	-0.259	-0.112	-0.910*	-0.989**
	T %	-0.438	0.899*	-0.647	0.438	0.091
	G %	0.316	-0.88*	0.969**	-0.477	0.044
	C %	-0.859	0.367	-0.21	0.935*	0.838
	GC %	-0.678	-0.202	0.426	0.647	0.893*
Mammals	A %	0.924*	-0.548	-0.688	0.273	0.186
	T %	-0.73	0.998**	0.171	-0.952*	-0.917*
	G %	-0.763	0.184	0.969**	0.095	0.213
	C %	0.326	-0.852	0.145	0.960**	0.963**
	GC %	0.154	-0.76	0.328	0.916*	0.943**

Note: One asterisk (*) and two asterisks (**) indicate correlations at a level of 0.05 and 0.01, respectively

Furthermore, positive correlation was observed in pisces between ENC and GC, ENC and GC3, GC and GC3%. In aves, positive correlation was observed between GC and GC3%. In mammals, positive correlation was observed between GC and GC3 (**Table 4.4.4**). These findings further suggest that the nucleotide compositions resulting from mutation pressure and natural selection are the main factors for synonymous codon usage in ATP6 gene.

Table 4.4.4 Correlation coefficients among ENC, CAI, GC contents

		Pisces		Aves		Mammals	
SL No	Correlation between	Correlation coefficient	p value	Correlation coefficient	p value	Correlation coefficient	p value
1	ENC and CAI	-0.483	0.41	0.621	0.263	0.06	0.924
2	ENC and GC %	0.966**	0.007	0.575	0.31	0.947*	0.014
3	ENC and GC3 %	0.917*	0.028	0.781	0.119	1	0
4	GC and GC3 %	0.962**	0.009	0.893*	0.041	0.943*	0.016
5	GC1 and GC3 %	0.785	0.116	-0.244	0.693	0.276	0.654
6	GC2 and GC3%	0.553	0.334	0.676	0.211	0.786	0.115
7	GC12 and GC3 %	0.827	0.084	0.097	0.876	0.549	0.338

Note: One asterisk (*) and two asterisks (**) indicate correlations at a level of 0.05 and 0.01, respectively.

In case of pisces, aves and mammals, the frequencies of those nucleotides were not same indicating that other factors such as natural selection might have played a role in codon usage pattern in ATP6 gene.

To determine the extent of mutation pressure against natural selection in the codon usage pattern in mitochondrial ATP6 gene, the neutrality plot was drawn. Neutrality plot is the regression of GC12 (average of GC1 and GC2) on GC3. In pisces, regression coefficient of GC12 on GC3 for MT-ATP6 gene is 0.360 which reveals that the relative neutrality is 36% and relative constraint is 64% for GC3 **(Figure 4.4.11a).** The GC12 was influenced by mutation pressure and natural selection with a ratio of 0.360/0.64= 0.56 in pisces. In aves, the regression coefficient of GC12 on GC3 is 0.027 indicating the relative neutrality of 2.7% and relative constraint of 97.3% for GC3 **(Figure 4.4.11b).** The GC12 in aves was affected by mutation pressure and natural selection with a ratio of 0.027/0.973=0.0277. In mammals, the regression coefficient of GC12 on GC3 is 0.151 which suggests that the relative neutrality is 15.1% and relative constraint is 84.9% for GC3 **(Figure 4.4.11c).** The GC12 was influenced by mutation pressure and natural selection with a ratio of 0.151/ 0.849=0.177 (Sueoka 1988) in mammals. These results suggest that natural selection played a major role while mutation pressure played a minor role in shaping the codon usage pattern of MT-ATP 6 gene.

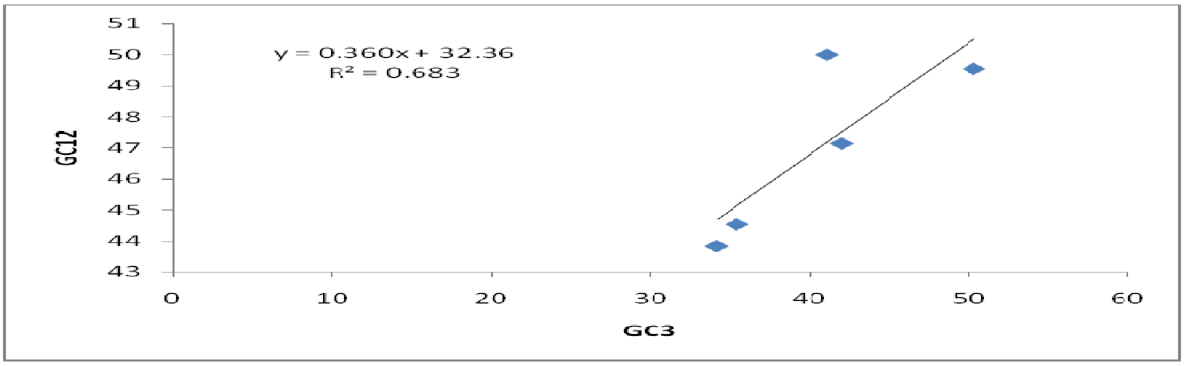


Figure 4.4.11 (a) Neutrality plot of GC12 with GC3 in different species of pisces for ATP6 gene

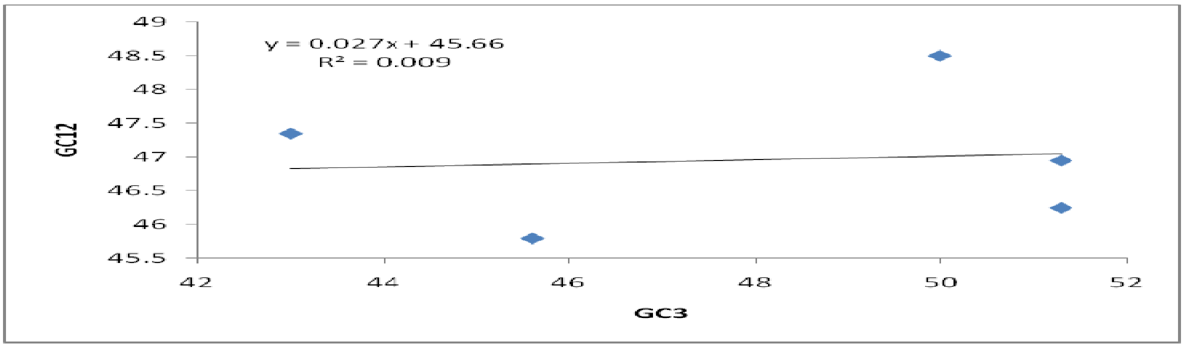


Figure 4.4.11(b) Neutrality Plot of GC12 with GC3 in different species of aves for ATP6 gene

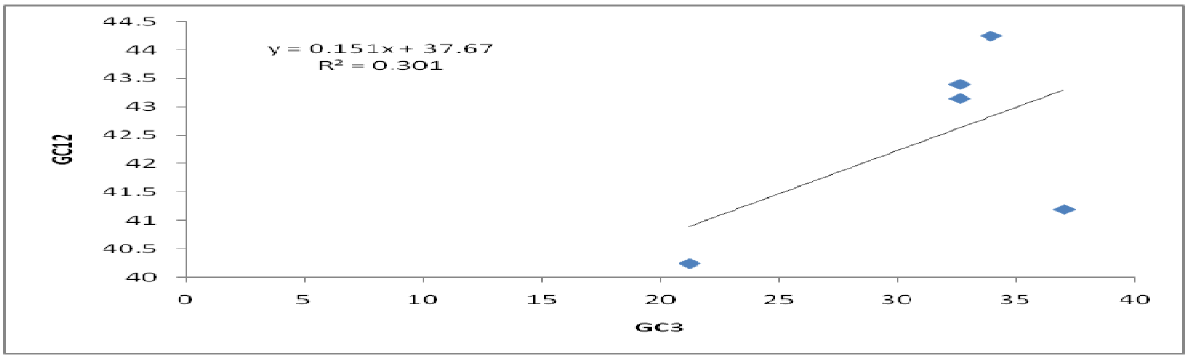


Figure 4.4.11(c) Neutrality plot of GC12 with GC3 in different species of mammals for ATP6 gene

In ATP8 gene, significant positive correlation was observed between C% and C3% in pisces and T% and T3%, G% and G3% in aves, while in mammals, no significant correlation was found. But significant negative correlation was observed between some of heterogeneous nucleotide comparisons in pisces, aves and mammals (**Table 4.4.5**). These results suggest that mutation pressure and natural selection might affect the codon usage pattern in different species of pisces, aves and mammals for ATP8 gene to some extent.

Table 4.4.5 Correlation between overall nucleotide composition (%) and its composition at 3rd codon position in pisces, aves and Mammals for ATP8 gene

	Nucleotide	**A3 %**	**T3 %**	**G3 %**	**C3 %**	**GC3%**
Pisces	A %	0.866	-0.896^{*}	-0.661	0.270	0.077
	T %	-0.243	0.659	0.608	-0.963^{**}	-0.113
	G %	-0.934^{*}	0.772	0.742	-0.040	0.054
	C %	0.096	-0.466	-0.639	0.967^{**}	0.015
	GC %	-0.563	0.143	-0.023	0.781	0.054
Aves	A %	-0.181	-0.633	-0.216	0.892^{*}	0.860
	T %	-0.502	0.973^{**}	-0.105	-0.529	-0.668
	G %	0.513	-0.636	0.927^{*}	-0.328	0.285
	C %	0.352	-0.727	-0.331	0.670	0.521
	GC %	0.588	-0.961^{**}	0.198	0.396	0.585
Mammals	A %	0.676	-0.201	0.056	-0.812	0.075
	T %	-0.907^{*}	0.862	0.828	-0.107	0.021
	G %	-0.730	0.652	0.876	-0.100	0.007
	C %	0.081	-0.474	0-.784	0.847	-0.083
	GC %	0.685	-0.692	-0.070	-0.059	0.790

Note: One asterisk (*) and two asterisks (**) indicate correlations at a level of 0.05 and 0.01, respectively

Furthermore, from the **Table 4.4.6**, correlation between ENC and various GC contents suggest that both mutation pressure and natural selection influenced the codon usage pattern for MT-ATP8 gene in pisces, aves and mammals.

From the neutrality plot (GC12 and GC3) of pisces as shown in **Figure 4.4.12 (a),** the regression coefficient of GC12 on GC3 for MT-ATP8 gene is 0.193, which reveals that the relative neutrality is 19.3% and relative constraint is 80.7% for GC3. The GC12 was influenced by mutation pressure and natural selection with a ratio of 0.193/0.807= 0.239. In aves, the regression coefficient of GC12 on GC3 is 0.106 indicating the relative neutrality of 10.6% and relative constraint of 89.4% for GC3 as shown in **Figure 4.4.12 (b).** The GC12 in aves was affected by mutation pressure and natural selection with a ratio of 0.106/0.894=0.118.

Table 4.4.6 Correlation coefficients among ENC, CAI, GC contents for ATP8 gene

		Pisces		Aves		Mammals	
SL No	Correlation between	Correlation coefficient	p value	Correlation coefficient	p value	Correlation coefficient	p value
1	ENC and CAI	-0.019	0.976	0.437	0.462	0.701	0.187
2	ENC and GC %	0.774	0.125	0.758	0.138	-0.032	0.960
3	ENC and GC3 %	0.995**	0	0.864	0.059	0.485	0.407
4	GC and GC3 %	0.824	0.086	0.585	0.300	0.790	0.112
5	GC1 and GC3 %	-0.044	0.944	0.231	0.709	-0.214	0.730
6	GC2 and GC3%	0.714	0.175	-0.029	0.963	-0.417	0.485
7	GC12 and GC3 %	0.379	0.530	0.127	0.838	-0.764	0.133

Note: One asterisk (*) and two asterisks (**) indicate correlations at a level of 0.05 and 0.01, respectively

In mammals, the regression coefficient of GC12 on GC3 is 0.271 which suggests that the relative neutrality is 27.1% and relative constraint is 72.9% for GC3 as shown in **Figure 4.4.12 (c).** The GC12 was influenced by mutation pressure and natural selection with a ratio of 0.271/ 0.729=0.371 (Sueoka 1988) in mammals. These results indicate that natural selection played a major role while mutation pressure played a minor role in shaping the codon usage pattern of MT-ATP8 gene.

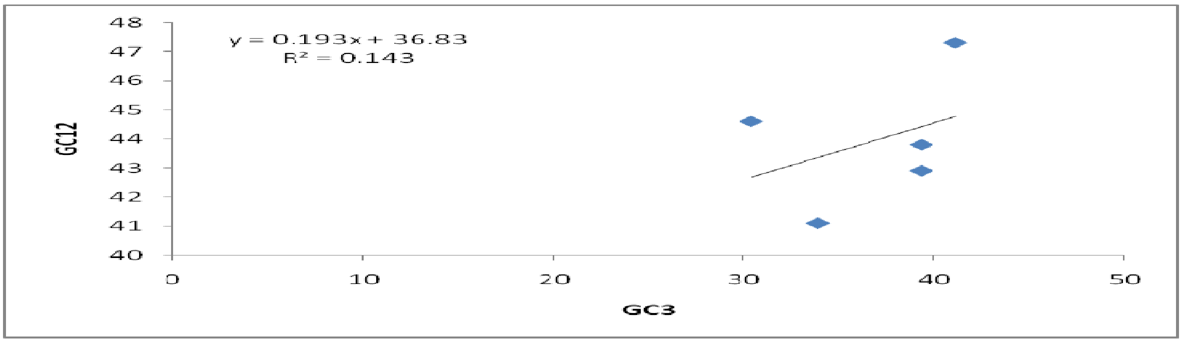


Figure 4.4.12 (a) Neutrality plot of GC12 with GC3 in different species of pisces for ATP8 gene

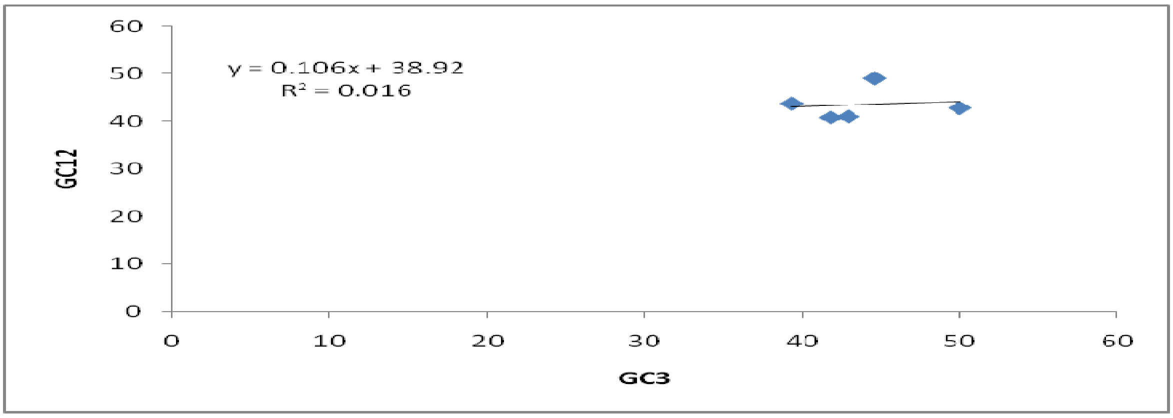


Figure 4.4.12 (b) Neutrality plot of GC12 with GC3 in different species of aves for ATP8 gene

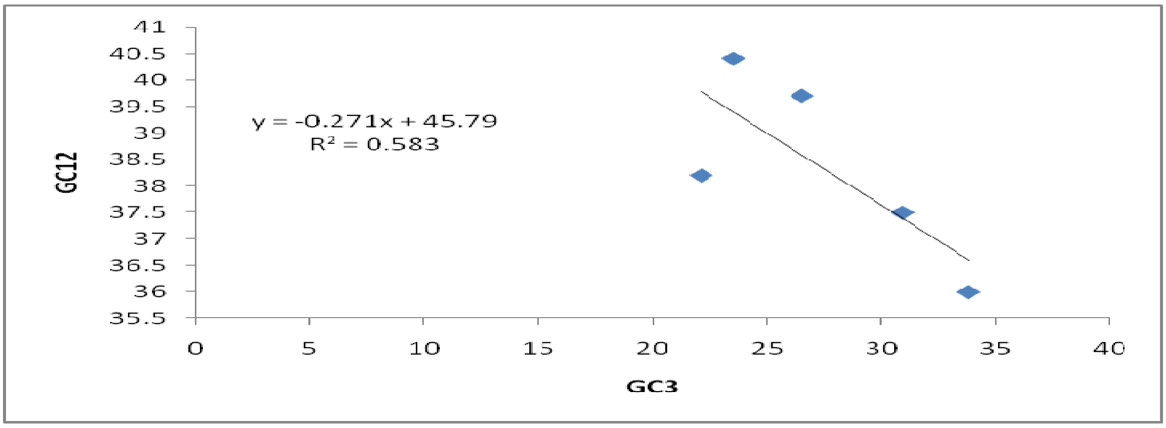


Figure 4.4.12 (c) Neutrality plot of GC12 with GC3 in different species of mammals for ATP8 gene

4.4.7 Correspondence analysis (COA) of MT-ATP gene among pisces, aves and mammals

COA is a multivariate statistical technique and widely used to determine the trends in codon usage variation among different species for MT-ATP6 gene. Therefore, we performed COA using RSCU values of codons. Based on the RSCU values of 60 synonymous codons, MT-ATP6 gene for different species was represented by the points in super dimensional axes and carried on to the correspondence analysis.

In pisces, aves and mammals, the contributions of the axes are shown in **Figure 4.4.13** and we observed four main contributors *i.e.* axis1, axis2, axis3 and axis4 for ATP6 gene.

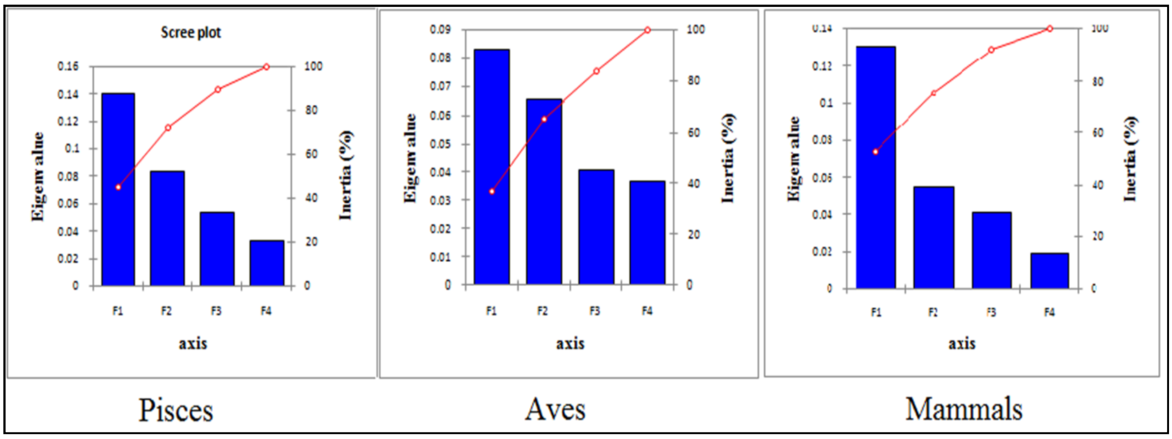
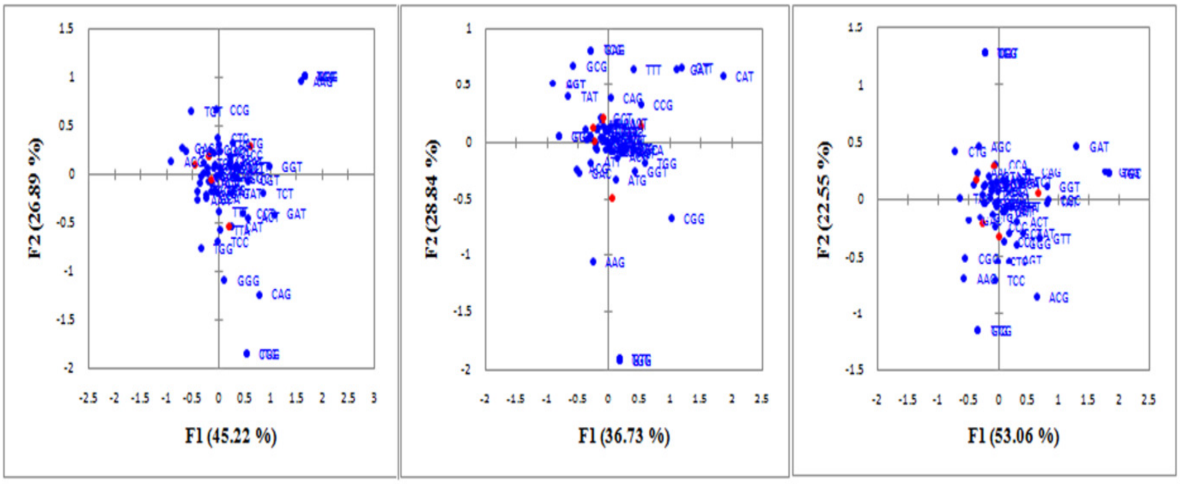


Figure 4.4.13 Contributions of the axes for MT-ATP6 in pisces, aves and mammals

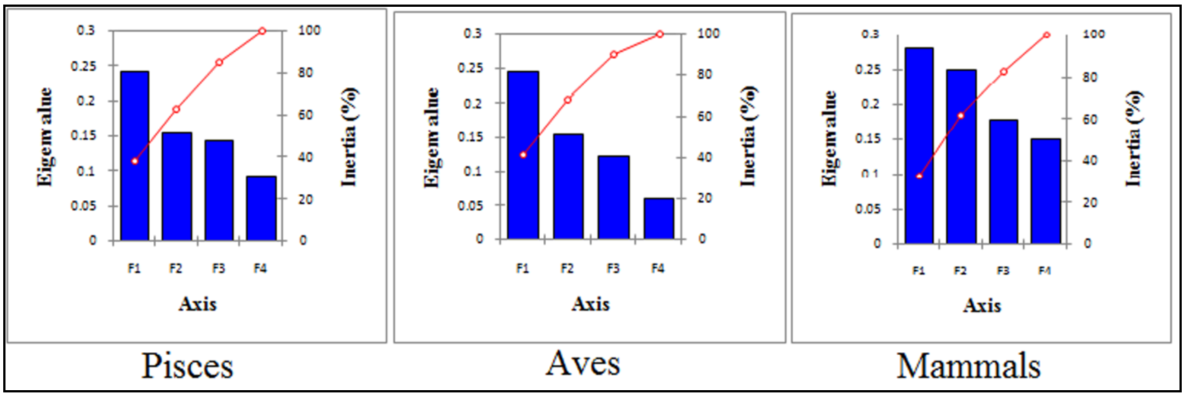
In ATP6 gene, the first two-main-dimensional coordinates, axis1 and axis2 (**Figure 4.4.14**) can explain 45.22% and 26.89% respectively of the total variation, which led to the first axis as the major contributor to the codon bias in pisces. In aves, axis1 contributes 36.73% of the total variation while axis2 (figure 2) contributes 28.84% respectively of the total variation, which led to the first axis as the major contributor to the codon bias. In mammals, axis1 and axis2 (Figure 2) can explain 36.73% and 28.84% respectively of the total variation, which revealed the first axis as the major contributor to the codon bias. The positions of most codons are more close to axes with a concentrate distribution (**Figure 4.4.14**), indicating that the base composition for mutation bias might correlate to the codon usage bias. A few codons are in a discrete distribution, indicating that there are many other factors which influence the codon usage e.g. role of natural selection



Pisces Aves Mammals

Figure 4.4.14 Correspondence analysis of the synonymous codon usage for ATP6 gene in pisces, aves and mammals

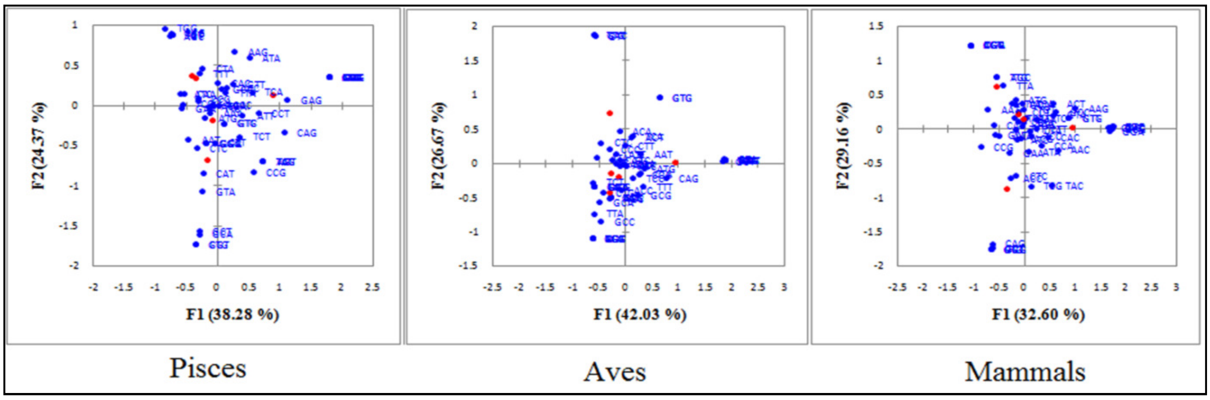
In ATP8 gene, the contributions of axes in pisces, aves and mammals were shown in **Figure 4.4.15** and the four main contributors to codon usage were axis1, axis2, axis3 and axis4 for ATP 8.



Pisces Aves Mammals

Figure 4.4.15 Contributions of the axes for MT-ATP8 gene in pisces, aves and mammals

In pisces, a major trend in the first axis (F1) accounted for 38.28% of total synonymous codon usage variation, and the second major trend in the second axis (F2) accounted for 24.37% of the total variation **(Figure 4.4.16).** In aves, the first axis is the major contributor (F1) which accounted for 38.28% of total synonymous codon usage variation, and the second axis (F2) accounted for 24.37% of the total variation. In mammals a major trend in the first axis (F1) accounted for 32.60% of total synonymous codon usage variation, and the second major trend in the second axis (F2) accounted for 29.16% of the total variation. The position of most codons is very close to axes with a concentrate distribution indicating that the base composition for mutation bias might correlate to the codon bias. A considerable amount of the genes are in a discrete distribution, which indicates that there are many other factor that might affect codon usage in ATP8 gene.



Pisces Aves Mammals

Figure 4.4.16 Correspondence analysis of the synonymous codon usage for ATP8 gene in pisces, aves and mammals

4.4.8 Amino acid composition in different species of pisces, aves and mammals for ATP protein

In ATP6 protein, the amino acid frequencies of the encoded proteins in different species of pisces, aves and mammals were estimated. The frequency of leucine residue was the highest in the proteins of pisces, aves and mammals while tyrosine, cysteine, histidine, lysine, aspartate and glutamate residues were lower in the proteins of MT-ATP6 protein as shown in **Figure 4.4.17.**

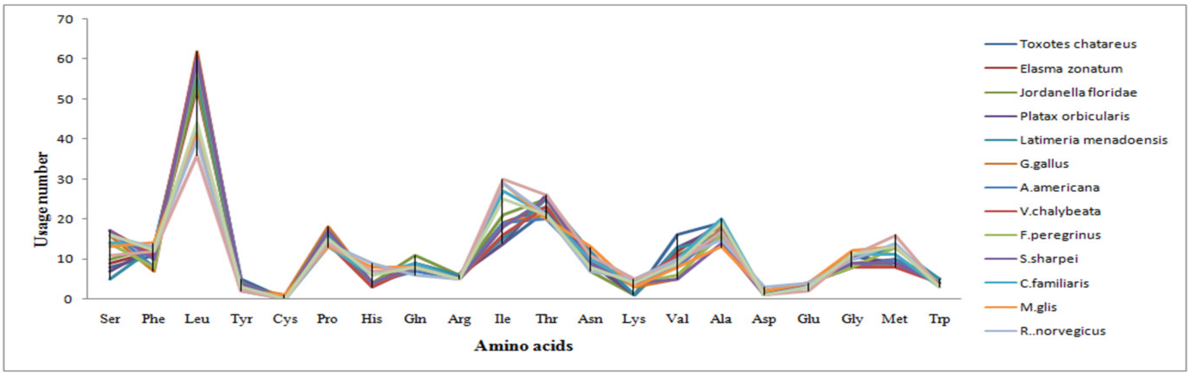


Figure 4.4.17 Comparison of amino acids in different species of pisces, aves and mammals for ATP6 protein

In ATP8 protein**,** the amino acid frequencies of the encoded proteins in different species of pisces, aves and mammals were estimated. The frequencies of leucine, proline and theonine residue were higher than other amino acids in composition of MT-ATP8 protein in pisces, aves and mammals. But tyrosine, cysteine, histidine, arginine, aspartate and glycine residues were lower in the proteins as shown in **Figure 4.4.18.**

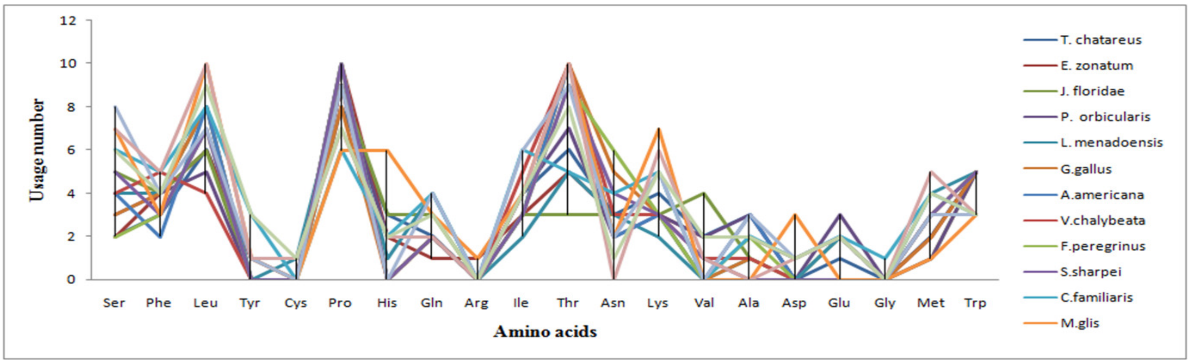


Figure 4.4.18 Comparison of amino acids in different species of pisces, aves and mammals for ATP8 protein

4.5 Comparative analysis of codon usage pattern in mitochondrial genes among the chordates

The pattern of codon usage differs among 13 mitochondrial protein-coding genes of pisces, aves and mammals. Here the average of RSCU values of codons in different species of pisces, aves and mammals was compared for each gene.

In ND1 gene, 26 codons were more frequently used in pisces, out of which 13 codons end with A, 2 codons end with T and 11 codons end with C. In aves, out of 28 more frequently used codons, 13 codons end with A, 3 codons end with T, 1 codon ends with G and 13 codons end with C. In mammals, 28 codons were more frequently used, out of which 14 codons end with A, 1 codons end with T and 13 codons end with C.

In ND2 gene, 30 codons were more frequently used in pisces, out of which 13 codons end with A, 3 codons end with T, 1 codon ends with G and 13 codons end with C. In aves, out of 28 more frequently used codons, 13 codons end with A, and 15 codons end with C. In mammals, 21 codons were more frequently used, out of which 14 codons end with A and 7 codons end with C.

In ND3 gene, 27 codons were more frequently used in pisces, out of which 12 codons end with A, 5 codons end with T, 1 codon ends with G and 9 codons end with C. In aves, out of 23 more frequently used codons, 13 codons end with A, and 10 codons end with C. In mammals, 31 codons were more frequently used, out of which 13 codons end with A, 8 codons end with T, and 10 codons end with C.

In ND4 gene, 28 codons were more frequently used in pisces in which 12 codons end with A, 6 codons end with T, and 10 codons end with C. In aves, out of 26 more frequently used, 10 codons end with A, and 16 codons end with C. In mammals, 26 codons were more frequently used, out of which 12 codons end with A, 3 codons end with T, and 11 codons end with C.

Table 4.5.1 RSCU values of codons for different mitochondrial genes in ND gene among pisces, aves and mammals

	MT-ND1			MT-ND2			MT-ND3			MT-ND4		
Codon	Pisces	Aves	Mammals	Pisces	Aves	Mammals	Pisces	Aves	Mammals	Pisces	Aves	Mammals
TCA	**2.568**	**1.968**	**2.568**	**2.28**	**1.956**	**2.832**	**1.476**	**2.568**	**1.896**	**1.992**	**1.464**	**2.268**
TCT	0.684	0.54	0.744	**1.032**	0.456	0.924	**1.008**	0.324	**1.32**	**1.092**	0.936	**1.2**
TCC	**1.692**	**2.016**	**1.728**	**1.476**	**2.52**	**1.536**	**2.532**	**2.22**	**1.62**	**1.752**	**1.668**	**1.92**
TCG	0.048	0.132	0.192	0.68	0.144	0.048	0	0	0	0.18	0.108	0.12
AGC	0.564	**1.26**	0.516	0.888	0.888	0.456	0.408	0.396	0.72	0.528	**1.716**	0.48
AGT	0.42	0.096	0.264	0.168	0.036	0.204	0.204	0.684	0.468	0.432	0.096	0
TTC	**1.248**	**1.488**	**1.104**	**1.648**	**1.668**	**1.032**	0.932	**1.068**	0.872	0.908	**1.524**	0.74
TTT	0.752	0.512	0.896	0.912	0.332	0.968	0.964	0.632	**1.132**	**1.092**	0.48	**1.26**
TTA	0.9	0.504	**1.272**	0.868	0.576	**1.104**	**1.14**	0.328	**1.428**	0.912	0.636	**1.164**
TTG	0.084	0.048	0.072	**1.1**	0.012	0.048	0	0.084	0.228	0.096	0	0
CTA	**2.376**	**2.832**	**2.604**	**2.028**	**2.784**	**2.88**	**1.572**	**3.084**	**1.236**	**2.16**	**3.324**	**3**
CTC	**1.308**	**1.404**	**1.188**	**1.44**	**1.608**	0.972	**1.668**	**1.32**	**1.98**	**1.14**	**1.128**	0.708
CTG	0.312	0.552	0.132	0.36	0.468	0.168	0.492	0.588	0.324	0.6	0.528	0.264
CTT	0.996	0.672	0.768	**1.152**	0.54	0.864	**1.116**	0.7204	0.84	**1.092**	0.396	0.864
TAC	0.988	**1.48**	**1.18**	**1.272**	**1.508**	**1.172**	0.404	**1.4**	0.932	0.004	**1.404**	**1.068**
TAT	**1.012**	0.52	0.82	0.732	0.496	0.828	**1.604**	0.608	**1.068**	0.4	0.608	0.936
TGT	0	0.148	0.812	0.012	0.016	0.004	0.416	0.416	**1.008**	0.796	0.556	0.528
TGC	0	**1.868**	**1.208**	**1.2**	**2**	0.4	**1.604**	**1.208**	**1.008**	**1.204**	**1.444**	**1.472**
CCA	**1.496**	**1.8**	**1.896**	**1.272**	**1.88**	**2.12**	**1.272**	**1.92**	**1.488**	**1.416**	**1.192**	**2.808**
CCC	**1.48**	**1.6**	**1.232**	**1.8**	**1.56**	**1.288**	**1.856**	**1.312**	**1.384**	**1.808**	**2.056**	0.432
CCG	0.096	0.136	0.08	0	0.08	0.04	0	0.224	0	0.024	0.016	0.432
CCT	0.912	0.472	0.792	0.92	0.472	0.568	0.904	0.544	**1.144**	0.808	0.768	0.432
CAT	0.38	0.668	0.7	0.376	0.596	0.248	0.408	0.004	0	0.692	0.336	0.212
CAC	**1.62**	**1.34**	**1.3**	**1.624**	**1.404**	**1.752**	0.804	0.4	0	**1.308**	**1.664**	**1.8**
CAA	**1.56**	**1.664**	**1.72**	**1.876**	**1.588**	**1.756**	**1.8**	**1.668**	**1.536**	**1.868**	**2**	**2**
CAG	0.44	0.336	0.28	0.124	0.412	0.244	0.212	0.332	0.464	0.136	0.016	0.02
CGA	**2.616**	**2.456**	**2.36**	**2.88**	**2.264**	**2.626**	**2.408**	**3.2**	0.832	**2.128**	**2.808**	**1.624**
CGC	0.608	0.784	0.8	0.56	**1.064**	0.612	0.832	0.432	**2.416**	**1.88**	**1.208**	**2.416**
CGG	0.2	0.408	0.472	0.16	0.2	0	0.432	0.04	0.04	0.024	0.024	0.04
CGT	0.608	0.384	0.4	0.4	0.464	0.012	0.432	0.432	0.832	0.024	0.024	0.04
ATG	0.128	0.5	0.192	0.452	0.496	0.156	**1.132**	0.18	0.124	0.932	0.652	0.268
ATA	**1.64**	**1.5**	**1.808**	**1.548**	**1.504**	**1.844**	0.872	**1.832**	**1.876**	**1.068**	**1.348**	**1.74**
ATC	0.98	**1.276**	0.968	0.916	**1.388**	0.986	0.8	**1.636**	**1**	0.208	**1.368**	0.844
ATT	**1.02**	0.724	**1.032**	**1.084**	0.612	0.714	**1.2**	0.364	1	1	0.632	**1.156**
ACA	**1.496**	**1.536**	**1.856**	**1.312**	**1.528**	**1.506**	0.96	**1.392**	**1.816**	**1.44**	**1.672**	**1.56**
ACC	**1.68**	**1.872**	**1.24**	**1.68**	**1.648**	0.984	**2.096**	**1.696**	**1.032**	**1.48**	**1.848**	**1.56**
ACG	0.304	0.144	0.168	0.08	0.104	0.188	0.088	0.072	0.088	0	0	0.008
ACT	0.484	0.456	0.744	0.92	0.72	0.58	0.848	0.688	**1.056**	**1.072**	0.488	0.904
AAC	**1.028**	**1.608**	**1.02**	**1.432**	**1.872**	**1.16**	**1.004**	**1.8**	**1.04**	**2**	**1.736**	**1.168**
AAT	0.696	0.392	0.98	0.568	0.128	0.54	0.604	0.212	0.96	0.008	0.268	0.764
GAC	**1.668**	**1.7**	**1.536**	**1.204**	**2**	0.8	1.46	1.868	1.8	0.808	**1.208**	**1.208**
GAT	0.332	0.3	0.464	0.804	0.02	0.008	0.54	0.132	0.212	**1.204**	0.812	0.812
GTA	**1.464**	**1.704**	**1.352**	**1.408**	**1.912**	**1.314**	**1.336**	**3.736**	**1.736**	**2.6**	**2.488**	**2.2**
GTC	**1.28**	**1.584**	**1.544**	**1.576**	**1.664**	0.848	**1.072**	0.296	**1.064**	0.56	**1.368**	**1.128**
GTG	0.296	0.352	0.344	0.456	0.288	0.27	0.2	0.032	0.008	0.224	0.184	0
GTT	0.968	0.368	0.76	0.552	0.136	0.818	1	0.032	**1.208**	0.68	0.024	0.68
GCA	**1.52**	**1.248**	**1.576**	**1.232**	**1.312**	**1.288**	**1.864**	**1.632**	**1.568**	**1.464**	0.88	**2**
GCC	**1.592**	**2.104**	**1.832**	**1.856**	**1.912**	**1.332**	**1.344**	**2.056**	**1.744**	**1.728**	**2.256**	**1.24**
GCG	0.12	0.048	0	0.096	0.16	0.048	0.088	0	0.328	0.12	0.112	0
GCT	0.744	0.616	0.584	0.808	0.608	0.574	0.592	0.312	0.352	0.688	0.744	0.76
AAA	**1.78**	**1.772**	**1.836**	**1.592**	**1.936**	**1.58**	**1.604**	**1.604**	**1.568**	0.4	0	0.4
AAG	0.224	0.228	0.168	0.408	0.064	0.12	0.416	0.416	0.432	0.004	0	0.004
GAA	**1.604**	**1.532**	**1.7**	**1.764**	**1.58**	**1.224**	**1.772**	**1.864**	**1.932**	**1.468**	**1.472**	**1.8**
GAG	0.396	0.468	0.3	0.236	0.42	0.476	0.228	0.136	0.068	0.54	0.536	0.216
GGA	**1.304**	**1.44**	**1.584**	**1.48**	**1.712**	**2.088**	**1.096**	**2.88**	**1.928**	**1.552**	0.536	**1.8**
GGC	**1.504**	**1.496**	**1.52**	**1.336**	**1.952**	0.882	**2.504**	0.8	0.936	**1.456**	**2.464**	**1.2**
GGG	0.824	0.64	0.512	0.552	0.272	0.056	0	0.16	0.664	0	0.664	0.2
GGT	0.368	0.416	0.368	0.64	0.064	0.232	0.4	0.16	0.464	0.992	0.736	0.8
TGG	0.156	0.204	0.232	0.14	0.288	0.09	0.24	0.4	0.1	0.416	0.412	0
TGA	**1.856**	**1.804**	**1.768**	**1.86**	**1.712**	**1.61**	**1.76**	**1.6**	**1.9**	**1.604**	**1.204**	0

Bold indicates more frequently used codons

In ND41 gene, 26 codons were more frequently used in pisces, out which 11 codons end with A, 5 codons end with T and 10 codons end with C. In aves, out of 26 more frequently used, codons 10 codons end with A, and 16 codons end with C. In mammals, 24 codons were more frequently used, out of which 12 codons end with A, 3 codons end with T and 11 codons end with C.

In ND5 gene, 30 codons were more frequently used in pisces in which 6 codons end with A/T, 7 codons end with G and 11 codons end with C. In aves, out of 29 more frequently used codons, 5 codons end with A, 6 codons end with T, 7 codons end with G and 12 codons end with C. In mammals, 31 codons were more frequently used, out of which 7 codons end with A, 6 codons end with T, 7 codons end with G and 11 codons end with C.

In ND6 gene, 23 codons were more frequently used in pisces, out of which 2 codons end with A, 5 codons end with T, 5 codons end with G and 11 codons end with C. In aves, out of 20 more frequently used codons, 1 codon ends with A, 6 codons end with T, 5 codons end with G and 8 codons end with C. In mammals, 25 codons were more frequently used, out which 3 codons end with A, 6 codons end with T/G, and 10 codons end with C.

Table 4.5.2 RSCU values of codons for different mitochondrial genes in ND gene among pisces, aves and mammals

	MT-ND4l			MTND5			MT-ND6		
Codon	Pisces	Aves	Mammals	Pisces	Aves	Mammals	Pisces	Aves	Mammals
TCA	**1.992**	**1.464**	**2.268**	**1.596**	**1.74**	**2.304**	0.804	0.792	**1.104**
TCT	**1.092**	0.936	**1.2**	**1.548**	**2.412**	**1.224**	**1.92**	**1.548**	**1.284**
TCC	**1.752**	**1.668**	**1.92**	0.168	0.06	0.096	0.384	0.216	0
TCG	0.18	0.108	0.12	0.936	0.576	**1.128**	0.072	0	0
AGC	0.528	**1.716**	0.48	**1.392**	**1.116**	0.816	**2.052**	**2.112**	**1.704**
AGT	0.432	0.096	0	0.336	0.108	0.456	0.78	**1.308**	**1.92**
TTC	0.908	**1.524**	0.74	**1.088**	**1.556**	**1.2**	**1.2**	0.4	**1.44**
TTT	**1.092**	0.48	**1.26**	0.912	0.444	0.8	0.012	0.004	0.56
TTA	0.912	0.636	**1.164**	0.72	0.54	**1.152**	0.6	0.3	0.444
TTG	0.096	0	0	0.084	0.024	0.084	0	0	0.12
CTA	**2.16**	**3.324**	**3**	**2.496**	**2.928**	**2.28**	0.204	0.936	0.156
CTC	**1.14**	**1.128**	0.708	**1.224**	**1.608**	**1.128**	**4.896**	**3.984**	**4.176**
CTG	0.6	0.528	0.264	0.396	0.372	0.18	0	0.48	0
CTT	**1.092**	0.396	0.864	**1.092**	0.528	**1.128**	0.3	0.3	**1.116**
TAC	0.004	**1.404**	**1.068**	**1.416**	**1.776**	**1.272**	**1.628**	**1.708**	**1.172**
TAT	0.4	0.608	0.936	0.584	0.224	0.732	0.376	0.296	0.828
TGT	0.796	0.556	0.528	**1.64**	**1.596**	**1.512**	**1.88**	**1.868**	0.208
TGC	**1.204**	**1.444**	**1.472**	0.36	0.404	0.488	0.12	0.14	1.4
CCA	**1.416**	**1.192**	**2.808**	**1.184**	**1.576**	**1.672**	0.824	0.544	0.712
CCC	**1.808**	**2.056**	0.432	**1.328**	**1.808**	**1.16**	**2.912**	**2.84**	**3.16**
CCG	0.024	0.016	0.432	0.24	0.128	0.12	0.208	0.576	0.088
CCT	0.808	0.768	0.432	**1.248**	0.464	**1.064**	0.056	0.024	0.04
CAT	0.692	0.336	0.212	0.532	0.344	0.724	0.84	0.628	**1.108**
CAC	**1.308**	**1.664**	**1.8**	**1.468**	**1.656**	**1.276**	**1.16**	**1.372**	0.892
CAA	**1.868**	**2**	**2**	**1.672**	**1.792**	**1.784**	**1.692**	**1.372**	**1.816**
CAG	0.136	0.016	0.02	0.328	0.212	0.216	0.312	0.632	0.184
CGA	**2.128**	**2.808**	**1.624**	**3**	**2.368**	**1.928**	**1.224**	0.904	0.64
CGC	**1.88**	**1.208**	**2.416**	0.744	**1.464**	**1.36**	**2.152**	**2.512**	**2.416**
CGG	0.024	0.024	0.04	0	0.088	0.248	0.304	0.384	0.584
CGT	0.024	0.024	0.04	0.256	0.104	0.464	0.328	0.216	0.36
ATG	0.932	0.652	0.268	**1.696**	**1.804**	**1.732**	**1.668**	**1.404**	**1.668**
ATA	**1.068**	**1.348**	**1.74**	0.304	0.196	0.268	0.34	0.612	0.34
ATC	0.208	**1.368**	0.844	**1.496**	**1.732**	**1.824**	**1.736**	**2**	**2**
ATT	1	0.632	**1.156**	0.992	**1.448**	0.9	**1.504**	**1.428**	**1.32**
ACA	**1.44**	**1.672**	**1.56**	**1.008**	0.552	**1.1**	0.496	0.572	0.68
ACC	**1.48**	**1.848**	**1.56**	**1.72**	**1.472**	**1.72**	0.496	0.632	0.392
ACG	0	0	0.008	**1.496**	**1.832**	**1.192**	**2.016**	**2.256**	**2.144**
ACT	**1.072**	0.488	0.904	0.08	0.056	0.12	0.576	0.272	0.304
AAC	**2**	**1.736**	**1.168**	0.712	0.648	0.968	0.912	0.84	**1.184**
AAT	0.008	0.268	0.764	**1.368**	**1.704**	**1.268**	**1.452**	**1.74**	0.872
GAC	0.808	**1.208**	**1.208**	0.632	0.296	0.732	0.548	0.26	**1.128**
GAT	**1.204**	0.812	0.812	**1.792**	**1.812**	**1.804**	**1.308**	**1.604**	**1.86**
GTA	**2.6**	**2.488**	**2.2**	0.208	0.188	0.2	0.692	0.396	0.14
GTC	0.56	**1.368**	**1.128**	**1.544**	**2.016**	**2.064**	**2.672**	**1.208**	**1.88**
GTG	0.224	0.184	0	**1.256**	**1.168**	0.84	0.008	0.816	**1.128**
GTT	0.68	0.024	0.68	0.288	0.384	0.28	0.008	0.024	0.016
GCA	**1.464**	0.88	**2**	0.928	0.432	0.808	0.536	0.416	0.216
GCC	**1.728**	**2.256**	**1.24**	**1.496**	**1.296**	**1.304**	**1.016**	0.416	0.672
GCG	0.12	0.112	0	**1.896**	**2.112**	**1.744**	**2.6**	**1.208**	**1.664**
GCT	0.688	0.744	0.76	0.048	0.072	0.136	0.024	0.024	0.016
AAA	0.4	0	0.4	0.544	0.528	0.808	0.416	0.816	0.872
AAG	0.004	0	0.004	**1.344**	**1.68**	**1.156**	**1.144**	**1.2**	**1.004**
GAA	**1.468**	**1.472**	**1.8**	0.656	0.324	0.844	0.064	0.408	1.008
GAG	0.54	0.536	0.216	**1.732**	**1.76**	**1.676**	0.804	0.604	**1.2**
GGA	**1.552**	0.536	**1.8**	0.268	0.24	0.328	0.404	0.604	0.012
GGC	**1.456**	**2.464**	**1.2**	**1.568**	**2.128**	**1.768**	**1.48**	0.976	0.416
GGG	0	0.664	0.2	**1.424**	**1.48**	**1.248**	**1.696**	**1.528**	0.416
GGT	0.992	0.736	0.8	0.592	0.264	0.376	0.84	0.736	0.808
TGG	0.416	0.412	0	0.424	0.12	0.608	0.016	0.024	0.808
TGA	**1.604**	**1.204**	0	0.504	0.268	0.18	0.268	0	0

Bold indicates more frequently used codons

In COI gene, 31 codons were more frequently used in pisces, out of which 13 codons end with A, 6 codons end with T, and 12 codons end with C. In aves, out of 24 more frequently used, 12 codons end with A, and 12 codons end with C. In mammals, 29 codons were more frequently used, out of which 14 codons end with A, 9 codons end with T, and 6 codons end with C.

In COII gene, 19 codons were more frequently used in pisces, out of which 7 codons end with A, 4 codons end with T, and 8 codons end with C. In aves, out of 17 more frequently used, 7 codons end with A, and 10 codons end with C. In mammals, 17 codons were more frequently used, out of which 8 codons end with A, 5 codons end with T, and 4 codons end with C.

In COIII gene, 31 codons were more frequently used in pisces, out of which 12 codons end with A, 5 codons end with T and 14 codons end with C. In aves, out of 27 more frequently used codons 13 codons end with A, and 14 codons end with C. In mammals, 27 codons were more frequently used, out which 13 codons end with A, 6 codons end with T, and 8 codons end with C.

In CYB gene, 29 codons were more frequently used in pisces, out of which 13 codons end with A, 1 codon ends with T and 15 codons end with C. In aves, out of 28 more frequently used, 13 codons end with A, and 15 codons end with C. In mammals, 29 codons were more frequently used, out of which 14 codons end with A, 1 codon ends with T, and 14 codons end with C.

Table 4.5.3 RSCU values of codons for different mitochondrial genes in CO and CYB gene among pisces, aves and mammals

	MT-COI			MT-COII			MT-COIII			MT-CYB		
Codon	Pisces	Aves	Mammals	Pisces	Aves	Mammals	Pisces	Aves	Mammals	Pisces	Aves	Mammals
TCA	**2.064**	**2.36**	**2.136**	**1.584**	**1.97**	**2.16**	**2.232**	**1.67**	**2.496**	**2.388**	**2.7**	**2.832**
TCT	**1.068**	0.85	**1.332**	**1.308**	0.37	**1.272**	**1.368**	0.78	**1.164**	0.794	0.7	**1.128**
TCC	**1.584**	**1.91**	**1.668**	**1.476**	**2.65**	**1.152**	**1.092**	**1.87**	**1.164**	**1.992**	**2.1**	**1.128**
TCG	0.216	0.26	0.192	0.24	0.17	0.06	0.084	0.23	0.18	0.372	0.1	0.156
AGC	**1.08**	0.53	0.432	**1.164**	0.7	**1.152**	0.804	**1.43**	0.816	0.432	0.3	0.612
AGT	0	0.08	0.192	0.24	0.06	0.204	0.444	0.07	0.18	0.06	0	0.144
TTC	**1.112**	**1.48**	**1.096**	0.88	**1.42**	0.836	**1.34**	**1.43**	0.94	**1.48**	**1.6**	**1.288**
TTT	0.888	0.52	0.904	**1.12**	0.58	**1.168**	0.66	0.57	**1.06**	0.52	0.4	0.712
TTA	0.744	0.66	**1.38**	0.768	0.43	**1.188**	0.864	0.77	**1.152**	0.756	0.2	**1.164**
TTG	0.168	0.26	0.108	0.084	0	0.144	0.216	0.34	0.036	0.18	0.1	0.168
CTA	**1.944**	**2.83**	**2.076**	**2.532**	**2.63**	**2.448**	**2.04**	**2.44**	**2.748**	**2.28**	**3.1**	**2.592**
CTC	**1.224**	**1.14**	0.756	**1.008**	**1.55**	0.744	**1.2**	**1.69**	0.828	**1.416**	**1.6**	**1.236**
CTG	0.408	0.59	0.384	0.3	0.58	0.216	0.396	0.47	0.288	0.372	0.4	0.112
CTT	**1.488**	0.53	**1.296**	**1.308**	0.82	**1.272**	**1.272**	0.29	0.96	0.996	0.6	0.696
TAC	**1.252**	**1.41**	0.8	**1.244**	**1.35**	0.78	**1.36**	**1.74**	0.996	**1.296**	**1.7**	**1.132**
TAT	0.748	0.59	**1.2**	0.76	0.65	**1.22**	0.64	0.26	**1.004**	0.704	0.3	0.868
TGT	0.02	0.22	**1.2044**	0.412	0.4	0.408	0.212	0.02	0.34	0	0.5	0.18
TGC	**2**	**1.8**	0.812	**1.6**	**1.6**	**1.6**	**1.8**	**1.6**	**1.668**	**2**	**1.5**	**1.82**
CCA	**1.184**	**1.64**	**1.496**	**1.336**	**2.03**	**2.24**	0.968	**2.05**	**1.552**	**1.712**	**2.2**	**2.112**
CCC	**1.096**	**1.59**	**1.024**	**1.536**	**1.38**	0.872	**1.968**	**1.4**	0.792	**1.328**	**1.5**	0.952
CCG	0.44	0.28	0.112	0.224	0.12	0	0	0.13	0.136	0.16	0.1	0.152
CCT	**1.296**	0.5	**1.368**	0.92	0.47	0.872	**1.064**	0.42	**1.504**	0.8	0.2	0.792
CAT	0.66	0.46	0.832	0.728	0.4	0.888	0.404	0.51	0.76	0.464	0.3	0.588
CAC	**1.34**	**1.54**	**1.168**	**1.272**	**1.6**	**1.116**	**1.596**	**1.5**	**1.24**	**1.536**	**1.7**	**1.412**
CAA	**1.644**	**1.54**	**1.684**	**1.656**	**1.93**	**1.548**	**1.768**	**1.66**	**1.624**	**1.932**	**1.9**	**1.932**
CAG	0.368	0.46	0.316	0.356	0.07	0.452	0.236	0.34	0.38	0.068	0.2	0.068
CGA	**2.296**	**2.66**	**2.752**	**2.136**	**1.9**	**2.536**	**2.72**	**2.62**	**3.2**	**2.064**	**2.5**	**2.104**
CGC	0.192	0.67	0	**1.2**	**1.67**	0.652	**1.124**	0.42	0.32	**1.04**	**1**	**1.312**
CGG	0.192	0.1	0.296	0.272	0.16	0.136	0.32	0.16	0	0.408	0.4	0.208
CGT	**1.328**	0.58	0.976	0.4	0.27	0.664	0.48	0	0.48	0.52	0.1	0.408
ATG	0.808	0.44	0.436	0.602	0.52	0.34	0.536	0.59	0.304	0.61	0.4	0.464
ATA	**1.192**	**1.56**	**1.564**	**1.332**	**1.46**	**1.66**	**1.472**	**1.41**	**1.704**	**1.28**	**1.6**	**1.536**
ATC	0.824	1.25	0.692	0.864	1.45	0.812	0.744	1.62	0.86	**1.132**	**1.5**	**1.136**
ATT	**1.172**	0.75	**1.308**	**1.136**	0.55	**1.188**	**1.256**	0.38	**1.144**	0.868	0.5	0.864
ACA	**2.088**	**1.51**	**2.008**	**2.24**	**1.59**	**2.328**	**2.064**	**1.73**	**1.864**	**1.648**	**1.4**	**1.896**
ACC	0.856	1.75	0.808	0.8	**1.54**	0.96	**1.232**	**1.53**	**1.032**	**1.72**	**2.1**	**1.44**
ACG	0.176	0.09	0.128	0.28	0.1	0.056	0.08	0.22	0.16	0.144	0.1	0.032
ACT	0.864	0.66	**1.056**	0.68	0.77	0.672	0.632	0.52	0.944	0.496	0.5	0.624
AAC	**1.104**	**1.67**	0.972	**1.74**	**1.52**	**1.156**	**1.076**	**1.52**	**1.392**	**1.44**	**1.6**	**1.52**
AAT	0.9	0.33	**1.032**	0.26	0.26	0.844	0.94	0.49	0.608	0.56	0.4	0.484
GAC	**1.408**	**1.42**	**1.096**	**1.528**	**1.48**	**1.464**	**1.308**	**1.54**	0.868	**1.492**	**1.8**	**1.52**
GAT	0.592	0.58	0.904	0.472	0.52	0.536	0.692	0.46	**1.132**	0.508	0.2	0.48
GTA	**1.784**	**1.88**	**2.328**	**1.456**	**1.92**	**1.6**	**1.496**	**1.78**	**2.168**	**1.432**	**1.8**	**2.04**
GTC	0.864	1.18	0.728	**1.032**	**1.18**	**1.016**	**1.168**	**1.3**	**1.008**	**1.176**	**1.6**	**1.072**
GTG	0.512	0.44	0.32	0.384	0.32	0.304	0.256	0.42	0.304	0.208	0.2	0.08
GTT	0.848	0.49	0.632	**1.12**	0.58	**1.064**	**1.072**	0.5	0.544	**1.176**	0.4	0.808
GCA	**1.28**	**1.22**	**1.248**	**1.68**	**1.02**	0.976	**1.576**	**1.32**	**1.192**	**1.656**	**1.2**	**1.68**
GCC	**1.552**	**1.93**	**1.376**	**1.328**	**2.36**	**1.832**	**1.568**	**1.67**	**1.976**	**1.56**	**2.3**	**1.6**
GCG	0.136	0.02	0.136	0	0	0.288	0	0.08	0.096	0.096	0.1	0.1
GCT	**1.024**	0.83	**1.232**	**1.008**	0.62	0.904	0.832	0.51	0.72	0.672	0.4	0.68
AAA	**1.904**	1.62	**1.864**	**1.6**	**1.82**	**1.8**	**1.604**	**1.53**	**1.868**	**1.604**	**1.8**	**1.828**
AAG	0.104	0.38	0.14	0.4	0.18	0.2	0.404	0.47	0.144	0.396	0.2	0.172
GAA	**1.748**	**1.62**	**1.688**	**1.864**	**1.73**	**1.656**	**1.828**	**1.57**	**1.848**	**1.592**	**1.8**	**1.732**
GAG	0.252	0.38	0.312	0.136	0.27	0.344	0.172	0.44	0.16	0.408	0.2	0.268
GGA	**1.296**	**1.98**	**1.976**	**1.648**	**2.02**	**1.536**	**1.48**	**2.35**	**1.624**	**1.696**	**2**	**2.216**
GGC	**1.448**	**1.22**	0.864	**1.256**	**1.13**	**1.248**	**1.488**	**1.24**	**1.52**	**1.536**	**1.3**	**1.232**
GGG	0.648	0.24	0.576	0.456	0.49	0.288	0.312	0.13	0.272	0.456	0.3	0.128
GGT	0.608	0.55	0.584	0.616	0.38	0.968	0.728	0.28	0.584	0.32	0.3	0.416
TGG	0.168	0.18	0.168	0.08	0.16	0	0.132	0.26	0.096	0.196	0.2	0.132
TGA	**1.832**	**1.82**	**1.832**	**1.92**	**1.84**	**2**	**1.868**	**1.74**	**1.904**	**1.804**	**1.8**	**1.868**

Bold indicates more frequently used codons

In ATP6, out of 60 codons, 28 codons were more frequently used in pisces, out of which 13 codons end with A, 2 codons end with T and 13 codons end with C. In aves, out of 28 more frequently used, 13 codons end with A, and 15 codons end with C. In mammals, 27 codons were more frequently used, out of which 14 codons end with A, 5 codons end with T, and 8 codons end with C.

In ATP8, out of 60 codons, 23 codons were more frequently used in pisces, out of which 9 codons end with A, 5 codons end with T, 1 codon ends with G and 8 codons end with C. In aves, out of 17 more frequently used codons 7 codons end with A, 2 codons end with T, 1 codon ends with G and 7 codons end with C. In mammals, 17 codons were more frequently used, out of which 10 codons end with A, 6 codons end with T, and 1 codon ends with C.

Table 4.5.4 RSCU values of codons for different mitochondrial genes in ATP gene among pisces, aves and mammals

	MT-ATP6			MT-ATP8		
Codon	Pisces	Aves	Mammals	Pisces	Aves	Mammals
TCA	**1.14**	**1.896**	**2.328**	**1.392**	**2.328**	**2.424**
TCT	0.78	0.456	**1.224**	**1.74**	**1.236**	0.908
TCC	0.588	**1.86**	0.72	**2.496**	**2.136**	**1.032**
TCG	0.12	0.072	0	0.024	0.012	0.768
AGC	**3.132**	**1.728**	**1.032**	0.42	0.312	0.336
AGT	0.24	0	0.696	0.024	0.012	0.348
TTC	**1.08**	**1.8**	0.952	**1.168**	**1.72**	0.652
TTT	0.92	0.204	**1.048**	**1.432**	0.28	**1.348**
TTA	0.984	0.54	**1.332**	**1.404**	0.324	0.876
TTG	0.072	0.024	0.048	0.204	0	0.528
CTA	**2.004**	**2.784**	**2.976**	**2.22**	**2.688**	**2.532**
CTC	**1.38**	**1.596**	0.576	0.756	**1.44**	0.552
CTG	0.456	0.36	0.216	0.204	0.312	0.156
CTT	**1.116**	0.708	0.828	**1.236**	**1.272**	**1.356**
TAC	**1.448**	**1.568**	**1.136**	0.8	0	0.548
TAT	0.552	0.432	0.872	0.008	0	**1.464**
TGT	0	0.004	0.004	0.4	0	0.404
TGC	0	0.4	0.4	0.004	0	0.404
CCA	**1.696**	**1.944**	**1.608**	**1.616**	**1.72**	**1.128**
CCC	**1.584**	**1.568**	**1.112**	**1.352**	**1.776**	0.76
CCG	0.144	0.096	0.112	0.184	0.08	0.272
CCT	0.576	0.384	**1.152**	0.856	0.432	**1.84**
CAT	0.6	0.2	1.068	0.412	0	0.672
CAC	**1.4**	**1.8**	0.936	**1.468**	0	0.932
CAA	**1.832**	**1.848**	**1.84**	**1.404**	**1.6**	**1.9**
CAG	0.168	0.16	0.164	0.604	0.412	0.104
CGA	**2.264**	**2.4**	**2.72**	0.8	0	0.8
CGC	**1.2**	**1.52**	0.68	0.008	0	0.008
CGG	0.136	0.32	0.48	0.008	0	0.008
CGT	0.408	0.16	0	0.008	0	0.008
ATG	0.688	0.424	0.416	**1.6**	**1.132**	0.992
ATA	**1.312**	**1.58**	**1.584**	0.412	0.868	**1.012**
ATC	0.912	**1.636**	0.732	0.932	**1.524**	0.7
ATT	**1.088**	0.364	**1.268**	**1.072**	0.476	**1.3**
ACA	**1.656**	**1.752**	**1.728**	**1.784**	0.992	**2.048**
ACC	**1.608**	**1.816**	**1.392**	**1.352**	**2.096**	0.912
ACG	0.104	0.192	0.12	0	0	0.6
ACT	0.64	0.232	0.784	0.856	0.904	0.448
AAC	**1.244**	**1.832**	**1.1**	**1.332**	**1.636**	0.804
AAT	0.756	0.168	0.9	0.672	0.368	0.804

Continued

	MT-ATP6			MT-ATP8		
Codon	Pisces	Aves	Mammals	Pisces	Aves	Mammals
GAC	**1.208**	**1.404**	**1.472**	0.004	0	0.412
GAT	0.812	0.612	0.544	0.4	0	**1.208**
GTA	**1.32**	**1.424**	**2**	0.816	0.808	**1.2**
GTC	**1.152**	**2.192**	0.968	**1.6**	0.016	0.408
GTG	0.544	0.16	0.328	0.032	0.016	0.016
GTT	0.968	0.232	0.712	**1.208**	0.808	0.016
GCA	**1.096**	**1.136**	**1.368**	0.28	**1.744**	0.28
GCC	**2.096**	**2.472**	**1.56**	**3.2**	0.68	0.944
GCG	0	0.104	0.08	0.288	0.024	0.408
GCT	0.808	0.312	0.984	0.288	0.816	0.808
AAA	**1.604**	**1.768**	**1.62**	**1.468**	**1.868**	**1.292**
AAG	0.416	0.232	0.38	0.332	0.132	0.584
GAA	**1.9**	**1.9**	**1.8**	**1.404**	0	**1.2**
GAG	0.1	0.1	0.204	0.608	0	0.212
GGA	**1.56**	**1.616**	**1.536**	0	0	0.208
GGC	**1.76**	**1.72**	**1.24**	0	0	0.008
GGG	0.32	0.104	0.64	0	0	0.008
GGT	0.36	0.56	0.568	0	0	0.8
TGG	0.16	0.3	0	0	0	0
TGA	**1.84**	**1.7**	**2**	**2**	**2**	**2**

Bold indicates more frequently used codons

These results suggest that compositional constraint under mutation pressure played an important role in shaping the codon usage patterns in mitochondrial genes among pisces, aves and mammals.