



Article

De Novo Assembly and Characterization of the Transcriptome of an Omnivorous Camel Cricket (*Tachycines meditationis*)

Jun-Hui Lu, De-Long Guan , Sheng-Quan Xu * and Huateng Huang *

College of Life Sciences, Shaanxi Normal University, Xi'an 710119, China

* Correspondence: xushengquan@snnu.edu.cn (S.-Q.X.); huanghuateng@snnu.edu.cn (H.H.)

Abstract: *Tachycines meditationis* (Orthoptera: Rhaphidophoridae: Tachycines) is a widely distributed insect in eastern Asia. This species is common in urban environments, and its unique omnivorous diet may contribute to its success in various habitats. However, molecular studies on the species are scarce. Here, we obtained the first transcriptome sequence of *T. meditationis* and performed preliminary analyses to test whether the evolution of coding sequences fits the expectations based on the species' ecology. We retrieved 476,495 effective transcripts and annotated 46,593 coding sequences (CDS). We analysed the codon usage and found that directional mutation pressure was the leading cause of codon usage bias in this species. This genome-wide relaxed codon usage pattern in *T. meditationis* is surprising, given the potentially large population size of this species. Moreover, despite the omnivorous diet, the chemosensory genes of this species do not exhibit codon usage deviating significantly from the genome-level pattern. They also do not seem to experience more gene family expansion than other cave cricket species do. A thorough search for rapidly evolved genes using the dN/dS value showed that genes associated with substance synthesis and metabolic pathways, such as retinol metabolism, aminoacyl-tRNA biosynthesis, and fatty acid metabolism, underwent species-specific positive selection. While some results seem to contradict the species ecology, our transcriptome assembly provides a valuable molecular resource for future studies on camel cricket evolution and molecular genetics for feeding ecology in insects, in general.

Keywords: camel cricket; chemosensory genes; codon usage pattern; transcriptome



Citation: Lu, J.-H.; Guan, D.-L.; Xu, S.-Q.; Huang, H. De Novo Assembly and Characterization of the Transcriptome of an Omnivorous Camel Cricket (*Tachycines meditationis*). *Int. J. Mol. Sci.* **2023**, *24*, 4005. <https://doi.org/10.3390/ijms24044005>

Academic Editor: Xesús Feás

Received: 2 January 2023

Revised: 12 January 2023

Accepted: 16 January 2023

Published: 16 February 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/>).

1. Introduction

Tachycines meditationis (Orthoptera: Rhaphidophoridae: Tachycines), also known as a camel cricket or cave cricket, is widely distributed throughout eastern and central China [1]. In urban environments, the success of *T. meditationis* is comparable to that of cockroaches, making it one of the most commonly observed insects [2]. However, molecular biology studies on this species are scarce [3]. The genomic and functional landscapes of *T. meditationis* and the underlying genetic basis for its adaptation to urban environments remain poorly studied.

In this study, we conducted the first transcriptome sequencing of *T. meditationis*. Although transcriptome sequencing provides less coverage than whole-genome sequencing, it has several advantages, including low cost, efficiency, and targeting, specifically for functional coding sequences (CDS) [4]. Transcriptomic analysis can reveal an organism's expressed gene sequences, annotations, and expression levels [5,6]. For research that is less concerned with other genomic factors (e.g., repetitive elements), the transcriptome provides sufficient data for probing functional components of a given organism's genome [7].

We analysed codon usage patterns based on de novo transcriptome assembly [8]. Codon bias has been proven to be related to a variety of biological functions, such as translation molecular mechanisms [9] and tRNA content [10]. Analysing genome-wide codon usage patterns is one of the most efficient and powerful ways to quantitatively measure the influence of different evolutionary forces, such as mutation pressure, selection

constraints, and genetic drift in a species [11,12]. The observation that the most common codons in highly expressed genes match the most abundant tRNAs makes codon usage bias a classic example of selection pressure operating at the genome level [13]. Owing to the weak selection for translational efficiency [14], the population-genetic theory predicts that the effect of selection would be more apparent in a larger effective population size. Given that this camel cricket species is widely distributed and successfully adapted to the urban environment, we expect its large effective population size to make the selection an efficient and dominant force in shaping the codon usage pattern.

To characterise gene functions, we annotated genes and estimated their dN/dS values to identify the top genes/pathways that may undergo positive selection [8,15,16]. In particular, we focused on chemosensory genes. Chemosensory genes play a very important role in the daily behavior of creatures, such as predation [17], reproduction [18], and avoidance of natural enemies [19]. In insects, chemosensory genes play a crucial role in identifying volatile and nonvolatile components [20,21], such as sugars and alkaloids in plants. Chemosensory genes help organisms perceive their external environment and respond accordingly [22]. The chemosensory supergene family includes odorant-binding proteins (OBP), chemosensory proteins (CSP), odorant receptors (OR), ionotropic receptors (IR), odorant-degrading enzymes (ODE), sensory neuron membrane proteins (SNMP), and gustatory receptors (GR) [23–25]. Previous studies have found that changes in the insect chemosensory system lead to changes in the adaptability of their diet [26]. The German cockroach, *Blattella germanica*, is an example of this theory; it has an extremely omnivorous diet and the most extensive chemosensory gene repertoire known for arthropods, with at least 897 IR and GR genes [27]. *T. meditationis* also has an omnivorous diet, which might contribute to its success indoors. Hence, we expect the evolution of this diet to be associated with the expansion of chemosensory gene families.

Interestingly, neither of the predictions based on camel cricket ecology—strong codon usage bias shaped by selection and expanded chemosensory gene families—is supported by our transcriptome analysis. This highlights how little is known about the genetic background and molecular biology of this species. Future investigations into the population genetics and functional genomics of camel crickets will aid in understanding the variety of molecular mechanisms underlying the processes of insects colonising urban environments.

2. Results

We obtained two sets of 9.32 Gb transcriptomic sequencing data for *T. meditationis*. A total of 476,495 raw transcripts, 111,197 unigenes, and 46,593 CDSs were identified. The average lengths of the transcripts, unigenes, and CDSs were 572.3 bp, 368.8 bp, and 488.7 bp, respectively.

The percentages of the four nucleotides in all CDSs were similar: Adenine (A; 26.1%), Thymine/Uracil (T/U; 23.9%), Guanine (G; 25.5%), and Cytosine (C; 24.5%). G and C accounted for 50% of the total nucleotide content. Most CDSs had a GC content between 30% and 70% (Figure 1a). Approximately 1000 CDSs had a GC content of less than 30% or more than 70%, and more than half of the CDSs had a GC content between 50% and 70%. Overall, the CDSs were slightly GC-biased in *T. meditationis*. Figure 1b shows the GC content at three codon positions. The GC content in the third position was higher than that in the first and second positions, and the overall GC content resembled the composition at the first codon position (Figure 1b). In addition, the frequencies of sixteen types of dinucleotides were analysed at different coding positions. The ratio between the observed frequency and the theoretical frequency, which is the mathematical product of the corresponding nucleotide content, was calculated (Figure 1c). Four dinucleotides (CU, GA, GC, and UU) showed relatively higher usage in the first and second positions, with AG, CG, UA, and UG having the lowest usage.

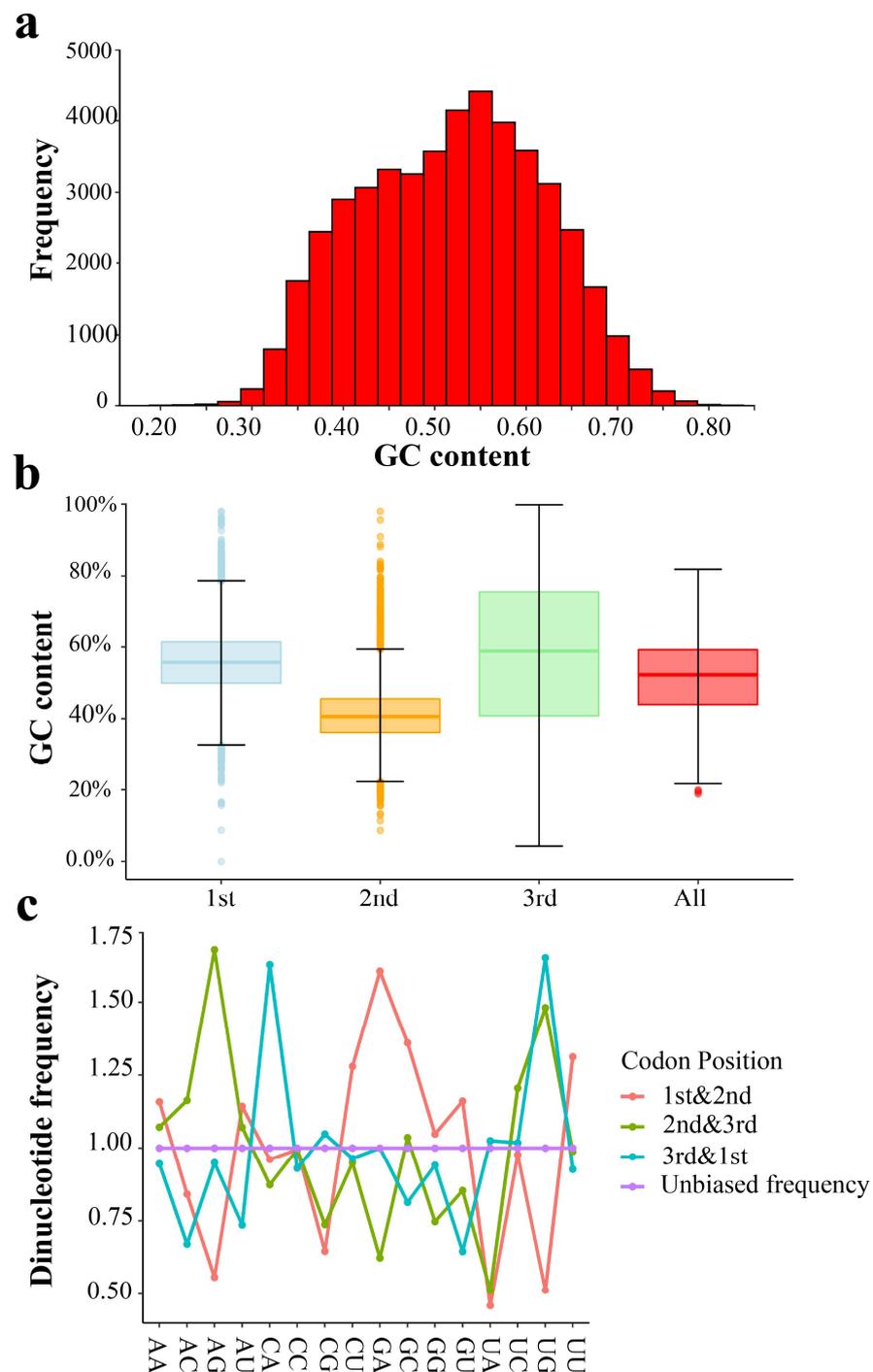


Figure 1. Base composition of the assembled transcriptomic data. (a) Distribution of GC content levels among coding sequences (CDS). (b) Box plot showing GC content variation among different codon positions and the genome average (All). (c) Dinucleotide frequencies at different codon position combinations (1st and 2nd, 2nd and 3rd and 3rd and 1st). The straight line represents the theoretical equal frequency (1/16).

2.1. Codon Usage Bias in *T. mediterraneus*

At the genome level, no codon (Table 1) has an RSCU (relative synonymous codon usage) value above 1.5 or below 0.5, two thresholds for preferred and avoided codons often used in the literature. The 33 codons with RSCU > 1 (Table 1) were slightly UC-biased at the last base: A (4), G (7), U (10), and C (12). It seems that the CDS in *T. mediterraneus* preferred codons ending in UC.

Table 1. Overall relative synonymous codon usage (RSCU) values of all CDSs in *T. meditationis*.

AA	Codon	RSCU	AA	Codon	RSCU	AA	Codon	RSCU	AA	Codon	RSCU
Ala	GCU	1.122	His	CAC	1.003	Pro	CCA	1.076	Ser	UCG	0.768
	<u>GCG</u>	<u>0.687</u>		CAU	0.997		CCC	1.014		UCU	1.187
	GCC	1.301	Ile	AUU	1.190	Gln	CCU	1.182	Thr	ACC	1.223
GCA	0.891	<u>AUA</u>		<u>0.465</u>	CCG		0.727	ACA		1.021	
Cys	UGU	1.010	Lys	AUC	1.345	Arg	CAA	0.849	Trp	<u>ACG</u>	<u>0.674</u>
	UGC	0.990		AAA	0.784		CAG	1.151		ACU	1.082
Asp	GAU	1.023	Leu	AAG	1.216	Ser	AGA	1.187	Tyr	GUU	1.048
	GAC	0.977		<u>CUA</u>	<u>0.392</u>		AGG	0.813		GUG	1.282
Glu	GAG	1.067	Met	CUC	1.148	Asn	CGA	0.772		GUC	1.126
	GAA	0.933		CUG	1.478		CGC	1.460		<u>GUA</u>	<u>0.545</u>
Phe	UUU	0.847	Asn	CUU	0.982	Ser	CGG	0.668		UGG	1.000
	UUC	1.153		UUA	0.793		CGU	1.100		UAC	1.083
Gly	GGU	1.153	Met	UUG	1.207	Asn	AGC	1.060		UAU	0.917
	<u>GGG</u>	<u>0.519</u>		AUG	1.000		AGU	0.940			
	GGC	1.388		AAC	1.062		UCA	1.048			
	GGA	0.941		AAU	0.938		UCC	0.998			

Notes: All codons except for stop codons were included in this analysis. AA: amino acids. Codons with RSCU > 1.25 are shown in bold, codons with RSCU < 0.75 are underlined.

This preference could be because of a selection or mutation bias. To disentangle these two factors, we calculated the number of preferred codons using GC-conservative codon pairs and gene expression data (i.e., FPKM values). No GC-conservative codon pairs showed significant usage changes correlating with gene expression—all *p* values were greater than 0.1 (Supplementary Table S1), suggesting that the weak codon bias we observed is unlikely due to translational selection.

Our conclusion was further supported by analysis of the variations in codon usage bias among genes. The ENC (effective number of codons) plot showed that the ENC values of most genes were between 40 and 60, indicating a relaxed codon bias, and did not deviate significantly from the value expected based on the GC content of the genes (Figure 2a). The deviation ratios, calculated as $(ENC_{exp} - ENC_{obs})/ENC_{exp}$, mostly fell between 0 and 0.1 (Figure 2b). The correlation between ENC value and gene expression was $p < 0.001$, but with an $R^2 = 0.005$, suggesting that gene expression can explain very little variation in codon usage bias among genes. Correlations between the ENC values and indices for protein properties (e.g., protein length and amino acid profile) were insignificant. The first two principal components (PC1 and PC2) were used for RSCU values; they accounted for 24% and 6% of the total variation, respectively. The only strong correlation we found was between PC1 and GC3 ($R^2 = 0.9615$, $p < 2.2 \times 10^{-16}$), suggesting a dominant effect of GC content on codon usage bias (Figure 2c,d).

2.2. Evolution of Chemosensory Genes in *T. meditationis*

Based on annotation, we identified 46 putative OBP genes, 22 CSPs, 38 ORs, 13 IRs, 6 SNMPs, and 5 GRs in *T. meditationis* (Figure 3a). There were more genes in OBP and OR than in GR and SNMP, consistent with the sizes of these gene families in most insects [7,26]. The ENC and GC12/GC3 values of the six gene families in *T. meditationis* were further compared with those of the whole gene sequence, and the results showed that the codon bias of the chemosensory genes did not differ from the codon bias of the whole genome (Figure 3b).

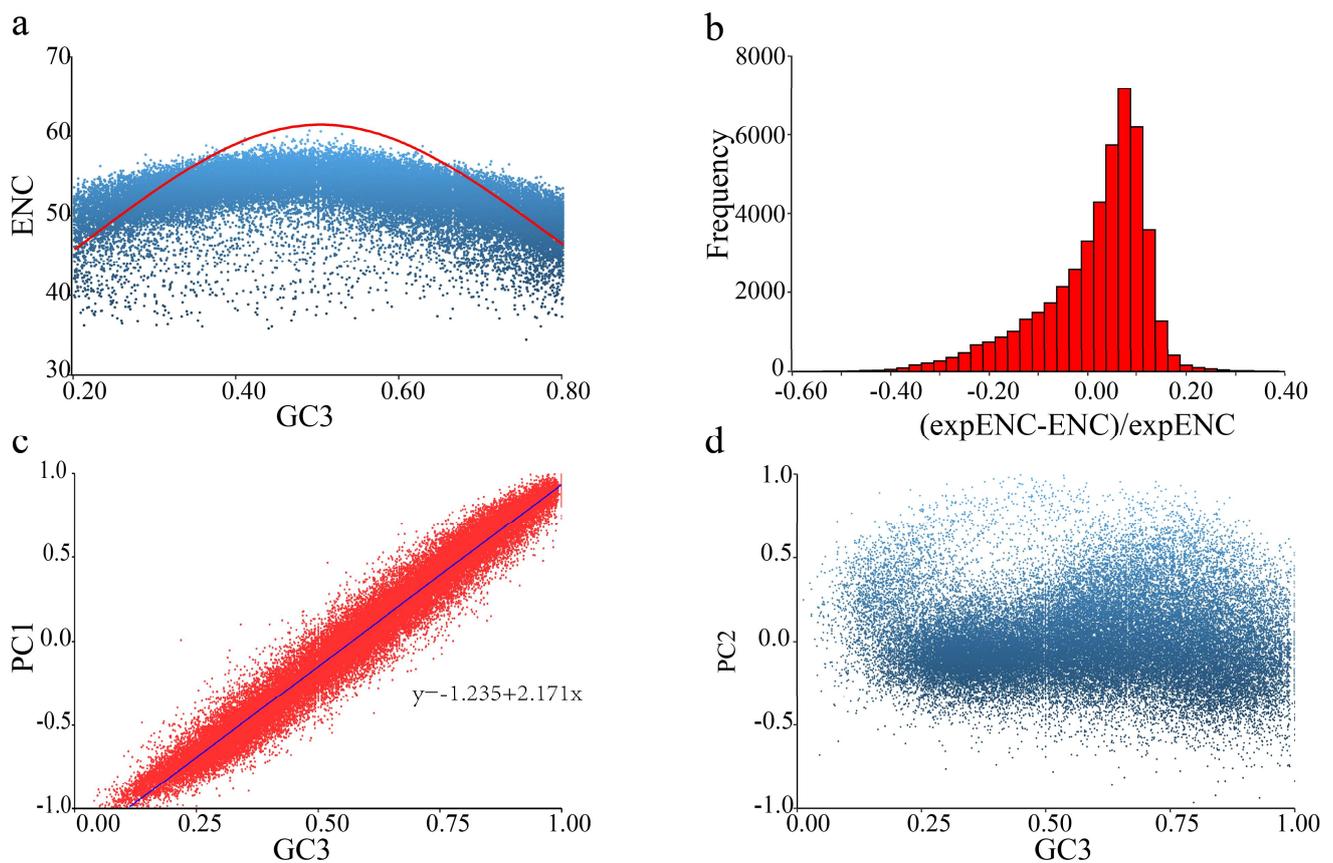


Figure 2. Codon usage bias of CDS. (a) The ENC-GC3 plot. GC3 is the GC content at the third codon position, and the solid line represents the expected curve when codon usage bias is only affected by mutation pressure. (b) Distribution of deviations from the expected ENC. (c) Correlation of PC1 and GC3. (d) Correlation of PC2 and GC3. PC1 and PC2 are the first two principal components of the codons' RSCU matrix.

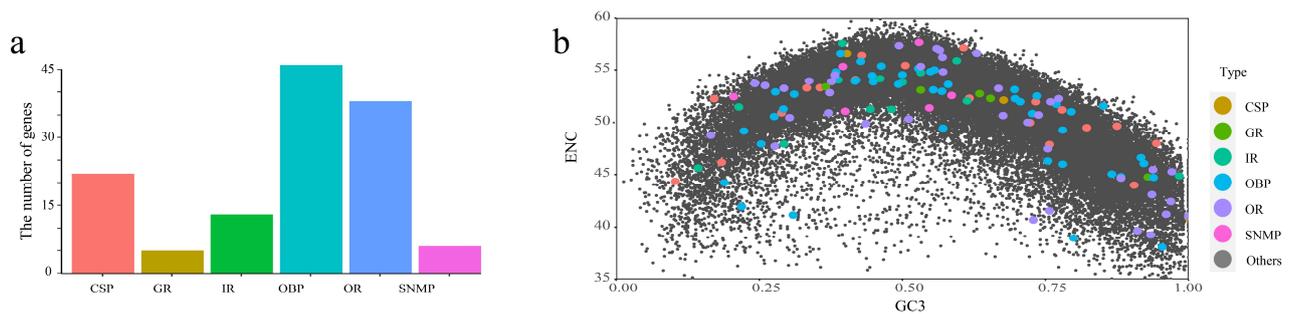


Figure 3. Codon usage bias of chemosensory genes in *T. meditationis*. (a) Number of identified chemosensory genes. Odorant-binding protein (OBP), chemosensory protein (CSP), odorant receptor (OR), ionotropic receptor (IR), sensory neuron membrane protein (SNMP), and gustatory receptor (GR). (b) The ENC-GC3 plot of chemosensory genes. Chemosensory genes are coloured according to (a). The gray points represent other CDS.

Blasting the transcripts against two other cave cricket transcriptomes, we obtained a similar profile in the other two species: 109 putative chemosensory genes in *Ceuthophilus sp.* (370 BPs, 23 CSPs, 32 ORs, 8 IRs, 6 SNMPs, and 3 GRs) and 105 putative genes in *Neonetus sp.* (410 BPs, 21 CSPs, 31 ORs, 4 IRs, 6 SNMPs, and 2 GRs).

We constructed a phylogenetic tree with the OR gene sequences of *T. meditationis*, two species of camel crickets (*Ceuthophilus sp.* and *Neonetus sp.*), and two species of locust

(*Ceracris kiangsu* and *Locusta migratoria* [28]; Figure 4). The migratory locust (Table 2) had the largest repertoire, which is unsurprising, as this species has an assembled whole-genome sequence. The results showed that most of the OR genes of *T. meditationis* and camel crickets clustered together, suggesting a contrast between the two suborders. There are several species-specific OR clusters for *T. meditationis*; however, this species does not have more clusters than the two cave crickets (Figure 4). We found similar results for five other gene families (Figures S1–S5).

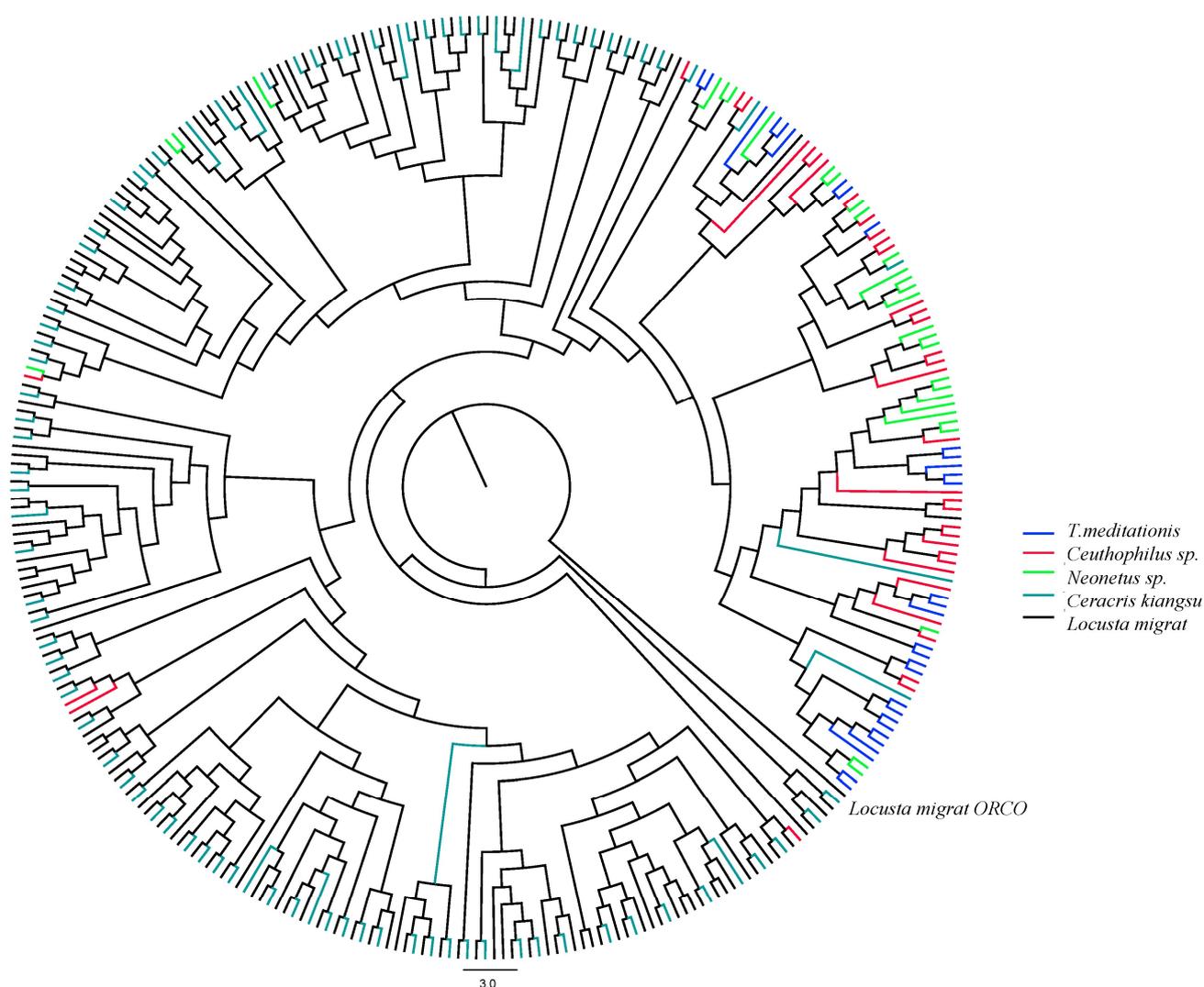


Figure 4. A maximum-likelihood phylogenetic tree of the OR gene family. Different colours represent different species of Orthoptera. The tree is rooted with the ORCO gene in *Locusta migratoria*.

Table 2. The number of chemosensory genes of four orthoptera insects.

Species	OR	GR	IR	CSP	OBP	SNMP
<i>Oedaleus asiaticus</i> [29]	60	-	6	-	15	3
<i>Ceracris nigricornis</i>	71	-	8	10	20	3
<i>Tachycines meditationis</i>	38	5	13	22	46	6
<i>Locusts migratoria</i> [28]	142	28	32	-	-	-

2.3. Genes under Positive Selection in *T. meditationis*

Including transcriptome datasets from two other species of cave crickets, OrthoFinder [30] identified 1186 single-copy orthologous genes across the three species. After translational

alignment, the dN/dS, dN, and dS values for each gene were calculated in PAML using the free-ratio model (M1), which allowed all species branches to have independent values. The distribution of dN/dS of *T. meditationis* is shown in Figure S6—most genes are less than 1. In the *T. meditationis* branch, we identified 124 genes with dN/dS > 1. We conducted KEGG enrichment analysis of these 124 genes and found only seven pathways. The most enriched pathways were the metabolism of xenobiotics by cytochrome P450, other types of O-glycan biosynthesis, retinol metabolism, ascorbate and aldarate metabolism, and drug metabolism-cytochrome P450 (Figure 5a).

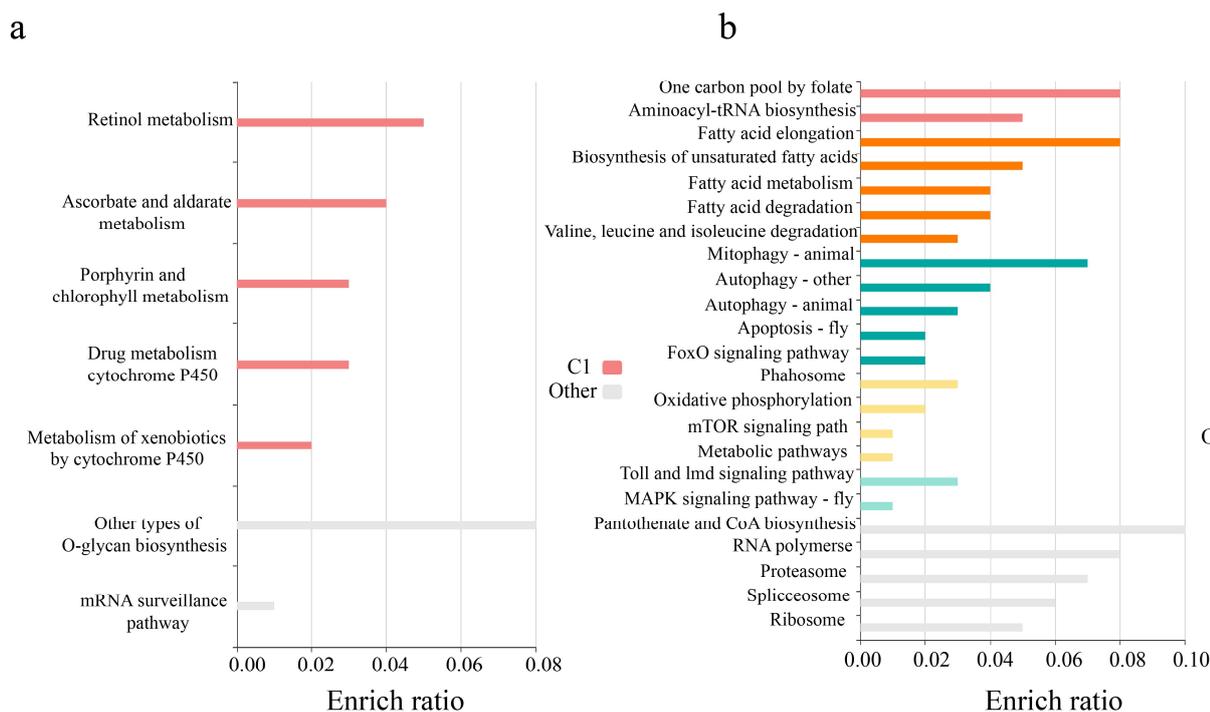


Figure 5. Results of KEGG enrichment analysis genes undergone species-specific positive selection in *T. meditationis*. Genes were identified by (a) the free-ratio model or (b) the branch-site model in PAML.

The free-ratio model assumes one dN/dS is shared across all sites in a gene. However, positive selection might only affect a few sites along particular lineages. Hence, we employed the branch-site models in PAML to scan for genes that might be affected by positive selection on a small proportion of sites on the *T. meditationis* branch. One hundred twenty-two genes significantly differ between Model A and the corresponding null model ($p < 0.05$). We further performed KEGG enrichment analysis with these genes. We found twenty-three related pathways, mainly concentrated in one carbon pool by folate, biosynthesis of unsaturated fatty acids, mitophagy-animal, pantothenate, and CoA biosynthesis (Figure 5b).

3. Discussion

T. meditationis is a widely distributed species of camel crickets in southeastern China. These nocturnal insects can reside outdoors in dark and damp environments, are often observed indoors, and feed on food debris. Hence, it earned the Chinese common name “Zhao Ma” (stove horse). Despite its abundance in urban environments, studies on this species are limited. Most studies have focused on the taxonomy of this group. In this study, we assembled and annotated a transcriptome dataset of *T. meditationis*, which is the first genome-scale dataset for this species. Although the transcriptome only provides information on expressed genes at the time of the experiment, this rich dataset can be used for various analyses. In this study, analyses of codon usage and chemosensory gene

families revealed unexpected results. Below, we discuss these results, possible caveats, and future research directions.

3.1. Weak Codon Usage in *T. meditationis*

Codon usage—the frequency of synonymous codons—has been the subject of molecular evolution for decades [31,32]. The match between codon frequencies and tRNA abundance in the cell has been well documented in microbial genomes [33] (e.g., *Escherichia coli* and *Saccharomyces cerevisiae*) since the 1980s. Advances in molecular engineering techniques have further enabled researchers to demonstrate that codon choices can affect the translational efficiency and accuracy of genes [34]. In addition to selection pressure from the translation process, mutation pressure, mutation biases or GC-biased gene conversion favouring G and C over A and T alleles, can also contribute to codon usage bias [35]. Transcriptome data provide the exact information needed to disentangle the evolutionary forces behind codon usage bias—the CDS for calculating codon frequencies and gene expression levels to assess the influence of translational selection.

We found that codon usage bias is weak in *T. meditationis*, in general. Most of the variation among the genes can be explained by the GC content at the third codon position, which is most likely due to mutation pressure. Eukaryotic genomes lacking a strong intensity of codon usage bias are common. Some have codon usage that is only weakly, or not at all, correlated with the expression level of genes or tRNA abundance (e.g., in humans [36]). Several hypotheses have been proposed to explain this variation among species [37]. Insight from population genetics has shown translational selection to be so weak that it becomes effective only in species with a large effective population size [38]. In species with a small effective population size, nonoptimal codon choices are only mildly deleterious and can still be fixed by genetic drift. Analyses across species reported higher bias in eukaryotes with shorter generation times, presumably species with large N_e (effective population size) [39] and more codons that have usage correlating with gene expression in large N_e animals [40].

Surprisingly, we found a very weak codon usage bias (Table 1) in *T. meditationis*, a commonly observed insect species with a short generation time and wide distribution range. In fact, excluding mutation pressure in examining only the conserved GC codon pairs, we found no preferred codon. One possible explanation is the decoupling of the effectiveness of this species and census population size. The population size (and range) of this species might have only recently expanded as humans provided a new habitat type; the low effective population size reflects the genetic diversity before the expansion. Another possibility is that *T. meditationis* harbours a deep population structure, which can also reduce the effective population size. According to the Orthoptera Species File, the taxonomy of the genus *Tachycines* has quickly evolved in the past few years—more than 30 new species have been reported since 2018 according to the Orthoptera Species File [41]. The distribution range of *T. meditationis*, its population dynamics, and possible cryptic species diversity would be worth examining with a phylogeographic approach [42].

3.2. Chemosensory Gene Evolution in Camel Crickets

Similar to other animals, chemoreception is critical for insect survival. Chemosensory genes are crucial for foraging, sensing food sources, and preventing poisoning. They also play a role in inter- and intra-species communication, such as mate choice and predator avoidance. The evolution of chemosensory genes is an excellent system to investigate how species adapt to new environments. Studies have found that chemosensory genes can respond rapidly to selection. Among the global lines of *Drosophila melanogaster*, chemosensory gene families show the strongest selection signals in genome-wide analyses [22]. Many studies have shown that chemosensory genes have played a crucial role in adapting to the environment [43,44]. For example, studies revealed a large-scale expansion of the chemosensory genes in German cockroaches [45], suggesting a vital role for chemosensory gene evolution in adapting to a wide range of climates and food sources during the global

expansion of this species [44]. Comparisons across sister species of orchid bees also detected elevated signals of divergent selection among chemosensory receptors [46]. In addition to sequence evolution, gene family sizes of chemosensory genes have been correlated with diet breadth. For example, host-specific beetle species have fewer chemosensory genes than polyphagous species do [47]. Comparisons of different species of mosquitos and *Drosophila* discovered a similar trend [48,49].

Given the omnivorous diet and presumably recent adaptation to the urban environment, we expected *T. meditationis* to have enlarged species-specific chemosensory gene repertoires, especially for OR and IR genes. However, all three camel cricket species in the family Rhaphidophoridae had similar repertoire sizes. While some species-specific gene clusters were present in the gene family tree, *T. meditationis* did not seem to have more gene duplication events than the other two species (Figure 4a). It should be noted that transcriptome data are limited by the specific developmental stage and tissue used for mRNA extraction. For example, the bamboo locust (*Ceracris kiangsu*) transcriptome data from the antenna has a much higher proportion of ORs (91 ORs, 13 IRs, 13 OBPs, 6 CSPs, and 2 SNMPs) [50]; hence, some OR genes might be missing from the transcriptome data (Table 2). A thorough comparison of gene family sizes among species requires whole-genome sequencing.

Here, we focused on the one feature of *T. meditationis*—its omnivorous diet. However, this species also has other relatively uncommon characteristics in the suborder Ensifera, such as apterous and lacking acoustic communications. Comparing its genomic data with other cricket species' genomes [51,52] could provide insights into how these creatures sense the environment and communicate among individuals.

4. Materials and Methods

4.1. *T. meditationis* Transcriptome Sequencing and Assembly

The two *T. meditationis* specimens used in this study were collected in Xian, Shaanxi, China and stored in liquid nitrogen until mRNA extraction. Using the NEBNext UltraTM DNA Library Prep Kit (New England Biolabs, Ipswich, MA, USA), two cDNA libraries (T1 and T2) were constructed from two *T. meditationis* individuals following the manufacturer's instructions and sequenced on an Illumina HiSeq 2500 high-throughput sequencing platform (Illumina, San Diego, CA, USA). Raw sequencing data were uploaded to the NCBI database (project number: PRJNA912169, SRA number: SRR22734937).

After quality filtering, reads from each library were assembled in Trinity 2.4.0 (<http://trinityrnaseq.github.io/>) (accessed on 21 July 2020) with default parameters [53]. The transdecoder function was used to extract transcripts with open reading frames (ORF) [54]. Using BLAST (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) (accessed on 28 July 2021), we compared the transcripts assembled from the two libraries and only retained those that exist in both libraries (i.e., having a blastx match in the other library with an E-value < 1×10^{-50}) [55].

4.2. CDS Identification and Annotation

Unigenes were annotated by blasting against the following five databases: NR (NCBI nonredundant protein sequences) [56], Pfam (Protein family) [57], COG/egg-NOG (clusters of orthologous groups of proteins) [58], KEGG (Kyoto Encyclopedia of Genes and Genomes) [59], and GO (Gene Ontology) [60]. The BLAST parameters *x*, *p*, and *r* were set with an E-value threshold of 1×10^{-5} . Unigenes with a BLAST match in at least one database were used for further analyses. We ran ORFfinder (<http://www.Geneinfinity.org/sms/sms>) (accessed on 5 August 2020) on these sequences to confirm the existence of the CDS. To ensure the quality of annotated genes, we deleted sequences shorter than 200 bp, containing more than 10% internal N gaps, or having more than one internal stop codons.

4.3. Codon Usage Analysis

Nucleotide composition of the whole transcriptome was determined using DAMBE <http://dambe.bio.uottawa.ca/DAMBE/dambe.aspx> (accessed on 10 August 2020) [61]. Compositions at each codon position and dinucleotide composition were calculated using CodonW. The correlation between GC12 and G3 was examined [62].

We calculated the RSCU value and ENC using the standard codon table in CodonW to quantify the codon usage bias for each CDS [63]. RSCU values show the exact frequency of each synonymous codon in the CDS compared to the expected frequency if all codons for an amino acid are used equally. Hence, they provide the most detailed information on codon usage for each CDS. The ENC reflects the extent to which codon usage deviates from random selection [64]. That is, each CDS has one ENC value and 61 RSCU values (64 codons, excluding the 3 stop codons).

4.4. Genome-Wide Codon Usage Bias

We identified which codons were overrepresented or underrepresented at the genome level by calculating the summary RSCU value of each codon. To examine the similarity between different codons, we conducted principal component analysis (PCA) on the codons' RSCU matrix, in which rows and columns correspond to codons and CDSs, respectively.

Next, we estimated the number of preferred codons defined in terms of gene expression—those used more frequently in genes with high expression than in genes with low expression [65]. We analysed GC-conservative pairs of synonymous codons to exclude the potential influence of GC-biased conversion [66], that is codon pairs that are XYA/XYT or XYC/XYG; there are 17 such pairs in the standard genetic code. GC-biased conversion did not affect the relative frequency of the two codons in a pair. We calculated the relative usage of 17 codon pairs for each gene and correlated it with gene expression. Following Galtier et al. (2017) [40], we defined preferred codons as those with an absolute correlation coefficient (r) > 0.05 and p -value < 0.001.

4.5. Variation of Codon Bias among Genes

We explored the variation in codon usage bias among CDS to investigate the evolutionary forces responsible for codon bias in *T. meditationis*. When mutation pressure is the only force acting on the third codon position, GC3 should determine the gene's level of codon bias [67]. In contrast, codon bias should correlate with the expression level with translational selection, as highly expressed genes should be under stronger selection for translation efficiency [63].

Hence, for ENC, we generated an ENC plot, which is a scatter plot showing the association between the ENC value and GC3 (G or C content in the third codon position). The expected ENC (expENC) based on GC3 was obtained by submitting 1000 GC values equally spaced between 0.001 and 1.000 to DAMBE [68]. The smooth curve connecting these expected values was compared to the observed ENC from the empirical CDS. We then tested the correlation between ENC and gene product properties. In addition to the gene expression level, we obtained the amino acid composition, aromaticity (AROMO) value, and hydrophobicity (GRAVY) value using CodonW (<https://sourceforge.net/projects/codonw>) (accessed on 12 August 2020) for each CDS. The GRAVY and AROMO values reflect the percentage of hydrophobic and aromatic amino acids in proteins that affect protein folding. We used principal components (PCs) to represent high-dimensional amino acid composition data.

Similar analyses were conducted for the RSCU values. We performed PCA to reduce the dimension of the data (each gene had 61 RSCU values) and correlated the first two PCs with gene properties (e.g., GC3, expression level, and GRAVY).

4.6. Chemosensory Gene Analysis

Based on annotation information, we identified genes belonging to six chemosensory gene families. We examined the codon bias of these genes. To study gene family

evolution, we downloaded the chemosensory genes identified in bamboo grasshoppers (*Ceracris kiangsu*; accession number: SRR1648029). We also searched NCBI for additional transcriptome datasets for the Rhaphidophoridae (cave crickets) family and identified two datasets: *Neonetus sp.* AD-2015 (cave weta; accession number: SRR2230582) and *Ceuthophilus sp.* AD-2013 (another camel cricket species; accession number: SRR921579). Unigenes were assembled and identified using the same approach as that used for our dataset. Chemosensory genes in these two cave cricket datasets were identified by blasting the chemosensory genes in our dataset (E-value < 1×10^{-10}).

Sequences of each gene family were pooled, and sequence alignments were obtained for each gene family using Mafft [69] with the default setting. Phylogenetic trees were constructed using IQtree [70] based on the GTR + F model and visualised in FigTree (v1.4.3) (<http://tree.bio.ed.ac.uk/software/figtree/> (accessed on 1 August 2020)). The phylogenetic tree was rooted using the default IQtree parameters.

4.7. dN/dS Analysis to Identify Genes under Positive Selection

Orthologous gene sequences from closely related species can help identify genes that have experienced species-specific positive selection during evolution. Orthofinder [30] with default parameters was used to identify orthologous genes between *T. meditationis* and the other two cricket species. Alignments of single-copy orthologous genes from three species were generated by Mafft [69], and bmge (g: 0.5, the value of g indicates the maximum proportion of gap and columns; genes exceeding this proportion are deleted) were used to prune the result of multiple sequence alignment [71].

The *codeml* program in PAML [72] was used to test for positive selection. First, we fitted the free-ratio model (M1) to calculate the branch-specific ω values for each species. Genes with $\omega > 1$ for *T. meditationis* were selected for further gene set enrichment analysis using KOBAS [73]. Since positive selection might only affect a few sites along particular lineages, we also applied a second test, using branch-site models that allow ω to vary between sites in the gene and across branches on the phylogeny [74]. Specifically, we compared two ML models in the second test: (1) a null model assuming two site classes, one with dN/dS < 1 and one with dN/dS = 1 (model = 2, NSsites = 2, fix_omega = 1, omega = 1), and (2) an alternative model with one site class having dN/dS > 1 (model = 2, NSsites = 2, fix_omega = 0). The branch leading to *T. meditationis* was selected as the foreground, whereas the others were kept as the background. The fit of the two models was compared using a likelihood ratio test with one degree of freedom, and a gene set enrichment analysis was conducted with significant genes.

5. Conclusions

This study characterised a de novo transcriptome of *T. meditationis*, an insect commonly observed in urban environments. Contrary to expectations grounded in the ecology and diet of this species, we did not find evidence for intense selection of codon usage or enlargement of the chemosensory gene repertoires. However, in combination with the transcriptomes of other species of the same family, we identified genes that might undergo species-specific positive selection and enrich biological pathways. Our dataset is a valuable molecular resource for further improving our understanding of the evolution of camel crickets, and our results will facilitate future studies on how insects' diet adapts to urban environments in general.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms24044005/s1>.

Author Contributions: S.-Q.X. and H.H. conceived and designed the research. J.-H.L. analyzed the transcriptome data and wrote the original draft. D.-L.G. helped with the sampling and guidance. H.H. revised the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This work was financially supported by the National Natural Science Foundation of China (No. 31872273) and the Fundamental Research Funds for the Central Universities (No. GK201903063, GK202105003).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: RNA-Seq data can be found with accession number PRJNA912169. The RNA-Seq data is publicly available on National Center for Biotechnology Information.

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

Abbreviations

AROMO	Aromaticity
CDS	Coding sequences
COG/egg-NOG	Clusters of orthologous groups of proteins
ENC	Effective number of codons
FPKM	Fragments per kilobase million
GC1	GC content at the first codon positions
GC2	GC content at the second codon positions
GC3	GC content at the third codon positions
GO	Gene Ontology
GRAVY	General average of hydrophobicity
KEGG	Kyoto Encyclopedia of Genes and Genomes
Nr	NCBI non-redundant protein sequences
PCA	Principal component analysis
Pfam	Protein family
RSCU	Relative synonymous codon usage
OBP	Odorant-binding protein
CSP	Chemosensory protein
OR	Odorant receptor
IR	Ionotropic receptor
SNMP	Sensory neuron membrane protein
GR	Gustatory receptor

References

1. Qin, Y.; Wang, H.; Liu, X.; Li, K. Divided the genus Tachycines Adelung (Orthoptera, Rhaphidophoridae: Aemodogryllinae; Aemodogryllini) from China. *Zootaxa* **2018**, *4374*, 451–475. [[CrossRef](#)]
2. Isaac, C.; Orue, P.O.; Iyamu, M.I.; Ehiaghe, J.I.; Isaac, O. Comparative Analysis of Pathogenic Organisms in Cockroaches from Different Community Settings in Edo State, Nigeria. *Korean J. Parasitol.* **2014**, *52*, 177–181. [[CrossRef](#)]
3. Qin, Y.; Liu, X.; Li, K. Review of the subgenus Tachycines (Gymnaeta) Adelung, 1902 (Orthoptera, Rhaphidophoridae, Aemodogryllinae, Aemodogryllini). *Zootaxa* **2019**, *4560*, 273–310. [[CrossRef](#)]
4. Yang, Y.; Li, Y.; Chen, Q.; Sun, Y.; Lu, Z. WGDdetector: A pipeline for detecting whole genome duplication events using the genome or transcriptome annotations. *BMC Bioinform.* **2019**, *20*, 1–6. [[CrossRef](#)]
5. Majewska, M.; Lipka, A.; Pauksztó, L.; Jastrzebski, J.P.; Myszczyński, K.; Gowkielewicz, M.; Jozwik, M.; Majewski, M.K. Transcriptome profile of the human placenta. *Funct. Integr. Genom.* **2017**, *17*, 551–563. [[CrossRef](#)]
6. Morey, J.S.; Huntington, K.A.B.; Campbell, M.; Clauss, T.M.; Goertz, C.E.; Hobbs, R.C.; Lunardi, D.; Moors, A.J.; Neely, M.G.; Schwacke, L.H.; et al. De novo transcriptome assembly and RNA-Seq expression analysis in blood from beluga whales of Bristol Bay, AK. *Mar. Genom.* **2017**, *35*, 77–92. [[CrossRef](#)]
7. Xue, P.; Zhao, X.; Qin, M.; Shi, Z.; Zhang, M.; Gu, W. Transcriptome Analysis of Male Drosophila melanogaster Exposed to Ethylparaben Using Digital Gene Expression Profiling. *J. Insect Sci.* **2017**, *17*, 87. [[CrossRef](#)]
8. 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](#)]

9. Bahiri-Elitzur, S.; Tuller, T. Codon-based indices for modeling gene expression and transcript evolution. *Comput. Struct. Biotechnol. J.* **2021**, *19*, 2646–2663. [[CrossRef](#)]
10. Rocha, E.P. Codon usage bias from tRNA's point of view: Redundancy, specialization, and efficient decoding for translation optimization. *Genome Res.* **2004**, *14*, 2279–2286. [[CrossRef](#)]
11. Eyre-Walker, A. An analysis of codon usage in mammals: Selection or mutation bias? *J. Mol. Evol.* **1991**, *33*, 442–449. [[CrossRef](#)]
12. Komar, A.A. The Yin and Yang of codon usage. *Hum. Mol. Genet.* **2016**, *25*, R77–R85. [[CrossRef](#)]
13. Ikemura, T. Codon usage and tRNA content in unicellular and multicellular organisms. *Mol. Biol. Evol.* **1985**, *2*, 13–34. [[CrossRef](#)]
14. Lloyd, A.T.; Sharp, P. CODONS: A Microcomputer Program for Codon Usage Analysis. *J. Hered.* **1992**, *83*, 239–240. [[CrossRef](#)]
15. Li, J.; Zhang, Z.; Vang, S.; Yu, J.; Wong, G.K.-S.; Wang, J. Correlation Between Ka/Ks and Ks is Related to Substitution Model and Evolutionary Lineage. *J. Mol. Evol.* **2009**, *68*, 414–423. [[CrossRef](#)]
16. Drummond, D.A.; Wilke, C.O. Mistranslation-Induced Protein Misfolding as a Dominant Constraint on Coding-Sequence Evolution. *Cell* **2008**, *134*, 341–352. [[CrossRef](#)]
17. Cole, N.C.; Wang, A.A.; Donovan, S.M.; Lee, S.-Y.; Teran-Garcia, M. Strong Kids the STRONG Kids Team Variants in Chemosensory Genes Are Associated with Picky Eating Behavior in Preschool-Age Children. *Lifestyle Genom.* **2017**, *10*, 84–92. [[CrossRef](#)]
18. Sun, H.; Bu, L.-A.; Su, S.-C.; Guo, D.; Gao, C.-F.; Wu, S.-F. Knockout of the odorant receptor co-receptor, orco, impairs feeding, mating and egg-laying behavior in the fall armyworm Spodoptera frugiperda. *Insect Biochem. Mol. Biol.* **2023**, *152*, 103889. [[CrossRef](#)]
19. Llopis-Giménez, A.; Carrasco-Oltra, T.; Jacquín-Joly, E.; Herrero, S.; Crava, C.M. Coupling Transcriptomics and Behaviour to Unveil the Olfactory System of Spodoptera exigua Larvae. *J. Chem. Ecol.* **2020**, *46*, 1017–1031. [[CrossRef](#)]
20. Wang, X.; Kang, L. Molecular Mechanisms of Phase Change in Locusts. *Annu. Rev. Entomol.* **2014**, *59*, 225–244. [[CrossRef](#)]
21. Benton, R. Multigene Family Evolution: Perspectives from Insect Chemoreceptors. *Trends Ecol. Evol.* **2015**, *30*, 590–600. [[CrossRef](#)]
22. Arguello, J.R.; Cardoso-Moreira, M.; Grenier, J.K.; Gottipati, S.; Clark, A.G.; Benton, R. Extensive local adaptation within the chemosensory system following *Drosophila melanogaster*'s global expansion. *Nat. Commun.* **2016**, *7*, 1–12. [[CrossRef](#)]
23. Nei, M.; Niimura, Y.; Nozawa, M. The evolution of animal chemosensory receptor gene repertoires: Roles of chance and necessity. *Nat. Rev. Genet.* **2008**, *9*, 951–963. [[CrossRef](#)]
24. Mombaerts, P. Seven-Transmembrane Proteins as Odorant and Chemosensory Receptors. *Science* **1999**, *286*, 707–711. [[CrossRef](#)]
25. Monahan, K.; Lomvardas, S. Monoallelic Expression of Olfactory Receptors. *Annu. Rev. Cell Dev. Biol.* **2015**, *31*, 721–740. [[CrossRef](#)]
26. Wada-Katsumata, A.; Robertson, H.M.; Silverman, J.; Schal, C. Changes in the Peripheral Chemosensory System Drive Adaptive Shifts in Food Preferences in Insects. *Front. Cell. Neurosci.* **2018**, *12*, 281. [[CrossRef](#)]
27. Robertson, H.M.; Baits, R.L.; Walden, K.K.; Wada-Katsumata, A.; Schal, C. Enormous expansion of the chemosensory gene repertoire in the omnivorous German cockroach *Blattella germanica*. *J. Exp. Zool. Part B Mol. Dev. Evol.* **2018**, *330*, 265–278. [[CrossRef](#)]
28. Wang, X.; Fang, X.; Yang, P.; Jiang, X.; Jiang, F.; Zhao, D.; Li, B.; Cui, F.; Wei, J.; Ma, C.; et al. The locust genome provides insight into swarm formation and long-distance flight. *Nat. Commun.* **2014**, *5*, 2957. [[CrossRef](#)]
29. Zhou, Y.-T.; Li, L.; Zhou, X.-R.; Tan, Y.; Pang, B.-P. Identification and expression profiling of candidate chemosensory membrane proteins in the band-winged grasshopper, *Oedaleus asiaticus*. *Comp. Biochem. Physiol. Part D Genom. Proteom.* **2019**, *30*, 33–44. [[CrossRef](#)]
30. Emms, D.M.; Kelly, S. OrthoFinder: Solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol.* **2015**, *16*, 157. [[CrossRef](#)]
31. Fages-Lartaud, M.; Hundvin, K.; Hohmann-Marriott, M.F. "Mechanisms governing codon usage bias and the implications for protein expression in the chloroplast of *Chlamydomonas reinhardtii*". *Plant J.* **2022**, *112*, 919–945. [[CrossRef](#)] [[PubMed](#)]
32. Dedon, P.C.; Begley, T.J. Dysfunctional tRNA reprogramming and codon-biased translation in cancer. *Trends Mol. Med.* **2022**. [[CrossRef](#)] [[PubMed](#)]
33. Lyu, X.; Yang, Q.; Li, L.; Dang, Y.; Zhou, Z.; Chen, S.; Liu, Y. Adaptation of codon usage to tRNA I34 modification controls translation kinetics and proteome landscape. *PLOS Genet.* **2020**, *16*, e1008836. [[CrossRef](#)] [[PubMed](#)]
34. Wan, X.-F.; Xu, N.; Kleinhofs, A.; Zhou, J. Quantitative relationship between synonymous codon usage bias and GC composition across unicellular genomes. *BMC Evol. Biol.* **2004**, *4*, 19. [[CrossRef](#)] [[PubMed](#)]
35. Wint, R.; Salamov, A.; Grigoriev, I.V. Kingdom-Wide Analysis of Fungal Protein-Coding and tRNA Genes Reveals Conserved Patterns of Adaptive Evolution. *Mol. Biol. Evol.* **2022**, *39*, msab372. [[CrossRef](#)] [[PubMed](#)]
36. Piovesan, A.; Vitale, L.; Pelleri, M.C.; Strippoli, P. Universal tight correlation of codon bias and pool of RNA codons (codonome): The genome is optimized to allow any distribution of gene expression values in the transcriptome from bacteria to humans. *Genomics* **2013**, *101*, 282–289. [[CrossRef](#)] [[PubMed](#)]
37. Liu, Y.; Yang, Q.; Zhao, F. Synonymous but Not Silent: The Codon Usage Code for Gene Expression and Protein Folding. *Annu. Rev. Biochem.* **2021**, *90*, 375–401. [[CrossRef](#)]
38. Sharp, P.M.; Li, W.-H. An evolutionary perspective on synonymous codon usage in unicellular organisms. *J. Mol. Evol.* **1986**, *24*, 28–38. [[CrossRef](#)]
39. Subramanian, S. Nearly Neutrality and the Evolution of Codon Usage Bias in Eukaryotic Genomes. *Genetics* **2008**, *178*, 2429–2432. [[CrossRef](#)]

40. Galtier, N.; Roux, C.; Rousselle, M.; Romiguier, J.; Figuet, E.; Glémin, S.; Bierne, N.; Duret, L. Codon Usage Bias in Animals: Disentangling the Effects of Natural Selection, Effective Population Size, and GC-Biased Gene Conversion. *Mol. Biol. Evol.* **2018**, *35*, 1092–1103. [[CrossRef](#)]
41. Zhu, Q.; Chen, H.; Shi, F. Remarks on the genus *Tachycines* Adelung, 1902 (Orthoptera: Rhaphidophoridae: Aemodogryllinae) with description of eight new species from caves in southern China. *Zootaxa* **2020**, *4809*, 71–94. [[CrossRef](#)]
42. Zhu, Q.; Zhou, Z.; Zheng, X.; Wang, T.; Ma, L.; Shi, F. Phylogeny and phylogeography of *Diestramima* cave crickets (Orthoptera: Rhaphidophoridae): Speciation driven by multiple dispersal and vicariance events. *Syst. Entomol.* **2022**, *47*, 179–201. [[CrossRef](#)]
43. Zhou, Y.-T.; Li, L.; Zhou, X.-R.; Tan, Y.; Pang, B.-P. Three Chemosensory Proteins Involved in Chemoreception of *Oedaleus asiaticus* (Orthoptera: Acridoidea). *J. Chem. Ecol.* **2019**, *46*, 138–149. [[CrossRef](#)]
44. Argüello, D.; Chavez, E.; Lauryssen, F.; Vanderschueren, R.; Smolders, E.; Montalvo, D. Soil properties and agronomic factors affecting cadmium concentrations in cacao beans: A nationwide survey in Ecuador. *Sci. Total. Environ.* **2019**, *649*, 120–127. [[CrossRef](#)]
45. He, M.; Ma, Y.-F.; Guo, H.; Liu, X.-Z.; Long, G.-J.; Wang, Q.; Dewey, Y.; Zhang, F.; He, P. Genome-wide identification and expression pattern analysis of novel chemosensory genes in the German cockroach *Blattella germanica*. *Genomics* **2022**, *114*, 110310. [[CrossRef](#)]
46. Brand, P.; Ramírez, S.R.; Leese, F.; Quezada-Euan, J.J.G.; Tollrian, R.; Eltz, T. Rapid evolution of chemosensory receptor genes in a pair of sibling species of orchid bees (Apidae: Euglossini). *BMC Evol. Biol.* **2015**, *15*, 1–16. [[CrossRef](#)]
47. Andersson, M.N.; Keeling, C.I.; Mitchell, R.F. Genomic content of chemosensory genes correlates with host range in wood-boring beetles (*Dendroctonus ponderosae*, *Agrius planipennis*, and *Anoplophora glabripennis*). *BMC Genom.* **2019**, *20*, 1–18. [[CrossRef](#)]
48. Arensburger, P.; Megy, K.; Waterhouse, R.M.; Abrudan, J.; Amedeo, P.; Antelo, B.; Bartholomay, L.; Bidwell, S.; Caler, E.; Camara, F.; et al. Sequencing of *Culex quinquefasciatus* Establishes a Platform for Mosquito Comparative Genomics. *Science* **2010**, *330*, 86–88. [[CrossRef](#)]
49. McBride, C.S. Rapid evolution of smell and taste receptor genes during host specialization in *Drosophila sechellia*. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 4996–5001. [[CrossRef](#)]
50. Li, R.; Wang, Y.-T.; Jiang, G.-F. The transcriptome analysis of the bamboo grasshopper provides insights into hypothermic stress acclimation. *Int. J. Biol. Macromol.* **2019**, *134*, 237–246. [[CrossRef](#)]
51. Sanno, R.; Kataoka, K.; Hayakawa, S.; Ide, K.; Nguyen, C.N.; Nguyen, T.P.; Le, B.T.N.; Kim, O.T.P.; Mineta, K.; Takeyama, H.; et al. Comparative Analysis of Mitochondrial Genomes in Gryllidea (Insecta: Orthoptera): Implications for Adaptive Evolution in Ant-Loving Crickets. *Genome Biol. Evol.* **2021**, *13*, evab222. [[CrossRef](#)] [[PubMed](#)]
52. Kataoka, K.; Minei, R.; Ide, K.; Ogura, A.; Takeyama, H.; Takeda, M.; Suzuki, T.; Yura, K.; Asahi, T. The Draft Genome Dataset of the Asian Cricket *Teleogryllus occipitalis* for Molecular Research Toward Entomophagy. *Front. Genet.* **2020**, *11*, 470. [[CrossRef](#)]
53. Haas, B.J.; Papanicolaou, A.; Yassour, M.; Grabherr, M.; Blood, P.D.; Bowden, J.; Couger, M.B.; Eccles, D.; Li, B.; Lieber, M.; et al. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Protoc.* **2013**, *8*, 1494–1512. [[CrossRef](#)] [[PubMed](#)]
54. Rombel, I.T.; Sykes, K.F.; Rayner, S.; Johnston, S.A. ORF-FINDER: A vector for high-throughput gene identification. *Gene* **2002**, *282*, 33–41. [[CrossRef](#)] [[PubMed](#)]
55. Camacho, C.; Coulouris, G.; Avagyan, V.; Ma, N.; Papadopoulos, J.; Bealer, K.; Madden, T.L. BLAST+: Architecture and applications. *BMC Bioinform.* **2009**, *10*, 421. [[CrossRef](#)]
56. Réau, M.; Lagarde, N.; Zagury, J.-F.; Montes, M. Nuclear Receptors Database Including Negative Data (NR-DBIND): A Database Dedicated to Nuclear Receptors Binding Data Including Negative Data and Pharmacological Profile. *J. Med. Chem.* **2018**, *62*, 2894–2904. [[CrossRef](#)]
57. Finn, R.D.; Bateman, A.; Clements, J.; Coggill, P.; Eberhardt, R.Y.; Eddy, S.R.; Heger, A.; Hetherington, K.; Holm, L.; Mistry, J.; et al. Pfam: The protein families database. *Nucleic Acids Res.* **2013**, *42*, D222–D230. [[CrossRef](#)]
58. Tatusov, R.L.; Galperin, M.Y.; Natale, D.A.; Koonin, E.V. The COG database: A tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res.* **2000**, *28*, 33–36. [[CrossRef](#)]
59. Kanehisa, M.; Goto, S.; Kawashima, S.; Okuno, Y.; Hattori, M. The KEGG resource for deciphering the genome. *Nucleic Acids Res.* **2004**, *32*, D277–D280. [[CrossRef](#)]
60. Ashburner, M.; Ball, C.A.; Blake, J.A.; Botstein, D.; Butler, H.; Cherry, J.M.; Davis, A.P.; Dolinski, K.; Dwight, S.S.; Eppig, J.T.; et al. Gene ontology: Tool for the unification of biology. *Nat. Genet.* **2000**, *25*, 25–29. [[CrossRef](#)]
61. Xia, X. DAMBE5: A Comprehensive Software Package for Data Analysis in Molecular Biology and Evolution. *Mol. Biol. Evol.* **2013**, *30*, 1720–1728. [[CrossRef](#)] [[PubMed](#)]
62. Parvathy, S.T.; Udayasuriyan, V.; Bhadana, V. Codon usage bias. *Mol. Biol. Rep.* **2021**, *49*, 539–565. [[CrossRef](#)] [[PubMed](#)]
63. Gupta, S.K.; Bhattacharyya, T.K.; Ghosh, T.C. Synonymous Codon Usage in *Lactococcus lactis*: Mutational Bias Versus Translational Selection. *J. Biomol. Struct. Dyn.* **2004**, *21*, 527–535. [[CrossRef](#)] [[PubMed](#)]
64. Wright, F. The ‘effective number of codons’ used in a gene. *Gene* **1990**, *87*, 23–29. [[CrossRef](#)]
65. Khelifi, A.; Meunier, J.; Duret, L.; Mouchiroud, D. GC Content Evolution of the Human and Mouse Genomes: Insights from the Study of Processed Pseudogenes in Regions of Different Recombination Rates. *J. Mol. Evol.* **2006**, *62*, 745–752. [[CrossRef](#)]
66. Chakraborty, S.; Nag, D.; Mazumder, T.H.; Uddin, A. Codon usage pattern and prediction of gene expression level in Bungarus species. *Gene* **2017**, *604*, 48–60. [[CrossRef](#)]

67. Duret, L. Evolution of synonymous codon usage in metazoans. *Curr. Opin. Genet. Dev.* **2002**, *12*, 640–649. [[CrossRef](#)]
68. Xia, X.; Xie, Z. DAMBE: Software Package for Data Analysis in Molecular Biology and Evolution. *J. Hered.* **2001**, *92*, 371–373. [[CrossRef](#)]
69. Kazutaka, K.; Misakwa, K.; Kei-ichi, K.; Miyata, T. MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* **2002**, *30*, 3059–3066. [[CrossRef](#)]
70. Nguyen, L.-T.; Schmidt, H.A.; Von Haeseler, A.; Minh, B.Q. IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. *Mol. Biol. Evol.* **2015**, *32*, 268–274. [[CrossRef](#)]
71. Criscuolo, A.; Gribaldo, S. BMGE (Block Mapping and Gathering with Entropy): A new software for selection of phylogenetic informative regions from multiple sequence alignments. *BMC Evol. Biol.* **2010**, *10*, 210. [[CrossRef](#)] [[PubMed](#)]
72. Yang, Z. PAML 4: Phylogenetic Analysis by Maximum Likelihood. *Mol. Biol. Evol.* **2007**, *24*, 1586–1591. [[CrossRef](#)] [[PubMed](#)]
73. Bu, D.; Luo, H.; Huo, P.; Wang, Z.; Zhang, S.; He, Z.; Wu, Y.; Zhao, L.; Liu, J.; Guo, J.; et al. KOBAS-i: Intelligent prioritization and exploratory visualization of biological functions for gene enrichment analysis. *Nucleic Acids Res.* **2021**, *49*, W317–W325. [[CrossRef](#)] [[PubMed](#)]
74. Wang, C.; Scharf, M.E.; Bennett, G.W. Behavioral and Physiological Resistance of the German Cockroach to Gel Baits (Blattodea: Blattellidae). *J. Econ. Entomol.* **2004**, *97*, 2067–2072. [[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.