biology

MDPI

Article

Analysis of Codon Usage of Speech Gene *FoxP2* among Animals

Tarikul Huda Mazumder [1], Ali M. Alqahtani [2], Taha Alqahtani [2], Talha Bin Emran [3], Afaf A. Aldahish [2] and Arif Uddin [4,*]

1 EduCare Academy, Silchar 788006, Assam, India; tariqulhmazumder@gmail.com
2 Department of Pharmacology, College of Pharmacy, King Khalid University, Abha 62529, Saudi Arabia; amsfr@kku.edu.sa (A.M.A.); ttaha@kku.edu.sa (T.A.); adahesh@kku.edu.sa (A.A.A.)
3 Department of Pharmacy, BGC Trust University Bangladesh, Chittagong 4381, Bangladesh; talhabmb@bgctub.ac.bd
4 Department of Zoology, Moinul Hoque Choudhury Memorial College, Hailakandi 788150, Assam, India
* Correspondence: arif.uddin29@gmail.com

Simple Summary: We evaluated codon usage bias in the *FoxP2* gene in fishes, birds, reptiles, and mammals. Fishes use C or G—ending codons, while birds, reptiles, and mammals employ T or A—ending codons. Apart from the nucleotide composition, natural selection and mutation pressure might influence the CUB. The ENC observed/ENC expected ratio demonstrated that mutation pressure influences *FoxP2* codon usage patterns. Natural selection may have had a key influence in shaping the CUB, although mutation pressure may have played a minor role. *FoxP2* gene codon usage is affected by the base composition under mutation bias.

Abstract: The protein-coding gene *FoxP2* (fork head box protein P2) plays a major role in communication and evolutionary changes. The present study carried out a comprehensive codon usage bias analysis in the *FoxP2* gene among a diverse group of animals including fishes, birds, reptiles, and mammals. We observed that in the genome of fishes for the *FoxP2* gene, codons ending with C or G were most frequently used, while in birds, reptiles, and mammals, codons ending with T or A were most frequently used. A higher ENC value was observed for the *FoxP2* gene indicating a lower CUB. Parity role two-bias plots suggested that apart from mutation pressure, other factors such as natural selection might have influenced the CUB. The frequency distribution of the ENC observed and ENC expected ratio revealed that mutation pressure plays a key role in the patterns of codon usage of *FoxP2*. Besides, correspondence analysis exposed the composition of the nucleobase under mutation bias affects the codon usage of the *FoxP2* gene. However, neutrality plots revealed the major role of natural selection over mutation pressure in the CUB of *FoxP2*. In addition, the codon usage patterns for FoxP2 among the selected genomes suggested that nature has favored nearly all the synonymous codons for encoding the corresponding amino acid. The uniform usage of 12 synonymous codons for FoxP2 was observed among the species of birds. The amino acid usage frequency for FoxP2 revealed that the amino acids Leucine, Glutamine, and Serine were predominant over other amino acids among all the species of fishes, birds, reptiles, and mammals.

Keywords: *FoxP2* gene; effective number of codons; relative synonymous codon usage; natural selection

Citation: Mazumder, T.H.; Alqahtani, A.M.; Alqahtani, T.; Emran, T.B.; A. Aldahish, A.; Uddin, A. Analysis of Codon Usage of Speech Gene *FoxP2* among Animals. *Biology* **2021**, *10*, 1078. https://doi.org/10.3390/biology10111078

Academic Editor: David G. Riley

Received: 6 September 2021
Accepted: 16 October 2021
Published: 21 October 2021

1. Introduction

The standard genetic code makes use of 64 codons to characterize the 20 standard amino acids, including 3 stop codons, during the translation of a protein. The redundancy of the genetic code implies that one amino acid may be encoded by more than one codon (except methionine and tryptophan) leading to the occurrence of synonymous codon usage bias (CUB) in the genome of an organism, which varies among diverse groups [1].

Mitochondrion 62 (2022) 111–121

Contents lists available at ScienceDirect

Mitochondrion

journal homepage: www.elsevier.com/locate/mito

Compositional features and pattern of codon usage for mitochondrial CO genes among reptiles

Supriyo Chakraborty [a,*], Priyanka Basumatary [a], Durbba Nath [a], Sunanda Paul [a], Arif Uddin [b]

[a] Department of Biotechnology, Assam University, Silchar 788011, Assam, India
[b] Department of Zoology, Moinul Hoque Choudhury Memorial Science College, Algapur, Hailakandi786150, Assam, India

ARTICLE INFO

Keywords:
Reptiles
Codon usage bias
MT-CO genes
Nucleotide composition
Evolutionary study

ABSTRACT

The phenomenon of non-random occurrence of synonymous nucleotide triplets (codons) in the coding sequences of genes is the codon usage bias (CUB). In this study, we used bioinformatic tool kit to analyze the compositional pattern and CUB of mitogenes namely COI, COII and COIII across different orders of reptiles. Estimation of overall base composition in the protein-coding sequences of COI, COII and COIII genes of the reptilian orders revealed an uneven usage of nucleotides. The overall count of A nucleotide was found to be the highest while the overall count of G nucleotide was the least. The CO genes across the three reptilian orders were prominently AT biased. Comparison of the GC proportion at each codon position displayed that GC1 percentage ranked the highest in all the three CO genes of the reptilian orders. SCUO values indicated weaker CUB, while considerable variation of SCUO values existed in the three CO genes across the studied reptiles. Relative synonymous codon usage (RSCU) values indicated that mostly the A ending codons were preferred. Based on the parameters namely neutrality plot, mutational responsive index and translational selection, we could conclude that natural selection was the major evolutionary force in COI, COII and COIII genes in the studied reptilian orders. However, correspondence analysis, parity plot and correlation studies indicated the existence of mutation pressure as well on the CO genes.

1. Introduction

Codon usage bias (CUB) can be illustrated as the unequal occurrence of synonymous codons in the coding sequences of genes (Sun et al., 2009). Multiple codons that encode the same amino acid were reported across species and in the genes of similar species. The frequently occurring codons are easily capable of pairing with the anticodons of tRNA pool, thereby refraining ribosomes from the recruitment of incorrect amino acids in the growing polypeptide chain (Sharp and Li, 1986a; Wei et al., 2014). Therefore, it is presumed that owing to the action of natural selection and mutation, equilibrium in the CUB phenomenon is maintained (Wei et al., 2014).

The prime factors influencing CUB are genetic drift, mutation pressure and natural selection (Behura and Severson, 2013) and the three forces act in equilibrium to opt for translationally adept codons (Akashi, 1995; McVean and Vieira, 2001). Pronounced codon bias can be observed in the genes with high expressivity, and this association is the consequence of selection pressure that ensures translational efficacy. Other factors that determine CUB are favored gene conversion (Marais et al., 2003), RNA and protein structure (Carlini et al., 2001; Oresic et al., 2003), GC content, synonymous substitution rate, origin of DNA replication (Wang et al., 2016), gene length etc. (Duret and Mouchiroud, 1999). Another influencer of CUB is the dinucleotide frequency in coding sequences. Lesser prevalence of CpG nucleotides pair was observed in vertebrates, and it was explained that the methylation of most cytosine bases is processed by methyl transferase enzyme followed by deamination of unpaired methylated cytosine to thymine (Shackelton et al., 2006). Codon usage profile of genes or species, when understood in detail can categorize a gene and help interpret its evolutionary trends. Further, using the concept of codon usage pattern it is possible to construct the efficiently expressed genes, cloning vectors (Wang et al., 2018) as well as vaccines that can elicit enhanced immune response (Paul et al., 2014).

GC biased mutation is believed to be the vital determinant of inconsistency across genomes. The segregation distortion due to GC preference over AT alleles occurs during recombination generating GC abundance in recombination hotspots (Figuet et al., 2015; Meunier and Duret, 2004). Crocodilians, turtles, birds and mammals share a common

* Corresponding author.
E-mail addresses: supriyoch_2008@rediffmail.com (S. Chakraborty), arif.uddin29@gmail.com (A. Uddin).

https://doi.org/10.1016/j.mito.2021.11.004
Received 7 April 2021; Received in revised form 2 November 2021; Accepted 10 November 2021
Available online 15 November 2021

04/01/23
Principal
M. H. C. M. Sc. College
Algapur, Hailakandi

Archives of Virology
https://doi.org/10.1007/s00705-020-04890-2

ORIGINAL ARTICLE

Composition, codon usage pattern, protein properties, and influencing factors in the genomes of members of the family *Anelloviridae*

Bornali Deb[1] · Arif Uddin[2] · Supriyo Chakraborty[1]

Received: 1 April 2020 / Accepted: 2 October 2020

Abstract
The present study was carried out on 62 genome sequences of members of the family *Anelloviridae*, as there have been no reports of genome analysis of these DNA viruses using a bioinformatics approach. The genes were found to be rich in AC content with low codon usage bias (CUB). Relative synonymous codon usage (RSCU) values identified the preferred codons for each amino acid in the family. The codon AGA was overrepresented, while the codons TCG, TTG, CGG, CGT, ACG, GCG and GAT were underrepresented in all of the genomes. A significant correlation was found between the effective number of codons (ENC) and base constraints, indicating that compositional properties might have influenced the CUB. A highly significant correlation was observed between the overall base content and the base content at the third codon position, indicating that mutations might have affected the CUB. A highly significant positive correlation was observed between GC12 and GC3 ($r = 0.904$, $p < 0.01$), which indicated that directional mutation pressure influenced all three codon positions. A neutrality plot revealed that the contribution of mutation and natural selection in determining the CUB was 58.6% and 41.4%, respectively.

Introduction

Amino acids play a crucial role in cellular metabolic activities of an organism. Amino acids are joined step by step to form proteins. In the standard genetic code, a set of 59 codons encode 18 standard amino acids. Here, methionine and tryptophan are the only two amino acids that are coded with a single codon while all other amino acids are encoded by more than one codons, thus making some codons seemingly redundant in transcript. A bias in synonymous substitution of codons resulting in preferential usage of a specific codon within a codon family is termed codon usage bias (CUB), and it is different for genes, genomes, transcriptomes, and species [6, 33, 51]. CUB is frequently observed in highly expressed genes, while genes that are expressed at a low level usually have less CUB [25]. The pattern of synonymous codon usage allows the identification of relevant isoacceptor tRNAs for efficient translation of a particular gene. Thus, genomes with highly expressed genes will have more bias in codon preference, leading to the formation of proteins with lower susceptibility to misfolding [2]. Studies have shown that variation in synonymous substitutions occurs between and within genes [23, 44]. Various researchers have pointed out that the study of CUB provides important information about the evolution of related organisms [55, 58]. Because viruses are replicated and their genes are translated in living host cells, investigation of codon usage patterns of viral genomes can potentially provide information about the interaction and co-evolution of viruses with their hosts [54].

Several theories have been propounded for the origin of CUB, two of which are the selection-mutation-drift theory and the neutral theory. In the selection-mutation-drift theory, the major determinants of CUB are mutational pressure,

Handling Editor: Sheela Ramamoorthy.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00705-020-04890-2) contains supplementary material, which is available to authorized users.

✉ Supriyo Chakraborty
supriyoch_2008@rediffmail.com

[1] Department of Biotechnology, Assam University, Silchar, Assam 788150, India

[2] Department of Zoology, Moinul Hoque Choudhury Memorial Science College, Algapur, Hailakandi, Assam 788150, India

Published online: 03 January 2021

Genomics 113 (2021) 2826–2838

Contents lists available at ScienceDirect

Genomics

journal homepage: www.elsevier.com/locate/ygeno

Free energy of mRNA positively correlates with GC content in chloroplast transcriptomes of edible legumes

Supriyo Chakraborty [a,*], Yengkhom Sophiarani [a], Arif Uddin [b]

[a] Department of Biotechnology, Assam University, Silchar 788011, Assam, India
[b] Department of Zoology, Moinul Hoque Choudhury Memorial Science College, Algapur, Hailakandi 788150, India

ARTICLE INFO

Keywords:
Chloroplast transcriptomes
Free energy
Codon usage bias
Mutation pressure
Natural selection

ABSTRACT

In the present study, the results of nucleotide composition analysis showed that the legume chloroplast (cp) transcriptomes were AT rich. From the neutrality plot, we observed that natural selection might have played a major role, while mutation pressure played a minor role in the CUB of cp transcriptomes. Highly significant ($p < 0.05$) negative correlation was found between mRNA free energy (mFE) and scaled chi-square for entire mRNA in *Cicer arietinum* and *Lens culinaris* suggesting that the release of higher energy by entire mRNA molecule might be associated with higher degree of codon usage bias in these two crop plants. Further, highly significant ($p < 0.01$, $p < 0.05$) positive correlation of mFE for entire mRNA was found with GC3 and that of mFE for 39 bases with GC, GC1, GC2 and GC3 contents among all the legumes. This indicated that higher GC content might induce the release of more free energy by cp transcriptomes.

1. Introduction

In standard genetic code, there are 64 codons of which 61 codons specify 20 standard amino acids while the remaining three codons (TAA, TAG, and TGA) are used as stop signals in protein synthesis. Each codon consists of three nucleotides and the sequence of codons in an mRNA molecule determines the sequence of amino acids in a growing polypeptide chain that terminates when a stop codon is encountered by the ribosome. The code determines the order in which amino acids are added to a polypeptide chain during protein synthesis. On the other hand, other genomic regions determine when and where these proteins are produced according to various gene regulatory codes and network. Codon usage bias (CUB) is a unique property of a genome with non-random variation and unequal occurrence of synonymous codons of amino acids in a transcript [22]. It shows species-specific variation [17]. Many researchers carried out experiments to explain the origin of codon usage bias. The findings of those studies revealed that CUB might be a product of mutation pressure or natural selection or both [28,34,42,44], where it also explained the accurate efficiency of translation, as well as mutation-selection-drift balance [3]. Many reports suggested that CUB is also affected by base compositional bias (GC% and GC skew), mutation pressure, natural selection, gene length, expression level, DNA replication, RNA stability, hydrophobicity and hydrophilicity of the protein [9]. Further it was mentioned that the variation in CUB of genes among different species is due to compositional constraints in the presence of mutation pressure and natural selection [45].

Several studies earlier reported on non-random usage of synonymous codon in many organisms. It was suggested that a positive correlation exists between the frequencies of abundant tRNAs and the frequently used codons and also between gene expression level and codon bias due to natural selection for efficient translation. It was reported that the codon usage in genes of some organisms like human and yeast, was related to specific locations of genes in the genome as a result of the mosaic patterns of GC content.

Legumes are the food plants that belong to the family Fabaceae (earlier known as Leguminosae). Especially legumes are grown for their protein-rich edible seeds for human as well as livestock consumption and also as soil-enriching green manure crops [4]. According to the recent findings it was reported that among various species of leguminosae, pea (*Pisum sativum* L.) is highly consumed in Asian countries, common bean (*Phaseolus vulgaris* L.) in Latin American and African countries, chickpea (*Cicer arietinum* L.) in India and lentil (*Lens culinaris* Med.) in countries of the Middle East [2]. Legumes played a pivotal role in Afro-Asian diets for protein and calorific values. Generally, legumes have significant amount of vitamins and minerals with high energetic value and also they are the sources of complex carbohydrates, protein

* Corresponding author.
E-mail address: supriyoch_2008@rediffmail.com (S. Chakraborty).

https://doi.org/10.1016/j.ygeno.2021.06.026
Received 31 December 2018; Received in revised form 1 January 2021; Accepted 16 June 2021
Available online 18 June 2021

Virology 561 (2021) 69–79

Contents lists available at ScienceDirect

Virology

journal homepage: www.elsevier.com/locate/virology

Analysis of codon usage of Horseshoe Bat Hepatitis B virus and its host

Bornali Deb [a], Arif Uddin [b], Supriyo Chakraborty [a,*]

[a] Department of Biotechnology, Assam University, Silchar, 788150, Assam, India
[b] Department of Zoology, Moinul Hoque Choudhury Memorial Science College, Algapur, Hailakandi, 788150, Assam, India

ARTICLE INFO

Keywords:
Codon usage bias
Horseshoe bat hepatitis B virus
Rhinolophus sinicus
Mutation pressure
Natural selection

ABSTRACT

In the present analysis, codon usage strategies and base distribution of Horseshoe bat hepatitis B virus (HBHBV) were analyzed and compared with its host *Rhinolophus sinicus*, as no work was yet reported. The magnitude of synonymous codon usage bias (CUB) in the virus and its host was low with higher proportion of the base C. Notably, 21 more frequently used codons, 19 less frequently used codons and 3 underrepresented codons (TCG, ACG and GCG) were found to be similar in both virus and its host coding sequences. Neutrality plot analysis reported greater role of natural selection in HBHBV (67.04%) and *R. sinicus* (76.90%) over mutation pressure. Base skewness and protein properties also influenced the CUB of genes. Further, codon usage analysis depicted, HBHBV and *R. sinicus* had many similarities in codon usage patterns that might reflect viral adaptation to its host.

1. Introduction

The standard genetic code forms the basis for translating genetic information encoded in DNA/RNA to produce protein through the usage of codons. Apart from three termination codons, a set of 61 sense codons codify 20 standard amino acids in polypeptides. Except two amino acids (tryptophan and methionine), 18 amino acids are synonymously encoded by more than one codon. Synonymous substitutions of codons and their usage in the coding sequences follow a certain biased pattern in each gene/genome/species (Bennetzen and Hall, 1982; Ikemura, 1981a). The event of biased usage of certain codons preferentially over other codons in coding sequence (cds) is defined as codon usage bias (CUB). Analysis of CUB is significant in the identification of genetic and evolutionary features of an organism/species. Previously CUB analysis in various organisms varying from single celled to multi celled was reported (Gu et al., 2004a; Liu et al., 2011).

According to the selection-mutation-drift theory, the major determinants framing CUB are natural selection, mutational pressure and genetic drift (Bulmer, 1991; Jenkins and Holmes, 2003). In the case of populations under the influence of intense genetic drift and less selection pressure, the patterns of codon usage tend to be highly affected with mutational bias (Sharp and Li, 1987; Palidwor et al., 2010). Other factors shaping CUB include GC constraints (McInerney, 1998), nucleotide skewness (Chakraborty et al., 2019), gene length (Eyre-Walker, 1996), mRNA structure, mRNA stability (Chen et al., 2004; Akashi, 1997; Mazumder and Chakraborty, 2015), protein properties (Deb et al., 2018) and geographical location of organism (Zhang et al., 2011). The role of base constraints along with evolutionary forces was considered to frame CUB across various species (Sharp and Li, 1986a; Sharp and Li, 1986b; Sharp et al., 1993). Genes with higher expression level usually possess higher CUB than those with lower expression level (Epstein et al., 2000; Gu et al., 2004b). Analysis of synonymous substitutions of codons assists in the identification of subsequent isoacceptor transfer RNAs (tRNAs) that subsequently lead the process of translation (Sharp et al., 1993; Sharp et al., 2005; Stenico et al., 1994). The codon usage associates well with the abundance of tRNA molecules in cell and the extent of gene expression, *i.e.*, optimization of codons occurs depending on the accessible tRNA pool that, indeed, regulates the translational process having a major impact on CUB and translational dynamics of transcripts (Ikemura, 1981a; Sharp and Li, 1986b; Post et al., 1979; Nakamura et al., 1980; Ikemura, 1981b; Gouy and Gautier, 1982; Osawa and Jukes, 1988). The codon corresponding to a specific tRNA species is known as optimized codon and it is linked to the increased rate of translation (Bulmer, 1991; Xia, 1998; Chevance et al., 2014; Presnyak et al., 2015), mRNA stability (Presnyak et al., 2015; Radhakrishnan et al., 2016), polyadenylation of transcripts (Zhou et al., 2018) and premature suppressed cleavage (Zhou et al., 2018). Biased codons help in reducing proof-reading expenses by decreasing the time and energy required to remove the non-cognate tRNAs (Bulmer, 1991). Less preferred codons increase the proof-reading task that might lead to a net reduction in protein content (Uddin et al., 2019a). Moreover, non-optimal codons were reported to be associated with initiation of translation (Tuller

* Corresponding author.
E-mail address: supriyoch_2003@rediffmail.com (S. Chakraborty).

https://doi.org/10.1016/j.virol.2021.05.003
Received 11 September 2020; Received in revised form 7 May 2021; Accepted 19 May 2021
Available online 10 June 2021

04/01/23
Principal
M. H. C. M. Sc. College
Algapur, Hailakandi

Computational Biology and Chemistry 92 (2021) 107433

Contents lists available at ScienceDirect

Computational Biology and Chemistry

journal homepage: www.elsevier.com/locate/cbac

Silencing lung cancer genes using miRNAs identified by 7mer-seed matching

Supriyo Chakraborty [a,*], Durbba Nath [a], Parvin A Barbhuiya [a], Yashmin Choudhury [a], Arif Uddin [b]

[a] Department of Biotechnology, Assam University, Silchar, 788011, Assam, India
[b] Department of Zoology, Moinul Hoque Choudhury Memorial Science College, Algapur, Hailakandi, 788150, Assam, India

ARTICLE INFO

Keywords:
Small cell lung cancer
Non small cell lung cancer
microRNA
Overexpressed genes

ABSTRACT

Lung cancer (LC) is the main cause of cancer-associated deaths in both men and women globally with a very high mortality rate. The microRNAs (miRNAs) are a class of noncoding RNAs consisting of 18–25 nucleotides. They inhibit translation of protein through binding to complementary target mRNAs. The non-coding miRNAs are recognised as potent biomarkers for detection, development and treatment of malignancy. In this study, we screened a set of 12 genes over expressed in small cell lung cancer, non small cell lung cancer and the genes involved in both categories and their binding sites for human miRNAs as no work was reported yet. Screening of human miRNAs revealed that a few genes showed numerous miRNA binding sites. Free energy values of mRNA sequences revealed that they might acquire compact folded structure causing complexity for miRNAs to interact. GC content in the target site was relatively higher than that of their flanks. It was observed through analysis of cosine similarity metric and compAI parameters that the genes related to lung cancer were encoded with non optimal codons and thus might be translationally less efficient for producing polypeptides. Gene ontology analysis was carried out to understand the diverse functions of these 12 genes.

1. Introduction

RNA is considered as an information mediator between DNA and encoded protein (Wang et al., 2009a). About 2% of the human RNAs are utilized in protein translation (Crick et al., 1961). The RNA which does not play any role in protein coding is referred to as non-coding RNA (ncRNA).Various researches shed light on the involvement of ncRNA in many cellular processes (Mattick and Makunin, 2006). The microRNAs (miRNAs) are a major class of ncRNAs. The miRNAs are endogenous short, single-stranded, non-coding RNA molecules of 19–23 nucleotides that are largely found in prokaryotic and eukaryotic cells (Jones-Rhoades et al., 2006). They act as post-transcriptional gene regulators and play a pivotal role in gene expression by silencing the genes either by degrading mRNA or blocking its translation (Panda et al., 2014). The miRNAs inhibit the mRNA synthesis by forming complexes outside the nucleus. The miRNAs bind to the 3′ UTR sequences (Kamaraj an et al., 2012) of their target genes and initiate either mRNA synthesis or cause translational repression of mRNAs. Drosha enzyme-mediated cleavage (Kamaraj an et al., 2012) and the formation of hair pin loop generates 80–100 nucleotides long stem loop precursor (pre-miRNA). It is exported to the cytoplasm by nucleocytoplasmic shuttler protein Exportin-5 and acted upon by Dicer enzymes (DCL) and DiGeorge Syndrome Critical Region 8 (DGCR8) to form miRNA duplex. The miRNA released by the helicase activity is loaded into the RNA Induced Silencing Complex (RISC), and guides the assembly to the target mRNA (Panda et al., 2014).

Many functions are said to be regulated by miRNAs which include differentiation, development, growth, apoptosis, cell cycle regulation, metabolism and stress reaction (Carleton et al., 2007; Petrocca et al., 2008; Ambros, 2003). Furthermore, it was shown earlier that miRNAs might be involved in the development of human diseases and their progression (Esteller, 2011). Several cancers, for instance, breast cancer and hepatocellular carcinoma were said to be associated with a few key miRNA regulons (Blenkiron et al., 2007; Zhang et al., 2018). Thus, many advantages are anticipated in recognizing the potential miRNAs in relation to diseases which can serve in acquiring the information about pathogenesis, diagnostic and therapeutic features of human ailments. Numerous studies were done on miRNA and lung cancer association (Sun et al., 2017; Fan et al., 2017; Wang et al., 2017; Hu et al., 2017).

* Corresponding author.
E-mail address: suprivoch_2008@rediffmail.com (S. Chakraborty).

https://doi.org/10.1016/j.compbiolchem.2021.107433
Received 20 December 2020; Received in revised form 19 March 2021; Accepted 3 April 2021
Available online 12 April 2021

04/05/23
Principal
M. H. C. M. Sc. College
Algapur, Hailakandi

Virus Research 292 (2021) 198248

Contents lists available at ScienceDirect

Virus Research

journal homepage: www.elsevier.com/locate/virusres

Genome-wide analysis of codon usage pattern in herpesviruses and its relation to evolution

Bornali Deb [a], Arif Uddin [b], Supriyo Chakraborty [a,*]

[a] Department of Biotechnology, Assam University, Silchar, 788011, Assam, India
[b] Department of Zoology, Moinul Haque Choudhury Memorial Science College, Algapur, Hailakandi, 788150, Assam, India

ARTICLE INFO

Keywords:
Codon
Usage bias
Herpesviruses
Mutation pressure
Natural selection

ABSTRACT

The preferential use of a specific codon, out of a group of synonymous codons encoding the same amino acid, in a gene transcript results from the bias in codon choice. Various evolutionary forces namely mutation pressure and natural selection influence the pattern of codon usage *i.e.* distinct for each gene/genome. We investigated the pattern of codon usage of eight human herpesvirus genomes and compared them with two other herpesvirus genomes namely murine herpesvirus 68 and bovine herpesvirus type 1.1 to elucidate its compositional features, pattern of codon usage across the genomes and report the differences of codon usage pattern of human herpesviruses from that of other two other viruses. We also identified the similarity of the codon usage of human herpesviruses with its host (human). The genes were found to be CG rich in HHV2, HHV3, HHV4, HHV6, HHV7 and BH genomes while TA rich in HHV1, HHV5, HHV8 and MH genomes. The codon usage bias (CUB) of genes was low. A highly significant correlation was found among compositional contents depicting the role of mutational pressure along with natural selection in framing CUB. Several more frequently used codons as well as less frequently used codons were identified to be similar between each human virus and its host (human), while murine herpesvirus 68 and bovine herpesvirus type 1.1 genomes did not possess similar adaptation strategy as human herpesviruses to human (host), thus we could conclude that viral CUB might have been shaped as per their host's nature for better surveillance. Neutrality plot revealed mutational pressure mostly influenced the CUB of HHV1, HHV8 and MH viruses, while natural selection had a major impact in the CUB of HHV2, HHV3, HHV4, HHV5, HHV6, HHV7 and BH genomes.

1. Introduction

From the very beginning of life on earth, a set of evolutionary processes has guided the genetic framework of each creature. Natural selection and mutation pressure are the two major attributes of the evolutionary pathway structured through biased codon usage in an organism/genome/species. The genetic code depicts that multiple codons can encode an amino acid, except tryptophan and methionine. Codon usage bias (CUB) is thus a universal phenomenon of synonymous substitution of codons in gene transcripts arising as a consequence of evolution, together with some other facets namely GC composition (Bibb et al., 1984), base skewness (Chakraborty et al., 2019), gene expression (Uddin and Chakraborty, 2019), gene length (Eyre-Walker, 1996) and properties of protein (Deb et al., 2018). CUB is said to effect mRNA transcription, its stability, translational and post- translational properties (Chamary et al., 2006; Rodriguez et al., 2012; Trotta, 2013). A deep understanding of CUB can reflect the role of selective forces in molecular mechanism, fitness strategies and genome evolution.

Previously, it was thought that synonymous substitution of codons was silent and would not alter the properties of encoded protein product and thus regarded as insignificant. On the contrary, some reports were shown to prove that the effects of synonymously substituted codons under evolutionary forces were related to geographic locations (Zhang et al., 2011) and host adaptive strategies. Biased preference of codons might influence mRNA splicing, protein structure, function and expression. Thus, the therapeutic target for many diseases could be identified and altered for disease resistance (Hunt et al., 2009). Hence, CUB studies in both prokaryotes and eukaryotes are thought to be functionally important (Hunt et al., 2014; Elmchi-Sarfaty et al., 2007). Earlier reports explained the major rationales for codon usage variation and highlighted the mutational bias hypothesis and the natural selection model. The mutational bias hypothesis states, codon usage variation is

* Corresponding author.
E-mail address: supriyoch_2008@rediffmail.com (S. Chakraborty).

https://doi.org/10.1016/j.virusres.2020.198248
Received 14 August 2020; Received in revised form 11 November 2020; Accepted 23 November 2020
Available online 27 November 2020

Principal
M. H. C. M. Sc. College
Algapur, Hailakandi

Gene 777 (2021) 145462

Contents lists available at ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Research paper

Understanding the codon usage patterns of mitochondrial CO genes among Amphibians

Parvin A. Barbhuiya [a,1], Arif Uddin [b,2], Supriyo Chakraborty [a,*,3]

[a] Department of Biotechnology, Assam University, Silchar 788150, Assam, India
[b] Department of Zoology, Moinul Hoque Choudhury Memorial Science College, Algapur, Hailakandi 788150, Assam, India

ARTICLE INFO

Keywords:
Codon usage bias
CO genes
Natural selection
Mutation pressure

ABSTRACT

A universal phenomenon of using synonymous codons unequally in coding sequences known as codon usage bias (CUB) is observed in all forms of life. Mutation and natural selection drive CUB in many species but the relative role of evolutionary forces varies across species, genes and genomes. We studied the CUB in mitochondrial (mt) CO genes from three orders of Amphibia using bioinformatics approach as no work was reported yet. We observed that CUB of mt CO genes of Amphibians was weak across different orders. Order Caudata had higher CUB followed by Gymnophiona and Anura for all genes and CUB also varied across genes. Nucleotide composition analysis showed that CO genes were AT-rich. The AT content in Caudata was higher than that in Gymnophiona while Anura showed the least content. Multiple investigations namely nucleotide composition, correspondence analysis, parity plot analysis showed that the interplay of mutation pressure and natural selection caused CUB in these genes. Neutrality plot suggested the involvement of natural selection was more than the mutation pressure. The contribution of natural selection was higher in Anura than Gymnophiona and the lowest in Caudata. The codons CGA, TGA, AAA were found to be highly favoured by nature across all genes and orders.

1. Introduction

Genetic code evolved over a long time scale. Vertebrate mitochondria use a deviant genetic code which is presumed to be the derivative of standard genetic code (Knight et al., 2001; Santos et al., 2004). Codon usage bias (CUB), a universal trend to use synonymous codons in a skewed manner, is a phenomenon witnessed in all forms of life. CUB varies between species and within individual organisms and is believed to present an evolutionary advantage (Grantham et al., 1980). CUB can affect transcription, mRNA stability, initiation and elongation of polypeptide chain, and post-translational modification of proteins (Chamary et al., 2006; Rodriguez et al., 2012; Trotta, 2013). Various aspects of how and why codon bias occurs remain unresolved. Explanation of underlying biases for synonymous usage of codons is often attributed to mutation bias, natural selection and genetic drift. However, the explanations of mutation and natural selection are not mutually exclusive, rather mutation and natural selection operate simultaneously and CUB, *in fact, reflects a balance between them* (Bulmer, 1991). *Knowing the* principles of CUB will deepen our understanding of the selective forces which take part not only in genome evolution, but also in creating and maintaining CUB. Codon usage is a crucial determining factor of gene expression due to its effects on transcription (Zhou et al., 2016). While some other studies showed that selection for efficient and proper translation is the major cause of synonymous CUB (Akashi, 1994; Drummond and Wilke, 2008; Gingold and Pilpel, 2011; Plotkin and Kudla, 2011).

Until relatively recently, wobble or synonymous mutation was thought to be silent as it would not alter the aminoacid at specific position in a protein and was consequently regarded as inconsequential. A number of studies have shown that synonymous mutations are under long-term evolutionary pressure and more importantly, synonymous mutations are found to cause diseases (Sauna et al., 2007). Several

Abbreviations: CUB, Codon usage bias;; ENC, Effective number of codons;; RSCU, Relative synonymous codon usage.

* Corresponding author.
E-mail address: supriyo123chakraborty@yahoo.in (S. Chakraborty).
[1] ORCID: 0000-0002-6512-1669.
[2] ORCID: 0000-0002-5173-9812.
[3] ORCID: 0000-0002-5996-2760.

https://doi.org/10.1016/j.gene.2021.145462
Received 15 August 2020; Received in revised form 18 December 2020; Accepted 20 January 2021
Available online 27 January 2021

Jyotishman Taye[1], Sushil Sinha[1] and A. F. Gulenur Islam Barbhuiya[1]

Turkish Online Journal of Qualitative Inquiry (TOJQI)
Volume 5, July 2021 : 3055 - 3062

Research Article

Witchcraft, Conflict and Mediation: The role of Semantics and Rhetoric for conflict resolution in the village Dhauliguri of Kokrajhar, Assam, India

Jyotishman Taye[1], Sushil Sinha[2] and A. F. Gulenur Islam Barbhuiya[3]

Abstract

Northeast India, a hub for various tribal communities, witnesses innumerable conflict in the past many decades. However, with the strong role of traditional leaders working under the customary laws, such conflict is kept under control and solved. Against this backdrop, the paper wants to attempt to highlight that traditional leaders can solve the conflict that arises out of witch-hunting practices in tribal areas of Assam. Witchcraft and its impact are deeply rooted in the form of insecurity across the tribal societies. Often labeling somebody witch also includes personal motives such revenge based, property dispute, male chauvinism or patriarchy, sexual orientation, or scapegoating for perpetrating witch-hunting. The victim can be blamed for illness or diseases to her family or neighbors, dying of the cattle in their neighborhood, or failure in crop cultivation during the harvest seasons (Islam & Ahmed, 2017). Using Max Gluchman analysis of the concept of semantics and rhetoric in the study of ritual and judicial processes and how judges manipulated culturally constituted notions to inform their rhetoric and finesse the ambiguity inherent in rules. Its emphasis on communication and dialogue enveloped by ritual and judicial processes is essential for conflict resolution. It has been argued that mediation and dialogue played a significant role in the litigation of conflict arising due to witchcraft. Although there is no doubt that beliefs in fear of witchcraft persisted, and witch-hunting is still widespread, however the atmosphere of conflict between parties can be resolved.

Key Words: Witchcraft, Conflict, Resolution, Tribals, North-East India.

Introduction

The conflict has been a core part of human society. Conflict of various forms occurs because of multiple reasons; however, mainly been because of human interaction. In the present paper, the conflict has been attached with arising from interpersonal relations and witch-hunting. Concerning conflict arising from witchcraft, different factors are involved. Sometimes personal grudge towards a person or family, and finally, such interpersonal jealously and hatred towards others may turn into labeling someone as a witch. Such interpersonal hostility can find in various societies. Witch-hunting may also reveal the underlying conflict between genders. The roots of this go very way past. We today are surrounded by various cases of witch-hunting, which are widespread because of people who harbor jealously, doubts, hatred either directly or indirectly because of a false rumor, or news and information, and so on. The National Crime Records Bureau (NCRB) reports that 768 women were murdered in India following accusations of being 'witches' in 2008-12 (National Crime Records Bureau, 2011). A report published by the North-east India

[1] Teaching Associate, Department of Anthropology, Gauhati University, Guwahati, Assam, India

[2] Retired Associate professor, Department of Anthropology, Gauhati University, Guwahati, Assam, India

[3] Assistant Professor, Department of Anthropology, M.H.C.M. Science Colllege, Hailakandi, Assam, India

Molecular Neurobiology
https://doi.org/10.1007/s12035-020-02068-0

Compositional Features and Codon Usage Pattern of Genes Associated with Anxiety in Human

Arif Uddin[1]

Received: 30 January 2020 / Accepted: 10 August 2020

Abstract
Codon usage bias (CUB) is the unequal usage of synonymous codon; some codons are more preferred than others. CUB analysis has applications in understanding the molecular organization of genome, genetics, gene expression, and molecular evolution. Bioinformatic approach was used to analyze the protein-coding sequences of genes involved in the anxiety to understand the patterns of codon usage as no work was reported yet. The improved effective number of codons (Nc) values ranged from 43.55 to 55.06, with a mean of 44.57, suggested that the overall CUB was low for genes associated with anxiety. The overall GC and AT content was 54.76 and 45.24, respectively. Relative synonymous codon usage (RSCU) analysis revealed that most frequently used codon ended mostly with C or G. The over-represented codons in genes associated with anxiety were CTG, ATC, GTG, AGC, ACC, and GCC, while under-represented codons were TTA, CTT, CTA, ATA, GTT, GTA, TCG, CCG, GCG, CAA, and CGT. Correlation analysis was performed between overall nucleotide composition and its 3rd codon positions, and observed highly significant ($p < 0.01$) correlation between them suggested that both mutation pressure and natural selection might affect the pattern of CUB. The highly significant correlation (0.598**, $p < 0.01$) was also observed between GC12 with GC3 suggested that directional mutation pressure might acted on all codon positions for genes associated with anxiety.

Keywords Codon usage bias · Anxiety mutation pressure · Natural selection · Gene expression level

Background

In modern lifestyles, stress and anxiety are two common psychiatric sign, and in small quantities, these are good. They are responsible for motivation and aids more productive. However, too much stress is harmful and persistent stress often leads to anxiety and unhealthy behaviors. Anxiety can be due to general poor health and specific physical or psychological sickness, namely, infection, heart disease, or depression. Anxiety is a disorder in central nervous system disorder [1, 2] and a common emotional state [3] and associated with uneasiness, discomfort, unpleasant, and concern or fear about some defined or undefined future menace [4]. It is the most common psychological illness and contributes one eighth of the total population [5]. The anxiety is a type of psychiatric disorders, and the prevalence is nearly 25% of the adult population at any time in their life. Its prevalence is 30.5% and 19.2% in women and men, respectively [6]. In young people, occurrence of anxiety is extremely high. In children, the prevalence rate of anxiety is 15.4%. The anxiety is distinguished into the "state" and the "trait" anxiety. The "state anxiety" occurs at a particular moment, and by the presence of an anxiogenic stimulus, it increases. In compare, "trait anxiety" is not due particular moment and is regarded as "enduring" feature of an individual [7].

Anxiety is due to neurotransmitters in the brain, and the brain synthesizes various neurotransmitters, namely, dopamine, serotonin, acetylcholine, adrenaline gamma amino butyric acid, and glutamate. Earlier report suggested that the anxiety is due to dysfunction of one or more neurotransmitters and their receptors. Gamma amino butyric acid (GABA) is one of the important inhibitory neurotransmitters in the mammalian brain, which regulates the neuronal excitability and thus serves as a "brake" during stress, and it is the brain's natural stress reliever [8]. Previous report suggested that GABAergic mechanism associated with the neurophysiology of anxiety [9]. Neurotransmitter serotonin involved in regulating emotional states in human [10]. The elevated level of noradrenaline is helpful in emergency situation or in fight/

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s12035-020-02068-0) contains supplementary material, which is available to authorized users.

✉ Arif Uddin
arif.uddin29@gmail.com

[1] Department of Zoology, Moinul Hoque Choudhury Memorial Science College, Algapur, Hailakandi, Assam 788150, India

Published online: 19 August 2020

Received: 23 July 2020 | Revised: 18 September 2020 | Accepted: 30 September 2020

DOI: 10.1002/arch.21750

RESEARCH ARTICLE

Archives of INSECT BIOCHEMISTRY AND PHYSIOLOGY WILEY

Analysis of mitochondrial protein-coding genes of *Antheraea assamensis*: Muga silkworm of Assam

Arif Uddin[1] | **Supriyo Chakraborty**[2]

[1]Department of Zoology, Moinul Hoque Choudhury Memorial Science College, Algapur, Assam, India

[2]Department of Biotechnology, Assam University, Silchar, Assam, India

Correspondence

Arif Uddin, Department of Zoology, Moinul Hoque Choudhury Memorial Science College, Algapur, Hailakandi 788150, Assam, India.
Email: arif.uddin29@gmail.com

Supriyo Chakraborty, Department of Biotechnology, Assam University, Silchar 788011, Assam, India.
Email: supriyoch_2008@rediffmail.com

Scientific editing by Yonggyun Kim.

Abstract

To understand the synonymous codon usage pattern in mitochondrial genome of *Antheraea assamensis*, we analyzed the 13 mitochondrial protein-coding genes of this species using a bioinformatic approach as no work was reported yet. The nucleotide composition analysis suggested that the percentages of A, T, G,and C were 33.73, 46.39, 9.7 and 10.17, respectively and the overall GC content was 19.86, that is, lower than 50% and the genes were AT rich. The mean effective number of codons of mitochondrial protein-coding genes was 36.30 and it indicated low codon usage bias (CUB). Relative synonymous codon usage analysis suggested overrepresented and underrepresented codons in each gene and the pattern of codon usage was different among genes. Neutrality plot analysis revealed a narrow range of distribution for GC content at the third codon position and some points were diagonally distributed, suggesting both mutation pressure and natural selection influenced the CUB.

KEYWORDS

codon usage bias, *Antheraea assamensis*, natural selection, mutation pressure

1 | INTRODUCTION

Codon usage bias (CUB) is a regular tendency for unequal usage of synonymous codons in messenger RNA (mRNA) molecules and occurs within the individual organisms or between the organisms (Grantham et al., 1980). From the beginning of its discovery in ribosomal protein genes of *Escherichia coli*, reliable explanations have been proposed and led to a hypothesis that CUB in bacteria was a genomic approach for the efficiency or accuracy of the optimal translation

Received: 29 October 2019 | Accepted: 5 January 2020

DOI: 10.1002/iub.2231

HYPOTHESIS

IUBMB LIFE WILEY

Similarities and dissimilarities of codon usage in mitochondrial ATP genes among fishes, aves, and mammals

Arif Uddin[1] | **Tarikul H. Mazumder**[2] | **Parvin A. Barbhuiya**[2] | **Supriyo Chakraborty**[2]

[1]Department of Zoology, Moinul Hoque Choudhury Memorial Science College, Hailakandi, Assam, India

[2]Department of Biotechnology, Assam University, Silchar, Assam, India

Correspondence
Arif Uddin, Department of Zoology,

Abstract

In this study, we used bioinformatic approach to analyze the compositional features and codon usage bias (CUB) of *ATP6* and *ATP8* genes among three groups, namely, fishes, aves, and mammals which thrive in three different habitats as no work was reported yet. The coding sequences of these genes were retrieved from the National Center for Biotechnology Information to explore the similarities and

Biochemical Genetics
https://doi.org/10.1007/s10528-020-10000-3

ORIGINAL ARTICLE

A Crosstalk on Codon Usage in Genes Associated with Leukemia

Supriyo Chakraborty[1] · Durbba Nath[1] · Sunanda Paul[1] · Yashmin Choudhury[1] · Yeongseon Ahn[2] · Yoon Shin Cho[2] · Arif Uddin[3]

Received: 27 March 2020 / Accepted: 18 September 2020

Abstract

Leukemia is the outcome of aggregation of damaged white blood cells. Several genes were reported to be associated with the pathogenesis of leukemia. These genes were computationally analyzed to decipher their codon usage bias (CUB) and to identify the prime factors influencing the codon usage profile as no work was reported yet. The mean values of synonymous codon usage order (SCUO) parameter indicated low CUB of the genes. Significant positive association of SCUO with overall GC and positional GCs might signal the presence of mutational pressure. However, neutrality plot suggested the dominant role of natural selection across the genes. Along with natural selection, the role of mutation pressure was also prominent and that might be responsible for lower CUB (SCUO = 0.19) of genes. Low translational speed might permit accuracy in the process. A strong inverse relationship of translational rate was observed with CUB of genes and folding energy.

Keywords Codon usage bias · Leukemia · Directional mutational pressure · Translational speed · Folding energy

Introduction

In living cells, the intrinsic feature of redundant genetic code is not engaged impartially, so the counts of specific synonymous codons diverge from expected codon usage (Peng et al. 2018; Xiang et al. 2015). The phenomenon of codon usage bias (CUB) is a typical feature of the protein coding sequences and appears to be distinct

✉ Supriyo Chakraborty
supriyoch_2008@rediffmail.com

[1] Department of Biotechnology, Assam University, Silchar, Assam 788011, India

[2] Department of Biomedical Science, Hallym University, Hallymdachak-gil, Chuncheon 24252, Gangwon-do, Korea

[3] Department of Zoology, Moinul Hoque Choudhury Memorial Science College, Algapur, Hailakandi, Assam 788150, India

Published online: 28 September 2020

Biotechnol Lett (2020) 42:1865–1875
https://doi.org/10.1007/s10529-020-02931-z

ORIGINAL RESEARCH PAPER

Codon usage trend in genes associated with obesity

Supriyo Chakraborty · Parvin A. Barbhuiya · Sunanda Paul · Arif Uddin · Yashmin Choudhury · Yeongseon Ahn · Yoon Shin Cho

Received: 29 February 2020 / Accepted: 29 May 2020 / Published online: 1 June 2020

Abstract Obesity is not only a social menace but also an economic burden as it reduces productivity and increases health care cost. We used bioinformatic tools to analyze the CUB of obesity associated genes and compared with housekeeping genes (control) to explore the similarities and differences between two data sets as no work was reported yet. The mean effective number of codons (ENC) in genes associated with obesity and housekeeping gene was 50.45 and 52.03 respectively, indicating low CUB. The relative synonymous codon usage (RSCU) suggested that codons namely CTG and GTG were over-represented in both obesity and housekeeping genes while under-represented codons were TCG, TTA, CTA, CCG, CAA, CGT, ATA, ACG, GTA and GCG in obesity genes and TCG, TTA, CCG, ATA, ACG, GTA, and GCG in housekeeping genes. t test analysis suggested that 11 codons namely TTA (Leu), TTG (Leu), CCG (Pro), CAC (His), CAA (Gln), CAG (Gln), CGT (Arg), AGA (Arg), ATA (Ile), ATT (Ile) and GCG (Ala) were significantly differed ($p < 0.05$ or $p < 0.01$) between obesity and housekeeping genes. Highly significant correlation was observed between GC12 and GC3 in obesity and housekeeping genes i.e. $r = 0.580^{**}$ and $r = 0.498^{**}$ ($p < 0.01$) respectively indicating the effect of directional mutation pressure present in all codon positions.

Keywords Codon usage bias · Obesity · Directional mutation pressure · GC richness

Introduction

All amino acids (except methionine and tryptophan) are determined by 2–6 codons, known as synonymous codons, in standard genetic code. Several studies have reported that the usage of synonymous codons in mature mRNA molecules is not equal. This asymmetrical occurrence of synonymous codons for an amino acid leading to the preferential recurrence of specific codons in mRNA molecules is known as codon usage bias or CUB (Brown et al. 2009). CUB is generally believed to be the product of both mutational and selection pressures acting on genes or genomes (Grantham et al. 1980). Recent research has found that CUB affects the efficiency and accuracy of protein

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10529-020-02931-z) contains supplementary material, which is available to authorized users.

S. Chakraborty (✉) · P. A. Barbhuiya · S. Paul · A. Uddin · Y. Choudhury
Department of Biotechnology, Assam University, Silchar, Assam 788011, India
e-mail: supriyoch_2008@rediffmail.com

Y. Ahn · Y. S. Cho (✉)
Department of Biomedical Science, Hallym University, Hallymdachak-gil, Chuncheon 24252, Gangwon-do, Korea
e-mail: yooncho33@hallym.ac.kr

Contents lists available at ScienceDirect

Mutation Research - Fundamental and Molecular Mechanisms of Mutagenesis

journal homepage: www.elsevier.com/locate/mut

Synonymous codon usage and context analysis of genes associated with pancreatic cancer

Supriyo Chakraborty [a,*], Sunanda Paul [a], Durbba Nath [a], Yashmin Choudhury [a], Yeongseon Ahn [b], Yoon Shin Cho [b], Arif Uddin [c]

[a] Department of Biotechnology, Assam University, Silchar 788011, Assam, India
[b] Department of Biomedical Science, Hallym University, Hallymdachak-gil, Chuncheon, Gangwon-do 24252, South Korea
[c] Department of Zoology, Moinul Hoque Choudhury Memorial Science College, Algapur, Hailakandi 788150, Assam, India

ARTICLE INFO

Keywords:
Pancreatic cancer
Codon usage bias
Mutation pressure
Natural selection

ABSTRACT

Pancreatic cancer is a fatal disorder which originates in pancreas. Its mortality rate is increasing with time. Some studies also reported that pancreatic cancer would be ranked 2nd by the year 2030. Codon usage bias (CUB) arises when synonymous codons for each amino acid are not used randomly in the coding sequences of genes. We used bioinformatic methods to analyze the compositional properties, codon context and codon usage trend of the genes associated with pancreatic cancer as no work was reported yet. From the base composition analysis, the pancreatic cancer genes were found to be GC-rich and at the 3rd codon position the G/C ending codons were more preferred to A/T ending ones. The CUB was low in genes associated with pancreatic cancer. Correspondence analysis proposed that other than base constraints, CUB might also be affected by some other factors such as natural selection. Moreover, results of correlation analysis indicated that CUB and various GC contents *i.e.* GC, GC1, GC2, GC3 played important role in the release of free energy by transcripts of the genes associated with pancreatic cancer. The low compAI values of coding sequences suggested a low translation rate of the genes.

1. Introduction

Condon usage bias (CUB) is the condition that arises when synonymous codons of amino acids are not used equally in the coding sequences of genes. CUB varies between genes within the same genome, even between different genomes and also for same gene across species [1]. CUB is an evolutionary footprint that helps us study the changes in a particular gene throughout evolution and it is a species specific characteristic. Owing to degeneracy of genetic code, more than one codon encodes the same amino acid which gives evolution a chance to regulate the translational accuracy/efficiency of the particular gene. Such codons are said to be synonymous codons which mostly vary for nucleotide at Wobble's position [1].

CUB is influenced by several factors such as gene expression level, protein length, codon location, secondary structure, evolution rate, base composition, environment, mutation pressure, natural selection, *etc.* But mutation pressure and/or natural selection have long been considered as the two main factors governing CUB of a particular gene reported by previous studies [1]. Codon bias might vary according to codon location [2]. In contrast to lowly expressed genes, CUB was reported to be high in the highly expressed genes [3]. For studying the impact of such influencing factors certain parameters are used to estimate CUB [1]. These parameters involve base compositional study, analysis of parity plot, neutrality plot, correspondence analysis, synonymous codon usage order (SCUO), relative synonymous codon usage (RSCU), folding energy, competition adaptive index (compAI) *etc.* [4,5].

CUB analysis can be used for understanding molecular mechanism and evolution of genes [1]. Hence, it may be considered as an evolutionary relic. CUB is a useful tool for determining the changes of evolution in different species and genes. Several studies reported CUB analysis could be used in designing of primers and transgenes, predicting gene function and its expression level, discovery of new gene, phylogenetic studies, *etc.* [6–10]. It also helps in understanding the molecular evolution in genes and evolutionary relationship among organisms. CUB analysis gives basic information that could be potentially used in genetic engineering research [11].

Uddin et al. [11] analyzed the codon usage pattern of ovarian cancer genes wherein they observed a low codon usage bias and GC-richness of

* Corresponding author.
E-mail address: supriyoch_2000@rediffmail.com (S. Chakraborty).

https://doi.org/10.1016/j.mrfmmm.2020.111719
Received 18 June 2019; Received in revised form 24 July 2020; Accepted 4 August 2020
Available online 11 August 2020

Planta (2020) 252:67
https://doi.org/10.1007/s00425-020-03470-7

ORIGINAL ARTICLE

Analysis of codon usage bias of chloroplast genes in *Oryza* species

Codon usage of chloroplast genes in *Oryza* species

Supriyo Chakraborty[1] · Sophiarani Yengkhom[1] · Arif Uddin[2]

Received: 25 March 2020 / Accepted: 15 September 2020

Abstract

Main conclusion **The codon usage bias in chloroplast genes of *Oryza* species was low and AT rich. The pattern of codon usage was different among *Oryza* species and mainly influenced by mutation pressure and natural selection.**

Abstract Codon usage bias (CUB) is the unequal usage of synonymous codons in which some codons are more preferred to others in the coding sequences of genes. It shows a species-specific property. We studied the patterns of codon usage and the factors that influenced the CUB of protein-coding chloroplast (cp) genes in 18 *Oryza* species as no work was yet reported. The nucleotide composition analysis revealed that the overall GC content of cp genes in different species of *Oryza* was lower than 50%, i.e., *Oryza* cp genes were AT rich. Synonymous codon usage order (SCUO) suggested that CUB was weak in the cp genes of different *Oryza* species. A highly significant correlation was observed between overall nucleotides and its constituents at the third codon position suggesting that both, mutation pressure and natural selection, might influence the CUB. Correspondence analysis (COA) revealed that codon usage pattern differed across *Oryza* species. In the neutrality plot, a narrow range of GC3 distribution was recorded and some points were diagonally distributed in all the plots, suggesting that

Received: 18 February 2020 | Revised: 7 August 2020 | Accepted: 31 August 2020

DOI: 10.1002/jcp.30050

ORIGINAL RESEARCH ARTICLE

Journal of Cellular Physiology WILEY

Codon usage pattern and evolutionary forces of mitochondrial ND genes among orders of class Amphibia

Parvin A. Barbhuiya[1] | Arif Uddin[2] | Supriyo Chakraborty[1]

[1]Department of Biotechnology, Assam University, Silchar, Assam, India

[2]Department of Zoology, Moinul Hoque Choudhury Memorial Science College, Hailakandi, Assam, India

Correspondence
Supriyo Chakraborty, Department of Biotechnology, Assam University, Silchar 788011, Assam, India.

Abstract

In this study, we used a bioinformatics approach to analyze the nucleotide composition and pattern of synonymous codon usage in mitochondrial ND genes in three amphibian groups, that is, orders Anura, Caudata, and Gymnophiona to identify the commonality and the differences of codon usage as no research work was reported yet. The high value of the effective number of codons revealed that the codon usage bias (CUB) was low in mitochondrial ND genes among the orders. Nucleotide composition analysis suggested

ISSN (Online): 2349-1183; ISSN (Print): 2349-9265
TROPICAL PLANT RESEARCH
The Journal of the Society for Tropical Plant Research
7(2): 396–402, 2020
DOI: 10.22271/tpr.2020.v7.i2.046

Research article

Morphological markers associated with pericarp colour and its inheritance pattern in black scented rice of Manipur

Warepam Jesmi Devi[1], Yaikhom Vivekananda[2], Arif Uddin[1], J. M. Laishram[2] and Supriyo Chakraborty[1]*

[1]Assam University, Silchar 788011, Assam, India
[2]College of Agriculture, Central Agricultural University, Imphal, Manipur, India

*Corresponding Authors: supriyoch_2008@rediffmail.com
[Accepted: 26 July 2020]

Abstract: Rice (*Oryza sativa*) is the most widely consumed staple food for a large section of world's population. Biotechnological developments have led plant breeders to develop for more efficient selection strategies instead of the traditional phenotype-based plant selection method. Rice varieties other than white are usually called red or black rice. Black scented rice is in great demand and mostly grown in Asian countries. The intensity of pericarp in black rice ranges from dark purple to light purple, and the genetic mechanism underlying this colour variation is yet to be identified. Therefore, the inheritance patterns of black rice cultivars and the phenotypic markers associated with black pericarp colour have been studied in this study. Phenotypic investigation of the black scented rice may aid in the breeding of anthocyanin-rich rice varieties and may provide insights into the potential enhancement of this valuable antioxidant in a variety of foodstuffs. The experiments were conducted on F_1 and F_2 individuals developed from the crosses between indigenous black scented rice (*Chakhao*) cultivars of Manipur and the local white landrace cultivars. The segregation ratio of F_2 individuals was analyzed with the chi-square formula. The F_2 population showed some of the morphological markers like purple coleoptiles, purplish-black colour at leaf tip and stalks and the formation of black rings (auricles) were associated with pericarp colours. This, in fact, revealed the inheritance pattern of pericarp colour in black scented rice. The present study provides useful information on the inheritance of pericarp colour in black scented rice of Manipur and possesses the potential for their further genetic improvement.
Keywords: Chakhao Rice - Anthocyanin - Phenotypic markers - Hereditary pattern.

[Cite as: Devi WJ, Vivekananda Y, Uddin A, Laishram JM & Chakraborty S (2020) Morphological markers associated with pericarp colour and its inheritance pattern in black scented rice of Manipur. *Tropical Plant Research* 7(2): 396–402]

INTRODUCTION

Rice as a cereal grain is the most widely consumed staple food for a large section of the world's population, particularly in Asian countries. Rice varieties other than white are usually called red or black rice. Black scented rice also known as specialty rice which has several unique properties like unique aroma, colour (red, purple, black), texture (glossiness and stickiness), chemical composition, aesthetic value, waxyness (very low amylose content) and because of its superior processing quality which have increased its demand in the market. In Asian countries, black rice is often consumed after mixing with white rice to enhance the flavour, colour, and nutritional value (such as high protein, total essential amino acids, vitamin B1 and minerals - Fe, Zn, Mn and P) (Yang *et al.* 2008). A spoonful of black rice bran contains not only more health- promoting anthocyanin as antioxidant than that found in a spoonful of blueberries, but also contains less sugar, more fibre and vitamin E. The purple/black colour of black scented rice is due to high anthocyanin content in the pericarp. Total anthocyanin content in Chakhao varieties was found to be 716 mg kg^{-1} of dried powder sample. And the total phenolic content was 539 mg / 100 g of the dried powder sample as Gallic acid equivalent in Chakhao varieties (Asem *et al.* 2015).

Received: 23 February 2020
Published online: 31 August 2020
https://doi.org/10.22271/tpr.2020.v7.i2.046

Contents lists available at ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Research paper

Analysis of compositional properties and codon usage bias of mitochondrial CYB gene in anura, urodela and gymnophiona

Parvin A Barbhuiya[a], Arif Uddin[b], Supriyo Chakraborty[a,*]

[a] Department of Biotechnology, Assam University, Silchar 788011, Assam, India
[b] Department of Zoology, Moinul Hoque Choudhury Memorial Science College, Algapur, Hailakand 788150, Assam, India

ARTICLE INFO

Keywords:
Codon usage bias
CYB
Mutation pressure and natural selection

ABSTRACT

We delineated the pattern of synonymous codon usage bias (CUB) and its determinants in mitochondrial CYB gene of respiratory chain across different amphibian groups namely orders anura, urodela and gymnophiona. We observed that CUB was low in CYB gene of amphibia. The gymnophionans had comparatively high bias followed by urodeles and anurans. The codons namely TCA, CCA, CAA, CGA, TGA, AAA and ACA were over-represented in all three orders. The codons such as GCC and TCC were over-represented in anura whereas in urodela, the over-represented codons were TTA, CTA, ATA, GTA, GAA, GGA and GCA. In gymnophiona, GCC, TTA, CTA, ATA, GTA, GAA and GGA codons were over-represented. The regression analysis between effective number of codons (ENC) and nucleobase at the 3rd position revealed that nucleobase A and C influenced CUB positively in order anura, while in urodela and gymnophiona, nucleobase A and T influenced the CUB positively. Mutation pressure and natural selection mutually illustrate the CUB of CYB gene (complex III gene) of amphibia as elucidated by correlation analysis between 3rd nucleotide in a codon and overall nucleotide content of the gene. However, neutrality plot showed that natural selection was the dominant evolutionary factor of CUB.

1. Introduction

Due to the degeneracy of the genetic code, considerable choice persists in usage of one codon over another. In practice, synonymous codon preferences vary from one organism to the other and this evolutionary phenomenon of bias in codon choice is termed as codon usage bias (CUB) (Bulmer 1991). This phenomenon varies in all organisms and at times different parts of the genome of the same organism have different CUB (Hershberg and Petrov 2008). Previous reports on codon usage suggest that genetic drift together with mutation and selection have been the predominant determinants of codon usage (Bulmer 1991). A general consensus opinion exists that the major cause for selection on codon bias is for efficient synthesis of protein by optimal codons which contribute fitness benefit (Shields et al., 1988; Stenico et al., 1994). In addition, many aspects that mediate such bias include tRNA abundance (Ikemura 1985; Duret 2000), GC (guanine-cytosine) content (Chen et al., 2013), protein structure (Saunders and Deane 2010), protein hydropathy (Romero et al., 2000a, Romero et al., 2000b; Mirsafian, et al., 2014), recombination (Hey and Kliman 2002), gene length (Duret and Mouchiroud 1999), gene expression level (Blake et al., 2003; Hiraoka et al., 2009), environmental stress etc (Lynn et al., 2002). Detailed research on codon usage allows to perceive our knowledge on genome evolution (Sharp and Matassi 1994), designing degenerate primers (Zhou et al., 2005) and for heterologous gene expression (Kane 1995). Codon usage pattern in various organisms from bacteria, virus to developed and complex eukaryotes were studied earlier (Lloyd and Sharp 1991; Sharp et al., 1993; Akashi 1997; Shackelton et al., 2006).

Base nucleotide composition constraint is thought to be the most significant factor which shapes CUB (Shang et al., 2011). The relationship between GC and synonymous codon usage was reported several times in the literature. It is a well-accepted fact that whole genomes may differ widely in their average G + C content (Muto and Osawa 1987). Moreover, G + C content of individual genes from various genomes may also vary (Jermiin et al., 1994). The variation in G + C content can occur equally at both synonymous and nonsynonymous codon sites (Jukes and Bhushan 1986). Earlier it was presumed that synonymous changes have no effect on protein and was thought as functionally neutral. But recent findings suggest that synonymous sites have functional roles and affect protein expression and conformation (Seligmann and Warthi 2017). Cancer (Gartner et al., 2013; Schutz et al., 2013) and numerous other prevalent diseases are linked to

Abbreviations: CUB, codon usage bias; ENC, effective number of codons; RSCU, relative synonymous codon usage

* Corresponding author.
E-mail address: supriyoch_2008@rediffmail.com (S. Chakraborty).

https://doi.org/10.1016/j.gene.2020.144762
Received 12 February 2020; Received in revised form 24 April 2020; Accepted 8 May 2020

Principal 04/07/23
M. H. C. M. Sc. College
Algapur, Hailakandi

Contents lists available at ScienceDirect

Molecular & Biochemical Parasitology

journal homepage: www.elsevier.com/locate/molbiopara

Analysis of codon usage pattern of mitochondrial ND genes in Platyhelminthes

Gulshana A. Mazumder[a], Arif Uddin[b], Supriyo Chakraborty[a,*]

[a] Department of Biotechnology, Assam University, Silchar, 788011 Assam, India
[b] Moinul Hoque Choudhury Memorial Science College, Algapur, Hailakandi, 788150 Assam, India

ARTICLE INFO

Keywords:
Codon usage bias
Evolutionary forces
Natural selection
Mutation pressure

ABSTRACT

Codon usage bias (CUB) is the nonrandom usage of synonymous codons in which some codons are more preferred to others.CUB can be determined by mutation pressure and selection. Various approaches have been used to understand the pattern of CUB in the mitochondrial ND (MT-ND or ND) genes involved in complex I of respiratory chain in five different classes of Platyhelminthes as no work was reported yet. The present study revealed that the CUB varies across MT-ND genes and the coding sequences showed the richness of A and T. Correspondence analysis implied the effect of mutational pressure and also the pattern of codon usage was different in different classes of platyhelminthes for MT-ND genes. Highly significant correlation was observed between overall nucleotide composition and its 3rd codon position in most of the homogeneous nucleotides such as A% and A3%, T% and T3%, G% and G3%, C% and C3%, GC% and GC3% and also some significant correlations observed among heterogeneous nucleotides in all the five classes for MT-ND genes suggested the role of mutational pressure as well as natural selection in affecting the CUB. Neutrality plot suggested that the contributions of natural selection and mutational pressure varied across different classes of platyhelminthes and also differed in different MT-ND genes.

1. Introduction

The standard genetic code describes the presence of 61 codons that are known to encode 20 amino acids, and each amino acid can be encoded by 1–6 synonymous codons [1]. The phylum Platyhelminthes, also known as flatworms, belongs to invertebrate. The mitochondria of Platyhelminthes follow Translation Table 14 of NCBI. Unlike that of standard genetic code, the mitochondrial genetic code of Platyhelminthes is distinguished by 63 sense codons and only 1 stop codon (TAG). In this genetic code, the amino acid serine is exceptionally encoded by 8 synonymous codons and two amino acids namely Tyr (TAT, TAC, and TAA) and Asn (AAT, AAC, AAA) are encoded by 3 codons each. Trp is encoded by 2 synonymous codons (TGA, TGG), whereas, Met (ATG) and Lys (AAG) are encoded by single codon each.

Codon usage bias (CUB) is the event where a specific codon is more preferred to its synonymous partners which codify the same amino acid [2]. The pattern of codon usage is found to be impacted by some factors like mutation pressure, natural selection, compositional bias (GC% and GC skew), level of gene expression as well as gene length, RNA stability, DNA replication, properties of protein such as hydrophobicity and hydrophilicity [3–6]. Among all the mentioned factors, the compositional constraints in presence of both natural selection and mutation pressure are reviewed as the major impacting factors on CUB [7–9]. Codon usage indices may be used to predict and optimize protein expression levels, to identify protein coding genes, and also to detect lateral gene transfer in organisms [10].

Platyhelminthes (flatworms) are bilaterally symmetrical, soft-bodied animals and consist of about 20,000 species worldwide. They are classified into five classes *viz.* Cestoda (tapeworm), Monogenea, Rabditophora, Trematoda (fluke) and Turbellaria. Most of its species are parasitic in nature and cause heavy damage to human health [11]. They are one of the major human parasitic groups. Turbellaria is a free living parasite and includes the planarian, *Dugesia*, which is mostly found in oceans, in fresh water, and in damp terrestrial habitats, and a very few species are parasitic [12]. The Cestoda are the parasites which are found in intestines of vertebrates, which may show structural as well as modified life history for parasitisation [13]. Characteristically flatworms are hermaphrodites (individual produces both eggs and sperms). Trematodas and Cestodes are found to shed their eggs continuously, while the Turbellarians are seasonally reproductive and, also display asexual reproduction [14]. In animal evolution, the structure of platyhelminthes or flatworms marks an important step. Their

* Corresponding author.
E-mail address: supriyoch_2008@rediffmail.com (S. Chakraborty).

https://doi.org/10.1016/j.molbiopara.2020.111294
Received 30 April 2020; Received in revised form 18 June 2020; Accepted 20 June 2020
Available online 24 June 2020

MITOCHONDRIAL DNA PART A
https://doi.org/10.1080/24701394.2020.1800661

Taylor & Francis
Taylor & Francis Group

ARTICLE

Codon usage pattern and its influencing factors for mitochondrial *CO* genes among different classes of Arthropoda

Riazul Islam Barbhuiya[a], Arif Uddin[b] and Supriyo Chakraborty[a]

[a]Department of Biotechnology, Assam University, Silchar, India; [b]Department of Zoology, M. H. C. M. Science College, Hailakandi, India

ABSTRACT

Analysis of codon usage bias (CUB) is very much important in perceiving the knowledge of molecular biology, the discovery of a new gene, designing of transgenes and evolution of gene. In this study, we analyzed compositional features and codon usage of *MT-CO* (*COI*, *COII* and *COIII*) genes among the classes of Arthropoda to explore the pattern of CUB as no research work was reported yet. Nucleotide composition analysis in *CO* genes suggested that the genes were AT-rich in all the four classes of Arthropoda. CUB was low in all the classes of Arthropoda for *MT-CO* genes as revealed from a high effective number of codons (ENC). We also found that the evolutionary forces namely mutation pressure and natural selection were the key influencing factors in CUB among *MT-CO* genes as revealed by correlation analysis between overall nucleotide composition and nucleotide composition at the 3rd codon position. Correspondence analysis suggested that the pattern of CUB was different among the classes of Arthropoda. Further, it was revealed from the neutrality plot that natural selection had a dominant role while mutation pressure exhibited a minor role in structuring the pattern of codon usage in all the classes of Arthropoda across *COI*, *COII* and *COIII* genes.

ABBREVIATIONS: CUB: codon usage bias; ENC: effective number of codons; RSCU: relative synonymous codon usage; COA: correspondence analysis

ARTICLE HISTORY
Received 26 March 2020
Accepted 18 July 2020

KEYWORDS
Codon usage bias; natural selection; mutation pressure; cytochrome oxidase gene; electron transport system

Introduction

Codon usage bias (CUB) is the phenomenon of uneven usage of synonymous codons in the course of the translation process of genes where the usage of some specific synonymous codons is more repeated than others. The more repeatedly used codons are named as optimal codons or preferred codons. The phenomenon of CUB is familiar in a variety of organisms ranging from unicellular (prokaryotic) to multicellular (eukaryotic) organisms (Akashi 1994). Codon usage pattern differs in various tissues of the same organism (Plotkin et al. 2004). It is considered that compositional properties with the influence of mutational pressure and natural selection influence the codon usage pattern of genes (Shields et al. 1988; Sharp et al. 1993; Stenico et al. 1994). On the other hand, a mutation occurs in wobble bases, where the changes in synonymous codons occur at the third codon position by keeping the same encoded amino acid in the same position and as a result protein's initial sequence is maintained (Biro 2008). The pattern of codon usage is governed by genetic drift and mutational pressure in some organisms while in others, it is governed by the balance between mutational pressure and natural selection (Bulmer 1991). In some genes, mutational pressure with an increased amount of any one of the four nucleotides plays a dominant part in structuring the usage pattern of synonymous codons (Sharp et al. 1993; Karlin and Mrázek 1996; Zhao et al. 2007; Zhong et al. 2007). In an open reading structure, mutational bias is indicated by very high or low content of C or G in the 3rd position of codon (Sueoka 1988). Some previous findings recommended that in the genes showing high expression level, CUB is governed by translational selection. Preferred codons can easily be traced in highly expressed genes by examining the abundance of tRNA molecules (Bibb et al. 1984; McEwan and Gatherer 1999).

Structurally, mitochondrial DNA is a double-stranded, closed molecule with a size of approximately 16.6 kb with genes for 22 tRNAs, 2 rRNAs, and 13 polypeptides (Chen et al. 2009). A higher rate of mutation occurs in mitochondrial DNA and the latter is inherited maternally (Taylor and Turnbull 2005). Due to the lack of introns and histones in the mitochondrial genome, the existence of reactive oxygen species (ROS) produced by the oxidative phosphorylation makes the mitochondrial DNA more vulnerable to mutation (Matsukage et al. 1975; Kunkel and Loeb 1981; Shay and Werbin 1992; Torri and Englund 1995; Singh et al. 2001; Modica-Napolitano and Singh 2004). It was reported that the DNA of mitochondria shows ten times more mutation than nuclear DNA (Wilson and Roof 1997; Shoubridge 2000). Due to its very small size, relatively packed gene content and greater mutation rate, mitochondrial DNA is considered to be the perfect tool for the study of evolution (Clark et al. 2007).

CONTACT Supriyo Chakraborty supriyoch_2008@rediffmail.com

Supplemental data for this article can be accessed here.

04/07/23
Principal
M. H. C. M. Sc. College
Algapur, Hailakandi

Biological and Medicinal Chemistry

In Silico Screening of Some Antiviral Phytochemicals as Drug Leads Against Covid-19

Working Paper

Monjur Ahmed Laskar, Moriom Begam, Manabendra Dutta Choudhury Assam University

Abstract

Background: COVID-19 caused by SARS-CoV-2 in December 2019 has become a pandemic

hazard to the community health. It is a respiratory difficulty causing fever, dry cough, fatigue,

shortness of breath, muscle aches and some instances lead to pneumonia. Coronaviruses have

European Journal of Molecular & Clinical Medicine
ISSN 2515-8260 Volume 07, Issue 03, 2020

Nutritional Status Of The Rabha Tribal Children Of Udalguri District Of Assam, India

Nabajit Kr. Das[1], A. F. Gulenur Islam Barbhuiya[2], Rijumoni Sarma[3]
[1]*Maulana Abul Kalam Azad Institute of Asian Studies, Kolkata, West Bengal, India*
[2]*M.H.C.M. Science Colllege, Hailakandi, Assam, India*
[3]*Directorate of Social Welfare, Guwahati, Assam, India*

ABSTRACT:

Tribal populations in India are considered to be socio-economically the most disadvantaged group and tribal children have very poor health indicators when compared to the rest of the population. In the present study an attempt has been made to assess the nutritional status of under- five years age group among the Rabha children of Udalguri district, Assam, India. For assessing the nutritional status of children, three anthropometric indices- height for age, weight for height and weight for age, which are considered as good indicators of assessment of nutritional status were adopted. Among the 225 children who were assessed for malnutrition, 93 (41.33%) were stunted, 62 (27.56%) wasted and 68 (30.22%) underweight as per WHO Growth Standards. The prevalence of malnutrition was found more common in girls as compared to boys. Education of the mothers was found to have a strong inverse relationship with all three measures of nutritional status. Education among the parents to improve nutritional status in the children as well as modification in the lifestyle can help to decline this health concern.

Keywords: Rabha children, India, malnutrition, stunted, wasted, underweight.

1. INTRODUCTION

Nutrition during the first five years of life has an impact not only on growth and morbidity during childhood, but also has a persistent impact on their physical and mental development and on their health status as adult. Nutritional deficiencies give rise to various morbidities, which in turn, may lead to increased mortality. Under nutrition is considered as major public health problem that closely associated with child mortality rates. The survivors of under nutrition have its pervasive effects that include acute morbidity as well as long-term impairment of cognitive & social development, physical work capacity, productivity, and economic growth (Jethy, 2008).

In India, 20% children of under five years children are suffer from wasting due to acute undernutrition. More than one third of the world's children who are wasted live in India. Forty three per cent of Indian children under five years are underweight and 48 percent are stunted due to chronic undernutrition, India accounts for more than 3 out of every 10 stunted children in the world (UNICEF, 2017). Despite rapid economic development along with increase in food production in recent decades and several nutritional intervention

This article was originally published in the

International Journal of

Ecology and Environmental Sciences

CODEN: IJESDQ ISSN: 0377-015X (Print); 2320-5199 (Online)

Homepage: http://www.nieindia.org/Journal/index.php/ijees/

Editors : **Brij Gopal, A. Raman and P.S. Pathak**

Published by the

National Institute of Ecology

Jaipur and New Delhi

For details, contact: Email: ijees1@gmail.com OR nieindia@gmail.com

International Journal of Ecology and Environmental Sciences 45 (1): 85-95, 2019
ISSN: 2320-5199 (Online)

Biomass Storage and Carbon Sequestration in Priority Bamboo Species in Relation to Village Physiography

ABUL FAZAL MAZUMDER[1], ASHESH KUMAR DAS[2] AND ARUN JYOTI NATH[3*]

Department of Ecology and Environmental Science, Assam University, Silchar 788011, Assam, India
[1] fazal.abz@gmail.com; [2] asheshkd@gmail.com; [3] arunjyotinath@gmail.com
* corresponding author

ABSTRACT

The necessity to enhance biotic carbon sink is one of the significance areas of research under the current scenario of anthropogenic climate change. Bamboos being managed in the rural landscape over the years could be an important strategy to promote carbon sink. Bamboos growing under natural conditions have been studied worldwide for their role in vegetation and soil carbon storage. However, village bamboos have received little attention. The present study was undertaken in Hailakandi district of Assam with the specific objectives: (i) to identify priority bamboo species in relation to village physiography, and (ii) to estimate biomass storage and carbon sequestration potential under at different ages of priority bamboo species in relation to village physiography. On the basis of village physiography, villagers have evolved their own priority species. The study shows that older clumps dominates over younger clump. The aboveground standing biomass stock was higher (0.25-24 Mg ha^{-1}) in flood unaffected villages than flood affected (0.14-15.75 Mg ha^{-1}) and riverside (0.38-15.44 Mg ha^{-1}) villages. Across different priority bamboo species, clump ages and village physiography, the estimated carbon sequestration rate ranges from 0.2 to 0.74 Mg ha^{-1} yr^{-1}. Although low in carbon sequestration rate, management of village bamboos can provide opportunity for long term carbon sink management. We suggest future studies to explore belowground biomass and soil organic carbon stock to represent the ecosystem carbon stock of village bamboos for better representation of such systems in carbon sink management.

Key Words: Carbon Sink; Biomass Models; Culm Density; Bamboo Flowering

INTRODUCTION

Combating greenhouse gas emission through reducing sources or enhancing sinks has been the priority theme of global research since mid-1990s. Since, direct CO_2 emission from land use change (LUC) alone contributes ~10% of total anthropogenic emission (Le Quere et al. 2016); it is one of the most important human-driven anthropogenic sources of atmospheric CO_2 (IPCC 2014). Anthropogenic activities such as burning of fossil fuels, land-use changes, and forestry activities are accelerating the rate of increase in atmospheric CO_2 concentration resulting in global warming and climate change during the recent times (Brahma et al. 2018). With a view to recognize the importance of agricultural soil in mitigating the greenhouse effect, role of managed agro-ecosystems in soil carbon sink management has been prioritized in COP 21 (Le Foll 2015). Increase in the concentration of greenhouse gases in the atmosphere and its hostile effects associated with climate change have increased the need for identification of systems with high carbon sink as a mitigation strategy. Tree-based systems such as farm forestry or agroforestry systems have the potential to sequester carbon in a short period of time (Nath et al. 2018a).

Recently, the carbon cycle has become an important global issue, and plants serve an important function in carbon storage. Numerous studies have been made on the role of woody trees species in carbon sequestration, but our knowledge of the potential of bamboos in biomass production and terrestrial carbon sequestration is very limited (Ly et al. 2012, Nath et al. 2015, 2018b, Thokcham and Yadav 29015, Yuen et al. 2017) as only a few

Archives of Virology
https://doi.org/10.1007/s00705-020-04533-6

ORIGINAL ARTICLE

Codon usage pattern and its influencing factors in different genomes of hepadnaviruses

Bornali Deb[1] · Arif Uddin[2] · Supriyo Chakraborty[1]

Received: 31 August 2019 / Accepted: 7 December 2019

Abstract
Codon usage bias (CUB) arises from the preference for a codon over codons for the same amino acid. The major factors contributing to CUB are evolutionary forces, compositional properties, gene expression, and protein properties. The present analysis was performed to investigate the compositional properties and the extent of CUB across the genomes of members of the family *Hepadnaviridae*, as previously no work using bioinformatic tools has been reported. The viral genes were found to be AT rich with low CUB. Analysis of relative synonymous codon usage (RSCU) was used to identify overrepresented and underrepresented codons for each amino acid. Correlation analysis of overall nucleotide composition and its composition at the third codon position suggested that mutation pressure might influence the CUB. A highly significant correlation was observed between GC12 and GC3 ($r = 0.910$, $p < 0.01$), indicating that directional mutation affected all three codon positions across the genome. Translational selection (P2) and mutational responsive index (MRI) values of genes suggested that mutation plays a more important role than translational selection in members of the family *Hepadnaviridae*.

Introduction

Amino acids are the building blocks of proteins, and the specific amino acids incorporated are determined by the genetic code. In the standard genetic code, a set of 61 codons encodes the 20 standard amino acids. Other than tryptophan and methionine, all amino acids are represented by more than one codon, resulting in codon redundancy. The condition of biased usage of some codons preferentially over other synonymous codons is known as codon usage bias (CUB), and it is specific for every genome [3, 29, 30, 53]. CUB differs among genomes as well as within the same genome, and studying these differences may help us to understand genome evolution among related species [66] as well as the relationship between host cells and viruses or immune reactions [62].

Various hypotheses have been proposed to explain the occurrence of CUB. In the neutral theory, mutational pressure at degenerate positions of a codon must be neutral, such that there is nonuniform usage of synonymous codons for a specific amino acid, indicating a lack of natural selection [48]. The level of gene expression has been shown to be associated with CUB [63, 64], whereas the selection-mutation-drift model postulates the importance of genetic drift, mutation pressure, and natural selection in the establishment of CUB [63, 64]. Natural selection of highly expressed genes can play an important role [31], and influences codon usage in various organisms [25]. Other notable determinants of innate CUB are base composition [5], skewness of bases [10], expression level of the gene [77], gene length [17], gene stability, replication [25, 39], translational selection [56], protein secondary structure [77] and hydrophobicity [13]. A previous study showed that variation in the tRNA pool and disparity in isochores of a cell are major determinants of CUB [4, 16].

Mutation is a major factor determining in configuring codon usage patterns in various viral genomes [52]. Investigation of constraints in codon usage provides information about molecular evolution of viruses and regulation of gene

Handling Editor: Carolina Scagnolari.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00705-020-04533-6) contains supplementary material, which is available to authorized users.

✉ Supriyo Chakraborty
supriyoch_2008@rediffmail.com

[1] Department of Biotechnology, Assam University, Silchar 788150, Assam, India

[2] Department of Zoology, Moinul Hoque Choudhury Memorial Science College, Algapur, Hailakandi 788150, Assam, India

Published online: 08 February 2020

Received: 22 April 2019 | Revised: 5 September 2019 | Accepted: 21 September 2019

DOI: 10.1002/mc.23124

Molecular Carcinogenesis WILEY

RESEARCH ARTICLE

Insights into the nucleotide composition and codon usage pattern of human tumor suppressor genes

Tarikul Huda Mazumder[1] | Arif Uddin[2] | Supriyo Chakraborty[1]

[1]Department of Biotechnology, Assam University, Silchar, Assam, India

[2]Department of Zoology, Moinul Hoque Choudhury Memorial Science College, Hailakandi, Assam, India

Correspondence
Arif Uddin, Department of Zoology, Moinul Hoque Choudhury Memorial Science College, Hailakandi 788150, Assam, India.
Email: arif.uddin29@gmail.com

Supriyo Chakraborty, Department of Biotechnology, Assam University, Silchar 788011, Assam, India.
Email: supriyoch_2008@rediffmail.com

Abstract

Tumor suppressor genes encode different proteins that inhibit the uncontrolled proliferation of cell growth and tumor development. To acquire clues for predicting gene expression level, it is essential to understand the codon usage bias (CUB) of genes to characterize genome which possesses its own compositional characteristics and unique coding sequences. We used bioinformatic tools to analyze the codon usage patterns of 637 human tumor suppressor genes as no work was reported earlier. The mean effective number of codons of these genes was 48, indicating low CUB. Our results exhibited a significant positive correlation among different nucleotide compositions and the codons ending with C base was most frequently used along with the most over-represented codon CTG and GTG codifying leucine and valine amino acid, respectively, in human tumor suppressor genes. The neutrality plot showed a significant positive correlation (Pearson, $r = 0.646$; $P < .01$) suggesting that mutation on GC bias might affect the CUB. However, the linear regression coefficient of GC_{12} on GC_3 in human tumor suppressor genes suggested that natural selection played a major role while mutation pressure played a minor role in the codon usage patterns of tumor suppressor genes in human. Our study would throw light into the factors that affect CUB and the codon usage patterns in the human tumor suppressor genes.

KEYWORDS

codon adaptation index, effective number of codons, relative synonymous codon usage, tumor suppressor gene expression

1 | INTRODUCTION

Tumor suppressor genes function as care takers of cells and maintain the coherence of the genome by mending the DNA damage. They avert uncontrolled cell division by repairing DNA damage and sometimes induce programmed cell death if the damaged DNA cannot be

acids in protein sequences. The degenerate property of genetic code ensures that more than one codon usually called synonymous codons encodes a particular amino acid except methionine and tryptophan. This redundancy in the genetic code might have evolved to increase the resistance of genes to mutation as the third codon position (wobble base) often changes but does not eventually affect the amino acid sequence in

Computational Biology and Chemistry 83 (2019) 107119

Contents lists available at ScienceDirect

Computational Biology and Chemistry

journal homepage: www.elsevier.com/locate/cbac

Research Article

Compositional features and codon usage pattern of *TP63* gene

Supriyo Chakraborty[a,*], Parvin A Barbhuiya[a], Gulshana A Mazumder[a], Bornali Deb[a], Arif Uddin[b,*]

[a] Department of Biotechnology, Assam University, Silchar, 788011, Assam, India
[b] Department of Zoology, Moinul Hoque Choudhury Memorial Science College, Algapur, Hailakandi, 788150, Assam, India

ARTICLE INFO

Keywords:
Codon usage bias
TP63 gene
Compositional features
Mutation pressure

ABSTRACT

The tumor protein p63encoded by the gene *TP63* acts as a homologue of p53 protein. *TP63* gene is the transformation factor with two initiation sites for transcriptional process and is related with stress, signal transduction and cell cycle control. The biasness in the preference of a few codons more frequently over other synonymous codons is the codon usage bias (CUB). Natural selection and mutational pressure are the two prime evolutionary forces acting on CUB. Here, the bioinformatic based analysis was performed to investigate the base distribution and CUB of *TP63*transcript variants (isoforms) as no work was performed earlier. Analysis of compositional features revealed variation in base content across *TP63* gene isoforms and the GC content was more than 50%, indicating GC richness of its isoforms. The mean effective number of codons (ENC), a measure of CUB, was 51.83, *i.e.* overall *CUB* of *TP63* gene was low. Among 13 isoforms of *TP63* gene, nature selected against the *CTA* codon in 8 isoforms and favored five over-represented (RSCU > 1.6) codons namely CTG, CAG, ATC, AAC and GCC during evolution. Correlation between overall nucleotide composition and its 3rd codon position revealed that both mutational pressure and natural selection moulded its CUB. Further, the correlation between ENC and aromaticity depicted that variation of CUB was related to the degree of aromaticity of p63 protein.

1. Introduction

The amino acid sequence in a protein is defined by the genetic code. As many as 61 codons encode just twenty amino acids and 3 codons act as termination signals in a growing polypeptide chain on the ribosomes. This reveals the degenerate property of the genetic code where in about two to six codons usually encode the same amino acid except two amino acids *i.e.* methionine and tryptophan. Such a bunch of codons codifying the same amino acid is named as synonymous codons. Very often, the usage of the synonymous codons for an amino acid in the mRNA transcripts is highly unequal, a trend noticed across every organism, and termed as codon usage bias or CUB (Grantham et al., 1981). In all domains of life, the genetic code is conserved none the less the direction of codon bias differs from organism to organism. The degree of CUB is by no means the same among genomes and genes (Hershberg and Petrov, 2008).

Many features of coding sequences namely gene length, GC content and the properties of encoded proteins like hydropathicity and aromaticity are linked to CUB (Bains, 1987; Bernardi and Bernardi, 1986; Lobry and Gautier, 1994; Lynn et al., 2002). Mutational biases and natural selection are the major evolutionary factors influencing the bias in codon utilization (Bulmer, 1991).

In the beginning, it was believed that synonymous mutations occurring in coding sequences have no effect as these do not alter the amino acid sequence in a protein and are referred to as "silent" mutations. However, further research revealed that CUB is associated with many cellular processes and might even affect human ailments (Bali and Bebok, 2015). CUB does have an effect on translation and mRNA degradation. Few researchers have found the link of synonymous mutation with amyotrophic lateral sclerosis, cystic fibrosis and Crohn's disease (Bali and Bebok, 2015; Bartoszewski et al., 2010; Lazrak et al., 2013). Studies on MDR1 gene (Multidrug Resistance1) have shown that synonymous mutation changes the substrate specificity resulting in a modified structure and function of the protein (Kimchi-Sarfaty et al., 2007). Similarly, there exists a few other good illustrations that show synonymous mutation does affect the protein function to a certain extent (Carpen et al., 2006; Matsuo et al., 2007). Further, about two decades ago in two bacterial studies (Komar AA et al., 1998; Komar Anton A et al., 1999), it was shown that synonymous codon substitution alters the translational efficiency of mRNA molecules through different cellular mechanisms. CUB also modulates the proper folding required for a particular protein to function effectively in the cell (Yu et al., 2015).

There are two evolutionary explanations for bias in codon

* Corresponding authors.
E-mail addresses: supriyoch_2008@rediffmail.com (S. Chakraborty), arif.uddin29@gmail.com (A. Uddin).

https://doi.org/10.1016/j.compbiolchem.2019.107119
Received 5 September 2017; Received in revised form 24 August 2019; Accepted 27 August 2019
Available online 28 August 2019
1476-9271/

Virus Research 264 (2019) 129–138

Contents lists available at ScienceDirect

Virus Research

journal homepage: www.elsevier.com/locate/virusres

Analysis of codon usage patterns and influencing factors in Nipah virus

Supriyo Chakraborty[a,*], Bornali Deb[a], Parvin A. Barbhuiya[a], Arif Uddin[b]

[a] Department of Biotechnology, Assam University, Silchar 788011, Assam, India
[b] Department of Zoology, Moinul Hoque Choudhury Memorial Science College, Algapur, Hailakandi 788150, Assam, India

ARTICLE INFO

Keywords:
Codon usage bias
Nipah virus
Natural selection
Mutation pressure

ABSTRACT

Codon usage bias (CUB) is the unequal usage of synonymous codons of an amino acid in which some codons are used more often than others and is widely used in understanding molecular biology, genetics, and functional regulation of gene expression. Nipah virus (NiV) is an emerging zoonotic paramyxovirus that causes fatal disease in both humans and animals. NiV was first identified during an outbreak of a disease in Malaysia in 1998 and then occurred periodically since 2001 in India, Bangladesh, and the Philippines. We used bioinformatics tools to analyze the codon usage patterns in a genome-wide manner among 11 genomes of NiV as no work was reported yet. The compositional properties revealed that the overall GC and AT contents were 41.96 and 58.04%, respectively *i.e.* Nipah virus genes were AT-rich. Correlation analysis between overall nucleotide composition and its 3rd codon position suggested that both mutation pressure and natural selection might influence the CUB across Nipah genomes. Neutrality plot revealed natural selection might have played a major role while mutation pressure had a minor role in shaping the codon usage bias in NiV genomes.

1. Introduction

Degeneracy or redundancy of the genetic code ensures that multiple codons codify the same amino acid except for two amino acids *i.e.* methionine and tryptophan. The codons that encode the same amino acid are called synonymous codons. Numerous previous studies have shown that the usage of these synonymous codons in mRNA molecules in varying frequencies leads to a phenomenon known as codon usage bias (CUB) (Hasegawa et al., 1979). The evolution of CUB is very complex and a highly debatable subject. Various evolutionary processes explain the origin of synonymous codon usage variation or CUB, among them the two most accepted theories are the neutral theory and the selection-mutation drift balance model theory (Duret and Mouchiroud, 1999; Sharp et al., 1986, 1993; Shields and Sharp, 1987). However, the impact of these evolutionary forces in different species remains undetermined (Akashi, 1997; Hershberg and Petrov, 2008). In addition, various biological factors have been identified to be associated with CUB such as GC composition, gene expression level, gene length, protein structure, tRNA abundance and its types, hydrophobicity and hydrophilicity of the protein (Bains, 1987; Bernardi and Bernardi, 1986; Gouy and Gautier, 1982; Ikemura, 1981; Tao and Dafu, 1998).

The relationships of codon usage between viruses and their hosts are fascinating as it has significance in overall viral existence, its codon adaptation to host, evasion of host's immune system by viral pathogen and their co-evolution.

CUB can provide significant insights relating to functional regulation of gene expression level, identification of horizontally transferred genes, optimization of protein expression level and adaptation of pathogens to certain specific hosts (Chaney and Clark, 2015; Lithwick and Margalit, 2005; Liu et al., 2012).

Nipah virus (NiV), an emerging zoonotic paramyxovirus, possesses high pathogenicity that causes fatal disease in both animals and humans (Wong et al., 2002). NiV is a single stranded RNA virus which belongs to genus *Henipavirus*, within the family *Paramyxoviridae* (Chua et al., 2000). Genome size varies from 18246 to 18252 nucleotides and the number of genes varies from 6 to 9 (http://www.ncbi.nlm.nih.gov). NiV was first identified during an outbreak of a disease that took place in Malaysia in 1998 (Lee et al., 1999). Outbreaks of NiV have occurred periodically since 2001 in India, Bangladesh, and the Philippines (Arankalle et al., 2011; Ching et al., 2015; Hossain et al., 2008; Hsu et al., 2004). In the mid of 2018, Nipah outbreak was again reported in Southern parts of India. The natural hosts of the virus are fruit bats of *Pteropodidae* family (Olson et al., 2002). It is transmitted through contact with NiV infected animal causing a prominent risk for epidemic outbreak or through consumption of contaminated foods. Human-to-human transmission has also been observed (Clayton, 2017; Escaffre et al., 2013). In the outbreaks of Malaysia and Singapore, pigs were reported to be the intermediate hosts, whereas in Bangladesh it was the

* Corresponding author.
E-mail address: supriyoch_2008@rediffmail.com (S. Chakraborty).

https://doi.org/10.1016/j.virusres.2019.01.011
Received 14 December 2018; Received in revised form 18 January 2019; Accepted 18 January 2019
Available online 18 January 2019

Received: 10 September 2018 | Revised: 10 March 2019 | Accepted: 20 March 2019

DOI: 10.1002/jez.b.22852

RESEARCH ARTICLE

WILEY JEZ-B MOLECULAR AND DEVELOPMENTAL EVOLUTION

Genome-wide comparison of codon usage dynamics in mitochondrial genes across different species of amphibian genus *Bombina*

Parvin A. Barbhuiya[1] | Arif Uddin[2] | Supriyo Chakraborty[1]

[1]Department of Biotechnology, Assam University, Silchar, Assam, India

[2]Department of Zoology, Moinul Hoque Choudhury Memorial Science College, Hailakandi, Assam, India

Abstract

The biological phenomenon where some synonymous codons are repeatedly preferred to others in gene transcripts is called codon usage bias (CUB). The analysis

Received: 17 November 2018 | Revised: 22 January 2019 | Accepted: 24 January 2019

DOI: 10.1002/jcp.28350

REVIEW ARTICLE

WILEY Journal of Cellular Physiology

A review on coronary artery disease, its risk factors, and therapeutics

Arup Kr. Malakar[1*] | Debashree Choudhury[1*] | Binata Halder[1*] | Prosenjit Paul[1*] | Arif Uddin[2] | Supriyo Chakraborty[1]

[1]Department of Biotechnology, Assam University, Silchar, Assam, India

[2]Department of Zoology, Moinul Hoque Choudhury Memorial Science College, Hailakandi, Assam, India

Correspondence
Supriyo Chakraborty, Department of Biotechnology Assam University, Silchar 788011, Assam, India.
Email: supriyoch_2008@rediffmail.com

Abstract

Coronary artery disease (CAD) is one of the major cardiovascular diseases affecting the global human population. This disease has been proved to be the major cause of death in both the developed and developing countries. Lifestyle, environmental factors, and genetic factors pose as risk factors for the development of cardiovascular disease. The prevalence of risk factors among healthy individuals elucidates the probable occurrence of CAD in near future. Genome-wide association studies have suggested the association of chromosome 9p21.3 in the premature onset of CAD. The risk factors of CAD include diabetes mellitus, hypertension, smoking, hyperlipidemia, obesity, homocystinuria, and psychosocial stress. The eradication and management of CAD has been established through extensive studies and trials. Antiplatelet agents, nitrates, β-blockers, calcium antagonists, and ranolazine are some of the few therapeutic agents used for the relief of symptomatic angina associated with CAD.

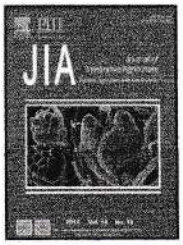
KEYWORDS

cardiovascular disease, coronary artery disease, low-density lipoproteins

Journal of Integrative Agriculture 2019, 18(12): 2771–2785

Available online at www.sciencedirect.com

ScienceDirect



RESEARCH ARTICLE

Deciphering codon usage patterns and evolutionary forces in chloroplast genes of *Camellia sinensis* var. *assamica* and *Camellia sinensis* var. *sinensis* in comparison to *Camellia pubicosta*

Sophiarani Yengkhom[1], Arif Uddin[2], Supriyo Chakraborty[1]

[1] *Department of Biotechnology, Assam University, Silchar 788011, Assam, India*
[2] *Department of Zoology, Moinul Hoque Choudhury Memorial Science College, Algapur, Hailakandi 788150, Assam, India*

Abstract

Codon usage bias (CUB) is a unique property of genome which refers to non-random usage of synonymous codons in coding sequences. The present study makes an attempt to find out the pattern of CUB in chloroplast (cp) genes among three tea species, i.e., *Camellia sinensis* var. *assamica* (Assam tea), *Camellia sinensis* var. *sinensis* (Chinese tea) and *Camellia pubicosta* (wild tea species) as no work on CUB was reported earlier. To understand the patterns of codon usage among the cp genes of three tea groups, we used bioinformatic tools to investigate the protein coding sequences of cp genes. In our present study, the mean nucleobase T was the highest whereas C was the lowest in all the three tea groups. The overall AT content was more than GC content, i.e., genes were AT rich. The scaled chi-square (SCS) value indicated that the CUB of cp genes was low. The codon CGT (Arg) was over-represented in *C. sinensis* var. *sinensis* whereas GGA (Pro) was over-represented in *C. pubicosta* species. Heatmap study revealed that most of the GC ending codons showed positive correlations between codon usage and GC3 while AT ending codons exhibited negative correlations. From neutrality plot analysis, it was evident that natural selection had played a major role, while mutation pressure exerted a minor effect in the CUB of cp genes in three tea groups. Highly significant (P<0.01) positive correlation was found between SCS and synonymous codon usage order (SCUO) of cp genes which suggested that high expression of cp genes was associated with high degree of CUB.

Keywords: *Camellia sinensis* var. *assamica*, *Camellia sinensis* var. *sinensis*, *Camellia pubicosta*, chloroplast genes, codon usage bias, mutation, natural selection

1. Introduction

In genetic code, more than one codon often encodes the same amino acid in a protein known as synonymous codons for a particular amino acid. Except two amino acids methionine and tryptophan, all other amino acids are encoded by 2–6 different codons in coding sequences of genes. Codon usage bias (CUB), a unique genomic feature, refers to non-random and unequal usage of synonymous

Received 14 January, 2019 Accepted 15 April, 2019
Correspondence Supriyo Chakraborty, E-mail: supriyoch_2008@rediffmail.com

doi: 10.1016/S2095-3119(19)62716-4

Genetica (2019) 147:231–248
https://doi.org/10.1007/s10709-019-00067-1

ORIGINAL PAPER

Compositional properties and codon usage pattern of mitochondrial ATP gene in different classes of Arthropoda

Riazul Islam Barbhuiya[1] · Arif Uddin[2] · Supriyo Chakraborty[1]

Received: 17 November 2018 / Accepted: 22 May 2019 / Published online: 31 May 2019

Abstract
Codon usage bias (CUB) is defined as the usage of synonymous codons unequally for an amino acid in a gene transcript. It is influenced by both mutation pressure and natural selection and is a species-specific property. In our current study, we used bioinformatic methods to investigate the coding sequences of mitochondrial adenosine triphosphate gene (MT-ATP) in different classes of arthropoda to know the codon usage pattern of the gene as no work was described earlier. The analysis of compositional properties suggested that the gene is AT rich. The effective number of codons revealed the CUB of both ATP6 and ATP8 gene was moderate. Heat map showed that the codons ending with AT were negatively associated with GC3 while the codons ending with GC were positively associated with GC3 in all the classes of arthropoda. Correspondence study revealed that the pattern of codon usage of ATP6 and ATP8 genes differed across classes. Neutrality plot suggested the codon usage bias of these two genes in phylum arthropoda was influenced by both mutation pressure and natural selection.

Keywords Codon usage bias · Mutation pressure · Natural selection · Arthropoda

Introduction

In a genetic code more than one codon often encodes a single amino acid and these codons are termed as synonymous codons for the amino acid except two amino acids in standard code namely methionine and tryptophan (Baker et al. 1970). The mitochondrial DNA of arthropoda follows the translation Table 5 of NCBI. In this code, out of 64 codons, 62 codons encode 20 amino acids and the remaining two codons i.e. TAA and TAG act as termination signal. The amino acid serine is encoded by eight codons whereas both methionine and tryptophan amino acids are encoded by two codons each in this code unlike in the standard genetic code.

The usage frequency of synonymous codons is not equal in coding sequences and this unequal frequency of codon usage leads to codon usage bias or CUB (Behura and Severson 2012). The unfairness in usage of codons confirms that the optimal codons, codons that are used repeatedly, can couple with the anticodons of the largely available tRNA genes (Sun et al. 2009). And it also tends to decrease the errors in processing and thereby reduces the chances of wrong inclusion of amino acids in a growing polypeptide chain. It ensures a balancing situation amongst the effects of natural selection (includes selection for translation, length and function of gene), mutation pressure (includes GC content and position of mutation in base) and random genetic drift (Bulmer 1991; Sharp and Li 1986a). Codon bias is a unique feature of the genome of an organism and among the genes of the same organism; it may vary significantly (Grantham et al. 1980; Marin et al. 1989; Prat et al. 2009a). It is well obvious that the variation in codon usage within the same family of synonymous codon governs the translational efficiency of a gene, thus having a strong impact in framing the genome evolution (Bentele et al. 2013). The up surged translational efficiency assists organisms to fit into the dynamic environmental conditions and is often correlated with their lifestyle (Botzman and Margalit 2011). Previously,

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10709-019-00067-1) contains supplementary material, which is available to authorized users.

✉ Supriyo Chakraborty
supriyoch_2008@rediffmail.com

[1] Department of Biotechnology, Assam University, Silchar, Assam 788011, India

[2] Department of Zoology, Moinul Hoque Choudhury Science College, Algapur, Hailakandi, Assam 788150, India

MITOCHONDRIAL DNA PART A
https://doi.org/10.1080/24701394.2019.1616701

Taylor & Francis
Taylor & Francis Group

RESEARCH ARTICLE

Mutation pressure and natural selection on codon usage in chloroplast genes of two species in *Pisum* L. (Fabaceae: Faboideae)

Debjyoti Bhattacharyya[a], Arif Uddin[b], Sudipa Das[a] and Supriyo Chakraborty[c]

[a]Department of Life Science and Bioinformatics, Assam University, Silchar, India; [b]Department of Zoology, Moinul Hoque Choudhury Memorial Science College, Algapur, India; [c]Department of Biotechnology, Assam University, Silchar, India

ABSTRACT

This study was attempted to focus on the pattern of codon usage bias (CUB) of chloroplast genes in two species of *Pisum viz. P. fulvum* and *P. sativum* and to identify the factors which influence CUB. Bioinformatic tools were used to understand codon usage pattern in the protein-coding sequences of *Pisum* chloroplast genomes. It was found that GC content was lower than AT content in the genes. Low synonymous codon usage order (SCUO) values of genes indicated low CUB in chloroplast genes of *Pisum* species. Heatmaps showed positive correlations of GC3 with all the GC and AT ending codons. Neutrality plot analysis revealed that natural selection might have played a prominent role over mutation pressure in sculpturing the CUB of chloroplast genes in these two taxa. Positive correlation between SCUO and mRNA free energy (mFE) suggested that higher energy release by entire mRNA was related to high degree of CUB. Further, highly significant ($p < .01$) negative correlation was found between parameters in pair i.e. mFE-GC, mFE-GC1, mFE-GC2 and mFE for entire mRNA-GC3. This pointed out that higher GC content might have influenced lesser energy release by mRNA molecules of chloroplast genes.

ARTICLE HISTORY
Received 10 February 2019
Accepted 3 May 2019

KEYWORDS
Codon usage bias; chloroplast gene; *Pisum fulvum*; *Pisum sativum*; Fabaceae

Introduction

The genus *Pisum* L., commonly known as pea, is a temperate crop native to south-western Asia. The genus belongs to the tribe Fabeae Rchb. under the subfamily Faboideae of Fabaceae family (Leguminosae). Globally, the genus *Pisum* is represented by three species *viz. P. abyssinicum* A. Braun, *P. fulvum* Sibth. & Sm. and *P. sativum* L. (Roskov and Zarucchi 2018). Notably, *P. sativum* has two subspecies, *P. sativum* subsp. *sativum* (domesticated pea) and *P. sativum* subsp. *elatius* (M. Bieb.) Asch. & Graebn. (Wild form) (Smýkal et al. 2017). However, the accurate number of species in the genus is always a debatable issue to the taxonomists and in a state of flux. In the Near East domestication region, pea is one of the important primeval cultivated crops (Smýkal et al. 2017), which is restricted to the Mediterranean basin and Western Asia. Wild pea is mostly distributed in northern part of Fertile Crescent (Smýkal et al. 2018). Later, its distribution extended

pod, often green or rarely golden or purple, bearing 2–10 spheroidal seeds.

Pisum abyssinicum is known as the Abyssinian Pea. The plants are annual, herbaceous climbers. The plant is native to Africa (Ethiopia) and Asia Temperate (Arabian Peninsula). The seeds are eaten as pulses. The plants are also used as fodder and forage.

Pisum fulvum, the Tawny Pea, is characterized by annual climbing habits bearing orange flowers. The plants flower from February to April with a peak in blooming in late March. *P. sativum* is commonly known as the garden pea, common pea or field pea. The plants produce white, pink or purple flowers from May to September. This is an old world species native to Asia and Europe. Later, it was cultivated worldwide as an important vegetable crop. *P. sativum* subspecies *elatius*, commonly known as wild pea is an annual growing up to 2 m. Although the seeds provide good source

International Journal of Research in Sociology and Anthropology (IJRSA)
Volume 2, Issue 2, 2016, PP 1-9
ISSN 2454-8677
http://dx.doi.org/10.20431/2454-8677.0202001
www.arcjournals.org

Haemoglobin Variants among the Bengali Muslims and the Meiteis in two Villages of Cachar District of Assam, India

A. F. Gulenur Islam Barbhuiya[1], Rekha Das[2] and Manwendra Singh Bartwal[3]

[1]Assistant Professor, Department of Anthropology, M. H. C. M. Science College, Hailakandi, Assam, India
[2]Professor, Department of Anthropology, Gauhati University, Guwahati, Assam, India
[3]Post Doctoral Research Fellow, AnSI, NWRC, Dehradun, Uttarakhand, India
[1]*connect2gulbarbhuiya@gmail.com*

Abstract: *Present paper documents the presence of haemoglobin variants along with haematogical parameters among the Bengali Muslims of Bhaurikandi Part-II and the Meiteis of Dakshin Mohanpur Part-V village of Cachar District of Assam. Capillary electrophoresis has been performed to detect the haemoglobin variants on 27 Bengali Muslims and 26 Meiteis from whom 2 ml intravenous blood was collected in EDTA vials. Haematological parameters has been analysed by using haematology analyzer. The results reveal the presence of haemoglobin variants like HbE and β-thalassemia among the Muslim population with a gene frequency of 0.0556 each. On other hand, only HbE trait is found to be present among the Meitei population with a gene frequency of 0.0962. Haemoglobin concentration level (Hb in g/dl) is found to be lower in haemoglobin variants in contrast to normal haemoglobins (HbAA) in both the communities.*

Keywords: *HbE, β-thalassemia, Haematology, Bengali Muslim, Meitei, Cachar, etc.*

1. INTRODUCTION

Haemoglobin (Hb), the red respiratory protein found in mammalian erythrocytes is one of the most informative molecules in primate blood [1]. In adult humans, the most common haemoglobin type is a tetramer known as haemoglobin-A consists of four separate polypeptide chains of amino acids and each subunit is composed of a protein chain tightly associated with a non-protein haeme group that fix oxygen. Mutations in the genes for the haemoglobin result in haemoglobin variants, some of which cause a group of hereditary disorders in humans termed as haemoglobinopathies. Haemoglobin variants are a part of the normal embryonic and fetal development which can also be pathologic mutant forms caused by variations in genetics of haemoglobin in a population. Haemoglobinopathies covers a group of hereditary disorders in which the structure (qualitative change) or the rate of synthesis (quantitative change) of one of the normal haemoglobin chain is altered [1]. Haemoglobinopathies causes varying degrees of microcytic anemia that can range from insignificant to life threatening [2]. Although haemoglobinopathy is a hereditary disorder but environment potentially can determine the genotypes and it does it through natural selection.

Haemoglobinopathies are the most commonly encountered hereditary abnormalities of blood posing a major genetic burden and public health problem in Southeast Asia and the Indian subcontinent [3]. HbS, HbE and HbD as well as thalassemia are the most widely distributed haemoglobin disorders in India out of which HbE is widely distributed in north-eastern states of India [4, 5, 6]. Thalassemia results from the reduced rate of synthesis one of the globin chains can cause the formation of abnormal haemoglobin molecules which in turn causes the anemia. The sickle cell haemoglobin (HbS) is structurally abnormal variant results in the substitution of the amino acid valine for glutamic acid at 6th position of β-globin polypeptide chain of the molecule. Haemoglobin E is another structurally abnormal variant with glutamic acid residue replaced by lysine at the 26th position of the β-globin polypeptide chain [7].

India is a land of different endogamous communities and non random (consanguinity) mating pattern leads to the co-heritance of β-thalassemia and other structural variants (D, E, S). Interaction of these structural variants of haemoglobin along with the reduced synthesis of globin chains result to the combination of two abnormalities, resulting in double heterozygosity of the disease in India [8]. Carriers of haemoglobinopathies are partially protected against morbidity and mortality of falciparum malaria. The resistance of HbAE red cells to Plasmodium falciparum is most likely the cause for its

IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-6, Issue-7, July 2018 E-ISSN: 2347-2693

Genetic Algorithm-Based Neural Network for Estimation of Scour Depth Around Bridge Abutment

A.K. Md. Fujail[1*], S.A. Begum[2], A.K. Barbhuiya[3]

[1*] Department of Computer Science, MHCM Science College, Hailakandi, Assam, India
[2] Department of Computer Science, Assam University, Silchar, Assam, India
[3] Department of Civil Engineering, National Institute of Technology, Silchar, Assam, India

*Corresponding Author: abul_fujail@yahoo.com, Tel.: +91-9401534700

Available online at: www.ijcseonline.org

Accepted: 19/Jul/2018, Published: 31/Jul/2018

Abstract — Scour depth at abutment is a major cause of bridge failure and significant issue towards maintenance cost of a bridge. Thus, early estimation of scour depth at abutment is essential for safe and cost-effective abutment structure design. Extensive research has been carried out to develop methods for predicting the depth of abutment scour. Despite various models presented by researchers to estimate the equilibrium local scour depth, an efficient technique with enhanced estimation capability will be more beneficial. The paper is aimed at investigating the applicability of soft computing (SC) models *viz.* artificial neural network, gene-expression programming (GEP) and hybrid techniques for estimation of scour depth around vertical, semi-circular and 45° wing-wall abutments using laboratory data compiled from published literature. The paper also emphasizes on further enhancement of the performances of the SC based models. On experimentations, the performance of multilayer perceptron (MLP) neural network for each type of abutment was found more effective than radial basis function network, GEP model and empirical equations. The generalization performance of optimal MLP network developed for each type of abutment was then improved with evolving connection weights of the MLP by Genetic Algorithm (GA-MLP). Finally, the hybrid model is validated with different types of validation techniques. The study demonstrates the suitability of the SC based hybrid methodology in improving the predictive accuracy of scour depth around different types of abutments.

Keywords— Scour depth, artificial neural network, genetic algorithm, hybrid technique, GEP

I. INTRODUCTION

Scour around bridge foundation takes place due to the erosion of soil by water stream [1]. Excessive scour may encounter huge maintenance costs or collapses of bridge. Therefore, accurate estimation of the maximum scour depth around abutment and pier is necessary for cost-effective design of bridge foundation. According to a survey report, repairing and maintenance of bridge damage required 50% of total expenditure, out of which 70% was spent to repair abutment scour [2]. Thus, it is essential to estimate reasonably accurate maximum scour depth at bridge abutment for safe and economic design of abutment foundation.

In the recent past, experimental investigation has been conducted and various empirical formulae [3-6] have been developed to predict clear water scour depth around abutments. Each of the developed formulae is suitable to a specific abutment condition and the results of each formula highly differ with each other for the same dataset. Thus, the estimated scour depths using empirical formulae are not reliable due to underestimation or overestimation which may cause bridge failure or increases construction cost.

To enhance the predictive accuracy, data-driven modeling tools based on soft computing (SC) techniques such as artificial neural networks (ANNs) [7-12], adaptive neuro-fuzzy inference system (ANFIS) [13-14], genetic programming (GP) [15-17] and gene expression programming (GEP) [18-20] have been recently employed to estimate scour depth around different types of hydraulic structures. The estimated results with SC methods have been reported to be effectively outperformed the empirical equations.

This paper presents a comparative analysis between MLP, radial basis function (RBF) neural network, GEP and empirical models for predicting scour depth around abutment using the same data set collected from different published literature. The main objective of this study is to enhance the performance of the available SC-based techniques in predicting scour depth around bridge abutment by

BANIJYA

ISSN: 0975-0010

A Peer Reviewed Research Journal of the Department of Commerce, Assam University, Silchar
UGC Approved Journal No. 63678

Vol. 10 No. 1 March, 2018

Dr A. S. Yarso : Managing Editor, Banijya

DEPARTMENT OF COMMERCE
ASSAM UNIVERSITY, SILCHAR
CACHAR, ASSAM, INDIA, PIN-788011

CONTENTS

UGC Approved Journal No. 63678

Brand Building Strategies for Fruit Juice : Some Issues

*Jarita Das
**Parag Shil

Abstract

Brand name is generally a first and foremost criterion for selection of a product. Use of branded products is in practice since antiquity. Brand is an important aspect of selecting a product from a varied number of options and thus has a vital role for the products' success in the market. Like brands of other products, fruit juice brands, viz., Tropicona and Real have emerged into the market that produce various fruit juice as a supplement of the fruits. The present study focuses on the various brand building strategies essential for success of fruit juice brands.

Keywords: Brand, Fruit juice, soft drinks, brand building

1. Introduction

A brand is a "name, term, design, symbol, or any other feature that identifies one seller's good or service as distinct from those of other sellers" (American Marketing Association, 1960). Top brands are constantly engaged in research and development services so that they could offer superior products than their competitors. This is a continuous process through which they could generate a good faith in their consumer and in the long run their brands are regarded as status symbol. Brand creates an impression on the consumer which is referred to as business capital. Thus, brand is the trademark as well as an identity which is the backbone in the field concerned (Prasad and Dev, 2000).

A brand differentiates a product in tangibles and intangibles forms. These dimensions distinguish a brand from its competitors and provide it equity. Brand equity in marketing industry describes value of having its popularity as consumers believe that a product with well known name is better than products with less well known. Brand equity is the "added value endowed by the brand to the product" (Farquhar, 1989). The present scenario of market shows little differences between the allied companies. There exists a perfect competition among the brands of allied companies. In present day competitive market, every company has its own technological advancements so that they could sustain in the market which makes the product differentiation very difficult (Kotler, 2000).

Juice is a beverage extracted from fruit or vegetable. It is the liquid naturally contained in fruit or vegetable tissue. Juice is prepared mechanically squeezing or macerating fresh fruits or vegetable flesh, often without the application of heat or solvents. Juices, being mechanically produced, can also be preserved for a long period of time. The juices thus produced maintain the essential physical, chemical, organoleptical and nutritional characteristics of the fruit from which it is made (Boukraâ, 2013). This has led to further demand for fruit juices in the market. The fruit juice industry on the whole is encountering new opportunities and challenges globally. Changing consumer demands and

*Research Scholar, Department of Commerce, Assam University, Silchar: jaritadas@gmail.com
**Associate Professor, Department of Commerce, Assam University, Silchar: drps007.aus@gmail.com

04/07/23
Principal
M. H. C. M. Sc. College
Algapur, Hailakandi

Genomics 111 (2019) 167–176

Contents lists available at ScienceDirect

Genomics

journal homepage: www.elsevier.com/locate/ygeno

Compositional dynamics and codon usage pattern of *BRCA1* gene across nine mammalian species

Supriyo Chakraborty[a,*], Tarikul Huda Mazumder[a], Arif Uddin[b,*]

[a] Department of Biotechnology, Assam University, Silchar 788011, Assam, India
[b] Department of Zoology, Moinul Hoque Choudhury Memorial Science College, Algapur, Hailakandi 788150, Assam, India

ARTICLE INFO

Keywords:
Cancer
Breast
Molecular genetics
Codon usage bias

ABSTRACT

The *BRCA1* gene is located on the human chromosome 17q21.31 and plays important role in biological processes. The aminoacyl-tRNA synthetases (AARS) are a family of heterogenous enzymes responsible protein synthesis and whose secondary functions include a role in autoimmune myositis. Our findings reveal that the compositional constraint and the preference of more A/T –ending codons determine the codon usage patterns in *BRCA1* gene while more G/C-ending codons influence the codon usage pattern of AARS gene among mammals. The codon usage bias in *BRCA1* and *AARS* genes is low. The codon CGC encoding arginine amino acid and the codon TTA encoding leucine were uniformly distributed in *BRCA1* and *AARS* genes, respectively in mammals including human. Natural selection might have played a major role while mutation pressure might have played a minor role in shaping the codon usage pattern of *BRCA1* and AARS genes.

1. Introduction

Genetic code is degenerate meaning that more than one codon encodes the same amino acid. Unequal usage of synonymous codons for encoding the same amino acid during translation of a gene transcript into a protein is a well-established phenomenon commonly known as codon usage bias (CUB). It is species specific and significantly differs among the genes of the same taxa [3,12,31,35]. The codon usage patterns have been analyzed since the outstanding efforts for the creation of the first molecular sequence databases were initiated [12]. The result of Grantham and his co-workers demonstrated that species specific genes share similar patterns of synonymous codon usage frequency as stated by the "genome hypothesis" [11,12]. Therefore, scanning the codon usage patterns of all the genes in an organism may obscure the underlying heterogeneity [2] and hence it is better to identify the trends of codon usage patterns within the genes of a species or between closely related species. Various factors responsible for codon usage bias in different organisms from lower prokaryotes to higher eukaryotes have been discussed earlier by researchers across the globe but till date the codon usage patterns within the genes of an organism during the course of evolution have been interpreted for varied explanations. In general, researchers reported that the compositional constraints under mutation pressure or natural selection have been considered as the major factors involved in the codon usage variation among different organisms [8,20,26,48].

The *BRCA1* gene in human is located on the chromosome 17q21.31 and comprises of 24 exons and its coding region encodes a protein of 1863 amino acids [33]. Multiple functions of BRCA1 attributed to its tumor activity include progression of cell cycle, DNA damage repair process and regulation of specific set of pathways as well as germ line mutations in its sequence. The predisposition of these functions of *BRCA1* gene to breast and ovarian cancer in affected individuals [36] has been discussed earlier but the comparative analysis of synonymous codon usage influencing the codon bias in *BRCA1* gene among mammals with reference to human has not been done so far.

Housekeeping genes are typically constitutive genes that carry out the maintenance of basic cellular functions, and are expressed in all cells of an organism under normal and patho-physiological conditions [9,22]. The *AARS* gene encodes the enzyme alanyl-tRNA synthetase and catalyzes the binding of alanine amino acid to the appropriate tRNA. The aminoacyl-tRNA synthetases are a family of heterogenous enzymes responsible protein synthesis and their secondary functions include a role in autoimmune myositis [18].

In this study, an attempt has been made to analyze the codon bias and codon context patterns in the coding sequences of *BRCA1* and compared with one house keeping gene (*AARS*) having same length across mammals using the codon bias measures like effective number of codons (ENC), relative synonymous codon usage (RSCU) and relative abundance of dinucleotides. Further, in order to understand the extent of selection pressure acting on the protein coding *BRCA1* and *AARS*

* Corresponding author.
E-mail addresses: supriyoch_2008@rediffmail.com (S. Chakraborty), arif.uddin29@gmail.com (A. Uddin).

https://doi.org/10.1016/j.ygeno.2018.01.013
Received 1 September 2017; Received in revised form 22 December 2017; Accepted 22 January 2018
Available online 23 February 2018

Journal	MSP No.	Dispatch: August 1, 2017	CE:
JCP	26095	No. of Pages: 9	PE: Diane Punger

Received: 8 June 2017 | Accepted: 11 July 2017

DOI: 10.1002/jcp.26095

REVIEW ARTICLE

WILEY Cellular Physiology

miRNAs and ovarian cancer: An overview

Bornali Deb MSc[1] | Arif Uddin MSc, PhD[2] | Supriyo Chakraborty PhD, PDF[1]

[1] Department of Biotechnology, Assam University, Silchar, India

[2] Department of Zoology, Moinul Hoque Choudhury Memorial Science College, Algapur, Hailakandi, India

Correspondence
Supriyo Chakroborty, PhD, PDF, Department of Biotechnology, Assam University, Silchar, Assam, India.
Email: supriyo123chakraborty@yahoo.in

Ovarian cancer (OC) is the sixth most common cancer in women globally. However, even with the advances in detection and therapeutics it still represents the most dangerous gynecologic malignancy in women of the industrialized countries. The discovery of micro-RNAs (miRNA), a small noncoding RNA molecule targeting multiple mRNAs and regulation of gene expression by triggering translation repression and/or RNA degradation, has revealed the existence of a new array for regulation of genes involved in cancer. This review summarizes the current knowledge regarding the role of miRNAs expression in OC. It also provides information about potential clinical relevance of circulating miRNAs for OC diagnosis, prognosis, and therapeutics. The identification of functional targets for miRNAs represents a major obstacle in our understanding of microRNA function in OC, but significant progress is being made. The better understanding of the role of microRNA expression in ovarian cancer may provide new array for the detection, diagnosis, and therapy of the OC.

KEYWORDS
gynecologic malignancy, miRNA, ovarian cancer

1 | INTRODUCTION

Ovarian cancer is one of the most common gynaecologic malignancies in women worldwide, with each year of about 230,000 new cases and almost 140,000 death per annum (Dahiya & Morin, 2010). Ovarian cancers usually develop in granulosa theca cells or germ cells, moreover >90% of ovarian cancers have an epithelial histology and were assumed to have been arisen from cells that cover ovarian surface or from the line of subsurface of inclusion cysts (Feeley & Wells, 2001). The risk factors for ovarian cancer include family history, age, and persistent ovulation with attributed carriers of mutated *BRCA1* hereditarily (Nikitin & Corney, 2008). Inspite of advances in detection and cytotoxic therapies, only 30% of advanced stage cancer patients survive nearly 5 years after initial stage prognosis (Greenlee, Hill-Harmon, Murray, & Thun, 2001), while other 70% patients have high mortality rate due to late stage diagnosis (Iorio et al., 2007). Only, 19% of ovarian cancer patients are diagnosed at its early stage (Iorio et al., 2007). Routine diagnostic procedures, pelvic examination, serum *CA125*, and transvaginal ultrasonography usually fail to detect early stage of cancer and thus more death rates (Kinose, Sawada, Nakamura, & Kimura, 2014). Primarily ovarian cancer patients have some early and specific symptoms shared in common genitourinary, gastrointestinal, and gynaecological conditions which are not usually proven useful for early diagnosis (Kinose et al., 2014). The basis of this poor prognosis is due to insidious symptomatic nature in early stage, tumour resistance to chemotherapy, and lack of robust and minimal invasive method at its early detection (Iorio et al., 2007). Hence, advanced approaches for detection for early stage of ovarian cancer is necessary for proper medication and treatment timely.

Ovarian carcinomas have four major histological subtypes *clear cell, endometrioid, serous, and mucinous, with serous being* most frequent (Iorio et al., 2007). Latest data on large scale analysis of ovarian cancer samples suggest acquisition of invasiveness accumulated at mesenchymal subtype of tumours to be associated with transforming growth factor-β (*TGF- β*), *TGF-β* is a multifunctional protein which induces epithelial to mesenchymal transition (EMT) that leads to metastasis and is associated with chemotherapy resistance in multiple cancers (Parikh et al., 2014). Recent investigations indicate that these

Molecular Neurobiology (2019) 56:1737–1748
https://doi.org/10.1007/s12035-018-1173-y

Codon Usage Pattern of Genes Involved in Central Nervous System

Arif Uddin[1] · Supriyo Chakraborty[2]

Received: 10 January 2018 / Accepted: 1 June 2018 / Published online: 19 June 2018

Abstract
Codon usage bias (CUB) is the non-uniform usage of synonymous codons in which some codons are more preferred to others in the transcript. Analysis of codon usage bias has applications in understanding the basics of molecular biology, genetics, gene expression, and molecular evolution. To understand the patterns of codon usage in genes involved in the central nervous system (CNS), we used bioinformatic approaches to analyze the protein-coding sequences of genes involved in the CNS. The improved effective number of codons (ENC) suggested that the overall codon usage bias was low. The relative synonymous codon usage (RSCU) revealed that the most frequently occurring codons had a G or C at the third codon position. The codons namely TCC, AGC, CTG, CAG, CGC, ATC, ACC, GTG, GCC, GGC, and CGG (average RSCU > 1.6) were over-represented. Both mutation pressure and natural selection might affect the codon usage pattern as evident from correspondence and parity plot analyses. The overall GC content (59.93) was higher than AT content, i.e., genes were GC-rich. The correlation of GC12 with GC3 suggested that mutation pressure might affect the codon usage pattern.

Keywords Central nervous system · Codon usage bias · Mutation pressure · RSCU

Background

Neurodegenerative diseases (ND) occur when a nerve cell in the central nervous system or peripheral nervous system loses its function over time and ultimately the cell dies. ND affect millions of people over the globe, among which Alzheimer's disease and Parkinson's disease are more prevalent. Reports have suggested that over five million Americans suffer from Alzheimer's disease, and at least 500,000 Americans suffer from Parkinson's disease. The risk of ND increases gradually and its incidence is dramatically related to age. Neurodegeneration is the process of neuropathological conditions and brain aging. It is well known that brain pathology and neurodegenerative diseases are the most important causes of death all over the world. The neurodegenerative disorders such as Parkinson disease (PD), Alzheimer's disease (AD), dementia, cerebrovascular impairment, and seizure disorders have been accounted for the major health problem in the twenty-first century. Neurodegenerative disorders are caused by the defects in some of the genes. AD is a progressive neurodegenerative disorder that accounts for a vast majority of age-related dementia and is known to be one of the most serious health problems in the modern world. AD is characterized by cognitive demur and the accumulation of Aβ deposits and neurofibrillary tangles in the brain. Genetically, the mutations in three genes (i.e., APP, PSEN1, and PSEN2) have been shown to cause AD [1]. Fronto temporal dementia (FTD) with parkinsonism brings about a mutation in the microtubule-associated protein tau [2]. PD is the second most common neurodegenerative disease of adult onset, characterized by a severe loss of dopaminergic neurons in the substantia nigra region of the brain and cytoplasmic inclusions. The mutation in α-synuclein leads to Parkinson disease (PD). The nigra region and cytoplasmic inclusions consist of insoluble protein aggregates in the form of Lewy bodies, causing difficulty in the progressive movement, namely the classic triad of tremor, bradykinesia, and rigidity. The PD occurs at an average age of 50 to 60 years [3–5]. The mutations of five genes have now been shown to cause parkinsonism in early onset such as α-

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s12035-018-1173-y) contains supplementary material, which is available to authorized users.

✉ Arif Uddin
arif.uddin29@gmail.com

✉ Supriyo Chakraborty
supriyoch_2008@rediffmail.com

[1] Department of Zoology, Moinul Hoque Choudhury Memorial Science College, Algapur, Hailakandi, Assam 788150, India

[2] Department of Biotechnology, Assam University, Silchar, Assam 788011, India

Received: 14 June 2018 | Accepted: 17 August 2018

DOI: 10.1002/jcp.27375

ORIGINAL RESEARCH ARTICLE

WILEY Cellular Physiology

Understanding molecular biology of codon usage in mitochondrial complex IV genes of electron transport system: Relevance to mitochondrial diseases

Arif Uddin[1] | Tarikul Huda Mazumder[2] | Supriyo Chakraborty[2]

[1]Department of Zoology, Moinul Hoque Choudhury Memorial Science College, Hailakandi, Assam, India

[2]Department of Biotechnology, Assam University, Dargakona, Silchar, Assam, India

Correspondence
Arif Uddin, Department of Zoology, Moinul Hoque Choudhury Memorial Science College, Hailakandi, Assam, India.
Email: arif.uddin29@gmail.com
Supriyo Chakraborty, Department of Biotechnology, Assam University, Dargakona, Silchar, Assam 788011, India.
Email: supriyoch_2008@rediffmail.com

Abstract

The mitochondrial cytochrome oxidase (CO) genes are involved in complex IV of the electron transport system, and dysfunction of CO genes leads to several diseases. However, no work has been reported on the codon usage pattern of these genes. We used bioinformatic methods to analyze the compositional properties and the codon usage pattern of the COI, COII, and COIII genes in fishes, birds, and mammals to understand the similarities and dissimilarities of codon usage in these genes, which gave an insight into the molecular biology of these genes. The effective number of codons (ENC) value of genes was high in different species of fishes, birds and mammals, which indicates that the codon bias of CO genes was low and the ENC values were significantly different among fishes, birds, and mammals, as revealed from the *t* test. The overall guanine and cytosine (GC) content in fishes, birds, and mammals was lower than 50% in all genes, indicating that the genes were AT-rich and significantly different among fishes, birds, and mammals. The TCA codon was overrepresented in fishes, birds, and mammals for the COI gene, in birds and mammals for the COII gene, but it was not overrepresented in others. Only three codons, namely CTA, CGA, and AAA, were overrepresented in all three groups for the COI, COII, and COIII genes, repectively. From the neutrality plot in fishes, birds, and mammals, it was observed that the slopes of the regression lines (regression coefficients) in the COI, COII, and COIII genes were <0.5, suggesting that natural selection played a major role, whereas mutation pressure played a minor role.

KEYWORDS
codon usage bias (CUB), cytochrome oxidase (CO) gene, electron transport system (ETS), mutation pressure, natural selection

1 | INTRODUCTION

Several metabolic enzyme systems function within mitochondria. These include the components of TCA (tricarboxylic acid) cycle enzymes and the β-oxidation pathway of fatty acids (H. Liu et al., 2014). Mitochondrial diseases are the results of either inherited or spontaneous mutations in mitochondrial DNA (mtDNA) or mDNA, which lead to altered functions of the proteins or RNA molecules that normally reside in mitochondria (Wallace, 1992). Mitochondrial dysfunction is involved in various diseases, such as cancer and neurodegenerative disorders, including Alzheimer's and Parkinson's diseases (Burté, Carelli, Chinnery, & Yu-Wai-Man, 2015). From previous studies, it was observed that mtDNA mutations have been associated with cancers (Modica-Napolitano & Singh, 2004). So, mutations in mtDNA might be expected to influence the gene expression and copy number of the mitochondrial genome. Different

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Research paper

DNA compositional dynamics and codon usage patterns of M1 and M2 matrix protein genes in influenza A virus

Himangshu Deka[a], Durbba Nath[a], Arif Uddin[b,*], Supriyo Chakraborty[a,**]

[a] Department of Biotechnology, Assam University, Silchar 788011, Assam, India
[b] Department of Zoology, Moinul Hoque Choudhury Memorial Science College, Hailakandi 788150, Assam, India

ARTICLE INFO

Keywords:
Influenza A virus
A/H3N2
Matrix protein
Codon usage
Neutrality
Mutation pressure

ABSTRACT

Influenza A virus subtype H3N2 has been a serious health issue across the globe with approximately 36 thousand annual casualties in the United States of America only. Co-circulation in avian, swine and human hosts has led to frequent mutations in the virus genome, due to which development of successful antivirals against the virus has become a formidable challenge. Recently, focussed research is being carried out targeting the matrix proteins of this strain as vaccine candidates. This study is carried out to unravel the key features of the genes encoding the matrix proteins that manoeuvre the codon usage profile in the H3N2 strains. The findings reveal differential codon choice for both matrix protein 1 and matrix protein 2. The overall codon usage bias is less pronounced in both the datasets which is evident from higher value of effective number of codons (> 55). Comparison of the codon usage for both the genes under study with that of humans revealed that the viral codon usage is not fully optimized for the human host conditions. Both the genes enrolled in the study showed variation which was reflected in almost all the indices used for codon usage studies. Neutrality analysis revealed a weak role of mutation pressure while selection was the major contributor towards codon usage.

1. Introduction

A member of *Orthomyxoviridae*, influenza A virus (IAV) has many subtypes circulating among a diverse range of hosts including human. Among the most important IAV subtypes infecting the humans, H3N2 has been one of the most severe strains. The seasonal IAV subtype H3N2 (A/H3N2) has been a major concern on global basis due to its high variability and reassortment. Last century had witnessed a H3N2 pandemic in the form of the Hong Kong flu during 1968–1969. This particular strain was, in fact, a reassortant of H2N2 subtype due to antigenic shift. Every season about 36,000 humans die in the United States of America as a result of H3N2 infection. The A/H3N2 mainly circulates among swine, avian and human hosts which have been reported to co-circulate with A/H1N1 and avian-origin A/H9N2 in pigs (Campitelli et al., 1997; Peiris et al., 2001).

Influenza A virus exerts tremendous virulence to warm blooded animals (birds and mammals) (Holmes et al., 2005; Rambaut et al., 2008). Based on the antigenic property of surface glycoproteins *viz.* haemagglutinin (HA) and neuraminidase(NA), the virus is classified into several subtypes (Lekcharoensuk et al., 2010). The genome consists of negative sense RNA fragmented into 8 single strands coding for minimum 11 proteins. The virus genome goes through consistent modifications by means of point mutations as well as genetic recombination and reassortment (Chan et al., 2006). The approximate molecular weight of the viral segments ranges from 1×10^6 to $2–4 \times 10^5$ (Palese and Schulman, 1976).

The dire need for an effective vaccine has prompted many workers to try different targets within the IAV genetic framework. Much effort has been laid on the surface proteins hemagglutinin (HA), neuraminidase (NA) and also the ion channel protein M2 (Mosier et al., 2016; Xiong et al., 2015). But due to high plasticity of IAV genome, development of an effective vaccine against these highly mutable surface proteins has been challenging. Alternative targets that are being tried out in recent times include nucleocapsid protein (NP) and integral matrix protein 1 (M1) (Antrobus et al., 2014; Berlanda Scorza et al., 2016). It makes sense in research to gain better understanding of the genetic features of these potential drug targets. Codon usage bias analysis is a useful technique which gives insights for understanding the underlying factors influencing the genetic architecture of an organism and also evolution at molecular level.

Codon usage bias is a well known phenomenon of unequal usage of synonymous codons in coding sequences of genes and has been

* Corresponding author.
** Corresponding author.
E-mail addresses: arif.uddin29@gmail.com (A. Uddin), supriyoch_2008@rediffmail.com (S. Chakraborty).

https://doi.org/10.1016/j.meegid.2018.10.015
Received 20 September 2018; Received in revised form 11 October 2018; Accepted 23 October 2018
Available online 24 October 2018

Received: 18 September 2018 | Accepted: 15 October 2018

DOI: 10.1002/jcb.28039

RESEARCH ARTICLE

WILEY [redacted]

Chronic obstructive pulmonary disease: A crosstalk on nucleotide compositional dynamics and codon usage patterns of the genes involved in disease

Durbba Nath[1] | **Himangshu Deka**[1] | **Arif Uddin**[2] | **Supriyo Chakraborty**[1]

[1]Department of Biotechnology, Assam University, Silchar, India

[2]Department of Zoology, Moinul Hoque Choudhury Memorial Science College, Algapur, India

Correspondence
Supriyo Chakraborty, Department of Biotechnology, Assam University, Silchar 788011, Assam, India.
Email: supriyo123chakraborty@yahoo.in; supriyoch_2008@rediffmail.com

Abstract

Chronic obstructive pulmonary disease (COPD), a lung disease, affects a large number of people worldwide, leading to death. Here, we analyzed the compositional features and trends of codon usage of the genes influencing COPD to understand molecular biology, genetics, and evolutionary relationships of these genes as no work was reported yet. Coding sequences of COPD genes were found to be rich in guanine-cytosine (GC) content. A high value (34-60) of the effective number of codons of the genes indicated low codon usage bias (CUB). Correspondence analysis suggested that the COPD genes were distinct in their codon usage patterns. Relative synonymous codon usage values of codons differed between the more preferred codons and the less-preferred ones. Correlation analysis between overall nucleotides and those at third codon position revealed that mutation pressure might influence the CUB of the genes. The high correlation between GC12 and GC3 signified that directional mutation pressure might have operated at all the three codon positions in COPD genes.

KEYWORDS

chronic obstructive pulmonary disease, codon usage bias, effective number of codons, GC content, nucleotide composition, relative synonymous codon usage

1 | INTRODUCTION

Degeneracy of genetic code ensures that more than one codon codifies an amino acid except for methionine and tryptophan. These codons for an amino acid are known as synonymous codons.[1] Codon usage in the genomes of all organisms has been found to be unequal, showing increased preference of a few codons over others. This phenomenon of tilted codon usage in mature messenger RNA (mRNA) molecules is termed as codon usage bias (CUB).[2] The paradigm of codon bias among the genes of the same organism may vary owing to the amino acid composition in protein and mRNA structure.[3] Codons are usually selected for translational efficiency by a variety of selection mechanisms such as DNA folding, mRNA secondary structure, use of codons with common transfer RNAs (tRNAs), or by the interaction between codon and complementary anticodon.[4] To avoid missense errors and ribosome sequestering, optimal codons related to more abundant tRNAs are selected under the selection process, which in turn assists polypeptide chain elongation.[5] CUB in mammalian major histocompatibility complex genes was found to be influenced by gene function.[6] The tRNA abundance of the preferred codons was also found to correspond with more gene copy numbers than the less-preferred codons.[7] The compositional constraints in the

MITOCHONDRIAL DNA PART A
https://doi.org/10.1080/24701394.2018.1532414

Taylor & Francis
Taylor & Francis Group

RESEARCH ARTICLE

A cross-talk on compositional dynamics and codon usage patterns of mitochondrial *CYB* gene in Echinodermata

Masuk Ahmed Barbhuiya[a], Arif Uddin[b] and Supriyo Chakraborty[a]

[a]Department of Biotechnology, Assam University, Silchar, India; [b]Department of Zoology, Moinul Hoque Choudhury Memorial Science College, Hailakandi, India

ABSTRACT

Codon usage bias (CUB) refers to a phenomenon in which some synonymous codons are used in mature mRNA at a higher frequency than other members codifying the same amino acid. CUB is mainly determined by mutation pressure and natural selection. We used bioinformatic tools to analyze the protein coding sequences of mitochondrial *CYB* gene in different classes of Echinodermata to understand the patterns of codon usage. The ENC values of *CYB* gene in five different classes of Echinodermata were 41.64, 30.33, 43.63, 41.11, and 41.33, which suggested that the CUB of this gene was low. The relative synonymous codon usage (RSCU) values showed that the patterns of over-represented and under-represented codons were different among different classes. Correspondence analysis indicated that the plots of *CYB* gene were different across classes, suggesting that the pattern of codon usage was also different among five classes under study. Highly significant correlation ($p < .01$) between overall nucleotide composition and its 3rd codon position indicated that both mutational pressure and natural selection had an influence on the codon usage bias of *CYB* gene. Furthermore, PR-2 bias plot analysis showed that both mutation pressure and natural selection might have affected the pattern of codon usage in *CYB* gene of Echinodermata.

ARTICLE HISTORY
Received 28 May 2018
Accepted 19 September 2018

KEYWORDS
Codon usage bias; mutation pressure; natural selection; Echinodermata

Introduction

The genetic code is the protocols that direct the translation of mRNA molecule into 20 amino acids, the basic units of proteins in living cells of an organism (Tomita et al. 1999). The genetic code consists of codons, which are three-letter chains of nucleotides (Sutcliffe 1978). Each codon codifies one specific amino acid (Yang et al. 2000). The genetic code possesses the property of degeneracy, i.e., more than one codon encodes most of the amino acids (Lin and Beal 2006). Codons that determine the same amino acids are called synonymous and they usually differ in the last base of the triplet (Page and Holmes 2009). In the standard genetic code two amino acids namely methionine and tryptophan are encoded by a single codon (Judson and Haydon 1999). But the mitochondrial DNA (mtDNA) of the phylum Echinodermata follows the genetic code of translation table 9 of NCBI (Telford et al. 2000). Out of 64 codons, 62 sense codons encode 20 amino acids and the remaining two codons (TAA, TAG) act as termination signals in protein synthesis unlike standard genetic code (Prescott 1994). In translation table 9, the amino acid serine is encoded by eight codons whereas methionine and tryptophan are encoded by two codons each in the mitochondria of Echinodermata (Osawa et al. 1992).

In most organisms, synonymous codons are not equally used (Wright 1990). Codons which are used in higher frequency are referred to as optimal codons (Sharp and Li 1987). Codon usage bias (CUB) points to a phenomenon in which some synonymous codons occur in higher frequency than others of the same group (Bulmer 1991). Codon usage bias (CUB) is a regular phenomenon in many species (Wright 1990) namely *Escherichia coli* (Blattner et al. 1997), *Arabidopsis thaliana* (Chiapello et al. 1998), *Xanthophyllomyces dendrorhous* (Baeza et al. 2015), *Taenia saginata* (Jeon et al. 2007), *Megalobrama amblycephala* (Duan et al. 2015), metazoans (Duret 2002), and even human beings (Consortium 2001). Analysis of codon usage bias is thus of prime importance to predict the gene expression levels within an organism. Recent studies described that some special synonymous codons can influence protein folding and cause errors in folding (Drummond and Wilke 2008). Moreover, studies have reported that the inherent links between codon usage and amino acids that affect the proteome of cells in an organism (Costanzo et al. 2010). Hence, the understanding of codon usage bias also plays a key role in determining gene functions (Gogarten et al. 2002).

The genome-wide study of codon usage patterns helps us in understanding the basic features of the molecular organization of genomes. In general, composition biased mutation pressure and selection pressure for accurate and efficient translation are the main reasons for codon usage (Plotkin and Kudla 2011). An equivalence in the codon usage pattern indicated some degree of biological relationship, environmental adaptation, and evolution among *Plasmodium* species (Kellis et al. 2003). Recent studies have revealed that patterns of CUB and nucleotide

CONTACT Supriyo Chakraborty supriyoch_2008@rediffmail.com Department of Biotechnology, Assam University, Silchar, Assam, 788011, India

Molecular & Biochemical Parasitology 225 (2018) 15–26

Contents lists available at ScienceDirect

Molecular & Biochemical Parasitology

journal homepage: www.elsevier.com/locate/molbiopara

Preference of A/T ending codons in mitochondrial ATP6 gene under phylum Platyhelminthes
Codon usage of ATP6 gene in Platyhelminthes

Gulshana A Mazumder[a], Arif Uddin[b], Supriyo Chakraborty[a,*]

[a] Department of Biotechnology, Assam University, Silchar, 788011, Assam, India
[b] Moinul Hoque Choudhury Memorial Science College, Algapur, Hailakandi, 788150, Assam, India

ARTICLE INFO

Keywords:
Codon usage bias
Mutation pressure
Natural selection and synonymous codon usage order

ABSTRACT

Unequal usage of synonymous codons in the gene transcript for an amino acid is known as codon usage bias (CUB). It is a unique property of gene as well as genome. Mutation and natural selection are known to be the major factors that influence CUB. Other factors encompass gene expression level, GC content, codon position, recombination rate, RNA stability and gene length. CUB analysis helps in in-depth understanding of the molecular biology, genetics and genome evolution in a species. We used bioinformatic methods to explore the pattern of CUB in MT-ATP6 gene in different classes of platyhelminthes. The analysis is based on genetic code of translational table 14 of National Center of Biotechnology Information (NCBI) where the codon AAA codes for asparagine and TAA for tyrosine amino acid. The synonymous codon usage order (SCUO), an index of CUB, values in different classes namely cestoda, monogenea, rabditophora, trematoda and turbellaria of platyhelminths were found to be 0.43, 0.32, 0.49, 0.40 and 0.36, respectively which suggest that the codon usage bias of ATP6 gene was low (SCUO < 0.50). Highly significant correlation ($p < 0.001$) was found between SCUO and various GC contents indicating that GC composition had an influence on CUB. From the relative synonymous codon usage (RSCU) analysis on codons, we found most of the over-represented codons in all the classes were A/T ending types, which suggested that the preferred codons were influenced by compositional constraints. The PR2 plot revealed asymmetric usage of AT and GC bases among the four fold degenerate codon families with greater usage of G and T over A and C. Highly significant correlation ($p < 0.001$) was found between overall nucleotide composition and its 3rd codon position suggesting that both natural selection and mutation pressure might have influenced the CUB among different classes. Neutrality analysis revealed that natural selection might play a major role. Mutational responsive index (MRI) and translational selection (P2) values elucidated that selection for translational efficiency moderately affected the codon usage bias in MT-ATP6 gene.

1. Introduction

Unlike standard genetic code the mitochondrial genetic code of phylum Platyhelminthes follows the genetic code given in translation table 14 of NCBI. It consists of 63 sense codons and a single stop codon (TAG), and 18 amino acids are encoded by of 2–8 synonymous codons. The amino acid serine is encoded by 8 codons and both asparagine and tyrosine are encoded by 3 codons each. Amino acid tryptophan is encoded by 2 codons whereas lysine and methionine are encoded by one codon each. This genetic code is uniquely characterized by AAA codon encoding asparagine and TAA codon encoding tyrosine [1].

In 1960s, the idea cropped out that most of amino acids were coded by more than one codon (2–6), with difference at the third codon position. Later in 1980's with the introduction of whole genome sequencing, it was stated that synonymous codons are not used in equal frequency in mRNA molecules [2]. Codons which encode the same amino acid are known as synonymous codons, and the usage of these codons in an mRNA is a non-random process. The phenomenon wherein some codons are used more often than others within the same family leads to the manifestation of codon usage bias (CUB) [3,4]. Numerous studies revealed that codon usage pattern is mainly influenced by natural selection and mutational pressure [5–8]. CUB might play an important role in modulating the expression level of exogenous genes in a genome and also in predicting the optimum host range of the exogenous genes [9].

Various factors that influence the codon usage bias in different

* Corresponding author.
E-mail address: supriyoch_2008@rediffmail.com (S. Chakraborty).

https://doi.org/10.1016/j.molbiopara.2018.08.007
Received 20 May 2018; Received in revised form 17 August 2018; Accepted 22 August 2018
Available online 24 August 2018
0166-6851/

MITOCHONDRIAL DNA PART A
https://doi.org/10.1080/24701394.2018.1472772

Taylor & Francis
Taylor & Francis Group

RESEARCH ARTICLE

Transcript free energy positively correlates with codon usage bias in mitochondrial genes of *Calypogeia* species (Calypogeiaceae, Marchantiophyta)

Sudipa Das[a], Arif Uddin[b], Debjyoti Bhattacharyya[a] and Supriyo Chakraborty[c]

[a]Department of Life Science and Bioinformatics, Assam University, Assam, India; [b]Department of Zoology, Moinul Hoque Choudhury Memorial Science College, Algapur, Assam, India; [c]Department of Biotechnology, Assam University, Assam, India

ABSTRACT

The present study attempts to focus on the pattern of codon usage bias (CUB) in mitochondrial genes among four species of *Calypogeia viz. C. arguta, C. integristipula, C. neogaea* and *C. suecica* along with the factors influencing their CUB as no work was reported yet. The four taxa exhibit noticeable difference in their morphological features, habitat preference and growth forms. To understand the patterns of codon usage in *Calypogeia* mitochondrial genomes, we used bioinformatic tools to analyze the mitochondrial protein-coding sequences of four *Calypogeia*. In our current study, the mean nucleobase T was the highest while G was the lowest in all four species. The overall GC content was higher than AT content *i.e.* genes were AT rich. The CUB was low in mitochondrial genes of *C. arguta, C. integristipula, C. neogaea* and *C. suecica*, respectively as revealed from low synonymous codon usage order (SCUO) values. Heat map study revealed that most of the GC and AT ending codons showed positive correlations between codon usage and GC3. From the neutrality plot, the slope of the regression line indicated that natural selection might have played a major role over mutation pressure in shaping the CUB of mitochondrial genes in these four taxa. Highly significant ($p < .01$) correlation was found between mRNA free energy (mFE) and SCUO for entire mRNA which suggested that the release of higher energy by entire mRNA is associated with higher degree of codon usage bias. Further, highly significant ($p < .01$) negative correlation of mFE for entire mRNA was found with GC, GC1, GC2 and GC3 in all the four taxa. This indicated that higher GC content might have induced the release of lesser energy by mRNA molecules.

ARTICLE HISTORY
Received 28 March 2018
Accepted 28 April 2018

KEYWORDS
Codon usage bias; mitochondrial gene; Calypogeia; Bryophyte

Introduction

The genus *Calypogeia* Raddi, commonly known as 'Pouchwort', is a leafy liverwort (Jungermanniopsida, Marchantiophyta) belonging to the family Calypogeiaceae Arnell. Globally, the genus is represented by about 90 species distributed worldwide (Buczkowska and Bączkiewicz 2011) with main centres of diversity in tropical and subtropical regions. The genus is morphologically characterized by monoecious to less often dioecious, irregularly branched plants, obliquely inserted incubous leaves with rounded or apiculate to retuse apices, reniform or subquadrate to orbicular underleaves with retuse to bifid apices, absence of perianth and presence of cylindrical, subterranean, pendulous marsupium densely covered with rhizoids. *Calypogeia* is considered as one of the most difficult genera of liverworts (Schuster 1966; Szweykowski 2006) because of their simple morphological characters and high phenotypic plasticity which have made the valid distinction difficult between two species (Schuster 1966, Szweykowski 1984). While in one hand environmental modifications have misled in creation of new distinct species, on the other hand truly distinct species were unrecognized (Buczkowska 2004). Experimental methods like isozyme and molecular markers associated with taxonomy helped unmask some of the hidden taxa up to certain extent. The species of *Calypogeia* are significant ecologically as well as for the chemical constituents present in them. Some species yield considerable amount of an essential oil, the main component of which is 1,4-dimethyl azulene. Besides this, a number of other sesquiterpenoids like (+)-δ- Candinene have been isolated and identified in *Calypogeia* (Asakawa et al. 2012). From the ecological point of view, species of *Calypogeia* contribute as a significant component of wood inhabiting species in boreal conifer forests. *C. integristipula* and *C. suecica* are found to grow as wood-inhabiting species in boreal conifer logs of natural forests in Europe. There, they act as the best indicators of pristine boreal forests characterized by a natural input of windthrows (Chopra 2005).

In the present study, four species of *Calypogeia viz. C. arguta* Nees & Mont. Ex Nees, *C. integristipula* Steph., *C. neogaea* (R. M. Schust) Bakalin and *C. suecica* (Arnell & J. Perss.) Müll. Frib. are considered for analysis of the codon usage pattern in mitochondrial genes.

C. arguta, commonly known as notched pouchworts, prefers to grow on loamy or clayey acidic soils forming low to extensive patches (Common 1988). The plants of the species are pale green, slender with *ca.* 1.2 mm ×1 mm leaves having a pair of small, divergent teeth at the apex of each

CONTACT Supriyo Chakraborty supriyoch_2008@rediffmail.com Department of Biotechnology, Assam University, Assam, India

Received: 25 December 2017 | Accepted: 16 March 2018
DOI: 10.1002/jcp.26607

REVIEW ARTICLE

WILEY Cellular Physiology

Role of miRNAs in lung cancer

Arif Uddin[1] | **Supriyo Chakraborty**[2]

[1] Department of Zoology, Moinul Hoque Choudhury Memorial Science College, Algapur, Hailakandi, Assam, India

[2] Department of Biotechnology, Assam University, Silchar, Assam, India

Correspondence
Arif Uddin, Department of Zoology, Moinul Hoque Choudhury Memorial Science College, Algapur, Hailakandi-788150, Assam, India.
Email: arif.uddin29@gmail.com
Supriyo Chakraborty, Department of Biotechnology, Assam University, Silchar-788011, Assam, India.
Email: supriyoch_2008@rediffmail.com

Funding information
Department of Health Research, Ministry of Health and Family Welfare (MHFW), Government of India, Grant number: V.25011/441/HRD/2016-HR, 13/10/2016

Lung cancer (LC) is the leading cause of cancer-related deaths all over the world, among both men and women, with an incidence of over 200,000 new cases per year coupled with a very high mortality rate. LC comprises of two major clinicopathological categories: small-cell (SCLC) and nonsmall-cell lung carcinoma (NSCLC). The microRNAs (miRNAs) are small noncoding RNAs, usually 18–25 nucleotides long, which repress protein translation through binding to complementary target mRNAs. The miRNAs regulate many biological processes including cell cycle regulation, cellular growth, proliferation, differentiation, apoptosis, metabolism, neuronal patterning, and aging. This review summarizes the role of miRNAs expression in LC. It also provides information about the miRNAs as biomarker and therapeutic target for lung cancer. Understanding the role of miRNAs in LC may provide insights into the diagnosis and treatment strategy for LC.

KEYWORDS
lung cancer, miRNAs, small-cell and nonsmall-cell lung carcinoma

1 | INTRODUCTION

Cancer is a diverse group of diseases characterized by unrestrained cell growth enhancing tumor formation and metastasis (Lopez-Camarillo et al., 2012). The essential features of tumors are insensitivity to growth-inhibition signals, self-sufficiency in growth signals, limitless replicative potential, evasion of apoptosis, tissue invasion, sustained angiogenesis, and metastasis (Hanahan & Weinberg, 2000, 2011).

Lung cancer (LC) is the leading cause of cancer-related deaths all over the world, among both men and women, with an incidence of over 200,000 new cases per year and a very high mortality rate (Takamizawa et al., 2004). Indeed, lung cancer is responsible for more deaths than breast, colon, and prostate tumors combined (Jemal et al., 2007). LC comprises of two major clinicopathological categories: small-cell (SCLC) and nonsmall-cell lung carcinoma (NSCLC) (Travis et al., 2011). SCLC accounts for around 12% of all cases, it is more aggressive than NSCLC, and it frequently metastasizes (Esposito, Conti, Ailavajhala, Khalil, & Giordano, 2010). NSCLC tumors consist mainly of three subtypes: adenocarcinoma (40% of lung cancers), squamous cell carcinoma (25% of lung cancers), and large cell carcinoma (10% of lung cancers) (Guz et al., 2014). NSCLC is less aggressive and more common, accounting for at least 88% of all lung cancer cases (Guz et al., 2014). The B-Raf mutations occur in about 1–2% of NSCLC and this mutation causes a permanent activation of the mitogen activated protein kinase (MAPK) pathway which promotes growth and proliferation of tumor. Previous studies reported that B-Raf inhibitors have shown hopeful results in BRAF mutated in NSCLC patients (Caparica et al., 2016).

The discovery of role of the miRNAs (miRNAs) in the regulation of gene expression has broadly shed glow on cancer biology (Lopez-Camarillo et al., 2012). The miRNAs control various processes such as cellular growth, proliferation, differentiation, regulation of cell cycle,

Abbreviations: AD, adenocarcinoma; ADC, AIDS-defining cancers miRNA mircro-ribonucleic acid; ADR, AIDS-defining cancers; AUC, area under the curve; CI, confidence interval; DFI, disease-free interval; FDR, false discovery rate; FFPE, formalin-fixed paraffin-embedded; HR, hazard ratio; hsa-miRNA, *Homo sapiens* miRNA; IHC, immunohistochemistry; LC, lung cancer; miRNAs, micro-ribonucleic acids; NLP, normal lung parenchyma; NSCLC, non-small-cell lung carcinoma; OS, overall survival; ref, reference; RFS, recurrence/relapse-free survival; SCC, squamous cell carcinoma.

Journal of Molecular Liquids 257 (2018) 144–154

Contents lists available at ScienceDirect

Journal of Molecular Liquids

journal homepage: www.elsevier.com/locate/molliq

Cybotactic nematic phase of achiral unsymmetrical bent-core liquid crystals – Quelling of polar ordering and the influence of terminal substituent moiety

Sourav Patranabish [a], Golam Mohiuddin [b], Nazma Begum [c], Atiqur R. Laskar [c], Santanu Kumar Pal [b], Nandiraju V.S. Rao [c], Aloka Sinha [a,*]

[a] Department of Physics, Indian Institute of Technology Delhi, Hauz Khas, New Delhi 110 016, India
[b] Department of Chemical Sciences, Indian Institute of Science Education and Research Mohali, Sector-81, SAS Nagar, Knowledge City, Manauli 140306, India
[c] Department of Chemistry, Assam University, Silchar, Assam 788011, India

ARTICLE INFO

Article history:
Received 26 October 2017
Received in revised form 13 February 2018
Accepted 21 February 2018
Available online xxxx

Keywords:
Bent-core
Liquid crystals
Cybotactic
Ferronematic

ABSTRACT

Nematic phase of bent-core liquid crystals (LCs) exhibiting cybotactic clusters (N_{Cyb}) have gained significant importance owing to its promising ability to demonstrate macroscopic biaxiality and the ferronematic phase. In this context, three achiral unsymmetrical four-ring bent-core LC compounds, bearing a long alkyloxy chain and differing only in the terminal substituent moiety (methyl, chloro, nitro), are designed and synthesized followed by their optical, dielectric, electro-optic and structural investigations. The presence of N_{Cyb} in the methyl and chloro substituted compounds was confirmed via dielectric spectroscopy and X-ray diffraction observations. The absence of ferroelectric behaviour in any of these compounds, even in the cybotactic nematic phase and in presence of polar substituent moieties (chloro and nitro), is attributed to the increased alkyloxy chain length and antiparallel molecular arrangement. The density functional theory (DFT) optimized molecular structure along with the experimental observations further substantiates these findings. The study establishes that cybotactic clusters and polar end moiety, although being a prerequisite for ferroelectric-like nature, do not necessarily result in a ferronematic phase.

1. Introduction

Shape-dependent achiral bent-core liquid crystals (LCs) exhibiting nematic phase have drawn special attention among the researchers in recent years and have been studied extensively due to their unusual shape and reduced symmetry offering widespread applications in electro-optic devices. The distinctive features of the bent-core nematic phase are its polarity [1], chirality [2–4] and biaxiality [5] despite the fact that they are formed by achiral molecules. The nematic phase exhibited by a majority of the bent-core LCs, possessing short range smectic-like cybotactic clusters (N_{Cyb}), is a hotly debated topic in liquid crystal science. The bent-core LCs possessing N_{Cyb} phase are potential candidates to exhibit macroscopic biaxiality [6–12], the ferroelectric nematic (ferronematic) phase [13–16] (ferroelectric switching), and negative bend-splay elastic anisotropy [17] that have promising applications in faster [12] and energy sustainable next-generation displays. Smectic C type cybotactic clusters in the nematic phase were first identified in 1,2,4-oxadiazole based bent-core LCs [18]. In a similar series of 1,2,4-oxadiazole derivatives, ferroelectric switching in the nematic phase was reported by Francescangeli et al., which was the first evidence of ferronematic phase in low molar mass thermotropic nematic LCs [13]. They also observed cybotactic clustering in the nematic phase of these LC compounds and the observed polar nature was attributed to the field-induced reorganization of polar cybotactic groups within the nematic phase with local biaxial ordering. Later, Shanker et al. also reported ferroelectric switching in a new series of 1,2,4-oxadiazole based bent-core LCs accompanied by the presence of cybotactic clusters [15]. Further, ferroelectric-like switching was observed by Ghosh et al. in four-ring achiral unsymmetrical bent-core compounds exhibiting nematic phase with cybotactic clusters of smectic C type molecular ordering [16]. Recently, Nafees et al. had reported the effect of replacing the polar substituent by a methoxy moiety in four-ring achiral bent-core compounds, resulting in an absence of polar ordering in the nematic phase although possessing smectic C type cybotactic clusters [19]. Thus, there is no subtle description whether the polar ordering in this class of LCs depends only on the cybotactic clusters or on the presence of polar end moieties and therefore needs a lucid explanation.

The bent-core LCs due to their reduced symmetry and transverse dipole moment, especially in the nematic phase, are strong candidates to

* Corresponding author.
E-mail address: aloka@physics.iitd.ac.in (A. Sinha).

https://doi.org/10.1016/j.molliq.2018.02.097
0167-7322/

IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-6, Issue-9, Sept. 2018 E-ISSN: 2347-2693

Hybrid Image Segmentation Model using KM, FCM, Wavelet KM and Wavelet FCM Techniques

A.H.M. Jaffar Iqbal Barbhuiya[1*], K. Hemachandran[2]

[1]Dept. of Computer Science, Assam University, Silchar, India
[2]Dept. of Computer Science, Assam University, Silchar, India

*Corresponding Author: jaffar.iqbal84@gmail.com, Tel.: +91-7002375667

Available online at: www.ijcseonline.org

Accepted: 24/Sept/2018, Published: 30/Sept/2018

***Abstract*—** An attempt has been made to study the DWT (Discrete Wavelet Transform) based K-Means (KM) clustering and DWT based Fuzzy C-Means (FCM) clustering methods for the segmentation of digital images. The segmentation results of Wavelet KM clustering and Wavelet FCM clustering are compared with the conventional KM clustering and FCM clustering techniques used for the segmentation. The images are split-up into identical areas using KM, FCM, wavelet KM and wavelet FCM algorithms. The algorithms are tested on different image formats available in the literature. The proposed methods are analyzed using discrete wavelet transform (DWT) for enhancing the digital images and various image features like regions, colors and shapes are considered to validate the proposed work. The segmentation results exhibit that the objects in various image clusters of wavelet KM and wavelet FCM performs better as compared to traditional KM and FCM clustering algorithm with respect to CPU execution time, sensitivity analysis, segmentation accuracy and PSNR(Peak Signal to Noise Ratio).

***Keywords*—** Image segmentation, Clustering, K-Means (KM), Fuzzy C-Means (FCM), Wavelet KM, Wavelet FCM, Discrete Wavelet Transform (DWT), CPU execution time, sensitivity analysis and segmentation accuracy

I. INTRODUCTION

Image clustering has been widely used in digital image processing for image segmentation, face detection, fingerprint recognition, object recognition, medical imaging and many more. Different authors have used clustering methods for image segmentation [1, 2, 3, 4, 5] and K-means (KM) and Fuzzy C-means (FCM) are most popular techniques [6, 7, 8, 9, 10, 11, 12]. The objective of wavelet based image segmentation is to extract the image features from the original image. The extracted image features are then classified into number of clusters by employing the K-means (KM) clustering and fuzzy C-means clustering (FCM) method. These methods are concisely presented in this section.

I.A. Segmentation using K-Means (KM) Clustering Algorithm

K-means clustering algorithm was first introduced by MacQueen in 1967 [13]. The method proceeds with a simple manner to categorize a specified data set via definite number of clusters which draws an attention. The basic concept is to derive k centroids for each and every cluster. The objective of the KM algorithm is to reduce objective function or a squared error function. The objective function, X may be represented as [14]:

$$X = \sum_{y=1}^{n} \sum_{z=1}^{k} ||Iy - Jz||^2 \quad (1)$$

In the above equation, n is the number of pixels in an image, k is the total number of clusters, , I_y means the yth pixel data in an image, J_z represents the centroid of the zth cluster and $||I_y-J_z||$ is called the Euclidean distance between data points I_y and the cluster centroids J_z. In the preliminary steps of K-means clustering, first the computation of number of clusters k should be carried out and after that the task for the centroid of those clusters should be observed.

The conventional KM algorithm can be stated that [14]

Step 1: Choose k randomly as initial centroid.

Step 2: Generate k clusters assigning each data points to the closest centroid.

Step 3: Compute the cluster centroids J_z, where $z = 1, 2, 3, \ldots \ldots, k$, according to equation (1).

Step 4: Compute the Euclidean distance, $||I_y-J_z||$ between data points and clusters, where $y = 1, 2, 3, \ldots, n$ and $z = 1, 2, 3, \ldots \ldots, k$.

International Journal of Computer Engineering & Technology (IJCET)
Volume 9, Issue 5, September-October 2018, pp. 32–42, Article ID: IJCET_09_05_005
Available online at
http://www.iaeme.com/ijcet/issues.asp?JType=IJCET&VType=9&IType=5
Journal Impact Factor (2016): 9.3590(Calculated by GISI) www.jifactor.com
ISSN Print: 0976-6367 and ISSN Online: 0976–6375

AN EFFICIENT EDGE DETECTION APPROACH USING DWT

A.H.M. Jaffar Iqbal Barbhuiya
Department of Computer Science, Assam University, Silchar, India

ABSTRACT

In today's digital world, image edge detection is widely used in all over the globe. Digital image edge detection is used to detect the sharp changes in brightness of the images and the edge of a digital image. There are many classical techniques used to detect image edge detection like Canny, Sobel, Prewitt and Robert edge detector method. The proposed research work mainly focuses on the study of different edge detection methods which were previously derived and to propose an efficient method of digital image edge detection technique using Discrete Wavelet Transform (DWT). The results were compared with one of the best edge detector method: Canny Edge Detector. In the proposed image edge detection method, the first analytical result showed that an input image after applying the thresholding technique represents a segmentation of image edges and the second analytical results showed the 8-connected components of pixel connectivity in two dimensional images related to their neighbor pixels producing the enhanced edges. To evaluate the performance of the proposed algorithm with the existing algorithms, the two thresholding values were applied as T_low and T_high using hysteresis thresholding method. PSNR and width σ were also computed to detect the efficiency of the proposed algorithm.

Key words: Connectivity analysis, DWT, Edge detection, PSNR, Thresholding.

Cite this Article: A.H.M. Jaffar Iqbal Barbhuiya, An Efficient Edge Detection Approach using DWT. *International Journal of Computer Engineering and Technology*, 9(5), 2018, pp. 32-42.
http://www.iaeme.com/IJCET/issues.asp?JType=IJCET&VType=9&IType=5

1. INTRODUCTION

An edge of an object in an image gives the border or boundary of a particular object in an image that differentiates adjacent image regions. Digital image edge detection is used to detect the sharp changes in brightness of the images and the edge of a digital image. So, in digital image processing, an edge is a curve which ensures a path of significant changes in image intensity [1]. There are many classical techniques used to detect image edge detection like Canny, Sobel, Prewitt and Robert edge detector method.

This research work is organized as follows: in section II, a review of the edge detection methods are described, in section III, the proposed edge detection approach using Canny edge detector and DWT is presented, in section IV, evaluation results as well as discussions of the

THE BENGALI MUSLIMS AND SOCIAL LIFE: FIELD STUDY OF SAIDBOND VILLAGE OF HAILAKANDI DISTRICT, ASSAM, INDIA

A. F. Gulenur Islam Barbhuiya and Nazia Parveen
Department of Anthropology, M. H. C. M. Science College, Assam, India
connect2gulbarbhuiya@gmail.com, naziap463@gmail.com

ABSTRACT

Present study is an attempt to understand the socio-cultural life of the Bengali Muslims of Saidbond Part-I village of Hailakandi district in Assam, India. Observation, interview and household census method was followed in the present study. The study reveals that the Bengali Muslims of Saidbond Part-I village of Hailakandi district follow Sunni sect of Islam. Their family structure is patrilineal in nature and marriage by negotiation is the prevailing practice among them. They generally follow community endogamy. Rice is the staple food and agriculture is their mainstay of livelihood.

Key Words: Bengali Muslims, Hailakandi, Field work, Socio-culture, etc.

INTRODUCTION

Field research or fieldwork is the collection of information outside a laboratory, library or workplace setting. The approaches and methods used in field research vary across disciplines. For example, biologists who conduct field research may simply observe animals interacting with their environments, whereas social scientists conducting field research may interview or observe people in their natural environments to learn their languages, folklore, and social structures. Field research involves a range of well-defined, although variable, methods: informal interviews, direct observation, participation in the life of the group, collective discussions, analyses of personal documents produced within the group, self-analysis, results from activities undertaken off- or on-line, and life-histories. Although the method generally is characterized as qualitative research, it may (and often does) include quantitative dimensions.

The field science, Anthropology shares an interest to study the social behaviour of man in its institutionalized form. As anthropology is the ordered knowledge of man, it includes the physical characters and diversity of form, social organization, kinship system, marriage, religion, agriculture, material culture, etc. in short any subject related to man is the concern anthropology. But sometimes it may not be possible to have detailed investigation about people for many reasons. Therefore, some aspects are selected, information are collected in the broader perspective of the people and their culture. To know the way people behave one must have to study their society in site. The field investigation makes a film impression in the minds of field workers about the various aspects of culture of the people in which the investigation is done. It bridges the gap between theoretical knowledge and practical experience (Barbhuiya and Deori, 2012).

IMPACT FACTOR : 5.7631(UIF) REVIEW OF RESEARCH
UGC APPROVED JOURNAL NO. 48514 ISSN: 2249-894X
VOLUME - 8 | ISSUE - 2 | NOVEMBER - 2018

SOCIO-CULTURAL PROFILE OF THE BENGALI MUSLIM AND THE MEITEI POPULATION LIVING IN TWO NEIGHBOURING VILLAGES OF CACHAR DISTRICT OF ASSAM, INDIA

A. F. Gulenur Islam Barbhuiya
Assistant Professor, Department of Anthropology,
M. H. C. M. Science College, Hailakandi, Assam, India.

ABSTRACT

Present study is an attempt to understand the socio-cultural life of the Bengali Muslims and the Meiteis living in two neighbouring villages of Cachar district in Assam, India. Observation as well as interview method was followed in the present study. The study reveals that although both the Bengali Muslims and the Meiteis living in close geographical proximity, they exhibit wide differences than similarities in socio-cultural characteristics which may be due to their distinct bio-cultural identity.

KEYWORDS: *Bengali Muslim, Meitei, Cachar, Field work, Socio-cultural, etc.*

INTRODUCTION

Assam, a state of India is inhabited by numerous endogamous communities of diverse ethnic and socio-cultural setting. The field science, anthropology known as the science of human being, includes the physical characters and diversity of form, social organization, kinship system, marriage, religion, agriculture, material culture, etc. in short any subject related to man is the concern of anthropology. To know the way people behave one must have to study their society in site. The field investigation makes a film impression in the minds of field workers about the various aspects of culture of the people in which the investigation is done. It bridges the gap between theoretical knowledge and practical experience (Barbhuiya & Deori, 2012).

Cachar district is one of the largest districts of Assam covers an area of 3786 sq.km and lies between 90°4'E and 93°15'E latitude and 24°22'N and 25°8'N longitude (Srivastava et al., 2010). Cachar is considered as a plain district but a number of hills spread across and surrounding the district. It is one of the most economically backward districts of India which is largely due to geographical barrier with the rest of the country. The district has a population of 1,736,319 with a sex ratio of 958 females per 1000 males and a literacy rate of 80.36% ("District Census," 2011). Bengali is the official language in the district with majority of the people primarily speak in Sylheti (a dialect) for communication. There are different communities inhabiting in the district like Bengali Hindu, Bengali Muslim, Meitei, Brishnupriya, Dimasa Kachari (Burmans of Cachar), Rongmei Naga, Hmar, Khasi, etc (Barbhuiya & Das, 2013; Barbhuiya et al., 2016).

OBJECTIVE

The main objective of the present study is to understand the socio-cultural life of the Bengali Muslim and the Meitei community living in two neighbouring villages of Cachar district of Assam, India.

THE BENGALI MUSLIMS AND SOCIAL LIFE: FIELD STUDY OF SAIDBOND VILLAGE OF HAILAKANDI DISTRICT, ASSAM, INDIA

A. F. Gulenur Islam Barbhuiya and Nazia Parveen
Department of Anthropology, M. H. C. M. Science College, Assam, India
connect2gulbarbhuiya@gmail.com, naziap463@gmail.com

ABSTRACT

Present study is an attempt to understand the socio-cultural life of the Bengali Muslims of Saidbond Part-I village of Hailakandi district in Assam, India. Observation, interview and household census method was followed in the present study. The study reveals that the Bengali Muslims of Saidbond Part-I village of Hailakandi district follow Sunni sect of Islam. Their family structure is patrilineal in nature and marriage by negotiation is the prevailing practice among them. They generally follow community endogamy. Rice is the staple food and agriculture is their mainstay of livelihood.

Key Words: Bengali Muslims, Hailakandi, Field work, Socio-culture, etc.

INTRODUCTION

Field research or fieldwork is the collection of information outside a laboratory, library or workplace setting. The approaches and methods used in field research vary across disciplines. For example, biologists who conduct field research may simply observe animals interacting with their environments, whereas social scientists conducting field research may interview or observe people in their natural environments to learn their languages, folklore, and social structures. Field research involves a range of well-defined, although variable, methods: informal interviews, direct observation, participation in the life of the group, collective discussions, analyses of personal documents produced within the group, self-analysis, results from activities undertaken off- or on-line, and life-histories. Although the method generally is characterized as qualitative research, it may (and often does) include quantitative dimensions.

The field science, Anthropology shares an interest to study the social behaviour of man in its' institutionalized form. As anthropology is the ordered knowledge of man, it includes the physical characters and diversity of form, social organization, kinship system, marriage, religion, agriculture, material culture, etc. in short any subject related to man is the concern anthropology. But sometimes it may not be possible to have detailed investigation about people for many reasons. Therefore, some aspects are selected, information are collected in the broader perspective of the people and their culture. To know the way people behave one must have to study their society in site. The field investigation makes a film impression in the minds of field workers about the various aspects of culture of the people in which the investigation is done. It bridges the gap between theoretical knowledge and practical experience (Barbhuiya and Deori, 2012).

ENERGY SOURCES, PART A: RECOVERY, UTILIZATION, AND ENVIRONMENTAL EFFECTS
https://doi.org/10.1080/15567036.2017.1381787

Comparative study of coal combustion residues from pulp and paper mills of Assam

Firoza Sultana and Ruma Rano

Department of Chemistry, National Institute of Technology, Silchar, Assam, India

ABSTRACT

In the present work, the physicochemical, mineralogical, and morphological characteristics of coal combustion residues (CCRs) collected from two different Pulp and Paper Mills of Assam, India, have been investigated. CCRs from both the sources were mechanically sieved to obtain various size-classified fractions and their physical properties were determined using standard methods. X-ray diffraction (XRD), Fourier Transformation Infrared (FTIR) Spectroscopy, Scanning Electron Microscopy (SEM) etc. techniques were employed to study the mineralogical and morphological characteristics of each size classified fraction. The loss on ignition (LOI) values of the coarser fractions were found to be unexceptionally high, implying the presence of a huge amount of unburned carbon, and indicating the incomplete combustion of feed coal. From the utilization point of view, coarser fractions may be recycled (as domestic fuels) and less-coarse fractions may be used as adsorbents of toxic organic chemicals. Due to the larger water-holding capacities (WHCs), the coarser fractions may be used for soil amendment.

KEYWORDS

Coal combustion residues; loss on ignition; unburned carbon; water-holding capacity; bulk density; phenol adsorption

Introduction

The concept of sustainable development emerged in the early and mid-1980s as an attempt to bridge the gap between environmental concerns about the increasingly evident ecological consequences of human activities and sociopolitical concerns about human development issues (Robinson, 2004). Industrial ecology (IE) is a discipline that has gained traction over the past decades. To conserve and optimize natural resource consumption, the by-product generated in a particular industrial process should be assimilated by other industrial activities. This minimizes the overall material and energy consumption as well as reduces environmental impacts and economic costs (Quijorna et al., 2011). There are many environmental benefits regarding the use of combustion wastes, such as saving natural resource, energy, landfill space, and reduction of air pollutants and CO_2 emissions. Therefore, there is a continuing concern in establishing procedures in which they can be reused efficiently. The residues generated post combustion of coal in coal-fired power plants are collectively termed as coal combustion residues (CCRs). It is such an industrial by-product that if not put to valuable use, it is recognized as an environmental pollutant and is one of the most complex anthropogenic materials that can be characterized.

With increasing demand of energy and power in all sectors, there is a continuous increase in coal consumption in thermal power plants, which generates a huge quantity of CCRs. The problems associated with the low level of coal ash utilization include inconsistency in the quality of ash produced and unavailability of appropriate cost-effective technologies for its proper utilization (Bhattacharjee and Kandpal 2002). The wide variation of percentage contribution (by weight) to the total CCRs produced is

CONTACT Ruma Rano ranoruma@yahoo.co.in Department of Chemistry, National Institute of Technology, Silchar, Assam 788010, India.

Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/ueso.

Advances in Computational Sciences and Technology
ISSN 0973-6107 Volume 10, Number 2 (2017) pp. 235-246

http://www.ripublication.com

Intrusion Detection System for NSL-KDD Data Set using Vectorised Fitness Function in Genetic Algorithm

Partha Sarathi Bhattacharjee
Assam University, Silchar
Cachar, Assam, India.

Abul Kashim Md Fujail
M H C M Sc College
Hailakandi, Assam, India

Shahin Ara Begum
Assam University, Silchar
Cachar, Assam, India.

Abstract

With rapid increase in the use of network computers over last few decades, there has been increase in many different types of network attacks by intruders. To detect different network attacks, Genetic Algorithm (GA) based Intrusion Detection System (IDS) is employed in this paper. The objective is to find a suitable Vectorised Fitness function for chromosome evaluations to get a solution for IDS. To achieve this objective, GA based IDS with weighted Vectorised Fitness function is proposed and evaluated over the NSL-KDD data set. In the present work, Fuzzy membership function is used with Vectorised Fitness function in GA for efficient intrusion detections. The experimental results show that the proposed Fuzzy Vectorised GA performs better than the Vectorised GA and Weighted Vectorised GA in detecting network attacks for the considered NSL-KDD data set.

Keywords: IDS, Genetic Algorithm, Vectorised Fitness Function, Fuzzy Membership function, NSL-KDD Data Set.

I. INTRODUCTION

Intrusion Detection System (IDS) is used to monitor network traffic and suspicious activity and alerts the system or network administrator. In some cases, the IDS may

Mitochondrion 42 (2018) 54–73

Contents lists available at ScienceDirect

Mitochondrion

journal homepage: www.elsevier.com/locate/mito

Codon usage and expression level of human mitochondrial 13 protein coding genes across six continents

Supriyo Chakraborty[a,*], Arif Uddin[b,*], Tarikul Huda Mazumder[a], Monisha Nath Choudhury[a], Arup Kumar Malakar[a], Prosenjit Paul[a], Binata Halder[a], Himangshu Deka[a], Gulshana Akthar Mazumder[a], Riazul Ahmed Barbhuiya[a], Masuk Ahmed Barbhuiya[a], Warepam Jesmi Devi[a]

[a] Department of Biotechnology, Assam University, Silchar, Assam 788001, India
[b] Department of Zoology, Moinul Hoque Choudhury memorial Science College, Algapur, Hailakandi, Assam 788150, India

ARTICLE INFO

Keywords:
Codon usage bias
Evolutionary relationships
Mitochondrial genes
Mutation pressure
Natural selection

ABSTRACT

The study of codon usage coupled with phylogenetic analysis is an important tool to understand the genetic and evolutionary relationship of a gene. The 13 protein coding genes of human mitochondria are involved in electron transport chain for the generation of energy currency (ATP). However, no work has yet been reported on the codon usage of the mitochondrial protein coding genes across six continents. To understand the patterns of codon usage in mitochondrial genes across six different continents, we used bioinformatic analyses to analyze the protein coding genes. The codon usage bias was low as revealed from high ENC value. Correlation between codon usage and GC3 suggested that all the codons ending with G/C were positively correlated with GC3 but vice versa for A/T ending codons with the exception of *ND4L* and *ND5* genes. Neutrality plot revealed that for the genes *ATP6, COI, COIII, CYB, ND4* and *ND4L*, natural selection might have played a major role while mutation pressure might have played a dominant role in the codon usage bias of *ATP8, COII, ND1, ND2, ND3, ND5* and *ND6* genes. Phylogenetic analysis indicated that evolutionary relationships in each of 13 protein coding genes of human mitochondria were different across six continents and further suggested that geographical distance was an important factor for the origin and evolution of 13 protein coding genes of human mitochondria.

1. Introduction

A codon is a set of three nucleotides that encode a specific amino acid residue in a polypeptide chain or for the termination of translation process. Due to degeneracy of the genetic code, a single amino acid is encoded by more than one codon except for two amino acids viz. methionine and tryptophan in standard genetic code. These codons are known as synonymous codons for an amino acid. The synonymous codons are not used with equal frequencies in the mature mRNA and this unequal relative codon usage frequency leads to codon bias. Thus, codon usage bias is the phenomenon involving non-uniform usage of synonymous codons encoding the same amino acid during the translation of mRNA to protein (Behura and Severson, 2012). Codon bias is a unique property of the genome of an organism and species specific but may vary significantly among the genes within the same organism (Grantham et al., 1980; Marin et al., 1989; Prat et al., 2009). It is well evident that the variation in codon usage within the same synonymous codon family dictates the translational efficiency of a gene, thus having a great impact in shaping genome evolution (Bentele et al., 2013). The increased translational efficiency of genes helps organisms adapt to the changing conditions and is found related to the lifestyle of the organisms (Botzman and Margalit, 2011). Previously, many studies have reported synonymous codon usage bias in different organisms from simple prokaryotes to higher eukaryotes. It was further reported that many genomic factors such as gene length, GC content, gene expression level, structure of mRNA and its stability mediate the codon usage bias in different organisms (Chen et al., 2004; Mazumder and Chakraborty, 2015). However, in general two major factors namely mutational pressure and weak natural selection are involved in guiding the codon usage bias in different organisms (Butt et al., 2014; Hershberg and Petrov, 2008).

The codon bias study has acquired renewed attention of the scientific community with the inception of whole genome sequencing in different organism. It acquires significance in molecular biology for understanding the patterns of synonymous codon usage, analysing the level of gene expression, genome characterization and also for

* Corresponding authors.
E-mail addresses: supriyoch_2008@rediffmail.com (S. Chakraborty), arif.uddin29@gmail.com (A. Uddin).

https://doi.org/10.1016/j.mito.2017.11.006
Received 14 July 2016; Received in revised form 9 October 2017; Accepted 27 November 2017
Available online 12 December 2017

Molecular & Biochemical Parasitology 214 (2017) 24–27

Contents lists available at ScienceDirect

Molecular & Biochemical Parasitology

journal homepage: www.elsevier.com/locate/molbiopara

Analysis of codon usage pattern of mitochondrial protein-coding genes in different hookworms

Bornali Deb[a], Arif Uddin[b,*], Gulshana Akthar Mazumder[a], Supriyo Chakraborty[a,*]

[a] Department of Biotechnology, Assam University, Silchar 788011, Assam, India
[b] Department of Zoology, Moinul Hoque Choudhury Memorial Science College, Algapur, Hailakandi 788150, Assam, India

ARTICLE INFO

Keywords:
Codon usage bias
GC content
AT content
Mutation pressure
Natural selection

ABSTRACT

The phenomenon of unequal usage of synonymous codons encoding an amino acid in which some codons are more preferred to others is the codon usage bias (CUB) and it is species specific. Analysis of CUB helps in understanding evolution at molecular level and acquires significance in mRNA translation, design of transgenes and new gene discovery. In our current study, we analyzed synonymous codon usage pattern and the factors influencing it on mitochondrial protein coding genes of 6 different hookworms *i.e. Ancylostoma ceylanicum, Ancylostoma duodenale, Necator americanus, Ancylostoma tubaeforme, Ancylostoma caninum* and *Uncinaria sanguinis* as no work was reported yet. The effective number of codons for mitochondrial genes suggested that codon usage bias was high in most species. The GC content was lower than AT content *i.e.* genes were AT rich as indicated by nucleotide composition analysis. The overall nucleotide composition along with its composition at 3rd codon position and correspondence analysis suggested that both natural selection and mutation pressure might have affected the codon usage bias in mitochondrial genes. However, neutrality plot revealed that mutation pressure might have played a major role in *A. ceylanicum* while natural selection might have played the dominant role in *Ancylostoma duodenale, Necator americanus, Ancylostoma tubaeforme, Ancylostoma caninum* and *Uncinaria sanguinis.*

1. Introduction

It is well known that 61 codons out of 64 codons encode 20 standard amino acids in standard genetic code. But mitochondrial genetic code of nemathelminths follows translation table 14 of National Center for Biotechnology Information (NCBI). It consists of 63 sense codons and only one stop codon (TAG). The degeneracy level of serine is 8 with the codons TCT, TCC, TCA, TCG, AGT, AGC, AGA and AGG. The amino acid leucine has six codons namely CTA, CTC, CTG, CTT, TTA and TTG while arginine has four codons *viz.* CGA, CGC, CGG and CGT. In this genetic code, tryptophan has two codons *i.e.* TGG and TGA. But two amino acids namely methionine and lysine are encoded by single codon each.The different codons that encode the same amino acid are called synonymous codons. Previous studies reported that the non uniform usage of synonymous codons in which some codons are used more preferentially than others is due to a feature known as codon usage bias (CUB) [1,2]. Earlier investigations suggested that CUB varies within the genome as well as between the genomes and it may help in understanding the genome evolution among related species [3].

Several hypotheses have been proposed to elucidate the origin of CUB. Amongst these hypotheses, the neutral theory [4] and the selection-mutation-drift balance model [5] are very important. According to neutral theory, mutations at degenerate codon positions should be selectively neutral which would result in non-uniform preference of synonymous codons of an amino acid and hence failure of natural selection pressure. According to selection-mutation-drift model, CUB is assumed to be balanced by mutation pressure, genetic drift and weak selection. On the other hand, if the gene expression level is high indicating high selective pressure, then it may result from stronger CUB [6]. However, with the completion of whole genome projects of many organisms, these two hypotheses are no longer enough to explain the observed CUB. Several important factors have been reported to affect the CUB such as compositional bias (GC% and GC skew), mutation pressure, natural selection, expression level, gene length, RNA stability, replication, hydrophobicity and hydrophilicity of the protein [7–10]. Of these, the compositional constraints in the presence of natural selection and mutation pressure are considered as the major factors which may vary across species [5,6,11]. The study of codon usage is essential to predict and optimize protein expression levels, to recognize protein coding genes, and to detect lateral gene transfer [12]. In the case of

* Corresponding authors.
E-mail addresses: arif.uddin29@gmail.com (A. Uddin), supriyoch_2008@rediffmail.com (S. Chakraborty).

https://doi.org/10.1016/j.molbiopara.2017.11.005
Received 14 May 2017; Received in revised form 14 November 2017; Accepted 16 November 2017
Available online 16 November 2017
0166-6851

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Research paper

Codon usage pattern of complex III gene of respiratory chain among platyhelminths

Gulshana A. Mazumder[a], Arif Uddin[b], Supriyo Chakraborty[a,*]

[a] Department of Biotechnology, Assam University, Silchar 788011, Assam, India
[b] Moinul Hoque Choudhury Memorial Science College, Algapur, Hailakandi 788150, Assam, India

ARTICLE INFO

Keywords:
Codon usage bias
MT-CYB gene
Mutation pressure
Natural selection

ABSTRACT

Codon usage bias refers to the phenomenon where synonymous codons are used with unequal frequencies. To understand the patterns of codon usage in mitochondrial cytochrome B (MT-CYB) gene of phylum platyhelminthes we used bioinformatic approaches to analyze the protein coding sequences of five different classes - cestoda, monogenea, rabditophora, trematoda and turbellaria. It was found from nucleotide composition analysis that in all the classes, A/T-ended codons were preferred to G/C -ended codons. From box plot analysis GC1 was found to have highest response to codon usage bias. Correspondence analysis indicated that besides mutation other factors such as natural selection might also affect the codon usage pattern. Neutrality plot reveals that both mutation and natural selection played role in codon usage pattern in five classes of MT-CYB gene. Various factors namely nucleotide composition, natural selection and mutation pressure affected the codon usage pattern.

1. Introduction

The mitochondria are called the power house of the cell and contain one or more circular DNA molecules (*i.e.* mt DNA). They consist of four membranes and help in the production of ATP (energy) through respiratory chain system. Respiratory system consists of five protein complexes- complex I, II, III, IV and ATPase synthase (complex V), and the chain also requires two electron carriers- ubiquinone and cytochrome c. The mitochondrial respiratory chain supplies energy to the cell in the form of ATP. Mitochondrial encephalomyopathies are the diseases which may be caused by mutations in this respiratory chain (DiMauro and Schon, 2003). The complex I is formed by seven mitochondrial encoded proteins and complex II is formed by the nuclear encoded proteins while the complex III of Respiratory chain (RC) is encoded by mitochondrial CYB gene which forms cytochrome b (CYB) protein during oxidative phosphorylation. The complex III of RC catalyzes the transfer of electrons from ubiquinol to cytochrome c. It has been found that the respiratory complex III deficiency is caused by nonsense, missense, or frame shift mutation in CYB gene. This deficiency was observed in many patients (Fernández-Vizarra and Zeviani, 2015). CYB gene has been extensively used for phylogenetic research, and the phylogenetic usefulness of CYB gene in different vertebrates has been shown by many workers (Irwin et al., 1991; Moritz et al., 1992). CYB gene is mostly used in systematics because it has rapidly evolving codon position and variable region (Farias et al., 2001; Meyer and Wilson, 1990).

According to the standard genetic code 20 amino acids are encoded by 61 sense codons, where each amino acid is encoded by 1–6 codons. The phylum platyhelminthes belongs to invertebrate and the mitochondria of platyhelminths were discovered to use genetic code according to Translation Table 14 of NCBI. The mitochondrial genetic code of platyhelminthes consists of 63 sense codons and a single stop codon (TAG), and each amino acid consists of 1–8 synonymous codons unlike standard genetic code. The amino acid serine is encoded by 8 codons and both Tyr and Asn are encoded by 3 codons each. Amino acid Trp is encoded by 2 codons whereas Met and Lys are encoded by 1 codon each. Synonymous codons encoding the same amino acid are generally used unequally in genes and the codons which are more or less preferred than others are known to create codon usage bias (CUB) (Ikemura, 1982, 1985). CUB is a unique property of genome and is found to vary from species to species. Various hypotheses have been reported to explain the origin of codon usage bias. The neutral theory (Nakamura et al., 1997) and the selection-mutation drift balance models (Sharp and Li, 1986) are the most accepted ones among these. CUB is also influenced by GC content, gene length, gene expression level, hydrophobicity, aromaticity of the encoded proteins and many other features.

It has been reported that the mutation pressure occurs due to high

* Corresponding author.
E-mail address: supriyoch_2008@rediffmail.com (S. Chakraborty).

https://doi.org/10.1016/j.meegid.2017.10.018
Received 14 June 2017; Received in revised form 30 August 2017; Accepted 19 October 2017
Available online 21 October 2017

DOI: 10.1515/ap-2017-00
© W. Stefański Institute of Parasitology, PAS
Acta Parasitologica, 2017, 62(4), 000–000; ISSN 1230-2821

DE GRUYTER

Comparative analysis of codon usage pattern, expression level and its influencing factors in *Schistosoma japonicum* and *Ascaris suum*

Gulshana A. Mazumder[1], Arif Uddin[2] and Supriyo Chakraborty[1]*

[1]Department of Biotechnology Assam University, Silchar-788011, Assam, India;
[2]Moinul Hoque Choudhury Memorial Science College Algapur, Hailakandi-788150, Assam, India

Abstract

Schistosoma japonicum and *Ascaris suum* are considered as the major parasites of human which cause various life threatening diseases such as schistomiasis and ascariasis. The codon usage bias (CUB) is known as the phenomenon of more usage of a specific codon than the other synonymous codons for an amino acid. The factors that influence the codon usage bias are mutation pressure, natural selection, gene expression, gene length, GC content, RNA stability, recombination rates, codon position etc. Here we had used various bioinformatic tools and statistical analyses to understand the compositional features, expression level and codon usage bias in the genes of these two species.After estimating the effective number of codon (ENC) and codon adaptation index (CAI) in both the species, codon usage bias was found to be low and gene expression was high. The nucleobase A and T were used most often than C and G. From neutrality plot and correspondence analysis it was found that both natural selection and mutation pressure played an important role in shaping the codon usage pattern of both species. Moreover, natural selection played a major role while mutation pressure played a minor role in shaping the codon usage bias in *S. japonicum and A.suum.* This is the first report on the codon usage biology in *S. japonicum and A.suum*, and the factors influencing their codon usage bias. These results are expected to be useful for genetic engineering and evolutionary studies.

Keywords

Schistosoma japonicum, *Ascaris suum*, codon usage bias, compositional properties, expression level, neutrality plot

Introduction

The genetic code is degenerate *i.e.*, 64 codons code for 20 standard amino acids, meaning that most of the amino acids are encoded by more than one codon. The codons generally differ at the third position. The non-uniform usage of synonymous codons *i.e.*, some codons are more frequently used than others in protein coding genes is known as codon usage bias. The frequency of the usage of the synonymous codons is usually unequal within and among different organisms. It has been found that various factors such as mutational bias, selection, intron splicing, gene conversion, protein secondary structures, and DNA replication are strongly related to synonymous codon usage bias (Drummond and Wilke 2008, Kahali *et al.* 2007, Warnecke and Hurst 2007). The balance between selection and mutational bias in prokaryotes or unicellular eukaryotes also determines the codon usage (Gouy and Gautier 1982, Sharp *et al.* 2005). In multicellular eukaryotes, such as *Caenorhabditis elegans* and *Drosophila melanogaster*, it is mostly determined by the selection for translational efficiency (Stenico *et al.* 1994). The equilibria among various forces such as mutation pressure, translational selection and genetic drift are considered as important factors responsible for explaining the codon usage patterns(Shah and Gilchrist 2011). Various other factors known to influence codon usage bias within and among species include geneexpression level, GC content, RNA stability, recombination rates, gene length, codon position, and others (which may include environmental stress and population size) (Behura and Severson 2013).

In this study our objective is to analyze the codon usage bias in the nuclear genes of *S. japonicum* and *A. suum*. *S. japonicum* is found in China and the Philippines. It is known to cause schistomiasis which poses a great threat to human health. Schistomiasis is an infection which is caused by three known species of Schistosoma (*S. mansoni, S. japonicum and S. haematobium*). Among these *S. japonicum* is considered to be the most infectious(Jia Tie-Wu *et al.* 2007). Infection caused by this species causes katayama fever (Ishii

*Corresponding author: supriyoch_2008@rediffmail.com

Contents lists available at ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Research paper

Factors affecting the codon usage bias of SRY gene across mammals

Mitochondrion xxx (2017) xxx-xxx

Contents lists available at ScienceDirect

Mitochondrion

journal homepage: www.elsevier.com

Factors influencing codon usage of mitochondrial ND1 gene in pisces, aves and mammals

Arif Uddin[a], Monisha Nath Choudhury[b], Supriyo Chakraborty[b,*]

[a] Department of Zoology, Moinul Hoque Choudhury Memorial Science College, Algapur, Hailakandi 788150, Assam, India
[b] Department of Biotechnology, Assam University, Silchar 788011, Assam, India

ARTICLE INFO

Keywords:
Codon usage
NADH dehydrogenase 1 gene
Natural selection
Mutation pressure

ABSTRACT

Animal mitochondrial genome harbours 13 protein coding genes which regulate the process of respiration. The mitochondrial NADH dehydrogenase 1 (MT-ND1) gene, one of the 13 protein-coding genes, encodes the NADH dehydrogenase 1 enzyme of the respiratory chain. Analysis of codon usage bias (CUB) acquires importance for better understanding of the molecular biology, new gene discovery, design of transgenes and gene evolution. The MT-ND1 gene seems to be a good candidate for analyzing codon usage pattern, since no work has yet been reported. Moreover, it is still not clear which factors significantly influence the codon usage pattern. In the present study, comparative analysis of codon usage pattern, expression level and influencing factors for MT-ND1 gene from 100 different species each of pisces, aves and mammals were used for CUB analysis. Our result suggests that the gene is AT rich in pisces, aves and mammals and most of the nucleotides significantly differ among them as revealed from *t*-test. CUB was not remarkable as reflected by high value of effective number of codons and it also significantly differs among pisces, aves and mammals. Although we found that CUB is mainly influenced by natural selection and mutation pressure for MT-ND1 gene as suggested by correlation and correspondence analysis but neutrality plot further revealed that natural selection played a major role and mutation pressure played a minor role in codon usage pattern. Additionally, *t*-test analysis showed that the MT-ND1 gene has a wide significant discrepancy in codon choices in pisces, aves and mammals. This study has contributed to boost our understanding about the mechanism of distribution of the codons and the factors that may influence the evolution of the MT-ND1 gene.

1. Introduction

The MT-ND1 gene is one of the 13 protein-coding genes in mitochondrial DNA involved in respiration. It encodes the NADH dehydrogenase 1 enzyme and forms a subunit of complex I of the mitochondrial respiratory chain (Hauptmann et al., 2009). The mitochondrial DNA is covalently closed, double stranded structure with nearly 16.6 kb size, which encodes 2 rRNA, 22 tRNAs and 13 polypeptides (Chen et al., 2009). Each polypeptide encoded by mitochondrial gene is a subunit of one of four respiratory complexes in the electron transport chain (ETC) localized in the inner membrane of mitochondria (Braun et al., 1992). The mitochondrial DNA is maternally inherited and harbours higher rates of mutation (Taylor and Turnbull, 2005). The lack of introns in the mitochondrial genes and histones in packaging of mitochondrial genome makes mitochondrial DNA more prone to mutation due to the presence of reactive oxygen species (ROS) generated by oxidative phosphorylation in the mitochondria (Kunkel and Loeb, 1981; Matsukage et al., 1975; Modica-Napolitano and Singh, 2004; Shay and Werbin, 1992; Singh et al., 2001; Torri and Englund, 1995). Mutation rate in mitochondrial DNA is tenfold higher than nuclear DNA (Shoubridge, 2000; Wilson and Roof, 1997). Mitochondrial DNA is an excellent tool for evolutionary study due to its small size and relatively conserved gene content and high mutation rate (Clark et al., 2007). Unlike nuclear genetic code, mitochondrial genetic code in vertebrates is composed of 60 sense codons that represent 20 standard amino acids, and the remaining four codons that act as termination signals are TAA, TAG, AGA and AGG (Knight et al., 2001). The genetic code is universal with a few exceptions in mitochondrial DNA. In animal mitochondrial DNA only six codons have been reported which vary in the process of evolution and these codons are TGA, ATA, AAA, AGA, AGG and TAA. The codon TGA acts as termination codon in standard genetic code but in

* Corresponding author.
Email address: supriyoch_2008@rediffmail.com (S. Chakraborty)

http://dx.doi.org/10.1016/j.mito.2017.06.004
Received 17 June 2016; Received in revised form 19 May 2017; Accepted 26 June 2017
Available online xxx
1567-7249/

EJA
Eurasian Journal of Anthropology
ISSN: 2166-7411

Euras J Anthropol 8(1):1-12, 2017

Anaemia and associated factors among the Bengali Muslims of Cachar district in Assam, India

A. F. Gulenur Islam Barbhuiya

Department of Anthropology
M. H. C. M. Science College, Hailakandi, Assam, India

Article info

Received: 3 February 2017

Accepted: 5 November 2017

Key words

Bengali Muslims, Cachar, Anaemia, Age, Education, Income.

For correspondence

Gulenur Islam Barbhuiya

Moinul Hoque Choudhury Memorial Science College, Hailakandi, Assam, India.

E-mail:
connect2gulbarbhuiya@gmail.com

Abstract

Present study is an effort to observe the prevalence of anaemia with reference to age, education and income among the Bengali Muslims of Cachar District of Assam, India. The data have been collected by household census method and colour scale for haemoglobin among 362 Bengali Muslims (male-183, female-179) of 15 to 79 years of age from Ganganagar Part-I and Bhaurikandi Part-II village of Cachar District. The study reveals that half (50.0%) of the adult Bengali Muslim population is prone to different grades of anaemia. 20.7% of them are found to have mild anaemia; but a major proportion (27.1%) of them are suffering from moderate anaemia and 2.2% of them are found to be severely anaemic. Prevalence of anaemia is found to be significantly high (p-0.001) among the Bengali Muslim females (60.3%) compared to their male counterparts (39.9%). Anaemia is found to be more prevalent in older age groups in comparison to younger age groups of the community. Prevalence of anaemia decreases with the increase of educational status and per capita monthly income in the community. The interaction of different factors like age, education, income and overall socio-economic condition may have influenced the anaemic situation in the community.

Eur J Clin Microbiol Infect Dis
DOI 10.1007/s10096-016-2907-2

ORIGINAL ARTICLE

Incidence of *Cryptosporidium andersoni* in diarrheal patients from southern Assam, India: a molecular approach

G. Hussain[1,2] · S. Roychoudhury[2] · B. Singha[1] · J. Paul[3]

Received: 21 November 2016 / Accepted: 20 December 2016

Abstract The distribution and public health significance of *Cryptosporidium* species and genotypes in humans and bovine differ across geographical areas. *Cryptosporidium* species causes a disease known as cryptosporidiosis in humans and animals. To characterize the prevalence of cryptosporidiosis in humans in southern Assam, India, stool samples ($n = 1119$) of diarrhea patients were collected from different hospitals and from the community during the period January 2014 to July 2016. Fecal smears were examined microscopically for *Cryptosporidium* species using modified acid fast staining and were screened to ascertain the presence of *Cryptosporidium* antigen by enzyme-linked immunosorbent assay (ELISA). The genomic DNA of positive fecal samples were analyzed by nested polymerase chain reaction (PCR), which were subsequently genotyped by PCR-restriction fragment length polymorphism (RFLP), based on small subunit (SSU) 18S rRNA. It was found that the prevalence of *Cryptosporidium* spp. was high during the monsoon season. The average infection rate of *Cryptosporidium* spp. was found to be 2.4% (27/1119) microscopically. When subjected to nested PCR using amplification of the 18S rRNA gene, *Cryptosporidium* was found to be 8.57% (98/1119). Based on the 18S rRNA gene, two *Cryptosporidium* spp., namely *Cryptosporidium andersoni* (6.97%: 78/1119) and *Cryptosporidium parvum* (1.7%: 20/1119), were identified. *Cryptosporidium andersoni* infections were found to be of either zoonotic or anthroponotic origin. The prevalence was statistically significant ($p = 0.03$, $R^2 = 0.042$) considering age, gender, and cast.

✉ B. Singha
babysingha@gmail.com

1 Molecular Parasitology Laboratory, Gurucharan College, Silchar 788004, Assam, India

2 Department of Life Science and Bioinformatics, Assam University, Silchar 788011, Assam, India

3 School of Life Science, Jawaharlal Nehru University, 110067 Delhi, India

Introduction

Cryptosporidium species is the unicellular, microscopic, spore-forming, obligate intracellular organism of the apicomplexa family and is living like a parasite in the microvillus border of the gastrointestinal tract of a wide range of vertebrates, including man [1]. It is a common zoonotic enteric pathogen responsible for diarrheal diseases in humans and a variety of animals around the world [2]. *Cryptosporidium* has emerged as a significant human pathogen, particularly in children, elderly people, and immune-compromised patients, where an untreated infection might be life-threatening [3]. The parasite spreads through the fecal-oral route, through contaminated water or food, and often through contact with animals [3, 4]. However, previous studies suggested that cryptosporidiosis is a zoonotic as well as anthroponotic infection caused by different species, including *C. muris*, *C. felis*, *C. canis*, *C. canicolus*, *C. ubiquitum*, *C. viatorum*, *C. suis*, *C. meleagridis*, and *C. andersoni* [5–9].

The detection of oocysts of *Cryptosporidium* in stool using the Ziehl–Neelsen modified acid fast stain is generally known as Kinyoun, and it is the most commonly used conventional technique in detecting the *Cryptosporidium* species [10, 11]. So far, 30 *Cryptosporidium* species and over 70 subtypes of *Cryptosporidium* have been recognized in a variety of vertebrate hosts [7, 12]. There is a need to develop fast and high-throughput molecular techniques for the detection and identification of *Cryptosporidium* species and their subtypes infecting humans. Molecular techniques like restriction fragment length polymorphism (RFLP) analysis based upon the gene

Published online: 19 January 2017

International Journal of Mathematics And its Applications
Volume 5, Issue 4–B (2017), 151–165.
ISSN: 2347-1557
Available Online: http://ijmaa.in/

International Journal of Mathematics And its Applications

A Comparative Study of Jacobi Method and Givens Method for Finding Eigenvalues and Eigenvectors of a Real Symmetric Matrices

Research Article

Alanur Hussain Laskar[1]* and Samira Behera[2]

1 Department of Mathematics, M. H. C. M.Science College, Algapur-Hailakandi, Assam, India.

2 Department of Mathematics, Assam University, Silchar, Assam, India.

Abstract: The aim of this paper is to compare the Jacobi method and the Givens method for finding the eigenvalues and the corresponding eigenvectors of a real symmetric matrices. Finally, we have seen that with examples Givens method is non iterative and more efficient than Jacobi method, although it requires the given symmetric matrix into a tridiagonal matrix having the same eigenvalues.

Keywords: Eigenvalues and eigenvectors, Jacobi method, Givens method, Symmetric matrix, Bisection method.

1. Introduction

In literature [1–3], there exists several methods to compute the eigenvalues of a given real symmetric matrix. The eigenvalues plays important role in computer engineering and control engineering. The Jacobi method [1, 4] uses plane rotations in each step to compute the eigenvalues of a given real symmetric matrix. The rotation is applied till the off-diagonal elements zero. The principal diagonal elements are the eigenvalues of the matrix. In Givens method [1, 4] we tridiagonalise the given real symmetric matrix A by employing the orthogonal matrices. Tridiagonalise is that form in which the only non-zero elements are on the principal diagonal and the two diagonals just above and below of principal diagonal. Solving tridiagonal linear systems [5, 6] is one of the most important problems in scientific computing. It is involved in the solution of differential equations and in various areas of science and engineering applications such as control system and computer science. There are various numerical techniques available in the literature [7–9] which are useful for determining eigenvalues of a real symmetric matrices. In most of these methods, the given real symmetric matrix is converted into tridiagonal form. In this method, Sturm sequence and bisection method is used to determine the eigenvalues of a given real symmetric matrix. One of the leading methods for computing the eigenvalues of a real symmetric matrix is Givens method. In that method, after transforming the matrix into tridiagonal form say, 'S', the leading principal minors of $|S - \lambda I|$ form a Sturm sequence. Then, using bisection approach, change of sign in various Sturm sequence is observed. Further, based on this, eigenvalue can be determined by repeatedly using bisection method. In order to show the comparative result, we have considered the example for the illustration of Jacobi method and Givens method.

* *E-mail: alanur.hkd@gmail.com*

Contents lists available at ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Research paper

Compositional properties and codon usage of TP73 gene family

Parvin A. Barbhuiya[a], Arif Uddin[b], Supriyo Chakraborty[a,*]

[a] Departments of Biotechnology, Assam University, Silchar 788011, Assam, India
[b] Department of Zoology, Moinul Hoque Choudhury Memorial Science College, Algapur, Hailakandi 788150, Assam, India

ARTICLE INFO

Keywords:
TP73 gene
Codon usage bias
AT-ending codons
GC-ending codons

ABSTRACT

The TP73 gene is considered as one of the members of TP53 gene family and shows much homology to p53 gene. TP73 gene plays a pivotal role in cancer studies in addition to other biological functions. Codon usage bias (CUB) is the phenomenon of unequal usage of synonymous codons for an amino acid wherein some codons are more frequently used than others and it reveals the evolutionary relationship of a gene. Here, we report the pattern of codon usage in TP73 gene using various bioinformatic tools as no work was reported yet. Nucleotide composition analysis suggested that the mean nucleobase C was the highest, followed by G and the gene was GC rich. Correlation analysis between codon usage and GC3 suggested that most of the GC-ending codons showed positive correlation while most of the AT-ending codons showed negative correlation with GC3 in the coding sequences of TP73 gene variants in human. The CUB is moderate in human TP73 gene as evident from intrinsic codon deviation index (ICDI) analysis. Nature selected against two codons namely ATA (isoleucine) and AGA (arginine) in the coding sequences of TP73 gene during the course of evolution. A significant correlation ($p < 0.05$) was found between overall nucleotide composition and its composition at the 3rd codon position, indicating that both mutation pressure and natural selection might influence the CUB. The correlation analysis between ICDI and biochemical properties of protein suggested that variation of CUB was associated with degree of hydrophobicity and length of protein.

1. Introduction

TP73 gene encodes a protein that is related to the p53 tumor protein. The TP73 gene has high homology with the tumor suppressor p53 as well as with p63 (Tomkova et al., 2007). The TP73 gene in human is characterized by the presence of different isoforms or transcript variants of the protein created by alternative splicing and/or use of alternate promoters (Engelmann et al., 2015). Among all TP73 transcript variants, two major forms are transcriptionally active *i.e.* TAp73 (full-length version) and inactive DNp73 isoforms (Engelmann et al., 2015). Isoforms containing the transactivation domain are pro-apoptotic. But the amino-terminally truncated version DNp73 is an antiapoptotic gene and blocks the function of p53 and transactivates p73 isoforms as well (Di et al., 2013). Since DNp73 overrides the positive effects of full-length TAp73 hence it possesses a decisive role in cancer pathogenesis (Melino et al., 2002; Engelmann et al., 2015; Lucena-Araujo et al., 2015; Sabapathy, 2016). The imbalance between DNp73 and TAp73 may lead to tumorigenesis and resistance to chemotherapy in human cancers (Lucena-Araujo et al., 2015). TP73 plays the crucial role in the prognosis of many cancers as it is associated with the sensitivity of cancer cells and chemotherapy (Min et al., 2015). However, TP73 is infrequently mutated in cancers (Wang et al., 2007). It regulates many biological phenomena like cell cycle arrest, apoptosis, neurogenesis, immunity and inflammation. Lately, it is reported that TP73 is associated with metabolic disorder including diabetes and cancer. It is also relevant to hepatocellular lipid metabolism, glutathione homeostasis and the pentose phosphate pathway by regulating selective metabolic enzymes. TP73 gene is also found to regulate basal and starvation-induced fuel metabolism in the liver (He et al., 2015). Study on TP73-deficient mice has shown novel roles for p73 in neurogenesis, sensory pathways and homeostatic control (Tomkova et al., 2007). This gene is involved in various cancers such as pancreatic cancer (Rödicker et al., 2001), leukemia and lymphomas (Coates, 2006), hepatocellular carcinoma (Stiewe et al., 2004) and breast carcinoma (Domínguez et al., 2001).

Previous studies on mutated TP53 cancerous cell in human demonstrated that a significant increase in translation efficiency (TE) might take part in the selection of TP53 cancerous mutations. Moreover, coadaptation of mutations and the tRNA pool may proliferate the overall TP53 translational efficiency (Waldman et al.,

Abbreviations: ENC, effective number of codons; RSCU, relative synonymous codon usage; CUB, codon usage bias

* Corresponding author.
E-mail address: supriyoch_2008@rediffmail.com (S. Chakraborty).

https://doi.org/10.1016/j.gene.2018.10.030
Received 6 April 2018; Received in revised form 3 October 2018; Accepted 11 October 2018
Available online 12 October 2018