Translational effects of differential codon usage among intragenic domains of new genes in *Drosophila*

Jianming Zhang\(^a\), Manyuan Long\(^a\), Liming Li\(^b,\)*

\(^a\)Department of Ecology and Evolution, The University of Chicago, Chicago, IL 60637, USA
\(^b\)Department of Molecular Pharmacology and Biological Chemistry, Northwestern University, Institute of Neuroscience, The Feinberg School of Medicine, Searle 5-474, MC S205, 320 E. Superior St., Chicago, IL 60611, USA

Received 18 June 2004; received in revised form 7 February 2005; accepted 9 February 2005

Available online 2 March 2005

**Abstract**

Evolved codon usages often pose a technical challenge over the expressing of eukaryotic genes in microbial systems because of changed translational machinery. In the present study, we investigated the translational effects of intragenic differential codon usage on the expression of the new *Drosophila* gene, *jingwei* (*jgw*), a chimera derived from two unrelated parental genes: *Ymp* and *Adh*. We found that *jgw* possesses a strong intragenic differential usage of synonymous codons, i.e. the *Adh*-derived C-domain has a significantly higher codon bias than that of the *Ymp*-derived N-domain \((P=0.0023\) by *t*-test). Additional evolutionary analysis revealed the heterogeneous distribution of rare codons, implicating its role in gene regulation and protein translation. The in vitro expression of *jgw* further demonstrated that the heterogeneous distribution of rare codons has played a role in regulating gene expression, particularly, affecting the quality of protein translation.

© 2005 Elsevier B.V. All rights reserved.

**Keywords:** Codon usage; Intragenic codon bias; Evolution of new genes; Analysis of protein functions; Expression of insect genes; Short-chain dehydrogenase

1. Introduction

Nonrandom synonymous codon usage (codon bias) is observed in many species, likely caused by natural selection and mutational bias \([1,2]\). Codon bias plays a significant role in regulating gene expression in many organisms, especially in unicellular organisms, such as *E. coli* and *S. cerevisiae*, in which highly expressed genes show a strong codon bias \([1,3,4]\). During the translational process, the preferred codons in highly biased genes optimally bind to the most abundant isoaccepting tRNAs in those species, thus increasing translational efficiency by improving the speed and accuracy of protein synthesis \([1,3,5]\). Codon bias varies significantly among different genes in the same species and among homologous genes in different species \([6–8]\). The interspecies differential codon bias has been documented as a major difficulty in expressing foreign genes in conventional bacterial host cells \([9]\). In practice, adjusting tRNA pools in host cells is required to successfully express foreign genes \([10]\).

The role of differential codon usage is more complex in multicellular organisms. There is evidence that codon bias is the result of natural selection associated with gene expression. For example, *Sry* (*serendipity*) genes are expressed differently in two *Drosophila* species, *D. melanogaster* and *D. pseudoobscura*, which may be related to their change in level of codon usage bias \([11,12]\). Correlation between codon bias and gene expression level is also found in other organisms, such as *C. elegans* and *Arabidopsis* \([13]\). But in mammals, the degree of bias depends more on GC content than on the actual codon or tRNA abundance \([14,15]\).

In many cases, the force of natural selection on codon bias is weak. Therefore it is only effective in determining codon frequencies in organisms with a large effective population size \([16]\). For example, *Drosophila* generally has a much larger effective population size than that of...
mammals, hence natural selection would act more effectively on maintaining the biased codon usage in fruit flies.

Recent studies found that intragenic differential codon usage also exists in genes encoding multiple domains [17,18]. For instance, non-homogenous codon usage was observed in the luciferase of a unicellular marine organism, Gonyaulax polyedra: the preferred synonymous codons appear more often in the central part of three functional domains than in the boundary region [19]. Alternatively spliced genes in Drosophila were also shown to display a different codon bias between constant and alternative expression parts, although the significant association of this correlation needs to be statistically reanalyzed after eliminating GC content effects [20].

In the present study, we examined the evolutionary and translational effects of differential codon usage on a new gene in Drosophila—jingwei (jgw). The chimeric jgw gene in D. teissieri and D. yakuba was initially reported by Long et al. [21], and its genomic structure and evolutionary history were further identified [21,22]. It was demonstrated that the chimeric jgw gene was created when a processed alcohol dehydrogenase messenger RNA was reverse-transcribed and inserted into the third intron of yande (ynd) of D. melanogaster (Random primer fluorescein labeling kit, NEN Life Science Products) [22]. Positive clones were selected and their identities were confirmed by PCR and sequence analysis. For E. coli expression, jgw and Adh ORFs were subcloned to the pGEX-2TK expression vector in frame with the C-terminus of GST (Amersham Pharmacia).

2. Materials and methods

2.1. Cloning jgw and Adh genes

D. teissieri and D. yakuba cDNA libraries were constructed using the Uni-ZAP XR vector (Stratagene). Specifically, 0.5 mg of total RNA from each species was extracted from adult flies (RNAgent total RNA isolation system, Promega) according to the manufacturer’s manual. After mRNA separation (Separator Kit, Clontech), 0.5 µg mRNA was used to construct the cDNA library. The cDNA clones of jgw from D. teissieri and D. yakuba were isolated by hybridizing the corresponding cDNA libraries using a fluorescein-labeled probe comprised of the first three exons of Ymp of D. melanogaster (Random primer fluorescein labeling kit, NEN Life Science Products) [22]. Positive clones were selected and their identities were confirmed by PCR and sequence analysis. For E. coli expression, jgw and Adh ORFs were subcloned to the pGEX-2TK expression vector in frame with the C-terminus of GST (Amersham Pharmacia).

2.2. Expression of JGW in BL21 and BL21-CodonPlus-RIL

The pGEX protein expression and purification system (Amersham Pharmacia) were used to express jgw genes. Expression plasmids containing the fusion gene GST–jgw were transformed into E. coli BL21 (Stratagene). Transformants containing the GST–jgw expression plasmids were grown overnight at 25 °C in 2YTG medium supplemented with 100 µg/ml of ampicillin. The overnight culture was diluted 10 times in fresh 2YTG medium supplemented with 100 µg/ml ampicillin and grown for

Fig. 1. Gene structure and evolutionary history of jingwei (jgw) in D. yakuba. The colored boxes represent exons. Ymp: yellow emperor, with three constitutive expressed exons shown in dark blue and other alternative spliced exons in different colors for Ymp-1 and Ymp-2. The two pairs of arrows indicate the direction of retroposition (from Adh to jingwei) and the direction of duplication (from Ymp to jingwei).
additional 3 hrs. IPTG (isopropyl-beta-D-thiogalactopyranoside) was then added to a final concentration of 0.1 mM. After 5 hrs of additional growth, cells were harvested and lysed by sonication. Fusion proteins were purified by incubation of the clarified total cell lysate with Glutathione Sepharose 4B (Amersham Pharmacia) after the removal of cell debris by centrifugation at 20,000×g for 15 min at 4 °C (Amersham Pharmacia pGEX expression manual). To improve the yield and functionality of the recombinant jgw protein, an E. coli BL21-CodonPlus-RIL strain (Stratagene) expressing E. coli minor tRNA genes argU, ileY, and leuW from a chloramphenicol-resistant vector [23] was also used for jgw expression. The growth and induction condition were the same as for the BL21 strain except chloramphenicol was added at a final concentration of 34 μg/ml for maintaining the extra tRNA genes in BL21-CodonPlus-RIL.

2.3. Codon bias analysis

Major or minor codons are defined as those that increase or decrease in frequency as a function of the codon bias level of Drosophila genes respectively. Major and minor codon tables were constructed based on data from D. teissieri and D. yakuba from the codon usage database by Nakamura, Y., Department of Plant Gene Research, Kazusa (http://www.kazusa.or.jp/codon), using methods described by Shields et al. [4] and Akashi [24]. In addition, we used the set of major codons defined by Akashi [24] as a reference to eliminate database bias caused by multiple entries of a single gene. GCUA (General Codon Usage Analysis) [25] was applied to analyze the codon bias for jgw, Ymp and Adh. The effective number of codons (ENC) [26] was used to measure how far the codon usage in a gene departs from equal usage of synonymous codons. ENC ranges from 20 (if only one codon is used for each amino acid) to 61 (if all synonymous codons are used equally). ENCs were calculated using the program CHIPS from the EMBOSS package (http://www.bgmp.mrc.ac.uk/Software/EMBOSS/) [25,27]. Statistical analysis was conducted using the R statistical package (http://www.r-project.org/).

2.4. Assay of enzymatic activity

Using the purified GST–jgw from the BL21-CodonPlus-RIL strain, we were able to perform biochemical analyses to characterize this novel chimeric protein. Since the major part of JGW was derived from alcohol dehydrogenase, we tested its dehydrogenase functionality by measuring its catalytic capability of converting different alcohols to corresponding aldehydes/ketones. Kinetic measurements were performed with various concentrations of alcohols and a fixed concentration (500 μM) of NAD+ throughout in a buffer containing 0.025 M MOPS pH 7.4, or 0.025 M CHES pH 9.4, and 0.1 M NaCl.

3. Results

3.1. Codon bias analysis

We first analyzed the major codon usage of Ymp. By comparing the constitutive exon region (the 5’ of Ymp) and the alternative splicing region (the 3’ side of Ymp), we found that the major codon usage is more biased at the 3’ terminal (Table 1). This was not expected since the usage of major codons is usually higher in constitutive exons than that in alternative exons due to the fact that constitutively expressed exons are translated more often than those in alternatively spliced exons [20].

We further analyzed the effective number of codons (ENC) of Adh, jgw and Adh-jgw (Adh-derived portion of the jgw gene) by implementation of the CHIPS program. A significant difference in codon bias between Adh and jgw was observed and summarized in Table 2. In comparison of Adh-jgw to the whole gene jgw, the P value (P=0.017) of a t-test showed a relaxation of codon bias in the Adh-jgw region. However, the difference in codon bias between Adh-jgw and Adh is statistically insignificant (P=0.44).

Next, we examined if there are any differences in minor codon usage in Adh and jgw, especially the usage of those rare codons (rare codons refer to codons whose usages are notably lower than those of other minor codons). It has been shown that different tRNA pools exist in different tissues and/or in different developmental stages in Drosophila [12]. If there is a homogenous tRNA pool, the protein products of genes with rare codon usage could be limited due to a slower translational speed [28]. Therefore, rare codon usage can strongly influence protein translation and it is important to determine the expression level of a particular protein. Indeed, a significant difference in rare codon usage was also observed in jgw, Ymp, Adh-jgw, and Ymp-jgw (Ymp-derived portion of jgw). For example, none of the Leu codons with A at the third position (NNA) is used in Ymp-jgw (the ratio of NNA/NNG is 62/38). Similarly, the ratio of Gly codons, GGA/GGG, is 25:1 in Ymp-jgw, but this ratio became 7:14 in Ymp-jgw. Table 3 shows the intragenic differential rare codon usage in jgw. The codon

<table>
<thead>
<tr>
<th>Gene</th>
<th>Constant exons</th>
<th>Alternative exons</th>
</tr>
</thead>
<tbody>
<tr>
<td>D. mel-Ymp-1</td>
<td>0.24</td>
<td>0.4</td>
</tr>
<tr>
<td>D. mel-Ymp-2</td>
<td>0.24</td>
<td>0.41</td>
</tr>
<tr>
<td>D. tei-Ymp-1</td>
<td>0.21</td>
<td>0.38</td>
</tr>
<tr>
<td>D. yak-Ymp-2</td>
<td>0.25</td>
<td>0.39</td>
</tr>
<tr>
<td>D. tei-jgw-Ymp</td>
<td>0.31</td>
<td></td>
</tr>
<tr>
<td>D. yak-jgw-Ymp</td>
<td>0.26</td>
<td></td>
</tr>
</tbody>
</table>

Codon usage is towards more bias at the 3’ terminus. D. mel: D. melanogaster; D. tei: D. teissieri; and D. Yak: D. yakuba. Constant exons: constant expressed exons; alternative exons: alternative spliced exons.
usage analysis indicates a significant differential codon bias within jgw gene.

3.2. In vitro expression

We observed that the fusion gene GST–jgw could be expressed in E. coli BL21 but could not be further purified, suggesting that the fusion protein might not be able to bind to Glutathione Sepharose 4B (Fig. 2). Further experiments for troubleshooting indicated that functional JGW products could not be produced despite the extensive changes in the expression conditions, such as modulating the concentration of IPTG (isopropyl-1-thio-β-D-galactopyranoside), induction time, media pH, and bacterial growth temperature (unpublished data). Since both Adh and Adh-jgw were expressed and purified well in the E. coli BL21 system under identical conditions, we inferred that the abundance of rare codons in the Ymp-derived region might be the reason for producing non-functional protein, perhaps by the misincorporation of amino acids at those rare codon sites in the E. coli cells. As shown in Table 3, rare codons CUA and CUU encoding Leu are frequently used at the 5' portion of jgw (Ymp-jgw) and other rare codons are also clustered in this region (Table 3). We hypothesized that these rare codons could be responsible for a dysfunctional expression.

To test this hypothesis, we transformed the plasmid containing GST–jgw into the E. coli BL21-CodonPlus-RIL strain (Stratagene), which contains E. coli CodonPlus-RIL as a chloramphenicol-resistant vector [23]. Those three extra tRNA genes argU, ileY, and leuW from a chloramphenicol-resistant vector [23]. Indeed, our subsequent experimental results proved that jgw could be successfully expressed in modified BL21-CodonPlus-RIL strains (Fig. 2). We were also able to purify JGW and demonstrate its capability to catalyze to ethanol and isopropanol using native gel staining method (data not shown).

3.3. Dehydrogenase functionality of JGW

Further enzyme kinetic analysis indicated that JGW preserved the short chain dehydrogenase functionality (Fig. 3). Similar to ADH, one of its parental proteins, JGW, is able to catalyze the conversion reaction from alcohols to corresponding aldehydes/ketones. However, ADH and JGW differ remarkably in their affinity to alcohols to correspondingly to ethanol and isopropanol using native gel staining method (data not shown).

The normality of ENC is confirmed by Kolmogorov–Smirnov Test.

Table 2
Effective number of codons of ymp, Adh, jgw and Adh-jgw

<table>
<thead>
<tr>
<th>Gene</th>
<th>D. mel</th>
<th>D. tei</th>
<th>D. yak</th>
</tr>
</thead>
<tbody>
<tr>
<td>ymp</td>
<td>61.0</td>
<td>59.1</td>
<td>54.2</td>
</tr>
<tr>
<td>Adh</td>
<td>31.5</td>
<td>28.7</td>
<td>28.5</td>
</tr>
<tr>
<td>Adh-jgw</td>
<td>30.9</td>
<td>32.6</td>
<td></td>
</tr>
<tr>
<td>jgw</td>
<td>40.5</td>
<td>41.7</td>
<td></td>
</tr>
</tbody>
</table>

The t-test shows the significant difference of ENC between ymp and Adh (P=0.0013), Adh-jgw and Adh (P=0.44), and jgw and Adh (P=0.0023). Adh-jgw denotes the Adh-derived part of jgw. The normality of ENC is confirmed by Kolmogorov–Smirnov Test.

The distribution of rare codons in jgw of D. teissieri and D. yakuba

<table>
<thead>
<tr>
<th>Rare codons</th>
<th>D. teissieri</th>
<th>D. yakuba</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ymp-jgw</td>
<td>Adh-jgw</td>
</tr>
<tr>
<td>Leu CUA</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>Leu CUU</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Pro CCU</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>His CAU</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Arg CGU</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>CGG</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>AGG</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Gly GGG</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Ile AUA</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>18</td>
<td>14</td>
</tr>
</tbody>
</table>

Adh-jgw: Adh-derived part of jgw; Ymp-jgw: Ymp-derived part of jgw. The rare codons are mainly distributed on the Ymp-derived domain of jgw.

Fig. 2. Crude protein extracts (top) and purified JGW (bottom) analyzed by SDS/PAGE and Coomassie-Blue staining. Arrows indicate the molecular weights of corresponding fusion proteins. Crude extracts and purified GST-fusion proteins from BL21 are in the left box, and those from BL21-CodonPlus-RIL are in the right box. The lanes from left to right in the left box: TADH (ADH of D. teissieri), YADH (ADH of D. yakuba), TJJA (ADH-derived domain of JGW of D. teissieri), YJA (ADH-derived domain of JGW of D. yakuba), TIGW (JGW of D. teissieri), and YJGW (JGW of D. yakuba). The lanes from left to right in the right box are: TADH, TJIA and TIGW.

Further enzyme kinetic analysis indicated that JGW preserved the short chain dehydrogenase functionality (Fig. 3). Similar to ADH, one of its parental proteins, JGW, is able to catalyze the conversion reaction from alcohols to corresponding aldehydes/ketones. However, ADH and JGW differ remarkably in their affinity to substrate binding (Fig. 3). For example, at pH 7.4, the Km values of JGW of D. teissieri towards ethanol (primary alcohol) and isopropanol (secondary alcohol) are 1.78 mM and 0.31 mM, respectively, whereas the Km values of ADH of D. teissieri for ethanol and isopropanol are 3.23 mM and 1.24 mM, respectively. This kinetic analysis therefore indicates that the alcohol dehydrogenase activity is retained in the novel chimerical enzyme, JGW. Moreover, the additional N-terminal domains from the captured
three Ymp exons are contributed to the regulation of JGW’s expression.

4. Discussions

4.1. Evolution of new genes and codon bias

Although the role of codon bias in protein translation has been appreciably acknowledged, the effect of differential intragenic codon usage in translation, especially the mechanism of creating such a bias, is poorly understood [1,8,10]. We have undertaken the task first time to examine the relation of the origination and evolution of new genes to the existence of differential intragenic codon usage. We have shown that the young chimeric gene, jgw, exhibits a very strong intragenic codon bias and we also demonstrated that such an intragenic codon bias has a profound effect on heterogeneous expression. It is likely that such a differential intragenic codon usage observed in jgw could be a common phenomenon for young chimeric genes that use gene fragments from distinct sources with different codon usages. New genes created by exon shuffling or domain shuffling could be subject to codon usage differentiation from the outset since they would have inherited different expression patterns from their distinct parental genes, such as differences in steady level abundance, tissue specific and/or development specific expressions. This differentiation might decrease during the long evolutionary process of the new genes as a consequence of substitutions in evolution towards a new function. Jgw inherited its testis specific expression pattern from one parental gene Ymp by recruiting the latter 5′ flanking regulatory region [22], thus is completely different from that of another parental gene Adh, which is a highly expressed in midgut and fat body.

If there exist any tissue-specific or developmental specific tRNA pools in codon biased species, genes whose expressions are tissue specific or developmental stage specific would take advantage of these specific tRNA pools relaxation for regulating gene expression. Consequently, a clear-cut pattern of differential codon usages would be found in those genes. There is, however, no immediate evidence from Drosophila to support the tissue-specific differential codon usage. For example, as Powell and Moriyama [12] pointed out, the avoidance of the three codons noted AUA, GGG, and UUA is shared by genes with different expression patterns, such as Adh, chorions, amylase, myosin and glyceraldehyde-3-phosphate dehydrogenase. However, the existence of differential tRNA pools is not impossible. A limited tRNA pool data from Drosophila indicated that the major species of tRNA might shift in different developmental stages by using different anticodons [30]. For instance, anticodon QUC is used in adults whereas GUC is used in the first and third instar. The consequence of the third nucleotide modification of anticodon is unknown [12,30,31]. In other multicellular eukaryotes, tissue-specific tRNA abundances have been found in tissues with highly specific functions and expression of a limited number of genes. For example, the concentrations of tRNAs match with the amino acid usage.

<table>
<thead>
<tr>
<th>Alcohols</th>
<th>TADH</th>
<th>TJGW</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ethanol</td>
<td>3.23</td>
<td>1.24</td>
</tr>
<tr>
<td>Isopropanol</td>
<td>1.78</td>
<td>0.31</td>
</tr>
<tr>
<td>Vmax (mM/min)</td>
<td>7.14</td>
<td>2.22</td>
</tr>
<tr>
<td></td>
<td>2.18</td>
<td>2.72</td>
</tr>
</tbody>
</table>

Fig. 3. Enzyme kinetics using substrates of primary (ethanol) and secondary (isopropanol) alcohols at pH 7.4 by JGW and ADH. The dehydrogenase activities of JGW and ADH were measured as described in Materials and methods and their corresponding values of $K_m$ and $V_{max}$ were measured by Double-reciprocal plot.
of fibroin in the posterior silk gland of the silkworm *Bombyx mori* L. and of hemoglobin in rabbit and human reticulocytes [32,33].

### 4.2. Translational selection

Carlini and Stephan [34] designed an elegant experiment to detect the signature of natural selection on the unpreferred codons. To examine whether changes from preferred codons to nonpreferred codons have any effect on ADH activity, they replaced various numbers of preferred Leu codons with CUA, the nonpreferred codon. They observed a significant decrease in ADH activity when they replaced one, six or ten preferred Leu codons with CUA, in the *Drosophila Adh* coding region. Moreover, the decreased activity is proportional to the number of preferred codons being replaced (Fig. 4). Presumably, the rare codons decrease the enzymatic activity by either decreasing the translation efficiency [35] or preventing correct folding according to an additive model. Taking Carlini and Stephan’s in vivo data [34] as reference, one can anticipate that only 37.7% of activities or products (Fig. 4) would be obtained as the expression of *jgw* in *E. coli* BL21 compared to its expression in BL21-Codon-Plus-RIL, based on the fact that 27 of the unpreferred codons in *D. teissieri* (26 in *D. yakuba*) are also rarely used in *E. coli* (Table 4). We notice that the change of codon usage of Leu in *D. melanogaster Adh* gene is one of the most biased in Carlini et al.’s experiments. However, our estimation of the reduction of active proteins is reasonable, because this number does not include the effects of clustered rare codons, which likely causes a more serious translational problem than those of randomly distributed. The translational effect of those rare codons on the expression of JGW is, however, eliminated by introducing new tRNA genes. Interestingly, the difference in total amount of expressed proteins was not significant in these two different *E. coli* strains (Fig. 2), however, the expressed protein in BL-21 was completely nonfunctional, supporting the notion that translational accuracy is an important target of natural selection [15].

Translational efficiency has two interrelated effects: translational speed and accuracy. The issue of selection for speed vs. accuracy is difficult to disentangle, and both may be affected by codon usage [16,18]. In bacteria and yeast, the corresponding of tRNA abundance to the genome codon usage indicates that high-level expression results in the depletion of the internal tRNA pools. Consequently, the translation of an unbiased mRNA is delayed. Most abnormal translation occurs during the waiting time for the “search” for the ternary complex (aminoacyl-tRNA-elongation factor Tu-GTP in bacteria) that matches the codon being translated; the longer the waiting time, the higher the probability of abnormality [9,36]. Hence genes translated rapidly are also translated more accurately. The data of expression of *jgw* in *E. coli* indicated that translational accuracy was a major target of natural selection. The fact that the selection for accuracy may account for at least some of the codon bias in *Drosophila* was also observed [17]. He examined the codon usage of 38 homologous genes from *D. melanogaster*, *D. pseudoobscura*, and *D. viridis* and found that in genes with weak codon bias, conserved amino acids had higher codon bias than nonconserved residues. In regions encoding important protein motifs, e.g. homeodomains and zinc finger domains, the frequency of preferred codons is higher than in the remainder of the gene, which suggested that selection for translational accuracy caused this bias. It was reasoned that selection would be the greatest to against to misincorporation of amino acids at crucial functional sites in a protein [17]. This pattern, however, was not observed in *E. coli* [37]. It may be due to the fact that the transcription–translation coupling mechanism affects the stability of mRNA.

### 4.3. Interspecies differential codon usage and gene expression

Overcoming interspecies differential codon usage is a challenge to protein functional and structural studies. Dysfunctional and aborted translation products were often observed in forced high-level expression of recombinant proteins in conventional host cells, such as *E. coli* and yeast. Thus, it is worthwhile to analyze intragenic codon usage before planning an experiment. Our study demonstrates that the detrimental effects of codon bias on heterogeneous expression can be eliminated by combining analysis of codon usage distribution and selection of an appropriate expression system [38]. In the case of JWG, the difficulty of obtaining soluble and functional recombinant JWG was overcome by using host bacteria with extra copies of the
cognate tRNA genes, such as the strain BL21-CodonPlus-RIL. An alternative strategy to solve the codon bias problem may be to replace the rare codons with preferred ones by site-directed mutagenesis [39].

4.4. Evolution of differential intragenic codon usage in jgw

The heterogeneous distribution of rare codons in the young chemical gene jgw provides a window to detect the functional role of synonymous codons in the evolution of new genes. Previous studies indicate that male specific genes show low codon bias and rapid evolution rates [40]. The faster evolutionary rates were primarily due to a higher nonsynonymous substitution ratio in the male-biased gene and driven by positive selection [40]. We have measured the GC contents in the region of jgw. The GC content of three introns of jgw are much lower than exons in both D. teissieri and D. yakuba, 33% in D. teissieri jgw and 38% in D. yakuba jgw. The 5′ flank region of jgw from D. teissieri also showed a lower GC content (39% in the total length of 800 base pairs). The GC content of jgw exons, however, is much higher, 57% for D. teissieri jgw and 55% for D. yakuba jgw. Those results with previous observations from population genetic analyses of the pattern of nucleotide variation surrounding the intron polymorphism[41] indicate that the codon usage of jgw likely under the influence of selection pressure.

Genome-wide codon usage analyses have revealed significant intragenic differential codon usage in E. coli, Yeast and Drosophila [42]. For most genes, their middle regions show higher codon biases than their corresponding termini do [43]. This observation implies a possible mechanism of creating and/or maintaining the intragenic codon bias in jgw, i.e., the differential strength of natural selection forces acting on different domains of the new gene.

The comparative structural modeling suggests that the major male specific protein classes, 28 out of 52, are conserved in Drosophila seminal fluid, despite the determined rapid evolution rate at the primary sequence level [44]. However, the comprehensive physiological consequences of such changes in amino acid residues at these critical sites have not been measured. The amplified fitness effects could not be ignored, such as the translational effects of synonymous codon usage we report in this study. Additionally, the broader dehydrogenase functions acquired by JGW with its taxon-specific evolutionary pattern imply its very candidacy for a physiological modifier.

Acknowledgements

We thank the members of Long’s lab for their experimental assistance and helpful discussions, especially Ana Llopart’s assistance for cloning jgw from D. teissieri, and Hong Qin, Janice Spofford, Todd Martin, and two anonymous reviewers for suggestions and discussions. This project was supported by US Army Medical Research and Material Command grant DAMD17-03-1-0282 to L.L and by a NIH R01 grant (GM65429-01A1) and a NSF CAREER award (0238168) to M.L.

References


