Current Trends in Biotechnology and Pharmacy
Vol. 20 (Supplementry Issue 1A), March 2026, ISSN 0973-8916 (Print)., 2230-7303 (Online)
10.5530/ctbp.2026.1s.27

Evaluating Selection and Mutation Influences on Codon Usage Patterns in the Chloroplast Genomes of *Amaranthus* species

Bevin Nishanth J, Romanshia Celes G A, G Atul Babu, and Suji Somasundaram*

Department of Biotechnology, SIMATS School of Engineering, SIMATS, Thandalam, Chennai-602 105, Tamil Nadu, India
Corresponding author: sujis.sse@saveetha.com

Abstract

Amaranthus is a genus characterized by taxonomic complexity and ecological diversity, gaining importance in agriculture and genomics. We report in-depth codon usage bias (CUB) analysis of 18 *Amaranthus* species chloroplast genomes to make inroads into the evolutionary forces responsible for their genomic structure. We analyzed various codon usage parameters, including relative synonymous codon usage (RSCU), effective number of codons (ENC), GC Content, nucleotide composition, and positional bias. Our results indicated a strong inclination towards A/T-ending codons and a generally AT-rich nucleotide composition, especially at the third codon position, across all species. The ENC values varied from 34.05 to 58.87, reflecting a moderate codon usage bias. Neutrality plot analysis revealed weak correlations between GC3 and GC12 (slope 0.095–0.141), implying that natural selection, rather than mutation pressure, predominantly influences codon preferences. ENC-GC3s plots corroborated this, with most genes positioned below the expected curve, and PR2-bias plots demonstrated an imbalance favoring T and G over A and C at the third codon position. Correspondence analysis (COA) further confirmed that both mutational and selective forces affect codon usage variation. We identified 31 high-frequency codons and 12 optimal codons, with species-specific codon preferences, indicating subtle evolutionary adaptations. This study offers new insights into the molecular evolution and codon optimization of *Amaranthus*, providing valuable information for phylogenetics, functional genomics, and biotechnological applications within the genus.

Keywords: *Amaranthus*, codon usage bias. Chloroplast genome, Optimal codons, High frequency codons, RSCU

Introduction

Amaranthus L. (Amaranthaceae Juss.) is a taxonomically complex and globally distributed genus comprising approximately 92 accepted species, approximately half of which are native to the Americas. It was first described in Linnaeus's *Species Plantarum* (1753) (http://www.plantsoftheworldonline.org), and the genus is now accepted as cosmopolitan in distribution, with species naturalized or introduced across all continents [1]. Members of this genus are ecologically diverse and economically significant, and are used as leafy vegetables, pseudocereal grains, ornamentals, and medicinal plants [1]. From a taxonomic perspective, *Amaranthus* is particularly challenging because of extensive phenotypic plasticity, high levels of interspecific hybridization, and inconsistent morphological traits. This has led to nomenclature inconsistencies and difficulties in classification. Traditional groupings into subgenera (*Acnida*, *Albersia*, and *Amaranthus*) have been questioned by recent molecular phylogenetic studies, highlighting the need for genomic approaches to better resolve evolutionary relationships within the genus [2].

Chloroplast genomes, recognized for their generally conserved structure and

Vol. 20 (Supplementry Issue 1A), March 2026, ISSN 0973-8916 (Print)., 2230-7303 (Online)
10.5530/ctbp.2026.1s.27

function, offer significant insights into evolutionary connections and gene expression dynamics within this genus[3]. Extensive research on chloroplast genomes across diverse plant species has highlighted the distinct features of these genomes, including their quadripartite structure, conserved gene arrangement, and codon usage bias (CUB) patterns, which are crucial for phylogenetic studies and for understanding evolutionary processes[4, 5]. Chloroplasts, which are essential organelles of green plants, are characterized by their unique genomic structure and play a fundamental role in the process of photosynthesis, yielding energy for various physiological activities in plants. The chloroplast gene consists of four parts: large single-copy (LSC) region, small single-copy (SSC) region, and two inverted repeat (IR) sets [6]. Despite the general conservation of chloroplast genomes, structural changes have occurred within and between species over the course of evolution, such as expansion and contraction of IR regions[7].

Codons are responsible for the internal transfer of genetic information from mRNA to proteins in an organism. While synthesizing proteins, 61 codons specify 20 amino acids through a phenomenon of codon degeneracy[8]. All amino acids except tryptophan and methionine are coded by more than one codon called synonymous codons. But non-uniform usage of these synonymous codons results in codon usage bias [9]. Both prokaryotic and eukaryotic organisms exhibit non-random usage of codons. Highly expressed genes preferably use certain synonymous codons, and these usually preferred or 'optimal codons' are abundantly present in such genes. In the vast majority of organisms, usage of certain synonymous codons is largely determined by natural selection [10, 9]. Analytical studies of codon usage tendencies of the codon genomes of Amaranthus can reveal the effects of mutation pressure and natural selection. Previous research has demonstrated that environmental adaptability and functional constraints play roles in codon usage bias (CUB). For example, the tendency toward AT-rich codons in chloroplast genomes is a common occurrence among plant species and is driven by selective pressures that enhance translational efficiency [11-13]. Analyzing codon biases in *Amaranthus* not only aids in understanding evolutionary trends, but also helps reveal the molecular mechanisms through which different species have adapted to their unique environments and lifestyles [14, 15]. Moreover, the use of complete chloroplast genome sequences is instrumental in deciphering the intricate balance between genomic stability and variability that drives the evolutionary dynamics of this genus. Chloroplast genomic data are essential for taxonomic classification, as evidenced by recent studies that combined genomic and phenotypic information to clarify taxonomic uncertainties [4, 16]. These studies also illuminate the complex interactions between chloroplast genomes and other genomic compartments, specifically the nuclear and mitochondrial genomes. Such comparative genomic methods offer vital insights into the evolution of significant traits, such as photosynthesis, growth efficiency, and environmental resilience, which are particularly pertinent to the *Amaranthus* genus because of their wide ecological adaptability and varied industrial uses.

Additionally, research has demonstrated that the distinct evolutionary rates of conserved structural components in chloroplast genomes offer valuable insights into species divergence within phylogenetic frameworks. These components also serve as excellent molecular markers for examining plant evolution [17, 18]. The chloroplast genomes of various *Amaranthus* species have revealed patterns of length variation, insertion – deletion events, and gene loss, all of which have significant ecological and evolutionary implications. This finding emphasizes the vital role of chloroplast genomics in both fundamental and applied biological research.

Vol. 20 (Supplementry Issue 1A), March 2026, ISSN 0973-8916 (Print)., 2230-7303 (Online)
10.5530/ctbp.2026.1s.27

Materials and Methods

Genomes and coding sequences

The chloroplast genomes(Cp genomes) of eighteen *Amaranthus* species, including *Amaranthus blitoides, Amaranthus blitum, Amaranthus capensis, Amaranthus crispus, Amaranthus cruentus, Amaranthus deflexus, Amaranthus hybridus, Amaranthus hypochondriacus, Amaranthus powellii, Amaranthus retroflexus, Amaranthus roxburghianus, Amaranthus spinosus, Amaranthus caudatus, Amaranthus albus, Amaranthus standleyanus, Amaranthus tuberculatus*, and *Amaranthus viridus*, were retrieved from the National Center for Biotechnology Information (NCBI). Table 1 lists the accession numbers of the 18 *Amaranthus* species. The selection of protein-coding sequences (CDS) from all chloroplast genomes was based on the following criteria [19, 20]: (1) Each CDS must begin with the initiation codon (ATG) and end with one of the termination codons (TAA, TAG, or TGA); (2) the sequence must contain more than three bases; (3) the length of the sequences should exceed 300; and (4) sequences containing an internal stop codon were excluded.

Indices of codon usage

Codon usage can be investigated from several valuable metrics: (1) Relative Synonymous Codon Usage (RSCU); (2) the Effective Number of Codons (ENC); (3) the GC composition, comprising total GC percentage and the GC third-position content at each codon position (initial, second, third codon positions) with particular reference to the third codon-position of the synonymous codons (GC3s); (4) the nucleotide composition (A, T, G, C) at the third codon-position specifically (A3, T3, G3, C3); and (5) the total amino acid count.

Table 1: Species of *Amaranthus* genus and their GenBank accession ID from NCBI

S. No.	Common Name	Scientific Name	CP genome	Acession ID
1	Tumbleweed Amaranth	*Amaranthus albus*	150,757 bp	MT526785.1
2	Mat Amaranth	*Amaranthus blitoides*	150,668 bp	MT526786.1
3	Livid Amaranth	*Amaranthus blitum*	150,775 bp	MW255966.1
4	Cape Amaranth	*Amaranthus capensis*	150,708 bp	MT526779.1
5	Pendant Amaranth	*Amaranthus caudatus*	150,523 bp	MG836508.1
6	Crispleaf Amaranth	*Amaranthus crispus*	150,539 bp	MT526778.1
7	Red amaranth	*Amaranthus cruentus*	150,757 bp	MG836507.1
8	Argentina Amaranth	*Amaranthus deflexus*	150,257 bp	MT526776.1
9	Smooth Amaranth	*Amaranthus hybridus*	150,759 bp	MT993471.1
10	Grain Amaranth	*Amaranthus hypochondriacus*	150,523 bp	MG836505.1
11	Powell's Amaranth	*Amaranthus powellii*	150,775 bp	PQ385849.1
12	Redroot Amaranth	*Amaranthus retroflexus*	150,710 bp	MW646089.1
13	Slender Amaranth	*Amaranthus roxburghianus*	149,969 bp	OQ354384.1
14	Spiny Amaranth	*Amaranthus spinosus*	150,523 bp	OP718299.1
15	Standley's Amaranth	*Amaranthus standleyanus*	150,582 bp	MT526780.1
16	Chinese amaranth	*Amaranthus tricolor*	150,718 bp	OM419215.1
17	Roughfruit Amaranth	*Amaranthus tuberculatus*	150,680 bp	MT559304.1
18	Verte Amaranth	*Amaranthus viridus*	150,452 bp	MW679034.1

Vol. 20 (Supplementry Issue 1A), March 2026, ISSN 0973-8916 (Print)., 2230-7303 (Online)
10.5530/ctbp.2026.1s.27

High frequency and optimal codons

Relative Synonymous Codon Usage (RSCU) is calculated by the ratio of observed codon frequency to expected frequency assuming all synonymous codons to be utilized uniformly, thus indicating codon usage bias [21]. If RSCU value is above 1, it indicates prominent positive bias in codon usage [22]. Those codons showing RSCU values above 1 are considered to be of the high-frequency codon type [23]. To examine codon usage tendencies, chloroplast genes were sorted by their Effective Number of Codons (ENC) values, and the top 10% of genes that possess the topmost and lowest ENC values were chosen to represent the dataset for high-expression genes and that for low-expression genes to generate, respectively. RSCU values of codons in these resulting datasets were computed and examined by ΔRSCU. Those codons were considered to be optimal that meet the conditions of showing ΔRSCU of at least 0.08 with that of RSCU value higher than 1 [23].

Neutrality plot analysis

To investigate neutrality, a scatter plot was created that plotted GC12 (the average of the first and second codon position's GC content) against GC3 and investigated how nucleotide composition varied at different codon positions. This approach allows for the consideration of how much mutation pressure and natural selection contribute to codon usage patterns [24]. An accurate correlation of GC12 to GC3, signified by a regression coefficient that is close to one, points towards mutation pressure as the driving force for codon usage. An unsound correlation that is described by a regression coefficient that is close to zero, on the other hand, suggests that natural selection has a stronger influence on codon preference [24, 25].

ENC-plot analysis

Effective Number of Codons (ENC) is another widely used measure for analyzing codon usage bias against random expectation. ENC ranges from 20 (high codon bias) to 61 (non-bias), indicating the level of non-uniform synonymous codon usage [26]. Higher ENC values suggest a lower bias in codon usage, whereas lower values indicate a stronger preference for certain codons [27]. An ENC threshold of 35 is frequently employed as a reference point for distinguishing strong codon preference from weaker or more balanced codon usage [28, 26]. For ENC-plot purposes, ENC values were plotted against GC3s, which are the proportion of guanine and cytosine nucleotides at the third position of synonymous codons. As tryptophan (Trp) and methionine (Met) do not have synonymous codons, we calculated GC3 excluding these residues. ENC-plot is utilized to evaluate the effect of base composition on codon usage and to detect other factors that affect codon bias. Theoretical ENC values were estimated on the basis of the standard equation proposed by Wright [26]:

$$ENC = 2 + GC3s + \frac{29}{GC3s^2 + (1 - GC3s)^2}$$

If observed ENC values follow the expected curve closely, mutation pressure is assumed to be the dominant force shaping codon usage. Conversely, large deviations indicate the effect of natural selection or other pressures of evolution.

$$ENC \text{ ratio} = \frac{ENC_{\exp} - ENC_{\text{obs}}}{ENC_{\exp}}$$

The ENC ratio, defined as the relative difference between observed and expected ENC values, provides an additional metric for evaluating codon usage bias [29, 30].

Parity rule 2 plot analysis

Nucleotide composition at the 3rd position of a codon is essential to define codon usage patterns [31]. PR2 plot analysis involves plotting the value of A3/(A3 + T3) on

Vol. 20 (Supplementry Issue 1A), March 2026, ISSN 0973-8916 (Print)., 2230-7303 (Online)
10.5530/ctbp.2026.1s.27

the y-axis against G3/(G3 + C3) on the x-axis. The technique is used to reveal underlying patterns of codon organization [32]. In theory, when mutation pressure is the predominant influence on chloroplast gene sequences, the proportions of adenine to thymine and guanine to cytosine at the third codon position will tend to reach equilibrium, resulting in both axes centered around the value 0.5 (when A equals T and G equals C). Deviation from these fundamental values suggests that other evolutionary forces, such as natural selection, may be driving codon usage patterns [33].

Correspondence analysis (COA)

Correspondence Analysis (COA) is an advanced statistical technique routinely applied to study variation in codon usage among genes [34, 35]. The method employed in this work used Relative Synonymous Codon Usage (RSCU) values for all codons except termination signals, tryptophan, and methionine. The coding regions were mapped onto a 59-dimensional space in which each coordinate was a distinct synonymous codon. Notably, the first principal axis identified in this analysis explained the largest portion of codon usage variability, while succeeding axes represented progressively less variance [36].

Results:

Codon Usage Patterns (CUP)

Analysis of Codon Composition

To examine codon preferences among the 18 *Amaranthus* species, various codon usage metrics were assessed. For *Amaranthus* chloroplast genes, the GC content was observed to be below 50% across all studied species (Fig. 1). When comparing codon positions, the first position displayed a higher GC proportion than the second and third, with the third position having the lowest GC values among all codon sites. Sequence analysis identified an unequal representation of the four nucleotides, with thymine emerging as the most abundant, followed by adenine, guanine, and cytosine across the 18 species analyzed (Fig. 2). This uneven distribution, particularly the notably greater occurrence of AT bases at the third position, contributes to distinct codon usage trends in these species.

Relative Synonymous Codon Usage (RSCU) Analysis

To evaluate the preference for synonymous codons, the RSCU values were determined (Supplementary Table S1). The codon usage patterns were largely consistent

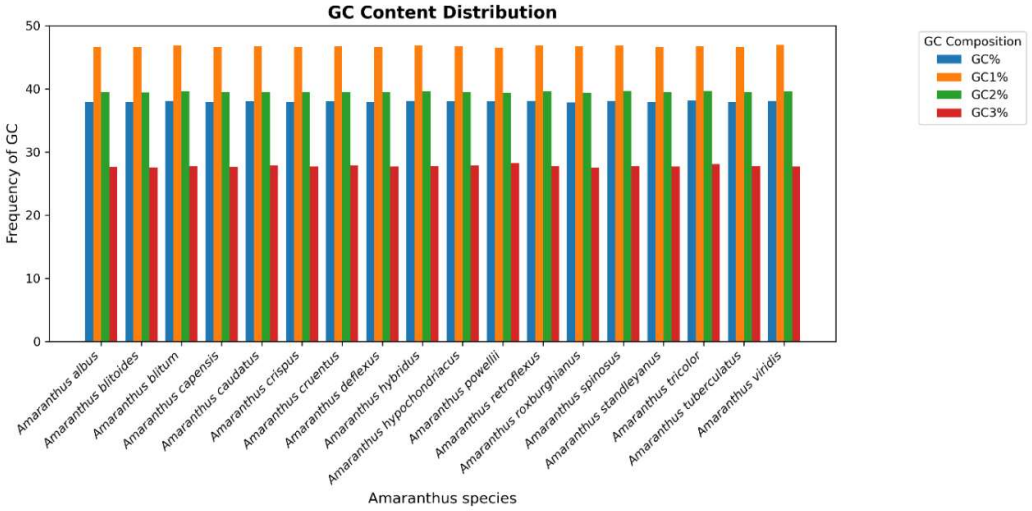

Fig. 1: GC content distribution across *Amaranthus* species

Vol. 20 (Supplementry Issue 1A), March 2026, ISSN 0973-8916 (Print)., 2230-7303 (Online)
10.5530/ctbp.2026.1s.27

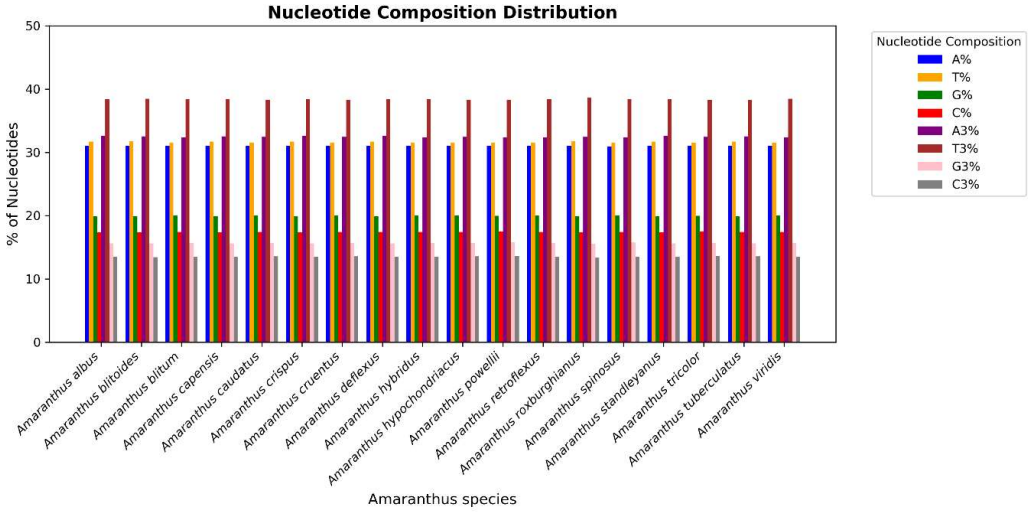

Fig. 2: Distribution of nucleotide composition in *Amaranthus* species

across all 18 *Amaranthus* species. A total of 29 codons were found to have RSCU values greater than 1, indicating that they were frequently used. Notably, 27 of these codons ended with either A or T, making up approximately 93.1%, which suggests a strong inclination towards A/T-ending codons. In contrast, 31 codons had RSCU values less than 1, with 28 ending in G or C, representing 90.3%, indicating a negative bias against G/C-ending codons. Codons such as GCT, TCT, TTA, ATT, AGA, GGA, and TAT were frequently used across multiple species, emphasizing the conserved nature of A/T-ending codon preference.

ENC Values and Codon Bias Strength

The ENC values for chloroplast genes in *Amaranthus* ranged from 34.05 to 58.87, with an average of 47.11, indicating a generally weak bias in codon usage (ENC > 35). This suggests that mutation pressure significantly influences codon usage patterns in *Amaranthus*, which is consistent with trends seen in other dicot species.

Optimal Codon Identification

Using ΔRSCU values comparing highly and lowly expressed genes, 12 optimal codons were identified (Supplementary Table S2), of which 11 ended in A or T. These included TCA, ACT, TAT, CAT, AAT, GAT, AGA, GGA, TGT, ATT, ACA, and CCT—codons commonly found across most species (Supplementary Table S3). This strong bias towards A/T-ending optimal codons is consistent with the overall genomic AT richness and mutational pressure. While most optimal codons were shared, certain species exhibited unique preferences. For instance, TTG was uniquely optimal in *A. powellii*, suggesting lineage-specific adaptation.

Neutrality Plot Analysis

Neutrality plots were developed for chloroplast genes of all 18 *Amaranthus* species to analyze the correlation between GC content at codon positions one and two (GC12) and the third position (GC3). This analysis aimed to distinguish the respective contributions of mutational pressure and natural selection in shaping codon usage patterns (Fig. 3). Correlation coefficients (r) for GC12 and GC3 were generally low across species, from 0.0632 in *A. caudatus* to 0.0955 in *A. blitoides*, indicating a poor linear relationship. Slopes of regression lines varied from 0.0951 to 0.1415, which indicate a

Vol. 20 (Supplementry Issue 1A), March 2026, ISSN 0973-8916 (Print)., 2230-7303 (Online)
10.5530/ctbp.2026.1s.27

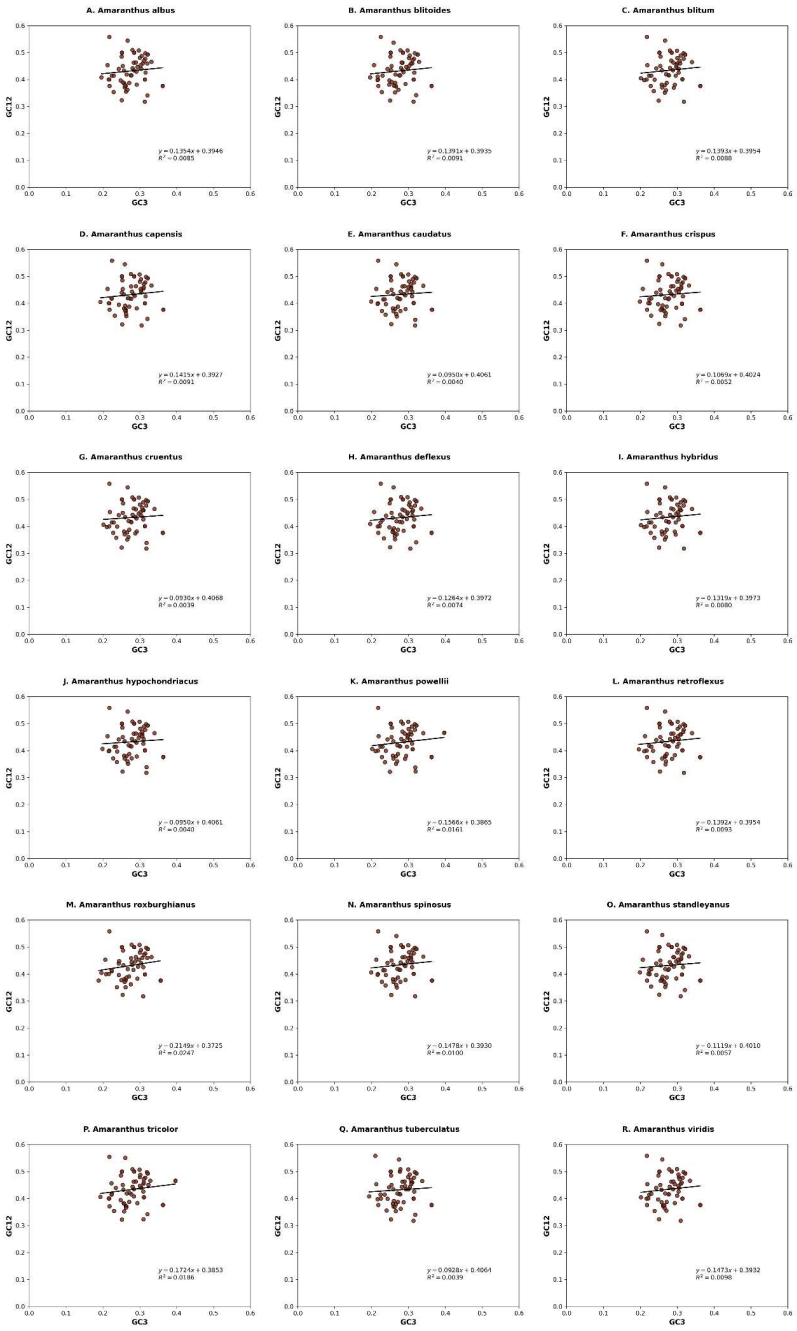

Fig. 3: Neutrality plot analysis of chloroplast genes in 18 *Amaranthus* species

Vol. 20 (Supplementry Issue 1A), March 2026, ISSN 0973-8916 (Print)., 2230-7303 (Online)
10.5530/ctbp.2026.1s.27

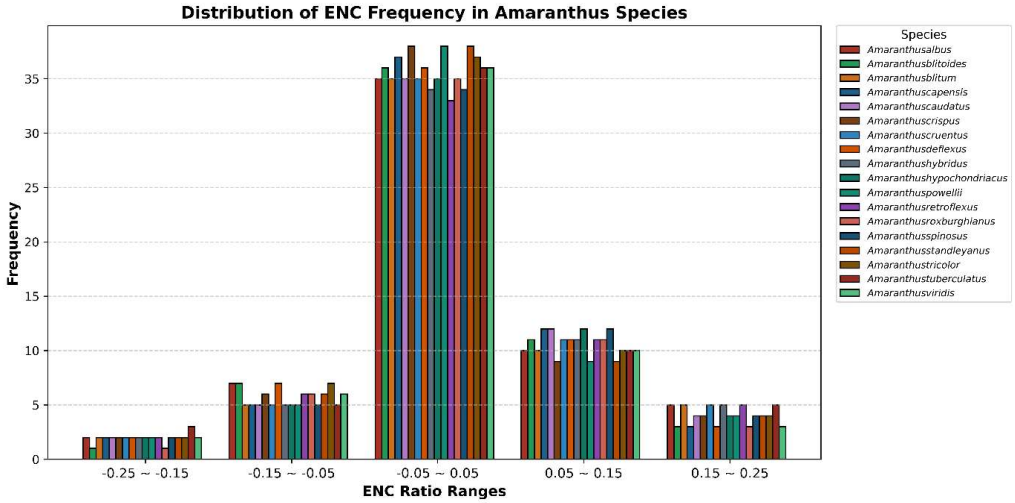

Fig. 4: Distribution of ENC ratio frequency across chloroplast genes in 18 *Amaranthus* species

contribution of mutation pressure by 9.51% and 14.15%. Highest mutation pressure was observed in *A. capensis* (14.15%), followed by *A. blitum* (13.96%) and *A. blitoides* (13.91%), while the lowest was observed in *A. caudatus* (9.51%). These findings imply that mutation pressure makes only a modest contribution to codon usage patterns and that natural selection and other mechanisms of evolution are probably the major driving forces behind codon usage bias in *Amaranthus* chloroplast genes.

ENC-Plot Analysis

For examining the correlation between codon usage bias and GC content at the third codon position (GC3s) in the chloroplast genes among 18 *Amaranthus* species, ENC-plot analysis was used. All plots displayed a reference curve of the expected ENC value under neutrality as the baseline (Fig. 4). In general, data points for most species were positioned below this expected ENC curve, suggesting that factors beyond just GC3s content influence codon usage bias. This means that other evolutionary forces such as natural selection, translational selection, and gene expression levels also play a part in selecting codons in *Amaranthus*. On average, 60–64% of genes in all *Amaranthus* species had ENC ratio deviations between –0.05 and +0.05 (Fig. 5), with the highest being in *A. powellii* at 65.52%, followed by *A. crispus* and *A. standleyanus* at 64.41%. This suggests a moderate level of codon usage bias and that GC3 content alone cannot account for the patterns of codon usage in *Amaranthus* chloroplast genomes. Overall, these results indicate that selectional pressures exert highly significant effects on codon usage bias in *Amaranthus* chloroplast genes compared to mutational pressures.

PR2-Plot Analysis

PR2-bias plots were used to determine the effect of mutation and selection on nucleotide use at the third codon position of the 18 Amaranthus chloroplast genes (Fig. 6). Each plot was produced by graphing A3 / (A3 + T3) on the Y-axis and G3 / (G3 + C3) on the x-axis. In all organisms, the genes were disproportionately scattered over the PR2-plane, with most of the data points concentrated in the bottom-right quadrant. This pattern indicated a high degree of skewness in base usage, tending to favor T over A and G over C at the third position of

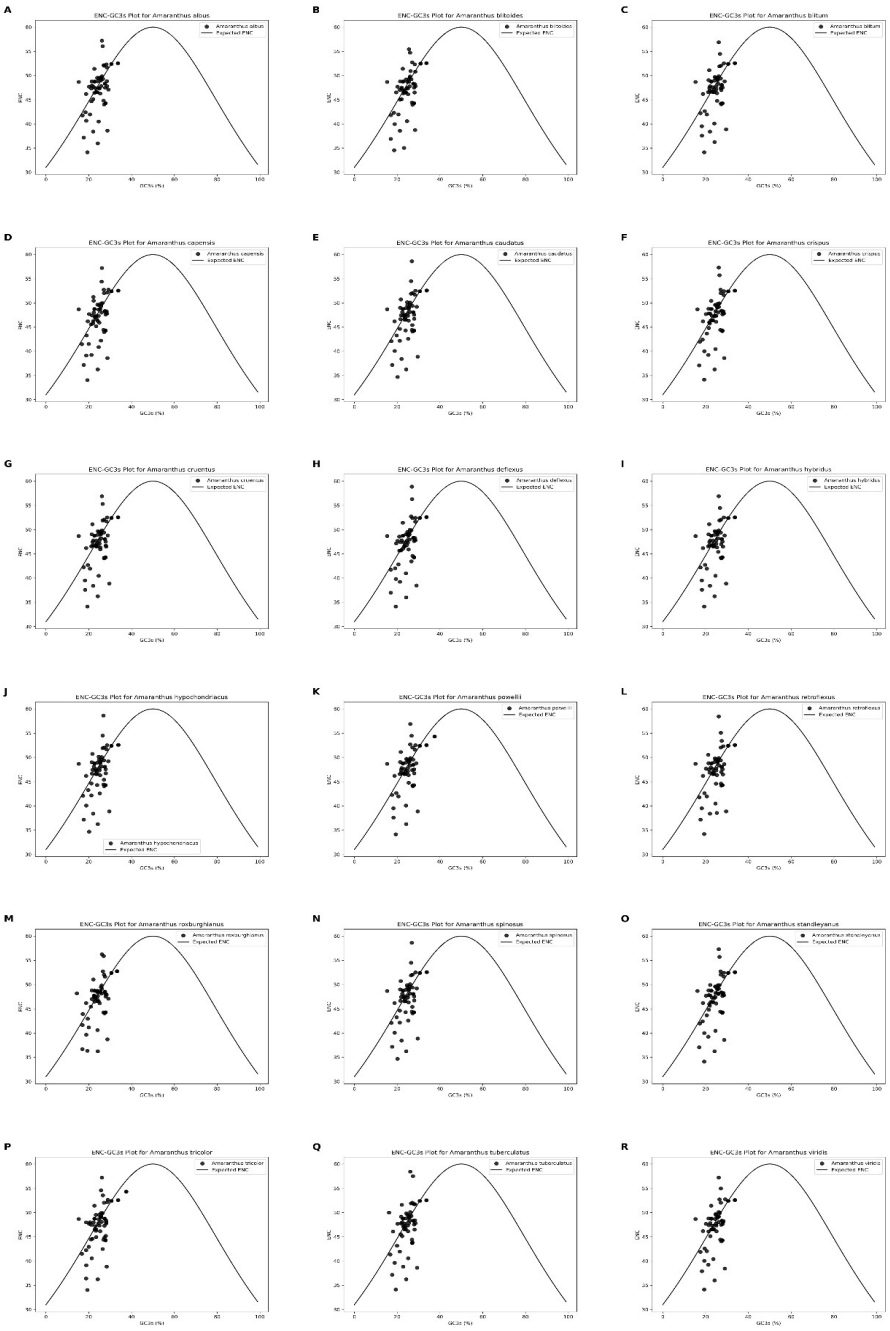

Fig. 5: ENC-GC3s plot analysis of chloroplast genes in 18 *Amaranthus* species

Nishanth et al.

Vol. 20 (Supplementry Issue 1A), March 2026, ISSN 0973-8916 (Print)., 2230-7303 (Online)
10.5530/ctbp.2026.1s.27

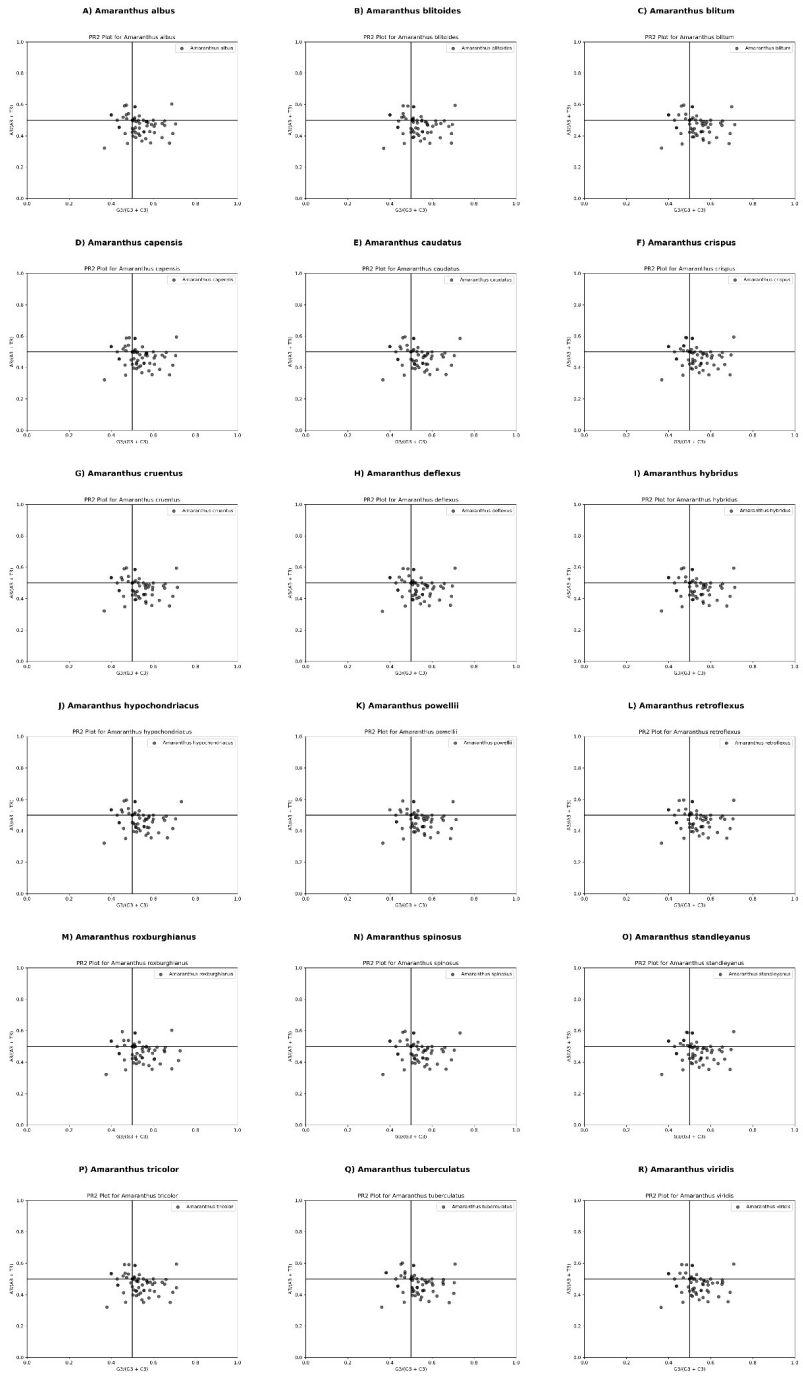

Fig. 6: Parity Rule 2 (PR2) bias plots of chloroplast genes in 18 *Amaranthus* species

Chloroplast Genomes of *Amaranthus species*

Vol. 20 (Supplementry Issue 1A), March 2026, ISSN 0973-8916 (Print)., 2230-7303 (Online)
10.5530/ctbp.2026.1s.27

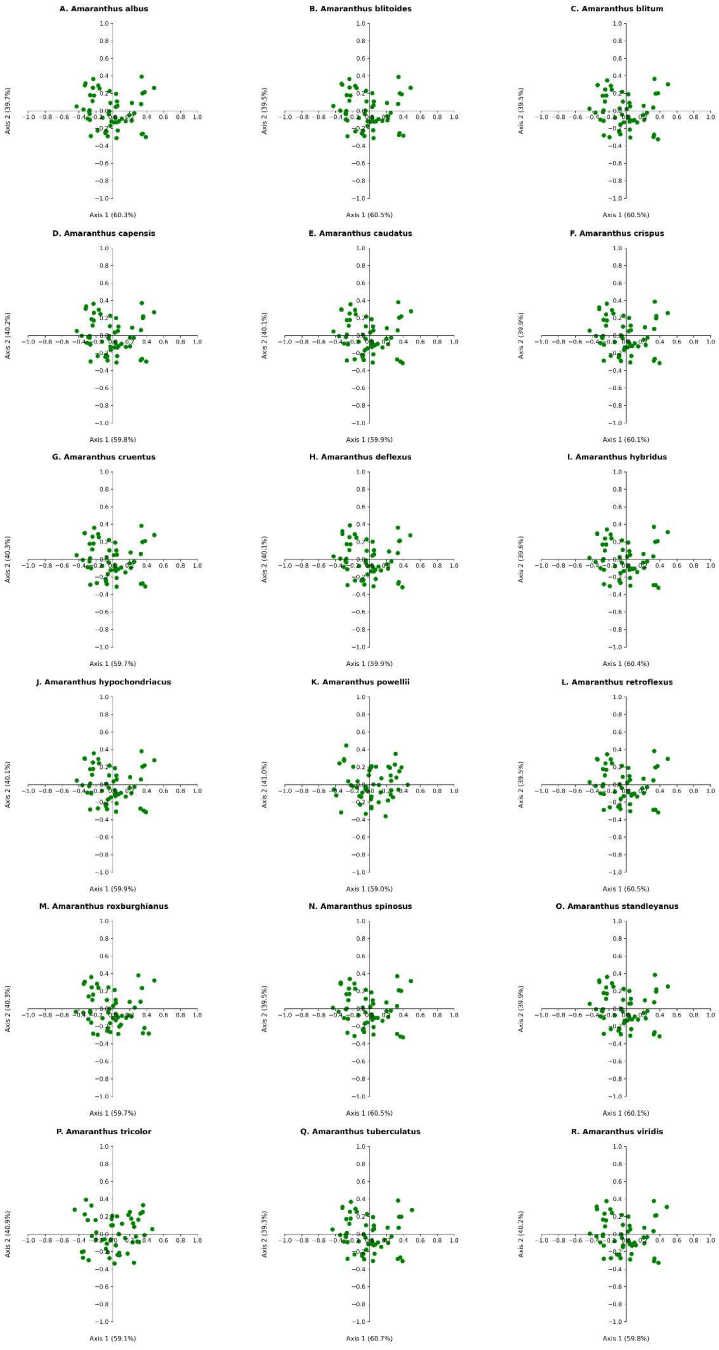

Fig. 7: Correspondence analysis (COA) plots based on relative synonymous codon usage (RSCU) of chloroplast genes in 18 *Amaranthus* species

Vol. 20 (Supplementry Issue 1A), March 2026, ISSN 0973-8916 (Print)., 2230-7303 (Online)
10.5530/ctbp.2026.1s.27

the codon. The mean AT-bias values throughout species varied from 0.465 to 0.468, while GC-bias values ranged between 0.536 and 0.538, suggesting a uniform T and G base preference at the third codon position in the genus. For instance, A. albus had a mean AT bias of 0.468 and a GC bias of 0.537, showing a bias in the order T > A and G > C. The observed imbalances highlighted the influence of mutation pressure, natural selection, and other evolutionary forces on codon usage. The significant departure from the midpoint of 0.5 on both axes indicates that non-neutral processes influence codon usage in the chloroplast genes of *Amaranthus.*

Correspondence Analysis (COA)

The RSCU values of coding sequences were compared among 18 *Amaranthus* species in this study to examine the differences in codon usage (Fig. 7). The first axis (Axis 1) explained the largest proportion of the total variation for all the species, i.e., 59.0% to 60.7%, while the second axis (Axis 2) explained 12.9% to 20.7% of the variation. Gene plots were characterized by a dispersed pattern along the axes, which indicates that base composition alone cannot explain codon usage bias. It indicates that natural selection and gene expression levels, in addition to mutational pressure, contribute to codon usage patterns in chloroplast genomes of *Amaranthus*.

Discussion

Codon usage bias (CUB) is a widespread phenomenon noticed in different organisms ranging from prokaryotes and eukaryotes to plant chloroplast genomes. This indicates a tendency to preferentially use certain synonymous codons over others coding for the same amino acid. Knowledge about CUB is essential for research on the molecular evolution of species as well as for optimizing gene expression in biotechnological processes [37].

In this research, have analyzed the chloroplast genomes of 18 *Amaranthus* species to investigate the patterns of codon usage and their evolutionary impact on codon usage bias (CUB). Our results indicated that there was a general GC content of below 50% with a preference for A-ending or U-ending codons at the third position. This AT-skewed trend is consistent with that of other dicotyledonous organisms, including Populus alba and the family Euphorbiaceae, in which this bias for A/U-ending codons has been observed [38]. Angiosperms also exhibit a bias towards A and T for the third codon position, which is related to mutational pressures that prefer certain nucleotide substitutions. This observation is supported by previous studies highlighting the significance of GC content in shaping codon usage patterns across various plant taxonomic groups [39]. Examination of nucleotide distributions revealed that thymine (T) was the most prevalent, followed by adenine (A), guanine (G), and cytosine (C). This significant AT bias underscores an evolutionary tendency to favor A/T-ending codons, implying possible advantages in expression efficiency and translational optimization. Similar findings have been reported in other studies, which highlight that codon preference is influenced not only by mutational pressures, but also by gene expression levels and translational efficiency [39, 40].

RSCU analysis identified high-frequency codons across 18 species, revealing that 29 out of 31 frequent codons ended in A or T, exceeding 93%. This pattern underscores the AT bias and suggests an adaptive mechanism within *Amaranthus*, aligning with patterns observed in other monocots and dicots. Research indicates that synonymous codons favoring A or T are preferred in genes with high expression, suggesting a co-evolutionary link between codon usage and intrinsic characteristics of the chloroplast genome [41, 40]. Additionally, 12 best codons were found consistently in the majority of *Amaranthus* species, 11 of which

Vol. 20 (Supplementry Issue 1A), March 2026, ISSN 0973-8916 (Print)., 2230-7303 (Online)
10.5530/ctbp.2026.1s.27

were A or T ending, which supports the use of AT-rich codons for the third codon position. These best codons were TCA, ACA, TAT, CAT, AAT, GAT, AGA, GGA, TGT, ATT, ACT, and CCT.

ENC calculation gave a numeric value of codon bias, which was found to be mild codon usage bias in *Amaranthus* chloroplast genes with ENC values ranging from 34.05 to 58.87. Additionally, ENC-plot analysis demonstrated that codon usage bias in *Amaranthus* species does not strictly conform to the patterns anticipated solely based on GC3 content, as the majority of genes showed observed ENC values that fell below the theoretical curve. Other influences, including natural selection, gene expression, and translational efficiency, play significant roles in regulating codon usage patterns [39].

To determine evolutionary pressures influencing codon usage bias, neutrality plot analysis was performed and revealed that there was only a very small relationship between the GC12 and GC3 values, with slopes varying from 0.095 to 0.141 when observed. This reveals that mutation pressure is responsible for just 9.5% to 14.1% of the observed codon usage bias. Additionally, analysis of ENC-plot showed that the majority of genes were below the neutral expectation, suggesting that GC3 content itself is not sufficient to fully account for codon bias. About 62% to 63% of the genes in each species had ENC deviations within ±0.05, suggesting a moderate preference for particular codons. These findings indicate that natural selection contributes more than mutation pressure to pattern codon usage in *Amaranthus*, as observed in other plant species, e.g., Epimedium, in which natural selection was found to be the dominating factor behind CUB [38].

PR2-bias plot analysis revealed genes to be skewed, predominantly clustering in the lower-right quadrant, indicating thymine (T) and guanine (G) nucleotide preference at the third codon position. The pattern signified that selection pressure, instead of neutral mutations, played a considerable role in determining codon preference. A similar observation has also been reported in other plant species, reinforcing the influence of natural selection on codon usage bias [42].

Studies of comparative codon usage have shown variations between monocots and dicots, with dicots generally having a bias for AT-rich codons in the third position. We observed this trend especially in our analysis of the *Amaranthus* species, which prefers A- or U-ending codons. This bias influences the efficiency of gene expression and is useful for the planning of genetic engineering and breeding strategies for crops [15]. These results are relevant to the molecular evolution of *Amaranthus* and provide a basis for future studies that seek to maximize gene expression in this genus.

Conclusion

In the chloroplast genomes of 18 *Amaranthus* species, codon usage patterns were largely consistent, although certain species exhibited distinct evolutionary trends in their codon usage preferences. The overall codon usage bias across all species was minimal, with a strong tendency towards A/T-ending codons at the third position.

In the *Amaranthus* genus, codon usage revealed 29 frequency high codons (RSCU > 1), more than 93% of which ended in A or T, suggesting a strong bias towards A/T-ending codons. This consistent AT-rich codon preference reflects underlying mutational and selective pressures associated with genomic composition. The identification of 12 common optimal codons across most species underscores a conserved codon usage pattern, while the presence of species-specific optimal codons, such as TTG in *Amaranthus powellii*, indicates subtle lineage-specific adaptations. Together, these findings suggest that codon usage bias in *Amaranthus* is shaped by both evolutionary conservation and species-specific factors.

Through multivariate analyses, such as neutrality plots and ENC-GC3s plots, it was found that natural selection played a

Vol. 20 (Supplementry Issue 1A), March 2026, ISSN 0973-8916 (Print)., 2230-7303 (Online)
10.5530/ctbp.2026.1s.27

more significant role than mutation pressure in shaping codon usage patterns. Only 9.5–14.1% of the codon bias was contributed by mutation pressure, and most of the variation resulted from selectional constraints and other sources. PR2 plot analysis also corroborated this by demonstrating biased nucleotide usage at the third codon position where T was overwhelmingly preferred over A and G over C, reflecting selective pressure. This research provides a comprehensive analysis of codon usage bias in the chloroplast genomes of Amaranthus species. It advances our understanding of chloroplast genome evolution and provides insight into the phylogenetic relationships of dicot plants. Additionally, these findings are a worthwhile input for subsequent research on the optimization of genes and expression improvement in Amaranthus plants.

Conflict of interest

The authors declare no conflict of interest.

References

1. Das, S., *Amaranthus: a promising crop of future*. 2016: Springer.
2. Iamonico, D., (2023). *Nomenclature survey of the genus Amaranthus (Amaranthaceae): 12 questions about Amaranthus caudatus.* Plants. 12(7): p. 1566.
3. Han, J., C. Lin, T. Zhu, Y. Liu, J. Yan, Z. Qi, and X. Yan, (2025). *Comprehensive Chloroplast Genomic Insights into Amaranthus: Resolving the Phylogenetic and Taxonomic Status of A. powellii and A. bouchonii.* Plants. 14(5): p. 649.
4. Chaney, L., R. Mangelson, T. Ramaraj, E.N. Jellen, and P.J. Maughan, (2016). *The complete chloroplast genome sequences for four Amaranthus species (Amaranthaceae).* Applications in Plant Sciences. 4(9): p. 1600063.
5. Stetter, M.G. and K.J. Schmid, (2017). *Analysis of phylogenetic relationships and genome size evolution of the Amaranthus genus using GBS indicates the ancestors of an ancient crop.* Molecular phylogenetics and evolution. 109: p. 80-92.
6. Deng, Y., Y. Luo, Y. He, X. Qin, C. Li, and X. Deng, (2020). *Complete chloroplast genome of Michelia shiluensis and a comparative analysis with four Magnoliaceae species.* Forests. 11(3): p. 267.
7. Walker, J.F., R.K. Jansen, M.J. Zanis, and N.C. Emery, (2015). *Sources of inversion variation in the small single copy (SSC) region of chloroplast genomes.*
8. Ang, K.S., S. Kyriakopoulos, W. Li, and D.-Y. Lee, (2016). *Multi-omics data driven analysis establishes reference codon biases for synthetic gene design in microbial and mammalian cells.* Methods. 102: p. 26-35.
9. Ikemura, T., (1985). *Codon usage and tRNA content in unicellular and multicellular organisms.* Molecular biology and evolution. 2(1): p. 13-34.
10. Gouy, M. and C. Gautier, (1982). *Codon usage in bacteria: correlation with gene expressivity.* Nucleic acids research. 10(22): p. 7055-7074.
11. Cheng, H., J. Li, H. Zhang, B. Cai, Z. Gao, Y. Qiao, and L. Mi, (2017). *The complete chloroplast genome sequence of strawberry (Fragaria× ananassa Duch.) and comparison with related species of Rosaceae.* PeerJ. 5: p. e3919.
12. Duan, H., Q. Zhang, C. Wang, F. Li, F. Tian, Y. Lu, Y. Hu, H. Yang, and G. Cui, (2021). *Analysis of codon usage patterns of the chloroplast genome in Delphinium grandiflorum L. reveals a preference for AT-ending codons as a result of major selection constraints.* PeerJ. 9: p. e10787.
13. Zhang, R., L. Zhang, W. Wang, Z. Zhang, H. Du, Z. Qu, X.-Q. Li, and H. Xiang, (2018). *Differences in codon usage bias between photosynthesis-related genes and genetic system-related genes of chloroplast genomes in cultivated and wild solanum species.* International Journal of Molecular Sciences. 19(10): p. 3142.
14. Rong, X., L. Huang, and J. Shen, (2024). *Analysis of Synonymous Codon Usage Bias in the Chloroplast Genome of Rhododendron farrerae.*

Vol. 20 (Supplementry Issue 1A), March 2026, ISSN 0973-8916 (Print)., 2230-7303 (Online)
10.5530/ctbp.2026.1s.27

15. Wang, Z., Q. Cai, Y. Wang, M. Li, C. Wang, Z. Wang, C. Jiao, C. Xu, H. Wang, and Z. Zhang, (2022). *Comparative analysis of codon bias in the chloroplast genomes of theaceae species.* Frontiers in Genetics. 13: p. 824610.
16. Viljoen, E., D.A. Odeny, M.P. Coetzee, D.K. Berger, and D.J. Rees, (2018). *Application of chloroplast phylogenomics to resolve species relationships within the plant genus Amaranthus.* Journal of Molecular Evolution. 86(3): p. 216-239.
17. Chi, X., F. Zhang, Q. Dong, and S. Chen, (2020). *Insights into comparative genomics, codon usage bias, and phylogenetic relationship of species from Biebersteiniaceae and Nitrariaceae based on complete chloroplast genomes.* Plants. 9(11): p. 1605.
18. Li, B. and Y. Zheng, (2018). *Dynamic evolution and phylogenomic analysis of the chloroplast genome in Schisandraceae.* Scientific Reports. 8(1): p. 9285.
19. Sharp, P.M. and E. Cowe, (1991). *Synonymous codon usage in Saccharomyces cerevisiae.* Yeast. 7(7): p. 657-678.
20. Wang, Z., B. Xu, B. Li, Q. Zhou, G. Wang, X. Jiang, C. Wang, and Z. Xu, (2020). *Comparative analysis of codon usage patterns in chloroplast genomes of six Euphorbiaceae species.* PeerJ. 8: p. e8251.
21. Sharp, P.M. and W.-H. Li, (1986). *An evolutionary perspective on synonymous codon usage in unicellular organisms.* Journal of molecular evolution. 24: p. 28-38.
22. Sharp, P.M. and W.-H. Li, (1987). *The codon adaptation index-a measure of directional synonymous codon usage bias, and its potential applications.* Nucleic acids research. 15(3): p. 1281-1295.
23. Liu, H., Y. Lu, B. Lan, and J. Xu, (2020). *Codon usage by chloroplast gene is bias in Hemiptelea davidii.* Journal of Genetics. 99: p. 1-11.
24. Sueoka, N., (1988). *Directional mutation pressure and neutral molecular evolution.* Proceedings of the National Academy of Sciences. 85(8): p. 2653-2657.
25. Sueoka, N., (1999). *Two aspects of DNA base composition: G+ C content and translation-coupled deviation from intra-strand rule of A= T and G= C.* Journal of Molecular Evolution. 49: p. 49-62.
26. Wright, F., (1990). *The 'effective number of codons' used in a gene.* Gene. 87(1): p. 23-29.
27. WU, Y.-q., Z.-y. LI, D.-q. ZHAO, and T. Jun, (2018). *Comparative analysis of flower-meristem-identity gene APETALA2 (AP2) codon in different plant species.* Journal of Integrative Agriculture. 17(4): p. 867-877.
28. Jiang, Y., F. Deng, H. Wang, and Z. Hu, (2008). *An extensive analysis on the global codon usage pattern of baculoviruses.* Archives of virology. 153: p. 2273-2282.
29. Kawabe, A. and N.T. Miyashita, (2003). *Patterns of codon usage bias in three dicot and four monocot plant species.* Genes & genetic systems. 78(5): p. 343-352.
30. Zhang, W.J., J. Zhou, Z.F. Li, L. Wang, X. Gu, and Y. Zhong, (2007). *Comparative analysis of codon usage patterns among mitochondrion, chloroplast and nuclear genes in Triticum aestivum L.* Journal of Integrative Plant Biology. 49(2): p. 246-254.
31. Wan, X.-F., D. Xu, A. Kleinhofs, and J. Zhou, (2004). *Quantitative relationship between synonymous codon usage bias and GC composition across unicellular genomes.* BMC evolutionary biology. 4: p. 1-11.
32. Sueoka, N., (1999). *Translation-coupled violation of Parity Rule 2 in human genes is not the cause of heterogeneity of the DNA G+ C content of third codon position.* Gene. 238(1): p. 53-58.
33. Xiang, H., R. Zhang, R.R. Butler III, T. Liu, L. Zhang, J.-F. Pombert, and Z. Zhou, (2015). *Comparative analysis of codon usage bias patterns in microsporidian genomes.* PloS one. 10(6): p. e0129223.
34. James, F.C. and C.E. McCulloch, (1990). *Multivariate analysis in ecology and systematics: panacea or Pandora's box?* Annual review of Ecology and Systematics: p. 129-166.
35. Sharp, P.M. and K.M. Devine, (1989). *Codon usage and gene expression level in*

Vol. 20 (Supplementry Issue 1A), March 2026, ISSN 0973-8916 (Print)., 2230-7303 (Online)
10.5530/ctbp.2026.1s.27

Dictyosteiium discoidtum: highly expressed genes do [prefer [optimal codons. Nucleic Acids Research. 17(13): p. 5029-5040.
36. Zhou, M., W. Long, and X. Li, (2008). *Patterns of synonymous codon usage bias in chloroplast genomes of seed plants.* Forestry Studies in China. 10: p. 235-242.
37. Kokate, P.P., S.M. Techtmann, and T. Werner, (2021). *Codon usage bias and dinucleotide preference in 29 Drosophila species.* G3. 11(8): p. jkab191.
38. Wang, Y., D. Jiang, K. Guo, L. Zhao, F. Meng, J. Xiao, Y. Niu, and Y. Sun, (2023). *Comparative analysis of codon usage patterns in chloroplast genomes of ten Epimedium species.* BMC genomic data. 24(1): p. 3.
39. Wang, L., H. Xing, Y. Yuan, X. Wang, M. Saeed, J. Tao, W. Feng, G. Zhang, X. Song, and X. Sun, (2018). *Genome-wide analysis of codon usage bias in four sequenced cotton species.* PloS one. 13(3): p. e0194372.
40. Zhao, Y. and S. Zhang, (2024). *Comparative Analysis of Codon Usage Bias in Six Eimeria Genomes.* International Journal of Molecular Sciences. 25(15): p. 8398.
41. Ahmad, T., G. Sablok, T.V. Tatarinova, Q. Xu, X.-X. Deng, and W.-W. Guo, (2013). *Evaluation of codon biology in citrus and Poncirus trifoliata based on genomic features and frame corrected expressed sequence tags.* DNA research. 20(2): p. 135-150.
42. Yang, X., Y. Wang, W. Gong, and Y. Li, (2024). *Comparative Analysis of the Codon Usage Pattern in the Chloroplast Genomes of Gnetales Species.* International Journal of Molecular Sciences. 25(19): p. 10622.