

Abstract

Codon bias is the unequal use of synonymous codons, some codons are more preferred than others. Synonymous codons encoding amino acids range from two to six, usually differing in the third position. However, the patterns of synonymous codon usage differ among the genes of an organism and among different organisms. Some codons that encode an amino acid are more frequently used than others by the genome.

In the present study we have attempted to analyze the pattern of codon usage bias of mitochondrial protein-coding genes across a diverse group of vertebrates *i.e.* pisces, aves and mammals to identify the commonality and the difference, if any, in base composition and other related parameters of codon usage pattern in mitochondria. The mitochondrial genes are a subset of the frequently expressed genes in eukaryotes, and its genome is usually ideal as the molecular marker for species identification, systematic phylogeny, and evolutionary studies. The choice of mitochondrial DNA in all these studies is primarily due to its small size, easy amplification and conserved gene content, lack of recombination, maternal inheritance pattern and high evolutionary rate. The metazoan mitochondrial DNA is circular, 16 kb in size, covalently closed and consists of 37 genes. The majority of the mitochondrial proteins are encoded by nuclear genes while only 2 rRNA, 22 tRNAs and 13 proteins involved in the respiratory chain are encoded by mitochondrial genomes. Further the genetic code of mitochondria differs from that of standard genetic code. The standard genetic code consists of 64 codons, wherein 61 sense codons encode 20 standard amino acids and three codons namely TAA, TAG and TGA act as termination signal. In mitochondrial genetic code there are four termination codons such as TAA, TAG, AGA and AGG. Patterns of codon usage in nuclear genomes are extensively studied whereas studies on mitochondrial genomes or genes are very scarce.

The oxidative phosphorylation is one of the most important biochemical processes operating in mitochondria in which aerobic eukaryotic cell uses oxygen to synthesize ATP. The 13 protein-coding genes of mitochondria are universal and encode for the protein subunits of the different complexes of oxidative phosphorylation. The oxidative phosphorylation is the multienzymatic system which creates the proton gradient required for ATP synthesis (or heat generation). The seven mitochondrially encoded proteins (ND1, ND2, ND3, ND4, ND4L, ND5, ND6) form the complex I in which ND1 and ND2 play an essential structural role

between the membrane-embedded and peripheral arms of the complex whereas the role of ND2, ND4, and ND5 is to transfer electron.

Analysis of the pattern of codon usage in mitochondrial protein-coding genes is of special interest to understand how the energy consumption of pisces, aves and mammals, influences the codon usage against quick environmental alterations during the course of evolution. The pisces, aves and mammals survive in three different habitats and so their means of respiration and energy demand are also unusual. Therefore, this study helps in understanding the factors influencing the process of gene evolution.

Analysis of codon usage is a useful technique to understand the genetic and evolutionary relationship of different species belonging to diverse habitats. Moreover, mitochondrial genes are very significant and suitable tools for these kinds of studies. In the current study, we investigated the codon usage pattern in thirteen mitochondrial protein-coding genes among pisces, aves and mammals thriving in different habitats to understand the pattern of codon usage. Moreover, this study gives insight in to the factors influencing the codon usage pattern among the species under study.

In the present study, analysis of codon usage pattern, expression level and influencing factors for mitochondrial protein-coding genes from 15 different species of pisces, aves and mammals were used for codon usage bias (CUB) analysis. We retrieved sequences from National Centre for Biotechnology Information (NCBI), Bethesda, USA and then CUB analysis was done using several parameters such as effective number of codons (ENC), codon adaptation index (CAI), relative synonymous codon usage (RSCU). A perl script was used for calculating nucleotide contents and different CUB parameters. The statistical analysis was done using SPSS software and correspondence analysis was done. Heat map was generated using XLSTAT.

In our current study, hierarchial clustering using RSCU values of the codons in 13 protein-coding genes among pisces, aves and mammals suggest that pattern of codon usage differs among 13 protein-coding genes and among pisces, aves and mammals. The ENC value > 50 in all 13 genes among pisces, aves and mammals indicates low codon usage bias of genes. The ENC value of all genes was lower in mammals (except ND6) followed by pisces and aves. The expression level of mitochondrial ND1, ND2, ND4, ND5, COI, COII, CYB and ATP6 genes was very high. The expression level of ND6 and COIII genes was moderately high. The expression level of ND3, ND4l and ATP8 genes was low. Correlation analysis

between ENC and CAI suggests that no significant relationship exists between codon usage bias and expression level. We found GC content was high in aves followed by pisces and mammals for all 13 genes. The GC content is lower than AT content *i.e.* 13 protein coding genes are AT rich. Correlation among compositional constraints and correlation between ENC with various GC contents suggest that both mutation pressure and natural selection affect the codon usage pattern in 13 protein-coding genes of mitochondria in pisces, aves and mammals. Neutrality plot suggests that natural selection played a major role while mutation pressure played a minor role in shaping the codon usage pattern in mitochondrial protein-coding genes except the ND6 gene in pisces where mutation pressure played a dominant role while natural selection played a minor role in codon usage pattern. Correspondence analysis also suggests that the patterns of codon usage are different among genes and vary among species. The study also revealed that both natural selection and mutation pressure affect the codon usage pattern in 13 protein-coding genes of mitochondria in pisces, aves and mammals.