Supplemental Material

Rapid evolution of ovarian-biased genes in the yellow-fever mosquito (*Aedes aegypti*)

Carrie A. Whittle[1], Cassandra G. Extavour[1,2*]

[1] Department of Organismic and Evolutionary Biology, Harvard University, 16 Divinity Avenue, Cambridge MA 02138, USA

[2] Department of Molecular and Cellular Biology, Harvard University, 16 Divinity Avenue, Cambridge MA 02138, USA

Supplemental Note 1: Identification of Optimal Codons

In order to study sex-biased expression on codon usage, we first had to determine whether *A. aegypti* had optimal codons that were shaped by selection. For this, we studied embryos, which comprise highly complex non-sexual structures, that have diverse gene expression profiles (e.g. DIEZ-ROUX *et al.* 2011; MITTMANN AND WOLFF 2012; COMBS AND EISEN 2013; SCHWAGER *et al.* 2015; DONOUGHE AND EXTAVOUR 2016), and have proven effective for the identification and study of the molecular evolution of highly versus lowly expressed genes in a taxon (SUBRAMANIAN AND KUMAR 2004), including for the identification of optimal codons (WHITTLE AND EXTAVOUR 2016). After mapping RNA-seq reads to the CDS list of *A. aegypti*, we identified 12,837 CDS expressed in *A. aegypti* embryos. Using these CDS, we determined the relative synonymous codon usage (RSCU; Sharp et al., 1986) of the 5% highest and 5% lowest expressed CDS in embryos, and calculated $\Delta RSCU = RSCU_{\text{Mean 5\% Highest Expressed CDS}} - RSCU_{\text{Mean 5\% Lowest Expressed}}$ for each codon per amino acid, a method with precedent for identification of codons preferred under high expression (DURET AND MOUCHIROUD 1999; CUTTER *et al.* 2006; INGVARSSON 2008; WANG *et al.* 2011; WHITTLE *et al.* 2011; WHITTLE AND EXTAVOUR 2015). The codon with the largest, statistically significant and positive ΔRSCU per amino acid was taken as the primary optimal codon per amino acid. The results are shown in Table 1. Based on ΔRSCU, we identified 18 primary optimal codons in *A. aegypti*; fifteen optimal codons ended in C and three in G. For further stringency in the identification of the optimal codon list, we conducted correspondence analysis for the *A. aegypti* CDS to predict optimal codons, as described previously (WHITTLE AND EXTAVOUR 2016). This method identifies putative optimal codons based on their distribution across the principle axis (see Methods), which needs to be verified using expression data (PEDEN 1999), as per Table 1. The predictions

from correspondence analysis agreed with our results from ΔRSCU, where these same 18 codons were identified as optimal. Thus, both comparative expression analyses and correspondence analysis support a model of expression-mediated selection for optimal codon usage in highly expressed CDS in this mosquito, which spans each of the 18 amino acids with synonymous codons.

We found that ΔRSCU tended to vary with degeneracy level of amino acids. For instance, the largest values were for six-fold degenerate amino acids; Arg. Leu and Ser had values of +0.670, +0.786, and +0.575 respectively. In contrast, four and three fold amino acids including Ala, Gly, Ile, Pro, Thr and Val had ΔRSCU between +0.294 and 0.579 , whilst the lowest values were typically (with one exception, Lys) observed for two-fold amino acids with Asn, Asp, Cys, Gln, Glu, His, Lys, Phe, and Tyr, having values of +0.227, +0.089, +0.088, +0.238, +0.060, +0.086, +0.309, +0.231, +0.149 respectively. This pattern suggests that selection on codon usage is greater for amino acids with higher degeneracy, and is notably consistent with results for certain other organisms such as spiders (*Parasteatoda tepidariorum*), fruit flies (*Drosophila melanogaster*), crickets (*Gryllus bimaculatus*), milkweed bugs (*Oncopeltus fasciatus*), amphipod crustaceans (*Parhyale hawaiensis*) and roundworms (*Caenorhabditis elegans*) (McVean and Vieira 1999; Cutter *et al.* 2006; Whittle and Extavour 2015; Whittle and Extavour 2016). Collectively, the results across these invertebrates suggest that expression-mediated selection is stronger in more degenerate amino acids, a trend that can be speculated to result from the requirement for a greater number of alternate tRNAs to encode more degenerate amino acids and competition within cells during translation (Whittle and Extavour 2016). Two-fold amino acids can usually be encoded by a single tRNA due to degeneracy (wobble rules) at the

third codon site (IKEMURA 1985; PERCUDANI 2001), thus limiting putative competition (WHITTLE AND EXTAVOUR 2016).

Codon Usage is Shaped by Selection

For further stringency in our conclusion that codon usage is shaped by selection in Table 1, we assessed whether the differences in codon usage between high and low expressed genes could be explained by mutational biases. For this, we compared GC content of synonymous 3rd codon positions (GC3s) and GC content of introns. We identified 36,272 introns for this study, with the criteria that they be >20bp and contained within the genes expressed in embryos. Evidence suggests that short introns in a genome are nearly entirely selectively neutral, while longer introns may be subject to selection, particularly in regulatory flanking regions (FARLOW *et al.* 2012). Thus, we divided the *A. aegypti* introns into short and long classes divided by the median (130bp). For long introns, we removed the 50 bp from each end, as these sequences are most likely to contain regulatory regions (WILLIFORD AND DEMUTH 2012). The results showed differences in GC content of short versus long introns (differences in mean GC < 0.03, MWU-test P<0.001; Figure 1), and thus we included both intron classes for comparison to the GC3s of CDS (exons). We report that the GC3s for the CDS with the highest 5% expression in embryos (Average GC3s=0.618±0.005), intermediate 5th to 95th expression level (0.564±0.001), and the CDS with lowest 5% expression (GC3s=0.536±0.003), were each markedly and statistically significantly higher than the GCi content of their associated introns (for both short and long introns), which averaged between 0.340 and 0.380 (Figure 1; MWU-test P <0.001 for all contrasts). This result is inconsistent with mutational bias shaping optimal codon usage. Further, unlike GC3s, which increased with expression, GCi was not elevated with respect to transcription level (and slightly decreased for short introns, Figure 1), also pointing against a

mutational-bias-expression level relationship. Thus, we conclude that mutational biases do not underlie the preferred usage of GC3 codons with high expression (Table 1), and rather that this usage is best explained by expression-linked selection. Accordingly, since selection for optimal codon usage exists in this taxon, it is feasible to assess the link between sex-biased gonad expression in *A. aegypti* and selection on codon usage.

Supplemental Note 2: Sex-Biased Effects on dN/dS and Fop Depends on Fold-Bias, CDS Length, and Expression Level

To evaluate in detail how sex-bias influences gene evolution in *A. aegypti*, we examined dN, dS, dN/dS and Fop for each combination of the three classes of fold-bias and the two classes of CDS lengths defined in the main text. As shown in Table S2, long ovary-biased CDS exhibited statistically significantly higher dN and dN/dS than testis-biased CDS for those with ≥2-5 fold bias, ≥5-10 fold bias, and ≥10 fold bias. Thus, at longer CDS lengths, ovary-biased CDS exhibited consistently strong patterns of faster protein evolution. For shorter CDS, which were fewer in number, a different trend was observed, wherein testis-biased CDS exhibited statistically significantly higher dN/dS in the ≥10 fold bias class (MWU-test $P<0.05$), and had a tendency (note, not statistically significant) towards higher dN/dS in the ≥5-10 fold-bias class, suggesting greater divergence of testis-biased CDS in these subsets (Table S2). We observed that dS was consistently lower in the testis-biased CDS than in ovary-biased CDS, perhaps suggesting a lower mutation rate in genes linked to these male tissues (see Supplementary Note 4). Nonetheless, these results support a model wherein dN/dS is persistently elevated in ovary-biased genes (Table S2, Figure 3), and in a relatively small subset of testis-biased genes (Table S2).

As noted above, within the highest fold-bias classes for short CDS (≥10 fold, and ≥5 fold <10), dN/dS was higher, and dS was lower, in testis- than ovary-biased CDS (MWU-test $P<0.05$). Thus, one could speculate that higher dN/dS in the testis-biased genes in those two cases may be partly caused by a lower dS value from stronger selection on synonymous codon usage (observed in testis-biased genes, see Figure S3). However, all six classes in Table S2

showed lower dS in testis-biased genes, including those classes with elevated dN/dS in ovary-biased genes, thus excluding any such effect. To further exclude a role of codon usage patterns on our dN/dS results, in addition to the Nei-Gojobori (1986) method in MEGA (KUMAR *et al.* 2012), we also used PAML to measure dN, dS, and dN/dS, based on a substitution model that specifically includes maximum likelihood and codon usage bias (YANG 2007). Using this method, we found nearly identical results to those reported in Table S2 for all categories of gene expression (data not shown), and the values of dN/dS across all 9,389 CDS under study were highly correlated between the two estimation methods (for dN/dS, Spearman's R=0.95, $P<2.2X10^{-7}$, Figure S2). In summary, ovary-biased CDS exhibited higher dN in five of six categories of fold-bias and CDS-lengths, and higher dN/dS in four of six classes, but exhibited lower dN/dS than testis-biased CDS in short CDS with 5-fold or higher bias. In this regard, the rapid functional divergence (dN/dS, ELLEGREN AND PARSCH 2007) of ovary-biased genes is not universal across all bias and CDS length classes, suggesting a complex pattern of evolution of sex-biased gonad genes when examined at a fine scale rather than genome-wide.

With respect to Fop, the ovary-biased CDS had statistically significantly lower Fop than testis-biased CDS for each of the ≥2-5 fold bias, ≥5-10 fold-bias, and ≥10 fold-bias classes for both short and long CDS (Figure S3). Thus, selection on optimal codon usage was consistently reduced for ovary-biased CDS as compared to testis-biased CDS regardless of the magnitude of fold-bias or CDS length. Close inspection of Table S2 reveals that in five of six combinations of fold-bias (3 classes) and CDS length (2 classes), the dN was higher in ovary- than testis-biased CDS (statistically significant for four of five). In terms of selective sweeps, the net rate of amino acid changes (dN) should be directly linked to reductions in Fop (BETANCOURT AND PRESGRAVES 2002), and in this regard, the combined results from Table S2 and Figure S3 are

largely consistent with events of selective sweeps in ovary-biased genes. As a secondary test of the link between protein sequence and codon usage evolution, we measured Spearman Rank correlations between dN and Fop. As shown in Table S3, dN was negatively correlated to Fop for each fold-bias and length class for testis- and for ovary-biased genes ($P \leq 2.6X10^{-6}$ for all correlations). Given the consistently higher dN in ovary-biased genes (Table S2), this translates into the potential for an overall higher number of events (amino acid fixations) of selective sweeps in ovary-biased genes. In this regard, these results concur with trends observed at the genome wide level in Aedes (Figure 3), and resemble those reported in Drosophila (BETANCOURT AND PRESGRAVES 2002), except that in Aedes the female- (ovary-) biased genes exhibit high dN and low Fop, rather than the male-biased genes. The trends also concur with Neurospora where genes with upregulation in female-sex-organs had higher sequence divergence and lower Fop (WHITTLE AND JOHANNESSON 2013).

Supplemental Note 3: Extremely Highly Expressed Testis-Biased Genes Evolve Faster than Ovary-biased Genes

As an additional assessment of sex-biased gonad genes, we examined the most extremely highly expressed sex-biased genes in the genome (≥1,000 FPKM), regardless of fold-bias level (with a cutoff minimum of ≥2 fold-bias). Due to the smaller sample sizes (N between 17 and 107), we used randomization tests of 1,000 permutations to measure P-values for this assessment, which are more powerful than MWU-tests for small samples (see Materials and Methods). As shown in Table 2, testis-biased CDS had higher dN/dS than ovary-biased CDS for both short and long categories (Randomization tests $P<0.05$). However, importantly, both testis- and ovary-biased CDS had higher dN/dS and lower Fop than unbiased CDS (note only short unbiased genes were statistically assessed, see Table 2). The low dN/dS and dN in unbiased CDS as compared to sex-biased CDS, despite the reduced dS of the unbiased CDS (Table 2), suggests exceptionally strong purifying selection in the former gene set. Fop was also lower in testis- and ovary-biased CDS than in unbiased CDS (randomization-tests $P<0.05$, Table 2). In this regard, within the relatively small subset of sex-biased CDS that were extremely highly transcribed in this taxon (N=220 of 9,839 genes, Table 2), testis-biased and ovary-biased genes exhibited rapid evolution. It is worth noting that for Table 2, using branch-site analysis we found evidence of positive selection in 16.7% of the testis-biased CDS (pooled short and long) and 7.6% in the ovary-biased CDS in *A. aegytpi*, suggesting positive selection contributes to faster evolution of the male-gonad CDS in this small specific subset of highly transcribed genes.

Functional annotation of the extremely highly expressed CDS was assessed using the GO tool DAVID (Huang da et al., 2009). As shown in Table S4, the unbiased CDS with ≥1,000 FPKM contained a concentration of ribosomal protein genes, which are typically among the most

highly expressed CDS in a genome (Wang et al. 2011). Notably, however, the testis-biased and ovary-biased genes included also ribosomal protein genes (N=10 and 28, respectively), suggesting that at least some ribosomal protein genes are differentially expressed between the gonads of the two sexes (Tables S4). While ribosomal protein genes are known as an essential component of the ribosome and needed for translation, data also indicate that many, seemingly redundant, gene copies of the approximately 80 core (with some variation among species) cytoplasmic ribosomal proteins (KENMOCHI *et al.* 1998; PLANTA AND MAGER 1998; BARAKAT *et al.* 2001)(Ribosomal Protein Gene Database, http://ribosome.med.miyazaki-u.ac.jp/) exist in some eukaryotic genomes, and that these may be involved in tissue-specific gene regulation or processes. This concurs with results showing that ribosomal protein gene copies exhibit differential expression across tissues in plants, yeast and certain animals, and implications that they are involved in tissue-specific gene regulation (UECHI *et al.* 2002; KOMILI *et al.* 2007; WHITTLE AND KROCHKO 2009). Ribosomal protein genes have also been linked to a number of specific biological processes in eukaryotes, including tumor suppression, apoptosis, development, ageing and post-transcriptional gene regulation Here, we speculate that these testis-biased and ovary-biased genes, including ribosomal-protein genes, may be involved in genetic pathways related to sex-specific activities, possibly including maintenance of the male/female gonads, mating or fertilization, and that their role in these processes may accelerate the sequence evolution (SWANSON AND VACQUIER 2002) and reduce optimal codon usage (HAMBUCH AND PARSCH 2005) in these genes. This notion is in agreement with the following two observations: first, these genes were highly differentially expressed between the male/female gonads (≥2 fold and a median of 712.0 fold testis-bias and 13.2 fold ovary-bias), and second,

despite their marked expression levels, they exhibited greater dN/dS than highly transcribed unbiased genes (Table 2), consistent with altered rates of evolution.

Supplemental Note 4: Variation in dS

It is worth noting that dS was lower in the extremely highly expressed unbiased CDS than in the testis- and ovary-biased genes in Table 2, and was also lower for testis-biased than ovary-biased genes in Table S2. We found no notable differences in these dS estimates when measured using a substitution model that does or does not correct for codon usage (YANG 2007) (see Supplementary Note 2). Thus, by excluding any major effect of selection on codon biases (observed in Table 1) on dS, it may be inferred these (dS) values likely largely reflect the mutation rate (KIMURA 1983); this finding is therefore suggestive of the hypothesis that the mutation rate might vary within the genome of *A. aegypti*. Mutation rates are known to vary within the genome, including across chromosomes and sites, in numerous organisms (ELLEGREN *et al.* 2003; HODGKINSON AND EYRE-WALKER 2011; NESS *et al.* 2015), which may result from differential DNA repair or other mechanisms (SUPEK AND LEHNER 2015). Mutation rates can also evolve in some genomic regions due to selective pressures (SNIEGOWSKI 1997). Further, if mutations in the testis are more typically highly deleterious and subjected to greater purifying selection (as suggested by their lower genome-wide dN/dS, Figure 3), it is feasible that evolution has favored lowered mutation rates in testis-involved genes in this taxon

Supplemental Note 5: Optimal Codon Usage for Genes Expressed in Ovaries, Testes and Embryos

With regard to Fop for gene sets in Figure 4: values were lowest in ovary-specific genes (mean=0.350±0.004), co-occurring with the highest dN/dS, thus suggesting these two traits evolve in concert in Aedes, and are particularly well connected for ovary-biased genes. The largest Fop values were observed in genes expressed in all tissues (mean=0.407±0.001; MWU-test as compared to ovary-specific, $P<0.001$), consistent with purifying selection promoting optimal codons in highly, and broadly, expressed genes, and intermediate values were observed for all remaining categories.

Table S1. The genomic and transcriptomic datasets for Aedes used in the present study. Expression data are for *Aedes aegypti* (AKBARI *et al.* 2013).

Genomic Data	**Database**
Aedes aegypti	Ensembl Metazoa (metazoa.ensembl.org) Version: AaegL3.29 (see also, https://www.ncbi.nlm.nih.gov; GCF_000004015) (Downloaded August 2015)
Aedes albopictus	NCBI Bioproject: PRJNA270274; Assembly v1.1 scaffolds and GFF file (Downloaded March 1, 2016) (CHEN *et al.* 2015)

Transcriptome Data (Downloaded Nov. 2015)	**Database**	**ID**	**Tissue-Type**	**No. Reads**
Aedes aegypti	NCBI SRA: Bioproject PRJNA209388 (AKBARI *et al.* 2013)	SRX316321/ SRR923857	Non-Blood Fed Female Ovary	15,431,793 (single)
		SRX316320/ SRR923856	Non-Blood Fed Female Carcass (lacking ovaries)	16,267,418 (single)
		SRX316319/ SRR923701	Male Testes (plus accessory glands)	15,120,717 (single)
		SRX316546/ SRR923703	Male Carcass (lacking testes and accessory glands)	14,451,789 (single)
		SRX316562/ SRR923853	Embryo 8-12 hours	18,643,531 (single)

Table S2. Mean dN/dS, dN and dS and standard errors (SE) of CDS for testis- or ovary-biased expression in Aedes. The cutoff for identification as a CDS with biased expression was two-fold higher in one tissue than the other, and includes CDS with tissue-specific expression (wherein FPKM must be >1 in at least one tissue). Different letters indicate a statistically significant difference using MWU-tests between each pair of testis-biased and ovary biased genes per fold-class ($P<0.05$). $N_{Testis\ biased} = 2{,}927$; $N_{Ovary\text{-}biased} = 2013$.

		Short (≤250 Codons)					Long (>250 codons)			
Fold Bias		**Testis-Biased**		**Ovary biased**			**Testis-Biased**		**Ovary biased**	
		Mean	SE	Mean	SE		Mean	SE	Mean	SE
≥2 Fold <5	dN/dS	0.082a	0.008	0.088a	0.005	dN/dS	0.070a	0.002	0.095b	0.003
	dN	0.053a	0.008	0.073b	0.006	dN	0.059a	0.002	0.091b	0.004
	dS	0.596a	0.023	0.789b	0.024	dS	0.806a	0.009	0.897b	0.014
	N	177		273		N	1128		626	
≥5 Fold <10	dN/dS	0.110a	0.013	0.083a	0.005	dN/dS	0.067a	0.002	0.095b	0.005
	dN	0.069a	0.009	0.061a	0.004	dN	0.054a	0.002	0.084b	0.004
	dS	0.619a	0.035	0.750b	0.026	dS	0.778a	0.011	0.875b	0.020
	N	102		203		N	603		268	
≥10 Fold	dN/dS	0.120a	0.009	0.095b	0.005	dN/dS	0.077a	0.002	0.111b	0.004
	dN	0.078a	0.007	0.089a	0.006	dN	0.059a	0.002	0.101b	0.005
	dS	0.628a	0.033	0.853b	0.025	dS	0.730a	0.010	0.879b	0.017
	N	150		282		N	767		361	

Table S3. Spearman Rank Correlations between Fop and dN for testis- and ovary-biased CDS in Aedes. Data are subdivided by fold-bias and CDS lengths in *A. aegypti*.

Fold Bias	Length Class	Spearman's R (P-value)
Testis-Biased		
≥2 Fold <5	Short	-0.429 ($3.6X10^{-9}$)
≥5 Fold <10	Short	-0.406 ($2.6X10^{-6}$)
≥10	Short	-0.525 ($2.0X10^{-7}$)
≥2 Fold <5	Long	-0.262 ($2.0X10^{-7}$)
≥5 Fold <10	Long	-0.307 ($1.6X10^{-14}$)
≥10	Long	-0.361 ($2.0X10^{-7}$)
Ovary-Biased		
≥2 Fold <5	Short	-0.360 ($1.2X10^{-9}$)
≥5 Fold <10	Short	-0.394 ($8.1X10^{-9}$)
≥10	Short	-0.423 ($6.4X10^{-14}$)
≥2 Fold <5	Long	-0.221 ($2.7X10^{-8}$)
≥5 Fold <10	Long	-0.318 ($1.2X10^{-7}$)
≥10	Long	-0.331 ($1.5X10^{-10}$)

Table S4. Functional annotation of testis- and ovary biased genes and unbiased genes extremely high expressed (≥1,000FPKM) in *A. aegypti*. The gene ontology was determined using the system DAVID (Huang da et al. 2009) and the cluster with the greatest enrichment score is shown. P-values are from a modified Fisher's test, wherein lower values indicate greater enrichment.

Testis-biased ≥1,000 FPKM		Ovary-biased ≥1,000 FPKM		Unbiased ≥1,000 FPKM	
Enrichment Score: 5.78	P-Value	**Enrichment Score: 20.99**	P-Value	**Enrichment Score: 18.77**	P-Value
ribosomal protein	1.60E-09	ribosomal protein	4.70E-36	ribosomal protein	2.90E-31
Ribosome	2.90E-08	structural constituent of ribosome	1.90E-32	structural constituent of ribosome	5.70E-27
ribosome	1.60E-07	ribosome	3.90E-23	Ribosome	3.30E-21
structural constituent of ribosome	1.80E-07	translation	4.80E-23	translation	9.80E-21
ribonucleoprotein complex	3.00E-06	ribonucleoprotein complex	1.80E-22	structural molecule activity	1.30E-18
ribonucleoprotein	3.20E-06	Ribosome	4.60E-21	ribosome	2.00E-18
translation	3.30E-06	structural molecule activity	9.80E-19	ribonucleoprotein complex	7.80E-16
non-membrane-bounded organelle	7.20E-05	ribonucleoprotein	1.20E-15	non-membrane-bounded organelle	3.40E-11
intracellular non-membrane-bounded organelle	7.20E-05	intracellular non-membrane-bounded organelle	2.70E-12		3.40E-11
structural molecule activity	7.60E-04	non-membrane-bounded organelle	2.70E-12		

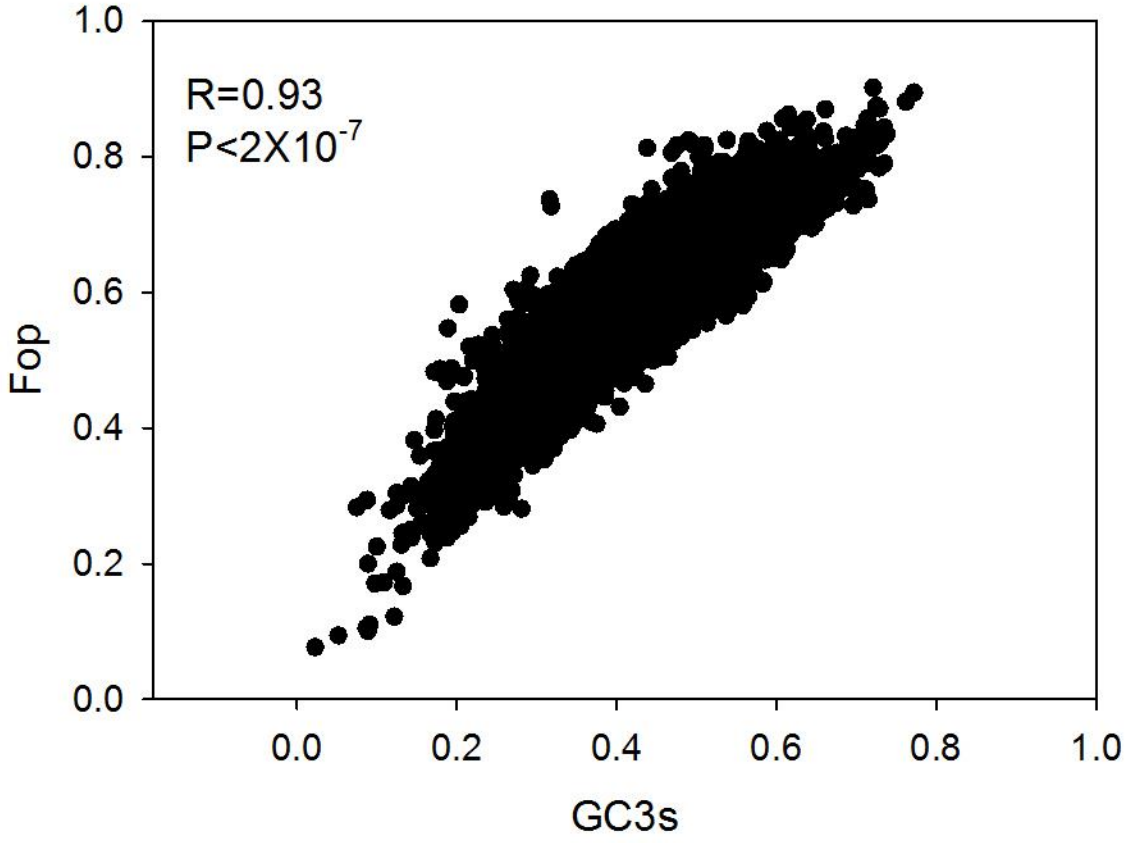

Figure S1. The relationship between GC content at synonymous third codon positions (GC3s) and the frequency of optimal codons (Fop) across all embryo expressed *A. aegypti* CDS. Spearman's Ranked correlation and P-value are shown.

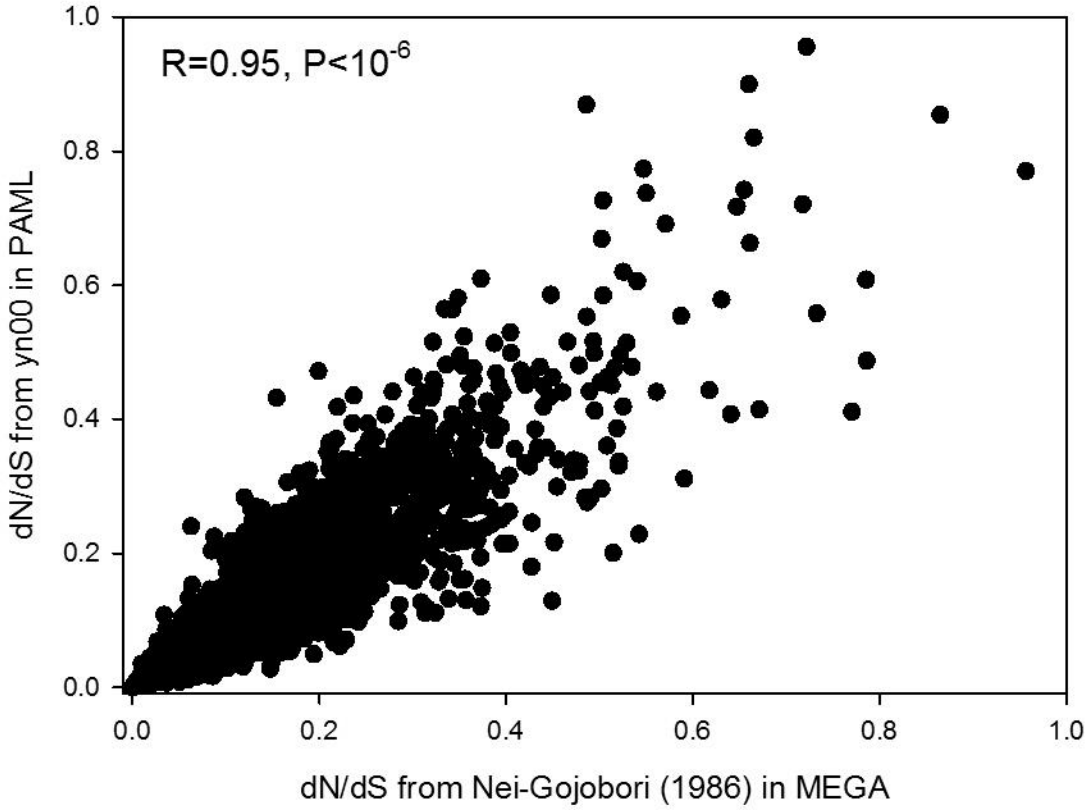

Figure S2. The dN/dS values obtained across 9,389 genes under study using the Nei-Gojobori method (NEI AND GOJOBORI 1986) in MEGA (KUMAR *et al.* 2012) and yn00 in PAML (YANG 2007). Spearman's R correlation value is shown.

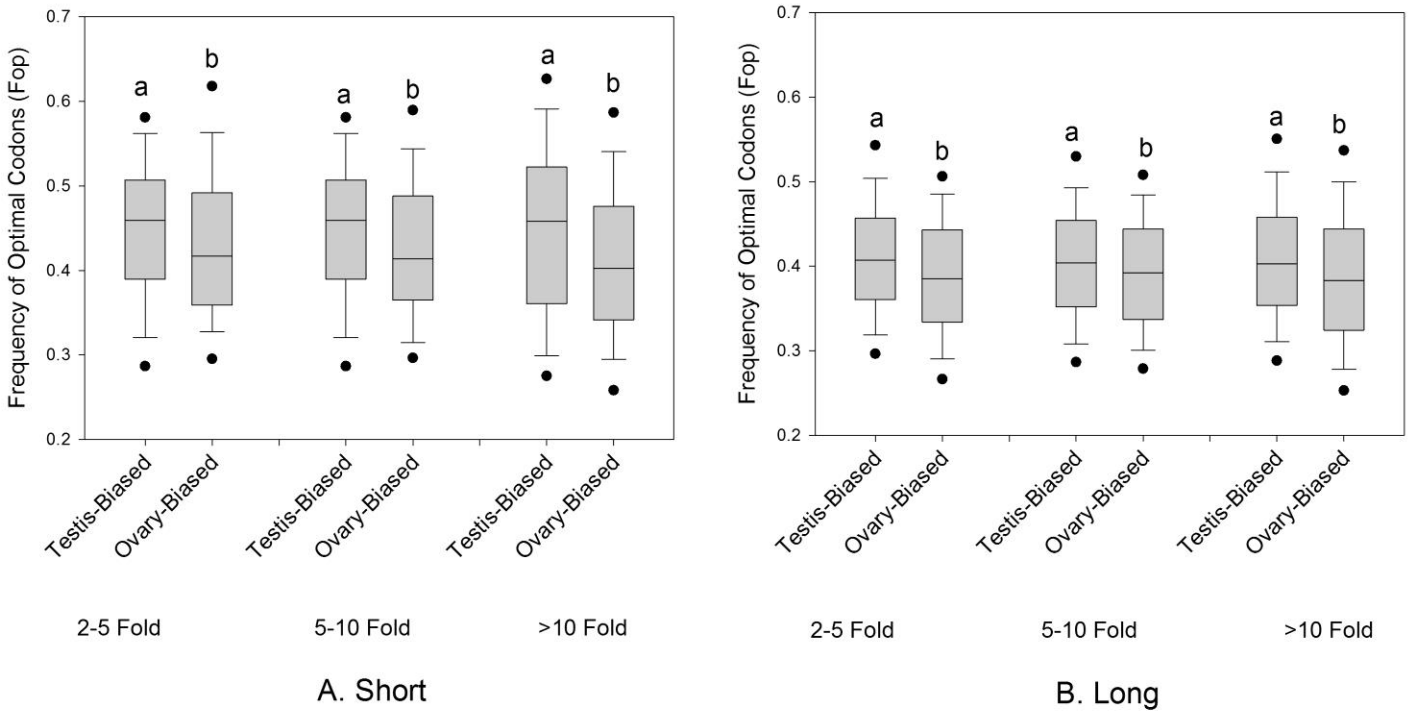

Figure S3. The frequency of optimal codons (Fop) for testis-biased and ovary-biased CDS relative to fold expression bias. A. Short CDS; B. Long CDS. Different letters within each fold-class indicates a statistically significant difference ($P<0.05$) using MWU-tests (P-values were <0.001).

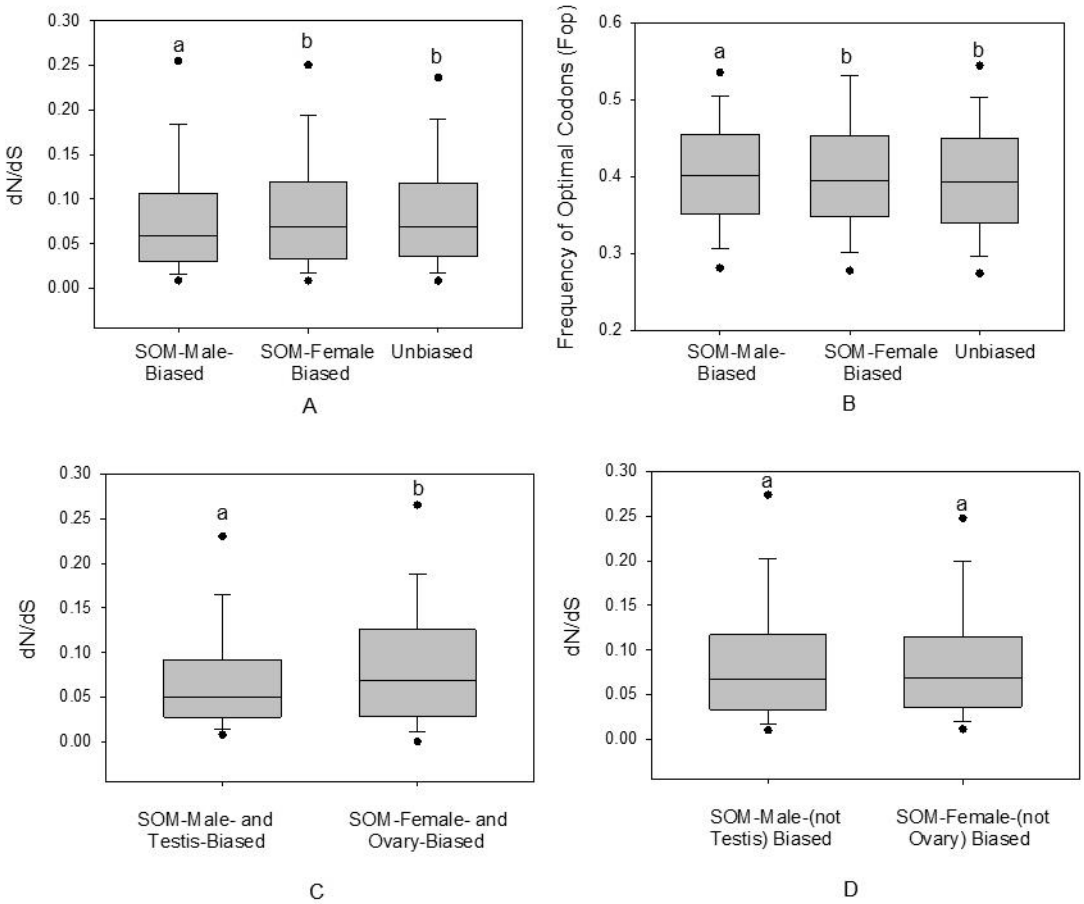

Figure S4. Box and whisker plot of for adult somatic (SOM) male-biased (N=3,743), SOM-female-biased (N=774) and unbiased (N=4,872) CDS in *A. aegypti* (note: SOM-male is the carcass minus testes and accessory glands, and SOM-female is the carcass minus ovaries, Table S1). A. dN/dS for all genes per category; B. Fop for all genes per category; C. dN/dS for genes that are both SOM-male-biased and testis-biased (N=1,621), and those that are both SOM female-biased and ovary-biased (N=264); D. dN/dS for SOM-male-biased genes (from A) after excluding those also having testis-biased expression (N=2,122), and SOM-female-biased genes after excluding those also having ovary-biased expression (N=510). Different letters above any two boxes in each figure indicate a statistically significant difference using MWU-tests ($P<0.05$). For panel D, the MWU test $P=0.400$. Fop was measured using *A. aegypti*, and dN/dS measured using *A. aegypti* and *A. albopictus*.

Supplementary References

Akbari, O. S., I. Antoshechkin, H. Amrhein, B. Williams, R. Diloreto *et al.*, 2013 The developmental transcriptome of the mosquito Aedes aegypti, an invasive species and major arbovirus vector. G3 (Bethesda) 3**:** 1493-1509.

Barakat, A., K. Szick-Miranda, I. F. Chang, R. Guyot, G. Blanc *et al.*, 2001 The organization of cytoplasmic ribosomal protein genes in the Arabidopsis genome. Plant Physiol 127**:** 398-415.

Betancourt, A. J., and D. C. Presgraves, 2002 Linkage limits the power of natural selection in Drosophila. Proceedings Of The National Academy Of Sciences Of The United States Of America 99**:** 13616-13620.

Chen, X. G., X. Jiang, J. Gu, M. Xu, Y. Wu *et al.*, 2015 Genome sequence of the Asian Tiger mosquito, Aedes albopictus, reveals insights into its biology, genetics, and evolution. Proc Natl Acad Sci U S A 112**:** E5907-5915.

Combs, P. A., and M. B. Eisen, 2013 Sequencing mRNA from cryo-sliced Drosophila embryos to determine genome-wide spatial patterns of gene expression. PloS one 8**:** e71820.

Cutter, A. D., J. D. Wasmuth and M. L. Blaxter, 2006 The evolution of biased codon and amino acid usage in nematode genomes. Molecular biology and evolution 23**:** 2303-2315.

Diez-Roux, G., S. Banfi, M. Sultan, L. Geffers, S. Anand *et al.*, 2011 A high-resolution anatomical atlas of the transcriptome in the mouse embryo. PLoS Biology 9**:** e1000582.

Donoughe, S., and C. G. Extavour, 2016 Embryonic development of the cricket *Gryllus bimaculatus*. Developmental Biology 411**:** 140-156.

Duret, L., and D. Mouchiroud, 1999 Expression pattern and, surprisingly, gene length shape codon usage in Caenorhabditis, Drosophila, and Arabidopsis. Proceedings of the National Academy of Sciences of the United States of America 96**:** 4482-4487.

Ellegren, H., and J. Parsch, 2007 The evolution of sex-biased genes and sex-biased gene expression. Nature Reviews Genetics 8**:** 689-698.

Ellegren, H., N. G. Smith and M. T. Webster, 2003 Mutation rate variation in the mammalian genome. Curr Opin Genet Dev 13**:** 562-568.

Farlow, A., M. Dolezal, L. Hua and C. Schlotterer, 2012 The genomic signature of splicing-coupled selection differs between long and short introns. Mol Biol Evol 29**:** 21-24.

Hambuch, T. M., and J. Parsch, 2005 Patterns of synonymous codon usage in Drosophila melanogaster genes with sex-biased expression. Genetics 170**:** 1691-1700.

Hodgkinson, A., and A. Eyre-Walker, 2011 Variation in the mutation rate across mammalian genomes. Nature reviews. Genetics 12**:** 756-766.

Ikemura, T., 1985 Codon usage and tRNA content in unicellular and multicellular organisms. Molecular Biology and Evolution 2**:** 13-34.

Ingvarsson, P. K., 2008 Molecular evolution of synonymous codon usage in Populus. BMC Evolutionary Biology 8**:** 307.

Kenmochi, N., T. Kawaguchi, S. Rozen, E. Davis, N. Goodman *et al.*, 1998 A map of 75 human ribosomal protein genes. Genome Res 8**:** 509-523.

Kimura, M., 1983 *The neutral theory of molecular evolution*. Cambridge University Press, Cambridge.

Komili, S., N. G. Farny, F. P. Roth and P. A. Silver, 2007 Functional specificity among ribosomal proteins regulates gene expression. Cell 131**:** 557-571.

Kumar, S., G. Stecher, D. Peterson and K. Tamura, 2012 MEGA-CC: computing core of molecular evolutionary genetics analysis program for automated and iterative data analysis. Bioinformatics 28**:** 2685-2686.

McVean, G. A., and J. Vieira, 1999 The evolution of codon preferences in Drosophila: a maximum-likelihood approach to parameter estimation and hypothesis testing. Journal of Molecular Evolution 49**:** 63-75.

Mittmann, B., and C. Wolff, 2012 Embryonic development and staging of the cobweb spider *Parasteatoda tepidariorum* C. L. Koch, 1841 (syn.: *Achaearanea tepidariorum*; Araneomorphae; Theridiidae). Development Genes and Evolution 222**:** 189-216.

Nei, M., and T. Gojobori, 1986 Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Molecular biology and evolution 3**:** 418-426.

Ness, R. W., A. D. Morgan, R. B. Vasanthakrishnan, N. Colegrave and P. D. Keightley, 2015 Extensive de novo mutation rate variation between individuals and across the genome of Chlamydomonas reinhardtii. Genome Res 25**:** 1739-1749.

Peden, J. F., 1999 Analysis of codon usage, pp. University of Nottingham.

Percudani, R., 2001 Restricted wobble rules for eukaryotic genomes. Trends in Genetics 17**:** 133-135.

Planta, R. J., and W. H. Mager, 1998 The list of cytoplasmic ribosomal proteins of Saccharomyces cerevisiae. Yeast 14**:** 471-477.

Schwager, E. E., Y. Meng and C. G. Extavour, 2015 *vasa* and *piwi* are required for mitotic integrity in early embryogenesis in the spider *Parasteatoda tepidariorum*. Developmental Biology 402**:** 276-290

Sniegowski, P., 1997 Evolution: setting the mutation rate. Curr Biol 7**:** R487-488.

Subramanian, S., and S. Kumar, 2004 Gene expression intensity shapes evolutionary rates of the proteins encoded by the vertebrate genome. Genetics 168: 373-381.

Supek, F., and B. Lehner, 2015 Differential DNA mismatch repair underlies mutation rate variation across the human genome. Nature 521: 81-84.

Swanson, W. J., and V. D. Vacquier, 2002 The rapid evolution of reproductive proteins. Nat Rev Genet 3: 137-144.

Uechi, T., N. Maeda, T. Tanaka and N. Kenmochi, 2002 Functional second genes generated by retrotransposition of the X-linked ribosomal protein genes. Nucleic Acids Res 30: 5369-5375.

Wang, B., Z. W. Shao, Y. Xu, J. Liu, Y. Liu *et al.*, 2011 Optimal codon identities in bacteria: implications from the conflicting results of two different methods. PLoS ONE 6: e22714.

Whittle, C. A., and C. G. Extavour, 2015 Codon and amino acid usage are shaped by selection across divergent model organisms of the Pancrustacea. G3: Genes, Genomes, Genetics 5: 2307-2321.

Whittle, C. A., and C. G. Extavour, 2016 Expression-linked patterns of codon usage, amino acid frequency and protein length in the basally branching arthropod *Parasteatoda tepidariorum*. Genome Biology and Evolution 8: 2722-2736.

Whittle, C. A., and H. Johannesson, 2013 Evolutionary dynamics of sex-biased genes in a hermaphrodite fungus. Molecular Biology and Evolution 30: 2435-2446.

Whittle, C. A., and J. E. Krochko, 2009 Transcript profiling provides evidence of functional divergence and expression networks among ribosomal-protein gene paralogs in *Brassica napus*. Plant Cell 21: 2203-2221.

Whittle, C. A., Y. Sun and H. Johannesson, 2011 Evolution of synonymous codon usage in *Neurospora tetrasperma* and *Neurospora discreta*. Genome Biology and Evolution 3: 332-343.

Williford, A., and J. P. Demuth, 2012 Gene expression levels are correlated with synonymous codon usage, amino acid composition, and gene architecture in the red flour beetle, *Tribolium castaneum*. Molecular Biology and Evolution 29: 3755-3766.

Yang, Z., 2007 PAML 4: phylogenetic analysis by maximum likelihood. Molecular Biology and Evolution 24: 1586-1591.