SUPPLEMENTARY MATERIAL

# Release from intralocus sexual conflict? Evolved loss of a male sexual trait demasculinises female gene expression

Jack G. Rayner,[1,*,†] Sonia Pascoal,[2,†] Nathan W. Bailey[1,*]

[1]School of Biology, University of St Andrews, St Andrews, Fife KY16 9TH, UK
[2]Department of Zoology, University of Cambridge, CB2 3EJ, UK

[†]These authors contributed equally to this work.
*__Corresponding authors:__ E-mails: jr228@st-andrews.ac.uk, nwb3@st-andrews.ac.uk

# SI. SUPPORTING METHODS

## Cricket rearing and production of lines

Pure-breeding *normal-wing* (*NW*) and *flatwing* (*FW*) lines from which RNA samples were collected were derived from individuals caught in Kauai, Hawaii in 2012 following the procedures detailed in Pascoal et al. (2016), and thereafter maintained under common garden conditions at 25 °C on a 12h:12h light:dark cycle in an incubator with calling males present. Lines used in the follow-up comparison of reproductive output of *NW* and *FW* genotypes of the two sexes were descended from the same lines from which RNA samples were collected, but had since been outcrossed by breeding *NW* and *FW lines* together to reduce effects of inbreeding. Pure-breeding *NW* and *FW* lines were subsequently re-established by mating heterozygous females with males of the desired morph, then backcrossing in the next generation and screening offspring to retain homozygous families. All crickets were reared under common-garden conditions, exposed to acoustic signals of other crickets in the environment. Note that *FW* and *NW* lines have not been found to differ in development time (Rayner & Bailey, unpublished data).

## Collection and sequencing of RNA samples

Head and body tissues were separated and the digestive tract discarded, then samples were stored in RNAlater at -20 °C. Immediately prior to RNA extraction, neural tissue was dissected from the head capsule, and thoracic muscle and testes or ovaries were dissected from the body cavity. Quantity and integrity of RNA samples were assessed using Qubit broad range (Invitrogen) and Agilent

Bioanalyzer assays, respectively. Total RNA was depleted with RiboZero and used in the ScriptSeq protocol (Epicentre), following manufacturer instructions. Samples were sequenced on an Illumina HiSeq 2000 with version 3 chemistry, generating 2×100 bp paired end reads. CASAVA v1.8.2 (Illumina) was used for basecalling and de-multiplexing of indexed reads. Adapter sequences were trimmed from fastq files using Cutadapt v1.2.1 (Martin 2011) and low quality bases were removed using Sickle v1.200 with a minimum window quality score of 20.

**Trinity Assembly Quality**

Across all 36 samples, sequencing produced 844,050,339 paired reads which passed trimming (Table S4). Trinity-identified genes and isoforms showing greater than 95% similarity were clustered using CD-hit-est (Li & Godzik 2006). Assembly completeness was assessed by testing for the presence of insect BUSCO groups of conserved expressed genes (Simão et al. 2015). Reads were aligned to the transcriptome using Bowtie2 (Langmead & Salzberg 2012) with strand-specific settings, and quantified in RSEM (Li & Dewey 2011) at the Trinity gene level using default parameters. Trinity-identified 'genes' without open reading frames (ORFs) of >100 amino acids in length were removed so that only transcripts with evidence of protein coding regions were included in the analysis.

Trinity *de novo* transcriptome assembly statistics, before and after ORF expression-based filtering are given in Table S5. BUSCO analysis of the unfiltered transcriptome showed 97.8% of 1658 conserved insect genes were present in the unfiltered transcriptome, 7.2% fragmented, and 90.6% complete; of which, after extracting the longest putative isoform for each gene, only 1.8% were duplicated.

Together, these statistics suggest the filtered assembly is of a good quality, at least comparable with previously published *de novo* assemblies for *T. oceanicus* and closely related species (Bailey et al. 2013; Berdan et al. 2016; Kasumovic et al. 2016; Pascoal et al. 2016; Pascoal et al. 2018a). Only the 42,496 transcripts passing ORF and expression filters were included in downstream analyses.

**Differential Expression Analyses**

Prior to constructing models, transcripts not expressed at a level above one count per million in a minimum of three samples were filtered from the dataset, as these were considered to have little empirical support and because their removal increased power to identify DE transcripts. After filtering, input counts were adjusted using trimmed mean of M-values (TMM) normalisation.

Separate models were initially constructed for somatic (including neural tissue and thoracic muscle from both sexes), and gonad (testes and ovaries) samples. Scaled normalisation procedures assume that no more than 50% of transcripts in a dataset are differentially expressed between any two groups of samples, which was likely to be violated in a single model including both reproductive and somatic tissues. Importantly, however, the TMM method is relatively robust against violations of this assumption (Robinson & Oshlack 2010). For downstream investigation of sex-biased patterns of gene expression it was convenient to combine gonad tissues from both sexes in a single model, despite the expectation that a very large proportion of transcripts would be differentially expressed between the sexes. We tested the validity of results from this combined gonad model by also constructing and comparing separate models for testes and ovaries samples. For downstream

analyses involving gonad tissues we used results from the model including gonad tissues from both sexes, as comparison of the identity of statistically DE transcripts indicated a high degree of overlap between separate and combined-sex models for ovaries (140 DE in ovaries model, 185 DE in combined sex model; 123 out of 140 (87.76%) DE transcripts shared between the two), while 16 transcripts were DE in testes samples in the separate sex model, versus 9 in the combined-sex model (6 out of 9 shared between the two).

Negative binomial generalised linear models (GLMs) were constructed in edgeR (Robinson et al. 2010), and tested using likelihood ratio tests. Once models had been constructed, pairwise contrasts were performed between groups of samples, as recommended by the authors of EdgeR for more complex experimental designs (Robinson et al. 2010), with a Benjamini and Hochberg-adjusted significance threshold of FDR<0.01. Contrasts were specified to examine the number of DE transcripts between morph genotypes for each of the tissues in each sex, as well as a sex comparison which was performed by contrasting average male and female expression values (i.e. samples for both morph genotypes were included for each sex). The approach of including both morph genotypes in sex comparisons was adopted to avoid statistical artefacts that could result from defining sex-biased transcripts using only normal-winged samples, i.e. using the same reference groups in both sex (*NW* male vs *NW* female) and morph (e.g. *NW* male vs *FW* male) comparisons (see: Mallard et al. 2018).

**Calculating Scaled Mass Index**

Scaled mass index (SMI) was calculated using the following equation:

110 $$\widehat{M}_i = M_i \left[\frac{L_0}{L_i}\right]^{b_{SMA}}$$

111 where for each individual, $i$, $M_i$ and $L_i$ represent mass (with total egg mass or testes
112 mass deducted) and pronotum length, respectively, $L_0$ is the mean pronotum length
113 within the population, and $b_{SMA}$ is a scaling component, here estimated by log-log
114 standardised major axis (SMA) regression of somatic mass on pronotum length.
115 Averages and SMA regression were calculated separately for each of the sexes.
116

## SII. SUPPORTING REFERENCES


Bailey, N.W., Veltsos, P., Tan, Y.F., Millar, A.H., Ritchie, M.G. & Simmons, L.W. (2013) Tissue-specific transcriptomics in the field cricket Teleogryllus oceanicus. *G3 (Bethesda)*, 3(2): 225–30.

Berdan, E.L., Blankers, T., Waurick, I., Mazzoni, C.J., & Mayer, F. (2016) A genes eye view of ontogeny: de novo assembly and profiling of the Gryllus rubens transcriptome. *Molecular Ecology Resources*, 16(6): 1478–1490.

Kasumovic, M.M., Chen, Z. & Wilkins, M.R., (2016) Australian black field crickets show changes in neural gene expression associated with socially-induced morphological , life- history , and behavioral plasticity. *BMC Genomics*, 17(827): 1–35.

Lun, A.T.L., Chen, Y., & Smyth, G.K., (2016) It's DE-licious: A recipe for differential expression analyses of RNA-seq experiments using quasi-likelihood methods in edgeR. *Methods in Molecular Biology*. DOI: 10.1007/978-1-4939-3578-9_19

Mallard, F., Jakšić, A.M. & Schlötterer, C. (2018) Contesting the evidence for non-adaptive plasticity. Nature, 555: E21.

Martin, M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*, 17(1): 10.

Pascoal, S., Liu, X., Fang, Y., Paterson, S., Ritchie, M.G., Rockliffe, N. et al. (2018a). Increased socially mediated plasticity in gene expression accompanies rapid adaptive evolution. *Ecol. Lett.*, 21(4):546-556.

Pascoal, S., Risse, J.E., Zhang, X., Blaxter, M., Cezard, T. et al. (2018b) Silent crickets reveal the genomic footprint of recent adaptive trait loss. bioRxiv https://doi.org/10.1101/489526.

141 Pascoal, S., Liu, X., Ly, T., Fang, Y., Rockliffe, N., Paterson, S. et al. (2016). Rapid
142 evolution and gene expression: a rapidly evolving Mendelian trait that silences
143 field crickets has widespread effects on mRNA and protein expression. *J. Evol.*
144 *Biol.*, 29(6): 1234–1246.
145 Robinson, M.D. & Oshlack, A. (2010) A scaling normalization method for differential
146 expression analysis of RNA-seq data. *Genome Biology*, 11(3).
147 Simão, F.A., Waterhouse, R.M., Ioannidis, P., Kriventseva, E.V., Zbodnov, E.M.
148 (2015). BUSCO: Assessing genome assembly and annotation completeness
149 with single-copy orthologs. *Bioinformatics*, 31(19): 3210–3212.
150

## SIII. SUPPORTING FIGURES

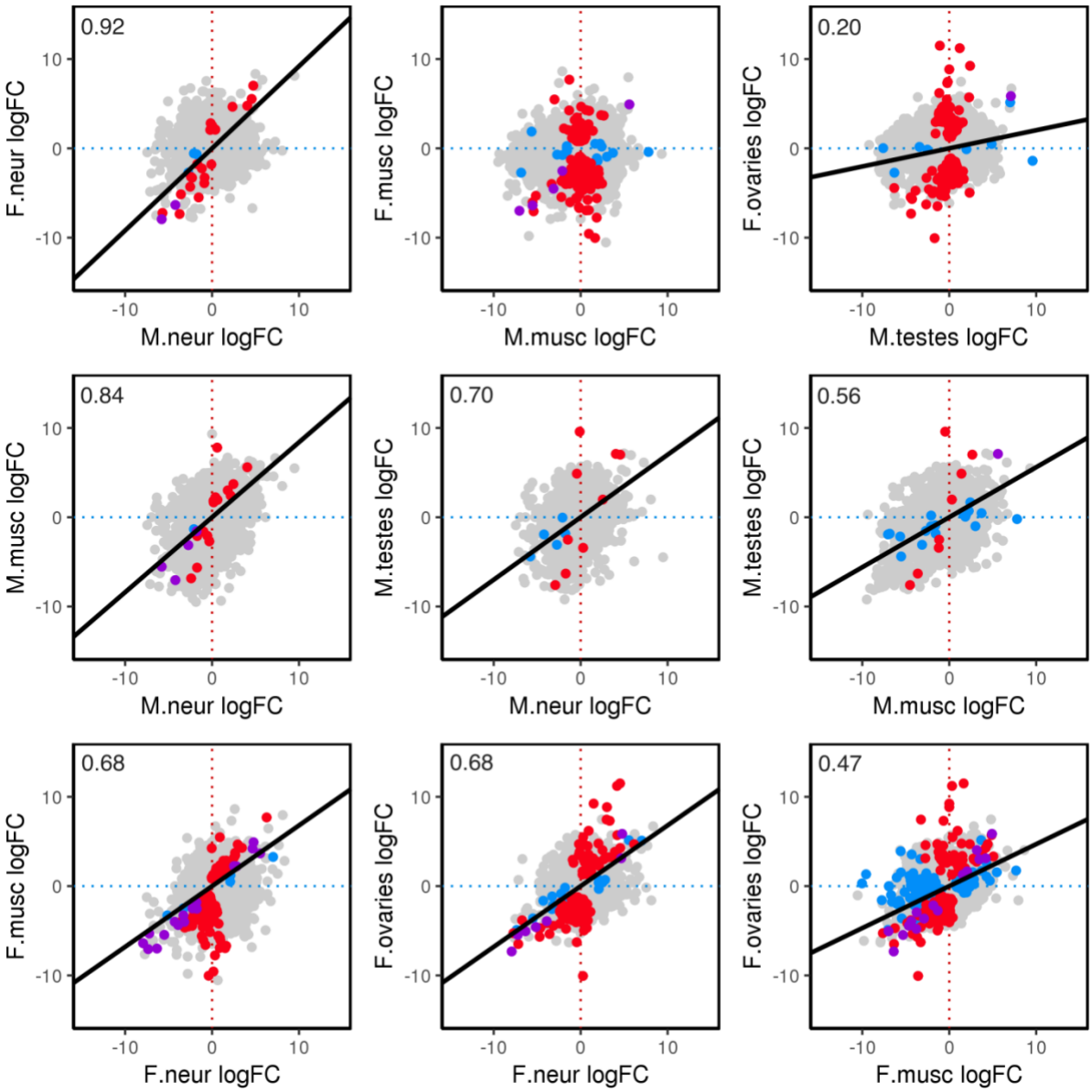


**Fig. S1.** Concordance of *flatwing*-associated gene expression changes in different sexes (top) and across male (middle row) and female (bottom row) tissues. Greater values along either axis indicate a transcript was more *FW*-biased. Transcripts DE in the x-axis comparison are highlighted in blue, transcripts DE in the y-axis comparison in red, and transcripts DE in both comparisons in purple. Lines and annotations denote coefficients from Spearman's rank correlation (P<0.01) across transcripts DE in either comparison.

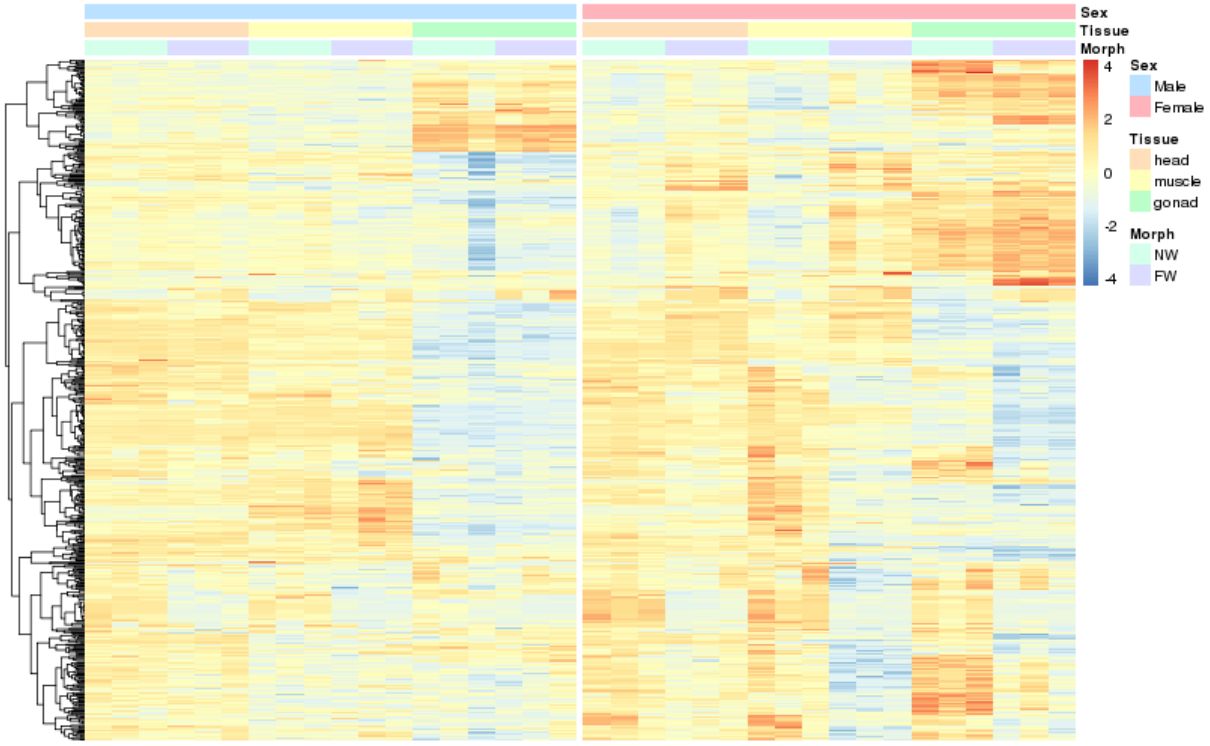


**Figure S2.** Heatmap showing expression values ($log_2$-transformed counts per million) across all samples for transcripts significantly (FDR<0.01) DE in any morph genotype comparison, z-scaled by row, for each of the morph genotypes and tissues. Note that expression values for somatic (neural and thoracic muscle) and gonad tissues are derived from separately normalised libraries, thus only relative (NW vs FW), not absolute, expression values can be compared between the two.

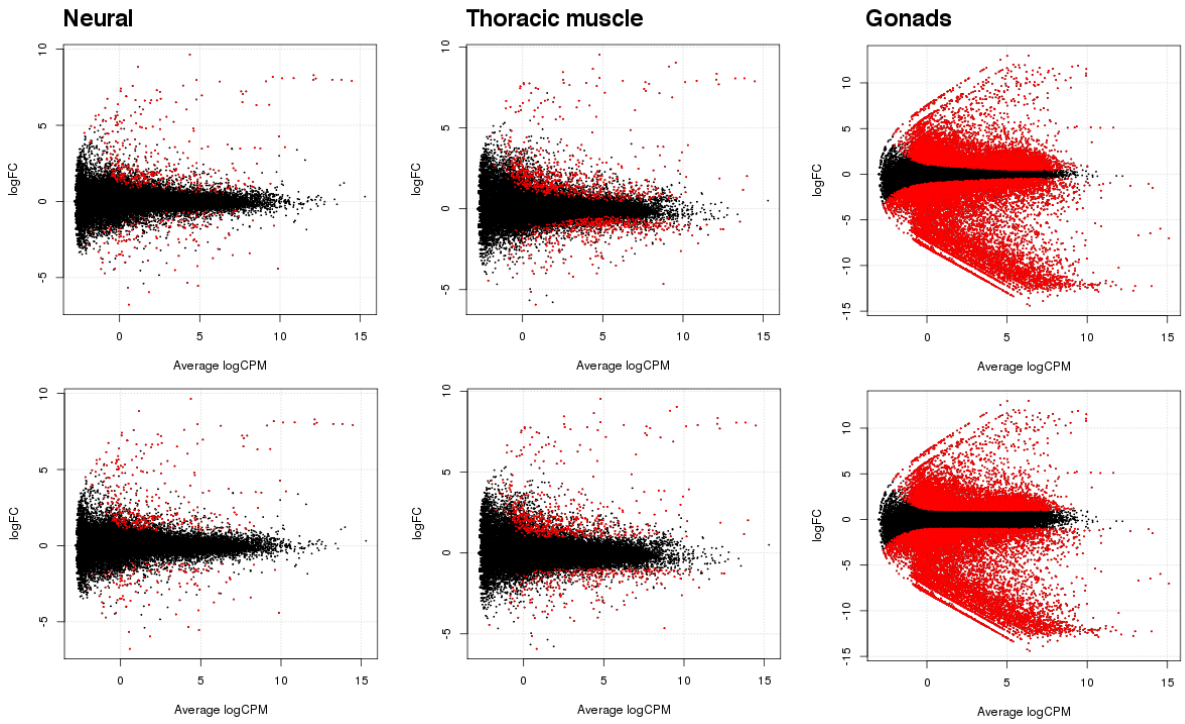


**Figure S3.** Volcano plots for sex comparisons between each of the tissues, showing expression level in $\log_2$ counts-per-million along the X-axis, and $\log_2$-fold change on the Y-axis. Red points are transcripts identified as DE at FDR<0.01. In the top row, no fold-change threshold is imposed for a transcript to be DE, whereas in the bottom row a $\log_2$-fold change of 1 (fold-change of 2) threshold is imposed.

# SIV. SUPPORTING TABLES

**Table S1.** Significantly enriched GO categories for each of the female genotypes. Male genotypes showed no significant GO enrichment.

| Genotype | GO Name | GO Category | FDR | Nr Test | Nr Reference | Non Annot Test | Non Annot Reference |
|---|---|---|---|---|---|---|---|
| Female.FW | endomembrane system | CELLULAR_COMPONENT | 1.82E-09 | 30 | 648 | 104 | 14195 |
| Female.FW | endoplasmic reticulum | CELLULAR_COMPONENT | 1.96E-07 | 15 | 160 | 119 | 14683 |
| Female.FW | ncRNA metabolic process | BIOLOGICAL_PROCESS | 0.005780891 | 9 | 118 | 125 | 14725 |
| Female.FW | tRNA aminoacylation for protein translation | BIOLOGICAL_PROCESS | 0.0179214 | 5 | 33 | 129 | 14810 |
| Female.FW | cytoplasm | CELLULAR_COMPONENT | 0.0179214 | 40 | 2291 | 94 | 12552 |
| Female.FW | ligase activity, forming aminoacyl-tRNA and related compounds | MOLECULAR_FUNCTION | 0.0179214 | 5 | 32 | 129 | 14811 |
| Female.FW | ligase activity, forming carbon-oxygen bonds | MOLECULAR_FUNCTION | 0.0179214 | 5 | 32 | 129 | 14811 |
| Female.FW | aminoacyl-tRNA ligase activity | MOLECULAR_FUNCTION | 0.0179214 | 5 | 32 | 129 | 14811 |
| Female.FW | tRNA aminoacylation | BIOLOGICAL_PROCESS | 0.0179214 | 5 | 33 | 129 | 14810 |
| Female.FW | amino acid activation | BIOLOGICAL_PROCESS | 0.0179214 | 5 | 33 | 129 | 14810 |
| Female.FW | tRNA metabolic process | BIOLOGICAL_PROCESS | 0.029074957 | 6 | 64 | 128 | 14779 |
| Female.FW | phenylalanyl-tRNA aminoacylation | BIOLOGICAL_PROCESS | 0.041903784 | 2 | 0 | 132 | 14843 |
| Female.FW | cell | CELLULAR_COMPONENT | 0.041903784 | 63 | 4622 | 71 | 10221 |
| Female.FW | phenylalanine-tRNA ligase activity | MOLECULAR_FUNCTION | 0.041903784 | 2 | 0 | 132 | 14843 |
| Female.FW | endodermal digestive tract morphogenesis | BIOLOGICAL_PROCESS | 0.041903784 | 2 | 0 | 132 | 14843 |
| Female.FW | cell part | CELLULAR_COMPONENT | 0.041903784 | 63 | 4622 | 71 | 10221 |
| Female.FW | intracellular part | CELLULAR_COMPONENT | 0.041903784 | 55 | 3827 | 79 | 11016 |
| Female.FW | UDP-glucose:glycoprotein glucosyltransferase activity | MOLECULAR_FUNCTION | 0.041903784 | 2 | 0 | 132 | 14843 |
| Female.FW | cellular process | BIOLOGICAL_PROCESS | 0.041903784 | 71 | 5412 | 63 | 9431 |
| Female.FW | intracellular | CELLULAR_COMPONENT | 0.049092922 | 55 | 3861 | 79 | 10982 |
| Female.NW | muscle cell development | BIOLOGICAL_PROCESS | 1.80E-12 | 15 | 87 | 91 | 14784 |
| Female.NW | striated muscle cell development | BIOLOGICAL_PROCESS | 1.80E-12 | 15 | 86 | 91 | 14785 |
| Female.NW | myofibril assembly | BIOLOGICAL_PROCESS | 8.63E-11 | 13 | 75 | 93 | 14796 |
| Female.NW | striated muscle cell differentiation | BIOLOGICAL_PROCESS | 8.63E-11 | 15 | 121 | 91 | 14750 |
| Female.NW | muscle cell differentiation | BIOLOGICAL_PROCESS | 1.25E-10 | 15 | 128 | 91 | 14743 |
| Female.NW | sarcomere | CELLULAR_COMPONENT | 1.67E-09 | 12 | 77 | 94 | 14794 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Female.NW | supramolecular complex | CELLULAR_COMPONENT | 1.67E-09 | 16 | 196 | 90 | 14675 |
| Female.NW | supramolecular polymer | CELLULAR_COMPONENT | 1.67E-09 | 16 | 194 | 90 | 14677 |
| Female.NW | actomyosin structure organization | BIOLOGICAL_PROCESS | 1.67E-09 | 13 | 103 | 93 | 14768 |
| Female.NW | myofibril | CELLULAR_COMPONENT | 1.99E-09 | 12 | 83 | 94 | 14788 |
| Female.NW | contractile fiber part | CELLULAR_COMPONENT | 1.99E-09 | 12 | 82 | 94 | 14789 |
| Female.NW | contractile fiber | CELLULAR_COMPONENT | 1.99E-09 | 12 | 83 | 94 | 14788 |
| Female.NW | A band | CELLULAR_COMPONENT | 1.23E-08 | 7 | 11 | 99 | 14860 |
| Female.NW | cellular component assembly involved in morphogenesis | BIOLOGICAL_PROCESS | 1.23E-08 | 13 | 129 | 93 | 14742 |
| Female.NW | supramolecular fiber | CELLULAR_COMPONENT | 1.23E-08 | 15 | 194 | 91 | 14677 |
| Female.NW | muscle structure development | BIOLOGICAL_PROCESS | 9.56E-08 | 16 | 272 | 90 | 14599 |
| Female.NW | sarcomere organization | BIOLOGICAL_PROCESS | 3.15E-07 | 9 | 55 | 97 | 14816 |
| Female.NW | skeletal myofibril assembly | BIOLOGICAL_PROCESS | 4.20E-07 | 5 | 3 | 101 | 14868 |
| Female.NW | organelle assembly | BIOLOGICAL_PROCESS | 1.34E-06 | 13 | 198 | 93 | 14673 |
| Female.NW | muscle thin filament assembly | BIOLOGICAL_PROCESS | 1.60E-06 | 5 | 5 | 101 | 14866 |
| Female.NW | adult somatic muscle development | BIOLOGICAL_PROCESS | 1.60E-06 | 6 | 14 | 100 | 14857 |
| Female.NW | skeletal muscle myosin thick filament assembly | BIOLOGICAL_PROCESS | 4.36E-06 | 4 | 1 | 102 | 14870 |
| Female.NW | striated muscle myosin thick filament assembly | BIOLOGICAL_PROCESS | 4.36E-06 | 4 | 1 | 102 | 14870 |
| Female.NW | supramolecular fiber organization | BIOLOGICAL_PROCESS | 6.40E-06 | 13 | 232 | 93 | 14639 |
| Female.NW | muscle contraction | BIOLOGICAL_PROCESS | 2.89E-05 | 6 | 26 | 100 | 14845 |
| Female.NW | actin cytoskeleton organization | BIOLOGICAL_PROCESS | 4.46E-05 | 13 | 278 | 93 | 14593 |
| Female.NW | myosin filament assembly | BIOLOGICAL_PROCESS | 5.12E-05 | 4 | 4 | 102 | 14867 |
| Female.NW | actin filament-based process | BIOLOGICAL_PROCESS | 7.67E-05 | 13 | 294 | 93 | 14577 |
| Female.NW | muscle system process | BIOLOGICAL_PROCESS | 8.63E-05 | 6 | 33 | 100 | 14838 |
| Female.NW | structural constituent of muscle | MOLECULAR_FUNCTION | 1.10E-04 | 5 | 17 | 101 | 14854 |
| Female.NW | non-membrane-bounded organelle | CELLULAR_COMPONENT | 5.72E-04 | 25 | 1251 | 81 | 13620 |
| Female.NW | intracellular non-membrane-bounded organelle | CELLULAR_COMPONENT | 5.72E-04 | 25 | 1251 | 81 | 13620 |
| Female.NW | myosin filament organization | BIOLOGICAL_PROCESS | 0.001041773 | 4 | 12 | 102 | 14859 |
| Female.NW | anatomical structure formation involved in morphogenesis | BIOLOGICAL_PROCESS | 0.001205333 | 14 | 450 | 92 | 14421 |
| Female.NW | myosin complex | CELLULAR_COMPONENT | 0.002402313 | 5 | 36 | 101 | 14835 |
| Female.NW | tissue development | BIOLOGICAL_PROCESS | 0.003135179 | 21 | 1035 | 85 | 13836 |
| Female.NW | somatic muscle development | BIOLOGICAL_PROCESS | 0.003667131 | 6 | 70 | 100 | 14801 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Female.NW | cytoskeleton organization | BIOLOGICAL_PROCESS | 0.004851486 | 15 | 588 | 91 | 14283 |
| Female.NW | actin cytoskeleton | CELLULAR_COMPONENT | 0.005463022 | 7 | 115 | 99 | 14756 |
| Female.NW | single-organism organelle organization | BIOLOGICAL_PROCESS | 0.007631604 | 14 | 541 | 92 | 14330 |
| Female.NW | actin-dependent ATPase activity | MOLECULAR_FUNCTION | 0.00827667 | 3 | 7 | 103 | 14864 |
| Female.NW | collagen trimer | CELLULAR_COMPONENT | 0.010052005 | 2 | 0 | 104 | 14871 |
| Female.NW | Z disc | CELLULAR_COMPONENT | 0.012809765 | 5 | 55 | 101 | 14816 |
| Female.NW | I band | CELLULAR_COMPONENT | 0.013560177 | 5 | 56 | 101 | 14815 |
| Female.NW | flight behavior | BIOLOGICAL_PROCESS | 0.019818808 | 4 | 31 | 102 | 14840 |
| Female.NW | flight | BIOLOGICAL_PROCESS | 0.021919929 | 3 | 11 | 103 | 14860 |
| Female.NW | developmental process | BIOLOGICAL_PROCESS | 0.026149506 | 38 | 3005 | 68 | 11866 |
| Female.NW | structural molecule activity | MOLECULAR_FUNCTION | 0.028546329 | 9 | 270 | 97 | 14601 |
| Female.NW | cytoskeletal part | CELLULAR_COMPONENT | 0.043255702 | 13 | 573 | 93 | 14298 |
| Female.NW | single-organism developmental process | BIOLOGICAL_PROCESS | 0.044807148 | 37 | 2977 | 69 | 11894 |
| Female.NW | anatomical structure development | BIOLOGICAL_PROCESS | 0.044807148 | 36 | 2864 | 70 | 12007 |

184

185 **Table S2.** Results from mixed models for proximate measures of reproductive
186 output, body condition and body size.

| | *N* | | $\chi^2_1$ | *df* | P-value |
|---|---|---|---|---|---|
| ln testes mass (mg)[1] | 139 | | | | |
| | | Morph | 8.800 | 1 | **0.003** |
| | | ln pronotum length | 0.875 | 1 | 0.350 |
| | | ln somatic mass | 33.841 | 1 | **<0.001** |
| Egg mass (mg)[2] | 145 | | | | |
| | | Morph | 0.011 | 1 | 0.916 |
| | | ln pronotum length | 1.190 | 1 | 0.275 |
| | | ln somatic mass | 2.688 | 1 | 0.101 |
| ln somatic SMI (mg)[1] | 284 | | | | |
| | | Sex | 14.071 | 1 | **<0.001** |
| | | Morph | 5.095 | 1 | **0.024** |
| | | Sex × Morph | 14.006 | 1 | **<0.001** |

187 All mixed models included a random effect of biological line, and pronotum length
188 and somatic mass measures were standardized. ln indicates natural log.
189 [1] LMM
190 [2] Negative binomial GLMM
191

192

**Table S3.** Results from LMMs for measures of pronotum length and somatic mass.

| | *N* | | $\chi^2_1$ | *df* | P-value |
|---|---|---|---|---|---|
| Pronotum length (mm) | 284 | | | | |
| | | Sex | 6.925 | 1 | **0.008** |
| | | Morph | 1.109 | 1 | 0.292 |
| | | Sex × Morph | 0.007 | 1 | 0.934 |
| Somatic mass (mg) | 284 | | | | |
| | | Sex | 5.521 | 1 | **0.019** |
| | | Morph | 0.000 | 1 | 0.998 |
| | | Sex × Morph | 6.493 | 1 | **0.011** |

All mixed models included a random effect of biological line.

**Table S4.** Summary of the number of sequenced reads before and after trimming.

| Sample | Untrimmed | Trimmed | R1/R2 |
|---|---|---|---|
| 12-M-head-NW | 61,006,572 | 60,574,658 | 30,135,313 |
| 13-M-musc-NW | 52,305,462 | 51,887,011 | 25,750,536 |
| 14-M-testes-NW | 32,528,962 | 32,272,950 | 16,016,686 |
| 15-M-head-NW | 45,129,482 | 44,800,976 | 22,253,685 |
| 16-M-musc-NW | 38,802,556 | 38,548,881 | 19,152,513 |
| 17-M-testes-NW | 45,673,402 | 45,345,163 | 22,526,496 |
| 18-M-head-NW | 45,144,822 | 44,759,087 | 22,255,296 |
| 19-M-musc-NW | 41,938,996 | 41,687,713 | 20,723,620 |
| 20-M-testes-NW | 46,192,918 | 45,807,083 | 22,753,029 |
| 21-M-head-FW | 41,271,576 | 41,009,761 | 20,390,600 |
| 22-M-musc-FW | 38,299,684 | 37,927,595 | 18,827,916 |
| 23-M-testes-FW | 34,296,850 | 34,040,366 | 16,913,615 |
| 24-M-head-FW | 38,464,592 | 38,194,020 | 18,981,007 |
| 25-M-musc-FW | 40,841,770 | 40,536,785 | 20,167,484 |
| 26-M-testes-FW | 56,988,020 | 56,698,696 | 28,215,093 |
| 27-M-head-FW | 35,047,354 | 34,844,351 | 17,324,645 |
| 28-M-musc-FW | 31,835,044 | 31,523,465 | 15,689,092 |
| 29-M-testes-FW | 35,387,996 | 35,209,052 | 17,526,777 |
| 30-F-head-NW | 42,161,524 | 41,907,832 | 20,851,634 |
| 31-F-musc-NW | 35,282,270 | 45,057,779 | 17,431,719 |
| 32-F-ovary-NW | 41,918,270 | 41,671,898 | 20,720,037 |
| 33-F-head-NW | 54,750,504 | 54,366,572 | 27,009,996 |
| 34-F-musc-NW | 36,707,644 | 36,478,463 | 18,133,920 |
| 35-F-ovary-NW | 61,241,274 | 60,568,124 | 30,140,953 |
| 36-F-head-NW | 40,529,344 | 40,286,894 | 20,029,845 |
| 37-F-musc-NW | 60,875,116 | 60,420,336 | 30,002,776 |
| 38-F-ovary-NW | 49,069,420 | 48,778,590 | 24,262,153 |
| 39-F-head-FW | 68,246,640 | 67,735,805 | 33,665,646 |
| 40-F-musc-FW | 53,574,414 | 53,152,845 | 26,401,475 |
| 41-F-ovary-FW | 60,275,904 | 59,672,380 | 29,544,754 |
| 42-F-head-FW | 49,036,456 | 48,319,730 | 23,843,272 |
| 43-F-musc-FW | 46,013,420 | 45,292,669 | 22,295,013 |
| 44-F-ovary-FW | 72,280,210 | 71,502,322 | 35,382,796 |
| 45-F-head-FW | 72,769,140 | 71,884,165 | 35,552,881 |
| 46-F-musc-FW | 50,181,870 | 49,502,672 | 24,456,228 |
| 47-F-ovary-FW | 58,764,358 | 58,047,104 | 28,721,838 |

197 **Table S5.** Summary statistics for full and filtered Trinity *de novo* transcriptome assemblies (longest isoform only)

| | Full assembly | Filtered clustered assembly |
|---|---|---|
| Total Trinity genes | 2,573,977 | 42,496 |
| N10 | 2,129 | 6,764 |
| N20 | 1,151 | 5,106 |
| N30 | 729 | 4,147 |
| N40 | 522 | 3,446 |
| N50 | 411 | 2,900 |
| Median contig length | 290 | 1,859 |
| Mean contig length | 413.35 | 2,293 |
| Total bases | 1,063,945,697 | 302,614,475 |

198

199