ELUCIDATING THE EVOLUTION OF SEX CHROMOSOME DOSAGE

COMPENSATION USING BEETLES

by

BALAN RAMESH

Presented to the Faculty of the Graduate School of

The University of Texas at Arlington in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT ARLINGTON

MAY 2021

Acknowledgements

I am grateful to the many people who entered my life since I began this dissertation.
Thank you to my close friends and lab mates in the Demuth Lab: Catherine Rogers,
Shana Pau, and Beena Margabandu. Marquerite Herzog deserves a special mention,
as we both started our research simultaneously and for sharing her experience in
research and life, guiding me to navigate life away from home. This lab group greatly
enriched my life with its intelligence, compassion, patience, and encouragement, and
my research would not have been possible without them. Thank you to the many friends
and colleagues I have met since starting at UTA, especially Thomas Firneno Jr.,
Fatema Ruma, Ayda Mirselahi, Adnan Qureshi, Nazmeen Qureshi, Suman Shrestha,
Danielle Rivera, Sruthi Srinivasan and Jees Augustine. To my committee members,
Drs. Esther Betran, Todd Castoe, Matt Fujita, and Sen Xu: I am grateful for the
encouragement, knowledge, and ideas you provided, which improved my research
immensely. Thank you to the biology department's administrative staff - Gloria
Burlingham, Linda Taylor, Chris Magno, Ashley Priest, and, especially, Mallory Roelke –
for all of the help. Special thanks to Rachel Wostl and Anna Williford for helping me with
the rewarding experience of teaching others. I am forever grateful for the support of my
parents, Ramesh and Geetha and my fiancé, Srividya. Finally, no one will ever replace
the unparalleled mentorship of my advisor, Jeff Demuth, which has inspired me and will
aspire towards.

Dedication

To my mom, Geetha. Thank you for your unwavering support, kindness, and inspiration as I pursued this journey.

Abstract

ELUCIDATING THE EVOLUTION OF SEX CHROMOSOME DOSAGE

COMPENSATION USING BEETLES

Balan Ramesh, Ph. D.

The University of Texas at Arlington, 2021

Supervising Professor: Jeffrey P. Demuth, PhD

Sex chromosomes are unique because of their role in sex determination. More importantly, sex chromosomes spend an unequal amount of time between the sexes which has significant consequences since it provides an opportunity to understand how evolutionary forces act on the genome. For example, in the XY system of sex determination, males are heterogametic and carry XY chromosomes, whereas females are homogametic and carry XX chromosomes. This difference in the gene dose of X-linked genes between the sexes can be catastrophic for interacting proteins if not balanced. In chapter 1, I introduce the background and current state of understanding for sex chromosome evolution and regulations. In chapter 2, I employ phylogenomics methods to established the phylogeny and divergence time among a group of closely related flour beetles, which provides an evolutionary framework for answering important questions about sex chromosome dosage balance between sexes and compensation considering ancestral state of expression. I find that the broad horned beetle (*Gnatocerus cornutus*) split from the flour beetles around 122 Million Years Ago (MYA) and the most recent split within flour beetles is around 14 MYA between *T.castaneum* and *T.freemani*. In chapter 3, I utilize an X to Autosome fusion in confused flour beetles

(*Tribolium confusum*) combined with the divergence times established in chapter 1 to study the evolution of dosage compensation in flour beetles by reconstructing the ancestral states of neoX-linked gene expression. I report that the neoX and X chromsomes are dosage balanced and compensated. In addition, the expression between the sexes is more tightly regulated and thus dosage balance seems to be more constrained comparing the ancestral autosomal expression with the current state of expression of the sex chromosomes. Lastly, in chapter 4 I employ single-cell RNA-Seq to explore the presence of meiotic sex chromosome inactivation, which affects the expression of X-linked genes during male meiosis. I find that the expression of X-linked genes is repressed or inactivated in most cell clusters hinting towards the presence of MSCI. Comparing the current state of expression with ancestral levels and differentiating expression profile between different types of cells, this dissertation exemplifies the use of comparative transcriptomics to address essential questions in the evolution of sex chromosomes and dosage compensation.

Table of Contents

Chapter 1

Introduction

In eukaryotic organisms, the two sexes often have different chromosomal compositions. Generally, one sex (called heterogametic) has a pair of morphologically distinct sex chromosomes, whereas the other sex (called homogametic) has two identical chromosomal pairs. In male heterogametic systems, the sex chromosomes are assigned as X and Y chromosomes (XX females and XY males), and in female heterogametic systems as Z and W chromosomes (ZZ males and ZW females). For simplicity, when discussing general aspects of sex chromosome evolution, I will refer to the XY chromosome systems, but the same evolutionary theories also apply to ZW systems.

In the XY system of sex determination, the X and Y chromosomes have evolved independently many times in plants and animals with different karyotypes. Sex chromosomes are thought to have originated from an pair of homologous chromosomes (autosomes). The morphological distinction between the X and Y chromosome is a by-product of the Y chromosome's degeneration that is present only in the heterogametic sex (i.e., XY males). Thus the lack of homology between the X and Y chromosome drastically reduced the genetic recombination on the Y chromosome. Evolutionary theory predicts that nonrecombining regions of a genome are more susceptible to an accumulation of deleterious mutations. In line with the theory, a predominant feature of many taxa is the erosion of genes from the Y chromosome. The pseudogenization of genes leaves primarily nonfunctional and repetitive DNA on the Y chromosome. In

contrast, due to recombination in females, X chromosomes retain most of their ancestral genes but have evolved mechanisms to comensate for their dosage imbalance between sexes and relative to autosomes.

First, the loss of genes on the Y chromosome leads to a gene dose deficiency for X-linked genes in males. Ohno (1967) suggested that this difference in gene dose between the sexes imposes a "peril of hemizygosity" because it may affect the stoichiometric equilibrium between members of multi-protein complexes that are crucial for cellular development and function (Bachtrog, 2006; Charlesworth et al., 2005; Furman et al., 2020; Gu and Walters, 2017). As a solution to this peril, Ohno suggested that "dosage compensation" (DC) mechanisms must evolve so that at the level of average gene expression, X=XX=AA. Although his paradigm has dominated sex chromosome research for over 50 years, mounting evidence demonstrates that Ohno's vision of DC is far from ubiquitous and maybe an exception rather than the rule.

Comparison of X to autosome ratios within and between sexes from a growing number of taxa can be categorized into four patterns (Gu and Walters., 2017). Note, however, that categorization of some species is still plagued by discrepancies potentially arising from different data types and/or tissues.

Type I (X=XX=Ancestral): *Complete dosage compensation*. The pattern envisioned by Ohno as the "rule" of DC. Under this pattern, males evolve mechanisms to *compensate* their haploid X-linked content to match ancestral diploid levels, while females evolve mechanisms to counter hyperexpression of the X so that they maintain ancestral diploid expression levels. Reported only in XY systems: aphids, stink bugs and milkweed bugs (Pal and Vicoso, 2015), fruit flies(Vicoso and Bachtrog, 2015),

stickleback fish (Leder et al., 2010), Anolis lizards (Rupp et al., 2016), and perhaps beetles (Mahajan and Bachtrog, 2015).

Type II (X=XX<Ancestral): *Dosage balance.* Under this pattern, females down-regulate X-linked gene expression so that the X is *balanced* between sexes, but both sexes have lower X-linked expression than the ancestral autosomal state. Reported in: worms (XY) (Wheeler et al., 2016) and placental mammals (XY) (Chen and Zhang, 2016; Julien et al., 2012), and butterflies (ZW)(Walters et al., 2015; Walters and Hardcastle, 2011).

Type III (X<XX=Ancestral): *Lack of dosage compensation.* On average, expression levels reflect copy number dosage. However, within the overall Type III pattern, local regulation of individual genes or gene regions may follow Type I or Type II. This type of compensation is reported in the sole (XY) (Graves, 2014), platypus (5X,5Y) (Julien et al., 2012), snakes (ZW) (Schield et al., 2019; Vicoso et al., 2013), birds (ZW) (Mank and Ellegren, 2009), *Gallus* (ZW) (Zimmer et al., 2016), and *Schistosoma* (ZW) (Vicoso and Bachtrog, 2011).

Type IV (X = Ancestral < XX): *Incomplete dosage compensation.* Here DC is incomplete because females do not counteract elevated X chromosome expression that presumably evolved to compensate for hemizygous males. Reported in: *Tribolium castaneum* (XY) (Prince et al., 2010).

Ohno's original DC hypothesis and decades of subsequent DC research are founded on comparing average X and autosome expression ratios (X: AA) within and between sexes. The use of the X: AA ratio coupled with the use of different tissues and/or methodology in different studies contributes to uncertainties in the classification

3

of the type of dosage compensation. For example, the use of tissue samples that include gonads can be problematic because gonads are prone to sex-biased expression. Also, the gene expression in gonads may further be accentuated by the lack of dosage compensation and mechanisms such as meiotic sex chromosome inactivation (MSCI) during spermatogenesis (Turner, 2007).

Further, if some autosomes are more likely to become sex chromosomes since their pattern of sex-biased expression is higher than the average autosome then using average autosomal expression as a proxy for ancestral expression can be dubious. So, to truly understand dosage compensation, comparing X-linked gene expression in the somatic tissue to their ancestral autosomal state to directly assess whether there has been regulatory "compensation" for the loss of Y-linked genes is much more desirable.

Using five species of flour beetles (*Tribolium brevicornis, Tribolium castaneum, Tribolium confusum, Tribolium freemani, and Gnatocerus cornutus*), this dissertation establishes the phylogeny with divergence estimates between these species using conserved orthologs (Chapter 2) and sheds light on the type of sex chromosome dosage compensation by estimating ancestral expression of the neoX chromosome in *T. confusum* (Chapter 3). Further, this disseratation adds that the dosage balance is more tightly conserved than dosage compensation. The final chapter explores the presence of MSCI to understand transcriptional dynamics of sex chromosomes in testes using single-cell RNA sequencing (scRNA-Seq).

Chapter 2

Divergence time estimation of genus *Tribolium* by extensive sampling of highly conserved orthologs

Balan Ramesh[1], Thomas J. Firneno, Jr.[1], and Jeffery P. Demuth[1]

[1]*Department of Biology, the University of Texas at Arlington, Texas 76019, USA.*

Abstract

Tribolium castaneum, the red flour beetle, is among the most well-studied eukaryotic genetic model organisms. *Tribolium* often serves as a comparative bridge from highly derived *Drosophila* traits to other organisms. Simultaneously, as a member of the most diverse order of metazoans, Coleoptera, *Tribolium* informs us about innovations that accompany hyper diversity. However, understanding the tempo and mode of evolutionary innovation requires well-resolved, time-calibrated phylogenies, which are not available for *Tribolium*. The most recent effort to understand *Tribolium* phylogenetics used two mitochondrial and three nuclear markers. The study concluded that the genus may be paraphyletic and reported a broad range for divergence time estimates. Here we employ recent advances in Bayesian methods to estimate the relationships and divergence times among *Tribolium castaneum*, *T. brevicornis*, *T. confusum*, *T. freemani*, and *Gnatocerus cornutus* using 1368 orthologs conserved across all five species and an independent substitution rate estimate. We find that the most basal split within *Tribolium* occurred ~86 Mya [95% HPD 85.90–87.04 Mya] and that the most recent split was between *T. freemani* and *T. castaneum* at ~14 Mya [95% HPD 13.55-14.00]. Our results are consistent with broader phylogenetic analyses of insects and suggest that Cenozoic climate changes played a role in the *Tribolium* diversification.

Introduction

Here we present a phylogenetic study of *Tribolium castaneum*, the red flour

beetle, and four closely related species. Beetles in the genus *Tribolium* have a long

history as model systems in ecology (e.g., Park, 1934, Park, 1948, Goodnight, 1990,

Agashe and Bolnick, 2010), evolution (e.g., Wade, 1977, Stevens 1989, Via and

Conner, 1995, Demuth and Wade, 2007, Agashe et al., 2011), behavior (e.g., Ghent,

1963, Sinnock, 1970, Konishi et al., 2020), genetics (e.g., Sokoloff, 1966, Sokoloff,

1977, Beeman and Brown, 1999, Lorenzen et al., 2005) and development (e.g.,

Beeman, 1987, Brown et al., 2009). Furthermore, *T. castaneum* is among the most well-

developed invertebrate systems for molecular genetic studies (Denell, 2008). The *T.

castaneum* genome is well-sequenced, assembled, and annotated (*Tribolium* Genome

Sequencing Consortium, 2008, Kim et al., 2010, Herndon et al., 2020), including an

atlas of functional genetic data based on large-scale RNAi knockdown screens

(Schmitt-Engel et al., 2015). The molecular toolkit includes virtually all modern

techniques for genome manipulation, including parental RNAi (Tomoyasu et al., 2008),

transposon-mediated insertional mutagenesis (Lorenzen et al., 2007, Trauner et al.,

2009), GAL4/UAS system for spatiotemporal expression analysis (Schinko et al., 2010,

Rylee et al., 2018), and CRISPR/Cas9 gene-editing system (Giles et al., 2015). Their

relatively short generation time (~1 month), ease of culture, importance as an

agricultural pest, and catalog of previous work contribute to the large and growing

community of researchers employing *Tribolium* in their work.

Despite *Tribolium's* utility as a model system, there has been little phylogenetic

work that would promote a higher resolution analysis of the tempo and mode of trait

7

evolution. For example, we generated the transcriptome data we employ in the following analyses to investigate the evolution of X-chromosome gene regulation following an X-autosome fusion observed in extant members of the *T. confusum* lineage (Smith, 1950, Smith, 1952, Smith and Brower, 1974). Understanding changes in expression of the neo-X linked genes requires comparing current expression levels to an estimate of the ancestral expression states when those genes were autosomal (i.e., before fusion). However, such analyses are impossible without the context provided by a well-resolved time-calibrated phylogeny. Furthermore, a more precise resolution of species divergence times will provide context for many studies of molecular evolution in this clade. For example, a recent investigation of opsin gene evolution (Wu et al., 2020) would have benefited from the improved dating provided by our analyses below.

The most recent analysis of the *Tribolium* phylogeny was more than a decade ago and included two mitochondrial and three nuclear markers (Angelini and Jockusch, 2008). The analysis included eight of the 36 species in the genus *Tribolium* and three other Tenebrionid beetles (*Gnatocerus cornutus*, *Latheticus oryzae*, *Tenebrio molitor*). To calibrate divergence times, they employed mitochondrial clocks based on crickets (Venanzetti et al., 1993) and milkweed beetles (Farrell, 2001). The results did little to constrain the temporal pattern of diversification among lineages. Divergence time estimates for the sister species *T. castaneum* and *T. freemani* ranged from 11.6–47.0 Mya, and for the older split between *T. castaneum* and *T. confusum,* the estimated time was 13.9–60.7 Mya. They did not estimate other ancestral splits.

The pattern of relationships within *Tribolium* is also uncertain, particularly with respect to the placement of *T. brevicornis.* Phylogenetic analysis using cytochrome

oxidase 1 and 16S rDNA sequences from eight *Tribolium* species did not resolve whether *T. brevicornis* is basal to the *T. castaneum* and *T. confusum* species groups, or more closely related to the *T. castaneum* species group (Meštrović et al., 2006). The more recent work by Angelini and Jockusch (2008) cast further doubt, as their consensus tree suggested that *T. brevicornis* was outside the clade containing *G. cornutus* and the other *Tribolium* species. Based on the apparent paraphyly of the genus *Tribolium*, Angelini and Jockush (2008) recommended that *T. brevicornis* be returned to the genus *Aphanotus* (Leconte, 1862). Thus, the monophyly of the genus *Tribolium* remains in some doubt.

Generating more precise date estimates and resolving the pattern of divergence for *Tribolium* species has important consequences for evolutionary studies. For example, the ~47 million year range in the divergence time estimate for *T. castaneum* and *T. confusum* is sufficiently broad to change inferences about the direction of gene expression evolution following the X-autosome fusion in the example presented above. Here, we present a high-resolution time-calibrated phylogeny of members of the genus *Tribolium* built using 6840 protein-coding genes (1368 orthologs conserved across each of five species). Our results are mostly congruent with those of the earlier study by Angelini and Jockusch (2008) but suggest that the genus *Tribolium* is not paraphyletic. Further, our study's divergence time estimates provide the best temporal parameters for *Tribolium* evolution to date, thus providing a useful framework for future work in this vital system.

Methods

Taxon Sampling, RNA Sequencing, and Ortholog Identification

We collected transcriptome data for five species of Tenebrionid beetles: *T. castaneum*, *T. freemani, T. confusum*, *T. brevicornis,* and *Gnatocerus cornutus.* All beetles were reared at large population sizes on 95% organic whole wheat flour with 5% Brewer's yeast at 27°C with 70% humidity. Briefly, each sample for RNA extraction consisted of ten adult individuals. In total, we sequenced two replicates per sex per species, with ten adults in each replicate. Samples were sequenced on the HiSeq 2000 platform using Illumina 150bp paired-end reads. We obtained ~40 Million reads per sample. We conducted *de novo* transcriptome assembly for each species using Trinity v2.8.4 (Haas et al., 2013), filtered transcript redundancy using cd-hit-est v4.8.1(Fu et al., 2012), and obtained amino acid sequences using TransDecoder v5.5.0 (Haas and Papanicolaou, 2016). We benchmarked these filtered transcriptomes using BUSCO v3.0.2 (Simão et al., 2015) and endopterygota lineage (odb10). Finally, we used the filtered transcriptomes to identify the common orthologs among the five species using Orthofinder v2.3.3 (Emms and Kelly, 2019).

Sequence Alignment and Data Partition

Using our Orthofinder results, we identified highly-conserved, complete, single-copy orthogroups. We used MAFFT v7.471 (Nakamura et al., 2018) to align the orthogroups. Orthogroups in which all orthologs had at least 99% of the sites present (i.e., a maximum of 1% of the total length as gaps was permitted) in all species were retained for subsequent steps. Each of the retained orthogroups was used as a separate partition in divergence time estimation.

Divergence Time Estimation

We estimated the phylogenetic relationships and associated divergence times of individuals using our orthogroup alignment via a StarBEAST2 template (Ogilvie et al., 2018) in BEAST2 v2.6.3 (Bouckaert et al., 2014). The site models were unlinked while the clock and the tree models were linked among all orthogroups. We performed analyses using an uncorrelated lognormal clock model, a Yule tree prior, and the Blosum62 substitution model. To model site variation using a gamma distribution, we set four gamma categories. The shape and proportion invariant parameter of the site model was set to estimate and zero, respectively.

We used the substitution rate of 6.43E-4 $\square$ 1.5E-5 amino acid substitutions per site per million years to calibrate the tree. This estimate is derived from a previous maximum likelihood (ML) analysis of 439 nuclear genes conserved among *T. castaneum, Drosophila melanogaster,* and *Anopheles gambiae,* and calibrated using fossils from the oldest known coleopteran (Savard et al., 2006). The BEAST analysis was run on CIPRES Science Gateway for 10,000,000 generations with sampling every 1000 generations, producing 10000 trees. The initial 10% of runs were discarded as burn-in, and the remaining 9000 runs were assessed locally using Tracer v1.7.1 (Rambaut et al., 2018) to examine the estimated sample size (ESS) value and the trace output for convergence. A maximum clade credibility (MCC) tree with median heights was created using Treeannotator v2.6.2 (Helfrich et al., 2018) for the remaining 9000 trees.

Results

BUSCO analysis of the *de novo* transcriptome assemblies indicated that our transcriptomes are > 97.5% complete, suggesting that our initial transcriptomics-based

gene sampling was unbiased. Using the predicted protein sequences from the filtered

transcriptome assemblies, OrthoFinder identified 4286 single-copy orthogroups, of

which 1368 had orthologs with at least 99% of the sites present among the five species.

The median length of the alignment for the orthogroups was 460 amino acid sites, with

few orthogroups having more than 2000 amino acid sites (Supplementary Table T1;

Supplementary Figure S1). Our filtered 1368 orthogroups includes proteins from the 28s

ribosomal DNA gene, which was one of the three nuclear genes included in Angelini

and Jockusch (2008). Subsequent phylogenetic analyses employed only these 1368

orthologs except where noted.

     After 10,000,000 generations of sampling in BEAST, with 10% burnin, the ESS

values for topology and divergence time estimates were 350.2 and 290.9 respectively

and had traces oscillating in the stationary distribution, indicating that the analysis

reached convergence. The resulting Bayesian phylogenetic tree rooted using midpoint

strongly supports (posterior probability = 1) monophyly of the genus *Tribolium* with

respect to *G. cornutus* (Figure 1), which was a point of uncertainty in the previous work

by Angelini and Jockusch (2008). Our divergence time analysis using the calibrated

substitution rate supports an early-Cretaceous divergence (121.95 Mya [95% HPD

121.21–122.68 Mya]) between *G. cornutus* and *Tribolium* species (Figure 1, Table 1).

The split between *T. brevicornis* and the rest of the *Tribolium*, as well the split

between *T. confusum* and the *T. castaneum* group, were estimated to have occurred

during the late-Cretaceous and Paleogene respectively (86.46 Mya [95% HPD 85.90–

87.04 Mya] and 65.5 Mya [95% HPD 64.98–65.96], respectively). The divergence

between *T. castaneum* and *T. freemani* was estimated to occur during the Neogene

(13.8 Mya [95% HPD 13.55-14.00]).

The amino acid sequences, XML file, and the result files from BEAST analysis

are available in the Dryad repository (https://doi.org/10.5061/dryad.pc866t1mz).

Discussion

Our results establish a reliable time of divergence among *Tribolium* species. We

included four *Tribolium* species from three subgroups - *brevicornis*, *confusum*, and

castaneum (Angelini and Jockusch, 2008), and *G. cornutus*. We analyzed 1368 highly

conserved nuclear protein-coding genes comprised of over 650,000 amino acid sites,

thereby significantly increasing the number of loci from previous studies. While the

topology of our phylogenetic tree is similar to the most parsimonious tree produced by a

combined analysis of the five genes (two mitochondrial and three nuclear) from the

previous research (Angelini and Jockusch, 2008), our expanded gene sampling

supports the monophyly of *Tribolium*, unlike the Maximum Likelihood and Bayesian

Inference consensus tree suggested by Angelini and Jockusch, (2008). However,

considering the genus *Tribolium* as a monophyletic group must be approached with

caution due to our limited taxonomic sampling.

To investigate the effect of including additional taxa on our inference of

monophyly, we reran our analysis pipeline with transcriptomes from two additional

Tenebrionid beetles sampled in the study by McKenna et al., (2019), *Neomida*

bicornis and *Lagria hirta*. The resulting 7-species tree (Supplementary Figure S2) is

consistent with the expected placement of *L. hirta* and *N. bicornis* relative to *T.*

castaneum and *G. cornutus* based on previous work (Kergoat et al., 2014a, McKenna et

al., 2019). *N. bicornis* is sister to *G. cornutus*, and that group is sister to all *Tribolium*

spp. *L. hirta* is an outgroup to all other species in the analysis. Thus, the 7-species

analysis reinforces our inference of monophyly for the genus *Tribolium*, which is also

consistent with morphological considerations that originally prompted members of the

brevicornis species group to be included in the genus *Tribolium* (Hinton 1948). Future

analyses of *T. brevicornis* status may benefit from sampling additional members of the

tribe Triboliini (e.g., *Latheticus oryzae*), which are even more closely related to the

genus *Tribolium*.

Including the additional Tenebrionid species reduces the number of orthogroups

that pass filtering by approximately 10 fold, from 1368 to 138. Consequently, although

the dating is similar, the 7-species analysis yields much broader confidence intervals.

Since a primary goal of our study was to produce more precise estimates for *Tribolium*

divergence, we include results of the 7-species analysis as Supplementary materials

(Supplementary Table T1, Table T2, Figures S2, and S3) and restrict further discussion

to the higher-resolution analysis of fewer species but a much larger set of orthogroups.

The divergence times of *Tribolium* estimated here fit well within the context of molecular

phylogenic studies that included more course sampling of beetle diversity. Our choice to

include only highly conserved genes had the potential to bias divergence estimates if

the retained genes represented consistently more slowly evolving genes than those

used in the estimate of substitution rate by Savard et al., (2006). However, the

consistency of our estimates with the broader Coleopteran (McKenna et al. 2019) and

Tenebrionid (Kergoat et al., 2014b) studies that used independent methods of

inference, including fossil calibration, suggests that our choice of filtering criteria and

substitution rate are unlikely to be substantially biasing on our inferences of divergence time. For example, previous estimates for the basal split between *Gnatocerus* and *Tribolium* range from (65 to 165 Mya; Table 1). Our results suggest that the divergence is likely intermediate (~122 Mya).   Other nodes in our analysis are similarly consistent with expectations from the other studies, but we provide much narrower confidence intervals (Table 1; see also Supplementary Table T2 for 7-species analysis, which provides an additional point of comparison).

Our estimated divergence dates suggest that the genus *Tribolium* began to diversify during the Mesozoic. However, given that much of the present-day species diversity arose subsequent to the split between *T. confu*sum and *T. castaneum,* our dates suggest that the radiation of *Tribolium* primarily occurred during the Cenozoic. Diversification of insects during the Cenozoic is well supported in Lepidopterans (Wahlberg et al., 2009; Toussaint et al., 2012; Halali et al., 2020), Dipterans (Esseghir et al., 2000; Junqueira et al., 2016), Hemipterans (Peccoud et al., 2010; Rui-ling et al., 2013), and even other Coleopterans (Matthews, 1980; Hidalgo-Galiana and Ribera, 2011). However, Kergoat et al., (2014b) found that net diversification rates in the Tenebrionidae were likely reduced in the Cenozoic following an increased extinction rate that was not fully compensated by speciation in the mid-Cretaceous (near our estimated date for the split between lineages leading to *Tribolium* and *Gnatocerus*). The drastic changes in the global temperatures and elevated tectonic activity during the Cenozoic era (Matthews, 1980) may also have played a role in the radiation and distribution of the *Tribolium* beetles. However, deducing a specific reason for the cause of divergence remains challenging.

Overall, our study provides high-resolution divergence estimates for *Tribolium* flour beetles, including the important model system *Tribolium castaneum*. Divergence times in the genus suggest that most of the species diversity arose during the Cenozoic era. The phylogenetic analysis provided here improves the utility of *T. castaneum* as a model system by providing a framework to understand the tempo of evolution within its relatively species-rich genus. Our results will contribute to future studies of genetic and organismal diversification focused on the hypotheses involved in the development and evolution of *Tribolium* beetles.
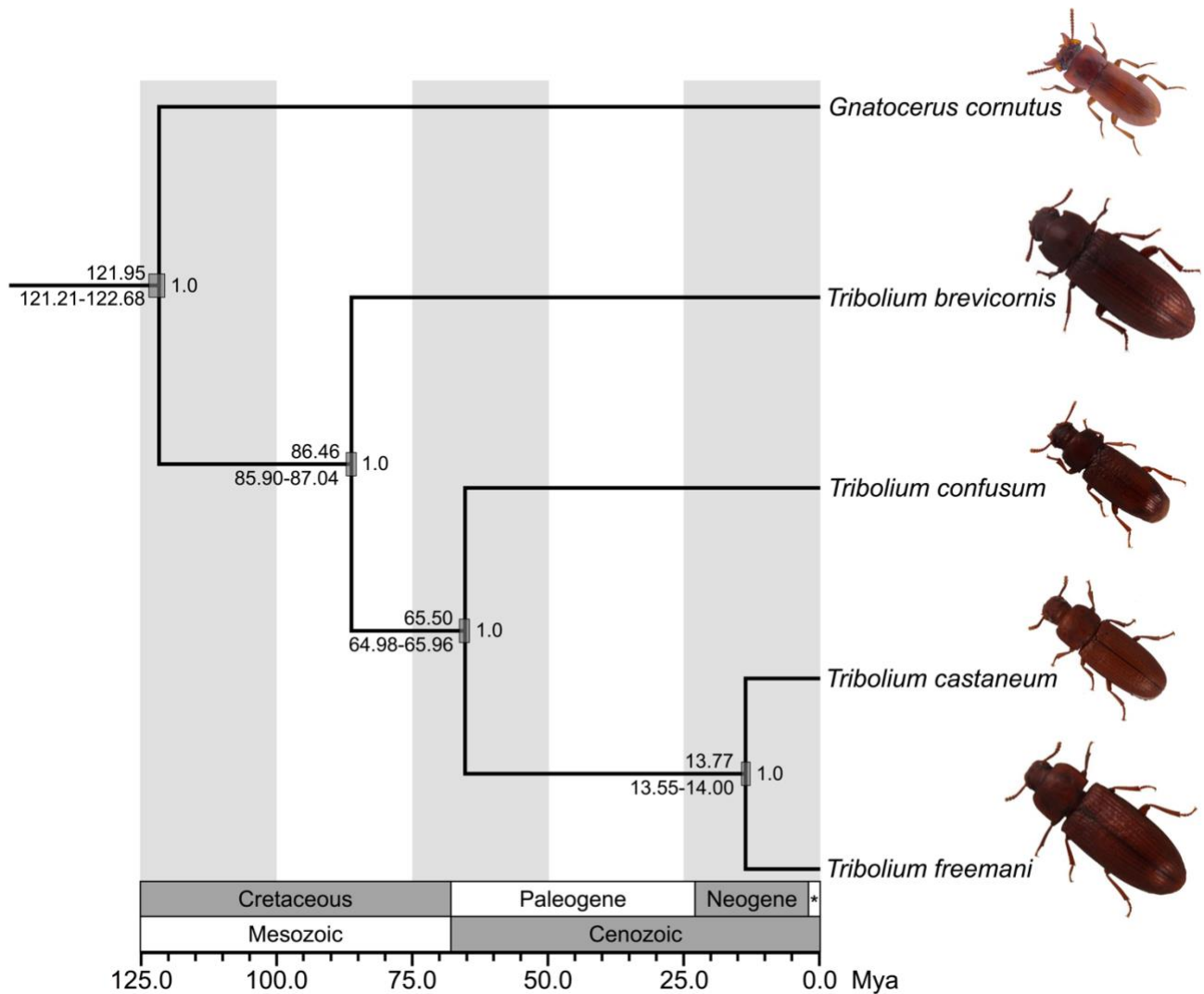
Acknowledgments

Figure 1. Chronogram of the four focal species based on a BEAST2 analysis of highly conserved nuclear protein-coding genes calibrated with amino acid substitution rate. * denotes the Quaternary period. Posterior probabilities are indicated to the right of the ancestral nodes. Median ages on the chronogram are provided above nodes, with 95% highest posterior densities (HPD) below and error bars representing the 95% HPD on the nodes. All *Tribolium* photos courtesy of Heath Blackmon (Texas A&M University, College Station, TX)

Table 1. Comparison of divergence times for nodes/clades shared across four studies. All the estimates are represented in Million Years ago (Mya). Brackets indicate 95% highest posterior densities (HPD).

Nodes/Clades	Divergence times	Previous studies
T. castaneum – T. freemani	13.77 [13.55 – 14.0]	11.6 – 47.0[a]
T. castaneum/T. freemani – T. confusum	65.50 [64.98.9 – 65.96]	13.9 – 60.7[a] 70 – 120[b]
T. castaneum/T. confusum – T. brevicornis	86.46 [85.9 – 87.04]	-
Tribolium – Gnatocerus	121.95 [121.21 – 122.68]	148 – 165[b] 65 – 129[c]

a: Angellini & Jockush, 2008.

b: Kergoat et al., 2014b. (95% HPD for "Yule crown" calibration)
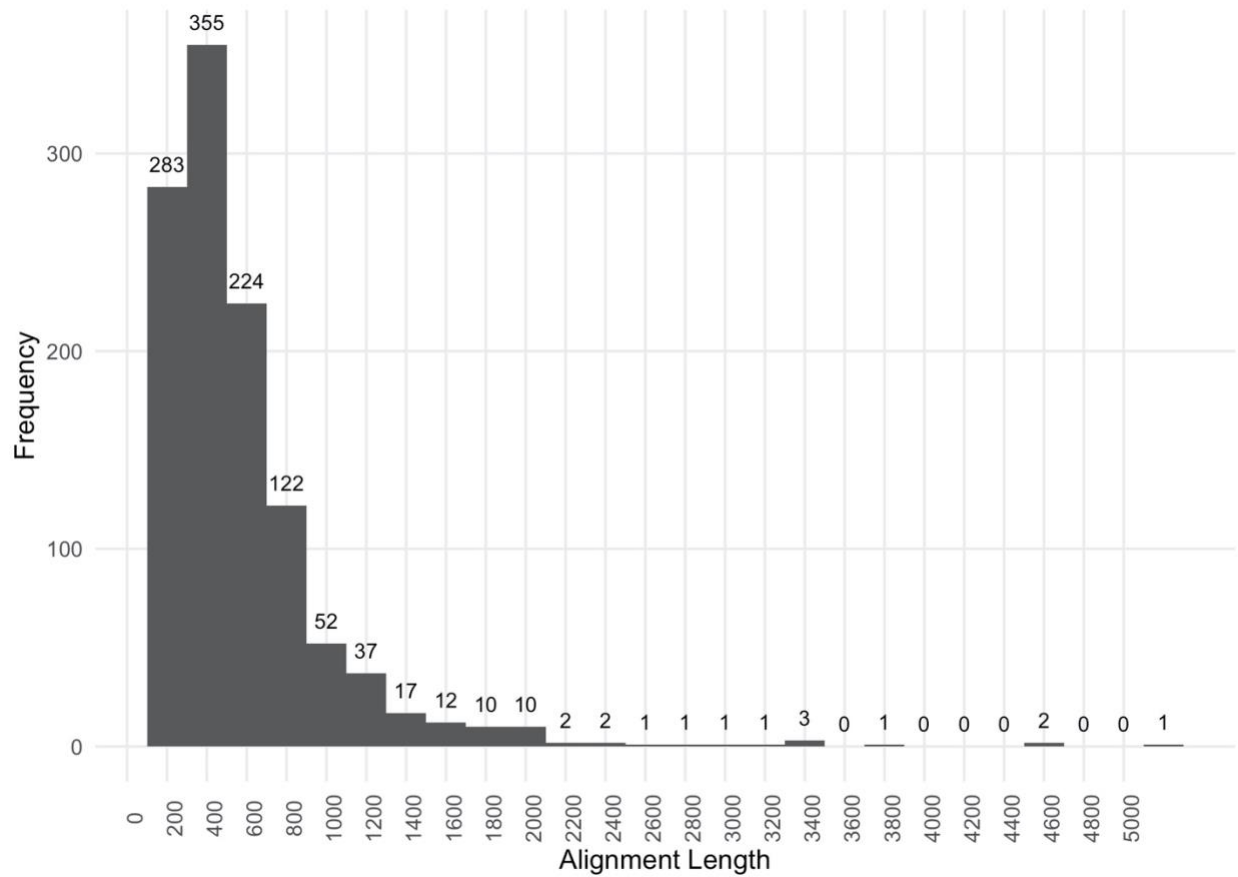
c: McKenna et al., 2019.

Figure S1. Frequency of Alignment Length from 1368 orthogroups for the 5-species tree. Each bin is 200 bp in width, and the value on the top of the bar indicates the number of orthogroups in a bin.
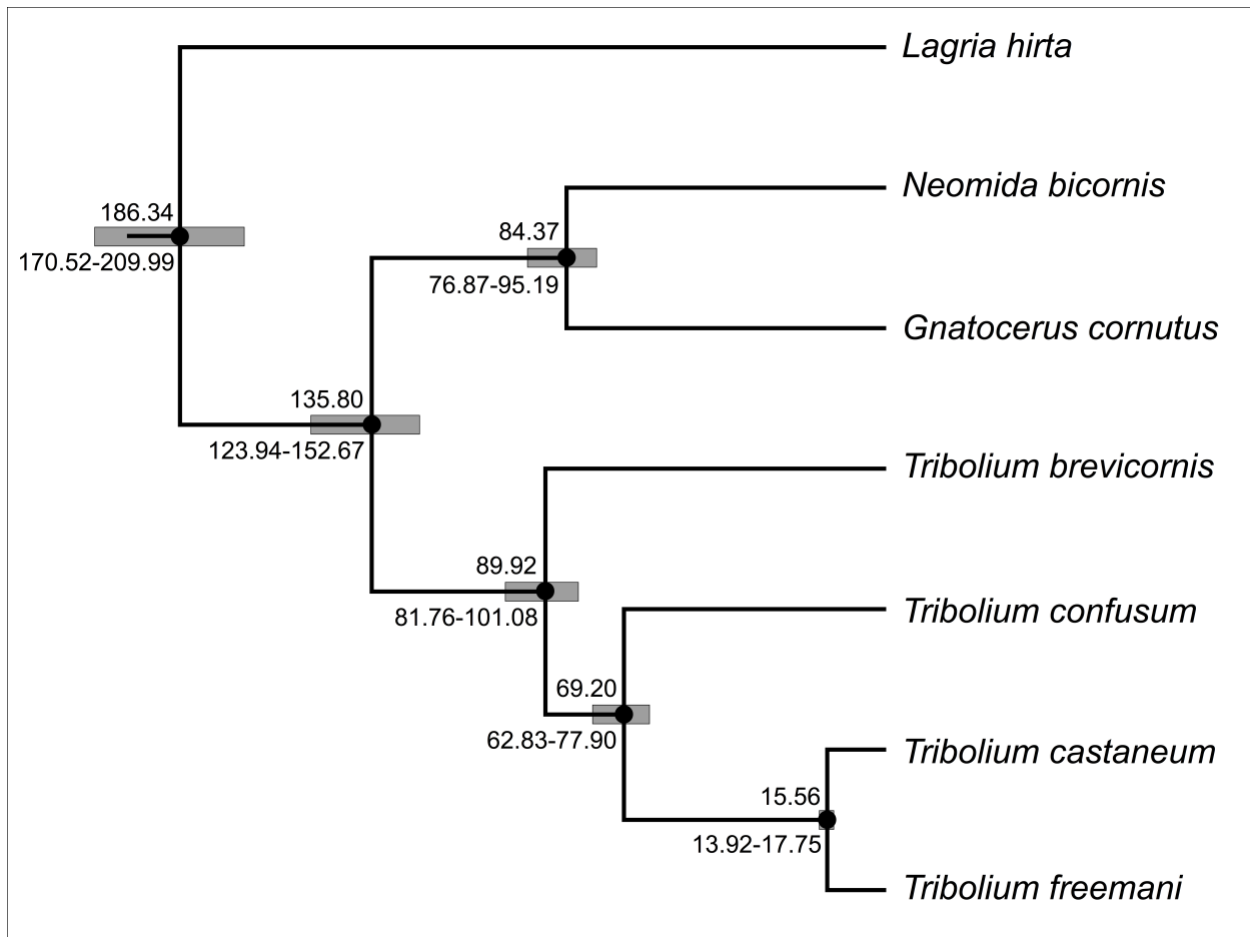
Figure S2. Dated Phylogenetic tree using expanded taxa includes *Lagria hirta* (L_hirt) and *Neomida bicornis* (N_bico) from McKenna *et al*., 2019. We used Orthofinder to identify the conserved orthogroups among all seven species and aligned them using mafft. We identified 138 orthogroups that aligned with up to 1% missing sites. The date estimates at each node are similar to the main chronogram (Figure 1), but the range is broader and less precise.

Figure S3. Frequency of Alignment Length from 138 orthogroups for the 7-species tree. Each bin is 200 bp in width, and the value on the top of the bar indicates the number of orthogroups in a bin.
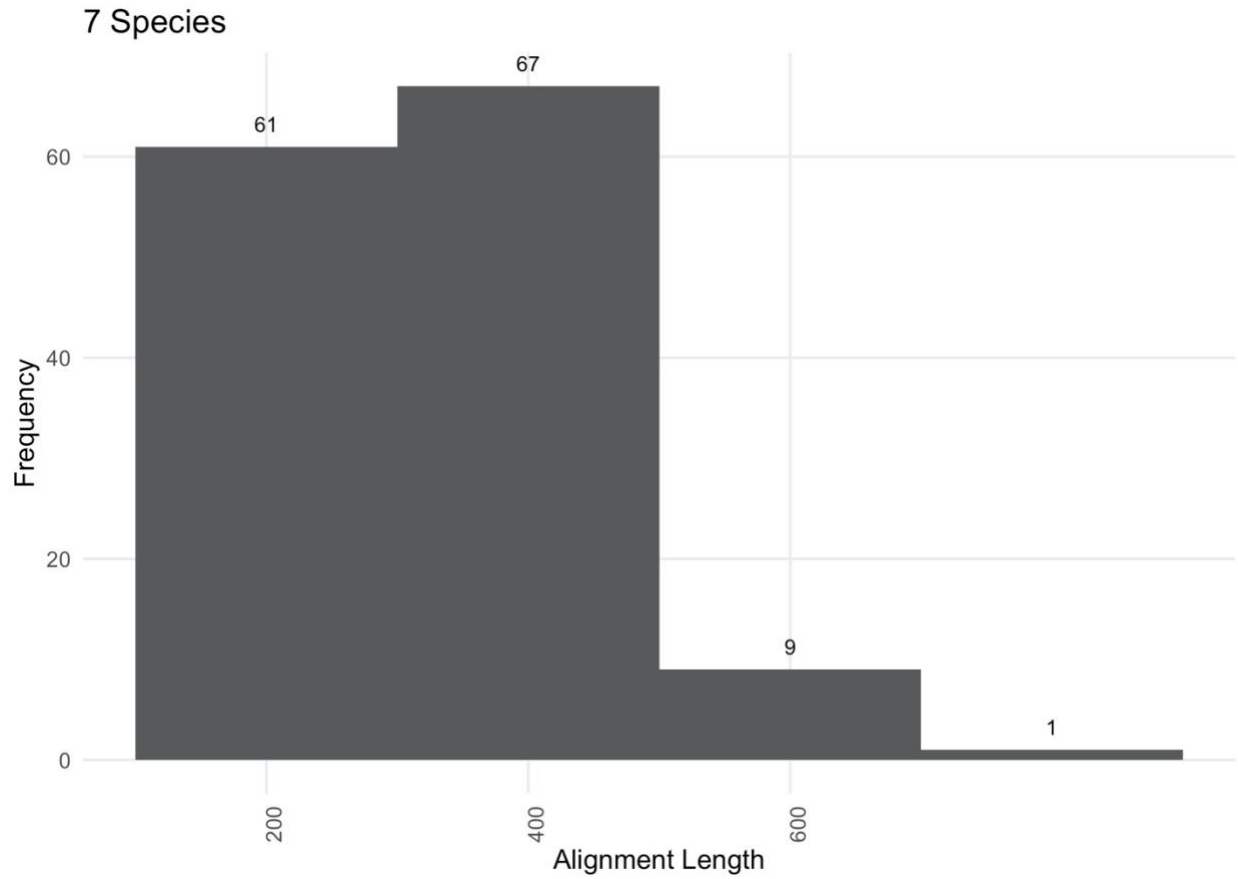
Supplementary Table 1. The annotation and function of all the orthogroups used in the

5-species tree and 7-species tree is available online at

https://doi.org/10.1016/j.ympev.2021.107084. Few lines from the 7-species sheet are

shown below. The rows highlighted in yellow are unique to the 7-species analysis.

og	aln_len	seq_ids	desc	T_cast_blast_hit_prot_id
OG0008919	267	G_corn_DN1126_c0_g1_i1.p4.p1,GDPQ 01015419.1.p1,GDMA01016952.1.p1,T_ brev_DN2176_c0_g1_i1.p1.p1,T_cast_D N524_c0_g1_i1.p3.p1,T_conf_DN2034_ c0_g1_i1.p1.p1,T_frem_DN2475_c0_g1 _i1.p1.p1	transmemb rane protein 199	XP_008199027. 1
OG0008946	203	G_corn_DN11837_c0_g1_i1.p2.p1,GDP Q01013683.1.p1,GDMA01016065.1.p1, T_brev_DN69993_c0_g1_i1.p1.p1,T_ca st_DN2144_c0_g1_i1.p2.p1,T_conf_DN 997_c0_g1_i1.p3.p1,T_frem_DN18755_ c0_g1_i1.p2.p1	UPF0454 protein C12orf49 homolog	XP_973740.1
OG0008953	177	G_corn_DN1191_c0_g2_i11.p2.p1,GDP Q01012439.1.p1,GDMA01008300.1.p1, T_brev_DN103461_c0_g1_i1.p1.p1,T_c ast_DN78913_c0_g1_i1.p1.p1,T_conf_D N695_c0_g1_i1.p5.p1,T_frem_DN5219_ c0_g1_i2.p3.p1	peptidyl- prolyl cis- trans isomerase- like 1	XP_971205.1
OG0008996	288	G_corn_DN1271_c0_g2_i3.p3.p1,GDPQ 01014586.1.p1,GDMA01009267.1.p1,T_	dnaJ homolog	XP_974891.1

		brev_DN957_c0_g1_i1.p2.p1,T_cast_D N3890_c0_g1_i1.p1.p1,T_conf_DN3266 _c0_g1_i5.p2.p1,T_frem_DN1780_c0_g 1_i14.p2.p1	subfamily C member 17	
OG0009007	390	G_corn_DN1289_c0_g1_i10.p1.p1,GDP Q01032073.1.p1,GDMA01016614.1.p1, T_brev_DN3161_c0_g1_i1.p1.p1,T_cast _DN3976_c0_g1_i1.p2.p1,T_conf_DN71 1_c0_g1_i3.p1.p1,T_frem_DN4171_c0_ g1_i1.p2.p1	vacuolar protein- sorting- associated protein 36	XP_973963.1

Supplementary Table T2. Comparison of divergence times for nodes/clades shared across four studies. All the estimates are represented in Million Years ago (Mya). Brackets indicate 95% highest posterior densities (HPD).

Clades	5-species tree	7-species tree	Previous studies
T.castaneum – T. freemani	13.77 [13.55 – 14.0]	15.6 [13.93 – 17.75]	11.6 – 47.0[a]
T. castaneum/T. freemani - T. confusum	65.50 [64.98.9 – 65.96]	69.20 [62.83 – 77.9]	13.9 – 60.7[a] 70 – 120[b]
T. castaneum /T. confusum - T. brevicornis	86.46 [85.9 – 87.04]	89.92 [81.76 – 101.08]	-
Tribolium - Gnatocerus	121.95 [121.21 – 122.68]	135.80 [123.94 – 152.67]	148 – 165[b] 65 – 129[c]
Gnatocerus/Tribolium - Lagria	-	186.34 [170.52 – 209.99]	160 – 175[b] 98 – 149[c]

a: Angellini & Jockush, 2008.

b: Kergoat et al., 2014b. (95% HPD for "Yule crown" calibration)

c: McKenna et al., 2019.

Chapter 3

Ancestral reconstruction of neoX chomosome expression states reveals the status and evolution of dosage compensation in flour beetles

Balan Ramesh[1] and Jeff Demuth[1]

[1]*Department of Biology, The University of Texas at Arlington, Arlington, Texas - 76013*

Summary

Ohno (1967) originally proposed that the sex difference in X-linked gene dose caused by the decay of Y-linked genes may impose a "peril of hemizygosity." To mitigate this peril, Ohno hypothesized that regulatory mechanisms must evolve to hyperexpress genes on the single X in males while maintaining unchanged diploid (XX) expression in females. Mounting evidence suggests that Ohno's paradigm is not universal, but our understanding remains unclear because quantifying expression evolution requires estimating the ancestral expression of X-linked genes, which is difficult or impossible in many systems.

Typically, studies assess dosage compensation by either comparing average male X: Autosome (AA), and or X_{male}: XX_{female} expression ratios. Inferring dosage compensation from the X:AA ratio is not ideal because, 1) it does not compare homologous genes, and 2) it assumes that the current average AA expression is a good proxy for the average ancestral expression of X-linked genes, which may be misleading if atypical expression is the reason some autosomes are more likely to become sex-linked. Similarly, comparison of the current X:XX expression ratio, which we will call dosage balance, does not provide clear insight about sex chromosome evolution without information about the ancestral states because balance can be achieved by upregulating the male X or downregulating female XX.

Here, we study the evolution of dosage compensation and balance by quantifying changes in gene expression across a time-calibrated phylogeny of five closely related beetle species. One species in the group, *Tribolium confusum*, harbors an X-Autosome fusion, which allows us to test the dosage compensation hypothesis (neoX=Ancestral)

26

directly. We find that, relative to their autosomal ancestral state, neoX-linked genes in *T. confusum* have evolved 2-fold higher expression per dose in males, but remained constant in females resulting in complete dosage compensation and dosage balance. Further, we observe that all flour beetle species included in the study are fully balanced and compensated (X=XX=Ancestral). This includes *T. castaneum* an important model system with previously uncertain DC status. Our results begin to suggest that the evolution of chromosome-wide dosage compensation mechanisms, as envisioned by Ohno, may at least be the rule in male heterogametic insects.

Results and Discussion

Sex chromosomes are thought to originate from a pair of ancestral autosomes after gaining a sex-determining locus (Bachtrog, 2006; Charlesworth, 1991; Charlesworth et al., 2005; Furman et al., 2020; Gu and Walters, 2017). In many cases, subsequent decay of Y chromosome gene content leaves XY males with only one functional copy for many X-linked genes, while XX females retain two copies. Ohno (1967) originally proposed that the sex difference in X-linked gene dose caused by the decay of Y-linked genes may impose a "peril of hemizygosity". To mitigate this peril, Ohno hypothesized that regulatory mechanisms must evolve to hyperexpress genes on the single X in males while maintaining unchanged diploid (XX) expression in females. This dosage compensation paradigm has dominated sex chromosome research for over 50 years, and although mounting evidence indicates that it not universal, many aspects of our understanding remains uncertain because quantifying expression evolution requires estimating the ancestral state of X-linked gene expression, which is difficult or impossible in many systems (Furman et al., 2020; Gu and Walters, 2017).

Often, studies assess dosage compensation by comparing 1) average expression of X-linked genes in males to the average expression of all autosomal genes (X:AA), and or 2) X-linked gene expression between males and females (X:XX; which, for the sake of clarity we will call dosage *balance*). Both measures pose challenges for the interpretation of dosage compensation and balance. First, inferring the evolution of dosage compensation from the X:AA ratio is not ideal because it does not compare homologous genes, and assumes that the current average AA expression is a good proxy for the average ancestral expression of X-linked genes. The latter assumption

could be particularly misleading if atypical expression is the reason some autosomes are more likely to become sex-linked. Second, testing dosage balance by comparing current X:XX ratios, will not reflect compensation of the ancestral expression level if females evolve reduced expression to match hemizygous male levels.

To test Ohno's hypothesis that dosage compensation evolves to maintain ancestral autosomal expression levels, we employ a phylogenetic comparative method to analyze the evolution of gene expression on the neoX chromosome of *Tribolium confusum*. The X-chromosome in *T. confusm* is comprised of an ancestral X region, which is conserved across flour beetles, and a neoX region, which remains autosomal in other lineages (Smith, 1952, 1950; Smith and Brower, 1974). To reconstruct ancestral gene expression, we sampled four additional species that phylogenetically flank the neoX origin: *Gnatocerus cornutus, Tribolium brevicornis, Tribolium freemani,* and the well-studied beetle model system *Tribolium castaneum*, which provides a high-quality reference genome. Aside from the neoX in *T. confusum*, all species in the analysis have the same karyotype and are expected to retain high levels of synteny based on previous analysis across a much wider phylogenetic breadth in coleopteran genomes (McKenna et al., 2016). For simplicity, throughout the results we refer to the region of the X chromosome shared by all species as the X-chromosome. The X-linked region that is unique to *T. confusum* is referred to as the neoX in *T. confusum* and chromosome 2 in the other four species, where it remains an autosome.

We conducted RNA-Seq on pooled somatic tissues (head and prothorax) from ten adult beetles of each sex for each species. Each sample type was replicated twice (5 species x 2 sexes x 2 replicates = 20 sequencing samples). All samples returned at

least 40M reads and are available through NCBI Sequence Read Archive

(SRR14070854 through SRR14070873) under the bioproject number PRJNA716663.

Since *T. castaneum* is the only species in our analysis with a reference quality genome,

we used it to assign chromosome annotations, but performed expression analysis on *de novo* assembled transcriptomes for all species to avoid potential biases due to read

mapping (see Supplementary Materials and Methods).

Karyotype data for *T. confusum* indicates that the neoX region is hemizygous

(i.e. the Y-chromosome is significantly reduced in chromosome spreads); however,

since there is no reference assembly for the species, we analyzed the number of

variable sites on the X, neoX and autosomes to establish zygosity.   If neoX linked

genes are heterozygous in males we expect to observe more variants than if they are

hemizygous because twice as many chromosomes will be sampled for diploid loci even

accounting for the fact that transcriptome data may include allelic biases. To test our

expectations we analyzed the number of variants on the neoX compared to the X

chromosomes and to autosomes in both male and female reads. We find that there is

no significant difference in the number of segregating sites between neoX and X (Mann-

Whitney U test, p-value = 0.688) but there are significantly fewer variant sites on the

neoX (and X) than on autosomes (Mann-Whitney U test, p-value = 0), indicating that the

neoX is indeed hemizygous in males (see Supplementary Materials and Methods).

Dosage compensation is achieved by hyperexpression of the X in males.

If dosage compensation evolves to maintain ancestral expression in males, then

we expect to see 2-fold increase in neoX-linked gene expression per gene dose (i.e.

neoX/1 : ancestral/2 = 2). To test this hypothesis, we trace the evolution of gene

expression for 14,583 orthogroups retained in all five species across a time calibrated

phylogenetic tree (Ramesh et al 2021). Ancestral expression states were estimated

using a weighted averaging approach as in previous analyses of vertebrate sex

chromosomes (see Supplementary Materials and Methods)

Our results show that average expression of neoX linked genes is not statistically

different from their ancestral states in males (Mann-Whitney U test, p-value = 0.79),

owing to 2-fold upregulation (Figure 2). Expression of neoX linked genes in females

also remains unchanged from their ancestral states (Mann-Whitney U test, p-value =

0.82). Indeed, the median expression levels of neoX genes in *T. confusum* are closer to

their inferred ancestral states ($\log_2$[neoX:Autosomal ancestor expression] = -0.10) than

the chromosome 2 homologs in *G. cornutus* (-0.25)*, T. brevicornis* (-0.51)*, and T.*

freemani (-0.16), where chromosome 2 remains autosomal (Figure 1a). Thus, the neoX

of *T. confusum* appears to have evolved complete dosage compensation and dosage

balance between the sexes by a mechanism of X hyperexpression in males.

Additionally, comparing the distribution of neoX expression with the expression of genes

from the rest of the X chromosome in *T. confusum* shows that they are virtually

indistinguishable from each other, which suggests that genes in the neoX region have

been fully integrated into existing mechanisms of compensation acting on the ancestral

X (Figure 3b and Supplemental Figure S15).

Dosage balance between sexes is more tightly constrained than dosage

compensation.

In mammals, X chromosome expression is balanced between sexes by X

chromosome inactivation (XCI) in females such that both sexes are functionally

hemizygous. Ohno originally proposed that XCI evolved as a way for females to compensate for the 2-fold upregulation that was necessary for males to avoid the peril of hemizygosity. However, recent studies show that placental mammals lack global upregulation of the X and only a subset of 'dosage sensitive' genes are upregulated to autosomal levels (Julien et al., 2012; Pessia et al., 2012). These findings cast doubt on Ohno's two-step hypothesis and highlight our lack of understanding about whether X chromosome expression evolution is driven more by the need to maintain dosage compensation relative to the ancestral state (X=Ancestral), as envisioned by Ohno, or to maintain dosage balance between male and female expression (X=XX).

If X-chromosome expression evolution is primarily driven by the need to maintain ancestral expression, we would expect there to be less variation between the neoX and ancestral state than between neoX expression in males and females. To test this prediction, for each neoX-linked gene we compared the deviations from their ancestral states to deviation between current expression in males and females. Counter to the prediction, our results show that the correlation in gene expression between males and females (Pearson's ρ= 0.90), is higher than the correlation between neoX and the ancestral state (Pearson's ρ = 0.82). The high degree of correlation in gene expression between males and females for chromosomal segments, the absence of any gene dose effect on neoX and X chromosome expression, coupled with the fact that the M:F ratios are similar to that of Autosomes (Figure 3b) suggest that *T. confusum* is completely dosage balanced.

To account for the possibility that the above result may be biased by the uncertainty inherent in ancestral state reconstruction, we expanded our analysis to

compare current X-chromosome expression to current autosomal expression across all

species in our analysis. As in the neoX comparisons above, we find that the median

expression of X chromosome between sexes (U(N_{male} = 5, N_{female} = 5) = 27, Z =

0.10445, p-value = 0.9168) is more tightly regulated than the median expression

between X and Autosomes (U (N_{AA} = 5, N_X = 5) = 22 , Z = 1.1489, p-value = 0.2506)

(Figure 3).

Flour beetles have complete sex chromosome dosage compensation and dosage balance.

Previous work in *T. castaneum* comparing current expression of X and

autosomal genes left the status of dosage compensation and dosage balance unclear.

An initial study by Prince et al. (2010), using microarray of whole-body adults, found that

males hyper express the X so that X=AA; however, females also appeared to

hyperexpress the X so that XX>X and XX>AA. Subsequently, Mahajan and Bachtrog

(2015) found complete compensation in RNA-Seq data from somatic tissue. Most

recently, a broad reanalysis of published RNA-Seq data found higher expression of X-

linked genes than autosomal genes in both males and females (Chen et al., 2020).

Making the parallel X to Autosome comparisons as conducted in the work cited above

reveals that all five of the flour beetle species all have complete dosage compensation

and balance (i.e., X=XX=AA, Figure 3). Furthermore, if we assume that the pattern of

neoX chromosome expression evolution in *T. confusum* is a reflection the neoX

incorporating existing (ancestral) X-chromosome regulatory machinery, our findings

suggest that compensation across all of these beetle species is achieved by 2-fold

hyperexpression of the X in males while simultaneously retaining ancestral diploid expression in females.

Gnatocerus cornutus, the broad horned flour beetle, is the only species in our analysis where some doubt may remain about complete dosage compensation in somatic tissues. The median expression of X compared to autosomes in males ($\log2$(X:AA) = -0.66) and females ($\log2$ (XX:AA) = -0.75) is more different than in the Tribolium species. It may be that the appearance of imperfect dosage compensation is due is actually a result of sexually antagonistic selection acting strongly to shape X-chromosome gene content and expression in this species since intrasexual conflict for secondary sexual traits such as mandibles in males and abdomen length in females is ongoing (Harano et al., 2010), but testing this hypothesis will require additional investigation.

Our analysis of the number of segregating sites suggests that female beetles may employ XCI. The analysis of segregating sites we employed to investigate hemizygosity of the neoX shows that the X and neoX have fewer segregating sites than autosomes in males, as we would expect for the reduced effective population size of hemizygous chromosomes; however, the same reduction in segregating sites is observed in females (i.e. the number of X-linked variants detected is the same in males and females Supplemental Figure S16). This pattern seems best explained by female beetles expressing only one X chromosome as observed in the XCI mechanism of marsupial mammals. In which case RNA-Seq would be effectively sampling the same number of X chromosomes in both sexes. We remain cautious about this conclusion due to the limitations of our sampling (pooled transcriptomes of multiple individuals), but

discerning whether females retain diploid ancestral expression by an XCI system, or whether a dosage compensation complex assembly is male specific as in *Drosophila* will be a promising avenue of future exploration.

Affects of gene family evolution, tissue type, and minimum expression thresholding

To analyze if the 5-way single copy orthologs in our phylogenetic analysis biases our estimation towards complete compensation, we split our analysis into the conserved 5-way single copy orthogroups (249 on X chromosome, 407 on Neo-X and 3334 on Autosomes) and those orthogroups that are not shared across all species (Figure 4). We find that both groups are dosage compensated but the tolerance in expression away from complete compensation is higher in the group that are not single copy orthologs which is as expected (see Supplementary Materials and Methods).

Previous studies that found a lack of complete compensation and balance in *T. castaneum*, included gonad tissues seems likely to explain differences among dosage compensation studies in beetles. *T. castaneum* appears to have complete compensation in studies that exclude gonads ((Mahajan and Bachtrog, 2015), present study), but lacks complete compensation in studies were testes and or ovaries are included (Chen et al., 2020; Prince et al., 2010). We now know that the *T. castaneum* X chromosome has significantly female-biased expression in ovaries (Prince et al., 2010; Whittle et al., 2020; Williford and Demuth, 2012) inflating average X expression in females relative to males and autosomes and, therefore, influencing the interpretation of incomplete dosage compensation in analyses where gonads are included. Consequently, we excluded gonads from the present study and recognize that the

complete compensation may only hold in the soma. This is consistent with other

studies that find variation among tissues in the degree of compensation and balance

(Marin et al 2015, Gu and Walters 2019)

Dosage compensation studies on humans provide an excellent example of how

the lower limit of expression (i.e.) highly sex/tissue-biased gene expression can have a

drastic impact in determining the type of dosage compensation. Here, we tested the

effect of highly biased genes on each chromosome contig by varying the minimum

threshold of expression. Our results indicate that the M:F ratio between neoX, X-

chromosome, and autosome is similar and comparable at different minimum threshold

levels in *T. confusum* (Supplemental Figure S10, S11 and S13). By including only

somatic tissues from adult beetles, we avoided the effect of tissue-specific biased gene

expression, thereby providing credibility to the type of dosage compensation in the flour

beetles. RNA-Seq studies with a library size of 20-30M provide sufficient statistical

power to identify almost all the differentially expressed genes present (Hart et al., 2013).

Our samples at ~40M reads are adequate to identify almost all the differences in gene

expression between the sexes though it includes reads from low expressed (LE)

transcripts. However, visualizing the expression profile of all the genes shows a clear

pattern of a bimodal curve, which is due to the presence of low expressed transcripts

and the active transcripts (Hart et al., 2013; Hebenstreit et al., 2011). Here, we use a

bimodal graph (see Supplementary Materials and Methods) to decide on the minimum

expression threshold (TMM > 1) thereby considering just the active genes in all the

analyses.

Conclusion

Our phylogenetic analysis of X chromosome expression evolution in flour beetles

indicate that these beetles have complete dosage compensation (X=XX=AA) in the

soma and that it is achieved by 2-fold upregulation in males (X=Ancestral) and

maintenance of XX=Ancestral expression in females. Furthermore, neoX and ancestral

X chromosomes in *T. confusum* do not differ in average expression. This suggests that

the ~65 million years since the X-autosome fusion has been sufficient for the

mechanisms of dosage compensation acting on the ancestral X to fully assimilate the

neoX.

The pattern of compensation and balance in flour beetle somatic tissues is

consistent with Ohno's paradigm for the evolution of dosage compensation. To our

knowledge, similar patterns are also known from all XY insects including: fruit flies

(Nozawa et al., 2016; Vicoso and Bachtrog, 2015; Zhou et al., 2013), mosquitoes (Jiang

et al., 2015; Rose et al., 2016), pea aphids (Richard et al., 2017), stink bugs (Vicoso

and Bachtrog, 2015) and milkweed bugs (Vicoso and Bachtrog, 2015), although dosage

compensation of the neoX region in some Strepsipterans may be incomplete (Mahajan

and Bachtrog, 2015).   Most of these examples likely represent the independent

evolution of dosage compensation, and although the mechanisms underlying the

patterns are virtually unknown, the consistency across such a wide phylogenetic sample

suggest that the Ohno's paradigm often holds in male heterogametic insects.

Finally, we also observed that the need to maintain balanced expression

between sexes is more constrained than the maintenance of diploid ancestral

expression levels. If dosage balance is actually the evolutionary priority, as opposed to

the peril of hemizygosity, the observation of sex chromosome balance without dosage
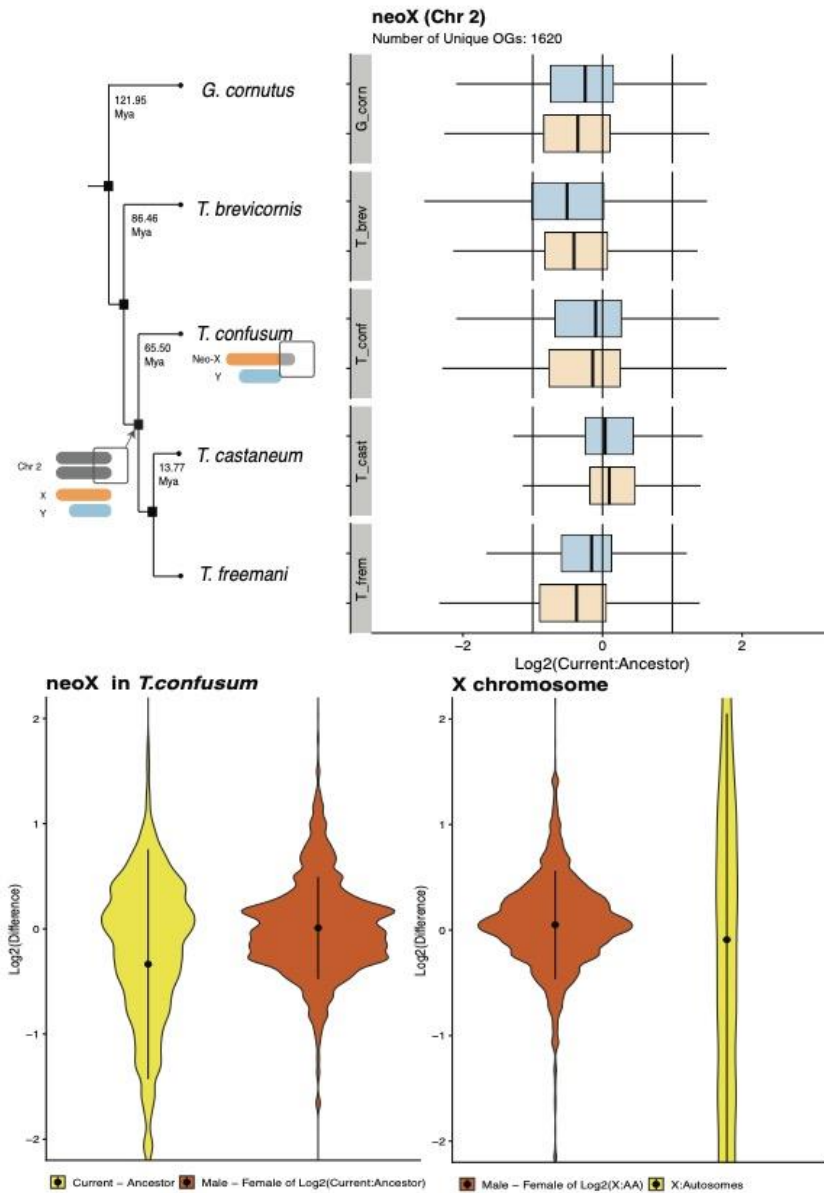
compensation becomes consistent.

Figure 1. Dosage compensation comparing the ratio of current to ancestral state between males (blue) and females (orange). a) The chromosomal cartoons illustrate the comparisons between the ancestral state and the current state in *T. confusum* males for neoX gene expression b) We depict the summary of the mean and the variance for

the difference between Male to Female as well as the neoX to Chromosome 2

combined across all 5 species to highlight the higher tolerance for dosage

compensation than dosage balance. The use of difference instead of ratio in Figure 1b

is for visualization purpose to highlight the tight constraint of dosage balance versus
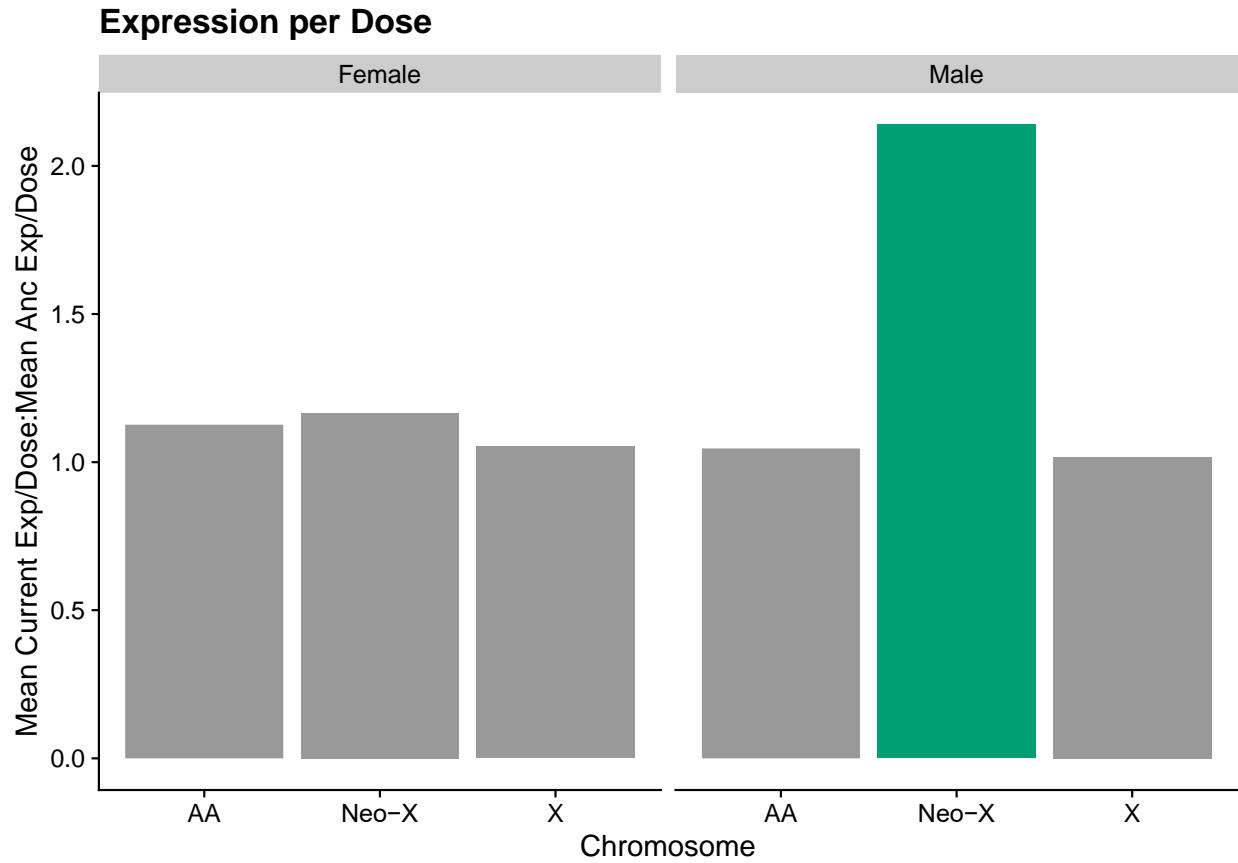
DC.

Expression per Dose

Figure 2. The ratio of mean expression per dose between the current the ancestral state for each chromosomal category in *T.confusum* is illustrated. The neoX chromosome is highlighted here to depict the increase in the expression following divergence from the common ancestor.
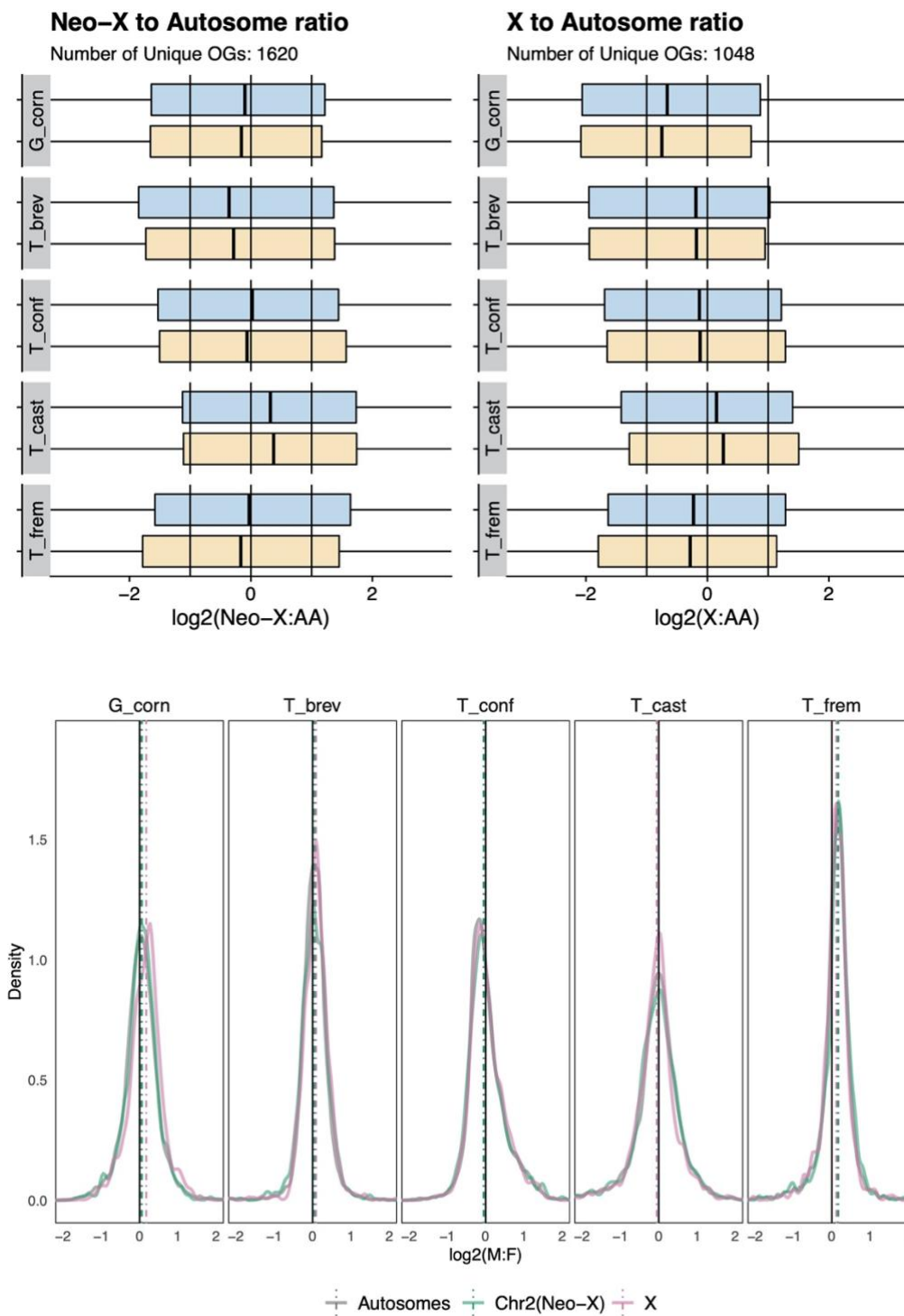
Figure 3. The expression ratio of X and neoX chromosome against autosome is shown here. The distribution and the five point statistics are all very close to 0 implying equal

expression between X/neoX chromosome and Autosome. Key reference log2 ratios are: -1 (2-fold increase in male expression over female); 0 (equal expression between sexes); 1(2-fold increase in female (orange) expression over male (blue)). b)Density plot for the log2(male to female ratio) for autosomes, Chr 2 (neoX in *T. confusum*) and X chromosome. b) Dosage balance comparison across chromosomes for each species is shown as a density plot for the log2(male to female ratio) for autosomes, Chr 2 (neoX in *T. confusum*) and X chromosome. The dotted dashed lines represent the median expression of each chromosome group.
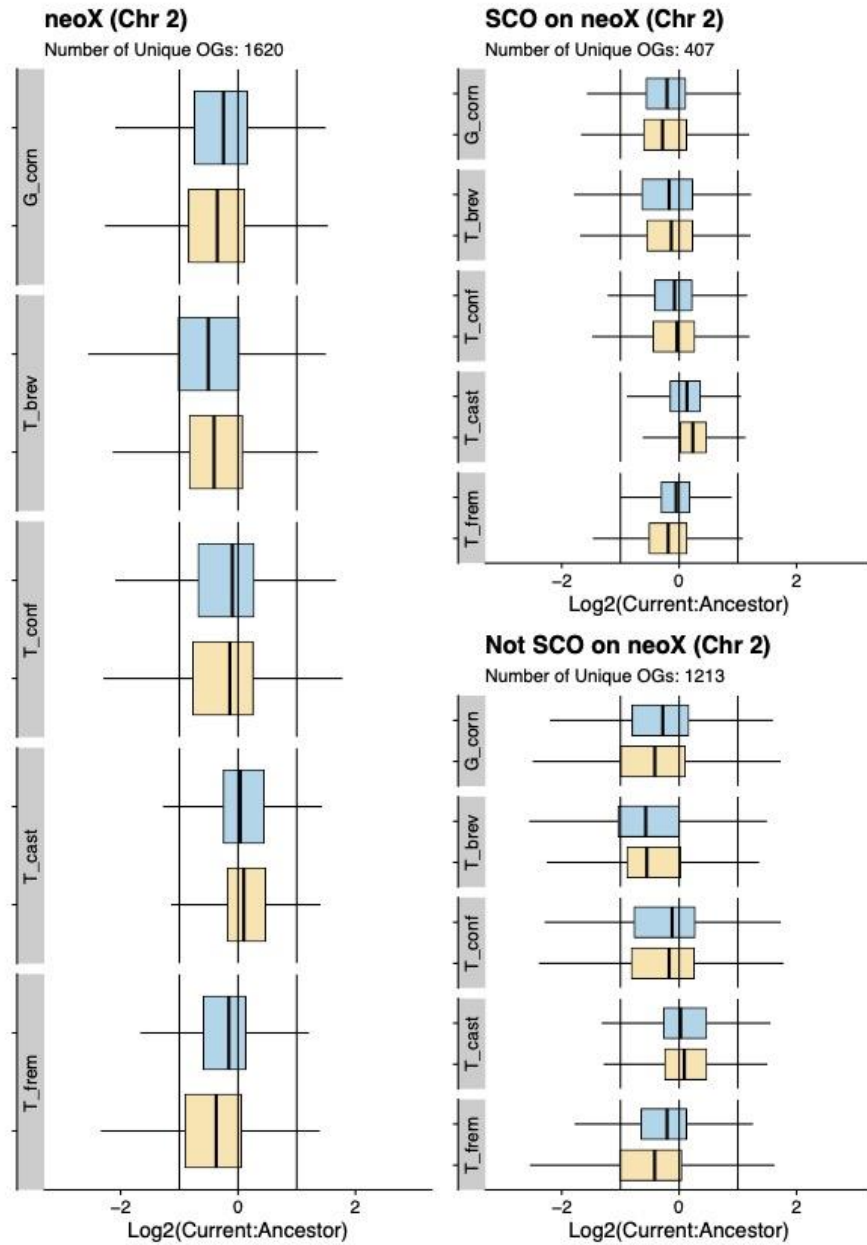
Figure 4. Ratio of Current to Ancestral expression for all orthogroups on neoX (Chr2) chromosome. Single Copy Orthogroups (SCO) and Orthogroups that are not SCO are shown as subplots. The total number of orthogroups in each plot is shown below the title.

Supplementary

Materials and Methods

Samples and RNA-Seq

  To extract the RNA, we dissected the head and prothorax from ten adult male beetles of *T. confusum* and pooled them together for one sequencing sample. Using this pooled sequencing sample, we used the Promega SV Total RNA Isolation System and followed the manufacturer's protocol to get the total RNA. We repeated the process for females as well as for *T. castaneum*, *G. cornutus*, *T. brevicornis,* and *T. freemani* such that we had two sequencing samples (replicates) per sex per species (2 sexes x 2 replicates x 5 species = 20 total sequencing samples). The quantity and quality of RNA were assayed using Qubit and Bioanalyzer prior to delivering the samples to the North Texas Genome Center/Novogene for RNA-Seq library preparation and sequencing. The libraries were sequenced on the Illumina HiSeq4000 platform (150 bp paired-end) at Novogene or NovoSeq platform (101 bp paired-end) at North Texas Genome Center. All the raw sequence reads are available through NCBI Sequence Read Archive (SRR14070854 through SRR14070873) under the bioproject number PRJNA716663.

De Novo Transcriptome Assembly, Orthology and Chromosomal Assignment

  Owing to the inadequate mapping of sequenced reads to reference *T. castaneum* assembly, we decided to construct a *de novo* assembly. We pooled the male and female sequencing samples for each species and used Trinity v2.8.6 (Haas et al., 2013) for the assembly. The Trinity assembly was then benchmarked using BUSCO v3.0.2 (Simão et al., 2015) using the endopterygota odb10 dataset. To create a non-redundant

transcriptome assembly, we tried various filtering pipelines (mentioned in Carruthers et al., 2018; Gahlan et al., 2012; Haak et al., 2018; Morandin et al., 2018; Moreno-Santillán et al., 2019; Thunders et al., 2017). A modified analyses from Moreno-Santillán and others (2019) provided a transcriptome that improved the number of single copy transcripts in the assembly and the steps we used are as follows. To identify the open reading frames and to obtain the coding strand sequences, we used Transdecoder v5.5.0 (37), retaining homology information using blast (blastp) against Tcas5.2 peptide sequences(Herndon et al., 2020). Using the predicted coding strand (CDS) sequence file, we used cd-hit-est v4.8.1 (Fu et al., 2012) to cluster sequences that are 95% similar together using the -c flag.

The clustered CDS sequences were then passed again to Transdecoder (Haas and Papanicolaou, 2017) to obtain the protein sequence file, which was then used to get the orthologs among the five species using OrthoFinder v2.3.11 (Emms and Kelly, 2019). We retained only the orthogroups, which had at least one transcript for each species for the subsequent analysis. To annotate each transcript in the retained orthogroups, we blasted the sequences against the *T. castaneum* reference peptide sequence file v5.2 (Herndon et al., 2020), and their chromosomal information was added using the *T. castaneum* reference gff file and a custom python script.

The Trinity assemblies, clustered transcriptomes, expression quantification files and scripts are available through dryad (https://doi.org/10.5061/dryad.g1jwstqqr).

Estimation of Gene Expression and Normalization

The CDS file from the second run of Transdecoder for each species was used as a reference transcriptome to quantify expression using salmon v1.1.0 (Patro et al.,

2017) for each sequencing sample. Using the annotation from the orthologs, we added chromosomal information to the quantified samples and kept the transcripts that were in the retained orthogroups. Using DESeq2 R package v1.26.0 (Love et al., 2014), we obtained the raw counts and passed the raw counts matrix and the length of the transcripts to edgeR R package v3.28.1 (Robinson et al., 2010) to normalize for library size and the length of the transcripts to obtain the normalized TMM values of expression.

Analysis of dosage balance between sexes (M:F ratio)

We assessed the pattern of dosage balance on neoX/X linked genes of *T. confusum* by contrasting it between males and females. We calculated the ratio of TMM values of neoX(X) for all genes against median AA for males and females, respectively. We log2 transformed the ratios and visualized expression using R statistical analysis software v3.6.1. We compared the dosage balance of neoX linked genes of *T. confusum* with current-X and against all the other species. Similarly, we repeated the steps to calculate the ratio for each species.

Analysis of dosage compensation among species (Current : Ancestral ratio)

To analyze dosage compensation, we estimated ancestral expression of neoX linked genes that are still autosomal in the related species with a non-homologous sex chromosome system, as in (Julien et al., 2012; Marin et al., 2017; Schield et al., 2019). Briefly, we retained all the 1:1 orthologs and normalized their expression by the median autosomal expression for each species. We calculated the ancestral expression as a weighted mean expression using the divergence time estimates (from the midpoint rooted chronogram from (Ramesh et al., 2021)) as weights such that the weighting of a

species is reciprocal to the distance from the focal ancestral node. We estimated the ratio of current to ancestral expression by dividing the median expression (between replicates) of each single copy orthogroup for each sex by the ancestral expression.

Analysis of the frequency of variants

As a proxy to identify the hemizygous regions/genes, we used the frequency of the variants along 100kb intervals to compare between Autosomes, X, and neoX in *T. confusum* . To identify variants from the RNA-Seq data, we used Trinity v2.11.0 (Haas et al., 2013) Trinity_gene_splice_modeler.py script to create a supertranscript of the *de novo* transcriptome assembly of *T. confusum* , where each gene with multiple transcripts is represented by a single sequence. With GATK pipeline v4.9.1, bowtie2 v2.3.4 (44), and samtools v1.6 (Li et al., 2009), we used Trinity's run_variant_calling.py script to call variants. Using the BLAST hits against *T. castaneum* transcriptome reference assembly v5.2 and its annotation file, we added the position and chromosome information for each variant. We used a custom R script to visualize the frequency of the variants in 100kb intervals along each chromosome. We performed a Mann-Whitney U test to check the statistical significance for the difference in the number of variants between the chromosomes. We repeated the steps to identify variants in *T. castaneum*.

Figure S1. X:AA Ratio of all orthologs among all 5 species. Single Copy Orthogroups (SCO) and Orthogroups that are not SCO are shown as subplots.The total number of orthogroups in each plot is shown below the title.
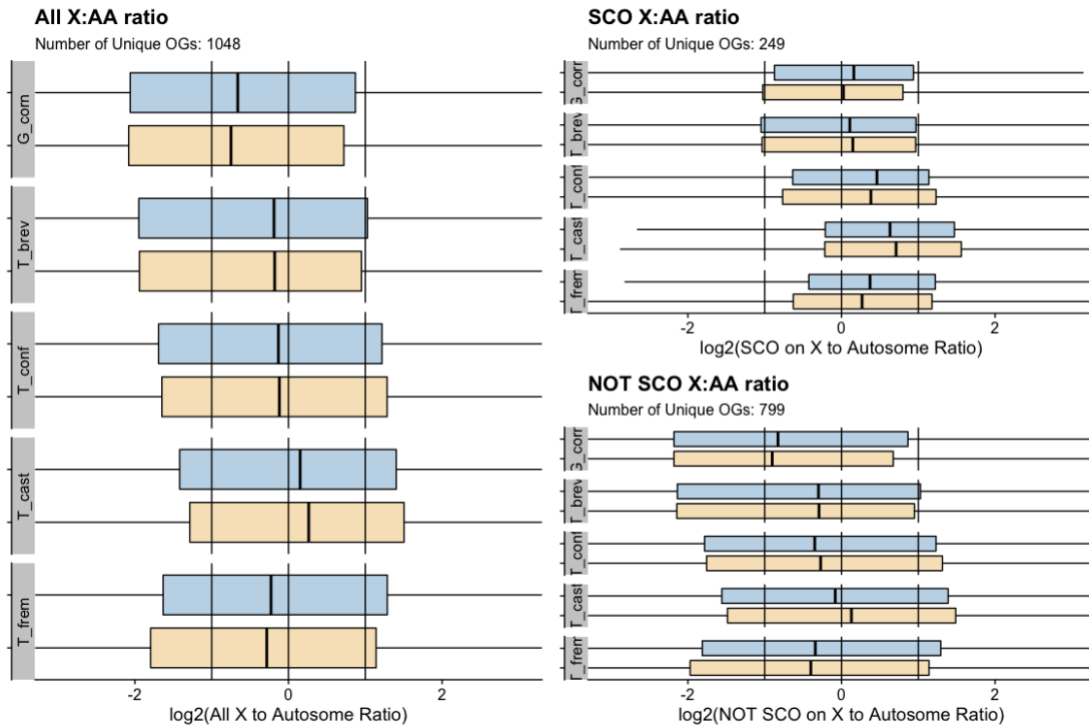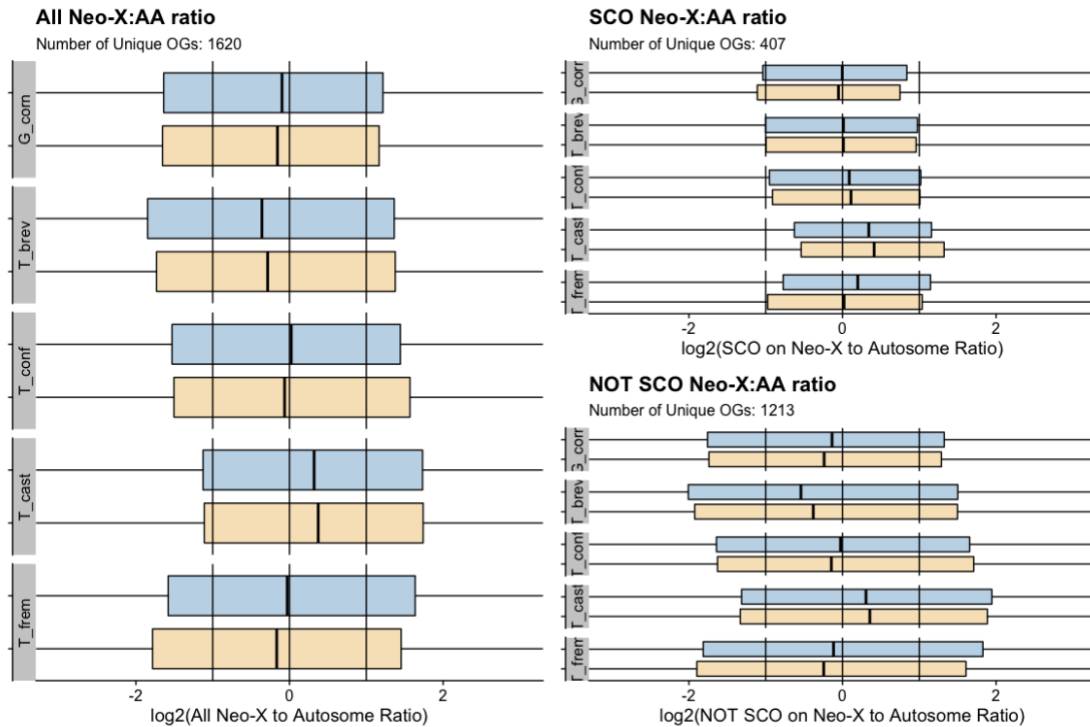
Figure S2. Neo-X:AA Ratio of all orthologs among all 5 species. Single Copy

Orthogroups (SCO) and Orthogroups that are not SCO are shown as subplots. The total

number of orthogroups in each plot is shown below the title.
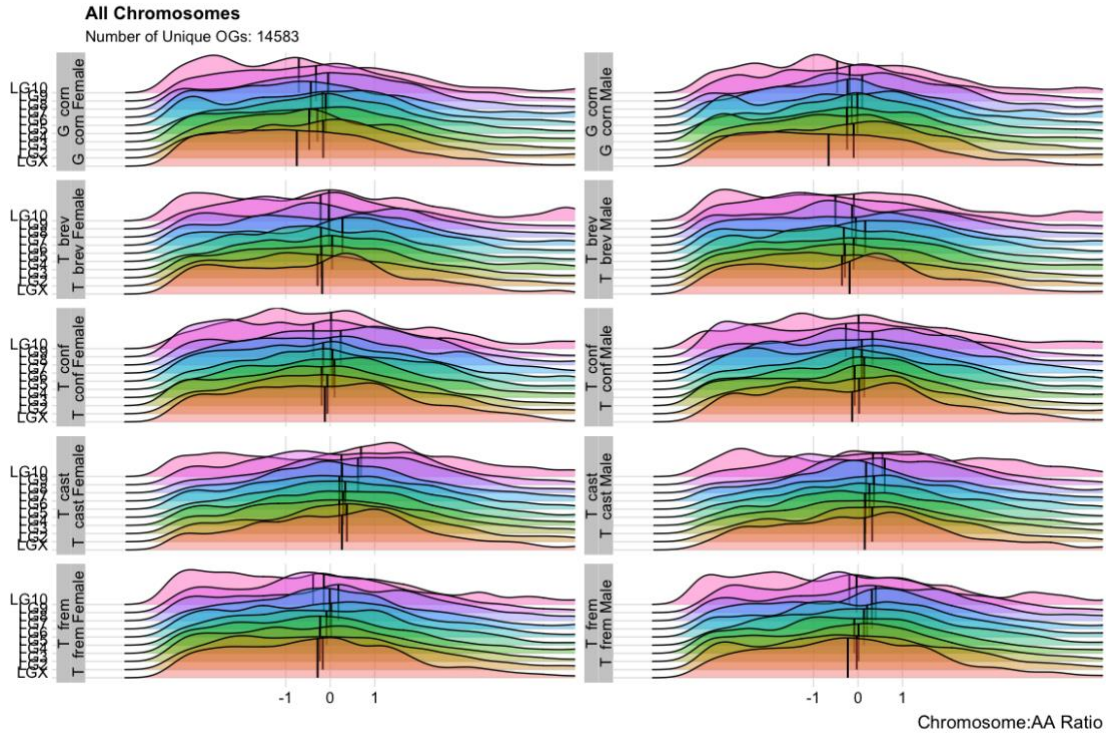
Figure S3. The density plot represents the expression across each chromosome for each species. The AA in the denominator is the Autosomal Median with all orthogroups included. The black line is the median expression and the the gray lines represents the upper and lower limit beyond which the difference is statistically significant.

Figure S4. Relatively Rapidly evolving orthogroups on Chromosome 2 present in *Gnatocerus cornutus* and *Tribolium castaneum*. Single Copy Orthogroups (SCO) and Orthogroups that are not SCO are shown as subplots. The AA in the denominator is the Autosomal Median with all orthogroups included. The total number of orthogroups in each plot is shown below the title.

Figure S5. Relatively Rapidly evolving orthogroups on Chromosome 2 present in *Tribolium brevicornis* and *Tribolium castaneum*. Single Copy Orthogroups (SCO) and Orthogroups that are not SCO are shown as subplots. The AA in the denominator is the Autosomal Median with all orthogroups included. The total number of orthogroups in each plot is shown below the title.
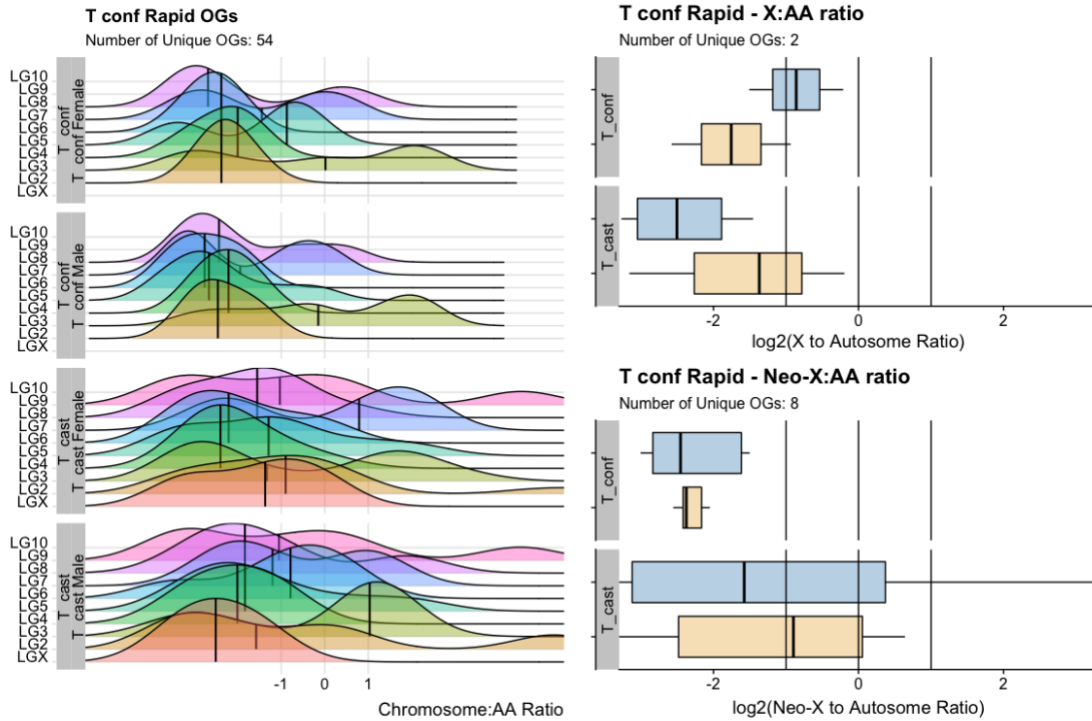
Figure S6. Relatively Rapidly evolving orthogroups on Neo-X present in *Tribolium confusum* and *Tribolium castaneum*. Single Copy Orthogroups (SCO) and Orthogroups that are not SCO are shown as subplots.The AA in the denominator is the Autosomal Median with all orthogroups included. The total number of orthogroups in each plot is shown below the title.
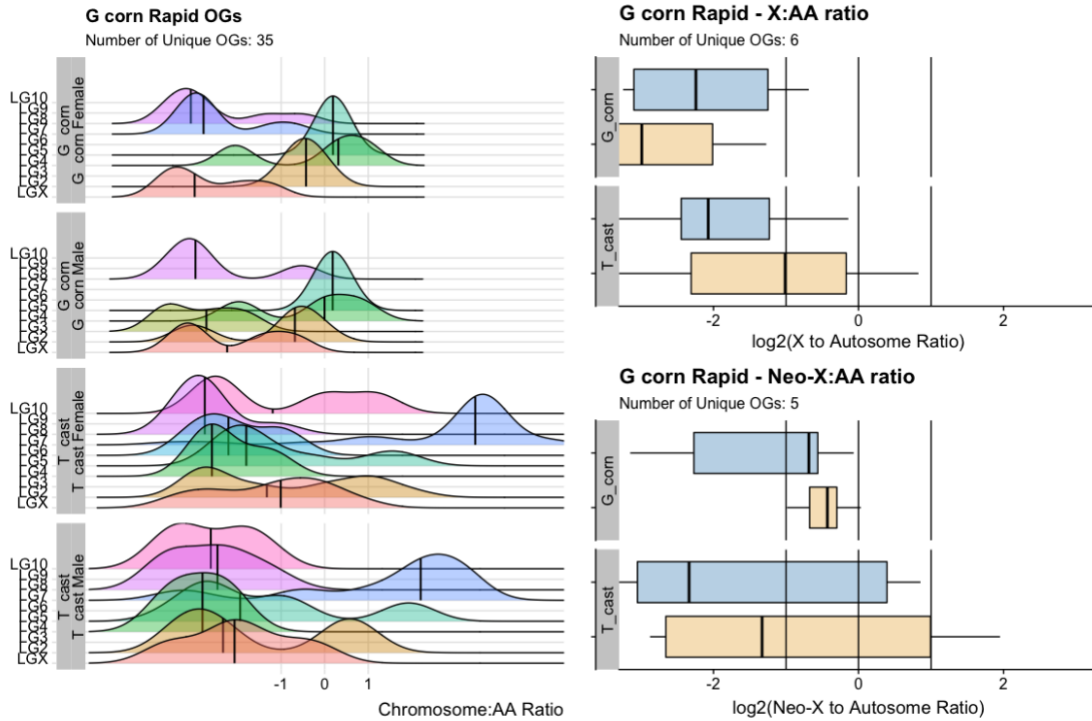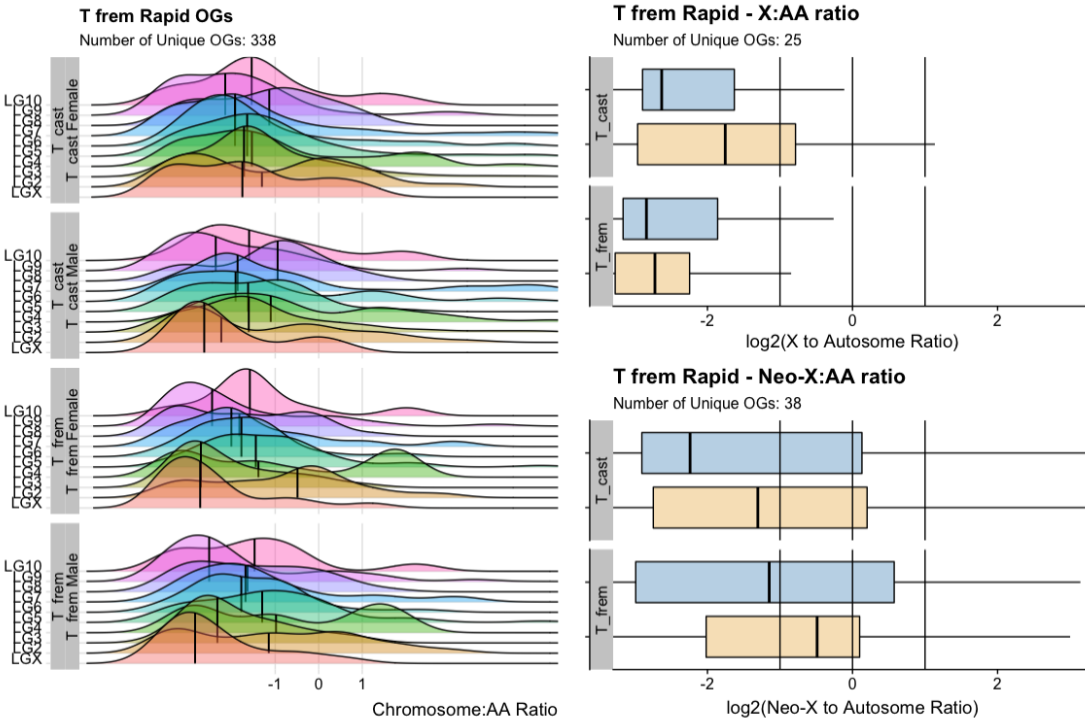
Figure S7. Relatively Rapidly evolving orthogroups on Chromosome 2 present in *Tribolium freemani* and *Tribolium castaneum*. Single Copy Orthogroups (SCO) and Orthogroups that are not SCO are shown as subplots. The AA in the denominator is the Autosomal Median with all orthogroups included. The total number of orthogroups in each plot is shown below the title.
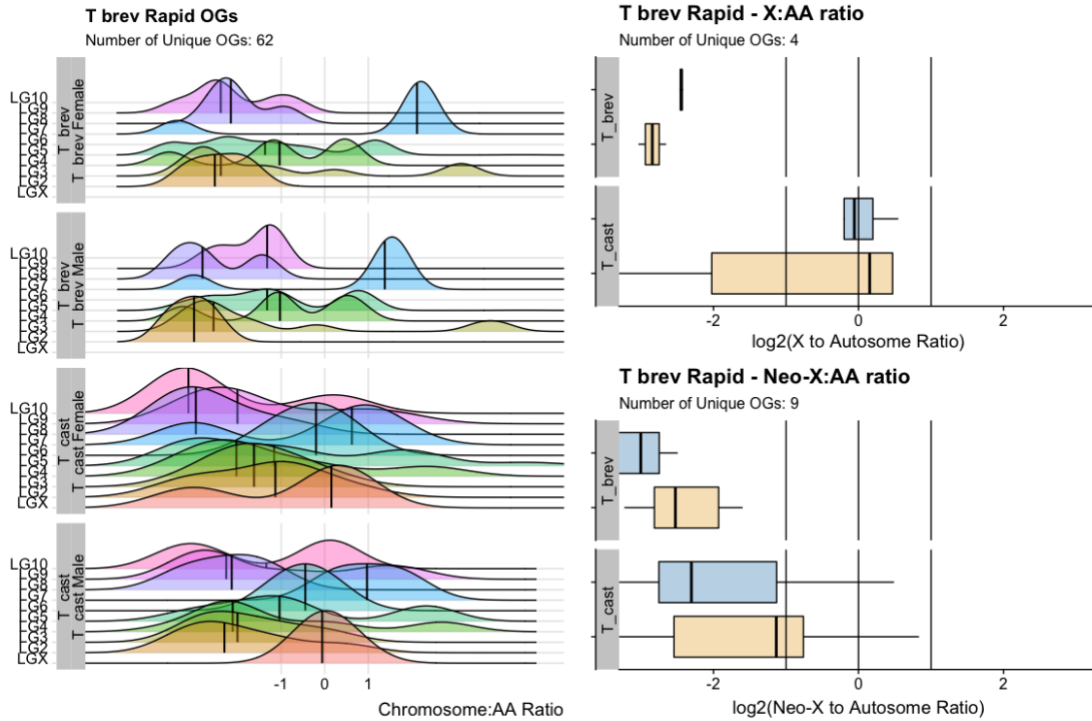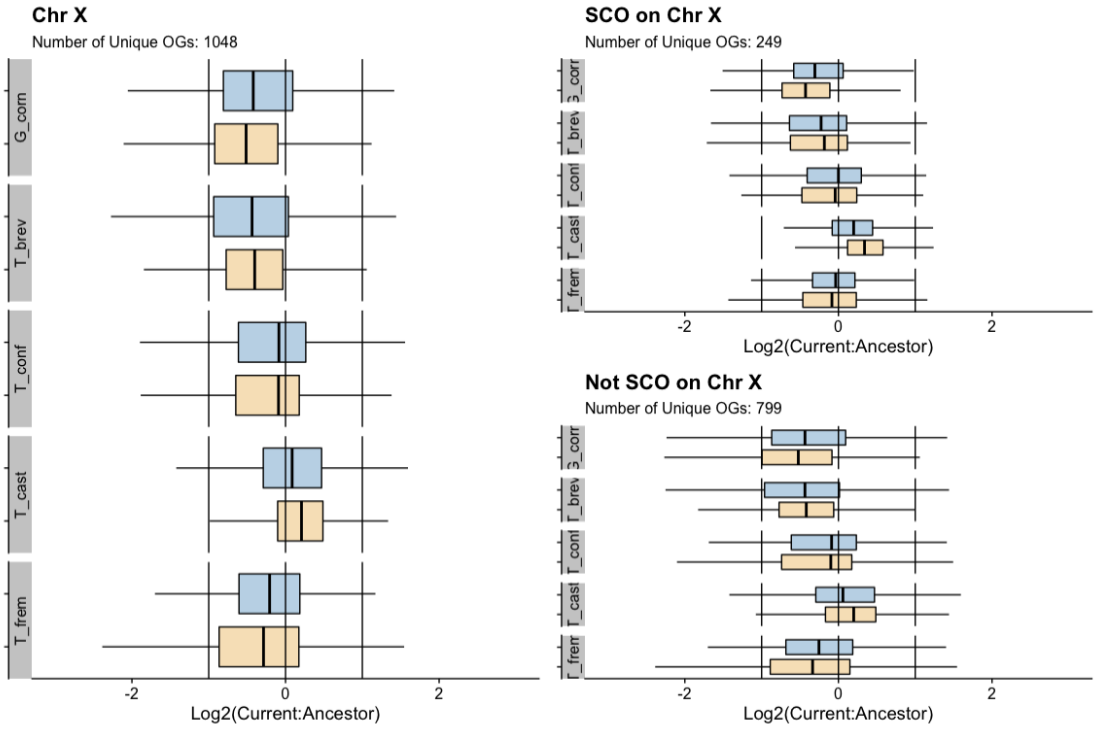
Figure S8. Ratio of Current to Ancestral expression for all orthogroups on X chromosome. Single Copy Orthogroups (SCO) and Orthogroups that are not SCO are shown as subplots. The total number of orthogroups in each plot is shown below the title.
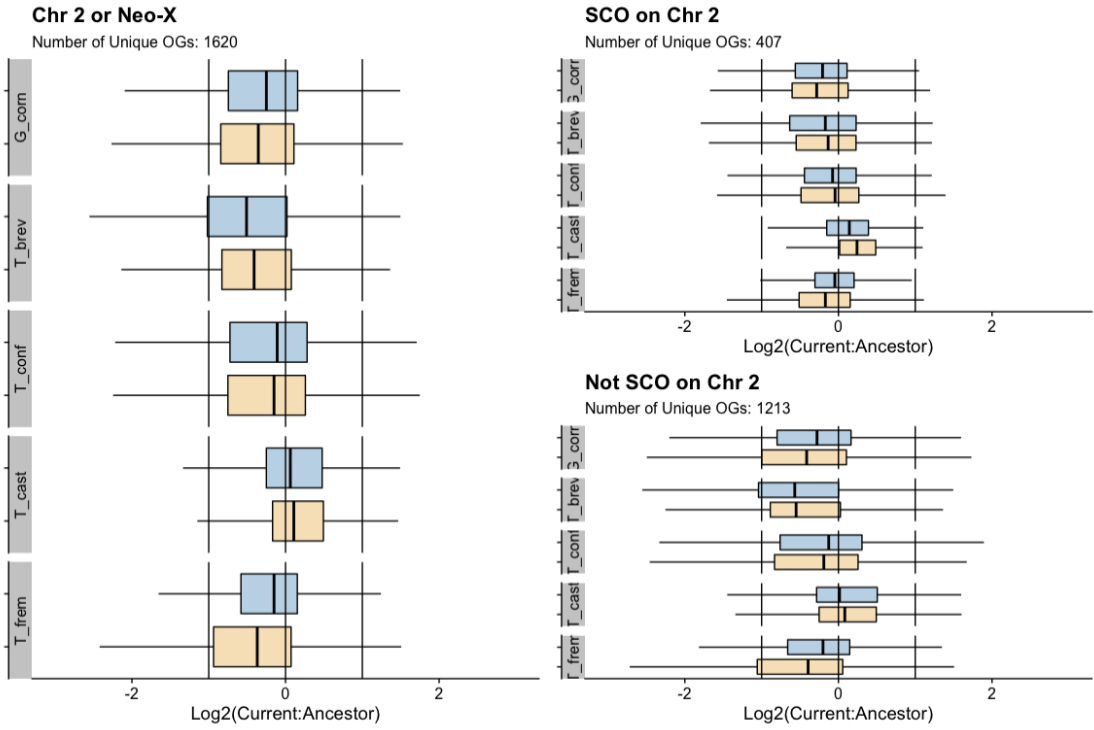
Figure S9. Ratio of Current to Ancestral expression for all orthogroups on Neo-X(Chr2) chromosome. Single Copy Orthogroups (SCO) and Orthogroups that are not SCO are shown as subplots. The total number of orthogroups in each plot is shown below the title.
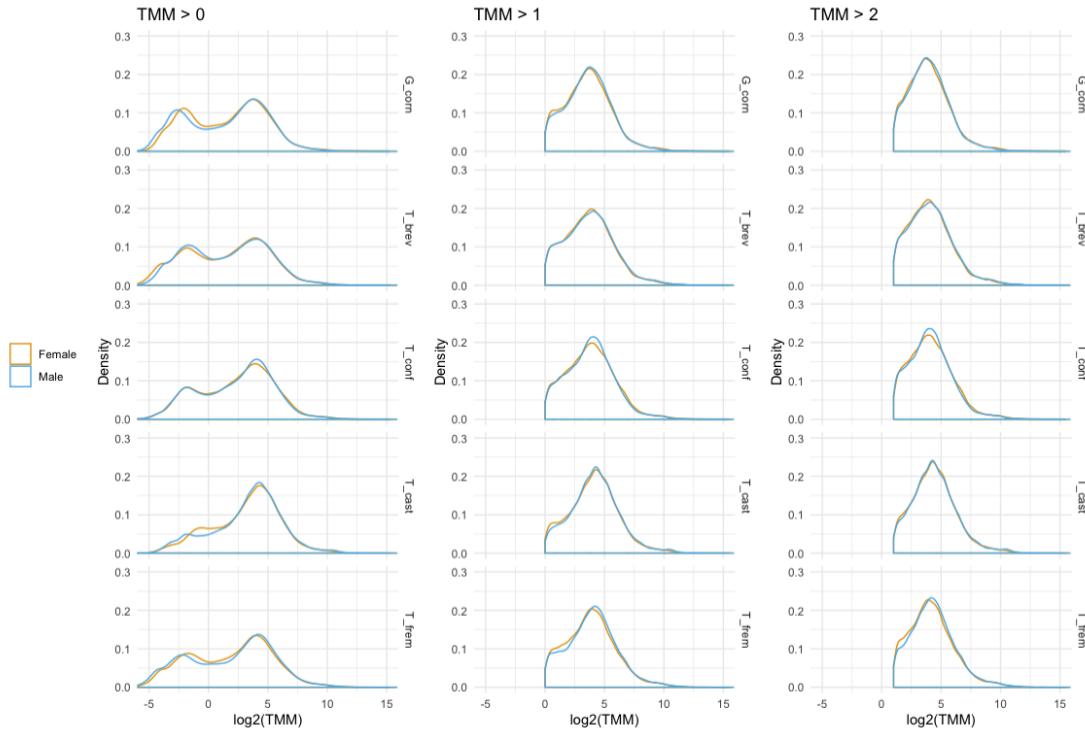
Figure S10. Minimum Expression Threshold across all 5 species at 2 levels (i.e) TMM >

1 and TMM > 2. The bimodal expression at TMM > 0 indicated the presence of low

expressed genes and active genes. The TMM > 1 removes all the low expression genes

retaining only the active ones.
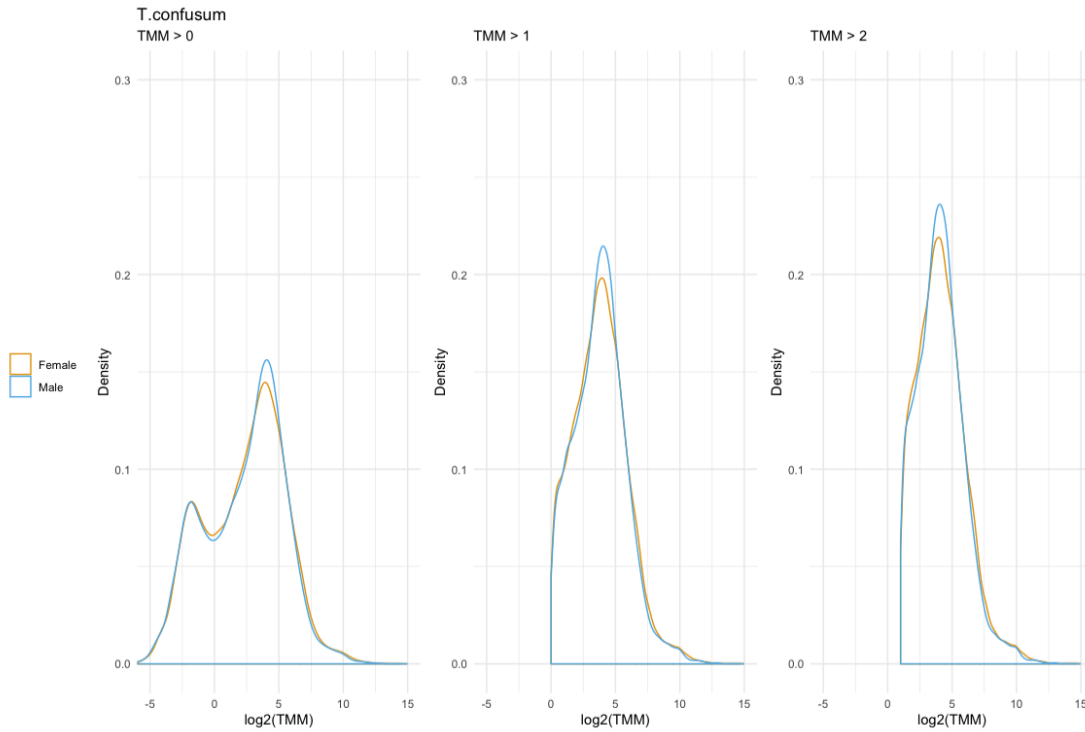
Figure S11. Minimum Expression Threshold for *Tribolium confusum* is highlighted at 2 levels (i.e) TMM > 1 and TMM > 2. The bimodal expression at TMM > 0 indicates the presence of low expressed genes and active genes. The TMM > 1 removes all the low expression genes retaining only the active ones.
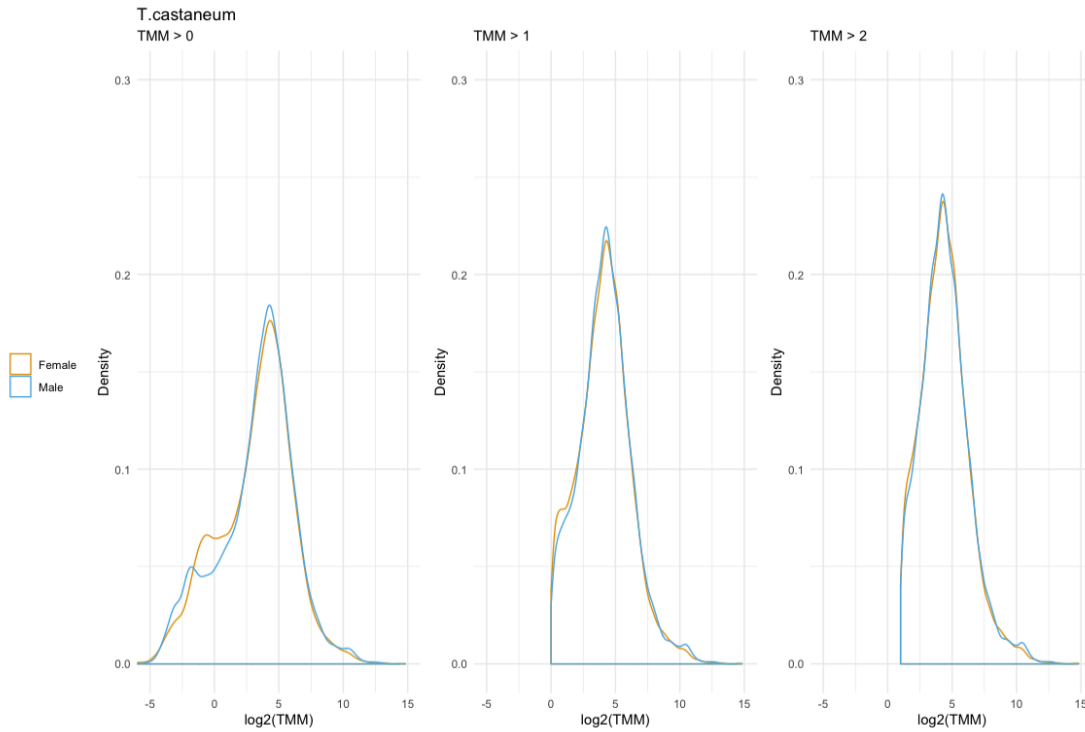
Figure S12. Minimum Expression Threshold for *Tribolium castaneum* is highlighted at 2 levels (i.e) TMM > 1 and TMM > 2. The bimodal expression at TMM > 0 indicates the presence of low expressed genes and active genes. The TMM > 1 removes all the low expression genes retaining only the active ones.
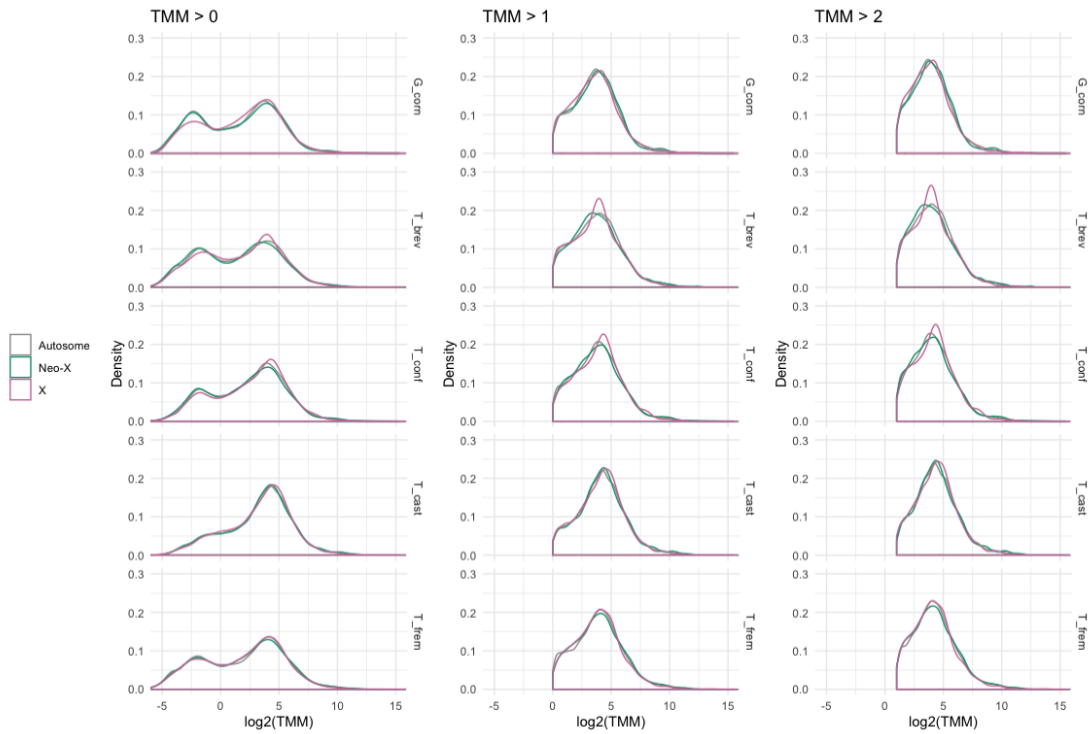
Figure S13. Minimum Expression Threshold for each chromosomal segment is shown at 2 levels (i.e) TMM > 1 and TMM > 2. The TMM > 1 removes all the low expression genes retaining only the active ones.
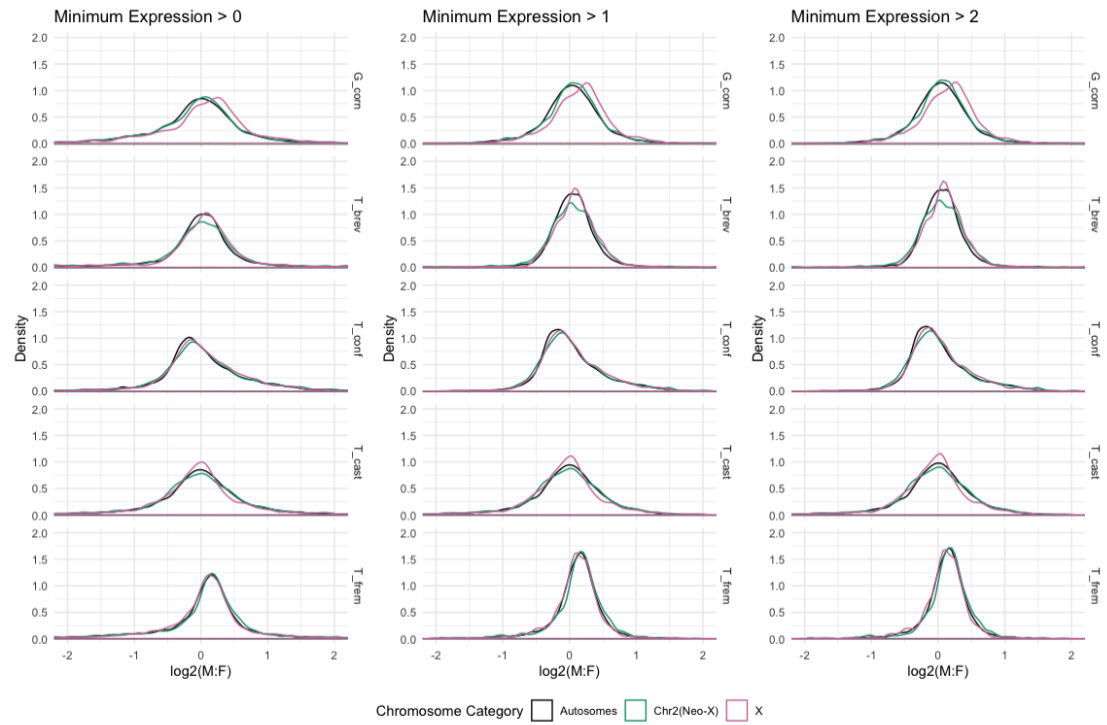
Figure S14. Minimum Expression Threshold for Male to Female ratio is shown at 2 levels (i.e) TMM > 1 and TMM > 2.
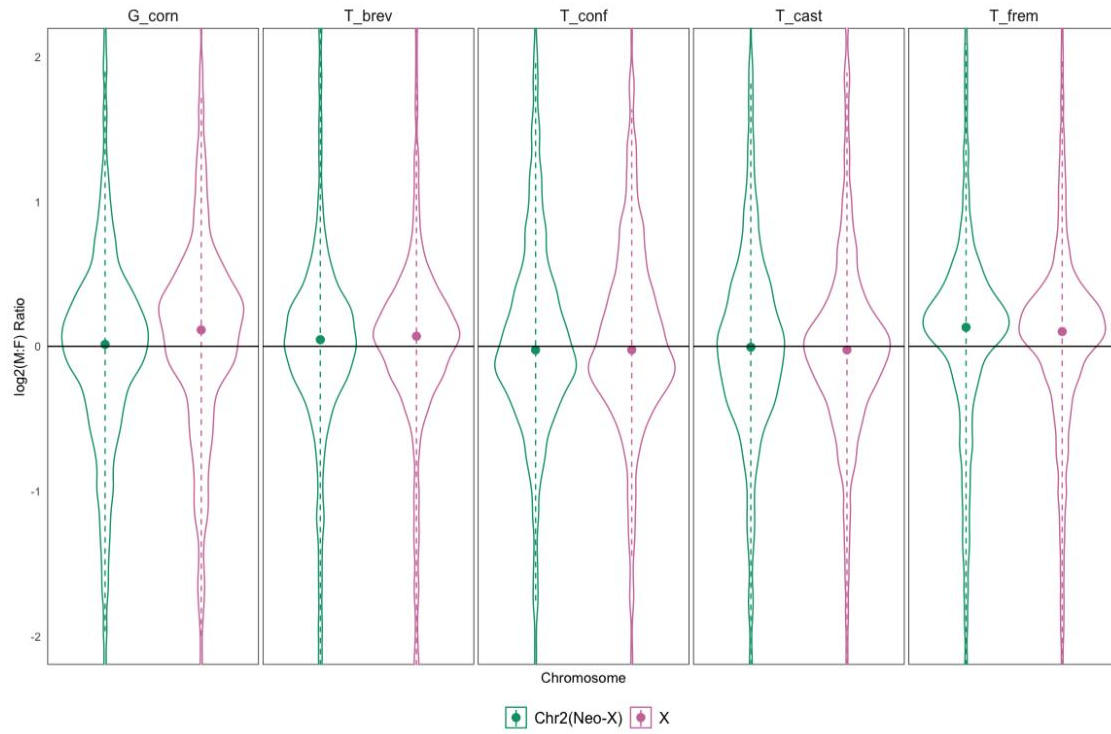
Figure S15. Violin Plots to highlight the M:F ratio of Neo-X(Chr 2) and X chromosome

across all 5 species at TMM > 1
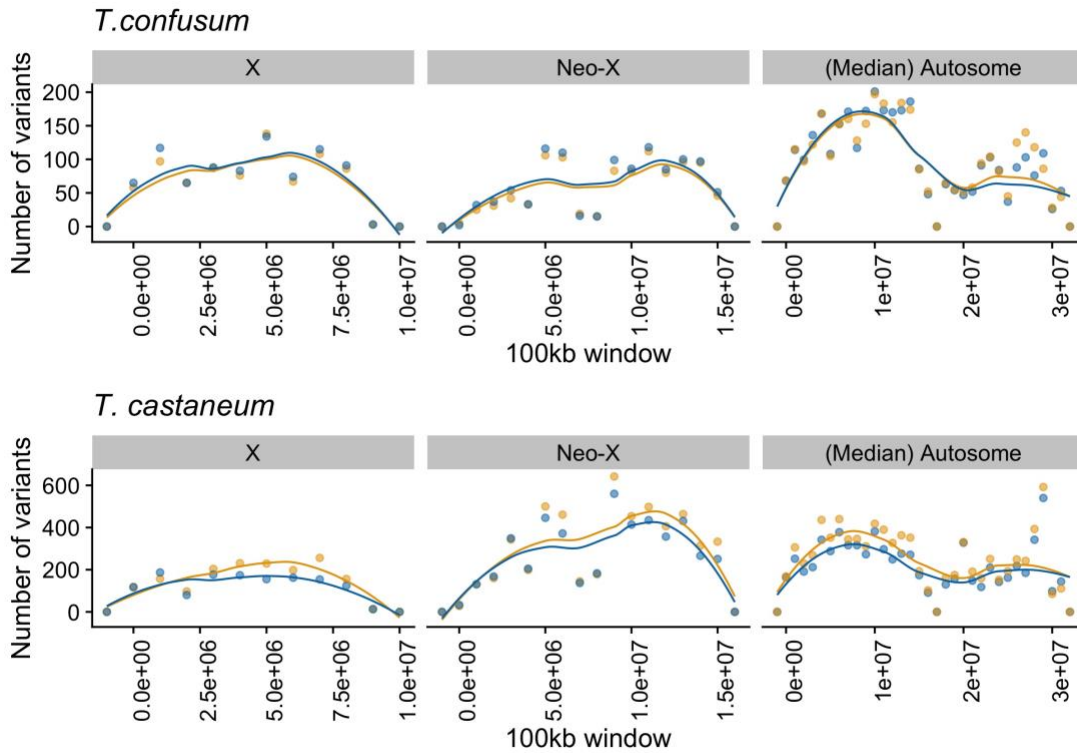
Figure S16. The number of variants in each 100kb interval is plotted for X and Neo-X of *Tribolium confusum* and *Tribolium castaneum*. For Autosomes, we represent the median number of variants and the trendline aids in comparing the number of variants across chromosomes.

Chapter 4

Single Cell RNA Sequencing reveals lack of X expression in *T.castaneum* testes in most cell clusters/types

Balan Ramesh[1] and Jeffrey Demuth[1]

[1]*Department of Biology, The University of Texas at Arlington, Arlington, Texas - 76013*

Abstract

The differing number of X chromosomes between sexes in the XY system

necessitates the need specialize regulatory mechanisms in somatic and germline cells.

Dosage compensation (DC) mechanisms, which may equalize average X-linked gene

expression with autosomes and or between the single X in males versus two X

chromosomes in females, are widely studied using somatic tissues. In male germline

tissues, the X-chromosome may experience meiotic sex chromosome inactivation

(MSCI), but this facet of sex chromosome regulation has not been as broadly studied

as dosage compensation. MSCI may be an exaptation of a more phylogenetically

ancient mechanism to avoid aneuploidies by meiotic silencing of unsynapsed chromatin

(MSUC), which is observed for autosomes of diverse eukaryotes. In MSCI the lack of

homology between heteromorphic sex chromosomes results in unpaired chromatin,

triggering MSUC. Although MSCI is well established in mammals, the presence or

absence of MSCI in meiotic cells of genetic model systems such as *Drosophila* and

Tribolium castaneum remains unclear. Here, we use single-cell RNA-seq (scRNA-seq),

on T. castaneum testes to investigate whether MSCI is present. Expression analysis

identifies 12 cell clusters, which is equal to the number of cell types defined by histology

studies of spermatogonial cell differentiation through mature spermatids. A broad

silencing of X expression across clusteres combined with a remarkable lack of

homologous gene expression compared to *Drosophila* spermatogenesis (the only other

insect with similarly detailed data) poses challenges to assigning correspondence

between computationally defined cell clusters and their specific meiotic stages.

However, our work provides the first account of single cell analysis from *Tribolium* and

clearly shows MSCI is present in the male germline. This work serves as a stepping

stone to understanding the forces governing sex chromosome regulation and evolution.

Introduction

Males and females carry unequal numbers of X chromosomes in XY sex determination systems. The loss of gene content on the Y chromosome by deletion and pseudogenization often results in X and Y chromosomes that lack homology across much or all of their length. This lack of homology may lead to transcriptional silencing during male meiosis, commonly referred to as meiotic sex chromosome inactivation (MSCI;(Turner, 2015, 2007). MSCI is the process in which the unsynapsed X and Y chromosomes are silenced during meiosis in spermatogenesis. In mammals, MSCI is initiated by compartmentalizing of the sex chromsomes within the nucleus into the sex- or XY-body. The XY-body is then transformed with epigenetic modifications that silence expression during the pachytene stage (McKee and Handel, 1993; Solari, 1974). MSCI is thought to be related to a broadly conserved process called meiotic silencing of unsynapsed chromatin (MSUC), which silences chromosomes that fail to pair with their homologous partner (Turner, 2007).

MSCI has evolved independently in XY and XO systems and is found in male germlines possessing heteromorphic sex chromosomes including grasshopper, nematode worm (Bean et al., 2004; Cabrero et al., 2007; Hense et al., 2007), and mammals. For example, in mice, it has been shown that MSCI down-regulates ~80% of X-linked genes in spermatocytes by ~10 fold (Mueller et al., 2008; Namekawa et al., 2006). MSCI has also been found in chickens which use the ZW system. Chicken MSCI is similar to XY MSCI (Schoenmakers et al., 2009) in that it also occurs during meiotic prophase and is marked by chromatin changes, implying a conserved mechanism based off MSUC. Chicken MSCI differs from XY MSCI though as the timing of when

MSCI occurs during pachytene is different. In eutherian mammals, MSCI is triggered by asynapsis (Baarends et al., 2005; Turner et al., 2005) while in chicken it is triggered by a homology search mechanism before the synapsis of Z and W (Namekawa et al., 2007). The transcriptional fates after meiosis are also different. In XY systems, the X and Y remains transcriptional suppressed during the post-meiotic period (Hense et al., 2007; Namekawa et al., 2007; Turner et al., 2006) while in ZW systems the suppression is lost by late diplotene (Namekawa et al., 2006).

Although failure of MSCI typically results in sterility, its ultimate purpose remains unclear (Turner, 2007). At least three hypotheses have been suggested. First, MSCI may be a form of host genome defense, either protecting against selfish genetic elements or preventing non-homologous recombination between the X and Y chromosomes (Hamilton, 1967; McKee and Handel, 1993; Meiklejohn and Tao, 2010; Namekawa and Lee, 2009). Second, MSCI may be advantageous as a suppressor of sex-ratio distorters (McKee and Handel, 1993). Finally, MSCI may be driven by sexually antagonist effects of X-linked genes. If the X chromosome accumulates female-beneficial genes which are detrimental to males it may be beneficial to reduce the X-linked genes' expression levels in the later stages of spermatogenesis (Wu and Yujun Xu, 2003). This latter hypothesis is consistent with the observed bias for testis expressed genes to be translocated from the X chromosome to autosomes in some species (Vibranovski et al., 2009).

Recent studies in *Drosphila* have cast doubt on the the ubiquity of MSCI because although expression is significantly suppressed below what can be explained by the absence of dosage compensation, the effect is established in premeiotic cells and

persists throughout meiotic cells. This potentially novel observation has been called "X suppression" (Meiklejohn et al., 2011). A recent study has provided evidence that X-linked genes have adapted for X suppression via the recruitment of strong testis-specific promoters (Landeen et al., 2016). The study identified and functionally validated a promoter element that drives strong expression in the testes, is enriched in the promoters of testes-specific genes on the X chromosome. Their results indicated that the X chromosome has evolved strong testis-specific promoters via gene-by-gene recruitment of sequence elements that counteract chromosome-wide transcriptional suppression (Landeen et al., 2016).

Single-cell RNA-Seq (scRNA-Seq) analyses of *Drosophila* male germline further show that both the sex chromosomes and autosomes are downregulated during and after meiosis, making effect of X suppression (or MSCI) challenging to differentiate from global downregulation in earlier studies using pooled cell populations . For example, using isolated testes tissue from different larval stages of *Drosophila,* Mikhaylova, and Nurminsky (2011) found no difference between Autosome and X-linked expression profile and suggested a lack of global MSCI. However, re-analysis of the microarray dataset from Mikhaylova and Nurminsky (2011) by Vibranovski and colleagues (2012) determined that the expression dataset was enriched with testis-biased genes, and analyzing non-biased gene dataset provided no evidence to support the lack of global MSCI. Recently, using scRNA-Seq, Witt and others (2021) classified the testes cells into somatic, pre-meiotic, meiotic, and post-meiotic cells and found that the cells are dosage compensated in somatic and pre-meiotic cells. Even with scRNA-seq, understanding DC and MSCI in meiotic cells are challenging since RNA molecules

transcribed in the previous cell stage might persist into later stages, skewing the expression ratio between the X chromosome and Autosome in meiotic cells (Witt et al., 2021).

Like *Drosophila*, *T.castaneum* is an insect model system with a high-quality reference genome whose dosage compensation mechanisms are well studied using various tissues. The type of DC in *T. castaneum* using whole body adults was first reported by Prince et al. (2010), and suggested dosage compensation was incomplete. In Chapter 3 (above) we analyzed transcriptomes from adult somatic tissues from multiple species and show that flour beetles have complete DC. The discrepancy in our findings thus appears largely to be contigent on the inclusion of gonads in the earlier study. A separate study that analyzed the transcriptomes of whole gonads (somatic and germline tissues included), shows the expression X-chromosome expression does not appear to be fully silenced, but also not dosage compensated (X:AA expression ratio =0.41; (Whittle et al., 2020). Unfortunately, the status of DC and MSCI specifically in germline tissues remains unclear because the X-linked expression observed by Whittle et al (2020), may be derived from somatic cells which were included in their gonad extractions.

The asynchronous timing of meiosis, coupled with spatial and morphological complexity, poses challenges to understanding stage specific gene regulation during *Tribolium* spermatogenesis. Spermatogenesis in *T.castaneum* testes is relatively well characterized (Dias et al., 2015), Fishman et al. 2017; Figure 1). The early sperm cells (spermatogonia) are small, round, and undergo first mitotic division to give rise to spermatocytes. The post-mitotic spermatocytes are relatively big and are organized in a

cyst, but meiosis occurs asynchronously. The secondary spermatocytes, which are found between meiosis I and II, are smaller than primary spermatocytes and have more condensed DNA. Early spermatids formed after meiosis, and a small gap between the acetylated tubulin-labeled axoneme and the nucleus are maintained throughout spermatogenesis. Later, the axoneme elongates and bends within the elliptical spermatid, while the tip of the axoneme forms a bulb. Following the oval form of the spermatid, the axoneme elongates and coils around the nucleus. After elongation, the spermatids, with round nuclei and long axonemes that appear as lollipops bundles, organized antiparallel to each other. Then, spermatids transition to ellipsis-shaped nuclei with a severely kinked neck region, and as the spermatids mature, the nuclei narrow and twist into an S-shape. Finally, the nuclei straighten in late spermatids, and spermatozoa have needle-shaped nuclei (Figure 1).

To better understand sex chromosome regulation in the *Tribolium* male germline, we employ high-throughput single-cell transcriptome sequencing from testes-derived heterogeneous cell populations. Based on whole transcriptome expression profiles, we identify 12 distinct cell type clusters. Reasuringly, this is similar to the number of stages described for adult spermatogenesis by by Fishman et al. (2017).   We find that X-linked genes are severely repressed/inactivated in most cell clusters; however, owing to a lack of homolgous gene expression with other model systems where meiotic marker genes are well defined, identifying the meiotic stage of each Tribolium testes cell cluster remains an active challenge. Our paper provides the first evidence of X-chromosome silencing in most cell types of the *Tribolium* male germline.

Materials and Methods

Preparation and sequencing of testis single-cell RNA-seq libraries

We used adult Tribolium castaneum-GA2 beetles in this study and adapted the single-cell preparation from (Mahadevaraju et al., 2021). Testes from 4 male beetles were dissected in cold aerated PBS. The resulting eight testes were placed in 200 ml of lysis buffer (100 ml 0.5% Trypsin EDTA + 100 ml of 4 mg/ml collagenase). The samples were incubated in the lysis buffer for 15 min in ice with gentle mixing every 30 seconds using a pipette. After incubation, we added 20 µL of 1% FBS to stop the enzymatic activity by gentle pipetting. The sample was filtered through a 40 mm cell strainer coated with 5µl of 0.04 % BSA solution followed by a 5 min centrifugation at 2000 rpm and 4°C. The resulting cell pellet was re-suspended in 15 µl of 0.04 % BSA solution before further processing. For cell counting, 5 µl of the single-cell suspension were mixed with 5 µl of the trypan blue dye, and the total cell number and the ratio between live and dead cells were analyzed using an automated cell counter (Bio-Rad Automated Cell Counter TC20™). This method yielded high numbers of single cells (~5 million live cells) with an average of 70–75% viability. We then submitted cells to the UT Southwestern genomics core for library preparation with the 10X Genomics chromium 3' kit v3 chemistry. Libraries were then sequenced using 150bp PE on Illumina NovaSeq at the North Texas Genome Center (UTA).

Processing of single-cell data

Illumina BCL files were converted into fastq files using Cellranger mkfastq. A reference genome was created with Cellranger v3.1.0 mkref, with the NCBI

T.castaneum genome. We used gffread v0.11.8 to convert the annotation (gff) file to gtf

format and used the reference genome, and the gtf file as a custom reference to run

Cellranger count demultiplexed the single-cell reads into a usable format for the Seurat

v3.0.6 R package (Satija et al., 2015). We kept all genes expressed in at least three

cells and all cells with at least 200 and at most 2500 genes for all the subsequent steps.

We ran Seurat NormalizeData, and ScaleData with default parameters, which

normalizes each gene count by the total gene count in a cell and log transforms

(natural-log transformed using log1p) the value after scaling utilizing a factor of 10000.

We then ran Seurat's default UMAP function and found clusters based on the first ten

principal components (resolution = 2). We determined the number of principal

components based on the standard deviation in each principal component illustrated in

the elbow plot (Supplementary).

Calculating relative RNA content from each cell type

As a proxy for RNA content, using the Unique Molecular Indices (UMIs) detected

from the X chromosome or autosomes in a cell cluster, we calculated RNA content per

cell as the sum of UMIs detected in X or Autosome divided by the number of cells in a

cluster. (Witt et al., 2021).

Results and Discussion

Our cell suspension from freshly dissected testes of adult *T.castaneum* yielded

~1600 cells/ul with ~70% viability. From this suspension, ~16,000 cells were used to

construct the Illumina library for sequencing (Supplementary Table S2).   We recovered

419.7 million reads of which 71.4% of the reads confidently mapped to the genome from

a total of ~8900 cells. On average, we mapped 47,192 reads per cell and detected a

median of 587 genes per cell.

Using t-Stochastic-Neighbor Embedding (t-SNE) from Cell Ranger v3.1.0, and

Uniform Manifold Approximation and Projection (UMAP) from Seurat v3.0.6, we reduced

the dimensionality of the gene/cell expression matrix to two primary axes and grouped

cells by their similarity of gene expression across hundreds of expression profiles. Both

UMAP and T-SNE grouped cells into twelve clusters based on transcriptome similarity

(Figure 2 and 3).   We also identified marker genes enriched in each cluster and top 10

marker genes differentially expressed among clusters (Supplementary Figure S3 and

Table S1). However, the identity of the cell type for each cluster remains elusive

because the *Tribolium* marker genes lack homology with known germline marker genes

from *Drosophila* spermatogenesis (Witt et al., 2021, 2019). It is worth noting that

Fishman and colleagues (2017) characterized the spermatogenesis into 12 distinct

stages, which is equal to the number of cell clusters in our analysis, starting from

spermatogonium to spermatozoa. Efforts to assign the meiotic stage of each gene

cluster are currently ongoing.

Even though the cell types are unknown, we analyzed whether X-linked genes

are expressed in various cell clusters. As a proxy, we used the UMI count for the

number of unique genes expressed in a given cell cluster and classified the genes into

either X chromosome or Autosome. All the cell clusters except cluster 11 represent no

genes from the X chromosome, suggesting a lack of X expression or inactivation

(Figure 3B). The developmental trajectory of *Drosophila* and *Tribolium castaneum* are

similar as they exhibit a continuous cluster of cellular differentiation and a cluster

separated from the others. This separate cluster was classified as somatic cells (Hub cells) in *Drosophila* based on the marker genes. If the separate cluster (cluster 11) in *Tribolium* is indeed a somatic cluster similar to *Drosophila*, it could explain previous bulk RNA-Seq expression profiles (Prince et al. 2010, Whittle et al. 2020). For example, the X:AA ratio of 0.41, found by Whittle et al 2020 could be explained by inclusion of dosage compensated somatic cells while the X-linked genes in germline cells are silenced.

Our results indicate that the X-chomosome is strongly suppressed or completely inactivated in the majority of *Tribolium* male germline. One potential explanation for such an early and sustained X-chromosome silencing may be the asynaptic pairing mechanism of X and Y chromosomes in polyphagan beetles such as *Tribolium* (Blackmon and Demuth 2014). Throughout the Coleoptera subphylum Polyphaga, the X and Y chromosome pair at a distance, held together by a protein scaffold that ensures proper meiotic segregation without synaptic pairing. Under these conditions we might expect MSUC to be active early as Anaphase I in spermatogenesis (Dutrillaux and Dutrillaux, 2009). To our knowledge, *Tribolium* is the first species where gene expression in this type of meiosis has been explored.

Ideally, assigning cell type to distinct cell stages and estimating the expression of X-linked genes in meiotic stages would aid in identifying the presence or absence of MSCI. However, another approach towards deducing MSCI would be to estimate the rate of RNA splicing and velocity. Estimating the RNA velocity will provide the direction of cellular differentiation of the spermatogonium and thus an approximation of meiotic cell clusters. We are currently working towards estimating the velocity.

Acknowledgements

Figures

Figure 1. (Fishman et al., 2017) characterized the *Tribolium* spermatogenesis and identified 12 distinct cell stages. The panel on the right by (Dias et al., 2015) describes the sperm ultrastructure (shown here) with an evolutionary perspective on the anti-parallel arrangements of sperm in cysts.
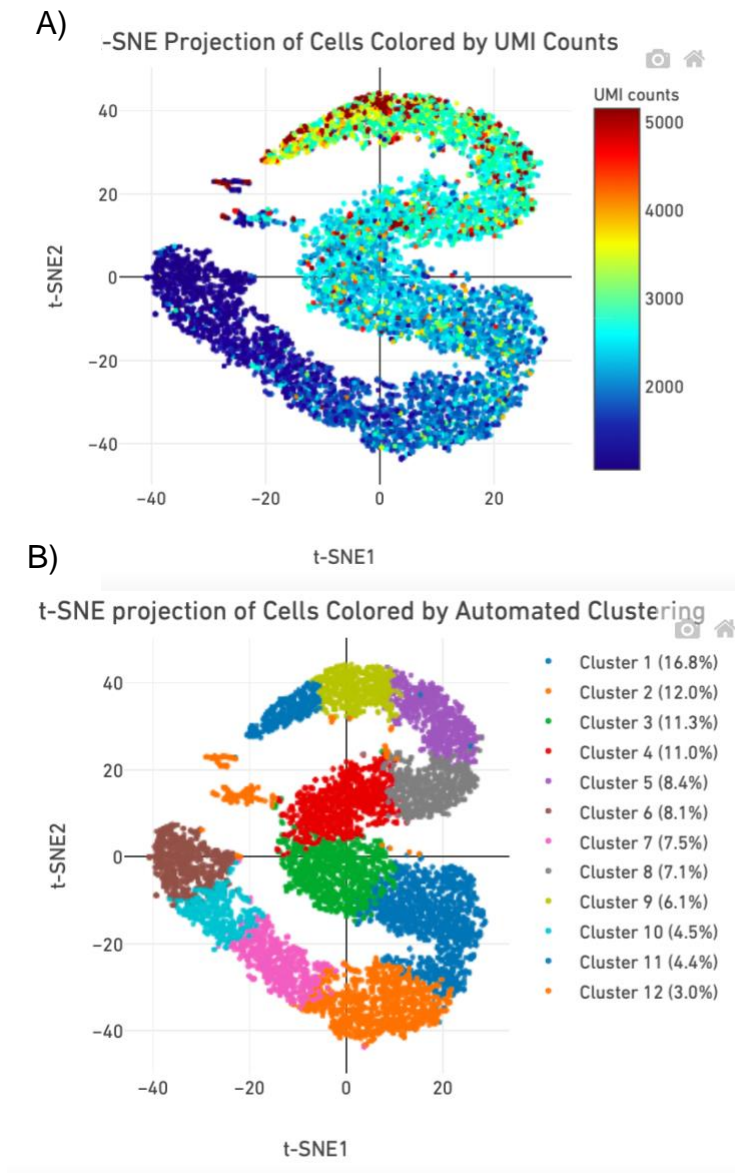
Figure 2 A) T-SNE plot of cell clusters from testes scRNA-Seq colored by UMI counts (a proxy for RNA content) showing a decreasing trend in the number of RNA. **B)** Reduced expression following meiosis in well documented in various species. Here cell clusters (from left to right 6,10,7,2,1) seem to have the lowest RNA content considering other cell clusters.
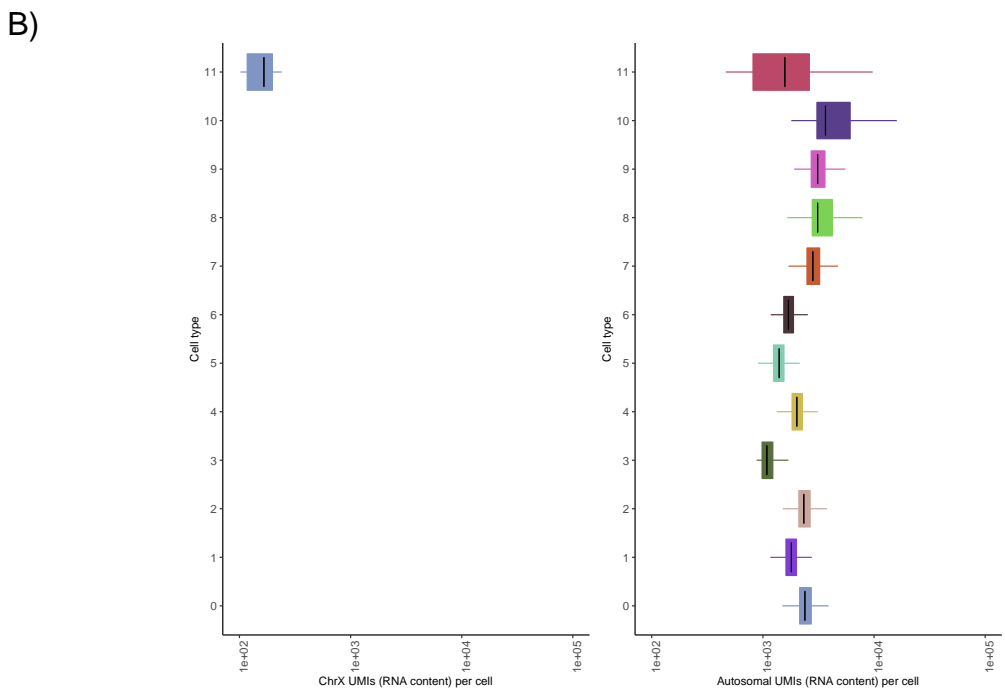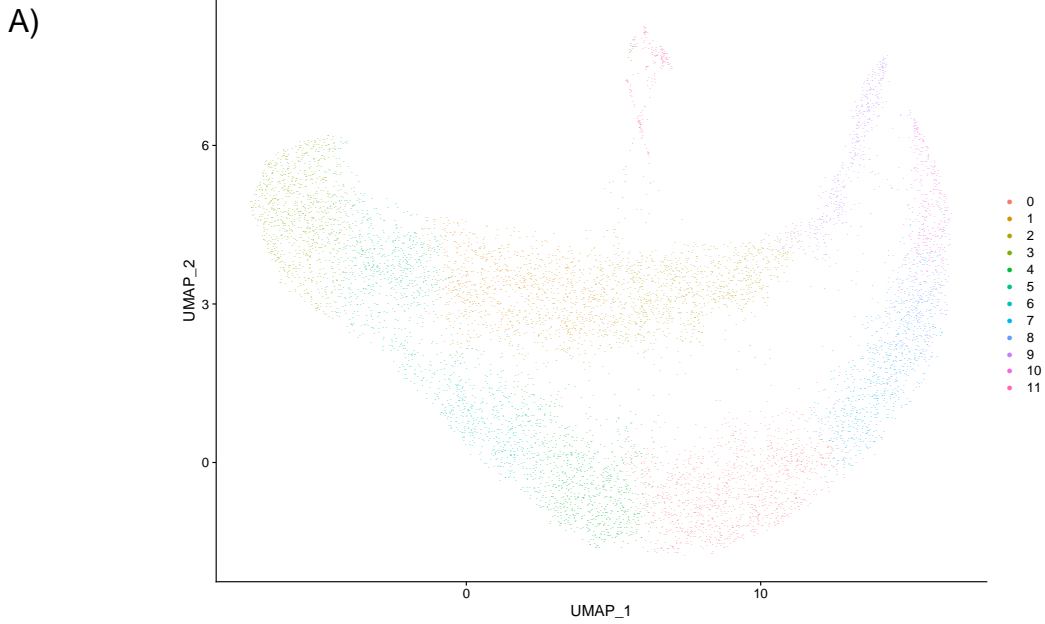
Figure 3. A) UMAP plot of cell clusters from testes scRNA-Seq. **B)** The panel below shows the UMI counts, a proxy for RNA content for each cell cluster from X chromosome and Autosome. All the cell clusters except cluster 11 shows no RNA content from X chromosome suggesting a lack of X expression.

81

Supplementary Figures
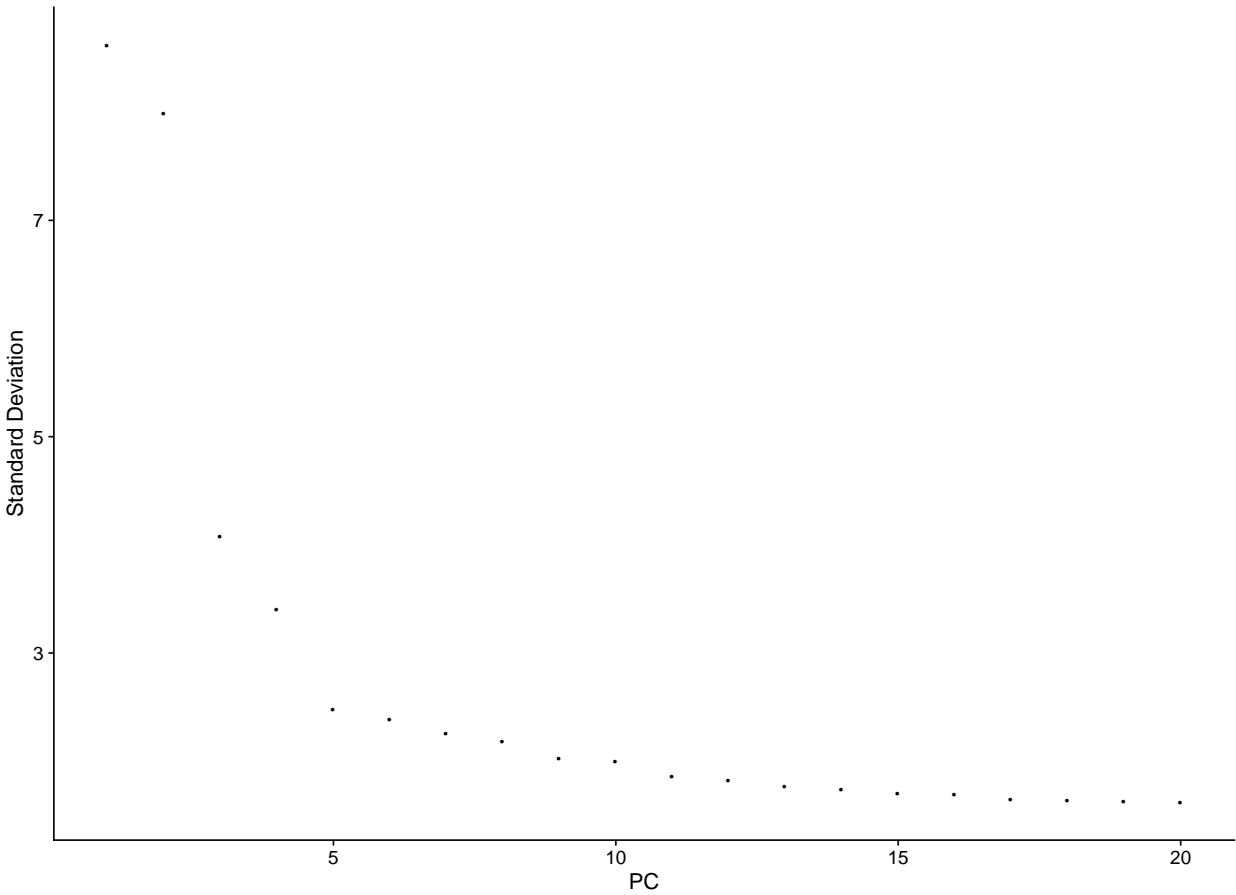
Figure S2. UMAP plot of cell clusters from testes scRNA-Seq.
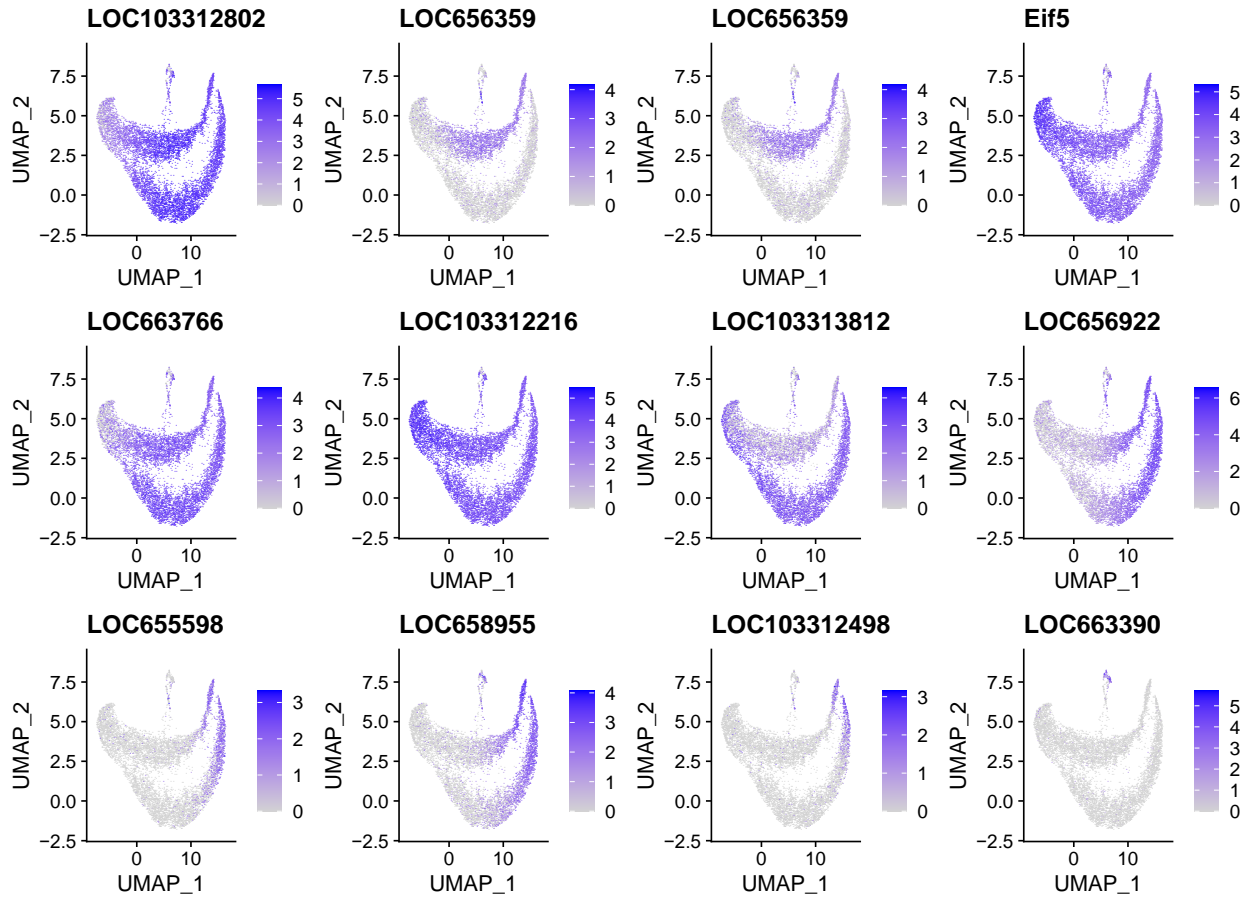
Figure S3. Marker genes expressed in each cluster is shown as a feature plot
embedded on top of the UMAP plot

Table S1. Top 10 marker genes from each cluster is shown below.

p_val	avg_logFC	pct.1	pct.2	p_val_adj	cluster	gene	Description	Chr
1.5E-289	0.478	1	0.95	1.5E-285	0	LOC103312802	serine/arginine repetitive matrix protein 4	LG4
3.4E-219	0.752	0.724	0.276	3.4E-215	0	LOC657276	protein disabled%2C transcript variant X1	LG7
7.4E-192	0.471	1	0.939	7.5E-188	0	LOC657733	dynein light chain D%2C transcript variant X1	LG6
6.4E-167	0.595	0.922	0.561	6.4E-163	0	LOC660169	translation initiation factor IF-3%2C mitochondrial	LG4
5.0E-164	0.516	0.999	0.662	5.1E-160	0	LOC658581	uncharacterized LOC658581%2C transcript variant X1	LG5
2.3E-162	0.531	0.988	0.664	2.3E-158	0	LOC662549	mitochondrial-processing peptidase subunit beta	LG6
1.6E-147	0.472	1	0.881	1.6E-143	0	LOC656353	mediator of RNA polymerase II transcription subunit 12%2C transcript variant X4	LG5
2.3E-142	0.550	0.844	0.436	2.3E-138	0	LOC103312701	uncharacterized LOC103312701	LG4
1.3E-131	0.511	0.913	0.617	1.3E-127	0	LOC103313917	No_Description	Chr
4.6E-103	0.524	0.688	0.375	4.6E-99	0	LOC659569	protein FAM60A%2C transcript variant X2	LG3
4.3E-159	0.782	0.66	0.284	4.3E-155	1	LOC656359	venom allergen 5	LG6
7.1E-155	0.819	0.631	0.29	7.2E-151	1	LOC107399257	No_Description	Chr
6.6E-91	0.398	0.965	0.866	6.6E-87	1	LOC655517	cold shock domain-containing protein E1%2C transcript variant X2	LG3

Table S1. Top 10 marker genes from each cluster is shown above (Cont.).

p_val	avg_logFC	pct.1	pct.2	p_val_adj	cluster	gene	Description	Chr
3.6E-86	0.341	0.991	0.911	3.6E-82	1	LOC103312481	uncharacterized LOC103312481	LG3
1.3E-84	0.368	0.955	0.909	1.3E-80	1	LOC664405	cilia- and flagella-associated protein 20	LG2
2.8E-63	0.413	0.834	0.694	2.8E-59	1	LOC662053	succinyl-CoA ligase subunit alpha%2C mitochondrial%2C transcript variant X5	LG7
1.9E-49	0.532	0.291	0.138	1.9E-45	1	LOC657144	RNA-binding protein 7	LG10
1.0E-31	0.409	0.493	0.376	1.0E-27	1	LOC659700	33 kDa inner dynein arm light chain%2C axonemal	Un
8.0E-23	0.346	0.448	0.348	8.0E-19	1	LOC656635	UPF0193 protein EVG1 homolog	LG8
4.5E-16	0.339	0.33	0.25	4.5E-12	1	LOC659395	kxDL motif-containing protein CG10681	LG7
0.0E+00	1.114	0.881	0.259	0.0E+00	2	LOC656359	venom allergen 5	LG6
0.0E+00	0.522	1	0.955	0.0E+00	2	LOC103312600	hypothetical protein	LG3
2.8E-253	0.492	1	0.953	2.8E-249	2	LOC103312802	serine/arginine repetitive matrix protein 4	LG4
5.9E-241	0.883	0.774	0.275	5.9E-237	2	LOC107399257	No_Description	Chr
2.4E-158	0.441	1	0.894	2.4E-154	2	LOC103315087	uncharacterized LOC103315087	LG3
1.2E-105	0.518	0.916	0.668	1.2E-101	2	LOC656504	EF-hand calcium-binding domain-containing protein 2%2C transcript variant X1	LG3
2.1E-98	0.616	0.474	0.197	2.1E-94	2	LOC103313284	sperm-tail PG-rich repeat-containing protein 2-like	LG5

Table S1. Top 10 marker genes from each cluster is shown above (Cont.).

p_val	avg_logFC	pct.1	pct.2	p_val_adj	cluster	gene	Description	Chr
3.4E-86	0.525	0.657	0.355	3.4E-82	2	LOC659700	33 kDa inner dynein arm light chain%2C axonemal	Un
4.4E-82	0.494	0.364	0.13	4.5E-78	2	LOC657144	RNA-binding protein 7	LG10
4.5E-79	0.433	0.908	0.666	4.5E-75	2	LOC655015	cytochrome c-type heme lyase	LG9
0.0E+00	0.802	1	0.992	0.0E+00	3	Eif5	eukaryotic translation initiation factor 5%2C transcript variant X2	LG4
0.0E+00	0.708	1	0.991	0.0E+00	3	LOC663790	outer dense fiber protein 3	LG7
1.1E-294	0.699	0.998	0.987	1.1E-290	3	LOC103313736	B1 protein	LG7
5.9E-291	0.811	0.991	0.96	5.9E-287	3	LOC662929	tubulin alpha-1 chain	LG2
1.9E-290	0.724	0.996	0.985	1.9E-286	3	LOC662329	tubulin-specific chaperone cofactor E-like protein%2C transcript variant X1	LG4
2.4E-176	0.708	0.94	0.927	2.4E-172	3	LOC103313914	uncharacterized LOC103313914%2C transcript variant X1	LG8
1.5E-152	0.839	0.832	0.75	1.5E-148	3	LOC656956	D-arabinitol dehydrogenase 1	LG8
6.3E-135	0.752	0.826	0.758	6.3E-131	3	LOC656062	thioredoxin reductase 1%2C mitochondrial%2C transcript variant X8	LG3
4.0E-112	0.888	0.71	0.598	4.1E-108	3	LOC657190	tubulin glycylase 3A-like	LG5
1.5E-37	0.699	0.529	0.491	1.5E-33	3	LOC657371	calpain-B%2C transcript variant X9	LG3

Table S1. Top 10 marker genes from each cluster is shown above (Cont.).

p_val	avg_logFC	pct.1	pct.2	p_val_adj	cluster	gene	Description	Chr
1.1E-90	0.375	0.993	0.881	1.1E-86	4	LOC663766	cys-loop ligand-gated ion channel subunit%2C transcript variant X3	LG4
2.2E-70	0.391	0.954	0.768	2.2E-66	4	LOC103313812	uncharacterized LOC103313812	LG7
7.7E-64	0.428	0.89	0.638	7.7E-60	4	LOC103313917	No_Description	Chr
5.0E-59	0.435	0.821	0.571	5.1E-55	4	LOC107398107	No_Description	Chr
1.1E-50	0.474	0.564	0.322	1.1E-46	4	LOC657276	protein disabled%2C transcript variant X1	LG7
1.4E-36	0.374	0.744	0.563	1.4E-32	4	LOC103312470	uncharacterized LOC103312470%2C transcript variant X1	LG3
2.4E-23	0.333	0.619	0.481	2.4E-19	4	LOC661529	ankyrin repeat and MYND domain-containing protein 2	Un
7.8E-20	0.337	0.354	0.234	7.8E-16	4	LOC107397581	No_Description	Chr
2.7E-19	0.342	0.279	0.171	2.8E-15	4	LOC103313461	DNA polymerase beta-like	LG6
7.8E-17	0.332	0.361	0.255	7.8E-13	4	LOC659986	nudC domain-containing protein 3	LG8
5.8E-168	0.391	1	0.993	5.9E-164	5	LOC103312216	hypothetical protein	LG3
3.5E-160	0.420	1	0.991	3.5E-156	5	LOC664168	enolase%2C transcript variant X1	LG9
4.2E-135	0.410	1	0.991	4.2E-131	5	LOC663790	outer dense fiber protein 3	LG7
5.0E-100	0.411	0.991	0.977	5.0E-96	5	LOC103312172	TPPP family protein CG45057	LG2
7.9E-75	0.404	0.962	0.947	8.0E-71	5	LOC662202	casein kinase I isoform gamma-3%2C transcript variant X3	LG8

Table S1. Top 10 marker genes from each cluster is shown above (Cont.).

p_val	avg_logFC	pct.1	pct.2	p_val_adj	cluster	gene	Description	Chr
2.5E-74	0.551	0.849	0.751	2.6E-70	5	LOC656956	D-arabinitol dehydrogenase 1	LG8
1.1E-61	0.396	0.94	0.928	1.1E-57	5	LOC103313914	uncharacterized LOC103313914%2C transcript variant X1	LG8
1.2E-57	0.393	0.933	0.919	1.2E-53	5	LOC663366	jmjC domain-containing protein 7	LG5
4.8E-28	0.431	0.666	0.62	4.9E-24	5	LOC663760	E3 ubiquitin-protein ligase MARCH5	LG9
9.7E-20	0.395	0.597	0.561	9.8E-16	5	LOC103314315	hypothetical protein	LG9
6.3E-77	0.547	0.969	0.774	6.3E-73	6	LOC103313812	uncharacterized LOC103313812	LG7
1.9E-58	0.581	0.824	0.581	1.9E-54	6	LOC107398107	No_Description	Chr
1.2E-57	0.341	0.998	0.966	1.2E-53	6	LOC659071	40S ribosomal protein S14%2C transcript variant X2	Un
3.9E-29	0.419	0.732	0.618	3.9E-25	6	LOC657690	acyl-protein thioesterase 1	LG3
1.1E-23	0.344	0.815	0.653	1.1E-19	6	LOC103313917	No_Description	Chr
6.5E-17	0.338	0.659	0.556	6.5E-13	6	LOC657342	farnesol dehydrogenase	LG6
3.2E-15	0.353	0.598	0.525	3.2E-11	6	LOC661145	dnaJ homolog subfamily B member 14	LG3
7.9E-11	0.407	0.327	0.241	8.0E-07	6	LOC107397581	No_Description	Chr
6.4E-10	0.335	0.406	0.32	6.4E-06	6	LOC660935	alpha-esterase like protein E3	LG8
2.2E-09	0.364	0.321	0.241	2.2E-05	6	LOC103313190	zinc finger protein 2-like	LG3
1.7E-229	1.253	1	0.718	1.7E-225	7	LOC656922	keratin-associated protein 10-8	LG3

Table S1. Top 10 marker genes from each cluster is shown above (Cont.).

p_val	avg_logFC	pct.1	pct.2	p_val_adj	cluster	gene	Description	Chr
8.1E-190	0.911	1	0.696	8.1E-186	7	LOC658581	uncharacterized LOC658581%2C transcript variant X1	LG5
4.7E-182	0.987	1	0.579	4.7E-178	7	LOC661355	nuclear protein 1	LG8
3.5E-173	0.915	1	0.706	3.5E-169	7	LOC103313149	venom allergen 5-like	LG5
6.0E-172	0.864	1	0.59	6.1E-168	7	LOC103313150	venom allergen 5-like	LG5
1.5E-170	0.963	1	0.682	1.5E-166	7	LOC661071	uncharacterized LOC661071%2C transcript variant X1	LG9
3.6E-170	0.925	0.984	0.537	3.6E-166	7	LOC103312599	uncharacterized LOC103312599	LG4
1.2E-166	0.902	0.982	0.481	1.2E-162	7	LOC660609	receptor expression-enhancing protein 5	LG7
4.6E-166	0.920	0.962	0.47	4.6E-162	7	LOC103312701	uncharacterized LOC103312701	LG4
2.3E-164	0.840	1	0.762	2.3E-160	7	LOC657573	uncharacterized LOC657573%2C transcript variant X1	LG5
1.9E-301	1.168	0.829	0.151	1.9E-297	8	LOC655598	TBC1 domain family member 19	LG4
9.7E-251	1.389	1	0.434	9.8E-247	8	LOC660132	bax inhibitor 1	LG9
9.9E-242	1.556	1	0.684	9.9E-238	8	LOC661071	uncharacterized LOC661071%2C transcript variant X1	LG9
2.0E-235	1.203	1	0.407	2.0E-231	8	LOC655849	mitochondrial carrier homolog 2	LG3
1.1E-224	1.283	1	0.708	1.1E-220	8	LOC103313149	venom allergen 5-like	LG5
4.4E-218	1.285	1	0.593	4.4E-214	8	LOC103313150	venom allergen 5-like	LG5
1.3E-209	1.104	0.988	0.394	1.3E-205	8	LOC658485	carbonic anhydrase 1	LG2

Table S1. Top 10 marker genes from each cluster is shown above (Cont.).

p_val	avg_logFC	pct.1	pct.2	p_val_adj	cluster	gene	Description	Chr
7.0E-207	1.100	0.996	0.464	7.0E-203	8	LOC656649	tubulin alpha 1-like protein	LG2
3.0E-202	1.225	1	0.721	3.0E-198	8	LOC656922	keratin-associated protein 10-8	LG3
1.6E-201	1.142	1	0.582	1.6E-197	8	LOC661355	nuclear protein 1	LG8
1.1E-266	1.518	0.987	0.37	1.2E-262	9	LOC658955	protein-methionine sulfoxide oxidase MICAL3%2C transcript variant X2	LG8
3.6E-263	1.608	0.989	0.421	3.6E-259	9	LOC103314312	cytochrome c oxidase subunit 7A%2C mitochondrial-like	LG9
8.7E-260	1.302	0.787	0.166	8.7E-256	9	LOC655212	peptidyl-prolyl cis-trans isomerase FKBP8%2C transcript variant X1	LG7
1.6E-234	1.625	1	0.594	1.6E-230	9	LOC103313150	venom allergen 5-like	LG5
1.6E-229	1.325	1	0.583	1.6E-225	9	LOC661355	nuclear protein 1	LG8
9.3E-229	1.411	0.989	0.409	9.3E-225	9	LOC655849	mitochondrial carrier homolog 2	LG3
1.9E-220	1.450	0.985	0.436	1.9E-216	9	LOC660132	bax inhibitor 1	LG9
2.6E-214	1.406	0.989	0.466	2.6E-210	9	LOC656649	tubulin alpha 1-like protein	LG2
3.6E-210	1.441	1	0.685	3.6E-206	9	LOC661071	uncharacterized LOC661071%2C transcript variant X1	LG9
7.0E-205	1.469	0.741	0.192	7.0E-201	9	LOC107398496	stabilizer of axonemal microtubules 2	LG8
0.0E+00	1.437	0.822	0.067	0.0E+00	10	LOC103312498	trichohyalin	LG9
9.2E-288	1.365	0.953	0.193	9.2E-284	10	LOC656809	proton-coupled amino acid transporter 1-like	Un

Table S1. Top 10 marker genes from each cluster is shown above (Cont.).

p_val	avg_logFC	pct.1	pct.2	p_val_adj	cluster	gene	Description	Chr
1.5E-278	1.458	0.94	0.189	1.5E-274	10	LOC107398496	stabilizer of axonemal microtubules 2	LG8
2.9E-239	1.347	0.937	0.22	3.0E-235	10	LOC664561	membrane-associated progesterone receptor component 1	LG4
5.2E-231	1.690	1	0.442	5.2E-227	10	LOC660132	bax inhibitor 1	LG9
1.8E-226	1.573	1	0.415	1.8E-222	10	LOC655849	mitochondrial carrier homolog 2	LG3
3.5E-210	1.484	1	0.471	3.5E-206	10	LOC656649	tubulin alpha 1-like protein	LG2
9.4E-203	1.618	1	0.599	9.4E-199	10	LOC103313150	venom allergen 5-like	LG5
3.6E-194	1.582	1	0.689	3.7E-190	10	LOC661071	uncharacterized LOC661071%2C transcript variant X1	LG9
1.9E-190	1.378	1	0.712	1.9E-186	10	LOC103313149	venom allergen 5-like	LG5
0.0E+00	2.997	0.533	0.023	0.0E+00	11	LOC663390	uncharacterized LOC663390%2C transcript variant X1	LG2
0.0E+00	2.898	0.417	0.008	0.0E+00	11	LOC659226	cathepsin L precursor	LG10
9.0E-301	2.783	0.719	0.058	9.0E-297	11	LOC662445	ribosomal protein S18	LG5
1.2E-288	2.817	0.704	0.058	1.2E-284	11	LOC659458	60S ribosomal protein L35	LG5
2.1E-285	3.530	0.799	0.083	2.1E-281	11	LOC656276	60S acidic ribosomal protein P2	LG6
1.7E-275	2.976	0.704	0.061	1.7E-271	11	LOC659536	60S ribosomal protein L38%2C transcript variant X2	LG8
2.3E-271	3.142	0.754	0.075	2.3E-267	11	LOC658148	60S ribosomal protein L37	LG9
2.7E-271	3.104	0.508	0.029	2.7E-267	11	LOC655420	ferritin subunit	LG2

Table S1. Top 10 marker genes from each cluster is shown above (Cont.).

p_val	avg_logFC	pct.1	pct.2	p_val_adj	cluster	gene	Description	Chr
6.2E-251	3.600	0.598	0.047	6.3E-247	11	LOC655492	ferritin heavy chain%2C transcript variant X2	LG2
2.5E-229	3.014	0.447	0.026	2.5E-225	11	LOC659790	gamma-interferon-inducible lysosomal thiol reductase	LG6

Table S2. Quality of the cells before and after library prep.

10X9_7380		Cell Counts			Library prep					Post library QC		
#	Sample Name	cells/ul	% viability	target	ul cells	ul H2O	ng/ul check	total ng	BC	Avg bp	ng/ul	nM
1	UTA	~1600	~70%	10K	10	36.6	1.38	13.8	G3	462	36.3	119.0

References

Agashe, D. and Bolnick, D.I., 2010. Intraspecific genetic variation and competition interact to influence niche expansion. Proceedings of the Royal Society B: Biological Sciences, 277(1696), pp.2915-2924.

Agashe, D., Falk, J.J. and Bolnick, D.I., 2011. Effects of founding genetic variation on adaptation to a novel resource. Evolution: International Journal of Organic Evolution, 65(9), pp.2481-2491. https://doi.org/10.1111/j.1558-5646.2011.01307.x

Angelini, D.R. and Jockusch, E.L., 2008. Relationships among pest flour beetles of the genus Tribolium (Tenebrionidae) inferred from multiple molecular markers. Molecular phylogenetics and evolution, 46(1), pp.127-141. https://doi.org/10.1016/j.ympev.2007.08.017

Baarends, W.M., Wassenaar, E., van der Laan, R., Hoogerbrugge, J., Sleddens-Linkels, E., Hoeijmakers, J.H.J., de Boer, P., Grootegoed, J.A., 2005. Silencing of Unpaired Chromatin and Histone H2A Ubiquitination in Mammalian Meiosis. MCB 25, 1041–1053. https://doi.org/10.1128/MCB.25.3.1041-1053.2005

Bachtrog, D., 2006. A dynamic view of sex chromosome evolution. Current Opinion in Genetics & Development 16, 578–585. https://doi.org/10.1016/j.gde.2006.10.007

Bean, C.J., Schaner, C.E., Kelly, W.G., 2004. Meiotic pairing and imprinted X chromatin assembly in Caenorhabditis elegans. Nat Genet 36, 100–105. https://doi.org/10.1038/ng1283

Beeman, R.W. and Brown, S.J., 1999. RAPD-based genetic linkage maps of Tribolium

castaneum. Genetics, 153(1), pp.333-338.

https://www.genetics.org/content/genetics/153/1/333.full.pdf

Beeman, R.W., 1987. A homoeotic gene cluster in the red flour beetle. Nature,

327(6119), pp.247-249. https://doi.org/10.1038/327247a0

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.H., Xie, D., Suchard, M.A.,

Rambaut, A. and Drummond, A.J., 2014. BEAST 2: a software platform for

Bayesian evolutionary analysis. PLoS Comput Biol, 10(4), p.e1003537.

https://doi.org/10.1371/journal.pcbi.1003537

Brown, S.J., Shippy, T.D., Miller, S., Bolognesi, R., Beeman, R.W., Lorenzen, M.D.,

Bucher, G., Wimmer, E.A. and Klingler, M., 2009. The red flour beetle, Tribolium

castaneum (Coleoptera): a model for studies of development and pest biology.

Cold Spring Harbor Protocols, 2009(8), pp.pdb-emo126.

https://doi.org/10.1101/pdb.emo126

Cabrero, J., Teruel, M., Carmona, F.D., Jiménez, R., Camacho, J.P.M., 2007. Histone

H3 lysine 9 acetylation pattern suggests that X and B chromosomes are silenced

during entire male meiosis in a grasshopper. Cytogenet Genome Res 119, 135–

142. https://doi.org/10.1159/000109630

Carruthers, M., Yurchenko, A.A., Augley, J.J., Adams, C.E., Herzyk, P., Elmer, K.R.,

2018. De novo transcriptome assembly, annotation and comparison of four

ecological and evolutionary model salmonid fish species. BMC Genomics 19, 32.

https://doi.org/10.1186/s12864-017-4379-x

Charlesworth, B., 1991. The evolution of sex chromosomes. Science 251, 1030–1033.

Charlesworth, D., Charlesworth, B., Marais, G., 2005. Steps in the evolution of
heteromorphic sex chromosomes. Heredity 95, 118–128.
https://doi.org/10.1038/sj.hdy.6800697

Chen, J., Wang, M., He, X., Yang, J.-R., Chen, X., 2020. The Evolution of Sex
Chromosome Dosage Compensation in Animals. Journal of Genetics and
Genomics S1673852720301740. https://doi.org/10.1016/j.jgg.2020.10.005

Chen, X., Zhang, J., 2016. The X to Autosome Expression Ratio in Haploid and Diploid
Human Embryonic Stem Cells. Mol Biol Evol 33, 3104–3107.
https://doi.org/10.1093/molbev/msw187

Demuth, J.P. and Wade, M.J., 2007. Population differentiation in the beetle Tribolium
castaneum. I. Genetic architecture. Evolution, 61(3), pp.494-509.
https://doi.org/10.1111/j.1558-5646.2007.00048.x

Demuth, J.P., De Bie, T., Stajich, J.E., Cristianini, N. and Hahn, M.W., 2006. The
evolution of mammalian gene families. PloS one, 1(1), p.e85.
https://doi.org/10.1371/journal.pone.0000085

Denell, R., 2008. Establishment of Tribolium as a genetic model system and its early
contributions to evo-devo. Genetics, 180(4), pp.1779-1786.
https://doi.org/10.1534/genetics.104.98673

Dias, G., Lino-Neto, J., Mercati, D., Dallai, R., 2015. The sperm ultrastructure and
spermiogenesis of Tribolium castaneum (Coleoptera: Tenebrionidae) with
evidence of cyst degeneration. Micron 73, 21–27.
https://doi.org/10.1016/j.micron.2015.03.003

Dutrillaux, A.M., Dutrillaux, B., 2009. Sex Chromosome Rearrangements in Polyphaga

    Beetles. Sex Dev 3, 43–54. https://doi.org/10.1159/000200081

Emms, D.M. and Kelly, S., 2019. OrthoFinder: phylogenetic orthology inference for

    comparative genomics. Genome biology, 20(1), pp.1-14.

    https://doi.org/10.1186/s13059-019-1832-y

Emms, D.M., Kelly, S., 2019. OrthoFinder: phylogenetic orthology inference for

    comparative genomics. Genome Biol 20, 238. https://doi.org/10.1186/s13059-

    019-1832-y

Esseghir, S., Ready, P.D. and Ben-Ismail, R., 2000. Speciation of Phlebotomus

    sandflies of the subgenus Larroussius coincided with the late Miocene-Pliocene

    aridification of the Mediterranean subregion. Biological Journal of the Linnean

    Society, 70(2), pp.189-219. https://doi.org/10.1111/j.1095-8312.2000.tb00207.x

Farnworth, M.S., Eckermann, K.N., Ahmed, H.M., Mühlen, D.S., He, B. and Bucher, G.,

    2020. The red flour beetle as model for comparative neural development:

    Genome editing to mark neural cells in Tribolium brain development. In Brain

    Development (pp. 191-217). Humana, New York, NY.

    https://doi.org/10.1007/978-1-4939-9732-9_11

Farrell, B.D., 2001. Evolutionary assembly of the milkweed fauna: Cytochrome oxidase I

    and the age of Tetraopesbeetles. Molecular phylogenetics and evolution, 18(3),

    pp.467-478.

Fishman, E.L., Jo, K., Ha, A., Royfman, R., Zinn, A., Krishnamurthy, M., Avidor-Reiss,

    T., 2017. Atypical centrioles are present in Tribolium sperm. Open Biol. 7,

    160334. https://doi.org/10.1098/rsob.160334

Fu, L., Niu, B., Zhu, Z., Wu, S. and Li, W., 2012. CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics, 28(23), pp.3150-3152. https://doi.org/10.1093/bioinformatics/bts565

Fu, L., Niu, B., Zhu, Z., Wu, S., Li, W., 2012. CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics 28, 3150–3152. https://doi.org/10.1093/bioinformatics/bts565

Furman, B.L.S., Metzger, D.C.H., Darolti, I., Wright, A.E., Sandkam, B.A., Almeida, P., Shu, J.J., Mank, J.E., 2020. Sex Chromosome Evolution: So Many Exceptions to the Rules. Genome Biology and Evolution 12, 750–763. https://doi.org/10.1093/gbe/evaa081

Gahlan, P., Singh, H., Shankar, R., Sharma, N., Kumari, A., Chawla, V., Ahuja, P., Kumar, S., 2012. De novo sequencing and characterization of Picrorhiza kurrooa transcriptome at two temperatures showed major transcriptome adjustments. BMC Genomics 13, 126. https://doi.org/10.1186/1471-2164-13-126

Ghent, A.W., 1963. Studies of behavior of the Tribolium flour beetles. I. Contrasting responses of T. castaneum and T. confusum to fresh and conditioned flours. Ecology, 44(2), pp.269-283. https://doi.org/10.2307/1932174

Gilles, A.F., Schinko, J.B. and Averof, M., 2015. Efficient CRISPR-mediated gene targeting and transgene replacement in the beetle Tribolium castaneum. Development, 142(16), pp.2832-2839. https://doi.org/ 10.1242/dev.125054

Goodnight, C.J., 1990. Experimental studies of community evolution I: The response to selection at the community level. Evolution, 44(6), pp.1614-1624. https://doi.org/10.1111/j.1558-5646.1990.tb03850.x

Graves, J.A.M., 2014. The epigenetic sole of sex and dosage compensation. Nat Genet 46, 215–217. https://doi.org/10.1038/ng.2903

Gu, L., Walters, J.R., 2017. Evolution of Sex Chromosome Dosage Compensation in Animals: A Beautiful Theory, Undermined by Facts and Bedeviled by Details. Genome Biology and Evolution 9, 2461–2476. https://doi.org/10.1093/gbe/evx154

Haak, M., Vinke, S., Keller, W., Droste, J., Rückert, C., Kalinowski, J., Pucker, B., 2018. High Quality de Novo Transcriptome Assembly of Croton tiglium. Front. Mol. Biosci. 5, 62. https://doi.org/10.3389/fmolb.2018.00062

Haas, B. and Papanicolaou, A.J.G.S., 2016. TransDecoder (find coding regions within transcripts). Google Scholar.

Haas, B., Papanicolaou, A., 2017. TransDecoder.

Haas, B.J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P.D., Bowden, J., Couger, M.B., Eccles, D., Li, B., Lieber, M. and MacManes, M.D., 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nature protocols, 8(8), pp.1494-1512. https://doi.org/10.1038/nprot.2013.084

Haas, B.J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P.D., Bowden, J., Couger, M.B., Eccles, D., Li, B., Lieber, M., MacManes, M.D., Ott, M., Orvis, J., Pochet, N., Strozzi, F., Weeks, N., Westerman, R., William, T., Dewey, C.N., Henschel, R., LeDuc, R.D., Friedman, N., Regev, A., 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference

generation and analysis. Nat Protoc 8, 1494–1512.

https://doi.org/10.1038/nprot.2013.084

Halali, S., Brakefield, P.M., Collins, S.C. and Brattström, O., 2020. To mate, or not to

mate: The evolution of reproductive diapause facilitates insect radiation into

African savannahs in the Late Miocene. Journal of Animal Ecology, 89(5),

pp.1230-1241. https://doi.org/10.1111/1365-2656.13178

Hamilton, W.D., 1967. Extraordinary Sex Ratios. Science 156, 477–488.

https://doi.org/10.1126/science.156.3774.477

Harano, T., Okada, K., Nakayama, S., Miyatake, T., Hosken, D.J., 2010. Intralocus

Sexual Conflict Unresolved by Sex-Limited Trait Expression. Current Biology 20,

2036–2039. https://doi.org/10.1016/j.cub.2010.10.023

Hart, T., Komori, H., LaMere, S., Podshivalova, K., Salomon, D.R., 2013. Finding the

active genes in deep RNA-seq gene expression studies. BMC Genomics 14,

778. https://doi.org/10.1186/1471-2164-14-778

Hebenstreit, D., Fang, M., Gu, M., Charoensawan, V., van Oudenaarden, A.,

Teichmann, S.A., 2011. RNA sequencing reveals two major classes of gene

expression levels in metazoan cells. Mol Syst Biol 7, 497.

https://doi.org/10.1038/msb.2011.28

Helfrich, P., Rieb, E., Abrami, G., Lücking, A. and Mehler, A., 2018, May.

TreeAnnotator: versatile visual annotation of hierarchical text relations. In

Proceedings of the Eleventh International Conference on Language Resources

and Evaluation (LREC 2018). https://www.aclweb.org/anthology/L18-1308.pdf

Hense, W., Baines, J.F., Parsch, J., 2007. X Chromosome Inactivation during

    Drosophila Spermatogenesis. PLoS Biol 5, e273.

    https://doi.org/10.1371/journal.pbio.0050273

Herndon, N., Shelton, J., Gerischer, L., Ioannidis, P., Ninova, M., Dönitz, J.,

    Waterhouse, R.M., Liang, C., Damm, C., Siemanowski, J. and Kitzmann, P.,

    2020. Enhanced genome assembly and a new official gene set for Tribolium

    castaneum. BMC genomics, 21(1), pp.1-13. https://doi.org/10.1186/s12864-019-

    6394-6

Herndon, N., Shelton, J., Gerischer, L., Ioannidis, P., Ninova, M., Dönitz, J.,

    Waterhouse, R.M., Liang, C., Damm, C., Siemanowski, J., Kitzmann, P., Ulrich,

    J., Dippel, S., Oberhofer, G., Hu, Y., Schwirz, J., Schacht, M., Lehmann, S.,

    Montino, A., Posnien, N., Gurska, D., Horn, T., Seibert, J., Vargas Jentzsch, I.M.,

    Panfilio, K.A., Li, J., Wimmer, E.A., Stappert, D., Roth, S., Schröder, R., Park, Y.,

    Schoppmeier, M., Chung, H.-R., Klingler, M., Kittelmann, S., Friedrich, M., Chen,

    R., Altincicek, B., Vilcinskas, A., Zdobnov, E., Griffiths-Jones, S., Ronshaugen,

    M., Stanke, M., Brown, S.J., Bucher, G., 2020. Enhanced genome assembly and

    a new official gene set for Tribolium castaneum. BMC Genomics 21, 47.

    https://doi.org/10.1186/s12864-019-6394-6

Hidalgo-Galiana, A. and Ribera, I., 2011. Late Miocene diversification of the genus

    Hydrochus (Coleoptera, Hydrochidae) in the west Mediterranean area. Molecular

    Phylogenetics and Evolution, 59(2), pp.377-385.

    https://doi.org/10.1016/j.ympev.2011.01.018

Hinton, H.E., 1948. A synopsis of the genus Tribolium Macleay, with some remarks on

the evolution of its species-groups (Coleoptera, Tenebrionidae). Bulletin of

entomological research, 39(1), pp.13-55.

https://doi.org/10.1017/S0007485300024287

https://doi.org/10.1006/mpev.2000.0888

Jiang, X., Biedler, J.K., Qi, Y., Hall, A.B., Tu, Z., 2015. Complete Dosage Compensation

in Anopheles stephensi and the Evolution of Sex-Biased Genes in Mosquitoes.

Genome Biol Evol 7, 1914–1924. https://doi.org/10.1093/gbe/evv115

Julien, P., Brawand, D., Soumillon, M., Necsulea, A., Liechti, A., Schütz, F., Daish, T.,

Grützner, F., Kaessmann, H., 2012. Mechanisms and Evolutionary Patterns of

Mammalian and Avian Dosage Compensation. PLoS Biol 10, e1001328.

https://doi.org/10.1371/journal.pbio.1001328

Junqueira, A.C.M., Azeredo-Espin, A.M.L., Paulo, D.F., Marinho, M.A.T., Tomsho, L.P.,

Drautz-Moses, D.I., Purbojati, R.W., Ratan, A. and Schuster, S.C., 2016. Large-

scale mitogenomics enables insights into Schizophora (Diptera) radiation and

population diversity. Scientific reports, 6, p.21762.

https://doi.org/10.1038/srep21762

Kergoat, G.J., Bouchard, P., Clamens, A.L., Abbate, J.L., Jourdan, H., Jabbour-Zahab,

R., Genson, G., Soldati, L. and Condamine, F.L., 2014b. Cretaceous

environmental changes led to high extinction rates in a hyperdiverse beetle

family. BMC evolutionary biology, 14(1), p.220. https://doi.org/10.1186/s12862-

014-0220-1

Kergoat, G.J., Soldati, L., Clamens, A.L., Jourdan, H., Jabbour-Zahab, R., Genson, G., Bouchard, P. and Condamine, F.L., 2014a. Higher level molecular phylogeny of darkling beetles (Coleoptera: Tenebrionidae). Systematic Entomology, 39(3), pp.486-499. https://doi.org/10.1111/syen.12065

Konishi, K., Matsumura, K., Sakuno, W. and Miyatake, T., 2020. Death feigning as an adaptive anti-predator behavior: further evidence for its evolution from artificial selection and natural populations. Journal of Evolutionary Biology. https://doi.org/10.1111/jeb.13641

Landeen, E.L., Muirhead, C.A., Wright, L., Meiklejohn, C.D., Presgraves, D.C., 2016. Sex Chromosome-wide Transcriptional Suppression and Compensatory Cis-Regulatory Evolution Mediate Gene Expression in the Drosophila Male Germline. PLoS Biol 14, e1002499. https://doi.org/10.1371/journal.pbio.1002499

Langdon, W.B., 2015. Performance of genetic programming optimised Bowtie2 on genome comparison and analytic testing (GCAT) benchmarks. BioData Mining 8, 1. https://doi.org/10.1186/s13040-014-0034-0

LeConte, J. L. 1862. Classification of the Coleoptera of North America. United States: Smithsonian Institution. p.233

Leder, E.H., Cano, J.M., Leinonen, T., O'Hara, R.B., Nikinmaa, M., Primmer, C.R., Merilä, J., 2010. Female-Biased Expression on the X Chromosome as a Key Step in Sex Chromosome Evolution in Threespine Sticklebacks. Molecular Biology and Evolution 27, 1495–1503. https://doi.org/10.1093/molbev/msq031

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., 1000 Genome Project Data Processing Subgroup,

2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Lorenzen, M.D., Doyungan, Z., Savard, J., Snow, K., Crumly, L.R., Shippy, T.D., Stuart, J.J., Brown, S.J. and Beeman, R.W., 2005. Genetic linkage maps of the red flour beetle, Tribolium castaneum, based on bacterial artificial chromosomes and expressed sequence tags. Genetics, 170(2), pp.741-747. https://doi.org/10.1534/genetics.104.032227

Lorenzen, M.D., Kimzey, T., Shippy, T.D., Brown, S.J., Denell, R.E. and Beeman, R.W., 2007. piggyBac-based insertional mutagenesis in Tribolium castaneum using donor/helper hybrids. Insect molecular biology, 16(3), pp.265-275. https://doi.org/10.1111/j.1365-2583.2007.00727.x

Love, M.I., Huber, W., Anders, S., 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15, 550. https://doi.org/10.1186/s13059-014-0550-8

Mahadevaraju, S., Fear, J.M., Akeju, M., Galletta, B.J., Pinheiro, M.M.L.S., Avelino, C.C., Cabral-de-Mello, D.C., Conlon, K., Dell'Orso, S., Demere, Z., Mansuria, K., Mendonça, C.A., Palacios-Gimenez, O.M., Ross, E., Savery, M., Yu, K., Smith, H.E., Sartorelli, V., Yang, H., Rusan, N.M., Vibranovski, M.D., Matunis, E., Oliver, B., 2021. Dynamic sex chromosome expression in Drosophila male germ cells. Nat Commun 12, 892. https://doi.org/10.1038/s41467-021-20897-y

Mahajan, S., Bachtrog, D., 2015. Partial Dosage Compensation in Strepsiptera, a Sister Group of Beetles. Genome Biology and Evolution 7, 591–600. https://doi.org/10.1093/gbe/evv008

Mank, J.E., Ellegren, H., 2009. SEX-LINKAGE OF SEXUALLY ANTAGONISTIC GENES IS PREDICTED BY FEMALE, BUT NOT MALE, EFFECTS IN BIRDS. Evolution 63, 1464–1472. https://doi.org/10.1111/j.1558-5646.2009.00618.x

Marin, R., Cortez, D., Lamanna, F., Pradeepa, M.M., Leushkin, E., Julien, P., Liechti, A., Halbert, J., Brüning, T., Mössinger, K., Trefzer, T., Conrad, C., Kerver, H.N., Wade, J., Tschopp, P., Kaessmann, H., 2017. Convergent origination of a Drosophila -like dosage compensation mechanism in a reptile lineage. Genome Res. 27, 1974–1987. https://doi.org/10.1101/gr.223727.117

Matthews, J.V., 1980. Tertiary land bridges and their climate: backdrop for development of the present Canadian insect fauna. The Canadian Entomologist, 112(11), pp.1089-1103. https:// 10.4039/Ent1121089-11

McKee, B.D., Handel, M.A., 1993. Sex chromosomes, recombination, and chromatin conformation. Chromosoma 102, 71–80. https://doi.org/10.1007/BF00356023

McKenna, D.D., Scully, E.D., Pauchet, Y., Hoover, K., Kirsch, R., Geib, S.M., Mitchell, R.F., Waterhouse, R.M., Ahn, S.-J., Arsala, D., Benoit, J.B., Blackmon, H., Bledsoe, T., Bowsher, J.H., Busch, A., Calla, B., Chao, H., Childers, A.K., Childers, C., Clarke, D.J., Cohen, L., Demuth, J.P., Dinh, H., Doddapaneni, H., Dolan, A., Duan, J.J., Dugan, S., Friedrich, M., Glastad, K.M., Goodisman, M.A.D., Haddad, S., Han, Y., Hughes, D.S.T., Ioannidis, P., Johnston, J.S., Jones, J.W., Kuhn, L.A., Lance, D.R., Lee, C.-Y., Lee, S.L., Lin, H., Lynch, J.A., Moczek, A.P., Murali, S.C., Muzny, D.M., Nelson, D.R., Palli, S.R., Panfilio, K.A., Pers, D., Poelchau, M.F., Quan, H., Qu, J., Ray, A.M., Rinehart, J.P., Robertson, H.M., Roehrdanz, R., Rosendale, A.J., Shin, S., Silva, C., Torson, A.S., Jentzsch,

I.M.V., Werren, J.H., Worley, K.C., Yocum, G., Zdobnov, E.M., Gibbs, R.A., Richards, S., 2016. Genome of the Asian longhorned beetle (Anoplophora glabripennis), a globally significant invasive species, reveals key functional and evolutionary innovations at the beetle–plant interface. Genome Biol 17, 227. https://doi.org/10.1186/s13059-016-1088-8

McKenna, D.D., Shin, S., Ahrens, D., Balke, M., Beza-Beza, C., Clarke, D.J., Donath, A., Escalona, H.E., Friedrich, F., Letsch, H. and Liu, S., 2019. The evolution and genomic basis of beetle diversity. Proceedings of the National Academy of Sciences, 116(49), pp.24729-24737. https://doi.org/10.1073/pnas.1909655116

McKenna, D.D., Wild, A.L., Kanda, K., Bellamy, C.L., Beutel, R.G., Caterino, M.S., Farnum, C.W., Hawks, D.C., Ivie, M.A., Jameson, M.L. and Leschen, R.A., 2015. The beetle tree of life reveals that Coleoptera survived end-Permian mass extinction to diversify during the Cretaceous terrestrial revolution. Systematic Entomology, 40(4), pp.835-880. https://doi.org/10.1111/syen.12132

Meiklejohn, C.D., Landeen, E.L., Cook, J.M., Kingan, S.B., Presgraves, D.C., 2011. Sex Chromosome-Specific Regulation in the Drosophila Male Germline But Little Evidence for Chromosomal Dosage Compensation or Meiotic Inactivation. PLoS Biol 9, e1001126. https://doi.org/10.1371/journal.pbio.1001126

Meiklejohn, C.D., Tao, Y., 2010. Genetic conflict and sex chromosome evolution. Trends in Ecology & Evolution 25, 215–223. https://doi.org/10.1016/j.tree.2009.10.005

Meštrović, N., Mravinac, B., Plohl, M., Ugarković, Đ. and Bruvo Mađarić, B., 2006. Preliminary phylogeny of Tribolium beetles (Coleoptera: Tenebrionidae) resolved by combined analysis of mitochondrial genes. http://fulir.irb.hr/id/eprint/1694

Misof, B., Liu, S., Meusemann, K., Peters, R.S., Donath, A., Mayer, C., Frandsen, P.B., Ware, J., Flouri, T., Beutel, R.G. and Niehuis, O., 2014. Phylogenomics resolves the timing and pattern of insect evolution. Science, 346(6210), pp.763-767. https://doi.org/10.1126/science.1257570

Morandin, C., Pulliainen, U., Bos, N., Schultner, E., 2018. De novo transcriptome assembly and its annotation for the black ant Formica fusca at the larval stage. Sci Data 5, 180282. https://doi.org/10.1038/sdata.2018.282

Moreno-Santillán, D.D., Machain-Williams, C., Hernández-Montes, G., Ortega, J., 2019. De Novo Transcriptome Assembly and Functional Annotation in Five Species of Bats. Sci Rep 9, 6222. https://doi.org/10.1038/s41598-019-42560-9

Mueller, J.L., Mahadevaiah, S.K., Park, P.J., Warburton, P.E., Page, D.C., Turner, J.M.A., 2008. The mouse X chromosome is enriched for multicopy testis genes showing postmeiotic expression. Nat Genet 40, 794–799. https://doi.org/10.1038/ng.126

Nakamura, T., Yamada, K.D., Tomii, K. and Katoh, K., 2018. Parallelization of MAFFT for large-scale multiple sequence alignments. Bioinformatics, 34(14), pp.2490-2492. https://doi.org/10.1093/bioinformatics/bty121

Namekawa, S.H., Lee, J.T., 2009. XY and ZW: Is Meiotic Sex Chromosome Inactivation the Rule in Evolution? PLoS Genet 5, e1000493. https://doi.org/10.1371/journal.pgen.1000493

Namekawa, S.H., Park, P.J., Zhang, L.-F., Shima, J.E., McCarrey, J.R., Griswold, M.D., Lee, J.T., 2006. Postmeiotic Sex Chromatin in the Male Germline of Mice. Current Biology 16, 660–667. https://doi.org/10.1016/j.cub.2006.01.066

Namekawa, S.H., VandeBerg, J.L., McCarrey, J.R., Lee, J.T., 2007. Sex chromosome silencing in the marsupial male germ line. Proceedings of the National Academy of Sciences 104, 9730–9735. https://doi.org/10.1073/pnas.0700323104

Nozawa, M., Onizuka, K., Fujimi, M., Ikeo, K., Gojobori, T., 2016. Accelerated pseudogenization on the neo-X chromosome in Drosophila miranda. Nat Commun 7, 13659. https://doi.org/10.1038/ncomms13659

Ogilvie, H.A., Bouckaert, R.R. and Drummond, A.J., 2017. StarBEAST2 brings faster species tree inference and accurate estimates of substitution rates. Molecular biology and evolution, 34(8), pp.2101-2114. https://doi.org/10.1093/molbev/msx126

Pal, A., Vicoso, B., 2015. The X Chromosome of Hemipteran Insects: Conservation, Dosage Compensation and Sex-Biased Expression. Genome Biol Evol 7, 3259–3268. https://doi.org/10.1093/gbe/evv215

Park, T., 1934. Observations on the general biology of the flour beetle, Tribolium confusum. The Quarterly Review of Biology, 9(1), pp.36-54.

Park, T., 1938. Studies in population physiology. VIII. The effect of larval population density on the postembryonic development of the flour beetle, Tribolium confusum Duval. Journal of Experimental Zoology, 79(1), pp.51-70.

Patro, R., Duggal, G., Love, M.I., Irizarry, R.A., Kingsford, C., 2017. Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods 14, 417–419. https://doi.org/10.1038/nmeth.4197

Peccoud, J., Simon, J.C., von Dohlen, C., Coeur d'acier, A., Plantegenest, M., Vanlerberghe-Masutti, F. and Jousselin, E., 2010. Evolutionary history of aphid-plant associations and their role in aphid diversification. Comptes Rendus Biologies, 333(6-7), pp.474-487. https://doi.org/10.1016/j.crvi.2010.03.004

Pessia, E., Makino, T., Bailly-Bechet, M., McLysaght, A., Marais, G.A.B., 2012. Mammalian X chromosome inactivation evolved as a dosage-compensation mechanism for dosage-sensitive genes on the X chromosome. Proceedings of the National Academy of Sciences 109, 5346–5351. https://doi.org/10.1073/pnas.1116763109

Prince, E.G., Kirkland, D., Demuth, J.P., 2010. Hyperexpression of the X Chromosome in Both Sexes Results in Extensive Female Bias of X-Linked Genes in the Flour Beetle. Genome Biology and Evolution 2, 336–346. https://doi.org/10.1093/gbe/evq024

Rambaut, A. and Drummond, A., 2008. FigTree: Tree figure drawing tool, version 1.2. 2. Institute of Evolutionary Biology, University of Edinburgh.

Rambaut, A., Drummond, A.J., Xie, D., Baele, G. and Suchard, M.A., 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. Systematic biology, 67(5), p.901. https://doi.org/ 10.1093/sysbio/syy032

Ramesh, B., Firneno, T.J., Demuth, J.P., 2021. Divergence time estimation of genus Tribolium by extensive sampling of highly conserved orthologs. Molecular

Phylogenetics and Evolution 159, 107084.
https://doi.org/10.1016/j.ympev.2021.107084

Richard, G., Legeai, F., Prunier-Leterme, N., Bretaudeau, A., Tagu, D., Jaquiéry, J., Le
Trionnaire, G., 2017. Dosage compensation and sex-specific epigenetic
landscape of the X chromosome in the pea aphid. Epigenetics & Chromatin 10,
30. https://doi.org/10.1186/s13072-017-0137-1

Robinson, M.D., McCarthy, D.J., Smyth, G.K., 2010. edgeR: a Bioconductor package
for differential expression analysis of digital gene expression data. Bioinformatics
26, 139–140. https://doi.org/10.1093/bioinformatics/btp616

Rose, G., Krzywinska, E., Kim, J., Revuelta, L., Ferretti, L., Krzywinski, J., 2016.
Dosage compensation in the African malaria mosquito Anopheles gambiae.
Genome Biol Evol evw004. https://doi.org/10.1093/gbe/evw004

Rui-ling, Z., Xiao-lei, H., Li-yun, J. and Ge-xia, Q., 2013. Molecular phylogenetic
evidence for paraphyly of Ceratovacuna and Pseudoregma (Hemiptera,
Hormaphidinae) reveals late Tertiary radiation. Bulletin of entomological
research, 103(6), pp.644-655. https:// 10.1017/S0007485313000321

Rupp, S.M., Webster, T.H., Olney, K.C., Hutchins, E.D., Kusumi, K., Wilson Sayres,
M.A., 2016. Evolution of dosage compensation in Anolis carolinensis , a reptile
with XX/XY chromosomal sex determination. Genome Biol Evol evw263.
https://doi.org/10.1093/gbe/evw263

Rylee, J.C., Siniard, D.J., Doucette, K., Zentner, G.E. and Zelhof, A.C., 2018.
Expanding the genetic toolkit of Tribolium castaneum. PloS one, 13(4),
p.e0195977. https://doi.org/10.1371/journal.pone.0195977

Satija, R., Farrell, J.A., Gennert, D., Schier, A.F., Regev, A., 2015. Spatial
reconstruction of single-cell gene expression data. Nat Biotechnol 33, 495–502.
https://doi.org/10.1038/nbt.3192

Savard, J., Tautz, D. and Lercher, M.J., 2006. Genome-wide acceleration of protein
evolution in flies (Diptera). BMC Evolutionary Biology, 6(1), p.7.
https://doi.org/10.1186/1471-2148-6-7

Schield, D.R., Card, D.C., Hales, N.R., Perry, B.W., Pasquesi, G.M., Blackmon, H.,
Adams, R.H., Corbin, A.B., Smith, C.F., Ramesh, B., Demuth, J.P., Betrán, E.,
Tollis, M., Meik, J.M., Mackessy, S.P., Castoe, T.A., 2019. The origins and
evolution of chromosomes, dosage compensation, and mechanisms underlying
venom regulation in snakes. Genome Res. 29, 590–601.
https://doi.org/10.1101/gr.240952.118

Schinko, J.B., Weber, M., Viktorinova, I., Kiupakis, A., Averof, M., Klingler, M., Wimmer,
E.A. and Bucher, G., 2010. Functionality of the GAL4/UAS system in Tribolium
requires the use of endogenous core promoters. BMC developmental biology,
10(1), pp.1-12. https://doi.org/10.1186/1471-213X-10-53

Schmitt-Engel, C., Schultheis, D., Schwirz, J., Ströhlein, N., Troelenberg, N., Majumdar,
U., Grossmann, D., Richter, T., Tech, M., Dönitz, J. and Gerischer, L., 2015. The
iBeetle large-scale RNAi screen reveals gene functions for insect development
and physiology. Nature communications, 6(1), pp.1-10.
https://doi.org/10.1038/ncomms8822

Schoenmakers, S., Wassenaar, E., Hoogerbrugge, J.W., Laven, J.S.E., Grootegoed, J.A., Baarends, W.M., 2009. Female Meiotic Sex Chromosome Inactivation in Chicken. PLoS Genet 5, e1000466. https://doi.org/10.1371/journal.pgen.1000466

Simão, F.A., Waterhouse, R.M., Ioannidis, P., Kriventseva, E.V. and Zdobnov, E.M., 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics, 31(19), pp.3210-3212. https://doi.org/10.1093/bioinformatics/btv351

Simão, F.A., Waterhouse, R.M., Ioannidis, P., Kriventseva, E.V., Zdobnov, E.M., 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics 31, 3210–3212. https://doi.org/10.1093/bioinformatics/btv351

Sinnock, P., 1970. Frequency dependence and mating behavior in Tribolium castaneum. The American Naturalist, 104(939), pp.469-476. https://doi.org/10.1086/282681

Smith, S.G. and Brower, J.H., 1974. Chromosome numbers of stored-product Coleoptera. Journal of the Kansas Entomological Society, pp.317-328. https://www.jstor.org/stable/25082656

Smith, S.G., 1950. The evolution of heterochromatin in the genus Tribolium (Tenebrionidae: Coleoptera). Chromosoma, 4(1), pp.585-610. https://doi.org/10.1007/BF00325793

Smith, S.G., 1950. The evolution of heterochromatin in the genus tribolium (Tenebrionidae: Coleoptera). Chromosoma 4, 585–610. https://doi.org/10.1007/BF00325793

Smith, S.G., 1952. The cytology of some tenebrionoid beetles (Coleoptera). Journal of Morphology, 91(2), pp.325-363. https://doi.org/10.1002/jmor.1050910206

Smith, S.G., 1952. The cytology of some tenebrionoid beetles (Coleoptera). J. Morphol. 91, 325–363. https://doi.org/10.1002/jmor.1050910206

Smith, S.G., Brower, J.H., 1974. Chromosome Numbers of Stored-Product Coleoptera 13.

Sokoloff, A., 1966. Genetics of Tribolium and related species.

Sokoloff, A., 1977. The biology of Tribolium with special emphasis on genetic aspects. Volume 3. Clarendon Press..

Solari, A.J., 1974. The behavior of the XY pair in mammals. International review of cytology 38, 273–317. https://doi.org/10.1016/S0074-7696(08)60928-6

Stevens, L., 1989. The genetics and evolution of cannibalism in flour beetles (genus Tribolium). Evolution, 43(1), pp.169-179. https://doi.org/10.1111/j.1558-5646.1989.tb04215.x

Thunders, M., Cavanagh, J., Li, Y., 2017. De novo transcriptome assembly, functional annotation and differential gene expression analysis of juvenile and adult E. fetida, a model oligochaete used in ecotoxicological studies. Biol Res 50, 7. https://doi.org/10.1186/s40659-017-0114-y

Tomoyasu, Y., Miller, S.C., Tomita, S., Schoppmeier, M., Grossmann, D. and Bucher, G., 2008. Exploring systemic RNA interference in insects: a genome-wide survey for RNAi genes in Tribolium. Genome biology, 9(1), p.R10. https://doi.org/10.1186/gb-2008-9-1-r10

Toussaint, E.F., Condamine, F.L., Kergoat, G.J., Capdevielle-Dulac, C., Barbut, J., Silvain, J.F. and Le Ru, B.P., 2012. Palaeoenvironmental shifts drove the adaptive radiation of a noctuid stemborer tribe (Lepidoptera, Noctuidae, Apameini) in the Miocene. PloS one, 7(7), p.e41377. https://doi.org/10.1371/journal.pone.0041377

Trauner, J., Schinko, J., Lorenzen, M.D., Shippy, T.D., Wimmer, E.A., Beeman, R.W., Klingler, M., Bucher, G. and Brown, S.J., 2009. Large-scale insertional mutagenesis of a coleopteran stored grain pest, the red flour beetle Tribolium castaneum, identifies embryonic lethal mutations and enhancer traps. BMC biology, 7(1), p.73. https://doi.org/10.1186/1741-7007-7-73

Turner, J.M.A., 2007. Meiotic sex chromosome inactivation. Development 134, 1823–1831. https://doi.org/10.1242/dev.000018

Turner, J.M.A., 2015. Meiotic Silencing in Mammals. Annu. Rev. Genet. 49, 395–412. https://doi.org/10.1146/annurev-genet-112414-055145

Turner, J.M.A., Mahadevaiah, S.K., Ellis, P.J.I., Mitchell, M.J., Burgoyne, P.S., 2006. Pachytene Asynapsis Drives Meiotic Sex Chromosome Inactivation and Leads to Substantial Postmeiotic Repression in Spermatids. Developmental Cell 10, 521–529. https://doi.org/10.1016/j.devcel.2006.02.009

Turner, J.M.A., Mahadevaiah, S.K., Fernandez-Capetillo, O., Nussenzweig, A., Xu, X., Deng, C.-X., Burgoyne, P.S., 2005. Silencing of unsynapsed meiotic chromosomes in the mouse. Nat Genet 37, 41–47. https://doi.org/10.1038/ng1484

Venanzetti, F., Cesaroni, D., Mariottini, P. and Sbordoni, V., 1993. Molecular

phylogenies in Dolichopoda cave crickets and mtDNA rate calibration. Molecular

Phylogenetics and Evolution, 2(4), pp.275-280.

https://doi.org/10.1006/mpev.1993.1026

Via, S. and Conner, J., 1995. Evolution in heterogeneous environments: genetic

variability within and across different grains in Tribolium castaneum. Heredity,

74(1), pp.80-90. https://doi.org/10.1038/hdy.1995.10

Vibranovski, M.D., Lopes, H.F., Karr, T.L., Long, M., 2009. Stage-Specific Expression

Profiling of Drosophila Spermatogenesis Suggests that Meiotic Sex Chromosome

Inactivation Drives Genomic Relocation of Testis-Expressed Genes. PLoS Genet

5, e1000731. https://doi.org/10.1371/journal.pgen.1000731

Vicoso, B., Bachtrog, D., 2011. Lack of Global Dosage Compensation in Schistosoma

mansoni, a Female-Heterogametic Parasite. Genome Biology and Evolution 3,

230–235. https://doi.org/10.1093/gbe/evr010

Vicoso, B., Bachtrog, D., 2015. Numerous Transitions of Sex Chromosomes in Diptera.

PLoS Biol 13, e1002078. https://doi.org/10.1371/journal.pbio.1002078

Vicoso, B., Emerson, J.J., Zektser, Y., Mahajan, S., Bachtrog, D., 2013. Comparative

Sex Chromosome Genomics in Snakes: Differentiation, Evolutionary Strata, and

Lack of Global Dosage Compensation. PLoS Biol 11, e1001643.

https://doi.org/10.1371/journal.pbio.1001643

Wade, M.J. and Johnson, N.A., 1994. Reproductive isolation between two species of

flour beetles, Tribolium castaneum and T. freemani: variation within and among

geographical populations of T. castaneum. Heredity, 72(2), pp.155-162.

https://doi.org/10.1038/hdy.1994.22

Wade, M.J., Johnson, N.A., Jones, R., Siguel, V. and McNaughton, M., 1997. Genetic

variation segregating in natural populations of Tribolium castaneum affecting

traits observed in hybrids with T. freemani. Genetics, 147(3), pp.1235-1247.

https://www.genetics.org/content/147/3/1235

Wahlberg, N., Leneveu, J., Kodandaramaiah, U., Peña, C., Nylin, S., Freitas, A.V. and

Brower, A.V., 2009. Nymphalid butterflies diversify following near demise at the

Cretaceous/Tertiary boundary. Proceedings of the Royal Society B: Biological

Sciences, 276(1677), pp.4295-4302. https://doi.org/10.1098/rspb.2009.1303

Walters, J.R., Hardcastle, T.J., 2011. Getting a Full Dose? Reconsidering Sex

Chromosome Dosage Compensation in the Silkworm, Bombyx mori. Genome

Biology and Evolution 3, 491–504. https://doi.org/10.1093/gbe/evr036

Walters, J.R., Hardcastle, T.J., Jiggins, C.D., 2015. Sex Chromosome Dosage

Compensation in Heliconius Butterflies: Global yet Still Incomplete? Genome Biol

Evol 7, 2545–2559. https://doi.org/10.1093/gbe/evv156

Wheeler, B.S., Anderson, E., Frøkjær-Jensen, C., Bian, Q., Jorgensen, E., Meyer, B.J.,

2016. Chromosome-wide mechanisms to decouple gene expression from gene

dose during sex-chromosome evolution. eLife 5, e17365.

https://doi.org/10.7554/eLife.17365

Whittle, C.A., Kulkarni, A., Extavour, C.G., 2020. Absence of a Faster-X Effect in

Beetles (Tribolium , Coleoptera). G3 10, 1125–1136.

https://doi.org/10.1534/g3.120.401074

Williford, A., Demuth, J.P., 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. https://doi.org/10.1093/molbev/mss184

Witt, E., Benjamin, S., Svetec, N., Zhao, L., 2019. Testis single-cell RNA-seq reveals

    the dynamics of de novo gene transcription and germline mutational bias in

    Drosophila. eLife 8, e47138. https://doi.org/10.7554/eLife.47138

Witt, E., Shao, Z., Hu, C., Krause, H.M., Zhao, L., 2021. Single-cell RNA-sequencing

    reveals pre-meiotic X-chromosome dosage compensation in Drosophila testis

    (preprint). Genetics. https://doi.org/10.1101/2021.02.05.429952

Wu, C.-I., Yujun Xu, E., 2003. Sexual antagonism and X inactivation – the SAXI

    hypothesis. Trends in Genetics 19, 243–247. https://doi.org/10.1016/S0168-

    9525(03)00058-1

Wu, M., Bao, R. and Friedrich, M., 2020. Evolutionary conservation of opsin gene

    expression patterns in the compound eyes of darkling beetles. Development

    Genes and Evolution, 230(5), pp.339-345.https://doi.org/10.1007/s00427-020-

    00669-2

Zhang, S.Q., Che, L.H., Li, Y., Liang, D., Pang, H., Ślipiński, A. and Zhang, P., 2018.

    Evolutionary history of Coleoptera revealed by extensive sampling of genes and

    species. Nature Communications, 9(1), pp.1-11. https://doi.org/10.1038/s41467-

    017-02644-4

Zhou, Q., Ellison, C.E., Kaiser, V.B., Alekseyenko, A.A., Gorchakov, A.A., Bachtrog, D.,

    2013. The Epigenome of Evolving Drosophila Neo-Sex Chromosomes: Dosage

Compensation and Heterochromatin Formation. PLoS Biol 11, e1001711.

https://doi.org/10.1371/journal.pbio.1001711

Zimmer, F., Harrison, P.W., Dessimoz, C., Mank, J.E., 2016. Compensation of Dosage-

Sensitive Genes on the Chicken Z Chromosome. Genome Biol Evol 8, 1233–

1242. https://doi.org/10.1093/gbe/evw075