

Research Article

Codon Usage Bias in Two Hemipteran Insect Species: *Bemisia tabaci* and *Homalodisca coagulata*

Jyotika Sharma, Supriyo Chakraborty, and Arif Uddin

Department of Biotechnology, Assam University, Silchar, Assam 788011, India

Correspondence should be addressed to Supriyo Chakraborty; supriyoch_2008@rediffmail.com

Received 23 May 2014; Accepted 9 September 2014; Published 13 October 2014

Academic Editor: Volodymyr Dvornyk

Copyright © 2014 Jyotika Sharma et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Codon bias is the nonuniform use of synonymous codons which encode the same amino acid. Some codons are more frequently used than others in several organisms, particularly in the highly expressed genes. The spectacular diversity of insects makes them a suitable candidate for analyzing the codon usage bias. Recent expansion in genome sequencing of different insect species provides an opportunity for studying the codon usage bias. Several works on patterns of codon usage bias were done on *Drosophila* and other related species but only few works were found in Hemiptera order. We analyzed codon usage in two Hemipteran insect species namely *Bemisia tabaci* and *Homalodisca coagulata*. Most frequent codons end with A or C at the 3rd codon position. The ENC (a measure of codon bias) value ranges from 43 to 60 (52.80) in *B. tabaci* but from 49 to 60 (56.69) in *H. coagulata*. In both insect species, a significant positive correlation was observed between A and A3%, C and C3%, and GC and GC3%, respectively. Our findings suggest that codon usage bias in two Hemipteran insect species is not remarkable and that mutation pressure causes the codon usage pattern in two Hemipteran insect species.

1. Introduction

Codon usage bias is the phenomenon of unequal usage of codons during transcription and translation of genes to proteins. Some codons are used more often than alternate synonymous codons in genes where such bias exists. Such codons are often referred to as optimized codons or preferred codons. The extent of codon bias is mainly caused by mutation and translational selection [1–4].

Repeatedly used codons are often termed optimal or major codons, whereas less repeatedly used ones are termed nonoptimal or minor codons. Less repeatedly used codons usually correspond to less abundant tRNAs in the cell than optimal codons do [5–9] and the translational machinery is more likely to pause there [10]. During elongation stage of translation, these processing errors can occur in the cell and are termed as premature termination [10]. Therefore, the intensity of purifying selection is expected to increase as peptide elongation proceeds [9, 11] and so the codon usage bias in a gene should also lead to a steady increase [11]. In addition, optimal codons may be selected for translational

effectiveness and so are more often used in highly articulated genes [2, 9, 12–16]. Therefore, codon usage bias may be accounted for by negative selection against nonoptimal codons and positive selection for optimal codons.

Analysis of codon usage bias is a well-established technique for understanding the protein coding sequences of genomes. As usage of synonymous codons during translation is nonuniform, identifying the patterns of codon usage bias is important towards understanding the mode of translational selection of protein coding genes among related species. The study of codon usage bias is gaining momentum with the initiation of whole genome sequencing of numerous organisms [17]. Molecular evolutionary investigations suggest that codon usage bias varies both within and between genomes and may have a major significance in understanding genome evolution among related species at molecular level [18]. The investigation of causes and consequences of the pattern of codon usage bias and the detection of selective forces that shape their evolution are of practical importance to studies of genome biology.

The analyses on codon usage patterns in insects are of great interest for several reasons. The exceptional diversity of insects within the animal kingdom makes insects suitable for studying the codon bias patterns at different evolutionary time scales. Moreover, the recent spurt in genome sequencing of different insect species complexes also provides an exceptional opportunity for studying the evolution of codon bias in coding sequences within specific families and genera. Although studies have been taken up on codon usage bias within the genus *Drosophila* or between a few other insect species [19–21], a comprehensive analysis of codon bias patterns between various Hemipteran insects sequenced genomes is lacking. The insect species analyzed in this study assume importance because they are relevant as vectors to the transmission of various plant diseases and these insect species are currently considered as important agricultural pests for many crop plants. The silver-leaf whitefly, *Bemisia tabaci*, which is also referred to as the sweet-potato whitefly is one of numerous whiteflies that had a major impact on tobacco crops. It has been transmitting geminiviruses such as lettuce infectious yellows virus, tomato yellow leaf curl virus, and African cassava mosaic virus for years and over many continents [22]. The glassy-winged sharpshooter, *Homalodisca vitripennis* (Hemiptera: Cicadellidae), which sucks xylem fluid is native to the southwestern United States. The discovery of the role of this insect in transmission of pathogenic bacterium *Xylella fastidiosa* which causes Pierce's disease of grapevines [23–25] has acquired importance in scientific study.

The codon usage bias is most prominent in highly expressed genes. Difference of codon optimization between genes provides differential efficiency as well as accuracy in the translation of genes [26–28]. The selection associated with translational efficiency/accuracy is often termed as "translation selection." The study of codon bias is gaining renewed attention of scientists across the globe with the advent of whole genome sequencing of numerous organisms [17]. Various factors such as gene expression level, gene length, composition bias (%GC content and GC skew), recombination rates, and RNA stability are known to influence codon bias in organisms [4, 29–31].

In this study, one of the major objectives is to understand the patterns of codon usage bias among the two Hemipteran insect species: *Bemisia tabaci* and *Homalodisca coagulata*. The Hemiptera order includes the majority of insect genomes that have been sequenced so far. Although studies have been conducted on codon usage bias within the genus *Drosophila* or between a few other insect species [19–21] comprehensive analysis of codon usage bias patterns between the sequenced genomes of these two species is lacking.

The goal of this study is to perform a comparative analysis of codon usage bias pattern among the sequenced genes of these two insect species belonging to the same order (Hemiptera). We have analyzed a few selected genes, which are available in NCBI, of the two insect species *Bemisia tabaci* and *Homalodisca coagulata* in this investigation. Our results provide useful insights on the patterns of codon usage bias that facilitate better understanding of the structure and evolution of gene coding sequences of these insects.

2. Materials and Methods

2.1. Sequence Data. The complete coding sequences of genes for different insect species were retrieved from the National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/Genbank/>). To minimize the sampling errors [32], we have analyzed only those CDS sequences that are 100 codons or more in length and have correct initial and termination codons.

2.2. Compositional Properties. General nucleotide composition (A, C, T, and G%) and nucleotide composition at the third position of each codon (A3, C3, T3, and G3%) were analyzed for insect CDSs using an in-house Perl script developed by SC. The GC and GC3 indices referred to the overall GC content in the gene sequence and at the third position of synonymous codon (excluding met, trp, and termination codons), respectively.

2.3. Measures of Synonymous Codon Usage Bias. A large number of indices for measuring codon usage bias have been proposed; some of the most relevant and widely used measures analyzed in this study are discussed below.

2.4. Relative Synonymous Codon Usage (RSCU). Relative synonymous codon usage (RSCU) is calculated as the ratio of the observed frequency of a codon to the frequency expected if all synonymous codons of a particular amino acid are used equally. RSCU value greater than 1.0 indicates that the corresponding codons are used more frequently than the expected frequency whereas the reverse is true for RSCU values less than 1.0 [12]. Consider

$$\text{RSCU}_{ij} = \frac{X_{ij}}{(1/n_i) \sum_{j=1}^{n_i} X_{ij}}, \quad (1)$$

where X_{ij} is the frequency of occurrence of the j th codon for i th amino acid (any X_{ij} with a value of zero is arbitrarily assigned as a value of 0.5) and n_i is the number of codons for the i th amino acid (i th codon family).

2.5. The Effective Number of Codons (ENC). The effective number of codons used by a gene (ENC) is generally used to measure the bias of synonymous codons [32]. The values of ENC range from 20 (when only one codon is used for each amino acid) to 61 (when codons are used randomly). If the calculated ENC is greater than 61 (because codon usage is more evenly distributed than expected), it is adjusted to 61. Consider

$$\text{ENC} = 2 + \frac{9}{F_2} + \frac{1}{F_3} + \frac{5}{F_4} + \frac{3}{F_6}. \quad (2)$$

2.6. Codon Adaptation Index (CAI). The codon adaptation index (CAI) [33] is a very widely used measure of codon bias in prokaryotes [34–36] as well as eukaryotes [37–39]. CAI is a measure of the relative adaptiveness of the codon usage of a gene towards the codon usage of highly expressed

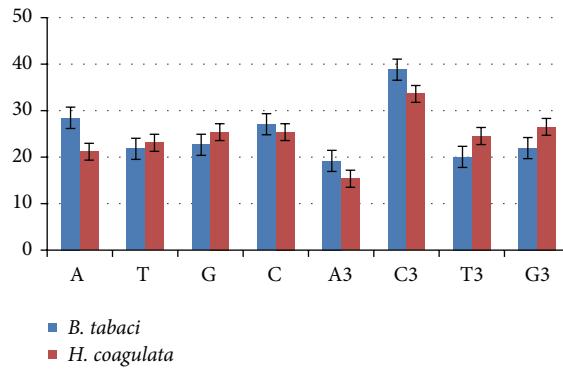


FIGURE 1: Nucleotide composition (overall and at 3rd codon position) in *B. tabaci* and *H. coagulata*.

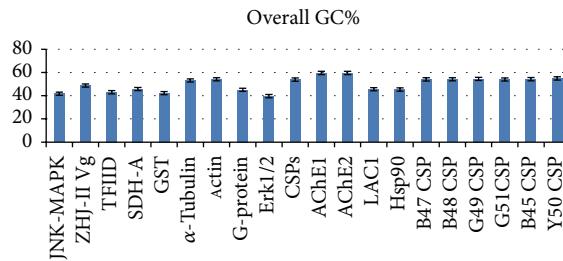


FIGURE 2: Distribution of overall GC% content in different genes of *Bemisia tabaci*.

genes. The relative adaptiveness (ω) of each codon is the ratio of the usage of each codon, to that of the most abundant codon within the same synonymous family. Nonsynonymous codons and termination codons (dependent on genetic code) are excluded from this analysis. CAI values range from 0 to 1, with higher values indicating a higher proportion of the most abundant codons [33]. The CAI is calculated as

$$\text{CAI} = \exp \left(\frac{1}{L} \sum_{k=1}^L \ln \omega_k \right), \quad (3)$$

where ω_k is the relative adaptiveness of the k th codon and L is the number of synonymous codons in the gene.

The analysis for codon usage bias was done by using a Perl script developed by SC.

2.7. Statistical Analysis. Correlation analysis was used to identify the relationship between overall nucleotide composition and each base at 3rd codon position. All the statistical analyses were done using the SPSS software.

3. Results

3.1. Compositional Properties. The genes of *Bemisia tabaci* and *Homalodisca coagulata* with their accession number, gene length, ENC, CAI, and overall GC(%), GC1(%), GC2(%), and GC3(%) are shown in Tables 1 and 2 (see Supplementary Material available online at <http://dx.doi.org/10.1155/2014/145465>). The overall nucleotide composition and nucleotide

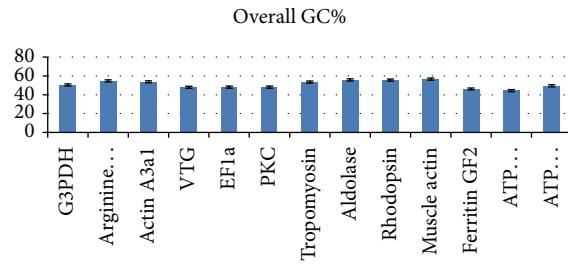


FIGURE 3: Distribution of overall GC content in different genes of *Homalodisca coagulata*.

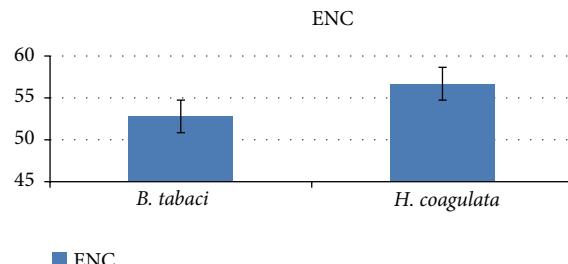


FIGURE 4: Comparison of ENC in both species *Bemisia tabaci* and *Homalodisca coagulata*.

composition at the third codon position of *Bemisia tabaci* and *Homalodisca coagulata* coding sequences are provided in Table 3 (Supplementary Material). In both species bases A and C occurred more frequently than T and G as shown in Figure 1. The nucleotide C occurred most frequently at the third codon position (average C3% = 38.79 ± 13.97 ; 33.60 ± 8.48) and A occurred least frequently (average A3% = 19.19 ± 7.69 ; 15.35 ± 5.05) as shown in Figure 1 in *B. tabaci* and *H. coagulata*, respectively. The overall nucleotide composition and the composition at the third codon position in *B. tabaci* and *H. coagulata* genes suggest that compositional constraint might be influencing the codon usage pattern of these species. The average GC and GC3% in *B. tabaci* are 50.14 ± 6.11 and 60.73 ± 17.72 , respectively. Figure 2 shows the distribution of the overall GC contents in *B. tabaci* genes. In *H. coagulata*, the average overall GC and GC3% are 51.07 and 60.10, respectively. Figure 3 shows the distribution of the overall GC contents in *H. coagulata* genes.

The ENC values range from 43 to 60 with a mean value of 52.80 in *B. tabaci* but from 49 to 60 with a mean value of 56.69 in *H. coagulata*, respectively (Figure 4). Higher ENC value means low bias. This result indicates that codon usage bias is not very remarkable in these species and is apparently maintained at a stable level. CAI is used to measure the level of gene expression. Although gene expression is more in *H. coagulata* than *B. tabaci*, the values are not statistically significant ($P > 0.05$) as shown in Figure 5.

3.2. Codon Usage in Insect Species. The overall RSCU values for the 59 codons in *Bemisia tabaci* and *Homalodisca coagulata* indicated that C occurred most frequently at the third codon position (as shown in Tables 5 and 6 in supplementary

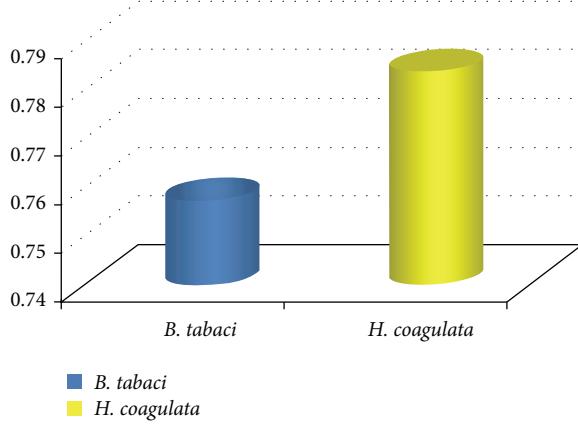


FIGURE 5: Comparison of CAI in *B. tabaci* and *H. coagulata*.

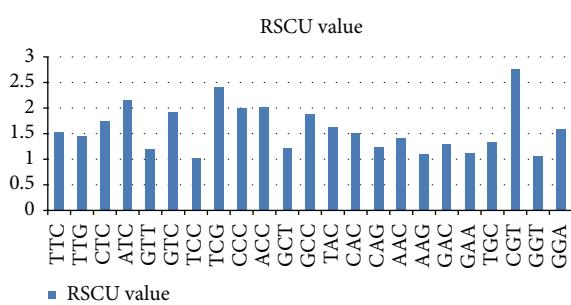


FIGURE 6: Codons used more frequently than expected (RSCU > 1) in *Bemisia tabaci* genes.

materials). In *B. tabaci* and *H. coagulata*, the most frequently used codons are the same and end with C at the 3rd position of codon. These codons are TTC, CTC, ATC, GTC, TCC, CCC, ACC, GCC, TAC, CAC, AAC, GAC, and TGC. In *H. coagulata*, the codon (GGC) has been found to be frequently used. These results indicate that the codon usage pattern in these two insect species is mostly contributed by compositional constraints. Figures 6 and 7 show the codons that are used more frequently than expected RSCU > 1 in *B. tabaci* and *H. coagulata*, respectively.

3.3. Amino Acids Contribute Differently to a Gene's Codon Usage Bias. Except methionine and tryptophan, the amino acid usage is different in the genes. Figure 8 shows that amino acid alanine is distinctly the single amino acid that accounts for the greatest usage and cysteine accounts for the least usage in the genes of both *B. tabaci* and *H. coagulata*, respectively.

3.4. Effect of Mutational Bias on Codon Usage Variation. To identify whether the evolution of codon usage bias in *B. tabaci* and *H. coagulata* had been driven by mutation pressure alone or whether the translational selection had also played a major role, we compared the correlation between general nucleotide composition (A, T, G, C, and GC%) and nucleotide composition at the third codon position (A3, T3, G3, C3, and GC3%) of *B. tabaci* and *H. coagulata* using the Spearman's

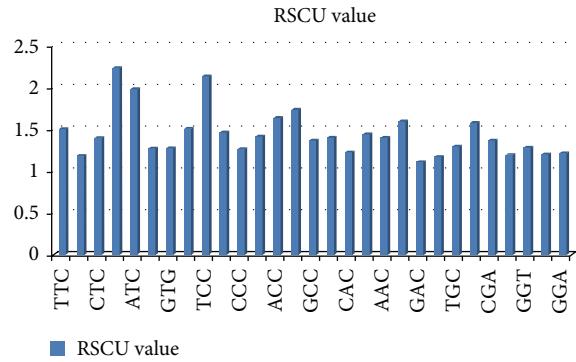


FIGURE 7: Codons used more frequently than expected (RSCU > 1) in *Homalodisca coagulata* genes.

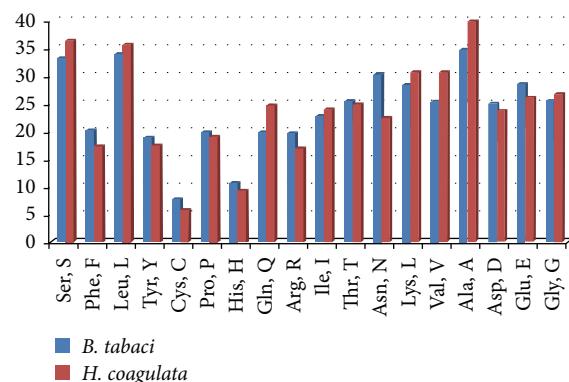


FIGURE 8: Distribution of amino acids in different genes of *Bemisia tabaci* ($n = 13$) and *Homalodisca coagulata* ($n = 20$).

rank correlation analysis method as shown in Tables 1 and 2, respectively.

In *B. tabaci*, a significant positive correlation was observed between A and A3% ($r = 0.682^{**}$, $P < 0.01$), C and C3% ($r = 0.880^{**}$, $P < 0.01$), and GC and GC3% ($r = 0.944^{**}$, $P < 0.01$) but significant negative correlation was observed between most of the heterogeneous nucleotide comparisons: A and GC3% ($r = -0.550^*$, $P < 0.05$), T and G3% ($r = -0.701^{**}$, $P < 0.01$), and T3% and C3% ($r = -0.869^{**}$, $P < 0.01$). In *H. coagulata*, a significant positive correlation was observed between A and A3% ($r = 0.841^{**}$, $P < 0.01$), C and C3% ($r = 0.846^{**}$, $P < 0.01$), and GC and GC3% ($r = 0.819^{**}$, $P < 0.01$) and significant negative correlation was observed between most of the heterogeneous nucleotide comparisons: A and GC3% ($r = -0.588^*$, $P < 0.05$), T and G3% ($r = -0.731^{**}$, $P < 0.01$), and T and C3% ($r = -0.308$, $P > 0.05$). These results suggest that compositional constraint under mutation pressure has contributed the pattern of codon usage bias in these two insect species within the order Hemiptera. However, a significant positive correlation between T and T3% ($r = 0.876^{**}$, $P < 0.01$) and G and G3% ($r = 0.304$, $P > 0.05$) and a negative correlation between GC and T3% ($r = -0.906^{**}$, $P < 0.01$) were found in *B. tabaci*. Similarly, a significant positive correlation between T and T3% ($r = 0.692^{**}$, $P < 0.01$) and G and G3% ($r = 0.599^*$,

TABLE 1: Correlation coefficients between overall nucleotide composition (A, T, G, C, and GC%) and nucleotide composition at 3rd position (A3, T3, G3, C3, and GC3%) in *Bemisia tabaci* genes.

Nucleotide	A3%	T3%	G3%	C3%	GC3%
A%	0.682**	0.420	-0.335	-0.336	-0.550*
T%	0.807**	0.876**	-0.701**	-0.869**	-0.862**
G%	-0.394	-0.222	0.304	0.058	0.324
C%	-0.817**	-0.850**	0.518*	0.880**	0.850**
GC%	-0.939**	-0.906**	0.717**	0.832**	0.944**

*,** Significant at $P < 0.05$ and $P < 0.01$, respectively.

TABLE 2: Correlation efficient between overall nucleotide composition (A, T, G, C, and GC%) and nucleotide composition at 3rd position (A3, T3, G3, C3, and GC%) in *Homalodisca coagulata*.

Nucleotide	A3%	T3%	G3%	C3%	GC3%
A%	0.841**	0.280	-0.071	-0.786**	-0.588*
T%	0.346	0.692**	-0.731**	-0.308	-0.533
G%	-0.269	-0.297	0.599*	0.049	0.236
C%	-0.797**	-0.346	-0.027	0.846**	0.621*
GC%	-0.896**	-0.654*	0.495	0.786**	0.819**

,* Significant at $P < 0.05$, $P < 0.01$, and $P < 0.001$, respectively.

$P < 0.05$) and a negative correlation between GC and T3% ($r = -0.654^*$, $P < 0.01$) in *H. coagulata* support the hypothesis that mutation pressure has been a major factor for the pattern codon usage bias in these species.

4. Discussion

The present investigation highlights the codon usage patterns in a comparative manner between the two important insect species, *Bemisia tabaci* and *Homalodisca coagulata*, belonging to the order Hemiptera. As synonymous codon usage is not uniform during translation process, the identification of the codon usage pattern is important to understand the translational selection of codons in protein coding genes in two species of Hemiptera.

Here we analyzed the synonymous codon usage bias in the two insect species belonging to the order Hemiptera. In this study, we found that the most frequent codons end with C in both species. This finding may be the result of compositional constraint that occurred in codon usage pattern in these insect species. Similar results were reported in other insect orders. In dipterans most frequent codons end with G or C mostly at 3rd codon position whereas in hymenopteran insect order, most frequent codons contain A or T at its 3rd codon position [40]. In *Drosophila* generally G and especially C are favored at synonymous sites in biased genes. However, some genes, even very tightly linked genes, can vary greatly in codon bias across species of *Drosophila*. Moreover, *Drosophila* genes are as biased or more biased than those in microorganisms [29]. In *B. tabaci*, the ENC value ranges from 43 to 60 with a mean of 52.80 and in *H. coagulata*, from 49 to 60 with a mean value of 56.69, respectively. Relatively high ENC values showed that codon bias was not remarkable in these species. But in *Drosophila* the distribution of effective number of codons was different

in short genes (300–500 bp) compared with longer genes [30]. In our study *H. coagulata* genes had greater gene expression than *B. tabaci* genes. In both of these species a highly significant positive correlation was observed between A and A3%, C and C3%, and GC and GC3%. In addition, a significant positive correlation was observed between T and T3% ($r = 0.8485^{**}$, $P < 0.01$), positive correlation between G and G3%, and a negative correlation between GC and T3%. The correlation between gene length and synonymous codon usage was negative in *Drosophila* [30]. These results suggested that mutation pressure had been a major factor for codon usage bias in these species. In *Drosophila*, the synonymous codon usage is well explained by tRNA availability and is probably influenced by developmental changes in relative abundance [15].

5. Conclusion

This is the first work on the comparative analysis of the pattern of codon usage in two insect species of Hemiptera order. This work is useful for understanding the evolution of pattern of codon usage in these species. Codon usage bias was not very remarkable in these species. Nucleotide constraint and compositional constraint are two significant factors as these affect the codon usage pattern in these species. The most frequent codons end with C at the 3rd position, most probably suggesting the role of the compositional constraint under mutation pressure. We find that mutation pressure is the main factor for the pattern of codon usage bias. However, further analysis would elucidate the role of any other factor that might be responsible for codon usage bias in insects.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

The authors are thankful to Assam University, Silchar, for providing necessary facilities for this research work.

References

- [1] M. Bulmer, "Are codon usage patterns in unicellular organisms determined by selection—mutation balance?" *Journal of Evolutionary Biology*, vol. 1, no. 1, pp. 15–26, 1988.
- [2] P. M. Sharp, M. Stenico, J. F. Peden, and A. T. Lloyd, "Codon usage: mutational bias, translational selection, or both?" *Biochemical Society Transactions*, vol. 21, no. 4, pp. 835–841, 1993.
- [3] R. M. Kliman and J. Hey, "Hill-Robertson interference in *Drosophila melanogaster*: reply to Marais, Mouchiroud and Duret," *Genetical Research*, vol. 81, no. 2, pp. 89–90, 2003.
- [4] H. Akashi, "Codon bias evolution in *Drosophila*. Population genetics of mutation-selection drift," *Gene*, vol. 205, no. 1-2, pp. 269–278, 1997.
- [5] T. Ikemura, "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," *Journal of Molecular Biology*, vol. 151, no. 3, pp. 389–409, 1981.
- [6] T. Ikemura, "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," *Journal of Molecular Biology*, vol. 151, no. 3, pp. 389–409, 1981.
- [7] T. Ikemura, "Codon usage and tRNA content in unicellular and multicellular organisms," *Molecular Biology and Evolution*, vol. 2, no. 1, pp. 13–34, 1985.
- [8] M. Bulmer, "Coevolution of codon usage and transfer RNA abundance," *Nature*, vol. 325, no. 6106, pp. 728–730, 1987.
- [9] H. Akashi, "Gene expression and molecular evolution," *Current Opinion in Genetics and Development*, vol. 11, no. 6, pp. 660–666, 2001.
- [10] C. G. Kurland, "Translational accuracy and the fitness of bacteria," *Annual Review of Genetics*, vol. 26, pp. 29–50, 1992.
- [11] A. Eyre-Walker, "Synonymous codon bias is related to gene length in *Escherichia coli*: selection for translational accuracy?" *Molecular Biology and Evolution*, vol. 13, no. 6, pp. 864–872, 1996.
- [12] P. M. Sharp and W.-H. Li, "An evolutionary perspective on synonymous codon usage in unicellular organisms," *Journal of Molecular Evolution*, vol. 24, no. 1-2, pp. 28–38, 1986.
- [13] D. C. Shields and P. M. Sharp, "Synonymous codon usage in *Bacillus subtilis* reflects both translational selection and mutational biases," *Nucleic Acids Research*, vol. 15, no. 19, pp. 8023–8040, 1987.
- [14] D. L. Hartl, E. N. Moriyama, and S. A. Sawyer, "Selection intensity for codon bias," *Genetics*, vol. 138, no. 1, pp. 227–234, 1994.
- [15] E. N. Moriyama and J. R. Powell, "Codon usage bias and tRNA abundance in *Drosophila*," *Journal of Molecular Evolution*, vol. 45, no. 5, pp. 514–523, 1997.
- [16] D. B. Carlini and W. Stephan, "*In vivo* introduction of unpreferred synonymous codons into the *Drosophila Adh* gene results in reduced levels of ADH protein," *Genetics*, vol. 163, no. 1, pp. 239–243, 2003.
- [17] J. B. Plotkin and G. Kudla, "Synonymous but not the same: the causes and consequences of codon bias," *Nature Reviews Genetics*, vol. 12, no. 1, pp. 32–42, 2011.
- [18] P. M. Sharp and G. Matassi, "Codon usage and genome evolution," *Current Opinion in Genetics and Development*, vol. 4, no. 6, pp. 851–860, 1994.
- [19] S. Vicario, E. N. Moriyama, and J. R. Powell, "Codon usage in twelve species of *Drosophila*," *BMC Evolutionary Biology*, vol. 7, no. 1, article 226, 2007.
- [20] A. Heger and C. P. Ponting, "Variable strength of translational selection among 12 *Drosophila* species," *Genetics*, vol. 177, no. 3, pp. 1337–1348, 2007.
- [21] F. G. Jørgensen, M. H. Schierup, and A. G. Clark, "Heterogeneity in regional GC content and differential usage of codons and amino acids in GC-poor and GC-rich regions of the genome of *Apis mellifera*," *Molecular Biology and Evolution*, vol. 24, no. 2, pp. 611–619, 2007.
- [22] J. K. Brown, D. R. Frohlich, and R. C. Rosell, "The sweetpotato or silverleaf whiteflies: biotypes of *Bemisia tabaci* or a species complex?" *Annual Review of Entomology*, vol. 40, no. 1, pp. 511–534, 1995.
- [23] A. C. Goheen, G. Nyland, and S. K. Lowe, "Association of a rickettsia-like organism with Pierce's disease of grapevines and alfalfa dwarf and heat therapy of the disease in grapevines," *Phytopathology*, vol. 63, pp. 41–45, 1973.
- [24] D. L. Hopkins and H. H. Mollenhauer, "Rickettsia-like bacterium associated with Pierce's disease of grapes," *Science*, vol. 179, no. 4070, pp. 298–300, 1973.
- [25] M. J. Davis, A. H. Purcell, and S. V. Thomson, "Pierce's disease of grapevines: Isolation of the causal bacterium," *Science*, vol. 199, no. 4324, pp. 75–77, 1978.
- [26] E. P. Rocha, "Codon usage bias from tRNA's point of view: redundancy, specialization, and efficient decoding for translation optimization," *Genome Research*, vol. 14, no. 12, pp. 2279–2286, 2004.
- [27] R. Hershberg and D. A. Petrov, "Selection on codon bias," *Annual Review of Genetics*, vol. 42, pp. 287–299, 2008.
- [28] P. M. Sharp, L. R. Emery, and K. Zeng, "Forces that influence the evolution of codon bias," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 365, no. 1544, pp. 1203–1212, 2010.
- [29] J. R. Powell and E. N. Moriyama, "Evolution of codon usage bias in *Drosophila*," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 94, no. 15, pp. 7784–7790, 1997.
- [30] E. N. Moriyama and J. R. Powell, "Gene length and codon usage bias in *Drosophila melanogaster*, *Saccharomyces cerevisiae* and *Escherichia coli*," *Nucleic Acids Research*, vol. 26, no. 13, pp. 3188–3193, 1998.
- [31] J. R. Powell, E. Sezzi, E. N. Moriyama, J. M. Gleason, and A. Caccone, "Analysis of a shift in codon usage in drosophila," *Journal of Molecular Evolution*, vol. 57, no. 1, pp. S214–S225, 2003.
- [32] F. Wright, "The 'effective number of codons' used in a gene," *Gene*, vol. 87, no. 1, pp. 23–29, 1990.
- [33] P. M. Sharp and W.-H. Li, "The codon adaptation index—a measure of directional synonymous codon usage bias, and its potential applications," *Nucleic Acids Research*, vol. 15, no. 3, pp. 1281–1295, 1987.
- [34] A. Eyre-Walker and M. Bulmer, "Reduced synonymous substitution rate at the start of enterobacterial genes," *Nucleic Acids Research*, vol. 21, no. 19, pp. 4599–4603, 1993.

- [35] G. Gutiérrez, J. Casadesús, J. L. Oliver, and A. Marin, “Compositional heterogeneity of the *Escherichia coli* genome: a role for VSP repair?” *Journal of Molecular Evolution*, vol. 39, no. 4, pp. 340–346, 1994.
- [36] G. Perrière, M. Gouy, and T. Gojobori, “NRSub: a non-redundant data base for the *Bacillus subtilis* genome,” *Nucleic Acids Research*, vol. 22, no. 25, pp. 5525–5529, 1994.
- [37] H. Akashi, “Synonymous codon usage in *Drosophila melanogaster*: natural selection and translational accuracy,” *Genetics*, vol. 136, no. 3, pp. 927–935, 1994.
- [38] D. R. Frohlich and M. A. Wells, “Codon usage patterns among genes for lepidopteran hemolymph proteins,” *Journal of Molecular Evolution*, vol. 38, no. 5, pp. 476–481, 1994.
- [39] B. R. Morton, “Codon use and the rate of divergence of land plant chloroplast genes,” *Molecular Biology and Evolution*, vol. 11, no. 2, pp. 231–238, 1994.
- [40] S. K. Behura and D. W. Severson, “Comparative analysis of codon usage bias and codon context patterns between dipteran and hymenopteran sequenced genomes,” *PLoS ONE*, vol. 7, no. 8, Article ID e43111, 2012.

