Experimental Research

ANNALS OF MEDICINE & SURGERY

OPEN

Molecular insights into codon usage analysis of mitochondrial fission and fusion gene: relevance to neurodegenerative diseases

Rekha Khandia, PhD[a,*], Megha Katare Pandey, PhD[c], Rajkumar Garg, PhD[b], Azmat Ali Khan, PhD[e], Igor Baklanov, PhD[f], Amer M. Alanazi, PhD[e], Prakash Nepali, MD[g,*], Pankaj Gurjar, PhD[h,i,*], Om Prakash Choudhary, PhD[d]

Abstract

Mitochondrial dysfunction is the leading cause of neurodegenerative disorders like Alzheimer's disease and Parkinson's disease. Mitochondria is a highly dynamic organelle continuously undergoing the process of fission and fusion for even distribution of components and maintaining proper shape, number, and bioenergetic functionality. A set of genes governs the process of fission and fusion. *OPA1, Mfn1*, and *Mfn2* govern fusion, while *Drp1*, *Fis1*, *MIEF1*, and *MIEF2* genes control fission. Determination of specific molecular patterns of transcripts of these genes revealed the impact of compositional constraints on selecting optimal codons. AGA and CCA codons were over-represented, and CCC, GTC, TTC, GGG, ACG were under-represented in the fusion gene set. In contrast, CTG was over-represented, and GCG, CCG, and TCG were under-represented in the fission gene set. Hydropathicity analysis revealed non-polar protein products of both fission and fusion gene set transcripts. AGA codon repeats are an integral part of translational regulation machinery and present a distinct pattern of over-representation and under-representation in different transcripts within the gene sets, suggestive of selective translational force precisely controlling the occurrence of the codon. Out of six synonymous codons, five synonymous codons encoding for leucine were used differently in both gene sets. Hence, forces regulating the occurrence of AGA and five synonymous leucine-encoding codons suggest translational selection. A correlation of mutational bias with gene expression and codon bias and GRAVY and AROMA signifies the selection pressure in both gene sets, while the correlation of compositional bias with gene expression, codon bias, protein properties, and minimum free energy signifies the presence of compositional constraints. More than 25% of codons of both gene sets showed a significant difference in codon usage. The overall analysis shed light on molecular features of gene sets involved in fission and fusion.

Keywords: Drp1, Fis1, mitochondrial dysfunction, mitochondrial fission and fusion, neurodegeneration, OPA1

Introduction

Mitochondria are also called cell powerhouses and produce ATP as the end product of a series of pathways involved in substrate oxidation. It also plays crucial functions, including amino acids' and steroids' biosynthesis, β-oxidation of fatty acids, and cytosolic calcium homoeostasis. In addition, it acts as a sensor for oxidative stress and plays a role in pathways like necrosis, apoptosis, and autophagy[1]. There are various diseases linked with mitochondrial dysfunction starting from common diseases like glaucoma, inflammation, neurodegenerative diseases, type 2 diabetes, cancers of the prostate and colon, cardiomyopathies, and dysrhythmias to less-known diseases like Freiderich's ataxia, Kearns-Sayre syndrome, Leber hereditary optic neuropathy,

Departments of [a]Biochemistry and Genetics, [b]Bioscience, Barkatullah University, [c]Translational Medicine Center, All India Institute of Medical Sciences, Bhopal, [d]Department of Veterinary Anatomy, College of Veterinary Science, Guru Angad Dev Veterinary and Animal Sciences University (GADVASU), Rampura Phul, Bathinda, Punjab, India, [e]Pharmaceutical Biotechnology Laboratory, Department of Pharmaceutical Chemistry, College of Pharmacy, King Saud University, Riyadh, Saudi Arabia, [f]Department of Philosophy, North Caucasus Federal University, Stavropol, Russia, [g]Government Medical Officer, Bhimad Primary Health Care Center, Government of Nepal, Tanahun, Nepal, [h]Centre for Global Health Research, Saveetha Medical College and Hospital, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, Tamil Nadu, India and [i]Department of Science and Engineering, Novel Global Community Educational Foundation, Hebersham, NSW, Australia

Sponsorships or competing interests that may be relevant to content are disclosed at the end of this article.

**Corresponding authors. Address: Government Medical Officer, Bhimad Primary Health Care Center, Government of Nepal, Tanahun 33910, Nepal, E-mail: dr.nepali.prakash@gmail.com (P. Nepali); Centre for Global Health Research, Saveetha Medical College and Hospital, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, Tamil Nadu 602117, India, and Department of Science and Engineering, Novel Global Community Educational Foundation, Hebersham, New South Wales, Australia. E-mail: pankajgurjar0103@gmail.com (P. Gurjar); Department of Biochemistry and Genetics, Barkatullah Universty, Bhopal, MP 462026, India. E-mail: bu.rekha.khandia@gmail.com (R. Khandia).*

Annals of Medicine & Surgery (2024) 86:1416–1425

Received 14 September 2023; Accepted 5 January 2024

Published online 15 January 2024

http://dx.doi.org/10.1097/MS9.0000000000001725

mitochondrial encephalopathy, lactic acidosis and strokes, and mitochondrial neuro-gastrointestinal encephalomyopathy[2]. Considering the high energy demand of neurons and limited regenerative capacities, impaired neurons' functioning might be detrimental to neuronal survival, and considerable evidence is present in support of neurodegeneration due to mitochondrial dysfunction[3]. Mitochondria can encode 13 polypeptides out of 92 required for oxidative phosphorylation. Nuclear DNA encodes other structural and assembly proteins. Mutations in mitochondrial or nuclear DNA, affecting oxidative phosphorylation, are highly detrimental to the tissues with high energy demands, including the central nervous system, skeletal muscle, and heart. Mitochondrial mutations are also evident in age-related neurodegenerative diseases such as Alzheimer's disease (AD) and Parkinson's disease (PD). Mitochondria is a highly dynamic organelle that continuously fuses and divides to ensure the even distribution of metabolites and mitochondrial DNA and decide the proper shape, number, and bioenergetic functionality. Multiple pieces of evidence suggest the involvement of mitochondrial dysfunction in early and casual neurodegeneration. Mitochondrial compartmentalization takes place by the fusion of the inner and outer mitochondrial membranes, and the proteins involved are outer membrane GTPases Mitofusins (*Mfn1* and *Mfn2*) and the inner membrane GTPase Optic atrophy 1 (*OPA1*)[4]. Dynamin-related protein 1 (*Drp1*/ *DNM1L*) and fission protein 1 (*Fis1*) are the key players in mitochondrial fission[5]. During the process of fission, other factors like mitochondrial fission factor (*MFF* gene), mitochondrial dynamics protein 49 (*Mid49*, *MIEF2* gene), and Mid51 (*MIEF1* gene) also support[6].

Unbalanced fusion results in mitochondrial elongation, while unbalanced fission results in the formation of tiny mitochondria. Dysregulated fusion-fission dynamics of mitochondria are involved in AD, PD, Huntington's disease (HD), and amyotrophic lateral sclerosis (*ALS*)[7]. Mutations in the *OPA1* gene lead to a rare inherited dominant optic atrophy (DOA) culminating in visual failure, deafness, encephalomyopathy, peripheral neuropathy, ataxia, and cardiomyopathy[8], and mutations in OPA1 account for at least 45% of all DOA cases[9]. Mice knocked out for *Mfn2* in mice's hippocampus and cortex region resulted in neurodegeneration, and the associated pathological changes were similar to those present during AD progression[10]. *Mfn2* is essential to postnatal cerebellum development[11]. *Mfn1* is required for membrane tethering, and functions of *Mfn1* and *Mfn2* overlap, and the fusogenic activity of *OPA1* requires *Mfn1*[12,13]. *Fis1* is a player of fission and is significantly enhanced in AD patient-derived fibroblasts. Mutation in *DRP1* is the cause of severe neurodevelopmental syndrome[14]. There is strong evidence of the genetic role behind the association of neurodegenerative disorders with mitochondrial dysfunctions. Multiple groups of researchers study and document the genes involved in fusion and fission; however, the genes have never been studied from the perception of codon usage. We were tempted to do this study because such studies on codon usage for fusion and fission-related genes are absent. The analysis will also help improve our knowledge of gene expression related to the said gene sets. The analysis will also help augment the associated diseases by modulating the gene expression using the knowledge of synonymous codon usage. Codon usage analysis gives insight into the gene expression level[15] and is linked with various protein properties, including gene length[16], gene composition[17], and nucleotide skew[18]. In the present study, we attempted to analyze the codon usage pattern of gene sets involved in mitochondrial fission and fusion and analyzed the codon that maximally influences codon usage. We also compared codon usage between the fission and fusion gene sets and tried to find the differences (if any). We also investigated the various parameters like overall nucleotide composition, nucleotide composition at different codon positions, gene expression level, nucleotide skew, and correlation among various gene parameters. The study is relevant to understanding various unanswered questions, like what forces alter the protein properties or modulate and fine-tune the gene expression or how the length of a gene can affect the nucleotide composition of a gene. Apart from differences at the molecular level in the fusion and fission gene set, we also tried to investigate the commonness between the two opposite processes. To investigate the same, we analyzed various parameters, including nucleotide composition, codon adaptation analysis (CAI), effective numbers of codons (ENc), relative synonymous codon usage (RSCU), and minimum free energy (MFE) of the transcripts. We then used various statistical tools to understand the relationship between different parameters. We investigated the most influential codons influencing the CUB in fission and fusion gene sets and tried to investigate the reason behind fine-tuning for the occurrence of specific codons. Furthermore, the study will provide insight into the mechanism of gene regulation, functions, and specific molecular features associated with genes responsible for mitochondrial fission and fusion related to neuronal health.

HIGHLIGHTS

- Mitochondrial dysfunction is the leading cause of neurodegenerative disorders like Alzheimer's disease, and Parkinson's disease. Determination of specific molecular patterns of transcripts of these genes revealed the impact of compositional constraints on selecting optimal codons.
- AGA codon showed a unique pattern of relative synonymous codon usage exhibiting imperative implications in translational regulation.
- Out of six synonymous codons encoding for leucine, five synonymous codons were used in a statistically significantly different manner in fission and fusion gene sets.
- More than 25% of codons of both gene sets showed a significant difference in codon usage.

Methods

Sequence retrieval

The gene set responsible for mitochondrial fusion contained *Mfn1* and Mf*n*2 and the *OPA1* gene, while the gene set responsible for mitochondrial fission contained *Drp1*/ *DNM1L*, *MFF*, *MIEF1*, and *MIEF2* genes. The transcripts were retrieved from the NCBI nucleotide repository for each gene. Both the confirmed isoforms and predicted isoforms were included in the study. Fourteen transcripts were studied for fusion-related gene sets encompassing 11, 01, and 02 transcripts, respectively, for *OPA1*, *MFN1*, and *MFN2* and will be called fusion gene set hereafter. Similarly, 45 transcripts were studied for fission-related gene sets encompassing 01, 22, 04, 04, and 14 transcripts for *Fis*, *MFF*,

MIEF1, *MIEF2*, and *DNM1L* genes, respectively, and will be called the Fission gene set hereafter. All the transcripts were triplet initiating with the start codon and terminating with stop codons. Any duplicate or ambiguous nucleotide-containing transcripts were avoided.

Compositional analysis

The amino acid and codon usage depends on the nucleotide composition and the variance in codon response to overall GC content. We performed a compositional analysis[19]. Overall nucleotide composition (%A, %T, %C, %G) and nucleotide composition at each of the three codon positions (%A1, %A2, %A3, %T1, %T2, %T3, %C1, %C2, %C3, %G1, %G2, %G3) was determined. Overall %GC composition and %GC composition at all three codon positions (%GC1, %GC2, %GC3) were also determined using CAIcal software developed by Puigbò and colleagues[20].

Determination of mutational bias

In the case of identical mutation rates in complementary strands, there will be A = T and G = C in complementary strands. Such deviations are the results of the processes of transcription and replication, both of which differentiate between complementary DNA strands. Therefore, the bias is measured as $B = [(G + T)-(A + C)]/(A + C + G + T)$, and if mutational machinery is absent, the value will come to zero[21].

Determination of nucleotide skew

The nucleotide skews help describe the overall pattern of nucleotide composition of genes[22]. An analysis by Green *et al.*[23] indicated that asymmetry in the frequencies of substitutions on the coding and non-coding strands of genes results in nucleotide-content asymmetry. Also, transcription-coupled repair affects both the AT and GC skew. An example includes the deamination of cytosine that results in GC skew due to depletion in cytosine and an enhancement in the thymidine pool, so ultimately, AT skew is also affected[24]. AT and GC skews affect gene expression levels in human genes[24], and a positive association between nucleotide skew and gene expression has been documented[25]. The formula used was (A-B)/(A + B), where A and B were respective nucleotides for each skew. AT and GC skews and other skews like purine, pyrimidine, amino, and keto skews were also calculated. Furthermore, parity analysis indicated the nucleotide disproportion at the third codon position. AT and GC bias is calculated using the formula {AT bias [A3/(A3 + T3)] and GC bias [G3/(G3 + C3)]}[26].

Odds ratio analysis

There are four individual nucleotides, and their combination makes 16 dinucleotides. Due to compositional differences, dinucleotides are absent in any gene in the expected ratio. The expected to the observed ratio of dinucleotide is called the odds ratio, and the ratio was calculated for all the transcripts present in the Fusion and Fission gene set. An odds ratio above 1.23 and below 0.78 are considered over and underrepresenting dinucleotides, respectively[27].

Relative synonymous codon usage calculation (RSCU)

A total of 64 codons are present; out of them, three are stop codons, while methionine and tryptophan are the amino acids encoded by a single codon. So apart from these, there are 59 codons, into which two or more two codons are coding for a single amino acid. Such codons are called synonymous codons. RSCU value indicates the relative frequency compared to the frequency of another codon coding for the same amino acid. RSCU of genes was determined using CAICal software developed byPuigbò and colleagues[20].

Codon adaptation index (CAI) analysis

The CAI is an index primarily used to predict gene expression level since it indicates the similarity of codon usage between the given gene set and reference gene set and to what extent a coding sequence represents the codon usage of any organism[28]. The CAI value ranges between 0 and 1, indicating higher selection towards the optimal codons to reach the efficient transcriptional level[29].

Codon usage bias (CUB)

The effective number of codons (ENc) is a non-directional measure of CUB. The highest and lowest values for ENc are 20 and 61, respectively. Value 20 indicates the highest bias, while 61 indicates the lowest[30]. An ENc value of 20 results from the highest bias when only one codon is used among all synonymous codons. In contrast, a value of 61 is obtained when all the synonymous codons are used equally to encode amino acids[31]. ENc values above 35 are considered moderately biased, and genes with ENc less than 35 are called biased genes.

General average hydropathicity (GRAVY) and aromaticity (AROMA)

Both of the properties GRAVY and AROMA are indicative of natural selection. The GRAVY and AROMA signify the frequencies of hydrophobic and aromatic amino acids, respectively, and indicators of amino acid usage[32]. Both the values were calculated using software COUSIN software[33].

Minimum free energy calculations

Due to attaining secondary and tertiary structure, RNA transcript possesses energy in its structure. When unfolded, a secondary structure releases free energy during transcription. The greater the minimum free energy, the higher the stability of RNA, and the less it is likely to be translated. The RNAfold program developed by Hofacker *et al.*[34] was used to determine the minimum free energy. The minimum free energy of the isoforms was calculated using the RNAfold web server (http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi). More bias is generally associated with high negative free energy.

Statistical analysis

For statistical analysis, Minitab version 17 was used. For paired *t*-test, Tukey's *t*-test was performed. PCA biplot analysis was done in ggplot2 software.

Results and discussion

The aetiology of neurodegenerative disorder remained largely elusive: how neurons are lost in disorders whose prevalence increases with an increase in age. Scientists have found an association between the mitochondrial function to answer this question. Mitochondria are a cell's powerhouse, and neurons are highly energy-consuming since they rely heavily on mitochondria for energy and perform specialized functions like membrane ion pumps, channel activity, and synaptic transmission. Mitochondria continuously undergo fission and fusion to maintain the appropriate number, morphology, mtDNA concentration, and subcellular organelle distribution and function[35,36]. Furthermore, mitochondrial fission is essential in dividing cells to ensure the rightful inheritance in daughter cells. The fission process is controlled by *Fis1*, *MFF*, *MIEF*, *MIEF2*, and *Drp1* genes[6], while fusion is controlled by *OPA1*, *Mfn1* and *Mfn2* support[4].

Considering the imperative role of mitochondrial dysfunction contribution to AD, HD, and PD, owing to the malfunction of genes, molecular insights into the genes' attributes, including nucleotide composition, nucleotide disproportion, protein properties, gene expression, and codon usage dinucleotide occurrences, transcript free energy, and mutational bias are envisaged. Since both the fission and fusion processes are essential to maintain appropriate mitochondrial dynamics, in the present study, we studied the gene sets encompassing the genes involved in the process of mitochondrial fission and fusion. The study helped determine the molecular features of both gene sets and compare the features to envisage similarities and dissimilarities. The information regarding codon usage and other molecular patterns will help us find the correlation between molecular patterns and functional aspects. Furthermore, codon usage analysis might help in gene augmentation and restoring the proper functioning of the gene by codon correction.

Compositional factors affect codon usage

Nucleotide A was most abundant in the fusion and fission gene sets, with mean values of 28.76 ± 5.71 and 30.04 ± 4.66, respectively. On the other hand, nucleotide C was least abundant in both the gene sets (22.63 ± 3.49 and 20.82 ± 4.29, respectively). Overall analysis indicated that the transcripts were AT-rich, and the %GC was 47.32 ± 7.84 and 45.35 ± 7.01 for fission and fusion gene sets. As a general rule, Hershberg and Petrov DA[37] proposed that the nucleotide composition affects the codon usage. In AT-rich genomes, optimal codons are AT-ending, and vice versa for GC-rich genomes. The same has been observed in the work of Majeed and colleagues[38], who analyzed RNAseq data from 93 790 assembled transcripts and found that in the AT-rich genome of Taxus contorta, preferred codons tend to end with A/T. In contrast, the avoided codons tend to end with G/C. Our results are in accordance with the rule, and since the transcripts are slightly AT-rich, optimal codons are AT-ending codons. The analysis revealed that compositional constraints affect the usage of optimal codons. Furthermore, in both the gene sets, the %GC composition was highest at the first codon position in comparison to the second and third codon positions, and our result is in concordance with the results of Choudhury *et al.*[39], who found a similar trend in SRY gene responsible for maleness in mammals.

Codon usage pattern analysis revealed variation in codon usage under compositional and selection forces

Codon usage pattern varies according to composition and other factors like mutational drift and selection pressure. RSCU-based matrix plots were constructed to see the overall trend of codon usage (Fig. 1A and B for fission and fusion gene sets, respectively). The figures show that GC-ending codons are under-represented or randomly used in the fusion and fission gene sets while AT-ending codons are generally over-represented or randomly used (discussed in the above section). In the fusion gene set, gene *MFN2* and in the fission gene set *FIS1*, *MIEF1*, and *MIEF2* genes displayed opposite properties, and AT-ending codons were either under-represented or randomly used. The same could be explained by the fact that in the fission gene set, in *FIS1*, *MIEF1*, and *MIEF2* transcripts, nucleotides A and T are dominant over nucleotides A and T. However, on the same basis, the under-representation of A and T cannot be explained in the MFN2 gene since, despite the abundance of A and T nucleotide (%A = 32.43 and %T = 27.40), AT-ending codons are either under-represented or randomly used. The results suggest the presence of other forces than compositional constraints acting on the MFN2 gene. The CTG codon has been found overexpressed in obesity, housekeeping[40], central nervous system genes[41], and Y-linked genes[42]. The over-representation of CTG can be understood because CpG dinucleotide is predisposed to mutation through the deamination of 5-methylcytosine, resulting in C→T transition. There is a report of 11 times more NCG mutating frequency to NTG reported in *BRCA1* or *BRCA2* mutant chicken DT40 cell line model[43]. A mutation that leads to the conversion of CpG to TpG is possibly the reason for the over-representation of CTG in many genes of eukaryotes. Hence, overexpression of CTG is a result of mutational forces. Our results concord with the results of other researchers, and codon CTG was over-represented in all fission transcripts except *DNM1L*, which is randomly used.

Similar to our result, the CTG codon is over-represented in more than 80% of genes involved in both primary immunodeficiency and cancer[44]. In brief, the over-representation of CTG in the fission gene set might be understood based on C→T transition in hypermutable CpG dinucleotide. In contrast, the over-representation of AGA in the fusion gene might be understood by its implication in transcriptional regulation. AGA and CCA codons were over-represented, and CCC, GTC, TTC, GGG, and ACG were under-represented in all transcripts of the fusion gene set except *MFN2*. Furthermore, GCG, CCG, and TCG were under-represented in a complete fusion gene set. Our study found the codon usage pattern exhibited by codon AGA very interesting. Within the fission gene set, AGA remarkably exhibited a different pattern. It was over-represented in all transcripts except *Fis1*, *MIEF1*, and *MIEF2* genes, where it is under-represented. On the other hand, in the fusion gene set, it is over-represented in all transcripts except the *MFN* gene transcripts. *MFN* gene produces two transcripts mentioned here as *MFN1* and MFN2. Between these two transcripts, codon AGA is over-represented in one transcript, while in the other, it is under-represented. We might link the specific behaviour of codon AGA with the fact that AGA codon runs have imperative implications in translational regulation by tRNAmethyltransferase[45]. Hence, the AGA codon number is meticulously maintained differently in different transcripts to maintain the optimal translational regulation. In the present study, within fusion and fission gene sets, both

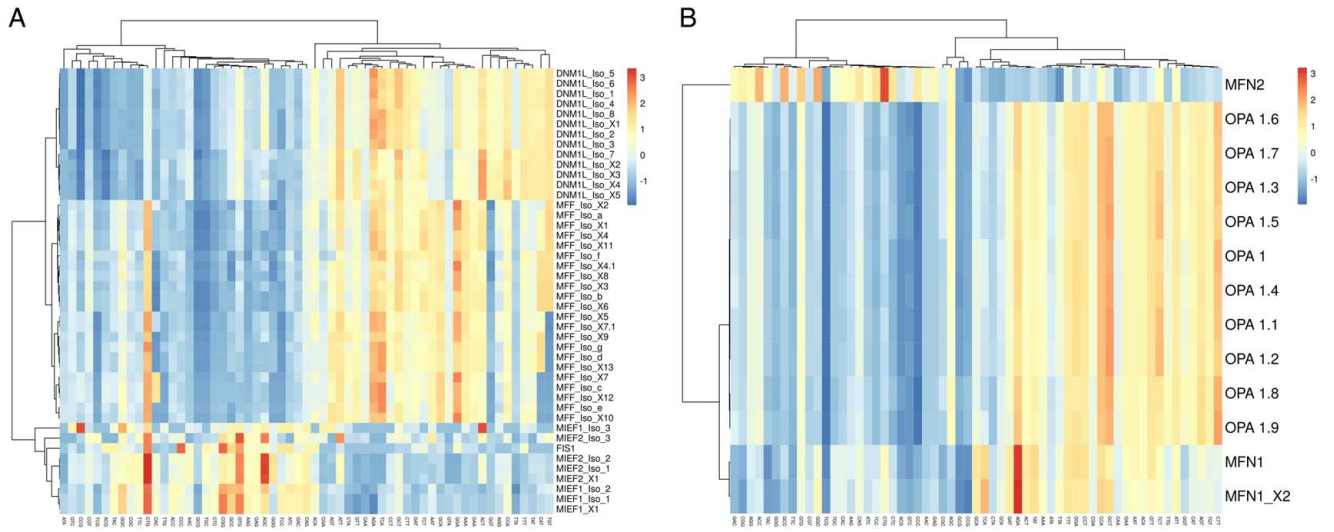

Figure 1. (A) Clustering of fission gene multivariate data based on relative synonymous codon usage (RSCU) values. (B) Clustering of fusion gene multivariate data based on RSCU values.

underrepresentation and over-representation of AGA codon indicate the presence of highly selective forces operating behind to keep the translation of these proteins in a highly regulated manner.

Biplot analysis to determine most influential codons influencing codon usage

We did a PCA biplot analysis to determine the codons that influence the codon usage the most, and the analysis revealed that CTG, TGT, GGT, and GTG codons have maximum loading values for fission gene sets among PC1 and PC2. Similarly, TCA, AGA, CTG, and CTT codons have maximum loading values for the fusion gene set among PC1 and PC2. Overall analysis revealed that the CTG codon is one of the most influential codons that affect codon usage in both the gene sets (Fig. 2A and B for fission and fusion gene sets, respectively). We compared codon usage between Fusion and fission gene sets (Table 1). Table 1 shows that despite being the most influencing codon, in both fission and Fusion gene sets, there is a statistically significant difference in codon usage ($P < 0.001$) for CTG. There were 15 codons out of 59 that significantly differed in codon usage in fission and Fusion gene sets. Other than CTG are TTA, CGA, CCA, TCA, CTA, ATA, GTT, CTT, TCT, GGT, CCC, GTC, ACC, and TTG. Hence, more than 25% of codons are statistically different in codon usage between fission and fusion gene sets. Furthermore, an interesting feature was observed that out of 15 codons showing significant differences in codon usage, 10 codons were A/T ending. The difference in codon usage pattern has already been demonstrated in self-renewing and differentiating human embryonic stem cells during differentiation[46].

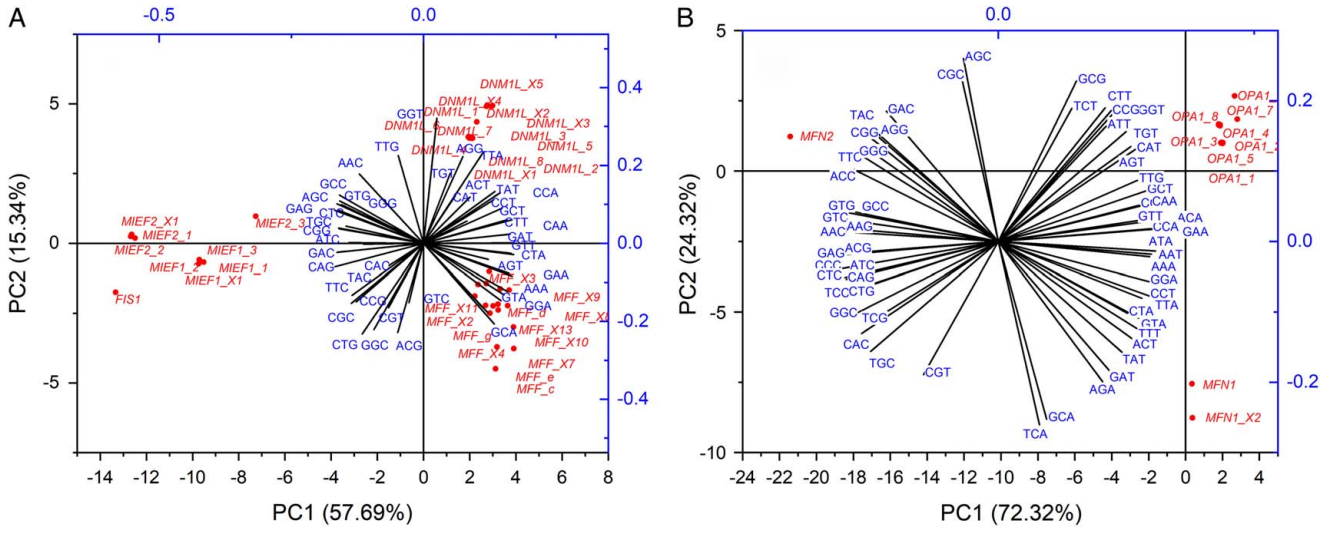

Figure 2. (A) PCA biplot for fission gene set revealing that CTG codon is having a maximum influence on codon usage. (B) PCA biplot analysis for fusion gene set revealed that TCA codon is having a maximum influence on codon usage.

Table 1

Result of paired Tukey's *t*-test analysis

S. No.	Codons	Fusion and FuFission	Fusion and HK	Fission and HK
1.	AAA	NS	**	***
2.	AGA	NS	*	*
3.	TTA	**	***	***
4.	GGA	NS	*	***
5.	CGA	*	**	NS
6.	GAA	NS	***	***
7.	GCA	NS	NS	**
8.	CCA	*	***	**
9.	TCA	**	NS	***
10.	CTA	***	NS	***
11.	ACA	NS	***	***
12.	ATA	**	***	**
13.	GTA	NS	***	***
14.	CAA	NS	***	***
15.	TTT	NS	***	**
16.	GTT	*	***	***
17.	CTT	**	***	***
18.	CAT	NS	NS	NS
19.	CGT	NS	NS	NS
20.	TGT	NS	NS	NS
21.	ATT[#]	NS	*	***
22.	GAT[#]	NS	**	***
23.	TAT[#]	NS	**	NS
24	GCT[#]	NS	**	**
25.	TCT[#]	*	*	NS
26.	ACT	NS	NS	*
27.	AAT	NS	***	***
28.	GGT	*	*	NS
29.	AGT	NS	**	**
30.	CCT	NS	NS	NS
31.	CCC	*	***	**
32.	TCC	NS	**	***
33.	ATC	NS	***	***
34.	GTC	**	***	NS
35.	AGC	NS	NS	NS
36.	CAC	NS	NS	**
37.	GGC	NS	*	NS
38.	CGC	NS	NS	*
39.	GAC	NS	**	***
40.	CTC	NS	*	***
41.	TTC	NS	***	**
42.	TAC	NS	**	*
43.	ACC	**	**	***
44.	AAC	NS	***	***
45.	TGC	NS	NS	***
46.	GCC	NS	**	***
47.	GGG	NS	**	***
48.	GCG	NS	NS	**
49.	CTG	***	***	NS
50.	CAG	NS	***	***
51.	AAG	NS	**	***
52.	ACG	NS	NS	NS
53.	GTG	NS	*	***
54.	TCG	NS	*	NS
55.	CCG	NS	NS	NS
56.	GAG	NS	***	***
57.	CGG	NS	NS	NS
58.	AGG	NS	NS	NS
59.	TTG	***	NS	***

[#]Five leucine-encoding codons presenting statistically significant codon usage between fission and fusion gene are presented as.
NS non-significant.
*Here, $P < 0.05$.
**$P < 0.01$.
***$P < 0.001$.

An analysis based on 1625 genomes, including over 14 million sequences, revealed that in archaea, bacteria, and eukaryotes, a single amino acid, arginine, is the major contributor to codon usage bias differences across these three domains of life speculated that domain-specific preference for arginine codons is linked with translation speed, which supports the notion of the presence of selective forces responsible for codon usage variation across genomes[47]. In the present study, we have five leucine-encoding codons out of six synonymous codons (TTA, CTT, CTA, CTG, TTG) that were quite differently used in fission and Fusion gene sets. Leucine is an integral part of the leucine zipper family (60–80 amino acid long protein domain), allowing for sequence-specific DNA binding and faster gene expression (Table 1)[48]. Hence, leucine codon usage appears tightly associated with gene expression regulation and is affected by translational selection.

The correlation of CUB with aromaticity and hydropathicity in fusion and fission gene set

Aromaticity and hydrophobicity determine the variation in codon usages[49]. We correlated CUB and protein properties hydropathicity and aromaticity in the fusion and fission gene set. In the fusion gene set, CUB negatively correlated with GRAVY ($r = -0.778$, $P < 0.001$), with no correlation with AROMA. CUB negatively correlated with AROMA ($r = -0.578$, $P < 0.001$) for the fission gene set but not with GRAVY. The correlation between codon bias and factors including aromaticity, aliphatic index, and hydropathy suggests natural selection acting on codon usage patterns[51,52]. The present study indicated the correlation of CUB with hydropathicity and aromaticity for fission and fusion gene sets, indicating that selection pressure is operative on codon usage in these genes.

Mutational force is dominant for the Fission gene set while selection for fusion

Mutation bias is when specific mutations occur more often than expected. Mutational bias is a product of evolution[53] and the evolutionary forces of mutation, gene flow, genetic drift, and natural selection. Elevated mutation rates are present in the extensively expressed gene; hence, there is an apparent positive correlation between mutational bias and gene expression[54]. The mutational bias for each set was calculated, and correlation analysis was done. A strong and significant Pearson correlation between mutational bias and expressivity ($r = 0.713$, $P < 10\text{-}8$) was found in the fission gene set. At the same time, there was no correlation between mutational bias and expressivity for the fusion gene set. In a set of ubiquitously expressed housekeeping genes ($n = 374$), a highly significant positive correlation between expression and the mutational bias (B) ($r = 0.28$; $P < 10 - 7$) was reported[24]. This correlation indicates the effect of selection pressure on the mutational bias. The present study shows that higher mutational force is acting on the fission gene set, while selection pressure is operating more in the fusion gene set, as evidenced by the mutational bias study.

Overall nucleotide compositional bias is different from bias at the third codon position

Overall, nucleotide composition determines the codon composition and composition, influencing proteins' physical properties. For example, codons with fewer Gs or Cs encode more hydrophobic amino acids[55]. Hence, compositional bias is an important aspect of studying the nature of protein encoded by any gene or transcript. Nucleotide skews can be subdivided into AT skew (between A and T), GC skew (between G and C), purine skew (between A and G), pyrimidine skew (between C and T), amino skew (between A and C) and keto skew (between T and G). The positive GC skew reveals G's richness over C, while the negative

Table 2

Correlation analysis of GC3 with other compositional parameters

	Fusion gene set				Fission gene set			
	Mean	STDEV	Pearson r value	*P*	Mean	STDEV	Pearson r value	*P*
Gene length	1179.623	522.6552	−0.51604	*	2523.6	356.6324	0.030412	NS
%A	28.76399	5.71197	−0.99467	***	30.04938	4.666742	−0.94351	***
%C	22.63508	3.497722	0.99584	***	20.8291	4.298483	0.90754	***
%T	23.90825	3.387046	−0.97604	***	24.59927	2.42886	−0.89219	***
%G	24.69268	12.95034	0.98954	***	24.52225	2.760818	0.93933	***
%G + C	47.32776	7.848387	0.99959	***	45.35135	7.015404	0.95445	***
%A1	28.11684	4.454528	−0.9708	***	29.89831	2.005982	−0.87736	***
%C1	26.09572	6.917679	0.96326	***	22.4986	2.529929	0.6776	***
%T1	13.45135	9.73207	−0.70528	**	16.26493	1.181649	0.27909	*
%G1	32.33609	13.72733	0.51336	*	31.33816	0.669756	0.56376	***
%G1 + C1	58.43181	5.107576	0.9845	***	53.83676	2.824593	0.77852	***
%A2	29.423	4.298914	−0.8636	***	36.19248	2.439331	−0.50523	***
%C2	23.69721	3.511072	0.38999	NS	18.50105	1.181175	−0.07012	NS
%T2	28.30127	5.464038	0.70018	**	29.40749	1.413467	0.13592	NS
%G2	18.57852	12.51249	0.55822	*	15.89898	1.17567	0.58224	***
%G2 + C2	42.27573	4.727453	0.91582	***	34.40003	1.219586	0.2938	*
%A3	28.75212	10.74327	−0.99963	***	24.05735	9.875308	−0.97623	***
%C3	18.11231	10.40818	0.9997	***	21.48764	9.947178	0.98659	***
%T3	29.97213	8.783096	−0.99932	***	28.12539	7.27323	−0.96554	***
%G3	23.16344	16.16978	0.99943	***	26.32962	7.199847	0.98363	***
CAI_59	0.703019	0.046545	0.99733	***	0.7286	0.064424	0.98215	***
Effective number of codon	47.43738	4.306638	−0.9654	***	49.95885	2.545394	−0.50486	***
GRAVY	0.358	0.191584	0.64646	**	−0.4526	0.146626	0.68599	***
AROMA	−0.38455	0.012369	0.16781	NS	0.0776	0.001875	0.59768	***
MFE	−357.077	178.761	−0.70977	***	−722.176	104.8033	−0.4601	***

AROMA, aromaticity; CAI, codon adaptation analysis; GRAVY, general average hydropathicity; MFE, minimum free energy; NS, non-significant.
Statistically significant $*P < 0.05$.
Statistically significant $**P < 0.01$.
Statistically significant $***P < 0.001$.

GC skew reveals the richness of C over G and vice versa[39]. In the present study, AT and GC skews generally have positive values in both gene sets, while in other skews like purine, pyrimidine, amino, and keto skews, a mixed trend was observed for both gene sets. Positive values of AT and GC skew indicated that the overall richness of G is indicated over C and A is over T nucleotide in both gene sets. The results suggest the presence of nucleotide compositional bias in both the fission and fusion gene sets. The mean value of AT bias at the third codon position was 0.48 ± 0.05 and 0.44 ± 0.05 for fission and fusion gene sets, respectively. A bias value greater than 0.5 indicates a preference for purine over pyrimidine and vice versa[56]. Hence, T will be preferred over A in both the gene sets at the third codon position. The value of GC bias was 0.57 ± 0.03 and 0.56 ± 0.04 for the fission and fusion gene set, respectively, suggestive of G dominance over C. Overall analysis indicated that overall compositional disproportion favors A in comparison to T and G in comparison to C; however, the results are slightly different at the third codon position, where T is dominant over A.

Compositional bias influence gene expression, codon bias, and protein properties

Guanine and cytosine content at the third codon position (GC3) might be used as an indicator of codon bias[57] and compositional bias[58]. The average %GC3 composition was 41.26 ± 17.11 and 47.82 ± 17.13 for fission and fusion gene sets. Correlation analysis of %GC3 with various compositional parameters, codon usage, MFE, and protein properties are given in Table 2. It is evident from the table that GC3 has a significant correlation with nucleotide composition at all three codon positions except for cytosine at the second codon position in both gene sets. A significant correlation of GC3 was also present with gene expression, codon bias, MFE, GRAVY (for fusion gene set), and AROMA (for fission gene set). The results indicated that both codon and compositional biases affect gene expression, codon bias, protein properties, and the energy stored in RNA's secondary and tertiary structure in both the fission and fusion gene sets.

MFE negatively influences gene expression

Besides its role in translation rate, CUB may also determine RNA secondary structure. mRNA stability resulting from mRNA secondary structure influences protein abundance, splicing, translation rate with a halt during protein synthesis, and gene expression level[59]. MFE is used to determine the stability of mRNA molecules, and lower free energy indicates superior stability, indicating that the more stable mRNA will have a higher expression level. Unstable mRNA molecules usually harbour non-optimal codons, resulting in substantial destabilization of protein expression, while stable mRNAs contain optimal codons[60]. The mean MFE and the standard deviation are given in Table 2. The result revealed that MFE has a significant negative

correlation with the compositional and codon bias (GC3) measure for both gene sets ($r = -0.504$; $P < 0.001$ and $r = -0.709$; $P < 0.001$). A negative correlation between the MFE and extracellular *ADAMT13* (A Disintegrin-like and Metalloprotease with Thrombospondin type-1 repeats, member-13 (*ADAMT13*) protein that has a crucial role in vascular haemostasis) protein expression levels have been documented in 175 nucleotide long mRNA[61]. We also observed a similar pattern of negative correlation of gene expression and MFE in the fission (-0.398; $P < 0.01$) and fusion (-0.663; $P < 0.01$) gene set. It indicates that codon bias also increases with an increase in MFE and reduces gene expression. MFE negatively influences gene expression. Regions of mRNA having high minimum free energy are expected to have a weak secondary structure; however, during evaluation, the structure of mRNA is not altered[62].

Overall codon bias is moderate

Codon usage pattern of a set of other genes involved in the central nervous system[41], in neurodegeneration with iron accumulation[63], and in ovarian cancer[64] showed low bias. The average ENc value was 47.44 ± 4.30 and 49.95 ± 2.54 for the fission and fusion gene set. The values suggest moderate codon bias in both the fission and fusion gene sets.

Conclusion

Mitochondria are ATP producers and have a role in pathways like necrosis, apoptosis, and autophagy. Neurons are highly energy-demanding cells with limited regenerative capability; hence, malfunctioning mitochondria might be detrimental to neuronal survival. Evidence is there in support of the linkage of neurodegeneration and mitochondrial dysfunction. Fission and fusion are critical processes to maintain the even distribution of metabolites and genetic material and maintain proper shape, number, and bioenergetic functionality. In the present study, we envisaged the codon usage for the genes associated with fission and fusion to see the molecular signatures associated with the gene set in these two imperative processes. The presence of over-represented or randomly used AT-ending codons in marginally AT-rich transcripts suggests the action of compositional constraints on codon usage. Based on the PCA biplot, the CTG codon was found to be one of the most influencing codons affecting codon usage in both the gene sets, which is in corroboration with other researchers' data who report CTG as one of the most influencing codons in many genes and gene sets. More than 25% of codons are statistically different in codon usage between fission and fusion gene sets, indicating distinct priorities for codon usage. In the case of amino acid leucine, the codon choice becomes highly specific for each gene set. Out of six synonymous codons encoding for leucine, five synonymous codons were used in a statistically significantly different manner in fission and fusion gene sets. Leucine is present in the leucine zipper family (a protein family associated with sequence-specific DNA binding and faster gene expression). The differential usage of five synonymous codons out of six in the fission and fusion gene set suggests different operative mechanisms adapted for gene expression. The observation further strengthens the notion that mutational bias is highly correlated with gene expression in the fission gene set, while it is not in the fusion gene set. AGA codon runs appear as a part of translational regulation machinery. Both underrepresentation and over-representation within the fusion and fission gene set suggest translational forces fine-tuning the presence of AGA codons to keep the various fission and fusion proteins' quantity at the required level. Out of six leucine-encoding codons, five were used in a statistically significantly different manner in fusion and fission gene sets; hence, the occurrence of leucine is tightly associated with gene expression regulation. The correlation of CUB and protein properties (GRAVY and AROMA) and mutational bias and gene expression further underscores the presence of selection force in shaping codon usage. In the present analysis, we compared the molecular features of two opposite processes, fission and fusion and found that the transcripts share many common molecular features despite opposite processes.

Ethical approval

Ethical approval is not required since the study used publicly available data from online databases and no human or animal sample has been used.

Consent

No patient consent required since publicly available data has been used.

Sources of funding

This work was funded by the Researchers Supporting Project Number (RSP2024R339) King Saud University, Riyadh, Saudi Arabia.

Author contribution

R.K.: concept or design, data collection, data analysis or interpretation, writing and editing the paper. M.K.P.: concept or design, data collection, data analysis or interpretation, writing:R.K.G.: concept or design, data collection, data analysis or interpretation, writing and editing A.A.K.: concept or design, data collection, data analysis or interpretation, writing. I.B.: concept or design, data collection, data analysis or interpretation, writing. A.M.A.: concept or design, data collection, data analysis or interpretation, writing. P. N.: concept or design, data collection, data analysis or interpretation, writing and editing. P.G.: concept or design, data collection, data analysis or interpretation, writing. O.P.C.: concept or design, data collection, data analysis or interpretation, writing and editing. All authors participated in the manuscript writing and read and approved the final version of the MS.

Conflicts of interest disclosure

There are no conflicts of interest.

Research registration unique identifying number (UIN)

No human subjects involved.

Guarantor

Prakash Nepali and Rekha Khandia.

Data availability statement

Will be made available upon reasonable request.

Provenance and peer review

Paper is not invited one.

References

[1] Nikoletopoulou V, Markaki M, Palikaras K, *et al*. Crosstalk between apoptosis, necrosis and autophagy. Biochim Biophys Acta 2013;1833:3448–59.
[2] Davis R, Williams M. Mitochondrial function and dysfunction: an update. J Pharmacol Exp Ther 2012;342:598–607.
[3] Johri A, Beal MF. Mitochondrial dysfunction in neurodegenerative diseases. J Pharmacol Exp Ther 2012;342:619–30.
[4] Zorzano A, Claret M. Implications of mitochondrial dynamics on neurodegeneration and on hypothalamic dysfunction. Front Aging Neurosci 2015;7:101.
[5] Kanfer G, Kornmann B. Dynamics of the mitochondrial network during mitosis. Biochem Soc Trans 2016;44:510–6.
[6] Zhu T, Chen J-L, Wang Q, *et al*. Modulation of mitochondrial dynamics in neurodegenerative diseases: an insight into prion diseases. Front Aging Neurosci 2018;10:336.
[7] Su B, Wang X, Zheng L, *et al*. Abnormal mitochondrial dynamics and neurodegenerative diseases. Biochim Biophys Acta 2010;1802:135–42.
[8] Sarzi E, Angebault C, Seveno M, *et al*. The human OPA1delTTAG mutation induces premature age-related systemic neurodegeneration in mouse. Brain 2012;135:3599–613.
[9] Fuhrmann N, Alavi MV, Bitoun P, *et al*. Genomic rearrangements in OPA1 are frequent in patients with autosomal dominant optic atrophy. J Med Genet 2009;46:136–44.
[10] Jiang S, Nandy P, Wang W, *et al*. Mfn2 ablation causes an oxidative stress response and eventual neuronal death in the hippocampus and cortex. Mol Neurodegener 2018;13:5.
[11] Chen H, McCaffery JM, Chan DC. Mitochondrial fusion protects against neurodegeneration in the cerebellum. Cell 2007;130:548–62.
[12] Cipolat S, Martins de Brito O, Dal Zilio B, *et al*. OPA1 requires mitofusin 1 to promote mitochondrial fusion. Proc Natl Acad Sci USA 2004;101:15927–32.
[13] Bertholet AM, Delerue T, Millet AM, *et al*. Mitochondrial fusion/fission dynamics in neurodegeneration and neuronal plasticity. Neurobiol Dis 2016;90:3–19.
[14] DuBoff B, Götz J, Feany MB. Tau promotes neurodegeneration via DRP1 mislocalization in vivo. Neuron 2012;75:618–32.
[15] Zhou Z, Dang Y, Zhou M, *et al*. Codon usage is an important determinant of gene expression levels largely through its effects on transcription. Proc Natl Acad Sci U S A 2016;113:E6117–25.
[16] Duret L, Mouchiroud D. Expression pattern and, surprisingly, gene length shape codon usage in Caenorhabditis, Drosophila, and Arabidopsis. Proc Natl Acad Sci U S A 1999;96:4482–7.
[17] Zahdeh F, Carmel L. Nucleotide composition affects codon usage toward the 3'-end. PLoS One 2019;14:e0225633.
[18] Chakraborty S, Deb B, Barbhuiya PA, *et al*. Analysis of codon usage patterns and influencing factors in Nipah virus. Virus Res 2019;263:129–38.
[19] Knight RD, Freeland SJ, Landweber LF. A simple model based on mutation and selection explains trends in codon and amino-acid usage and GC composition within and across genomes. Genome Biol 2001;2:RESEARCH0010.
[20] Puigbò P, Bravo IG, Garcia-Vallve S. CAIcal: a combined set of tools to assess codon usage adaptation. Biol Direct 2008;3:38.
[21] Sueoka N. Intrastrand parity rules of DNA base composition and usage biases of synonymous codons. J Mol Evol 1995;40:318–25.
[22] Perna NT, Kocher TD. Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes. J Mol Evol 1995;41:353–8.
[23] Green P, Ewing B, Miller W, *et al*. NISC Comparative Sequencing Program, E.D. Green, Transcription-associated mutational asymmetry in mammalian evolution. Nat Genet 2003;33:514–7.
[24] Majewski J. Dependence of mutational asymmetry on gene-expression levels in the human genome. Am J Hum Genet 2003;73:688–92.
[25] Bergman J, Betancourt AJ, Vogl C. Transcription-associated compositional skews in drosophila genes. Genome Biol Evol 2018;10:269–75.
[26] Nair RR, Mohan M, Rudramurthy GR, *et al*. Strategies and patterns of codon bias in molluscum contagiosum virus. Pathogens 2021;10:1649.
[27] Kunec D, Osterrieder N. Codon pair bias is a direct consequence of dinucleotide bias. Cell Rep 2016;14:55–67.
[28] Fu H, Liang Y, Zhong X, *et al*. Codon optimization with deep learning to enhance protein expression. Sci Rep 2020;10:17617.
[29] Almutairi MM. Analysis of chromosomes and nucleotides in rice to predict gene expression through codon usage pattern. Saudi J Biol Sci 2021;28:4569–74.
[30] Bahiri-Elitzur S, Tuller T. Codon-based indices for modeling gene expression and transcript evolution. Comput Struct Biotechnol J 2021;19:2646–63.
[31] Kumar U, Khandia R, Singhal S, *et al*. Insight into codon utilization pattern of tumor suppressor gene EPB41L3 from different mammalian species indicates dominant role of selection force. Cancers (Basel) 2021;13:2739.
[32] Chen Y, Li X, Chi X, *et al*. Comprehensive analysis of the codon usage patterns in the envelope glycoprotein E2 gene of the classical swine fever virus. PLoS One 2017;12:e0183646.
[33] Bourret J, Alizon S, Bravo IG. COUSIN (COdon Usage Similarity INdex): A Normalized Measure of Codon Usage Preferences. Genome Biol Evol 2019;11:3523–8.
[34] Hofacker IL, Fontana W, Stadler PF, *et al*. Fast folding and comparison of RNA secondary structures. Monatsh Chem 1994;125:167–88.
[35] Bossy-Wetzel E, Barsoum MJ, Godzik A, *et al*. Mitochondrial fission in apoptosis, neurodegeneration and aging. Curr Opin Cell Biol 2003;15:706–16.
[36] Saneto RP. Mitochondrial diseases: expanding the diagnosis in the era of genetic testing. J Transl Genet Genom 2020;4:384–428.
[37] Hershberg R, Petrov DA. General rules for optimal codon choice. PLoS Genet 2009;5:e1000556.
[38] Majeed A, Kaur H, Bhardwaj P. Selection constraints determine preference for A/U-ending codons in Taxus contorta. Genome 2020;63:215–24.
[39] Choudhury MN, Uddin A, Chakraborty S. Nucleotide composition and codon usage bias of SRY gene. Andrologia 2017;e12787:1–11.
[40] Chakraborty S, Nath D, Paul S, *et al*. A crosstalk on codon usage in genes associated with leukemia. Biochem Genet 2021;59:235–55.
[41] Uddin A, Chakraborty S. Codon usage pattern of genes involved in central nervous system. Mol Neurobiol 2019;56:1737–48.
[42] Nath Choudhury M, Uddin A, Chakraborty S. Codon usage bias and its influencing factors for Y-linked genes in human. Comput Biol Chem 2017;69:77–86.
[43] Zámborszky J, Szikriszt B, Gervai JZ, *et al*. Loss of BRCA1 or BRCA2 markedly increases the rate of base substitution mutagenesis and has distinct effects on genomic deletions. Oncogene 2017;36:746–55.
[44] Khandia R, Alqahtani T, Alqahtani AM. Genes common in primary immunodeficiencies and cancer display overrepresentation of codon CTG and dominant role of selection pressure in shaping codon usage. Biomedicines 2021;9:1001.
[45] Uddin A, Chakraborty S. Codon usage trend in mitochondrial CYB gene. Gene 2016;586:105–14.
[46] Bornelöv S, Selmi T, Flad S, *et al*. Codon usage optimization in pluripotent embryonic stem cells. Genome Biol 2019;20:119.
[47] Novoa EM, Jungreis I, Jaillon O, *et al*. Elucidation of codon usage signatures across the domains of life. Mol Biol Evol 2019;36:2328–39.
[48] Miller JB, McKinnon LM, Whiting MF, *et al*. Codon pairs are phylogenetically conserved: a comprehensive analysis of codon pairing conservation across the tree of life. PLoS One 2020;15:e0232260.
[49] Lobry JR, Gautier C. Hydrophobicity, expressivity and aromaticity are the major trends of amino-acid usage in 999 Escherichia coli chromosome-encoded genes. Nucleic Acids Res 1994;22:3174–80.
[50] Khandia R, Singhal S, Kumar U, *et al*. Analysis of Nipah virus codon usage and adaptation to hosts. Front Microbiol 2019;10:886.
[51] Munjal A, Khandia R, Shende KK, *et al*. Mycobacterium lepromatosis genome exhibits unusually high CpG dinucleotide content and selection is key force in shaping codon usage. Infect Genet Evol 2020;84:104399.
[52] Morla S, Makhija A, Kumar S. Synonymous codon usage pattern in glycoprotein gene of rabies virus. Gene 2016;584:1–6.
[53] Monroe JG, Srikant T, Carbonell-Bejerano P, *et al*. Mutation bias reflects natural selection in Arabidopsis thaliana. Nature 2022;602:101–5.
[54] Marek A, Tomala K. The contribution of purifying selection, linkage, and mutation bias to the negative correlation between gene expression and polymorphism density in yeast populations. Genome Biol Evol 2018;10:2986–96.

[55] Sawada R, Mitaku S. Biological meaning of DNA compositional biases evaluated by ratio of membrane proteins. J Biochem 2012;151:189–96.
[56] Zhang L-P, Cai Y-Y, Yu D-N, *et al*. Gene characteristics of the complete mitochondrial genomes of Paratoxodera polyacantha and Toxodera hauseri (Mantodea: Toxoderidae). PeerJ 2018;6:e4595.
[57] Shen W, Wang D, Ye B, *et al*. GC3-biased gene domains in mammalian genomes. Bioinformatics 2015;31:3081–4.
[58] Deka H, Chakraborty S. Insights into the usage of nucleobase triplets and codon context pattern in five influenza a virus subtypes. J Microbiol Biotechnol 2016;26:1972–82.
[59] Victor MP, Acharya D, Begum T, *et al*. The optimization of mRNA expression level by its intrinsic properties-Insights from codon usage pattern and structural stability of mRNA. Genomics 2019;111:1292–7.
[60] Presnyak V, Alhusaini N, Chen Y-H, *et al*. Codon optimality is a major determinant of mRNA stability. Cell 2015;160:1111–24.
[61] Edwards NC, Hing ZA, Perry A, *et al*. Characterization of coding synonymous and non-synonymous variants in ADAMTS13 using ex vivo and in silico approaches. PLoS One 2012;7:e38864.
[62] Coy JF, Sedlacek Z, Bächner D, *et al*. A complex pattern of evolutionary conservation and alternative polyadenylation within the long 3"-untranslated region of the methyl-CpG-binding protein 2 gene (MeCP2) suggests a regulatory role in gene expression. Hum Mol Genet 1999;8:1253–62.
[63] Alqahtani T, Khandia R, Puranik N, *et al*. Leucine encoding codon TTG shows an inverse relationship with GC content in genes involved in neurodegeneration with iron accumulation. J Integr Neurosci 2021;20:905–18.
[64] Uddin A, Paul N, Chakraborty S. The codon usage pattern of genes involved in ovarian cancer. Ann N Y Acad Sci 2019;1440:67–78.