

Article

Codon Usage Analyses Reveal the Evolutionary Patterns among Plastid Genes of Saxifragales at a Larger-Sampling Scale

De Bi ^{1,†}, Shiyun Han ^{2,3,†} , Jun Zhou ¹, Maojin Zhao ¹, Sijia Zhang ^{2,3} and Xianzhao Kan ^{2,3,*} ¹ Suzhou Polytechnic Institute of Agriculture, Suzhou 215000, China² Anhui Provincial Key Laboratory of the Conservation and Exploitation of Biological Resources, College of Life Sciences, Anhui Normal University, Wuhu 241000, China³ The Institute of Bioinformatics, College of Life Sciences, Anhui Normal University, Wuhu 241000, China

* Correspondence: xianzhao@ahnu.edu.cn; Tel.: +86-139-5537-2268

† These authors contributed equally to this work.

Abstract: Saxifragales is a 15-family order of early-divergent Eudicots with a rich morphological diversity and an ancient rapid radiation. Codon usage bias (CUB) analyses have emerged as an essential tool for understanding the evolutionary dynamics in genes. Thus far, the codon utilization patterns had only been reported in four separate genera within Saxifragales. This study provides a comprehensive assessment of the codon manipulation based on 50 plastid genes, covering 11 constituent families at a larger sampling scale. Our results first showed a high preference for AT bases and AT-ending codons. We then used effective number of codons (ENC) to assess a range of codon bias levels in the plastid genes. We also detected high-informative intrafamilial differences of ENC in three families. Subsequently, parity rule 2 (PR2) plot analyses revealed both family-unique and order-shared bias patterns. Most importantly, the ENC plots and neutrality analyses collectively supported the dominant roles of selection in the CUB of Saxifragales plastid genes. Notably, the phylogenetic affinities inferred by both ML and BI methods were consistent with each other, and they all comprised two primary clades and four subclades. These findings significantly enhance our understanding of the evolutionary processes of the Saxifrage order, and could potentially inspire more CUB analyses at higher taxonomic levels.



Citation: Bi, D.; Han, S.; Zhou, J.; Zhao, M.; Zhang, S.; Kan, X. Codon Usage Analyses Reveal the Evolutionary Patterns among Plastid Genes of Saxifragales at a Larger-Sampling Scale. *Genes* **2023**, *14*, 694. <https://doi.org/10.3390/genes14030694>

Academic Editor: Zhiqiang Wu

Received: 15 February 2023

Revised: 7 March 2023

Accepted: 10 March 2023

Published: 11 March 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: Saxifragales; plastid genes; phylogeny; codon usage

1. Introduction

The order Saxifragales is the representative of parts of the early-divergent branches of the Eudicot [1–3]. Fossil data have suggested a rapid radiation of this group [2,4], and currently, it includes ca. 2500 taxa from 15 families [5], exhibiting a high degree of diversity in morphology (e.g., succulents, shrubs, vines, and trees) [6–8]. Although the monophyly of this order has been widely accepted [9–14], its internal phylogenetic affinities were more elusive in earlier studies, which might be caused by an ancient, rapid radiation and limited data resources [15,16]. In particular, the circumscriptions of Saxifragales have been regarded as a major surprise in plant systematics [17]. Notably, based on a 50,845 bp concatenated dataset from nuclear, mitochondrial, and plastid sequences, Jian et al. [1] firstly inferred three major clades of Saxifragales: (1) Peridiscaceae, (2) Paeoniaceae + “woody group” (Altingiaceae, Cercidiphyllaceae, Daphniphyllaceae, and Hamamelidaceae), and (3) Crassulaceae alliance (Crassulaceae, Haloragaceae, Penthoraceae, and Tetracarpaceae) + Saxifragaceae alliance (Saxifragaceae, Grossulariaceae, Iteaceae, and Pterostemonaceae). Thus far, many analyses have supported this backbone phylogeny, and greatly contributed to insights into the interfamilial relationships among these lineages [18–20].

Saxifragales is considered to be a pretty good demonstration to explore diverse evolutionary patterns among angiosperms [17,20]. Along with rapid progress in molecular

methods, the evolutionary history of Saxifragales has undergone many intense explorations [18–23]. Therefore, plastid genome data exemplifies the advantages of affordability, high sequencing efficiency, and rich evolutionary information. Numerous studies on plastomes have been dedicated to explore genomic organizations, interspecific differences, intracellular gene transfer, RNA editing, genetic engineering, and evolutionary analysis [24–28]. At present, plastomes have been regarded as a powerful tool not only for phylogeny inference, but also for DNA barcoding of plants [24,29–33]. Furthermore, as for plastid genes, the patterns of synonymous codon usage could reflect the driving factors in evolution among different taxa [34–36].

Among the plastomes of Saxifragales, under the rapid development of DNA sequencing methodology, there have been several advances in understanding the nucleotide and structural variations in gene content, intron indels, plastomic tRNAs secondary structures, and hypervariable regions, etc. [20,21]. Furthermore, our previous work, involving 208 Saxifragales plastomes, has suggested several family- and lineage-specific markers for this order. However, the evolutionary dynamics of specific plastid genes in Saxifragales have not been systematically investigated yet.

As a ubiquitous phenomenon, codon usage bias (CUB), referring to the unbalanced employment of synonymous codons [37–40], can provide crucial information for deciphering the evolutionary patterns of genes [41–43]. Unravelling such bias could be intriguing, usually due to the comprehensive effects of mutation, selection, and genetic drift [44]. Most significantly, the preference of codon usage is suggested to differ among organisms and even within a certain gene [43,45]. Since the early report of codon catalog usage from Grantham et al. [46], a series of indices have been proposed to quantify CUB, containing effective number of codons (ENC) [47], relative synonymous codon usage (RSCU) [48], parity rule 2 (PR2) plot [49], neutrality plot [50], etc. Within Saxifragales, the patterns of codon usage have been reported for several taxa, including Crassulaceae (*Aeonium*, *Monanthes*, and *Crassula*), and Saxifragaceae (*Saxifraga*) [19,51,52]. Of note, our previous study on *Crassula* revealed a genera-specific distribution of ENC values in *matK* genes [52]. To reach a deeper understanding of Saxifragales plastomes, further explorations on a larger sampling scale will thus be crucial.

The present work aims to comprehensively assess the evolutionary dynamics of Saxifragales plastid genes, with all publicly available sequence data. Employing multiple indices, the goals of this study are to elucidate (1) the various codon bias patterns among plastid genes of Saxifragales, (2) the dominant factors in shaping the diverse CUB patterns, and (3) the phylogenetic affinities within the saxifrage order. Most importantly, it is hoped that this study will broaden and deepen our understanding of the Saxifragales evolution, as well as the patterns of codon usage at higher taxonomic levels.

2. Materials and Methods

2.1. Data Retrieval, Reannotation, Alignment, and Compositional Analysis

To obtain the widest sampling of available Saxifragales plastomes, 361 plastomes were acquired from the GenBank of NCBI database (<https://www.ncbi.nlm.nih.gov> accessed on 20 February 2023), including 24 sequences that were reported in our previous studies. Overall, our sampling involved 258 taxa from 64 genera in 11 families of Saxifragales, which are listed in Table S1 along with their accession numbers and genome sizes. Then, under the principles of (1) including the initial codon, the terminal codon, and no internal stop codons, and (2) being a multiple of three and longer than 300 bp in gene size, a total of 50 protein-coding genes (PCG) were filtrated from each plastome for further analyses. After careful checking and correcting of gene annotations by GeSeq [53], CPGAVAS2 [54], and BLAST [55], coding sequences (CDS) alignment was performed using MAFFT v7.505 [56]. The GC contents of the CDSs at different positions, including GC1, GC2, GC12, GC3, and total GC, were also calculated using DAMBE v7.2.25 [57].

2.2. The Calculations of Codon Usage Indices

Concurring with McInerney [58], Roth et al. [59], Wang et al. [60], and others, the termination codons were removed from the CDS matrix for further analyses. Using CodonW v1.4.4 for each gene, we focused on the following CUB parameters: RSCU, ENC, and GC content of synonymous codons at the 3rd position (GC3s) [61].

RSCU represents the proportion of the actual frequency to the expected frequency for a certain codon, and is routinely applied to detect differences in codon use among genes [44,48]. For this value, equality to, greater than, and less than 1 indicates no, positive, and negative bias, respectively.

ENC values are an excellent estimator for absolute synonymous codon bias and reflect the degree of preferential codons in genes [44,62]. With a range of 20–61, ENC values are inversely related to the extent of CUB [47]. In addition, considerable bias can be observed for values no higher than 35 [63,64].

2.3. Graphic Analyses

To visually assess the driving factors of the CUB patterns among plastid genes within Saxifragales, several plot analyses, including a PR2 plot, neutrality plot, and ENC–GC3s plot, were carried out using ggplot2 in R v4.0.1 [65].

The PR2 plot analysis is performed to estimate the influences of mutation and selection on the third position of codons, which is usually confined to the four-fold degenerate codons [42,44,49]. This plot is generated by GC bias [$G3/(G3 + C3)$] and AU bias [$A3/(A3 + U3)$] as the x - and y -axis, respectively [66,67]. In detail, the center-located points suggest no bias and equal effects from both mutation and selection, whereas the off-center splashes would manifest the corresponding orientations and degrees of bias [42,44,68]. More specially, if $A3 + T3$ is equal to $G3 + C3$, then the preferential utilization of codons might be entirely influenced by mutational pressure [69].

As a routine analysis to infer the determinants of CUB, ENC–GC3s plot is drawn by GC3s (x -axis), ENC values (y -axis), and a standard curve (calculated following Wright et al. [47]). In this plot, the splashes will fall near the curve when the CUB patterns are mainly driven by mutational effects. Instead, if the codon preference is determined by selection or other factors, the points will be distant from the curve [70–72].

Likewise, a neutrality plot, based on the regression analysis of GC12 against GC3, is also utilized to evaluate the influences from mutation and selection [50]. Furthermore, if the slopes of the regression line are close to 0, selection is suggested to be the decisive force governing the codon usage, while being close to 1 is indicative of the dominating impact of mutation in CUB.

2.4. Phylogenetic Inferences of Saxifragales

To reconstruct the evolutionary affinities within Saxifragales, phylogenetic analyses were performed among all 258 involved species by maximum-likelihood (ML) and Bayesian inference (BI) methods. With two taxa from Rosids as outgroups, a sequence matrix of 79 plastid genes from 363 individuals was built using MAFFT v7.505 for multiple alignments [56] and PhyloSuite v1.2.1 for concatenation [73].

In the ML inference, RAxML v8.2.12 was employed with 50 runs and 1000 bootstrap replicates using the GTRCAT model [74]. Convergence was then examined using the command “-I autoMRE”. After determining the best models in ModelTest-NG [75], MrBayes v3.2.7a was run to generate the Bayesian cladograms for 10,000,000 generations, with two runs using four chains each (sampling every 1000 generations). Furthermore, we tested the convergence of independent runs using Tracer v1.7.1 [76].

3. Results

3.1. Base Composition Comparisons of Plastid Genes among Saxifragales

To assess the impact of base compositions on codon bias in genes in the 361 individuals of the Saxifragales group, the GC content of 50 plastomic genes was counted individually.

Overall, Table S2 shows extensive variation in GC content (GC123) among the genes, from a low of 28.91% in the *ycf1* gene of *Gonocarpus micranthus* from Haloragaceae to a high of 48.31% in the *rps11* gene in *Crassula volkensii* from Crassulaceae. Significantly, the GC and GC3 contents of all investigated genes were less than 50%, suggesting that AT nucleotides and AT-ending codons are often preferred within Saxifragales. Interestingly, in the first and second positions of codons, despite the fact that most genes showed an AT bias, some genes exhibited a GC bias. For instance, across 361 individuals, the following genes harbored GC1 > 50%: *atpB*, *clpP*, *petA*, *psaA*, *psbB*, *psbC*, *psbD*, *rbcL*, *rps11*, and *rps12*, while *rps11* and *rps7* showed GC2 > 50%.

To explore the interfamilial differences, the comparison of base compositions among the 11 families was performed. Our findings, summarized in Table 1 with mean and standard deviation, revealed that the patterns of GC content among different families were remarkably similar. Notably, the average GC values among three major clades of Saxifragales were highest in the Paeoniaceae + woody clade, with Paeoniaceae at the top with 39.46%, followed by Altingiaceae at 39.29%, Daphniphyllaceae at 39.20%, and Hamamelidaceae at 39.15%. Moreover, all 11 families exhibited a similar trend of GC-distribution across the three codon positions, with GC1 > GC2 > GC3.

Table 1. Base compositions of 50 plastid genes among 11 families of Saxifragales.

Family	GC%	GC1%	GC2%	GC3%
Altingiaceae	39.29 ± 3.31	48.47 ± 5.45	40.05 ± 5.75	29.35 ± 3.21
Cercidiphyllaceae	38.67 ± 3.44	48.05 ± 5.48	39.65 ± 5.83	28.31 ± 3.59
Crassulaceae	38.42 ± 3.53	47.84 ± 5.56	39.54 ± 5.76	27.89 ± 3.69
Daphniphyllaceae	39.20 ± 3.25	48.43 ± 5.38	40.01 ± 5.57	29.15 ± 3.53
Grossulariaceae	38.83 ± 3.44	48.18 ± 5.52	39.82 ± 5.58	28.49 ± 3.74
Haloragaceae	38.79 ± 3.45	48.15 ± 5.51	39.78 ± 5.61	28.45 ± 3.57
Hamamelidaceae	39.15 ± 3.19	48.36 ± 5.51	40.11 ± 5.64	28.97 ± 3.34
Iteaceae	38.59 ± 3.59	48.04 ± 5.76	39.87 ± 5.83	27.87 ± 3.96
Paeoniaceae	39.46 ± 3.47	48.26 ± 5.34	39.86 ± 5.55	30.27 ± 3.60
Penthoraceae	38.75 ± 3.30	48.26 ± 5.56	39.88 ± 5.51	28.12 ± 3.57
Saxifragaceae	38.77 ± 3.43	48.23 ± 5.44	39.76 ± 5.77	28.33 ± 3.60

3.2. Codon Usage Bias Indices of Plastid Genes among Saxifragales

The aim of this study was to examine the codon preference patterns among 361 Saxifragales individuals by calculating the ENC values of the 50 plastid genes (Table S3). As shown in Table 2, the comparisons were initially performed at the intergenic level. The overall ENC values ranged from 39.23 ± 2.50 to 54.91 ± 3.38, with the lowest ENC found in the *rpl16* (39.23 ± 2.50) and the highest in the *pafI* gene (54.91 ± 3.38). The genes with lower ENC values, such as *rpl16*, *psbA*, and *rps8*, had stronger codon preferences, while those with higher ENC values, such as *pafI*, *ycf2*, and *rpl2*, showed weaker codon bias.

Table 2. The mean ± SD ENC values of 50 plastid genes of Saxifragales.

Gene	ENC Values	Gene	ENC Values	Gene	ENC Values	Gene	ENC Values	Gene	ENC Values
<i>rpl16</i>	39.23 ± 2.50	<i>ndhI</i>	44.48 ± 3.03	<i>atpF</i>	47.13 ± 3.09	<i>rps2</i>	48.28 ± 1.81	<i>rpoC1</i>	50.16 ± 0.63
<i>psbA</i>	39.67 ± 1.82	<i>rps7</i>	44.77 ± 0.67	<i>atpB</i>	47.21 ± 1.78	<i>ycf1</i>	48.53 ± 1.09	<i>rpoC2</i>	50.54 ± 0.77
<i>rps8</i>	39.92 ± 3.46	<i>atpI</i>	45.28 ± 2.24	<i>ndhC</i>	47.27 ± 5.15	<i>rps4</i>	48.72 ± 1.40	<i>petA</i>	50.78 ± 2.10
<i>rps14</i>	40.60 ± 1.56	<i>rps12</i>	45.56 ± 0.95	<i>ndhD</i>	47.66 ± 1.36	<i>cemA</i>	48.76 ± 2.68	<i>ndhJ</i>	50.86 ± 2.45
<i>petD</i>	41.32 ± 2.81	<i>ccsA</i>	45.77 ± 2.42	<i>accD</i>	47.74 ± 2.03	<i>psaA</i>	49.24 ± 1.19	<i>ndhK</i>	51.21 ± 1.53
<i>ndhG</i>	41.66 ± 3.06	<i>psbB</i>	46.21 ± 1.06	<i>rps3</i>	47.80 ± 2.36	<i>rpoB</i>	49.45 ± 0.93	<i>atpE</i>	52.19 ± 3.73
<i>rpl20</i>	42.05 ± 2.88	<i>rps11</i>	46.40 ± 3.25	<i>ndhB</i>	47.92 ± 0.67	<i>rpoA</i>	49.62 ± 1.27	<i>clpP</i>	52.68 ± 1.53
<i>ndhA</i>	43.59 ± 1.06	<i>atpA</i>	46.66 ± 1.27	<i>rpl14</i>	47.94 ± 3.61	<i>ndhE</i>	49.68 ± 4.29	<i>rpl2</i>	52.72 ± 1.15
<i>psbD</i>	44.30 ± 1.33	<i>rbcL</i>	46.80 ± 1.89	<i>ndhH</i>	48.09 ± 1.57	<i>matK</i>	49.78 ± 2.19	<i>ycf2</i>	52.98 ± 0.28
<i>petB</i>	44.42 ± 2.04	<i>psbC</i>	46.87 ± 1.06	<i>psaB</i>	48.20 ± 1.09	<i>pafII</i>	50.11 ± 1.95	<i>pafI</i>	54.91 ± 3.38

Then, to gain further insights into the interfamilial differences, the ENC values of 50 plastid genes were calculated as mean \pm SD for each family and compared. The results displayed a high degree of diversity among 11 families (Table S3 and Figure 1). The most striking heterogeneity was identified in *ndhE* between Paeoniaceae (mean value: 58.44) and other ten families (mean values: 46.43–50.59), indicating a considerably lower bias in the former. Meanwhile, evidence from some taxa also supported the existence of a higher codon bias, particularly in the *rps8* genes of 26 Crassulaceae taxa and *Boykinia aconitifolia* (Saxifragaceae), with ENC < 35, suggesting significant codon usage preferences.

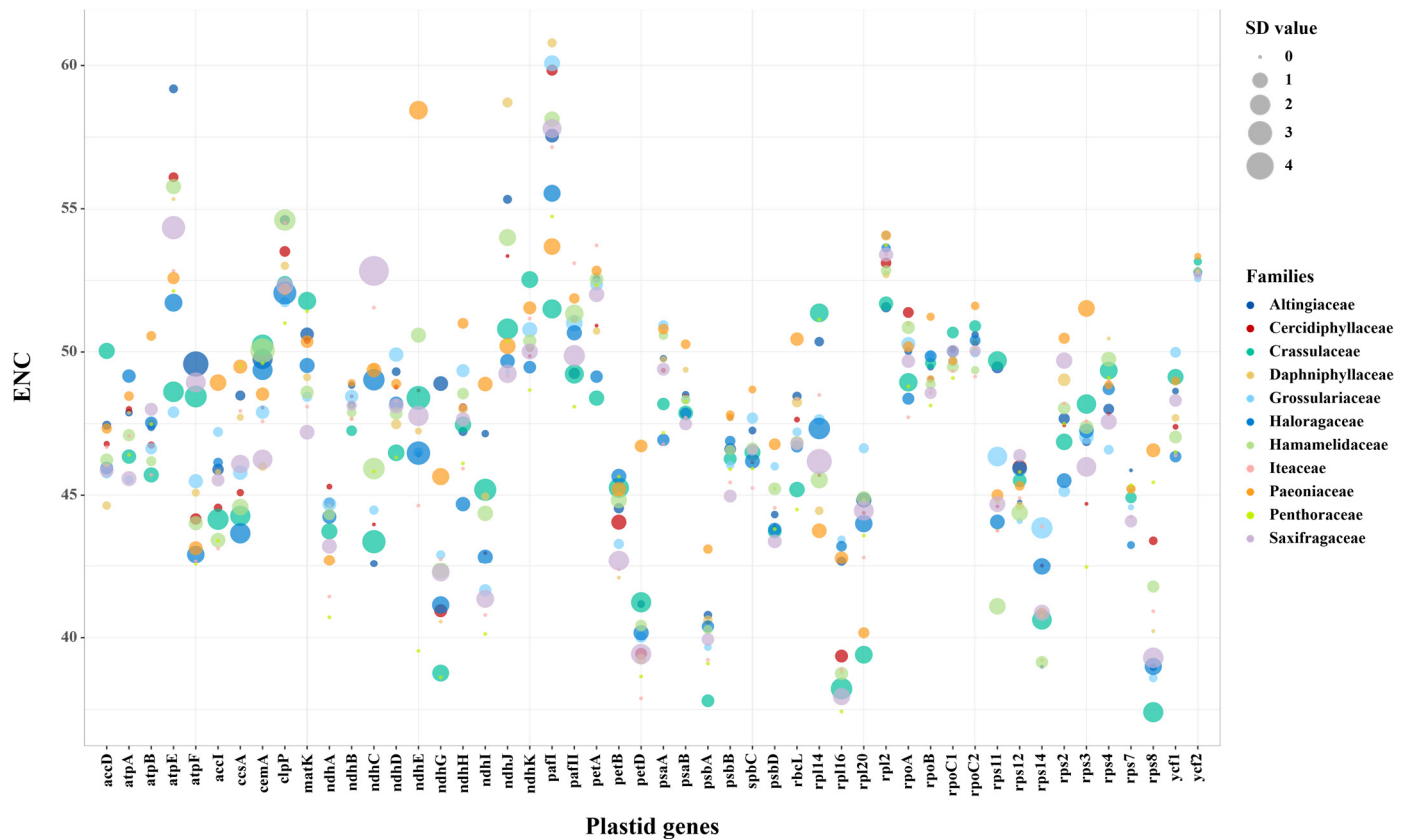


Figure 1. The mean \pm SD values of ENC for each gene in each family of Saxifragales. The center, size, and color of circle denote the corresponding average value, SD value, and family, respectively.

Furthermore, to explore intrafamilial diversities, we also compared the ENC values of genes within each family. Overall, a high level of variation was detected with the SD value ranging from 0 to 4.93. Remarkably, the most informative differences were detected in three families. For the *ndhE* genes in Haloragaceae, clade A and *Glischrocaryon aureum* had ENC values ranging from 43.20 and 45.57, while clade B had higher values of 48.78 to 50.49 (Figure 2a). Likewise, in the *cemA* genes of Hamamelidaceae, except for clade A with ENC of 50.88–53.68, clade B, C, D, E, and F similarly possessed lower a ENC range of 45.57–48.11 (Figure 2b). Strikingly, the most substantial intrafamilial difference was observed in the *atpF* genes of Altingiaceae (Figure 2c), where clade A could be clearly distinguished by its much higher ENC values of 51.42–52.07, compared to clade B and C, which had ENC values of 44.74–44.92.

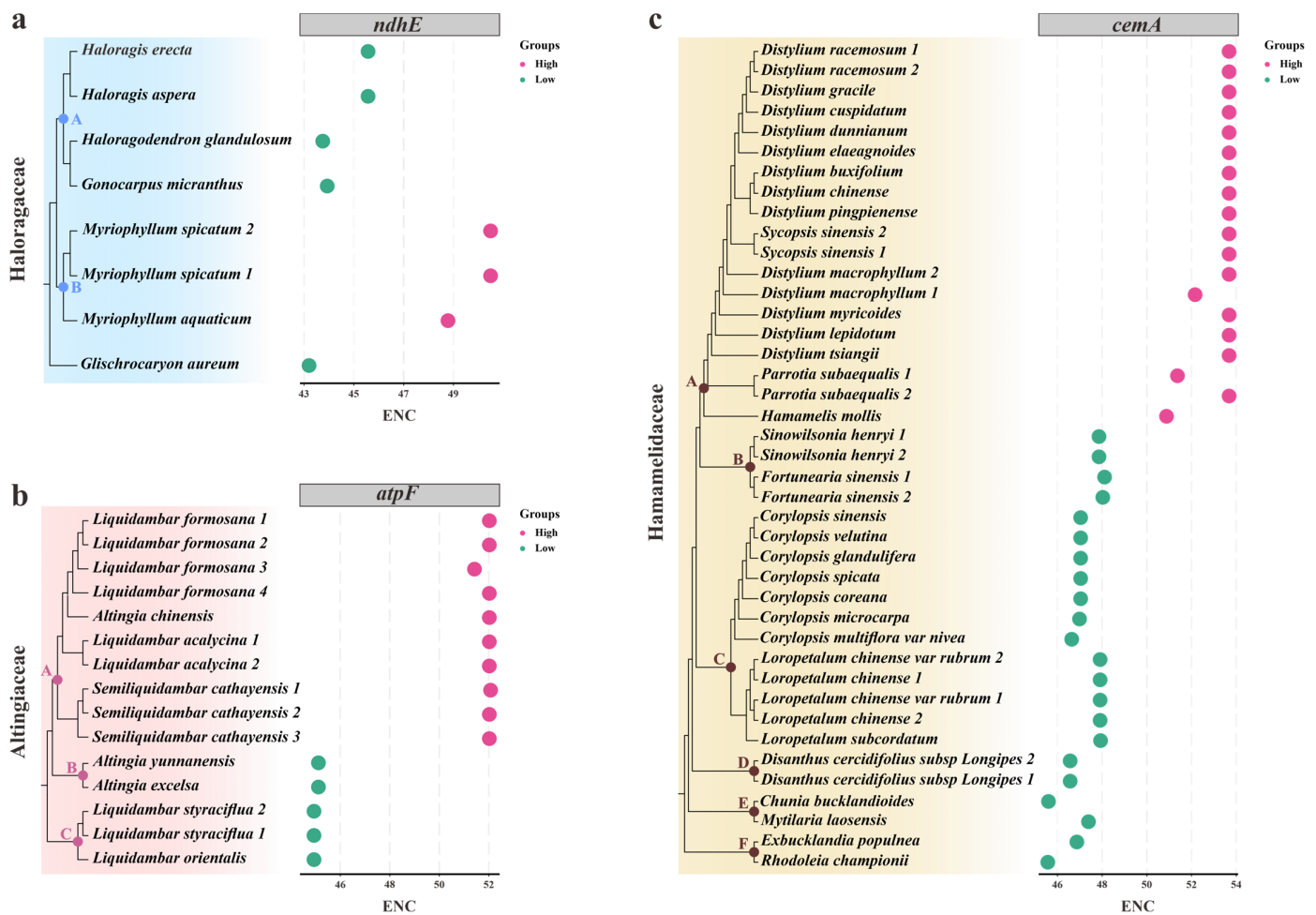


Figure 2. The different ENC patterns of plastid genes in three families. (a) The ENC values of *ndhE* genes in Haloragaceae taxa. (b) The ENC values of *atpF* genes in Altingiaceae taxa. (c) The ENC values of *cemA* genes in Hamamelidaceae taxa. Different clades are marked with dots in the nodes.

The RSCU index was also used to assess codon usage bias. We performed the calculations on a concatenated matrix of 50 plastid genes from 11 families (Table S4). In general, the plastid genes of Saxifragales harbored RSCU values ranging from 0.32 (CGC) to 2.16 (AGA). It was noted that 93.25% of the preferred codons with RSCU > 1 ended in AT, indicating an overwhelming bias for these codons in the CUB patterns. We further performed clustering analyses among the codons and families (Figure 3). Intriguingly, all families were clustered into the same branching patterns as those reported in the phylogeny trees of Saxifragales [18–20], with the exception of the positions of Iteaceae and Paeoniaceae.

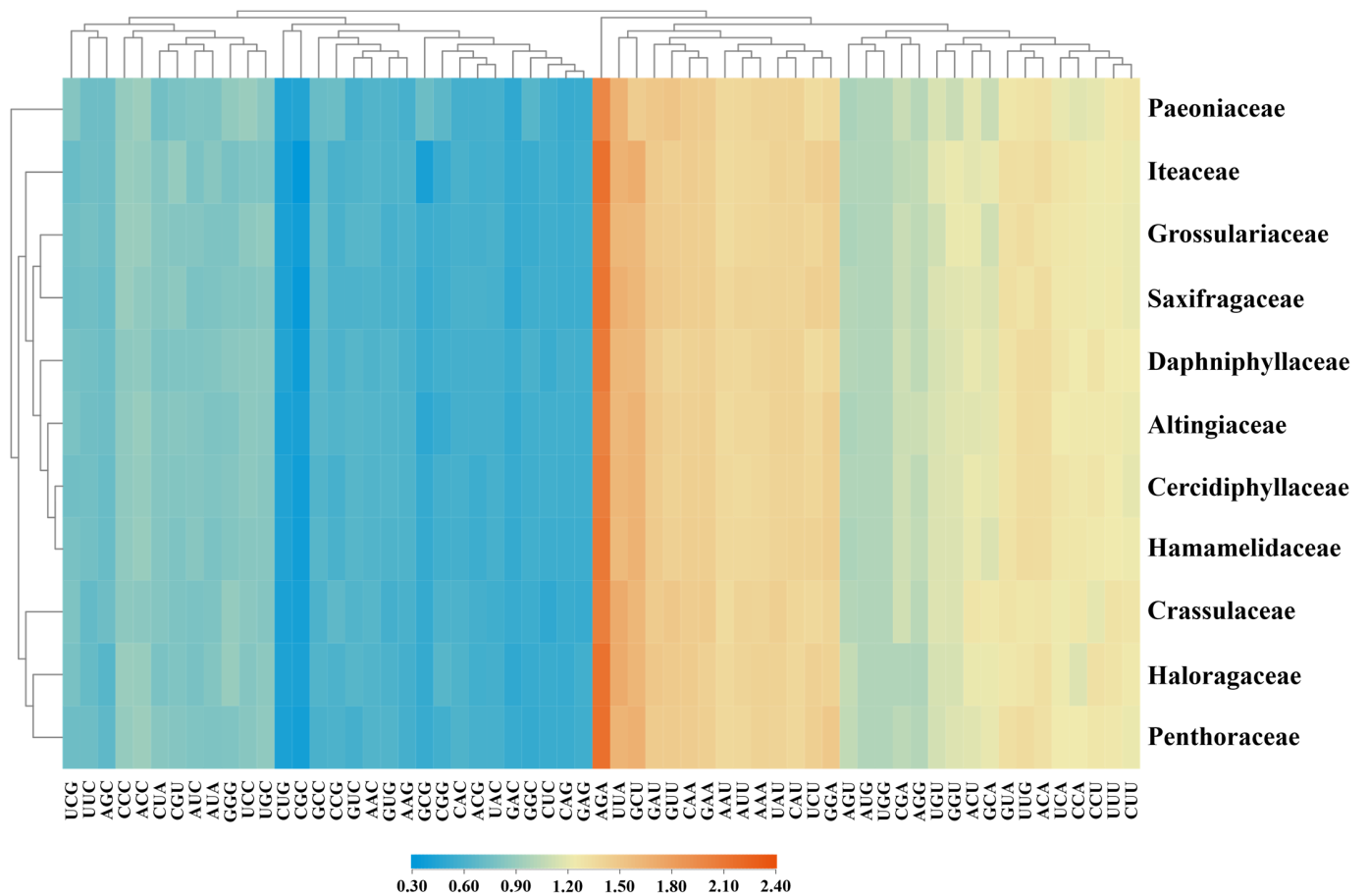


Figure 3. Heatmap of the RSCU values of the 50 plastid genes among Saxifragales. The horizontal axis represents the cluster of the RSCU values of codons, and the vertical axis represents the cluster of the constituent families.

3.3. Determinants in Codon Usage of Plastid Genes among Saxifragales

To understand the influences of mutation and selection on the CUB of plastid genes, PR2 bias plots were created focusing on the four-fold degenerate codons. Initially, the AT bias and GC bias at the third codon position were individually calculated in each gene (Table S5). The ranges of AT bias (0.29 ± 0.02 to 0.67 ± 0.07) and GC bias (0.21 ± 0.09 to 0.84 ± 0.12) among the 50 genes of Saxifragales revealed a relatively high level of diversity. Then, our plot analyses also allowed the identification of unequal codon usage in different genes and families (Figure S1). Particularly, visibly family-specific variations were observed in two genes. The *petA* gene showed a stronger G bias in Crassulaceae compared to other families (Figure 4a). In addition, the *rps11* gene exhibited a higher T bias in Paeoniaceae than in other families (Figure 4b).

Significantly, several common patterns have also been discovered among the Saxifragales group. As illustrated in Figure 5, all 11 families collectively presented the CT bias in *atpA*, *psbA*, and *psbC*, the CA bias in *clpP* and *rps7*, as well as the GA bias in *rpl14*.

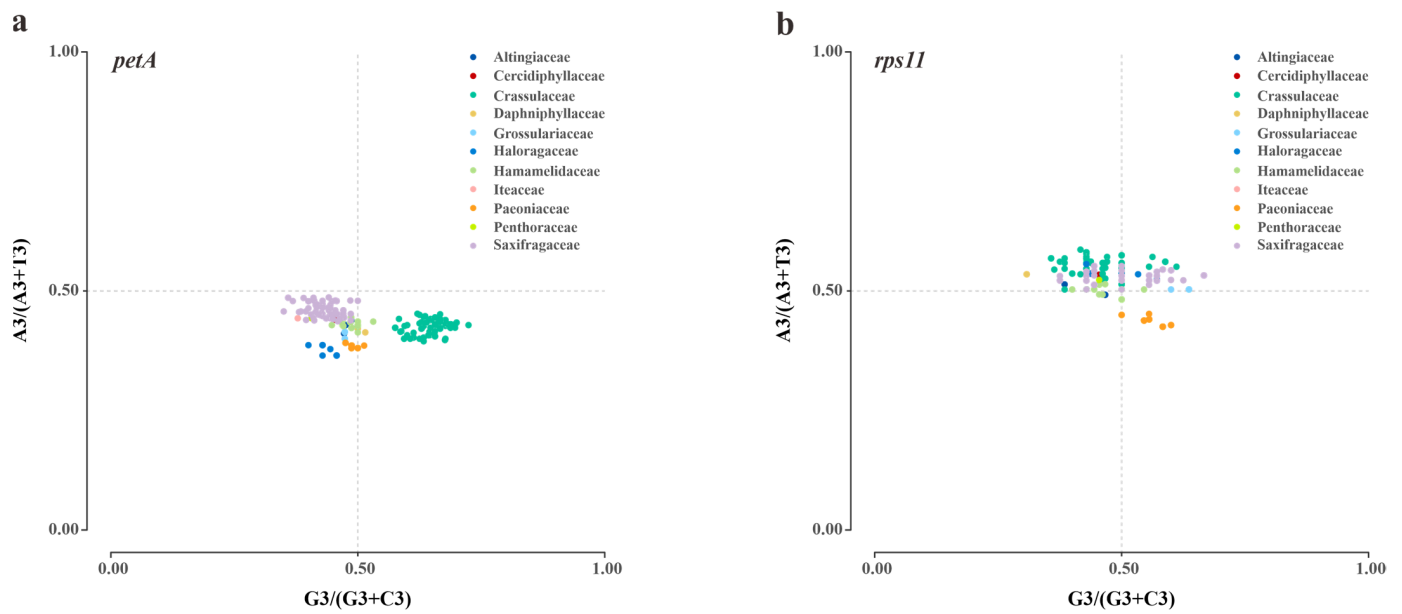


Figure 4. The PR2 bias plots of two plastid genes in Saxifragales. The colors denote the corresponding families. (a) *petA*; (b) *rps11*.

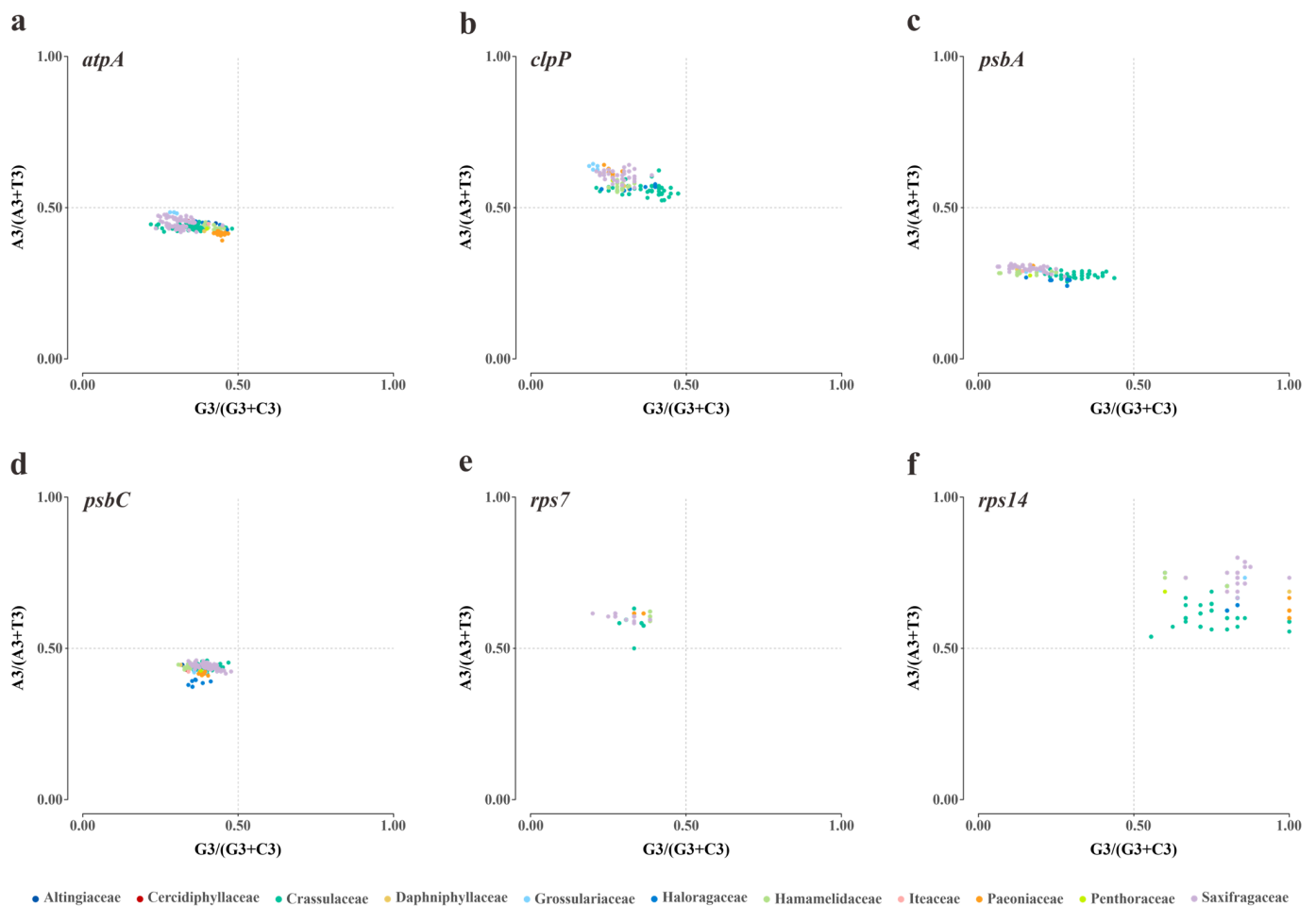


Figure 5. The PR2 bias plots of six plastid genes in Saxifragales. The colors denote the corresponding families. (a) *atpA*; (b) *clpP*; (c) *psbA*; (d) *psbC*; (e) *rps7*; (f) *rps14*.

With the aim of inferring the main governing factors in the codon use patterns of plastid genes, ENC-GC3s plot and neutrality plot analyses were performed across Saxifragales. The results of the ENC plots revealed a similar pattern of point distribution among the 11 families (Figure 6). The GC3s values were relatively narrow (ranging from 0.14 to 0.35), and most of the splashes lay distant from the expected curve, suggesting that selection might play a bigger role in codon bias than mutation. Moreover, the gradient of regression line in the neutrality plot (GC12 vs. GC3) denoted the balance between mutation and selection (Figure 7). Across Saxifragales, the GC12 (0.45 ± 0.05) and GC3 (0.25 ± 0.04) were relatively concentrated, with a weak correlation ($r = 0.02$, $p < 0.05$), suggesting that mutation had a limited impact on the CUB. More importantly, the slopes of all 11 families were close to 0, with non-negative values ranging from 0.00593 (Altingiaceae) to 0.136 (Hamamelidaceae). The findings from both of the neutrality and ENC plots support the dominant role of selection pressure in shaping the CUB of plastid genes within Saxifragales.

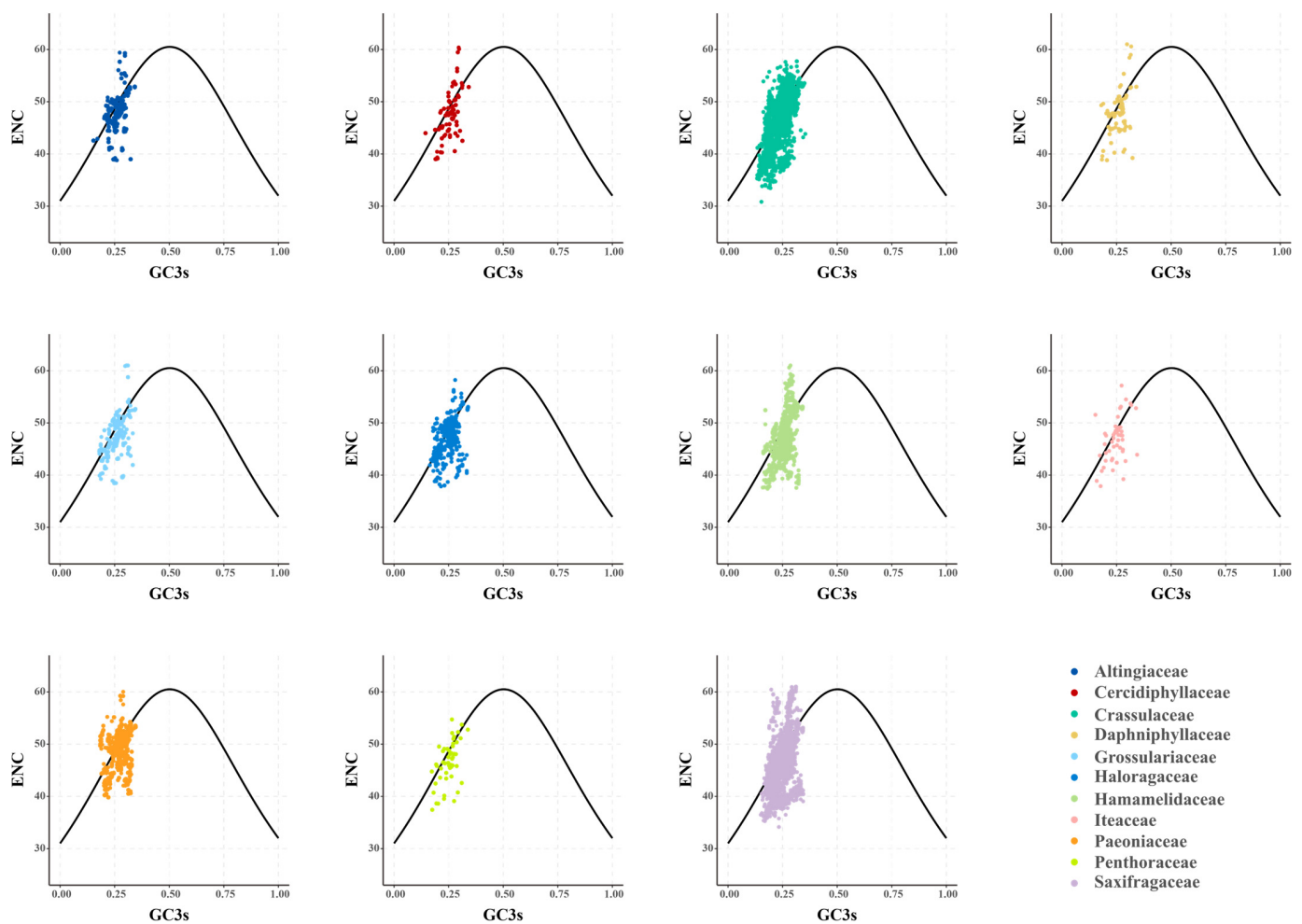


Figure 6. The ENC-GC3s plots of plastid genes in 11 Saxifragales families. The colors denote the corresponding families.

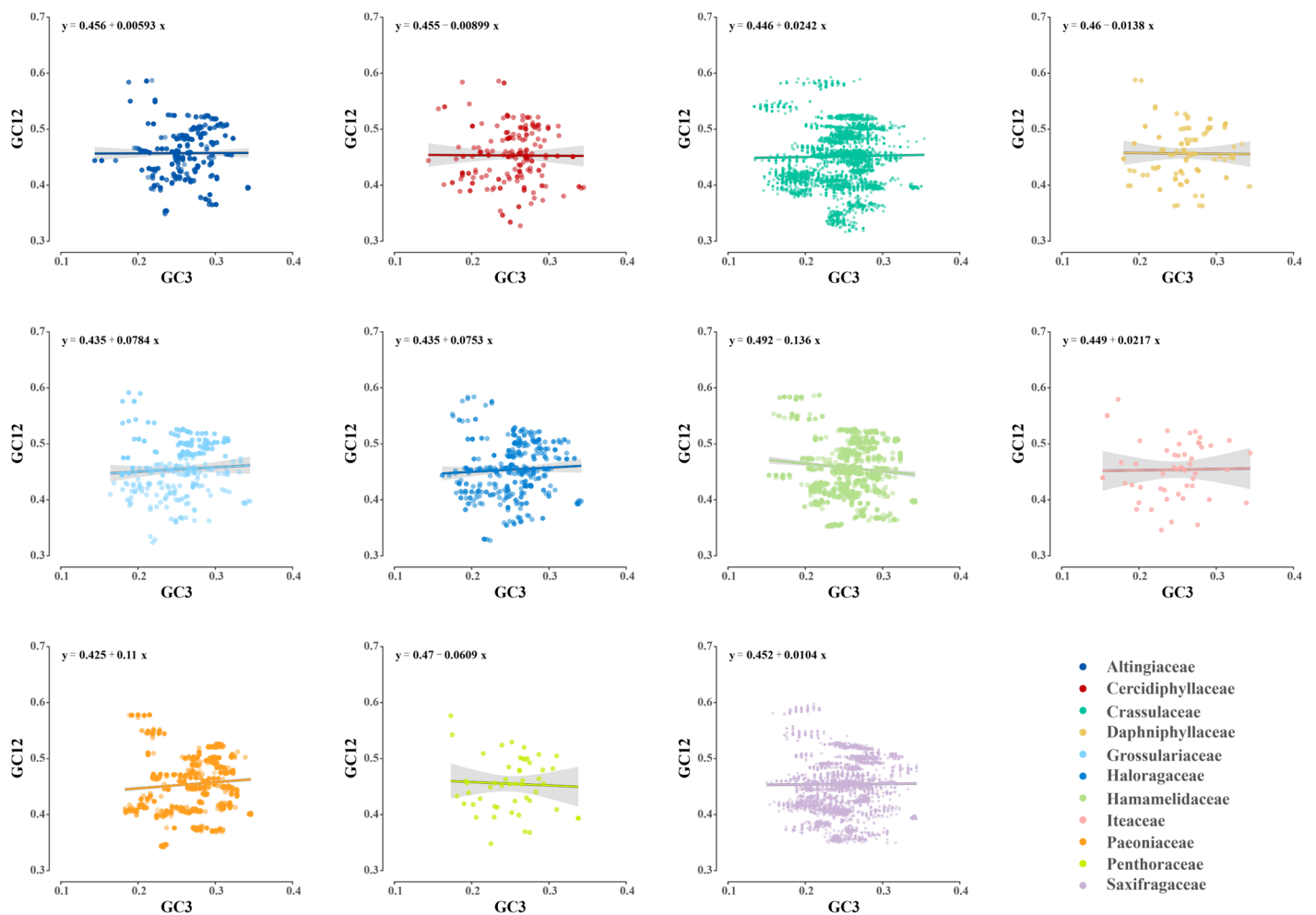


Figure 7. The neutrality plots of plastid genes in 11 Saxifragales families. The colors denote the corresponding families.

3.4. Phylogenetic Inferences of Saxifragales Based on Plastid Data

We are dedicated to uncovering the phylogenetic relationships within the Saxifragales order. Our analysis utilized a 364-taxon matrix to generate a 77,368 bp alignment for further analyses. Based on two approaches, the ML and BI inferences generally resulted in well-resolved phylogenetic cladograms that shared similar topologies (Figures 8 and S2).

Overall, our results suggested the existence of two main clades: Core Saxifragales (maximum likelihood bootstrap [BS] = 100 and Bayesian posterior probability [PP] = 1.00) and PWC (BS = 66, PP = 0.53). Two subclades of the former clade, the Crassulaceae alliance and the Saxifragaceae alliance, were both well recovered (BS = 100, PP = 1.00).

Within the Crassulaceae alliance, the Haloragaceae and Penthoraceae families are sister to the Crassulaceae, with robust supports (BS = 100, PP = 1.00). The families Saxifragaceae and Grossulariaceae (BS = 100, PP = 1.00), along with Itaceae, were observed to be nested within the Saxifragaceae alliance. In addition, it is now recognized that the PWC clade consists of five families, with Hamamelidaceae + (Cercidiphyllaceae + Daphniphyllaceae) and Altingiaceae forming the sister group to Paeoniaceae. Notably, with the rare exceptions of Cercidiphyllaceae + Daphniphyllaceae (BS = 66, PP = 0.93) and the PWC node (BS = 66, PP = 0.53), all inter-family affinities received strong node support.

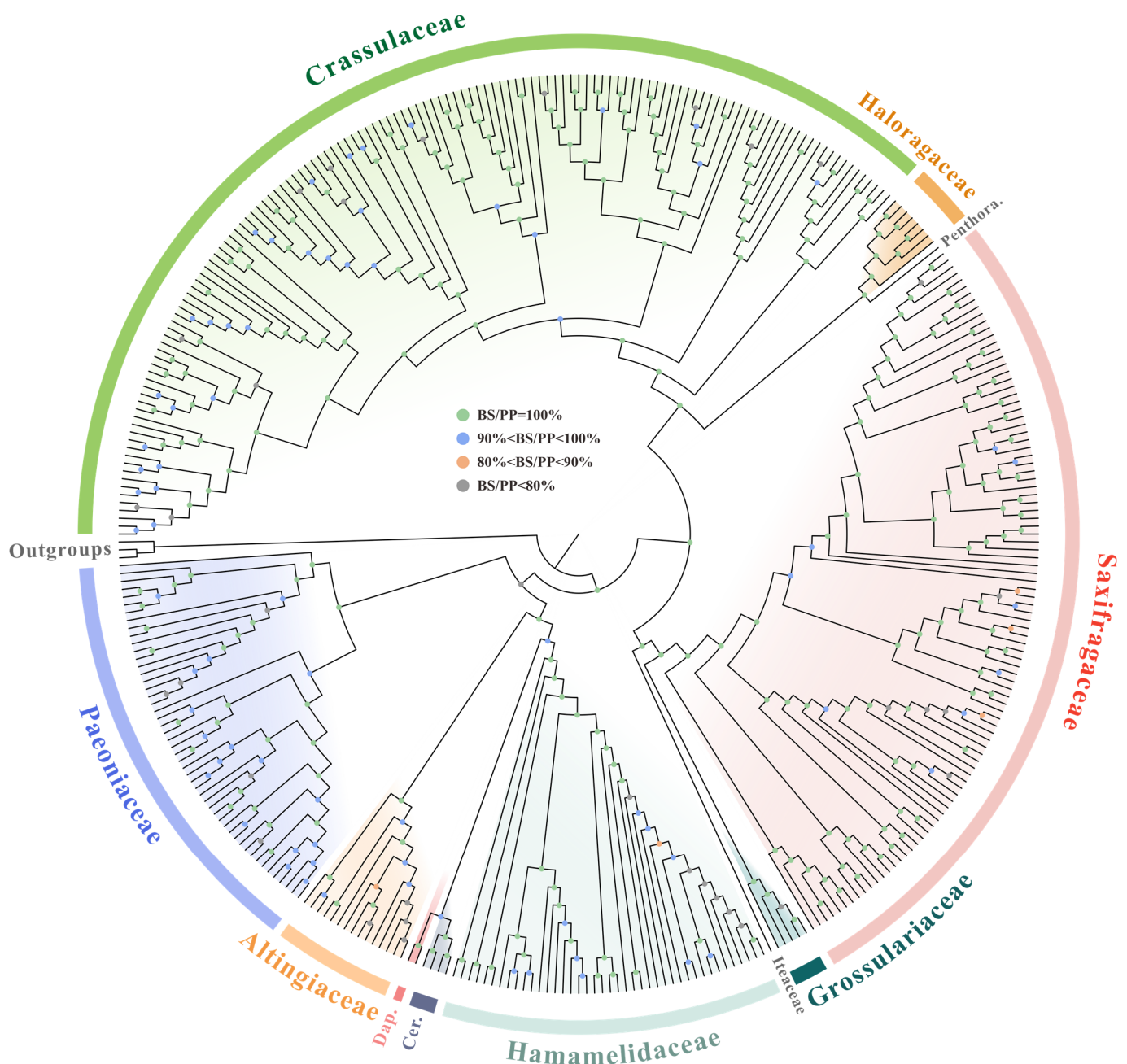


Figure 8. The simplified phylogenetic tree obtained by 79 PCGs among 361 Saxifragales plastomes, with BS and PP values shown by colored circles (the more detailed reconstruction is shown in Figure S2). Penthora. denotes Penthoraceae, Cer. denotes Cercidiphyllaceae, and Dap. denotes Daphniphyllaceae.

At a deeper phylogenetic level, we obtained high-resolution support (BS = 100, PP = 1.00) for the monophyly of all nine multiple-sample families. However, our analyses also revealed non-monophyly for some genera, including *Hylotelephium*, *Orostachys*, and *Sedum* in Crassulaceae, and *Distylium* in Hamamelidaceae.

In particular, our findings showed that *Pachyphytum compactum* was not assigned to its expected position based on the previous report, but rather clustered with *Echeveria lilacina* by weak support (BS = 55). Furthermore, at the deepest level, several individuals from the same taxon were not grouped together, but instead were nested within the clusters of other taxa, such as *Rhodiola quadrifida* (individual 2) and *Paeonia suffruticosa* (individual 5, 7, and 8).

4. Discussion

With the currently largest sampling of 361 Saxifragales plastomes, this study aimed to comprehensively evaluate the codon usage patterns of plastid genes in the Saxifrage order. Significantly, we employed many indices for analyses, including GC content at different positions of codons, RSCU, ENC, PR2 bias, ENC–GC3s, and neutrality plots. Most importantly, our findings presented here would considerably deepen our understanding of the evolutionary dynamics of plastid genes within the Saxifragales.

Interspecific codon usage heterogeneities are believed to be strongly influenced by GC content [77], which plays a vital role in the evolution of genome structure [34,78,79]. A high AT content has been widely recognized as an iconic characteristic of the plastome, which can be observed by a consistent pattern at the third codon position [44,80,81]. Similarly, all involved plastid genes in our investigation showed a tendency to utilize AT bases and AT-ending codons. It is noteworthy that a neutral effect of mutation on the third codon position leads to random usage of synonymous codons, resulting in a proportional usage of GC and AT [82,83]. Consequently, AT is strongly favored at the third bases of codons in the Saxifragales plastid genes, suggesting that selection pressure is responsible for the CUB patterns. Additionally, our findings revealed that ten genes in Saxifragales had GC1 > 0.5 and two genes had GC2 > 0.5, which is different from other studies where GC1, GC2, and GC3 values of plastid genes were all less than 0.5 [36,62,84,85]. Most notably, as proposed by Morton and So [86], stronger codon adaptation in plastid genes would lead to higher GC contents in the first and second bases of codons. Therefore, the 12 plastid genes mentioned above, particularly *rps11* with both GC1 and GC2 values over 0.5, might undergo higher codon adaptations during the evolution of the Saxifrage order. These findings will definitely provide a better understanding of the various evolutionary processes involving CUB across different plastid genes in the Saxifragales group.

The absolute synonymous codon bias can be accurately quantified by the ENC index, which measures the preferential degree of codon in a gene [19,44]. Importantly, it has been proposed that genes with higher expression tend to harbor lower ENC values and stronger CUB, and vice versa [60,87,88]. Moreover, Bulmer et al. [89] suggested that natural selection increases with the extent of gene expression. The plastid genes of Saxifragales exhibit a range of codon bias levels, both high and low. In particular, the highest bias was detected in *rps8* of 27 taxa (ENC < 35), which were mostly from Crassulaceae. Conversely, much weak bias was found in *ndhF* of Paeoniaceae with an average ENC > 55. Thereby, these findings not only highly mirrored the various expression levels of plastid genes among Saxifragales families, but also indicated that *rps8* might undergo more selection pressure in Crassulaceae. Additionally, the most substantial differences in ENC within each of the three families (Altingiaceae, Haloragaceae, and Hamamelidaceae) were found to harbor strong phylogenetic implications, which can further serve as clade-specific markers.

As another important index for CUB assessment, RSCU values could directly reflect the usage of synonymous codons. Thus far, it has been widely applied to explore codon use variations among genes. Across Saxifragales, the vast majority of the preferred codons were revealed to end with A/T bases. Notably, such bias might be congruent with its plastomic AT-rich base composition, and is also similar to the patterns of most dicot genomic genes [90–92]. In addition, the RSCU values of all NCG codons in Saxifragales families were less than 1, implying their relative rarity. Interestingly, it has been proposed that the infrequent usage of NCG codons might indicate a high level of methylation in the organism [93]. Our results demonstrated here could deepen our insights into the evolutionary patterns of plastid genes in the Saxifragales.

The silent (i.e., synonymous) sites of codons have been considered a powerful candidate to investigate evolutionary processes. To explain the preference for codon usage, several theoretical models have been proposed, comprising the neutral theory [94], the expression regulation theory [95], and the selection–mutation drift (SMD) theory [96]. The SMD model, which attributes CUB to the combined consequence of selection, mutation, and drift [96–99], has received widespread attention. To date, numerous studies have supported its applicability in elucidating codon bias [42,100–103]. Of particular interest are the

conclusions drawn for plastomic genes of angiosperms, where two different perspectives have emerged concerning the governing factors in CUB. Morton [104] argued that mutation and generic drift had a stronger influence, while another model suggests that selection plays a larger role, which has received support from multiple studies [105–107].

In this study of the Saxifragales group, the origin of the CUB in plastid genes appears to have been influenced by three factors, as suggested by the SMD model. One piece of evidence supporting this model is the presence of several splashes in the ENC–GC3s plots that are just on and near the curve, which suggests a partial role of mutation/drift [42,47]. However, the dominant role of selection is indicated by several factors, such as the obvious AT bias of the third codon position, the weak correlation between GC12 and GC3, the distant location of most splashes from the curve in the ENC plots, and the near zero slopes of neutrality plots. These findings collectively suggest that selection played a major role in shaping the CUB of Saxifragales plastid genes.

The rapid radiation in Saxifragales has resulted in highly divergent morphology, constituting a significant obstacle to disentangle its taxonomy [20]. Along with the great development of phylogenomic methodology, numerous efforts have been made to better understand its internal phylogenetic relationships [1,15–17,19,20,22]. Despite the placements of most constituent families having been well recovered to date, the relationships within the PWC clade are still a matter of debate [1].

Here, our phylogenetic inferences assigned 361 individuals into two main clades and four subclades. In general, the topology of the yielding trees was highly similar to those of the previous work [1,19,20]. Notably, most nodes received relatively high support, but some nodes had issues. For instance, we found that Paeoniaceae was a weak sister to the woody group, and a sister relationship existed between Cercidiphyllaceae and Daphniphyllaceae in the woody group, consistent with the findings of Han et al. [19] and Jian et al. [1]. However, these relationships were poorly resolved. In fact, the placement of Paeoniaceae has long been a taxonomic enigma, with inconsistent and weak support in previous analyses [1,7,15,16]. Additionally, according to the phylogeny analyses by multiple datasets of Jian et al. [1], the placement of Paeoniaceae resulted as an important inconformity among nuclear, mitochondrial, and plastid topologies of Saxifragales. Fishbein et al. [15] argued that the longest branch leading to this family could best explain such poor resolution, while Jian et al. [1] attributed the unrecovered relationships of the woody clade to an ancient, rapid radiation.

In addition, the unexpected phylogenetic positions of several taxa, as noted above, might be due to the following reasons: (1) species identification errors, (2) the current phylogenetic dataset's inability to distinguish between extremely closely related taxa, and (3) potential chloroplast capture events. Therefore, acquiring more data and increasing sampling efforts will be crucial for gaining a deeper understanding of the phylogenetic affinities within Saxifragales.

5. Conclusions

This study aimed to present a comprehensive analysis of the codon utilization of the plastid genes, focusing on a larger-sampling scale of Saxifragales. Evaluations of CUB were performed with a series of indices at the inter- and intra-family levels, involving a total of 11 families. Overall, a variety of codon bias was elucidated among the 50 plastid genes. Despite a general bias toward AT bases and AT-ending codons, the GC1 and GC2 were still found to be more than 50% in 12 genes. Furthermore, the ENC values exhibited a high level of diversity among the constituent families. Notably, significant phylogenetic implications were identified for the intrafamilial heterogeneities of ENC in Altingiaceae, Haloragaceae, and Hamamelidaceae. Moreover, PR2 plots revealed that the *petA* genes in Crassulaceae and the *rps11* genes from Paeoniaceae were more G biased and T biased, respectively, compared to the other ten families. Furthermore, six genes exhibited favored base patterns that were shared by all the examined Saxifragales lineages. More importantly, both the ENC plots and neutrality plots suggested that selection pressure predominantly governed

the CUB of Saxifragales plastid genes. Furthermore, the phylogenetic affinities inferred by both the ML and BI methods were consistent with each other, comprising two primary clades and four subclades. Altogether, our findings presented herein will convincingly deepen the current understanding of the molecular evolution of Saxifragales.

Supplementary Materials: The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/genes14030694/s1>, Figure S1: The PR2 plots of the 50 plastid genes among Saxifragales; Figure S2: The ML phylogenetic tree obtained by 79 PCGs among 361 Saxifragales plastomes; Table S1: List of the plastomes of Saxifragales involved in this study; Table S2: Nucleotide compositions of the plastid genes among Saxifragales involved in this study; Table S3: The ENC values of the plastid genes among Saxifragales; Table S4: The RSCU values of the concatenated plastid genes among Saxifragales; Table S5: The PR2 analysis of the plastid genes among Saxifragales.

Author Contributions: Conceptualization, X.K.; methodology, S.H.; software, J.Z. and S.Z.; validation, D.B. and M.Z.; formal analysis, S.H.; investigation, D.B., J.Z., and M.Z.; resources, S.Z.; data curation, S.H.; writing—original draft preparation, D.B. and S.H.; writing—review and editing, X.K.; supervision, X.K.; project administration, X.K.; funding acquisition, D.B. and X.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Basic Research Program (Natural Science Foundation) of Jiangsu Province, grant number BK20211078.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: We are grateful for computation resources from the High Performance Computing System at National Engineering Laboratory of Soil Pollution Control and Remediation Technologies, CAS Key Laboratory of Soil Environment and Pollution Remediation, Institute of Soil Science, Chinese Academy of Sciences. We kindly acknowledge the anonymous reviewers for the fruitful and critical comments.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Jian, S.; Soltis, P.S.; Gitzendanner, M.A.; Moore, M.J.; Li, R.; Hendry, T.A.; Qiu, Y.L.; Dhingra, A.; Bell, C.D.; Soltis, D.E. Resolving an ancient, rapid radiation in Saxifragales. *Syst. Biol.* **2008**, *57*, 38–57. [[CrossRef](#)] [[PubMed](#)]
- Magallon, S.; Crane, P.R.; Herendeen, P.S. Phylogenetic pattern, diversity, and diversification of eudicots. *Ann. Mo. Bot. Gard.* **1999**, *86*, 297–372. [[CrossRef](#)]
- Moore, M.J.; Soltis, P.S.; Bell, C.D.; Burleigh, J.G.; Soltis, D.E. Phylogenetic analysis of 83 plastid genes further resolves the early diversification of eudicots. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 4623–4628. [[CrossRef](#)]
- Hernández-Castillo, G.R.; Cevallos-Ferriz, S.R. Reproductive and vegetative organs with affinities to Haloragaceae from the Upper Cretaceous Huepac Chert Locality of Sonora, Mexico. *Am. J. Bot.* **1999**, *86*, 1717–1734. [[CrossRef](#)]
- Chase, M.W.; Christenhusz, M.J.; Fay, M.F.; Byng, J.; Judd, W.; Soltis, D.; Mabberley, D.; Sennikov, A.; Soltis, P. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. *Bot. J. Linn. Soc.* **2016**, *181*, 1–20. [[CrossRef](#)]
- Soltis, D.; Soltis, P.; Endress, P.; Chase, M.; Manchester, S.; Judd, W.; Majure, L.; Mavrodiev, E. *Phylogeny and Evolution of the Angiosperms*; University of Chicago Press: Chicago, IL, USA, 2018.
- Takhtadzhian, A.L.; Takhtajan, A. *Diversity and Classification of Flowering Plants*; Columbia University Press: New York, NY, USA, 1997.
- Cronquist, A.; Takhtadzhian, A.L. *An Integrated System of Classification of Flowering Plants*; Columbia University Press: New York, NY, USA, 1981.
- Morgan, D.R.; Soltis, D.E. Phylogenetic relationships among members of Saxifragaceae sensu lato based on *rbcL* sequence data. *Ann. Mo. Bot. Gard.* **1993**, *80*, 631–660. [[CrossRef](#)]
- Soltis, D.E.; Hibsich-Jetter, C.; Soltis, P.S.; Chase, M.W.; Farris, J.S. Molecular phylogenetic relationships among angiosperms: An overview based on *rbcL* and 18S rDNA sequences. In *Evolution and Diversification of Land Plants*, 1st ed.; Iwatsuki, K., Raven, P.H., Eds.; Springer Tokyo: Tokyo, Japan, 1997; pp. 157–178.
- Soltis, D.E.; Soltis, P.S. Phylogenetic relationships in Saxifragaceae sensu lato: A comparison of topologies based on 18S rDNA and *rbcL* sequences. *Am. J. Bot.* **1997**, *84*, 504–522. [[CrossRef](#)] [[PubMed](#)]

12. Soltis, D.E.; Soltis, P.S.; Nickrent, D.L.; Johnson, L.A.; Hahn, W.J.; Hoot, S.B.; Sweere, J.A.; Kuzoff, R.K.; Kron, K.A.; Chase, M.W. Angiosperm phylogeny inferred from 18S ribosomal DNA sequences. *Ann. Mo. Bot. Gard.* **1997**, *84*, 1–49. [[CrossRef](#)]
13. Hoot, S.B.; Magallon, S.; Crane, P.R. Phylogeny of basal eudicots based on three molecular data sets: *atpB*, *rbcl*, and 18S nuclear ribosomal DNA sequences. *Ann. Mo. Bot. Gard.* **1999**, *86*, 1–32. [[CrossRef](#)]
14. Soltis, D.E.; Soltis, P.S.; Chase, M.W.; Mort, M.E.; Albach, D.C.; Zanis, M.; Savolainen, V.; Hahn, W.H.; Hoot, S.B.; Fay, M.F. Angiosperm phylogeny inferred from 18S rDNA, *rbcl*, and *atpB* sequences. *Bot. J. Linn. Soc.* **2000**, *133*, 381–461. [[CrossRef](#)]
15. Fishbein, M.; Hibsich-Jetter, C.; Soltis, D.E.; Hufford, L. Phylogeny of Saxifragales (angiosperms, eudicots): Analysis of a rapid, ancient radiation. *Syst. Biol.* **2001**, *50*, 817–847. [[CrossRef](#)] [[PubMed](#)]
16. Fishbein, M.; Soltis, D.E. Further resolution of the rapid radiation of Saxifragales (angiosperms, eudicots) supported by mixed-model Bayesian analysis. *Syst. Bot.* **2004**, *29*, 883–891. [[CrossRef](#)]
17. Soltis, D.E.; Mort, M.E.; Latvis, M.; Mavrodiev, E.V.; O'Meara, B.C.; Soltis, P.S.; Burleigh, J.G.; Rubio de Casas, R. Phylogenetic relationships and character evolution analysis of Saxifragales using a supermatrix approach. *Am. J. Bot.* **2013**, *100*, 916–929. [[CrossRef](#)] [[PubMed](#)]
18. Ding, H.; Zhu, R.; Dong, J.; Bi, D.; Jiang, L.; Zeng, J.; Huang, Q.; Liu, H.; Xu, W.; Wu, L.; et al. Next-Generation Genome Sequencing of *Sedum. plumbizincicola* Sheds Light on the Structural Evolution of Plastid rRNA Operon and Phylogenetic Implications within Saxifragales. *Plants* **2019**, *8*, 386. [[CrossRef](#)]
19. Han, S.; Bi, D.; Yi, R.; Ding, H.; Wu, L.; Kan, X. Plastome evolution of *Aeonium* and *Monanthes* (Crassulaceae): Insights into the variation of plastomic tRNAs, and the patterns of codon usage and aversion. *Planta* **2022**, *256*, 35. [[CrossRef](#)]
20. Han, S.; Ding, H.; Bi, D.; Zhang, S.; Yi, R.; Gao, J.; Yang, J.; Ye, Y.; Wu, L.; Kan, X. Structural Diversities and Phylogenetic Signals in Plastomes of the Early-Divergent Angiosperms: A Case Study in Saxifragales. *Plants* **2022**, *11*, 3544. [[CrossRef](#)]
21. Dong, W.; Xu, C.; Cheng, T.; Zhou, S. Complete chloroplast genome of *Sedum. sarmentosum* and chloroplast genome evolution in Saxifragales. *PLoS ONE* **2013**, *8*, e77965. [[CrossRef](#)]
22. Dong, W.; Xu, C.; Wu, P.; Cheng, T.; Yu, J.; Zhou, S.; Hong, D.Y. Resolving the systematic positions of enigmatic taxa: Manipulating the chloroplast genome data of Saxifragales. *Mol. Phylogenet. Evol.* **2018**, *126*, 321–330. [[CrossRef](#)]
23. Tarullo, C.; Rose, J.; Sytsma, K.; Drew, B.T. Using a supermatrix approach to explore phylogenetic relationships, divergence times, and historical biogeography of Saxifragales. *Turk. J. Botany* **2021**, *45*, 5. [[CrossRef](#)]
24. Daniell, H.; Lin, C.-S.; Yu, M.; Chang, W.-J. Chloroplast genomes: Diversity, evolution, and applications in genetic engineering. *Genome Biol.* **2016**, *17*, 134. [[CrossRef](#)]
25. Liu, L.; Wang, Y.; He, P.; Li, P.; Lee, J.; Soltis, D.E.; Fu, C. Chloroplast genome analyses and genomic resource development for epilithic sister genera *Oresitrophe* and *Mukdenia* (Saxifragaceae), using genome skimming data. *BMC Genom.* **2018**, *19*, 235. [[CrossRef](#)]
26. Henriquez, C.L.; Abdullah; Ahmed, I.; Carlsen, M.M.; Zuluaga, A.; Croat, T.B.; McKain, M.R. Molecular evolution of chloroplast genomes in Monsteroideae (Araceae). *Planta* **2020**, *251*, 72. [[CrossRef](#)] [[PubMed](#)]
27. Shahzadi, I.; Abdullah; Mehmood, F.; Ali, Z.; Ahmed, I.; Mirza, B. Chloroplast genome sequences of *Artemisia maritima* and *Artemisia absinthium*: Comparative analyses, mutational hotspots in genus *Artemisia* and phylogeny in family Asteraceae. *Genomics* **2020**, *112*, 1454–1463. [[CrossRef](#)] [[PubMed](#)]
28. Mower, J.P.; Guo, W.; Partha, R.; Fan, W.; Levens, N.; Wolff, K.; Nugent, J.M.; Pabon-Mora, N.; Gonzalez, F. Plastomes from tribe Plantagineae (Plantaginaceae) reveal infrageneric structural synapomorphies and localized hypermutation for *Plantago* and functional loss of *ndh* genes from *Littorella*. *Mol. Phylogenet. Evol.* **2021**, *162*, 107217. [[CrossRef](#)] [[PubMed](#)]
29. Ren, T.; Yang, Y.; Zhou, T.; Liu, Z.-L. Comparative plastid genomes of *Primula* species: Sequence divergence and phylogenetic relationships. *Int. J. Mol. Sci.* **2018**, *19*, 1050. [[CrossRef](#)]
30. Hong, Z.; Wu, Z.; Zhao, K.; Yang, Z.; Zhang, N.; Guo, J.; Tembrock, L.R.; Xu, D. Comparative analyses of five complete chloroplast genomes from the genus *Pterocarpus* (Fabaceae). *Int. J. Mol. Sci.* **2020**, *21*, 3758. [[CrossRef](#)] [[PubMed](#)]
31. Li, P.; Lou, G.; Cai, X.; Zhang, B.; Cheng, Y.; Wang, H. Comparison of the complete plastomes and the phylogenetic analysis of *Paulownia* species. *Sci. Rep.* **2020**, *10*, 2225. [[CrossRef](#)] [[PubMed](#)]
32. Pacheco, T.G.; Lopes, A.d.S.; Welter, J.F.; Yotoko, K.S.C.; Otoni, W.C.; Vieira, L.d.N.; Guerra, M.P.; Nodari, R.O.; Balsanelli, E.; Pedrosa, F.d.O. Plastome sequences of the subgenus *Passiflora* reveal highly divergent genes and specific evolutionary features. *Plant Mol. Biol.* **2020**, *104*, 21–37. [[CrossRef](#)]
33. Chen, S.; Wang, T.; Shu, J.; Xiang, Q.; Yang, T.; Zhang, X.; Yan, Y. Plastid Phylogenomics and Plastomic Diversity of the Extant Lycophytes. *Genes* **2022**, *13*, 1280. [[CrossRef](#)]
34. Sheng, J.; She, X.; Liu, X.; Wang, J.; Hu, Z. Comparative analysis of codon usage patterns in chloroplast genomes of five *Miscanthus* species and related species. *PeerJ* **2021**, *9*, e12173. [[CrossRef](#)]
35. Suzuki, H.; Morton, B.R. Codon Adaptation of Plastid Genes. *PLoS ONE* **2016**, *11*, e0154306. [[CrossRef](#)]
36. Zhang, Y.; Nie, X.; Jia, X.; Zhao, C.; Biradar, S.S.; Wang, L.; Du, X.; Weining, S. Analysis of codon usage patterns of the chloroplast genomes in the Poaceae family. *Aust. J. Bot.* **2012**, *60*, 461–470. [[CrossRef](#)]
37. Ikemura, T. Correlation between the abundance of *Escherichia coli* transfer RNAs and the occurrence of the respective codons in its protein genes: A proposal for a synonymous codon choice that is optimal for the *E. coli* translational system. *J. Mol. Biol.* **1981**, *151*, 389–409. [[CrossRef](#)]

38. Wang, L.; Roossinck, M.J. Comparative analysis of expressed sequences reveals a conserved pattern of optimal codon usage in plants. *Plant. Mol. Biol.* **2006**, *61*, 699–710. [[CrossRef](#)] [[PubMed](#)]
39. Vicario, S.; Mason, C.E.; White, K.P.; Powell, J.R. Developmental stage and level of codon usage bias in *Drosophila*. *Mol. Biol. Evol.* **2008**, *25*, 2269–2277. [[CrossRef](#)] [[PubMed](#)]
40. Machado, H.E.; Lawrie, D.S.; Petrov, D.A. Pervasive strong selection at the level of codon usage bias in *Drosophila melanogaster*. *Genetics* **2020**, *214*, 511–528. [[CrossRef](#)]
41. Iriarte, A.; Lamolle, G.; Musto, H. Codon Usage Bias: An Endless Tale. *J. Mol. Evol.* **2021**, *89*, 589–593. [[CrossRef](#)]
42. Yang, J.; Ding, H.; Kan, X. Codon usage patterns and evolution of HSP60 in birds. *Int. J. Biol. Macromol.* **2021**, *183*, 1002–1012. [[CrossRef](#)]
43. Ding, H.; Bi, D.; Han, S.; Yi, R.; Zhang, S.; Ye, Y.; Gao, J.; Yang, J.; Kan, X. Mitogenomic Codon Usage Patterns of Superfamily Certhioidea (Aves, Passeriformes): Insights into Asymmetrical Bias and Phylogenetic Implications. *Animals* **2022**, *13*, 96. [[CrossRef](#)]
44. Parvathy, S.T.; Udayasuriyan, V.; Bhadana, V. Codon usage bias. *Mol. Biol. Rep.* **2022**, *49*, 539–565. [[CrossRef](#)]
45. Salim, H.M.; Cavalcanti, A.R. Factors influencing codon usage bias in genomes. *J. Braz. Chem. Soc.* **2008**, *19*, 257–262. [[CrossRef](#)]
46. Grantham, R.; Gautier, C.; Gouy, M.; Mercier, R.; Pavé, A. Codon catalog usage and the genome hypothesis. *Nucleic. Acids. Res.* **1980**, *8*, 197. [[CrossRef](#)]
47. Wright, F. The ‘effective number of codons’ used in a gene. *Gene* **1990**, *87*, 23–29. [[CrossRef](#)]
48. Sharp, P.M.; Li, W.-H. An evolutionary perspective on synonymous codon usage in unicellular organisms. *J. Mol. Evol.* **1986**, *24*, 28–38. [[CrossRef](#)]
49. Sueoka, N. Intrastrand parity rules of DNA base composition and usage biases of synonymous codons. *J. Mol. Evol.* **1995**, *40*, 318–325. [[CrossRef](#)]
50. Sueoka, N. Directional mutation pressure and neutral molecular evolution. *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 2653–2657. [[CrossRef](#)]
51. Chen, Z.; Yu, X.; Yang, Y.; Wei, P.; Zhang, W.; Li, X.; Liu, C.; Zhao, S.; Li, X.; Liu, X. Comparative Analysis of Chloroplast Genomes within *Saxifraga* (Saxifragaceae) Takes Insights into Their Genomic Evolution and Adaption to the High-Elevation Environment. *Genes* **2022**, *13*, 1673. [[CrossRef](#)] [[PubMed](#)]
52. Ding, H.; Han, S.; Ye, Y.; Bi, D.; Zhang, S.; Yi, R.; Gao, J.; Yang, J.; Wu, L.; Kan, X. Ten Plastomes of *Crassula* (Crassulaceae) and Phylogenetic Implications. *Biology* **2022**, *11*, 1779. [[CrossRef](#)] [[PubMed](#)]
53. Tillich, M.; Lehwick, P.; Pellizzer, T.; Ulbricht-Jones, E.S.; Fischer, A.; Bock, R.; Greiner, S. GeSeq—versatile and accurate annotation of organelle genomes. *Nucleic Acids. Res.* **2017**, *45*, W6–W11. [[CrossRef](#)]
54. Shi, L.; Chen, H.; Jiang, M.; Wang, L.; Wu, X.; Huang, L.; Liu, C. CPGAVAS2, an integrated plastome sequence annotator and analyzer. *Nucleic Acids Res.* **2019**, *47*, W65–W73. [[CrossRef](#)]
55. Johnson, M.; Zaretskaya, I.; Raytselis, Y.; Merezuk, Y.; McGinnis, S.; Madden, T.L. NCBI BLAST: A better web interface. *Nucleic. Acids. Res.* **2008**, *36*, W5–W9. [[CrossRef](#)]
56. Katoh, K.; Standley, D.M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **2013**, *30*, 772–780. [[CrossRef](#)]
57. Xia, X. DAMBE7: New and improved tools for data analysis in molecular biology and evolution. *Mol. Biol. Evol.* **2018**, *35*, 1550–1552. [[CrossRef](#)] [[PubMed](#)]
58. McInerney, J.O. Replicational and transcriptional selection on codon usage in *Borrelia burgdorferi*. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 10698–10703. [[CrossRef](#)]
59. Roth, A.; Anisimova, M.; Cannarozzi, G.M. Measuring codon usage bias. In *Codon Evolution: Mechanisms and Models*, 1st ed.; Cannarozzi, G.M., Schneider, A., Eds.; Oxford University Press: Oxford, UK, 2012; pp. 189–217.
60. Wang, L.; Xing, H.; Yuan, Y.; Wang, X.; Saeed, M.; Tao, J.; Feng, W.; Zhang, G.; Song, X.; Sun, X. Genome-wide analysis of codon usage bias in four sequenced cotton species. *PLoS ONE* **2018**, *13*, e0194372. [[CrossRef](#)]
61. Peden, J.F. Analysis of Codon Usage. Ph.D. Thesis, University of Nottingham, Nottingham, UK, 2000.
62. Li, G.; Zhang, L.; Xue, P. Codon usage pattern and genetic diversity in chloroplast genomes of *Panicum* species. *Gene* **2021**, *802*, 145866. [[CrossRef](#)] [[PubMed](#)]
63. Comeron, J.M.; Aguadé, M. An evaluation of measures of synonymous codon usage bias. *J. Mol. Evol.* **1998**, *47*, 268–274. [[CrossRef](#)] [[PubMed](#)]
64. He, Z.; Gan, H.; Liang, X. Analysis of synonymous codon usage bias in potato virus M and its adaption to hosts. *Viruses* **2019**, *11*, 752. [[CrossRef](#)] [[PubMed](#)]
65. Wickham, H. Ggplot2. *Wiley Interdiscip. Rev. Comput. Stat.* **2011**, *3*, 180–185. [[CrossRef](#)]
66. Galtier, N.; Lobry, J.R. Relationships between genomic G+C content, RNA secondary structures, and optimal growth temperature in prokaryotes. *J. Mol. Evol.* **1997**, *44*, 632–636. [[CrossRef](#)]
67. Sueoka, N. 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* **1999**, *238*, 53–58. [[CrossRef](#)] [[PubMed](#)]
68. Sueoka, N. Two aspects of DNA base composition: G+C content and translation-coupled deviation from intra-strand rule of A=T and G=C. *J. Mol. Evol.* **1999**, *49*, 49–62. [[CrossRef](#)]

69. Yu, X.; Liu, J.; Li, H.; Liu, B.; Zhao, B.; Ning, Z. Comprehensive analysis of synonymous codon usage patterns and influencing factors of porcine epidemic diarrhea virus. *Arch. Virol.* **2021**, *166*, 157–165. [[CrossRef](#)] [[PubMed](#)]
70. Xiang, H.; Zhang, R.; Butler III, R.R.; Liu, T.; Zhang, L.; Pombert, J.-F.; Zhou, Z. Comparative analysis of codon usage bias patterns in microsporidian genomes. *PLoS ONE* **2015**, *10*, e0129223. [[CrossRef](#)] [[PubMed](#)]
71. Li, G.; Pan, Z.; Gao, S.; He, Y.; Xia, Q.; Jin, Y.; Yao, H. Analysis of synonymous codon usage of chloroplast genome in *Porphyra umbilicalis*. *Genes Genom.* **2019**, *41*, 1173–1181. [[CrossRef](#)] [[PubMed](#)]
72. Indrabalan, U.B.; Suresh, K.P.; Shivamallu, C.; Patil, S.S. An extensive evaluation of codon usage pattern and bias of structural proteins p30, p54 and, p72 of the African swine fever virus (ASFV). *Virusdisease* **2021**, *32*, 810–822. [[CrossRef](#)] [[PubMed](#)]
73. Zhang, D.; Gao, F.; Jakovlić, I.; Zou, H.; Zhang, J.; Li, W.X.; Wang, G.T. PhyloSuite: An integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. *Mol. Ecol. Resour.* **2020**, *20*, 348–355. [[CrossRef](#)]
74. Stamatakis, A. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **2014**, *30*, 1312–1313. [[CrossRef](#)]
75. Darriba, D.; Posada, D.; Kozlov, A.M.; Stamatakis, A.; Morel, B.; Flouri, T. ModelTest-NG: A new and scalable tool for the selection of DNA and protein evolutionary models. *Mol. Biol. Evol.* **2020**, *37*, 291–294. [[CrossRef](#)]
76. Rambaut, A.; Drummond, A.J.; Xie, D.; Baele, G.; Suchard, M.A. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Syst. Biol.* **2018**, *67*, 901–904. [[CrossRef](#)]
77. Plotkin, J.B.; Kudla, G. Synonymous but not the same: The causes and consequences of codon bias. *Nat. Rev. Genet.* **2011**, *12*, 32–42. [[CrossRef](#)]
78. Liu, Q.; Xue, Q. Comparative studies on codon usage pattern of chloroplasts and their host nuclear genes in four plant species. *J. Genet.* **2005**, *84*, 55–62. [[CrossRef](#)] [[PubMed](#)]
79. Bellgard, M.; Schibeci, D.; Trifonov, E.; Gojobori, T. Early detection of G + C differences in bacterial species inferred from the comparative analysis of the two completely sequenced *Helicobacter pylori* strains. *J. Mol. Evol.* **2001**, *53*, 465–468. [[CrossRef](#)] [[PubMed](#)]
80. Clegg, M.T.; Gaut, B.S.; Learn, G.H., Jr.; Morton, B.R. Rates and patterns of chloroplast DNA evolution. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 6795–6801. [[CrossRef](#)] [[PubMed](#)]
81. Howe, C.J.; Barbrook, A.C.; Koumandou, V.L.; Nisbet, R.E.R.; Symington, H.A.; Wightman, T.F. Evolution of the chloroplast genome. *Philos. Trans. R Soc. Lond. B Biol. Sci.* **2003**, *358*, 99–107. [[CrossRef](#)] [[PubMed](#)]
82. Zhang, W.J.; Zhou, J.; Li, Z.F.; Wang, L.; Gu, X.; Zhong, Y. Comparative analysis of codon usage patterns among mitochondrion, chloroplast and nuclear genes in *Triticum aestivum* L. *J. Integr. Plant. Biol.* **2007**, *49*, 246–254. [[CrossRef](#)]
83. Yengkhom, S.; Uddin, A.; Chakraborty, S. Deciphering codon usage patterns and evolutionary forces in chloroplast genes of *Camellia sinensis* var *assamica* and *Camellia sinensis* var *sinensis* in comparison to *Camellia pubicosta*. *J. Integr. Agric.* **2019**, *18*, 2771–2785. [[CrossRef](#)]
84. Wang, Z.; Xu, B.; Li, B.; Zhou, Q.; Wang, G.; Jiang, X.; Wang, C.; Xu, Z. Comparative analysis of codon usage patterns in chloroplast genomes of six Euphorbiaceae species. *PeerJ* **2020**, *8*, e8251. [[CrossRef](#)]
85. Wang, Z.; Cai, Q.; Wang, Y.; Li, M.; Wang, C.; Wang, Z.; Jiao, C.; Xu, C.; Wang, H.; Zhang, Z. Comparative analysis of codon Bias in the chloroplast genomes of Theaceae species. *Front. Genet.* **2022**, *13*, 824610. [[CrossRef](#)]
86. Morton, B.R.; So, B.G. Codon usage in plastid genes is correlated with context, position within the gene, and amino acid content. *J. Mol. Evol.* **2000**, *50*, 184–193. [[CrossRef](#)]
87. Kandeel, M.; Ibrahim, A.; Fayed, M.; Al-Nazawi, M. From SARS and MERS CoVs to SARS-CoV-2: Moving toward more biased codon usage in viral structural and nonstructural genes. *J. Med. Virol.* **2020**, *92*, 660–666. [[CrossRef](#)]
88. Nair, R.R.; Raveendran, N.T.; Dirisala, V.R.; Nandhini, M.B.; Sethuraman, T.; Venkateswarulu, T.C.; Doss, G. Mutational pressure drives evolution of synonymous codon usage in genetically distinct *Oenothera* plastomes. *Iran. J. Biotechnol.* **2014**, *12*, e1156. [[CrossRef](#)]
89. Bulmer, M. Are codon usage patterns in unicellular organisms determined by selection-mutation balance? *J. Evol. Biol.* **1988**, *1*, 15–26. [[CrossRef](#)]
90. Zhang, R.; Zhang, L.; Wang, W.; Zhang, Z.; Du, H.; Qu, Z.; Li, X.-Q.; Xiang, H. Differences in codon usage bias between photosynthesis-related genes and genetic system-related genes of chloroplast genomes in cultivated and wild solanum species. *Int. J. Mol. Sci.* **2018**, *19*, 3142. [[CrossRef](#)]
91. Kawabe, A.; Miyashita, N.T. Patterns of codon usage bias in three dicot and four monocot plant species. *Genes Genet. Syst.* **2003**, *78*, 343–352. [[CrossRef](#)]
92. Campbell, W.H.; Gowri, G. Codon usage in higher plants, green algae, and cyanobacteria. *Plant. Physiol.* **1990**, *92*, 1–11. [[CrossRef](#)]
93. Shen, Z.; Gan, Z.; Zhang, F.; Yi, X.; Zhang, J.; Wan, X. Analysis of codon usage patterns in citrus based on coding sequence data. *BMC Genom.* **2020**, *21*, 234. [[CrossRef](#)] [[PubMed](#)]
94. Nakamura, Y.; Gojobori, T.; Ikemura, T. Codon usage tabulated from international DNA sequence databases: Status for the year 2000. *Nucleic Acids Res.* **2000**, *28*, 292. [[CrossRef](#)] [[PubMed](#)]
95. Gouy, M.; Gautier, C. Codon usage in bacteria: Correlation with gene expressivity. *Nucleic Acids Res.* **1982**, *10*, 7055–7074. [[CrossRef](#)] [[PubMed](#)]
96. Bulmer, M. The selection-mutation-drift theory of synonymous codon usage. *Genetics* **1991**, *129*, 897–907. [[CrossRef](#)] [[PubMed](#)]

97. Rao, Y.; Wu, G.; Wang, Z.; Chai, X.; Nie, Q.; Zhang, X. Mutation bias is the driving force of codon usage in the *Gallus gallus* genome. *DNA Res.* **2011**, *18*, 499–512. [[CrossRef](#)]
98. Guan, D.-L.; Ma, L.-B.; Khan, M.S.; Zhang, X.-X.; Xu, S.-Q.; Xie, J.-Y. Analysis of codon usage patterns in *Hirudinaria. manillensis* reveals a preference for GC-ending codons caused by dominant selection constraints. *BMC Genom.* **2018**, *19*, 542. [[CrossRef](#)] [[PubMed](#)]
99. Ray, S.K.; Baruah, V.J.; Satapathy, S.S.; Banerjee, R. Cotranslational protein folding reveals the selective use of synonymous codons along the coding sequence of a low expression gene. *J. Genet.* **2014**, *93*, 613–617. [[CrossRef](#)]
100. Somaratne, Y.; Guan, D.-L.; Wang, W.-Q.; Zhao, L.; Xu, S.-Q. The complete chloroplast genomes of two *Lespedeza* species: Insights into codon usage bias, RNA editing sites, and phylogenetic relationships in Desmodieae (Fabaceae: Papilionoideae). *Plants* **2019**, *9*, 51. [[CrossRef](#)] [[PubMed](#)]
101. Tang, D.; Wei, F.; Cai, Z.; Wei, Y.; Khan, A.; Miao, J.; Wei, K. Analysis of codon usage bias and evolution in the chloroplast genome of *Mesona. chinensis* Benth. *Dev. Genes Evol.* **2021**, *231*, 1–9. [[CrossRef](#)]
102. Wang, Y.; Jiang, D.; Guo, K.; Zhao, L.; Meng, F.; Xiao, J.; Niu, Y.; Sun, Y. Comparative analysis of codon usage patterns in chloroplast genomes of ten *Epimedium* species. *BMC Genom. Data* **2023**, *24*, 3. [[CrossRef](#)]
103. Wang, Y.; Yao, L.; Fan, J.; Zhao, X.; Zhang, Q.; Chen, Y.; Guo, C. The Codon Usage Bias Analysis of Free-Living Ciliates' Macronuclear Genomes and Clustered Regularly Interspaced Short Palindromic Repeats/Cas9 Vector Construction of *Stylonychia lemnae*. *Front. Microbiol.* **2022**, *13*, 785889. [[CrossRef](#)]
104. Morton, B.R. Context-dependent mutation dynamics, not selection, explains the codon usage bias of most angiosperm chloroplast genes. *J. Mol. Evol.* **2022**, *90*, 17–29. [[CrossRef](#)] [[PubMed](#)]
105. Duan, H.; Zhang, Q.; Wang, C.; Li, F.; Tian, F.; Lu, Y.; Hu, Y.; Yang, H.; Cui, G. 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* **2021**, *9*, e10787. [[CrossRef](#)]
106. Liu, H.; Lu, Y.; Lan, B.; Xu, J. Codon usage by chloroplast gene is bias in *Hemiptelea davidii*. *J. Genet.* **2020**, *99*, 8. [[CrossRef](#)] [[PubMed](#)]
107. Xu, C.; Cai, X.; Chen, Q.; Zhou, H.; Cai, Y.; Ben, A. Factors affecting synonymous codon usage bias in chloroplast genome of oncidium gower ramsey. *Evol. Bioinform.* **2011**, *7*, 271–278. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.