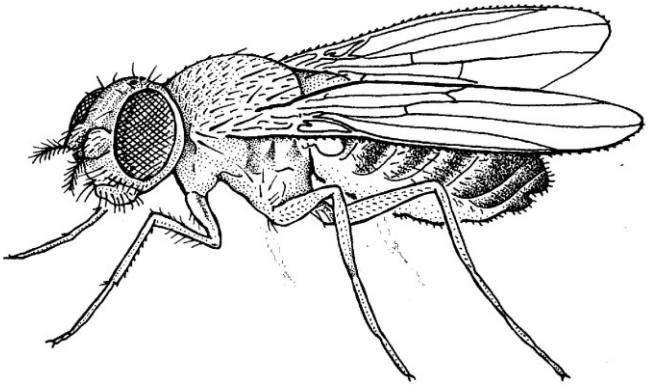
Drosophila Information Service



Number
104

December 2021

Prepared at the
Department of Biology
University of Oklahoma
Norman, OK 73019 U.S.A.

Preface

Drosophila Information Service (often called "DIS" by those in the field) was first printed in March, 1934. For those first issues, material contributed by *Drosophila* workers was arranged by C.B. Bridges and M. Demerec. As noted in its preface, which is reprinted in Dros. Inf. Serv. 75 (1994), Drosophila Information Service was undertaken because, "An appreciable share of credit for the fine accomplishments in *Drosophila* genetics is due to the broadmindedness of the original *Drosophila* workers who established the policy of a free exchange of material and information among all actively interested in *Drosophila* research. This policy has proved to be a great stimulus for the use of *Drosophila* material in genetic research and is directly responsible for many important contributions." Since that first issue, DIS has continued to promote open communication.

The production of this volume of DIS could not have been completed without the generous efforts of many people. Except for the special issues that contained mutant and stock information now provided in detail by FlyBase and similar material in the annual volumes, all issues are now freely-accessible from our web site: www.ou.edu/journals/dis. For the rare early issues that only exist as aging typed or mimeographed copies, some notes and announcements have not yet been fully brought on line. But we intend to fill in those gaps for historical purposes in the future.

We continue to encourage all researchers to consider submitting articles that use *Drosophila* for teaching as well as articles that report new techniques, research results, and interesting new mutations. In the interests of honoring the long-standing philosophy of open exchange of ideas, we sometimes accept articles that have unusual or limited perspectives. We thank the many contributors from around the world who sent material for this issue, and we invite your submissions for future annual issues as well as any suggestions you have for maintaining this as a useful *Drosophila* research community resource.

James N. Thompson, jr., Editor
David Ross Boyd Professor of Biology
Department of Biology
University of Oklahoma, Norman

Jenna J. Hellack, Associate Editor
Department of Biology, Emeritus
University of Central Oklahoma, Edmond
Department of Biology, Adjunct Professor
University of Oklahoma, Norman

Contributions and Inquiries should be sent to:

James N. Thompson, jr., Department of Biology, University of Oklahoma, Norman, OK 73019; Phone: (405) 325-2001 or 325-4821, FAX (405) 325-6202, email: jthompson@ou.edu

Printed copies of the recent volumes can be ordered from lulu.com.

List of Contributions

General Announcements

Research Notes

Teaching Notes

Other Reports

Research Notes

Mutations in Mcm10's C-terminal domain impact chromatin dynamics in *D. melanogaster*

Reubens, Michael C.[1,2], Ritu Dalia[1,3], Sidney E. Bedsole[1], Lucas T. Hopkins[1], and Tim W. Christensen[1*]. [1]Department of Biology, East Carolina University, Greenville, NC 27878, USA; Current positions: [2]Department of Cell and Molecular Biology, The Scripps Research Institute, La Jolla, CA 92037; [3]Asssistant Director of teaching laboratories, Drexel University, Philadelphia, PA 19104; [*]Correspondence: christensent@ecu.edu; Phone: 252-328-0162; Fax: 252-328-4178.

Abstract

Replication of the genome and the establishment of chromatin states are two processes essential for the maintenance of genomic stability and eukaryotic life. The linked nature of these processes can be observed by the many classically defined replication factors that have been shown to function in both replication and chromatin dynamics. Of these dual functioning proteins, Mcm10 has received increased attention due to correlations between miss-expression and tumorigenesis. This connection with cancer makes Mcm10 an interesting protein in which to study the connections between replication, chromatin formation, and oncogenesis. Historically Mcm10 research has centered on elucidating the function of this conserved protein in the initiation and elongation phases of replication. However, results from *S. cerevisiae* and *D. melanogaster* indicate that Mcm10 is involved in chromatin dynamics, despite disagreements on the regions of the protein required for the chromatic function. Here we report the results of three independent screens that indicate that the CTD of Mcm10 is required for the formation of endogenous chromatin states in the eye, brain, and salivary gland of *D. melanogaster*. These results further support Mcm10's role in the formation of chromatin and indicate that Mcm10's is involved in chromatin formation in cells undergoing variant cell cycles. These data suggest understanding this protein's chromatic function may aid in unraveling the essential nature of this enigmatic protein in DNA metabolism, as well as its ties to the mechanisms of oncogenesis. Keywords: Dominant suppression of variegation, heterochromatin formation, metaphase chromosomes, chromosome condensation, Mcm10, chromatin dynamics, endoreplication; List of commonly used abbreviations and acronyms: CTD: carboxy terminal domain; CNS: central nervous system; PEV: position effect variegation; TILLING: targeting induced local lesions in genomes; Mcm10: mini-chromosomal maintenance protein 10.

Introduction

Within the nucleus of eukaryotic cells the information required for cell viability, function, and cellular identity are contained in the DNA-protein complex known as chromatin. The replication and packaging of this genetic material through rounds of cellular division is essential in maintaining genome function and stability. Research has demonstrated that the processes of DNA replication and chromatin formation are intimately linked both spatially and temporally, and that errors during these key biological processes can manifest in genomic instability, loss of cellular identity, cellular transformation, and cellular death.[1,2] As such, a thorough understanding of these processes individually, and collectively, will be essential in understanding the processes of oncogenesis as well as the continued discovery of novel diagnostics and therapeutics.[3-5]

The linked nature of DNA replication and chromatin establishment can be observed in the effects that each process has on the other. Indeed, chromatin states must be deconstructed ahead of the processing replication fork, and subsequently re-established in the wake of the replication machinery. The shuttling of epigenetic information across the processing replisome allows for the potential establishment of new

chromatin states, aiding in processes such as cellular differentiation during development. However, it is of the utmost importance that these processes are also able to reliably preserve the epigenetic marks required to maintain differentiated cells.[2, 6] Likewise, chromatin states have been suggested to regulate the process of DNA replication by dictating locations of origin recognition, the timing of replication in specific genomic regions, and preventing the re-replication of the genome.[7] The interplay between these processes is not completely understood and remains an intense area of research.

Many well-established replication factors such as Orc2, PCNA, Cdc45, and Mcm2 have been demonstrated to function in chromatin dynamics, further illustrating the connected nature of these processes.[8-10] Another protein recently added to this list of dual functioning replication factors is mini-chromosomal maintenance protein 10 (Mcm10). Physical and genetic interactions with proteins involved in the initiation and elongation phases of DNA replication have traditionally defined Mcm10 as a DNA replication factor;[11-13] however, recent evidence of Mcm10's function in chromatin dynamics in *S. cerevisiae* and *D. melanogaster* have broadened our understanding of this protein's cellular utility. In *S. cerevisiae* a region of Mcm10's C-terminal domain (CTD) mediates the interactions between Mcm3 and Mcm7 with the essential silencing factor Sir2, which is required for gene silencing at both mating type loci and telomeres.[14, 15] Likewise, RNAi depletion in KC cells and PEV analysis using the hypomorphic $Mcm10^{Scim19}$ allele suggest that reduced Mcm10 protein levels result in aberrant chromatin formation in *D. melanogaster*.[11, 12]

Here we report the combined results of three independent screens conducted using 17 newly acquired C-terminal Mcm10 mutants in *D. melanogaster*. The results of the three independent chromatin assays demonstrate that the CTD of *Drosophila* Mcm10 is required for the establishment of chromatin states in the eye, central nervous system (CNS), and salivary glands of this highly utilized model system. These results further support the dual roles for Mcm10 in DNA replication and chromatin dynamics. Furthermore, the data presented here suggest that the expanded CTD conserved between flies and humans is highly sensitive to alterations in amino acid composition. The results of these experiments indicate the importance of researching Mcm10's roles in DNA replication as well as chromatin biology in order to understand the role that it plays in the process of oncogenesis, maintenance of the epigenome, and genomic stability.

Results

Identification of missense Mcm10 mutant collection

Studies in *S. cerevisiae* and *D. melanogaster* have suggested that Mcm10 is not only a DNA replication factor, but also functions in the formation of heterochromatin and chromatin condensation.[11, 12, 14, 15] Results in budding yeast tied the chromatin function of Mcm10 to a region of the CTD.[14, 15] However, results from *Drosophila* suggest that protein dosage impacts the function of Mcm10 in the formation of heterochromatin, but that this function is not dependent upon an intact CTD.[11, 12] The protein domains present in the expanded CTD of *Drosophila* Mcm10 are conserved across metazoan taxa, but are not present in budding yeast,[16-18] potentially explaining the discrepancy between these two model systems. To explore this hypothesis we conducted a TILLING (Targeting Induced Local Lesions in Genomes) screen on the C-terminal half of Mcm10.[19-22] Of the mutants generated through this screen 17 independent lines were able to propagate as homozygotes. Sequencing of the *Mcm10* (NM_136253) gene region in these lines verified that 16 possessed single point mutations that each altered one amino acid within the C-terminal half of Mcm10, and the remaining line encoded a petite 3 aa deletion within the CTD (Figure 1). The individual point mutants identified through this screen impact amino acid residues of varying levels of conservation, but none alter an amino acid residue that is conserved from yeast to humans. However, two of the three residues removed in the 3 aa deletion are conserved from yeast to humans (Figure 1).

16 Missense Mcm10 mutants dominantly suppress PEV

The ability of these lines to propagate as homozygotes suggested that the essential function of Mcm10 had been retained and provided an opportunity to investigate the impacts that the individual missense mutations had on the establishment of heterochromatin. Analysis of position effect variegation (PEV) using the white mottled four (w^{m4}) reporter line provides a high throughput screening method to evaluate the establishment of heterochromatin in the eye of adult male *Drosophila*. Using this system an inversion on the

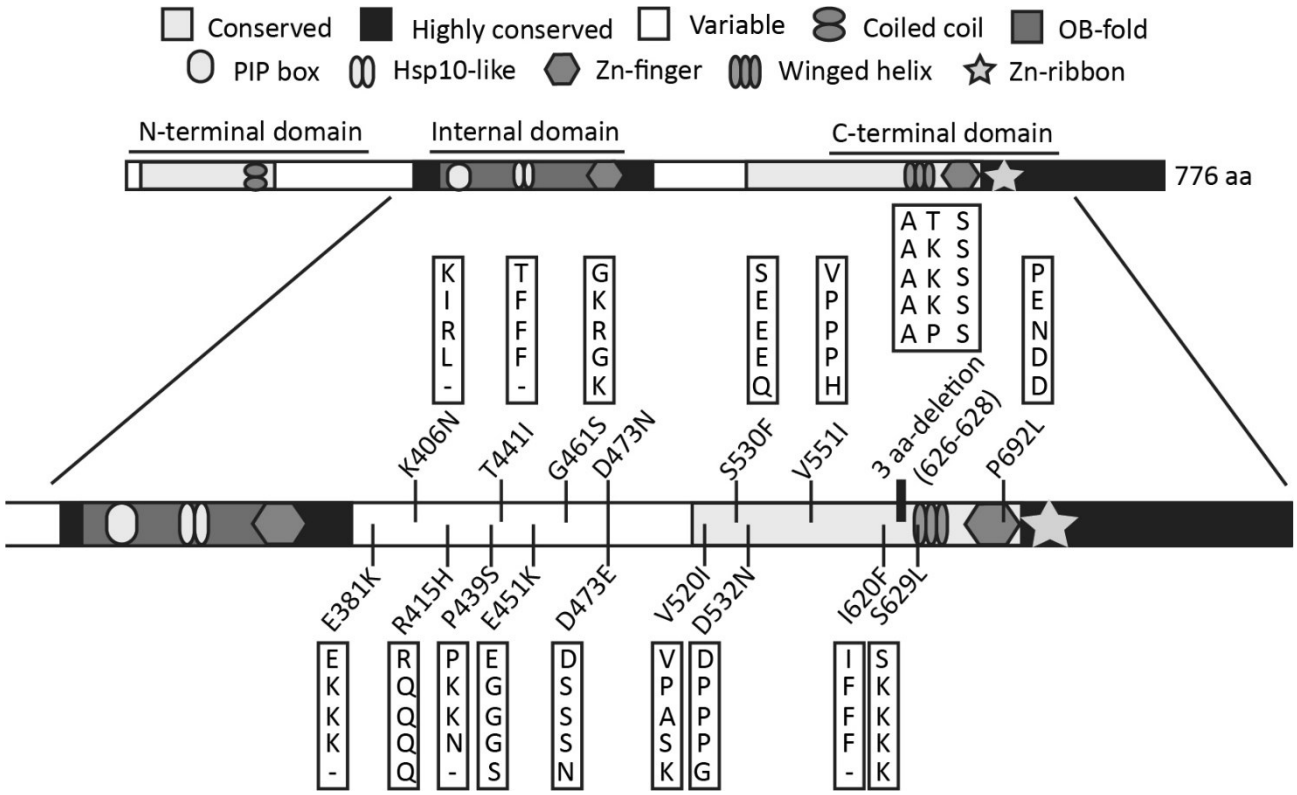


Figure 1. Novel Mcm10 C-terminal mutations. Schematic diagram of *D. melanogaster* Mcm10 protein domains, and the location of the novel mutations within the C-terminal domain. The boxes next to each label display the level of conservation for each amino acid based on a ClustalW alignment using *D. melanogaster* (top), *H. sapiens* (2nd from top), *M. musculus* (3rd from the top), *X. laevis* (4th from the top), and *S. cerevisiae* (bottom) protein sequences.

X chromosome places the white gene (w^+) near pericentric heterochromatin and creates a situation in which variegated w^+ expression results in a mosaic eye phenotype dependent upon the establishment pericentric heterochromatin within the mutant background.[23] Extraction of the drosopterin pigment allows for a quantitative measure of gene expression spectrophotometrically at 485 nm.[24]

Utilizing this technique we screened our mutant collection to assess the impact that each heterozygous mutant had on the formation of heterochromatin in the adult eye. If the CTD of *Drosophila* Mcm10 was indeed dispensable for its function in the formation of heterochromatin, we would expect to see no deviation from the w^{m4}*; Mcm10*$^+$*/CyO* control with respect to the amount of pigment in the eye. The results of our screen indicate that 16 out of 17 mutant lines were dominant suppressors of PEV (Figure 2), suggesting that the expanded CTD of *Drosophila* Mcm10 is involved in the establishment of heterochromatin. The only point mutation in our collection that failed to generate a statistically significant result was $Mcm10^{V520I}$ (Figure 2).

C-terminal mutations increase condensation defects in the larval brain

The PEV analysis results suggest that the majority of the missense mutations impact Mcm10's ability to establish heterochromatin in the canonical cell cycle of the eye. We sought to validate these results further using another canonical cycling cell type. In *Drosophila*, the central nervous system (CNS) of the third instar wandering larvae continues to be a highly utilized tissue for cytological studies investigating the mitotic phenotypes of mutations thought to impact cell-cycle progression and/or chromosome dynamics. The larval CNS is unique in that, unlike most other larval organs, it persists into the adult stage,[25] and it is comprised of two major cell types both undergoing canonical cell cycles.[26] Therefore, we utilized this tissue to screen our

mutant collection for defects in chromosome condensation following the established protocol for calculation of mitotic indices.[12]

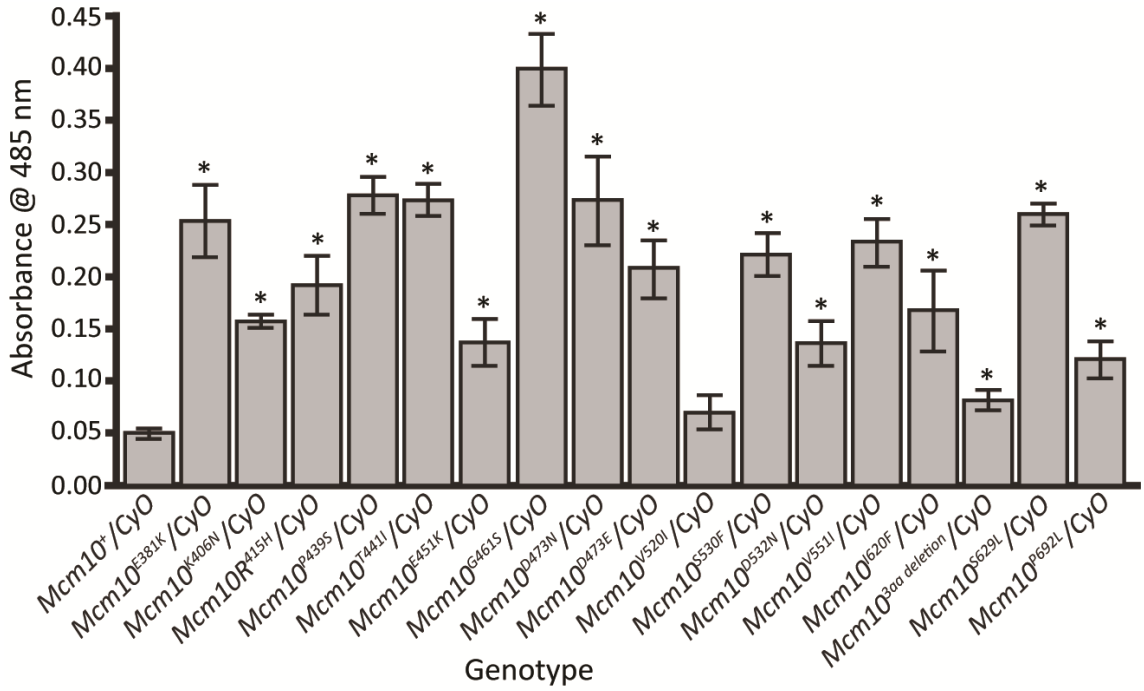


Figure 2. 16 missense mutants are dominant suppressors of PEV. Bar graph displaying the results of w^{m4} PEV screen. Each bar represents the average (± S.E.) of five biological replicates using five males per replicate. Asterisks denote statistical significance from *Mcm10*$^{+}$*/CyO* at a significance level of 0.05.

Similar to the results obtained from previous KC cell experiments[11] we observed no defects in cell cycle progression in the brain preparations (data not shown); however, mitotic chromosomes displaying signs of both over and under-condensed states of chromatin condensation were observed more often in all of the mutant lines than in the w^{1118} control (Figure 3A-C). Odds ratio analysis of the percent of nuclei displaying condensation defects out of the total mitotic chromosomes observed for each line *versus* the w^{1118} control indicated a significant increase in the occurrence of condensation defects in 12 of the 17 lines (Figure 3D). These results suggest a requirement for the CTD of Mcm10 in the formation of heterochromatin in this cell type as well. Interestingly *Mcm10*^{V520I} also failed to generate significant effects in the CNS screen (Figure 3D).

Mcm10's CTD functions in chromatin packaging in endoreplicating nuclei

Taken together the results of the previous two screens implied that the CTD of *Drosophila* Mcm10 is involved in the formation of heterochromatin and chromosome condensation in cells undergoing the canonical cell cycle. To address if the observed chromatin effects were cell cycle specific, we were interested in testing what impact the mutations would have on the packaging of DNA in a cell type that does not undergo the canonical cell cycle. The nuclei present in the salivary gland of the third instar wandering larvae utilize a variant cell cycle known as the endocycle. In endocycling cells euchromatic regions of the genome are amplified through multiple rounds of DNA replication without intervening cellular divisions, a process known as endoreplication. Through this process these nuclei are able to obtain as many as 2048 copies of euchromatic regions of the genome.[27] This copious amount of DNA is packaged in the nucleus by lining up sister chromatids in parallel orientation with interphase levels of condensation all sharing a single large chromocenter.[28-30]

To evaluate the impacts that our mutant collection had on the packaging of DNA in these polytene nuclei, we compared the mass of DNA per unit volume, or "packing ratio", of our w^{1118}; $Mcm10^{mut}/Mcm10^{mut}$ panel to that of the w^{1118}; $Mcm10^{+}/Mcm10^{+}$ control as previously described[31, 32]. The results of these comparisons indicated that 14 out of the 17 mutants in our collection significantly altered the packaging of DNA from that of the w^{1118} control (Figure 4). In this cell type $Mcm10^{P439S}$, $Mcm10^{E451K}$, and $Mcm10^{G461S}$ were the only mutant lines that did not generate statistically significant results. These results suggest that the CTD of Mcm10 is involved in the establishment of chromatin states in endoreplicating cells as well.

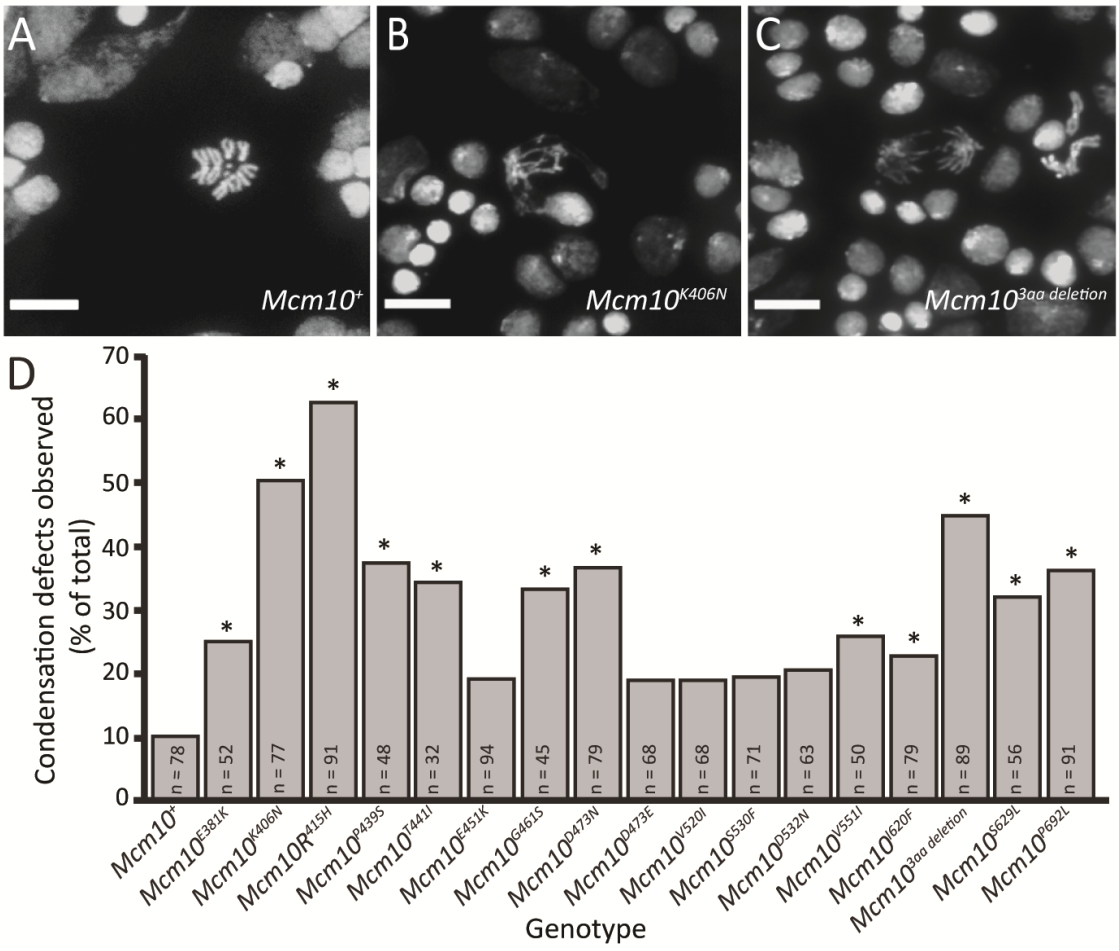


Figure 3. Mutations in Mcm10's CTD result in condensation defects in the CNS. (A-C).100× confocal micrographs of mitotic chromosomes observed in the CNS of the genotypes listed in the bottom right. Scale bars represent 10 μm. (A.) Representation of properly condensed mitotic chromosomes in which all four pairs of chromosomes are easily observed. (B-C) Representative images of condensation defects observed in the mutant collection from $Mcm10^{K406N}$ and $Mcm10^{3aadeletion}$, respectively. The image in B represents chromosomes from $Mcm10^{K406N}$ which are mis-condensed as observed by the heterogeneous compaction within the same chromosomes, which were not observed in the control. 3C represents one set of chromosomes which are aneuploid and undergoing anaphase (center), and another that displays heterogeneous compaction as indicated by the differential DAPI staining when compared to the control (right) (D). Bar graph showing the percentage of condensation defects observed in each mutant line. The numbers expressed in the bars represent the total number of mitotic chromosomes observed for each line. Asterisks denote statistical significance from the w^{1118} control at a significance level of 0.05 from the odds ratio analysis.

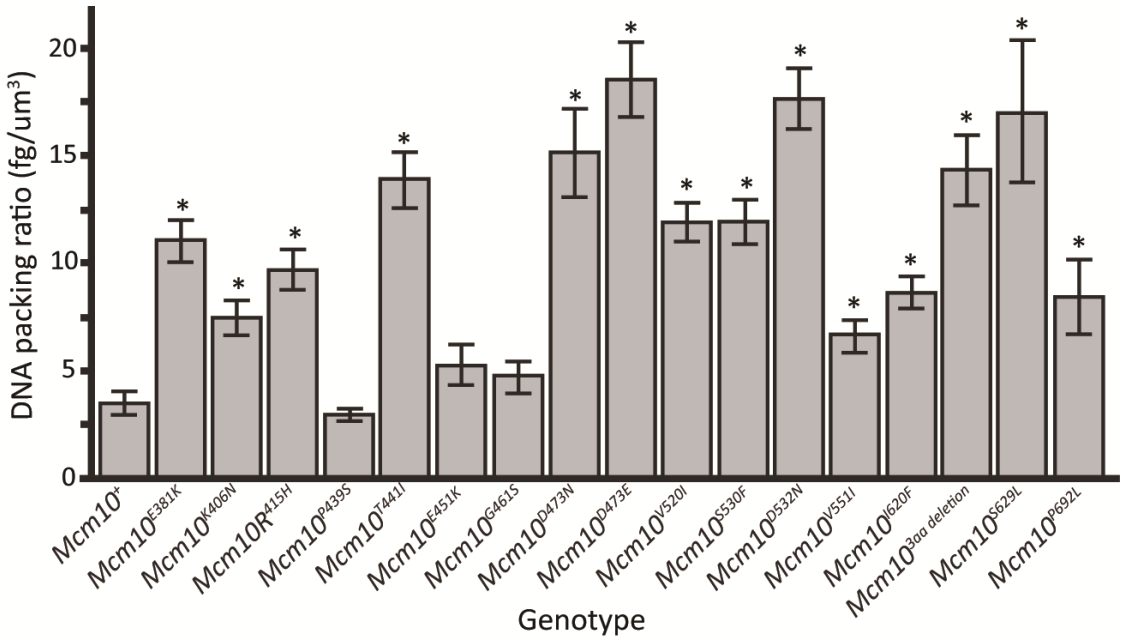


Figure 4. Mcm10 C-terminal mutations impact DNA packaging in endoreplicating nuclei. Bar graphs depicting the amount of DNA packaged per unit volume of nuclei for each mutant in the collection. The bars represent the average (± S.E.) from a distribution of 25 values for each line. Asterisks denote statistical significance from w^{1118} at a significance level of 0.05.

Discussion

Mcm10's involvement in the processes of DNA replication and chromatin dynamics make this conserved protein a very attractive candidate with which to study the connections of these biological processes. Indeed, both of these cellular processes have been suggested to be essential in maintaining genomic stability.[1] Furthermore, the recent studies showing a correlation between abnormal Mcm10 expression and many cancerous states make this protein even more attractive as a potential diagnostic tool and suggest that a better understanding of Mcm10's role in DNA metabolism may aid in our understanding of mechanisms underlying oncogenesis.[4, 5, 33] To date the majority of research on Mcm10 has been to elucidate the function of this conserved replication factor in the processes of initiation and elongation of DNA replication.[13] Only recently has research in *S. cerevisiae* and *Drosophila* brought to light the impact that this protein has on chromatin dynamics.[11, 12, 14, 15]

Previous results from *S. cerevisiae* tied the chromatic function of Mcm10 to a region of the CTD.[14, 15] However, results in *D. melanogaster* suggested that the function of Mcm10 in chromatin dynamics was dependent upon protein load, not an intact expanded CTD.[12] Results from the C-terminal mutations used in this study suggest that the expanded CTD present in *Drosophila*, and other metazoans, is involved in the formation of pericentric heterochromatin in cells undergoing the canonical cell cycle. Furthermore, mutants affecting the amino acid composition of the expanded CTD also resulted in increased occurrence of condensation defects in mitotic chromosomes of the CNS and had a significant impact on the packing ratio of endoreplicating salivary gland nuclei. The results of these three independent screens support the idea that Mcm10 plays a role in chromatin dynamics in *Drosophila* and suggest that the CTD is required for this protein to function optimally in these processes (Figure 5).

The results of the PEV and CNS screens presented here (Figure 2 and 3D) suggest that the CTD of Mcm10 is involved in the formation of endogenous heterochromatin states during the canonical cell cycle. It was interesting that $Mcm10^{V520I}$ failed to produce a significant deviation from the controls in both of these

screens (Figure 2 and 3D). It is possible that the substitution of one non-polar amino acid for another at this position reduced the impact on the overall protein chemistry and function. It is also possible that this residue is not involved in, or does not impact, the chromatic function of Mcm10, potentially explaining the lack of conservation observed at this residue (Figure 1). Further characterization of the mutant alleles used in this study, and their impact on Mcm10 gene function and known protein interactions, will be required to further infer the essential nature of individual residues with in the CTD of Mcm10. The discrepancy observed in the significance generated between the PEV screen and the CNS condensation screen with respect to $Mcm10^{E451K}$, $Mcm10^{D473E}$, $Mcm10^{S530F}$, and $Mcm10^{D532N}$ may represent legitimate differences in the protein interactions required for the formation of constitutive versus facultative heterochromatin in cells undergoing the canonical cell cycle. However, it is also possible that these differences could be artifacts generated by differing sensitivities in regards to the PEV and CNS screening methodologies, as the PEV assay is very sensitive to changes in gene expression.

The results from the packing ratio screen seem to contradict the results of the previous two screens reported from non-endoreplicating tissues. Based on the data from the salivary glands, 14 of the 17 mutant lines were able to package significantly more DNA per unit volume of the nuclei, suggesting that they are packaging DNA more tightly than the w^{1118} control (Figure 4). However, the way in which the DNA is replicated and packaged in endoreplicating cells is much different than in the canonical cycles of the eye and brain.[27-30] This could potentially explain the observed differential effects of the mutations from the PEV and CNS screens to those observed in the salivary glands. We also observed that all mutant lines used except $Mcm10^{E451K}$ contained more DNA per nucleus than the w^{1118} control (data not shown), so it is also possible that the increased packing ratios could be a result of over-replication. Whether or not this over replication was a result of unregulated replication or erroneous chromatin states on the substrate DNA due to the Mcm10 mutations would require further investigation. Regardless of the seemly conflicting results, the ability of the mutants to significantly alter the state in which the chromatin was packaged when compared to w^{1118} supports the idea that the CTD of Mcm10 is involved in the packaging of the DNA in this variant cell cycle as well.

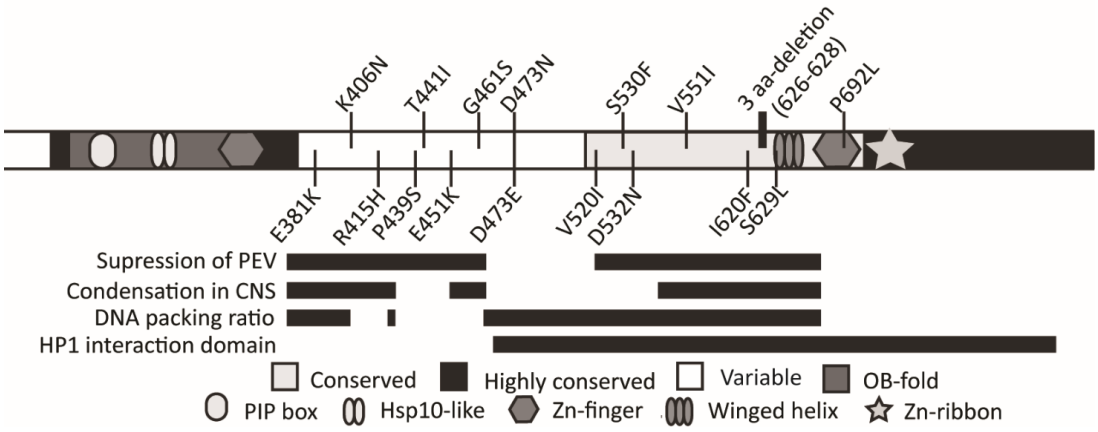


Figure 5. Mcm10's CTD is involved in chromatin dynamics in *Drosophila*. Figure summarizing the results obtained from three independent screens carried out with the novel C-terminal mutant collection, showing the regions of the CTD that produced significant results in each screen. The previously mapped HP1 interaction domain is also depicted on the figure.[12]

The combined results of the screens presented here support the data generated in budding yeast which ties the chromatic function of Mcm10 to a region of the CTD (Figure 5). Though the N-terminal and internal domains of Mcm10 are conserved from budding yeast to humans, the expanded C-terminal domain found in *Drosophila* and humans is not conserved in yeast.[16-18] Therefore, it is likely that the discrepancy between our previous results and the results presented here could be due to the region of the CTD required for its function

in chromatin dynamics residing upstream of the C-terminal-most 85 aa removed in the *Mcm10*^{d08029} truncation allele. In support of this claim the interaction between *Drosophila* Mcm10 and HP1 was mapped to the last 300 aa of the CTD (Figure 5), and the removal of the C-terminal 85 aa resulting from the *Mcm10*^{d08029} mutant allele did not alter this interaction as observed by yeast two-hybrid analysis.[12] Furthermore, the combined results of our current screens would suggest that the chromatin effects observed in KC cells depletion studies[11] and in the *Mcm10*^{Scim19} allele[12] may have been a result of reduced CTD availability. The results presented here demonstrate the utility of our mutant collection in studying the function of this conserved replication factor. Furthermore, these results suggest that further investigation of Mcm10's role in chromatin dynamics may aid in understanding this protein's roles in DNA metabolism and the mechanisms of oncogenesis.

Methods and Materials

Drosophila strains and culture

For all experiments presented, flies were cultivated at 25°C on *Drosophila* diet media K12 (U.S. Biologicals D9600-07B), and *Mcm10*$^{+}$ from the w^{1118} (FBst0003605) genetic background was used as our wild type control allele for all analyses reported (see individual materials and methods sections for complete information on the assay specific control genotype used). The PEV reporter line *In(1)w*^{m4h}*/ In(1)w*^{m4h}*; Su(var)205*^{5}*/In(2L)Cy, In(2R)Cy, Cy*^{1} (FBst0006234) was obtained from the Bloomington Stock center. The 17 missense mutants in our collection were obtained from a TILLING screen conducted on the EMS mutagenized Zuker collection through the Fred Hutchinson Cancer Research Center.[19, 20, 22]

Upon receiving the panel of TILLING mutants, the entire *Mcm10* gene region was sequenced in each strain to validate the presence of the reported point mutations or deletion, and the absence of any confounding mutations within the gene region. Once the indicated mutations were detected, w^{+}*; Mcm10*^{mut}*/Mcm10*^{mut} males from each mutant strain were then crossed to w^{1118}*; In(2LR)Gla/CyO, cl[4]* (a stock generated in our lab by crossing w^{1118} with *w; CyO, cl[4]/In(2LR)Gla* which was a gracious gift from Dr. Ross MacIntyre) virgin females. The resulting F_1 w^{1118}/w^{+}*; Mcm10*^{mut}*/CyO* virgin females and F_1 w^{1118}*; Mcm10*^{mut}*/CyO* males were then subsequently sibling crossed to establish a w^{1118}*; Mcm10*^{mut}*/Cyo* working stock for each mutant in the panel.

To remove potential second site mutations from the EMS treatment that may have resulted outside of the Mcm10 gene region, the established w^{1118}*; Mcm10*^{mut}*/CyO* mutant males were crossed to w^{1118}*; Df(2L)Exel6047, P{w*$^{+mC}$*=XP-U}Exel6047/CyO* (FBst0007529) virgin females. The resulting F_1 w^{1118}*; Mcm10*^{mut}*/Df(2L)Exel6047, P{w*$^{+mC}$*=XP-U}Exel6047* males and virgin females were then collected and allowed to sibling cross for eight generations. After eight generations w^{1118}*; Mcm10*^{mut}*/Mcm10*^{mut} true breeding stocks were established for all 17 lines, the presence of the mutations were once again verified by sequencing the *Mcm10* gene region, and the individual lines were cultivated as homozygotes until analyzed.

Position effect variegation screen

PEV analysis was conducted following a previously described protocol.[24] Briefly, w^{1118}*; Mcm10*^{mut}*/Mcm10*^{mut} male flies were crossed to *In(1)w*^{m4h}*/ In(1)w*^{m4h}*; Su(var)205*^{5}*/In(2L)Cy, In(2R)Cy, Cy*^{1} virgin females. The resulting F_1 *In(1)w*^{m4h}*; Mcm10*^{mut}*/ In(2L)Cy, In(2R)Cy, Cy*^{1} males were then collected the day of eclosion and aged for five days at 25°C. Once five aged males were collected, their heads were removed and homogenized in a 1:1 mixture of Chloroform (Sigma, Cat# C7559) and 0.1% ammonium hydroxide (Sigma, Cat#318612) using a MP Biomedicals FastPrep®24 homogenizer to insure equal homogenization between samples. The homogenate was then centrifuged for 2 minutes to separate phases. The absorbance of 100 µL of the aqueous phase was measured spectrophotometrically at 485 nm, using 0.1% ammonium hydroxide as a blank. Five biological replicates were completed for each mutant strain and the *In(1)w*^{m4h}*; Mcm10*$^{+}$*/ In(2L)Cy, In(2R)Cy, Cy*^{1} control to obtain the distributions used for analysis. The distribution from each strain was compared to the distribution of the control using two-tailed Student's t-test in JMP10 (SAS Institute Inc.) at a significance level of 0.05.

Screening larval CNS for condensation defects

Condensation defects were visualized as previously described for the generation of larval brain squashes and calculation of mitotic indices.[12] Briefly, third instar wandering larvae were harvested and their brains were dissected. The brains were swollen in 0.5% Sodium Citrate for 10 minutes, fixed, and transferred to a clean glass slide and overlaid with a siliconized coverslip. The slide and coverslip were then placed in machinist vise and 15 Nm of force was applied for 2 minutes. The brain preparations were then washed, stained with DAPI, sealed under a cover slip, and stored at 4°C until imaging.

Ten brains were prepared for each w^{1118}; $Mcm10^{mut}/Mcm10^{mut}$ experimental line as well as the w^{1118} control line, and 10 random well populated images were obtained from each using an Olympus IX81 Motorized Inverted Microscope with Spinning Disk Confocal at 60× magnification. Any mitotic chromosomes observed in the 10 random images, from the 10 prepared slides for each strain, were then further imaged at 100× to allow for scoring of the phenotype. All images were processed and analyzed using Slidebook™ software. The total number of mitotic chromosomes was counted from all 100 images for each stain and recorded, as well as the number of chromosomes observed displaying abnormally condensed chromosomes. Chromosomes were scored as improperly condensed if they were not packaged into stereotypically distinct mitotic chromosomes as observed in the w^{1118} controls, if they displayed heterogeneous states of compaction not observed in the w^{1118} control (*i.e.,* chromosome arms appeared thicker or thinner than the controls), or if they displayed an abnormal phenotype associated with perturbed chromosome packaging that was not observed in the w^{1118} control spreads (*i.e.,* diffuse unorganized chromosome structure, multiple centromeres, or altered patterns of DAPI staining intensity). The percentage of chromosomes displaying condensation defects across all 10 brains from each line was then calculated and compared to the percentage displaying condensation defects in the w^{1118} control strain using odds ratio analysis in JMP10 (SAS Institute Inc.) at a significance level of 0.05.

Packing ratio screen

DNA packing ratios were generated for each strain, and the w^{1118} control, as previously described.[31, 32] In summary, five pairs of salivary glands were collected from wandering third instar larvae for each strain and subsequently fixed, DAPI stained, and mounted on slides for confocal imaging. A 3D image of each gland pair was then obtained using a Zeiss LSM 700 laser scanning confocal microscope at 20× magnification using 2 µm virtual sections. Once the 3D images had been acquired for each gland pair they were converted to maximum intensity projection images using the Zen software package, and exported as TIF files for further analysis in Adobe Photoshop®. Using the counting and measurement features in Photoshop®, the average number of nuclei per gland pair and the average volume per nuclei was calculated for each strain and recorded.

To assess the mass of DNA present in each gland pair, 25 pairs of salivary glands were dissected from third instar wandering larvae from each strain. The individual gland pairs were then placed in 200 µL tubes, homogenized, and centrifuged. The samples were then heated to 37°C for 30 minutes, then heated to 85°C for 10 minutes, and then centrifuged. The concentration of DNA was determined for each sample using the Qubit® 2.0 Fluorometer following manufactures' established protocol for the Qubit® dsDNA HS Assay Kit (Invitrogen, Cat #Q32866). Distributions of packing ratios were calculated for each strain by using the average number of nuclei per gland pair, the average volume of nuclei per strain, and the distribution of DNA mass per gland pair for each strain. The distribution from each strain was compared to the distribution of the w^{1118} control using two-tailed Student's t-test in JMP10 (SAS Institute Inc.) at a significance level of 0.05.

Acknowledgments: Thanks to the Bloomington Stock Center and Dr. Ross MacIntyre for fly stocks, and the Fred Hutchinson Cancer Research Center for providing the TILLING services. Thanks to Dr. Tom Fink for assistance and maintenance of the microscopy core facility, and thanks to Dr. Elizabeth Ables for valuable guidance and assistance with the Zeiss LSM 700 laser scanning confocal microscope. This work was supported by National Institutes of Health R15 GM069875 (T.W.C.)

References:

1. Stillman B. Origin recognition and the chromosome cycle. FEBS Letters 2005; 579:877-84.
2. Alabert C, Groth A. Chromatin replication and epigenome maintenance. Nat Rev Mol Cell Biol 2012; 13:153-67.
3. Kastan MB, Bartek J. Cell-cycle checkpoints and cancer. Nature 2004; 432:316-23.

4. Das M, Prasad SB, Yadav SS, Govardhan HB, Pandey LK, Singh S, Pradhan S, Narayan G. Over expression of minichromosome maintenance genes is clinically correlated to cervical carcinogenesis. PLoS One 2013; 8:e69607.
5. Hua C, Zhao G, Li Y, Bie L. Minichromosome Maintenance (MCM) Family as potential diagnostic and prognostic tumor markers for human gliomas. BMC Cancer 2014; 14:526.
6. Raynaud C, Mallory AC, Latrasse D, Jégu T, Bruggeman Q, Delarue M, Bergounioux C, Benhamed M. Chromatin meets the cell cycle. Journal of Experimental Botany 2014.
7. Singh J. Role of DNA replication in establishment and propagation of epigenetic states of chromatin. Seminars in Cell & Developmental Biology 2014; 30:131-43.
8. Polo SE, Almouzni G. Histone metabolic pathways and chromatin assembly factors as proliferation markers. Cancer Letters 2005; 220:1-9.
9. Groth A, Rocha W, Verreault A, Almouzni G. Chromatin challenges during DNA replication and repair. Cell 2007; 128:721 - 33.
10. Moldovan G, Pfander B, Jentsch S. PCNA, the Maestro of the Repliation Fork. Cell 2007; 129:665-79.
11. Christensen TW, Tye BK. Drosophila MCM10 interacts with members of the prereplication complex and is required for proper chromosome condensation. Mol Biol Cell 2003; 14:2206 - 15.
12. Apger J, Reubens M, Henderson L, Gouge CA, Ilic N, Zhou HH, Christensen TW. Multiple Functions for Drosophila Mcm10 Suggested Through Analysis of Two Mcm10 Mutant Alleles. Genetics 2010; 185:1151-65.
13. Thu YM, Bielinsky A-K. Enigmatic roles of Mcm10 in DNA replication. Trends in Biochemical Sciences 2013; 38:184-94.
14. Liachko I, Tye BK. Mcm10 is required for the maintenance of transcriptional silencing in Saccharomyces cerevisiae. Genetics 2005; 171:503 - 15.
15. Liachko I, Tye BK. Mcm10 mediates the interaction between DNA replication and silencing machineries. Genetics 2009; 181:379 - 91.
16. Robertson PD, Warren EM, Zhang H, Friedman DB, Lary JW, Cole JL, Tutter AV, Walter JC, Fanning E, Eichman BF. Domain architecture and biochemical characterization of vertebrate Mcm10. J Biol Chem 2008; 283:3338 - 48.
17. Warren EM, Vaithiyalingam S, Haworth J, Greer B, Bielinsky AK, Chazin WJ, Eichman BF. Structural basis for DNA binding by replication initiator Mcm10. Structure 2008; 16:1892 - 901.
18. Warren EM, Huang H, Fanning E, Chazin WJ, Eichman BF. Physical interactions between MCM10, DNA, AND DNA polymerase {alpha}. J Biol Chem 2009; 284:24662-72.
19. Winkler S, Schwabedissen A, Backasch D, Bokel C, Seidel C, Bonisch S, Furthauer M, Kuhrs A, Cobreros L, Brand M, et al. Target-selected mutant screen by TILLING in Drosophila. Genome Res 2005; 15:718-23.
20. Henikoff S, Till BJ, Comai L. Tilling. Traditional Mutagenesis Meets Functional Genomics. Plant Physiology 2004; 135:630-6.
21. Cooper JL, Till BJ, Henikoff S. Fly-TILL: Reverse genetics using a living point mutation resource. Fly 2008; 2:300-2.
22. Koundakjian EJ, Cowan DM, Hardy RW, Becker AH, Koundakjian EJ, Cowan DM, Hardy RW, Becker AH. The Zuker collection: a resource for the analysis of autosomal gene function in Drosophila melanogaster. Genetics 2004; 167:203-6.
23. Schotta G, Ebert A, Dorn R, Reuter G. Position-effect variegation and the genetic dissection of chromatin regulation in Drosophila. Seminars in Cell & Developmental Biology 2003; 14:67-75.
24. Evans BA, Howells AJ. Control of drosopterin synthesis in Drosophila melanogaster: Mutants showing an altered pattern of GTP cyclohydrolase activity during development. Biochem Genet 1978; 16:13-26.
25. Truman JW. Metamorphosis of the central nervous system of Drosophila. Journal of Neurobiology 1990; 21:1072-84.
26. Hofbauer A, Campos-Ortega J. Proliferation pattern and early differentiation of the optic lobes in Drosophila melanogaster. Roux's Arch Dev Biol 1990; 198:264-74.
27. Edgar BA, Orr-Weaver TL. Endoreplication Cell Cycles: More for Less. Cell 2001; 105:297-306.

28. Hochstrasser M, Mathog D, Gruenbaum Y, Saumweber H, Sedat JW. Spatial organization of chromosomes in the salivary gland nuclei of Drosophila melanogaster. The Journal of Cell Biology 1986; 102:112-23.
29. Deng H, Zhang W, Bao X, Martin J, Girton J, Johansen J, Johansen K. The JIL-1 kinase regulates the structure of Drosophila polytene chromosomes. Chromosoma 2005; 114:173-82.
30. Lee HO, Davidson JM, Duronio RJ. Endoreplication: polyploidy with purpose. Genes & Development 2009; 23:2461-77.
31. Chmielewski JP, and Tim W. Christensen. Novel method for determining chromosome compaction and DNA content of salivary gland nuclei in Drosophila. . Drosophila Information Services 2011; 94.
32. Chmielewski J, Henderson L, Smith C, Christensen T. Drosophila Psf2 has a role in chromosome condensation. Chromosoma 2012; 121:585-96.
33. Thu YM, Bielinsky A-K. MCM10: One tool for all—Integrity, maintenance and damage control. Seminars in Cell & Developmental Biology 2014; 30:121-30.

Evaluation of anti-stress property of *Dracaena trifasciata* (Snake plant) on stress induced *Drosophila melanogaster.*

Sathish Kumar, B.Y. Postgraduate Department of Biotechnology. JSS College (Autonomous), Ooty road, Mysore – 570025, Karnataka, India. Email: bysathish@gmail.com.

Abstract

Stress is a normal human reaction that happens to everyone. In fact, the human body is designed to experience stress and react to it. When you experience changes or challenges (stressors), your body produces physical and mental responses. In medical parlance "stress" is defined as a perturbation of the body`s homeostasis. This demand on mind-body occurs when it tries to cope with incessant changes in life. A "stress" condition seems "relative" in nature. Anti-stress property of *Dracaena trifasciata* (Snake plant) was evaluated using the *Drosophila melanogaster* as a model organism. In the first group, control flies were taken as normal flies and were considered as control to compare with that of second group of stress induced flies by different concentrations of MTX, and third group of flies was reared on media containing plant sample of 0.1g along with MTX, and the last group of flies was reared on the media containing only 0.1g of plant sample. Then the flies were subjected for enzymatic assay using enzymes Catalase and SOD at a time. The result of the present study showed that the plant sample used may have the anti-stress property as it reduced the stress, which was demonstrated by the reduced activities of marker enzymes like Catalase and SOD in stress induced *Drosophila melanogaster*. Key words: Stress, *Drosophila*, Catalase, Superoxide dismutase

Introduction

Oxidative stress is a "privilege" of aerobic organisms. It can be induced by endogenous and exogenous factors (Dallman *et al.,* 2005). At the same time, the human body had never been designed to live in this complex, modern world with its many demands. Hence, life in the 21st Century is infinitely far more complex than it has ever been. One physiological response to stress is the increased activity of certain enzymes (Sorensen *et al.,* 2003). Over-production of the reactive oxygen species (ROS) superoxide (O_2-) and hydrogen peroxide (H_2O_2) are increasingly implicated in human disease and aging. ROS are also being explored as important modulating agents in a number of cell signaling pathways. Increasing attention has been devoted to developing catalase or peroxidase mimetic as a way to treat overt inflammation associated with the pathophysiology of many human disorders. It focused on recent development of catalytic scavengers of peroxides and their potential use as therapeutic agents for pulmonary, cardiovascular, neurodegenerative, and inflammatory disorders. Antioxidants, the free radical scavengers, however, are shown to be anticarcinogens.

They function as the inhibitors at both initiation and promotion/transformation stage of carcinogenesis and protect cells against oxidative damage (Sun *et al.,* 1993). The novel antioxidant enzyme was shown to reduce hydro peroxides and, more recently, peroxynitrite with the use of electrons provided by a physiological thiol like thioredoxin. Methotrexate decreases titers of reduced folates, interferes with DNA synthesis, and results in the arrest of rapidly proliferating cells; it results in stress in *Drosophila* flies (Barclay *et al.,* 1982).

Their defense against these free radicals is achieved by natural antioxidant molecules but also by antioxidant enzymes. Three important anti-oxidant enzymes are Cu/Zn- superoxide dismutase (Cu/Zn-SOD), catalase, and selenium-glutathione peroxidase. They are all necessary for the survival of the cell even in normal conditions. In addition, these three enzymes act in a cooperative or synergistic way to ensure a global cell protection. However, optimal protection is achieved only when an appropriate balance between the activities of these enzymes is maintained (Michiels *et al.,* 2005).

Reactive oxygen species (ROS) are defined as oxygen-containing species that are more reactive than O_2 itself, which include hydrogen peroxide and super oxide. Although these are quite stable, they may be converted in the presence of transition metal ions, such as Fe (II), to the highly reactive oxygen species (hROS). hROS may exist as free hydroxyl radicals (HO), as bound ("crypto") radicals, or as Fe(IV)-oxo (ferryl) species and the somewhat less reactive, non-radical species, singlet oxygen (Tiwari *et al.,* 2001*)*.

Catalytic activity is present in nearly all animal cells and organs and in aerobic microorganisms. Catalase activity varies greatly between tissues with highest activities in the liver, kidney, and erythrocyte, and lowest activity present in connective tissues. In eukaryotic cells the enzyme is concentrated in sub-cellular peroxisome organelles.

Catalase is a tetramer of four polypeptide chains, each over 500 amino acids long (Boon *et al.,* 2007). The optimum temperature also varies by species.

Stress was induced to the flies by adding Methotrexate to check the anti-stress property of the plant. Methotrexate is a structural analog of folic acid and acts by binding and inhibiting dihydrofolate reductase (DHFR), a key enzyme required for intracellular folate metabolism. It is an antimetabolite and antifolate drug, and it works by inhibiting the metabolism of folic acid. Methotrexate acts specifically during DNA and RNA synthesis and is cytotoxic during the S-phase of the cell cycle. Methotrexate is commonly used in combination with misoprostol to terminate early pregnancies, *i.e*., pregnancy in the early stages. It may also be used in case of missed miscarriage, in which fetal demise has occurred, but the body has not expelled the fetus (Mol *et al.,* 2008).

Reduced folate is involved in normal synthesis and metabolism of neurotransmitters in central nervous system. Methotrexate may exert a beneficial effect in psoriasis by mechanism other than inhibition of dihydrofolate reductase. Methotrexate inhibits neutrophil chemotaxis (Johnston *et al.,* 2005).

Addition of folic acid to methotrexate therapy should allow dermatologists to use methotrexate in a much better way and enhance patient compliance *(*Mol *et al.,* 2008*)*. The improved efficacy of high-dose methotrexate as compared to conventional dose methotrexate suggests that osteosarcoma may have intrinsic methotrexate resistance, which can be Sarcoma overcome by achieving a high extracellular drug concentration (Johnston *et al.,* 2005). Intracellular methotrexate undergoes polyglutamylation whereby the polyglutamylated methotrexate is preferentially retained in the cell and ultimately results in DHFR inhibition (Bertino*,* 1993). As a cytotoxic drug it may slow the rapid growth of cells in the synovial membrane that lines the joints (Sirotnak, 1985). Methotrexate is a chemotherapy drug used to treat leukemia, lymphomas, and osteosarcoma. It is also used in the treatment of AIDS and rheumatoid arthritis.

The genus *Dracaena* consists of more than 100 accepted species which are mainly distributed in the tropics and subtropics, especially in Africa, Australia, and Southern Asia. They are mainly succulent shrubs and trees, and a few are commonly grown as shrubby houseplants, especially the variegated forms. The complete chloroplast (CP) genomes of six species have recently been reported, showing that they can be used as a super-barcode for *Dracaena* spp. identification (Zhang, *et al*., 2020).

About six *Dracaena* plants, growing in China, Southeast Asia, West Africa, Arabian Peninsula, Yemen, India, and Macaronesia are the main sources of this resin (Zhang, *et al.*, 2020; Sun, *et al.*, 2019). The drug was a commercially important export, notably from the island of Socotra (Yemen) and it has widely been used in traditional medicines through the world for thousands of years as an efficacious remedy for the treatment of hemorrhage, dysentery, diarrhea, stomach and external ulcers, wounds, leucorrhea, fractures,

piles, diabetes, and even tumors (Zhang, *et al.*, 2020; Sun, *et al.*, 2019; Gupta *et al.*, 2008; Cai, *et al.*, 1979; Fan *et al.*, 2014). In fact, the resin is known to have remarkable anti-inflammatory and antioxidant effects and to enhance immune function, promote skin repair, stop bleeding, and enhance blood circulation (Sun, *et al.*, 2019; Gupta *et al.*, 2008; Cai, *et al.*, 1979; Machala, *et al.*, 2001; Silva, *et al.*, 2011). Dragon's blood is included in the Pharmacopoeia of the People's Republic of China, where it was first imported through the silk road during the Sui and Tang dynasties. Until the 1970s, the red resin of *Dracaena cochinchinensis* S. C. Chen, used originally by the Dai people, living in the south part of Chinese Yunnan province, for treating pain and stopping hemorrhages, was found and used as the substitute of the traditionally imported dragon's blood, called Long-Xue-Jie (Chinese dragon's blood) (Gupta *et al.*, 2008; Cai, *et al.*, 1979). In the above context, *Dracaena trifasciata* was taken to analyze the antistress properties present in it.

Materials and Methods

The Drosophila Stock Centre, Department of Studies in Zoology, University of Mysore, provided the stocks of wild type of *D. melanogaster*. Further, the stocks were cultured in our laboratory at 26°C. As the temperature decreases, the development time increases (Ashburner *et al.*, 2005). At higher temperature around 31 degrees, flies may become sterile and may result in death. They require a controlled temperature and humidity environment. Stocks kept at room temperature were transferred to fresh media at every 20 days or the flies (5 male and 5 female flies) were transferred to fresh media when overcrowding occurs.

Culturing of stress induced flies

Methotrexate is an antimetabolite that interferes with the way cells utilize essential nutrients, so this chemical agent was added to create stress. Into the bottle along with media, methotrexate was added in different concentrations in the range of 5 ppm, 10 ppm, 15 ppm, 20 ppm, 25 ppm. The catalase and SOD activities were increased in the flies cultured in the media containing MTX. This was confirmed by comparing the activity of the enzymes with the control flies cultured in normal media.

Enzyme collection

Different groups of flies were taken in different eppendorff tubes as methotrexate flies of different concentration from 5 ppm-25 ppm and also these stress induced flies along with plant sample. These were fully homogenized in 200 microlitres of fresh phosphate buffer of 50 mM for catalase assay of pH 7.0 and for SOD assay 250 mM phosphate buffer of pH 7.8. These were homogenized with the help of tissue homogenizer, which was kept in ice cold condition and centrifuged at 8000 rpm for 20 min in a cooling microfuge. After centrifugation, the supernatant was transferred to a fresh eppendorff tube and 100 microlitres of this supernatant served as the enzyme source for both Catalase and SOD enzymatic assays.

Assay of Catalase enzyme

Catalase enzyme (EC 1.11.1.6) activity method is essentially described by Beers and Sizer (1952). 2.9 ml of hydrogen peroxide was taken along with 0.1 ml of enzyme extract, then immediately mixed by inverting, and the absorbance was read at 240 nm in spectrophotometer. The absorbance decreases gradually. Then activity of Catalase was calculated to know the specific activity protein estimation, which was expressed in units/mg of protein.

Assay of SOD

SOD enzyme (EC 1.15.1.1) was assayed using a slightly modified procedure originally described by Beauchamp and Fridovich (1971). Mix a 3 ml of cocktail solution containing Phosphate buffer (0.8 ml), Methionine (1 ml), riboflavin (0.5 ml), EDTA(0.1 ml), NBT(0.5 ml) and volume made up to3 ml by adding distilled water. A blank was set without the enzyme and NBT to calibrate the spectrophotometer having buffer (1.0 ml), Methionine (1 ml), riboflavin (0.7 ml), EDTA (0.3 ml). Another control was prepared having NBT but no enzyme and is taken as a reference control which contains buffer (0.9 ml), Methionine (1 ml), riboflavin (0.5 ml), EDTA(0.1 ml), NBT(0.5 ml). These colored solution absorbances were read at 560 nm immediately

to know the activity, and later on to know the specific activity protein estimation was done by Lowry's method and units expressed in units/mg of protein.

Results

Rearing of flies on media containing Methotrexate resulted in the increased activity of SOD and Catalase. These enzymes are the marker enzymes for the oxidative stress. The activity of SOD and Catalase increases, with respect to the increased concentration of Methotrexate in the media when compared to control flies on normal media (no Methotrexate, *i.e.*, 0 ppm) (Table 1).

Table 1. Increased Catalase and SOD activity in flies exposed to different concentration of Methotrexate.

Concentration of MTX	0 ppm (Control)	5 ppm	10 ppm	15 ppm	20 ppm	25 ppm
Catalase Activity in units/mg of protein	17.15	27.92	32.34	43.79	65.13	83.08
SOD Activity in units/mg of protein	0.83	0.79	0.84	1.04	1.08	1.15

Enzyme activity in stress induced flies treated with plant sample

The activity of SOD is decreased in flies reared on the media containing different concentrations of Methotrexate in presence of plant sample. The elevated level of enzyme due to Methotrexate was decreased in the presence of plant sample (Table 2).

Table 2. Catalase and SOD activity of alone stress induced flies and MTX+0.1 gm of plant sample of different concentrations.

	Concentration of MTX	5 ppm	10 ppm	15 ppm	20 ppm	25 ppm
Catalase Activity in units/mg of protein	MTX alone	27.92	32.34	43.79	65.13	83.08
	MTX + 0.1gm plant sample	18.24	24.13	29.07	45.25	48.13
SOD Activity in units/mg of protein	MTX alone	0.79	0.84	1.04	1.08	1.15
	MTX + 0.1gm plant sample	0.71	0.69	0.88	0.96	1.02

The enzyme activity was different in the flies reared on the media containing only 0.5 gm of plant sample. There was increased Catalase activity compared with control flies, and the SOD activity was found to be decreased when compared to the control flies (Table 3).

Table 3. Enzyme activity variation in normal *D. melanogaster* flies and flies treated with plant sample alone.

	Catalase Activity in units/mg of protein	SOD Activity in units/mg of protein
Control flies	17.15	0.83
Flies treated with plant sample alone	17.95	0.79

Discussion

Catalase and SOD activity was measured in stress-induced flies along with the plant sample of 0.1 gm, but there was a decrease in the activity compared to the control flies. Oxidative stress has been implicated to play a role, at least in part, in pathogenesis of many disease conditions and toxicities in animals, and overproduction of reactive oxygen species and free radicals due to use of toxic chemicals showed elevated increased catalase and SOD activity, so decrease in the catalase and SOD activity showing that plant sample is effective in decreasing both catalase and SOD activity. SOD dismutases the highly reactive superoxide anion to the less reactive species H_2O_2 (Teixeira *et al.*, 1998). Catalase, a haeme-containing enzyme, scavenges hydrogen peroxide to water and molecular oxygen (Mates and Sanchez-Jimenez, 1999), and non-enzymic ascorbic acid, which is a water-soluble antioxidant forage free radical, protects the biological system from oxidative stress.

When flies were treated with *Dracaena trifasciata* plant sample alone, the activity of SOD was also decreased and the activity of catalase was decreased. Hence, further study has to be achieved to isolate active constituents from the plant that can be used for applied research. Taken together, our data suggest that the plant sample we used may have anti-stress property in it.

References: Ashburner, M., K.G. Golic, and R.S. Hawley 2005, Cold Spring Harbor Laboratory Press. pp. 162–4; Arun Kumar, N., and B.Y. Satish Kumar 2010. Dros. Inf. Serv. 93: 30-35; Barclay, B.J., B.A. Kunz, J.G. Little, and R.H. Haynes 1982, J. Biochem. Cell Biol. 60: 172–194; Beauchamp, C., and Fridovich 1971, Anal. Biochem. 44: 276-287; Beers, R.F., Jr., and I.W. Sizer 1952, J. Biol. Chem. 195 (1): 133-140; Boon, E.M., A. Downs, and D. Marcey 2007, *Catalase Structural Tutorial Text*.; Chae, H.Z., S.W. Kang, and S.G. Rhee 1999, J. Methods in Enzymology 61: 219-226; Dallman, M.F., N.C. Pecoraro, and S.E. La Fleur 2005, Brain Behav. Immun. 19: 275-280; Sirotnak, F.M., 1985, Cancer Research 45 (9): 3992-4000 Sarcoma 5; Johnston, A.G., J.E. Udjonsson, H. Sigmundsdottir, B.R. Ludviksson, and H. Valdimarsson 2005, Clin. Immunol. 114: 154-163; Bertino, J.R., 1993, Journal of Clinical Oncology 11(1): 5-14; Mates, J.M., and F. Sanchez-Jimenez 1999, J. Front. Biosci. 4: 339-345; Michiels, C., M. Raes, O. Toussaint, and J. Remacle 1994, Free Radic. Biol. Med. 17(3): 235-48; Mol, F., B.W. Mol, W.M. Ankum, F. Van der Veen, and P.J. Hajenius 2008, Hum. Reprod. Update 14(4): 309-19; Sorensen, J.G., T.N. Kristensen, and V. Loeschcke 2003, J. Ecology Letters 6: 1025-1037; Sowmya, M., and B.Y. Sathish Kumar 2010, Journal of Stress Physiology and Biochemistry 6(4): 18-27; Sun, Y., 1990, Free Radic. Biol. Med. 8(6): 583-99; Teixeira, H.D., R.I. Schumacher, and R. Meneghini 1998, J. Proc. Natl. Acad. Sci. USA. 95: 7872–7875; Tiwari, A.K., 2001, J. Curr. Sci. 81: 1179-1187; Valko, M., C.J. Rhodes, J. Moncol, M. Izakovic, and M. Mazur 2006, Chem. Biol. Interact.160(1): 1-40; Willner, P., 1986, J. Prog. Neuropsyopharmacol. Biol. Psychiatry 10: 677-690; Zhang, Z., Y. Zhang, M. Song, Y. Guan, and X. Ma 2020, Front. Pharmacol. 2020, 11: 51; Sun, J., J.-N. Liu, B. Fan, X.-N. Chen, D.-R. Pang, J. Zheng, Q. Zhang, Y.-F. Zhao, W. Xiao, P.-F. Tu, *et al.* 2019, J. Ethnopharmacol. 244: 439–444; Gupta, D., B. Bleakley, and R.K. Gupta 2008, J. Ethnopharmacol. 115: 361–380; Cai, X.T., and Z.F. Xu 1979, Yunnan Zhi Wu Yan Jiu 1: 1-10; Fan, J.-Y., T. Yi, C.-M. Sze-To, L. Zhu, W.-L. Peng, Y.-Z. Zhang, Z.-Z. Zhao, and H.-B. Chen 2014, Molecules 10650–10669; Machala, M., R. Kubínová, P. Hořavová, and V. Suchý 2001, Phytother. Res. 15: 114-118; Silva, B.M., R.P. Santos, L.S. Mendes, P.G. de Pinho, P. Valentão, P.B. Andrade, J.A. Pereira, and M. Carvalho 2011, Food. Res. Int. 44: 2182–2189.

A report on Drosophilid diversity from Tumakuru (Karnataka, India).

Shilpa, M.S., C.D. Sharanya, P. Swathi, and A.L. Kokilamani[1]. Department of Studies and Research in Zoology, Tumkur University, Tumakuru-572103, Karnataka, India. [1]Email: alkokilamani@yahoo.co.in

The fauna of *Drosophila* of the Drosophilidae family is been subjected to biological research as a model organism in understanding the principles of basic genetics, molecular biology, population genetics, developmental biology, and evolution from more than a century. The endemism and cosmopolitan nature of the distribution of *Drosophila* fauna have influenced the eco-distributional patterns of its various species (Carson, 1965). Therefore, it needs to have scientific study on biodiversity of *Drosophila,* which analyzes the species compositions and the patterns of distribution of various members of the genus *Drosophila* in different geographical regions of the world.

The vast array of vegetation and climatic conditions of the Indian subcontinent has attracted many species of *Drosophila*. Numerous studies on diversity of *Drosophila* fauna and influence of various ecological factors on its distribution have been undertaken in different parts of India (Hegde *et al*., 2001; Upadhyay and Singh, 2006; Harini and Pranesh, 2011; Achumi *et al*., 2014; Alwyn D'souza *et al*., 2019). During the present study an effort was taken to record the diversity of *Drosophila* species in Tumkur University campus, Tumakuru, Karnataka, India (30°20′16″N and 77°7′13″E). The study area is 15 km away from the State Reserve Forest Devarayan Durga, which abodes thick vegetation and rare fauna of the country. Despite its proximity to State Reserve Forest and dense vegetation, no records of drosophilids are available nor has any species inventory been undertaken. Hence, a preliminary study to record the drosophilid diversity was undertaken from January to May, 2019. During the study period every alternate day the flies were collected by bottle trapping method, in which the bottles containing smash of ripened fruit of *Musa paradisiaca* (Banana) were tied to twigs of trees 3 feet above the ground and left open in an orchard at random positions with equidistance. The trapped flies were carefully extracted after securing the mouth of the traps with a cotton cloth and preserved in 70% alcohol. They were identified according to their characters as described by Bock and Wheeler (1972), Sajjan and Krishnamurthy (1975), and Markow and O'Grady (2006).

Table 1. List of *Drosophila* species collected from the study area.

Sl. No.	Subgenus	Species
1	*Dorsilopha*	*D. busckii*
2		*D. nasuta*
3	*Drosophila*	*D. repleta*
4	*Phorticella*	*P. striata*
5	*Scaptodrosophila*	*D. nigra*
6		*D. ananassae*
7		*D. auraria*
8		*D. bipectinata*
9		*D. eugracilis*
10	*Sophophora*	*D. Kikkawai*
11		*D. melanogaster*
12		*D. orena*
13		*D. rajashekari*
14		*D. simulans*
15		*D. suzukii*
16		*D. varians*
17		*D. yakuba*
18	*Scaptomyza*	*S. flava*

A total of 18 species of 6 subgenera of *Drosophila* were obtained from the study area (Table 1) by bait trapping method. The trapped *Drosophila* species were counted in the following percentage order *P. striata* (29%) > *D. bipectinata* (16%) > *D. melanogaster* (15%) > *D. nasuta* (11%) > *D. albomicans* (6%) > *D. ananassae* (5%) > *D. simulans* (4%) > *D. eugracilis* (4%) > *D. suzuki* (3%) > *D. rajashekari* (2%) > *D. varians* (1%) > *D. nigra* (1%) > *D. auraria* (1%). Among 18 species recorded from the study area *P. striata* and *D. bipectinata* were found to be more abundant with 101 and 63 individuals, respectively.

Phorticella striata recorded from the study area has also been noticed from different parts of south India (Sajjan and Krishnamurthy, 1975; Dilip *et al*., 2014; Alwyn D Souza *et al*., 2019). As the subgenus *Sophophora* dominates North-eastern India (Achumi *et al*., 2013) and Southern India (Guru Prasad and Pankaj Pathak, 2011) with maximum species of *Drosophila* fauna, the present study area was also noticed to be dominated by the subgenus *Sophophora* with 12 species of

Drosophila. A rich diversity index (Simpson_1-D = 0.15135) was obtained; however, moderate richness and the evenness of the community was obtained (Shannon _H = 2.20604).

As a biological indicator, *Drosophila* species are highly sensitive to slight environmental modifications that are reflected by significant variation in their natural population size and structure, morphology, and even in their ecology (Parsons, 1991). Assessing species diversity is among the most sensitive and extensively adopted measures of biodiversity which could be robustly related with different levels of ecosystem organization (Colwell and Coddington, 1994). The study of diversity of *Drosophila* species signifies the wide biodiversity and environmental fluctuations of any ecosystem of the locality. In this regard the diversity of *Drosophila* species was recorded from Tumakuru district for the first time. In the present study the ecological factors of the study area witnessed the rich diversity of Drosophilid species. Further research on diversity of *Drosophila* species is needed to assess the biodiversity of the entire region.

References: Achumi, B., L. Phom, Zevelou, Mohamad Ayajuddin, P. Lal, and Sarat Chandra Y. 2014, Nagaland University Journal (A multidisciplinary approach) 7: 320-3374; Achumi, B., S.N. Hegde, P. Lal, and S.C. Yenisetti 2013, Journal of Insect Science 13: 117. Available online: www. Insectscience.org; D'souza, A., M. Manjunath, and V. Shakunthala 2019, International Journal of Entomology Research 4(2): 01-08; Bock, I.R., and M.R. Wheeler 1972, Univ. Texas Publ. 7213: 1-102; Carson, H.L., 1965, Chromosomal polymorphism in geographically wide spread species of *Drosophila*. In: *Genetics of Colonizing Species* (Baker, H.G., and G.L. Stebbins, eds.), pp. 503–531. Academic Press; Colwell, R., and J. Coddington 1994, Philosophical Transactions of the Royal Society Series B: Biological Sciences 345: 101-118; Dilip, A.S., D.S.S. Pranesh, R. Alexander, K. Avinash, S. Phalke, and M. Jayashankar 2014, Dros. Inf. Serv. 97: 29-30; Guru Prasad, B.R., and Pankaj Pathak 2011, Dros. Inf. Serv. 94: 93-95; Harini, B.P., and D.S.S. Pranesh 2011, Dros. Inf. Serv. 94: 9-11; Hegde, S.N., V. Vasudev, and M.S. Krishna 2001, Wildlife Conservation Management 1: 55-71; Markow, T.A., and P.M. O'Grady 2006, Elsevier Inc. Academic press pub., 1-247; Parsons, P.A., 1991, Global Ecology and Biogeography 1: 77-83; Sajjan, S.N., and N.B. Krishnamurthy 1975, Oriental Insects. 9(1): 117-119; Upadhyay, K., and B.K. Singh 2006, Proceeding Zoological Society Calcutta 59(2): 195-202.

Month-wise collection of drosophilids at the Chandigarh University site (Gharuan, India).

Singh, Divya, and Seema Ramniwas*. University Center for Research and Development, Chandigarh University, Gharuan, Mohali-140413, India; *Corresponding author; E-mail: *seema.ramniwas@gmail.com

Drosophila sample were collected from the Chandigarh University site (Gharuan) in year 2020. The site is located at (300m above sea level) Mohali, India. We collected sample by bait trapping methods containing fermenting fruits. Flies as well as pupae were collected in year 2020 from Chandigarh University (Figure 1). Plants, vegetation, guava, mango, and decaying fruits were used by *Drosophila* as a breeding site.

Climatic data for the collection site was obtained Ramniwas (104-0 from the Indian Institute of Tropical Meteorology; IITM; www.tropmet.res.in) and the average Maximum and Minimum temperatures of each month (Figure 2) were recorded.

We collected 13 species of *Drosophila* and collected samples were classified by Nikon microscope according to species (Table 1). The result of this study could be explained by the fact that *Drosophila* species differ in different months of the year. In the colder months availability of cold adapted species were significant, whereas in warmer months warm adapted species are more abundant. Changing climatic conditions affected the distribution pattern of species (Parmesan, 1996; Kumar 2006). Thus, the extent to which ectotherms can tolerate changes in their ambient thermal environments is critical in determining their distribution and abundance. Distribution of *Zaprionus indianus* throughout the whole year (January 2020 to February 2021) given in Table 1 is worth pointing out, because its presence is in the whole year except the two

very coldest months (December and January). This species (*Zaprionus indianus*) has been recorded as a pest on oranges, peaches, and figs in Brazil (Santos *et al*., 2003) and a pest alert has been issued by the Division of Plant Industry, Florida Department of Agriculture & Consumer Services (Steck, 2005). Climate changes resulting in higher temperatures of winters would allow large numbers of generations and population size of *Z. indianus* to survive. This increase in progeny size may have a potential effect on fruits and vegetables and ultimately causing the harm will be very large to farmers.

Figure 1. Collection of *Drosophila* (flies and pupae) from different sites of Chandigarh University using bait trap method.

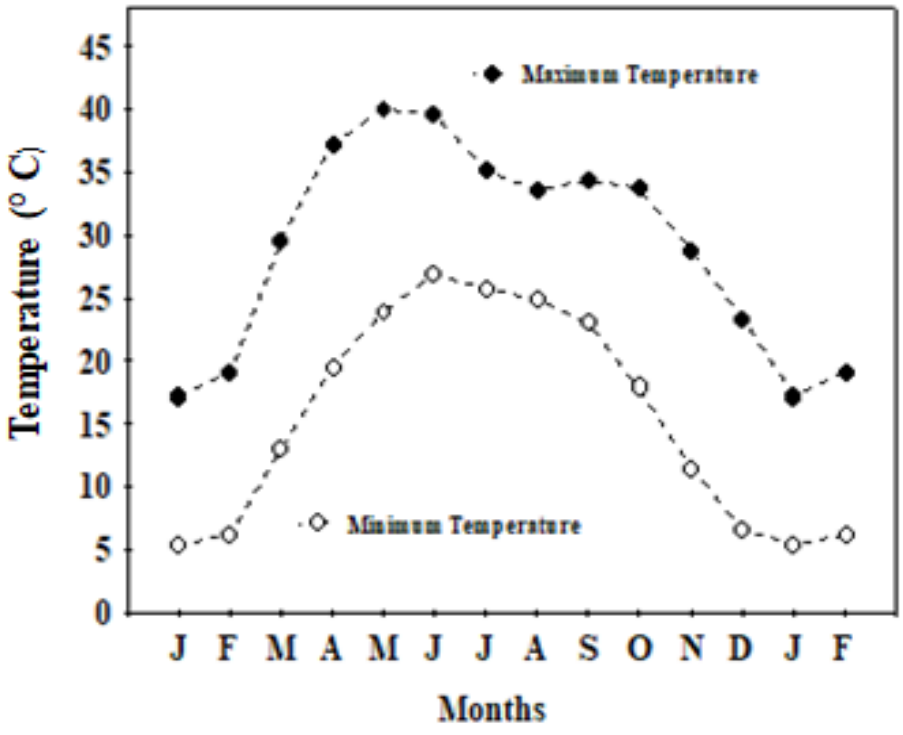


Figure 2. Maximum and minimum temperature of months (from January 2020 to February 2021) in which collection was done. J = January, F = February, M = March, A = April, M = May, J = June, J = July, A = August, S = September, O = October, N = November, D = December.

Table 1. Number of individuals of different species of Drosophila recorded in whole year January 2020 to December 2020.

Species	Jan	Feb	March	April	May	June	July	Sept	Aug	Oct.	Nov	Dec
Drosophila melanogaster	--	5	95	96	2	2	37	84	59	250	63	43
Drosophila ananassae	--	--	--	---	32	26	341	57	236	--	--	--
Drosophila kikkawai	7	19	36	--	--	--	5	464	369	7	4	8
Drosophila jambulina	--	--	--	--	--	74	34	46	79	14	5	9
Drosophila malerkotiana	--	--	--	--	--	--	34	365	25	--	--	--
Drosophila busckii	--	232	198	--	--	--	--	--	--	--	178	215
Drosophila punjabiensis	5	20	23	--	--	--	--	--	--	--	--	8
Drosophila biarmipes	--	--	--	8	6	--	--	6	9	--	--	--
Drosophila nepalensis	278	33	3	--	--	--	--	--	--	1	12	35
Drosophila takahashii	5	18	2	--	--	--	36	16	29	19	3	5
Drosophila immigrans	8	350	357	--	--	--	--	18	7	28	4	3
Drosophila bipectinata	--	--	--	--	16	24	13	3	16	--	--	--
Drosophila simulans	29	450	176	0	0	0	0	0	0	0	2	23
Zaprionus indianus	0	15	35	50	114	673	500	76	179	26	15	0

References: Kumar, S., 2006, Indian Academy of Sciences; Parmesan, C., 1996, Nature 382(6594): 765-766; Santos, J.F., T.T. Rieger, S.R.C. Campos, A.C.C. Nascimento, P.T. Félix, S.V.O. Silva, and F.M.R. Freitas 2003, Dros. Inf. Serv. 86: 92-95; Steck, G.J., 2005, Florida Department of Agriculture and Consumer Services, Division of Plant Industry. Pest Alert.

Co-existence and pre-pupal behavior of *Drosophila* species: wild *versus* lab.

Ramniwas, Seema, and Divya Singh*. University Center for Research and Development, Chandigarh University, Gharuan, Mohali-140413, India; *Corresponding author; E-mail: singh.genetics@gmail.com; Telephone: +918447576113.

Abstract

Conditions that influence pre-adult stages can be just as important as those affecting the adults. Behavioral responses of pre adult stages include burrowing into substrate, pupation site preference (PSP), and being active only through a restricted period of the day**.** The pupation site preferences in naturally coexisting species *Drosophila melanogaster, Drosophila immigrans,* and *Drosophila busckii* were studied. The study has revealed how pupation site preference differences contribute to species coexistence. Results can conclude that PSP plays a significant role in coexistence of *Drosophila* species. Keywords: *Drosophila melanogaster*, *Drosophila immigrans*, *Drosophila busckii*, pupation site preferences, coexistence.

Pupation height is a polygenic trait that responds effectively to bidirectional selection. Before pupation, larvae take up certain preferred positions as they can discriminate substrate moistures and avoid

burrowing and pupating in extremely dry or wet substrates that would negatively affect their emergence success (Wen *et al.*, 2016). Environment interactions with genetic makeup affect PSP in *Drosophila* species (Sokolowski, 1985; Joshi, 1997), but if coexisting *Drosophila* species show different pupation behaviors, how that informs on their coexistence has not been explored yet. Thus, it is important to study how pupation site preference differences contribute to species coexistence. Therefore, we tested the response of two *Drosophila* species (from a single locality) for PSP and life history traits of 2 coexisting species under field and laboratory conditions.

Pupae were collected from Chandigarh 30°45′N 76°47′E), India. The collections were made in February from nursery, and various *Drosophila* species (*D. immigrans, D. busckii, D. melanogaster*, *D. simulans, D. punjabiensis,* and *D. takahashii*) emerged from the collected papae sample. In total, 386 pupae of various *Drosophila* species were collected which coexist in the wild in February (18°C, RH 57%).

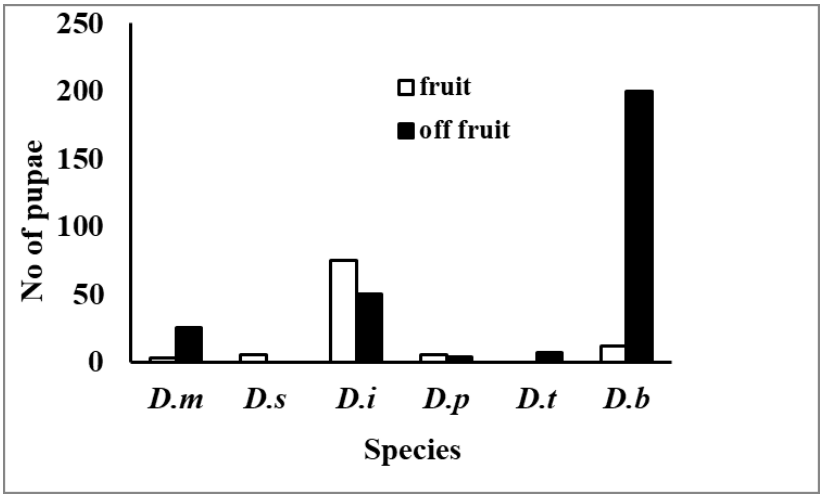


Figure 1. Total number of pupae collected from nursery in February from fruits and off fruits.

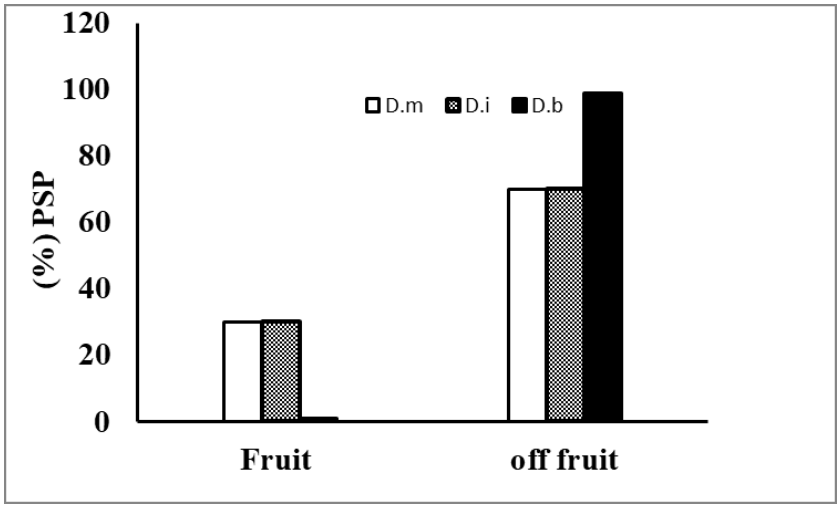


Figure 2. Preferences on fruit and off fruit of *D. melanogaste*r, *D. immigrans,* and *D. busckii* larvae in the absence of another *Drosophila* species.

From 386 pupae which were collected from fruit and off fruit in the month of February, 3 individuals of *D. melanogaster,* 12 of *D. busckii,* and 75 individuals of *D. immigrans* emerged from fruits and 25 individuals of *D. melanogaster*, 200 of *D. busckii,* and 50 of *D. immigrans* emerged from off fruits (Figure 1). The remaining 21 individuals were of *D. simulans* (5), *D. punjabiensis* (9), and *D. takahashii* (7).

We distinguished various pupae based on difference in size and in color. *D. melanogaster* pupae were brown-yellow in color and measured 3.49 ± 62.05 mm (N = 35). *D. immigrans* pupae were red-brown in color

and measured 4.61± 59.10 mm (N = 70), while *D. busckii* pupae were yellow-brown in color but small in size with measurement 2.10 ± 49.20 mm.

Number of pupae differs significantly on fruit and off fruit in different *Drosophila* species, except *D. busckii* and *D. melanogaster,* which show maximum pupae off fruit. All other collected species show maximum pupation on fruits. The number of flies eclosed showed that larvae choose different sites for pupation under different environmental conditions. Three studied species showed the same behavior for PSP in the wild as well as in the laboratory (Figure 2).

PSP is a variable, potentially adaptive trait in *Drosophila.* Despite the fact that more than 3000 species have been described under drosophilidae and about 1000 species belong to *Drosophila,* PSP has been analyzed in a few species. We found that distribution of pupae differs significantly among *D. melanogaster*, *D. immigrans,* and *D. busckii* when these species are grown together (or co-exist together) and independently (or exist independently) inside a vial (*i.e*., on food) in laboratory. *D. melanogaster* pupates 2 cm away from the food, *D. immigrans* pupates 4 cm away from the food and *D. busckii* pupates 0.5 cm away from the food (Figure 3). The variation displayed among sibling species sympatric species between closely related species (Vandal and Shivanna, 2007). Various drosophilids are distributed almost throughout the world and differences in pupation site preference offer potential adaptation in different *Drosophila* species.

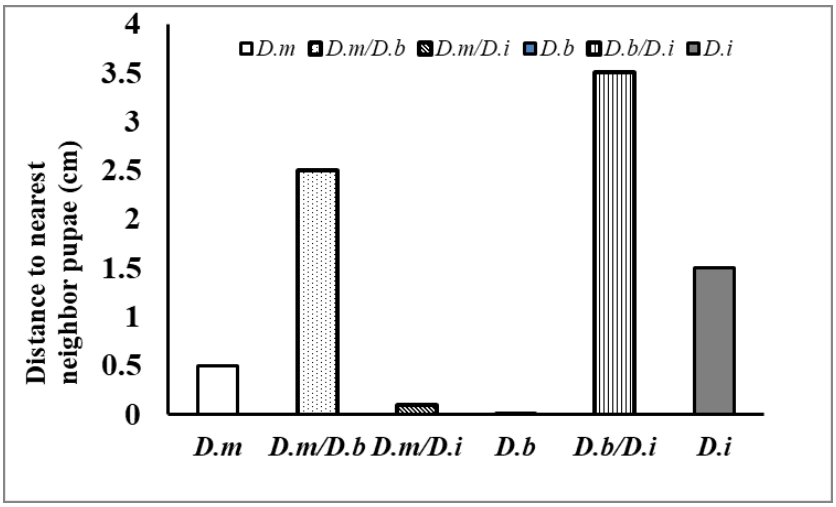


Figure 3. Distance measured to nearest pupae (between same species and between different species) in laboratory conditions inside a vial.

Temperature difference is sufficiently pronounced in various geographical regions and coexisting species adapt different mechanisms using the same resource at the same time and in the same place. Pupation in nature occurs either on host fruit and/or in the surrounding soil (Sokolowski, 1985; Rodriguez *et al*., 1992; Vandal *et al*., 2008), while in laboratory cultures, *Drosophila* larvae of some species prefer to pupate either on food medium or on glass walls of the culture bottle or in the cotton plug (Ramniwas and Kumar, 2019; Vandal *et al.,* 2003, 2008; Sisodia and Singh, 2005; Dillon *et al.,* 2009). These variations provide an adaptive significance to different drosophilids relative to prevailing environmental conditions. In this study, we found that despite their coexistence, *Drosophila* species show different pupation preferences.

References: Dillon, M.E., G. Wang, P.A. Garrity, and R.B. Huey 2009, Journal of Thermal Biology 34: 109-119; Joshi, A., 1997, Current Science 72: 555-562; Ramniwas, S., and G. Kumar 2019, Ethology Ecology & Evolution 31: 1-12; Rodriguez, L., M.B. Sokolowski, and J.S. Shore 1992, Journal of Evolutionary Biology 5: 61-70; Singh, B.N., and M.B. Pandey 1993, Behavior Genetics 23: 239-243; Sisodia, S., and B.N. Singh 2005, Journal of Genetics 84(2): 195-216; Sokolowski, M.B., 1985, Journal of Insect Physiology 31: 857-864; Vandal, N.B., S.A. Modagi, and N. Shivanna 2003, Indian Journal of Experimental Biology 41: 918-920; Vandal, N.B., G.S. Siddalingamurthy, and N. Shivanna 2008, Entomological Research 38: 188-194; Wen, Y., X. Jin, C. Zhu, X. Chen, T. Ma, S. Zhang, Y. Zhang, S. Zeng, X. Chen, Z. Sun, and X. Wen 2016, Journal of Insect Behavior 29 (4): 473-489.

Effects of inbreeding on the morphometric traits in males and females of *Drosophila melanogaster.*

Dubey, Pragya, and A.K. Singh. Genetics Laboratory, Department of Zoology, Institute of Science, Banaras Hindu University, Varanasi-221 005, India; Email: *aksbhu23@rediffmail.com*

Morphological features in almost all organisms play key roles in self-defence, mating success, territorial defense, and in achieving evolutionary fitness (Connolly and Cook, 1973; Markow and Sawka, 1992; Xiao, 2015). Vertebrates, like amphibians and reptiles, display a wide array of sexual dimorphism with variation in color, size, and structural adaptations (Andersson, 1994). Body size is the determining factor in deciding the parental role in fish, *Lamprologus callipterus* (Sato *et al.,* 2004). Structural organizations in the fruit fly, *Drosophila melanogaster,* are continuously shaped by the ever-changing environmental conditions during the course of evolution. In various species of *Drosophila*, males with larger body size have been observed to show higher mating success (Markow and Sawka, 1992). It has also been found that temperature present during the developmental time in *D. melanogaster* affected its behavioral ability of oviposition and access to food resource. Varying temperatures can also create range of body size in the fruit flies (Zamudio *et al.,* 1994). Effect of developmental temperature in *Drosophila* was also observed on morphometric features such as, wing area, body size, and flight capacity (Frazier *et al.,* 2008). In this short note, we present data regarding six morphometric features in both the sexes of *D. melanogaster* which were reared for five consecutive generations by transferring single pair of flies.

This experiment work was performed on the laboratory reared wild type stock, Oregon-R strain of *D. melanogaster.* The flies were maintained at 24 ±1°C temperature and 60-80% relative humidity with 12 h L/D cycle in the laboratory. Stock maintenance and all the experiments were conducted on standard food medium. The constituents of the standard food medium comprised yellow cornmeal, brown sugar, yeast powder (active yeast and yeast extract mixture), agar, nipagin, and propionic acid. The stocks were maintained by transferring fifteen pairs of 7-day-old virgin flies in culture bottles with standard food medium and being allowed to mate for 3 days and, thereafter, the flies were discarded to prevent any crowding of larvae and to avoid food limitation.

Our objective was to analyze the effect of inbreeding depression for five generations on six morphometric traits in the male and female of *D. melanogaster*. For the experiments, 30 cohorts were prepared where one male and female were introduced in each vial and allowed to mate for 3 days and, thereafter, they were discarded. On the emergence of F_1 progeny, virgin male and female flies were again collected and the next cross was set with the same protocol as described above with 1:1 sex ratio. This was continued for five generations and progeny obtained in the F_5 were utilized for morphometric measurements for both virgin males and females. The number of individuals used for the dimensions were 30 for each sex. Wing length (WL) was calculated as the absolute length between the anterior cross vein to the distal tip of the third longitudinal vein and was measured under the microscope at 50× magnification using ocular micrometer (1o.u. = 15µm). Wing width (WW) was scored as the distance between distal tips of second and fifth vein. Thorax length (TL) was measured as a distance from the anterior margin of the thorax to the posterior tip of the scutellum. Wing area (WA) was calculated as wing length × wing width. Wing/Thorax ratio (W/T) was calculated as the ratio of WL/ TL. Wing Aspect Ratio (WR) was calculated as the ratio of wing length^2/ wing area. W/T and WR both are regarded as indices of flight capacity. Thorax length had been established to be the most reliable estimate of body size in *Drosophila* (Robertson and Reeve, 1952).

To compare the difference for a particular morphometric trait between the two sexes, we applied unpaired student t-test assuming equal variance. The results of mean size of all the six traits in both the sexes of *D. melanogaster* are presented in Figure 1 as well as in Table 1. Based on the statistical analysis of the data (Table 3), it can be stated that there exists significant difference for these morphological features in the two sexes of this species. Coefficient of variation (C.V.) values for the concerned traits were also computed to see

which of the traits showed more consistency in the expression among the individuals of the same sex (Figure 2 and Table 2). There is comparatively higher variation in the thorax length and wing area in both the sexes of this species. Likewise, among females, the extent of variation was noted to be higher for wing by thorax ratio.

Table 1. Mean values of morphometric traits (N = 30) in males and females of *D. melanogaster*.

	WL	WW	*WA	TL	W/T	WAR
Male	1406.5 ± 5.96	1010 ± 8.17	1420.97 ± 14.41	900.50 ± 7.55	1.56 ± 0.01	1.39 ± 0.01
Female	1537 ± 9.37	1088 ± 7.29	1672.79 ± 17.10	986 ± 9.54	1.56 ± 0.02	1.41 ± 0.01

(Mean ± S.E.) in different sex ratios in Drosophila melanogaster. *unit = µm; *mm*

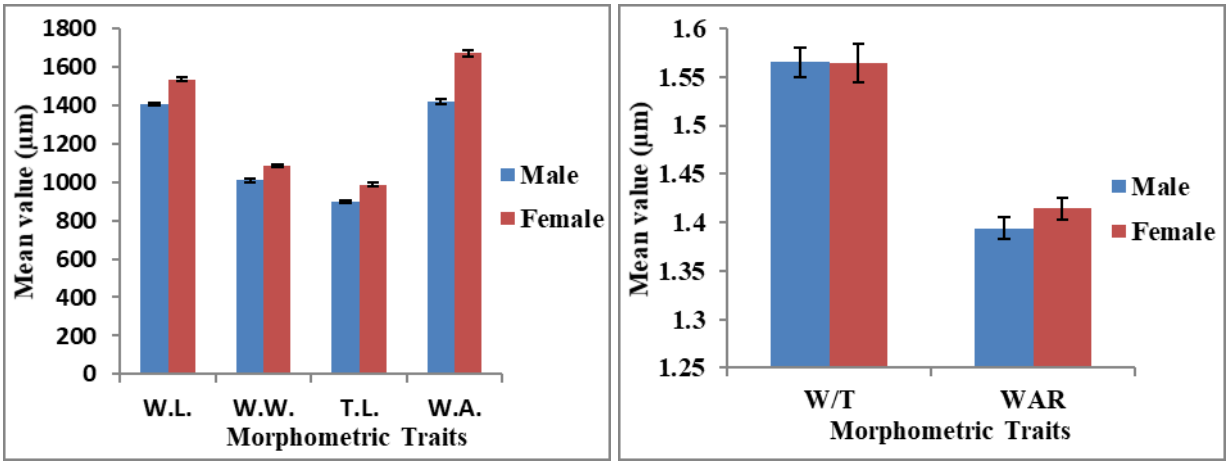


Figure 1. (a) Mean values of WL, WW, TL and WA; (b) Mean values of W/T and WAR in both males and females of *D. melanogaster*.

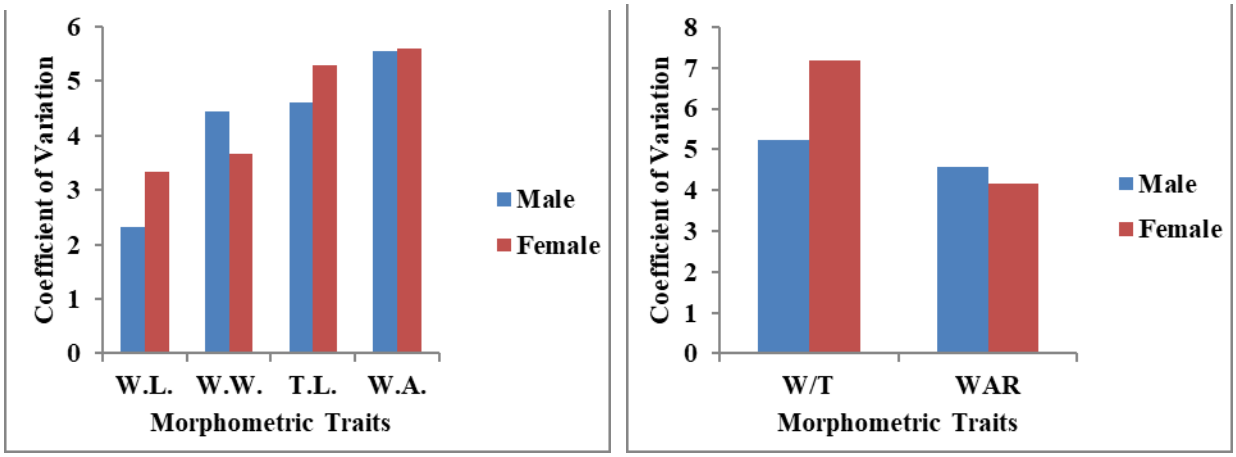


Table 2. Coefficient of variation (C.V.) observed for different morphological features in males and females of *D. melanogaster*.

	WL	WW	WA	TL	W/T	WAR
Male	2.32	4.44	5.56	4.60	5.23	4.57
Female	3.34	3.67	5.60	5.30	7.20	4.18

Figure 2. (a) Coefficient of variation in WL, WW, TL and WA; (b) Coefficient of variation in W/T and WAR in both males and females of *D. melanogaster*.

Table 3. Results of unpaired t-test for comparison of morphometric traits between males and females of *D. melanogaster*.

	WL	WW	WA	TL	W/T	WAR
t-value	-11.75***	-7.12***	-11.26***	-7.02**	0.05	-1.20*

d.f.58; *p<0.05, **p<0.01, ***p<0.001

Acknowledgment: One of the authors, Pragya Dubey, is thankful for BHU-UGC research fellowship as financial support to her to carry out this research.

References: Andersson, M., 1994, Princeton University Press, New Jersey; Connolly, K., and R. Cook 1973, Behaviour 44: 142-165; Frazier, M.R., J.F. Harrison, S.D. Kirkton, and S.P. Roberts 2008, J. Exp. Biol, 211: 2116-2122; Robertson, F.W., and E.C.R. Reeve 1952, J. Genet. 50: 414-448; Markow, T.A., and S. Sawka 1992, J. lnsect Behav. 5: 375-383; Sato, T., M. Hirose, M. Taborsky, and S. Kimura 2004, Ethology 110: 49–62; Xiao, Y.H., A.Z. Kosi, L.W. Zhang, T.R. Prentice, J.S. McElfresh, S.P. Chinta, Y.F. Zou, and J.G. Millar 2015, Insect Science 22: 840-852; Zamudio, K.R., R.B. Huey, and W.D. Crill 1995, Anim. Behav. 49: 671-677.

A preliminary study of *Drosophila* in a few localities of Western Ghats of Dharwad, Belagavi, and Uttara Kannada districts of Karnataka, India.

Srinath, B.S.*, S. Anurag, and N. Shivanna. Department of Zoology, Karnatak University Dharwad; *Email: srinathbs@kud.ac.in

Abstract

The present study deals with a preliminary and comparative analysis of *Drosophila* fauna from different localities of Western Ghats forests of three districts (Dharwad, Belagavi, and Uttara Kannada) in Karnataka state, India. It reveals 13 species of Drosophilids with Kalghatgi forest of Dharwad district having maximum abundance and species richness and Alnavar forest having minimum abundance and Khanapur forests having least species richness. Species such as *D. eugracilis*, *D. bipectinata,* and *D. malerkotliana* have maximum abundance in all the localities, whereas species such as *D. hypocausta* is reported for the first time from South India and *D. anomelani* for the first time from the Western Ghats forests of North Karnataka.

Introduction

Survey of *Drosophila* is still being carried out throughout India and has yielded a few new and rare species being discovered recently (Srinath and Shivanna, 2012, 2014a, b, 2017, and 2019; Achumi and Yenisetti, 2020; Fartyal *et al.,* 2017; Khan *et al.,* 2018; Bharadwaj *et al.,* 2019; Toda *et al.,* 2020). There is much scope for exploration of the faunal composition in unexplored regions. Western Ghats is one such habitat of interest, which is one of the biodiversity hotspots in India. These are a chain of mountains extending from the states of Gujarat, Maharashtra, Goa, Karnataka, Kerala, and eastern Tamil Nadu running parallel to India's western coast. It is approximately 30-50 km inland, and covers an area of around 140,000 km² in a 1,600 km long stretch (UNESCO data). The forests include evergreen, semi-evergreen, moist deciduous, and dry deciduous forests. Earlier studies on *Drosophila* fauna were much focused on Southern Western Ghats and parts of Mysore in Karnataka (Reddy and Krishnamurthy, 1968, 1970, 1971, 1973; Ranganath and Krishnamurthy, 1972; Gowda and Krishnamurthy, 1972; Prakash and Reddy, 1978; Prakash and Ramachandra, 2008). There are no reports or any studies on *Drosophila* from Northern Western Ghats of Karnataka except Nagaraj and Krishnamurthy (1980), whose study was mainly concentrated on forests of

Dandeli and Ambika Nagar. Hence, in order to fill this lacuna, a primary survey was initiated in a few Western Ghats forest localities of Dharwad, Uttara Kannada, and Belagavi districts.

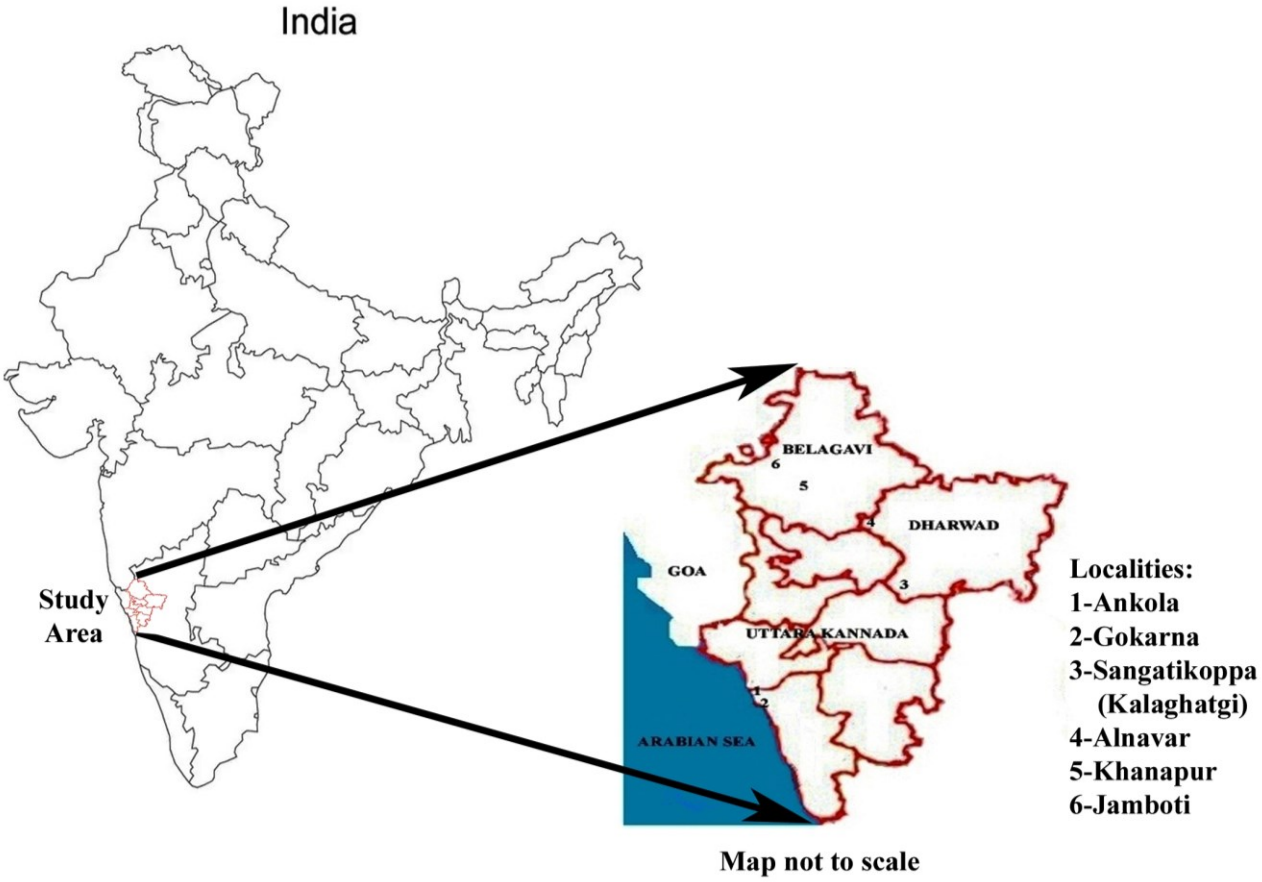


Figure 1. Map showing collection of Drosophilids from different localities of Western Ghats.

Table 1. List of localities selected for *Drosophila* collection from the study area.

Locality	District	Geographical position
Sangatigoppa (Kalghatgi)	Dharwad	15.072388 N, 74.881596 S
Alnavar	Dharwad	15.439585 N, 74.715079 S
Jamboti	Belagavi	15.687708 N, 74.355299 S
Nagargali (Khanapur)	Belagavi	15.423760 N, 74.601243 S
Gokarna	Uttara Kannada	14.546166 N, 74.317497 S
Ankola	Uttara Kannada	14.675590 N, 74.314257 S

Materials and Methods

Study Area

The study area included the Western Ghats forests of Belagavi, Dharwad, and Uttara Kannada district (Table 1) together comprising an area of 18713.874 km² (Karnataka Forest Department) (Figure 1).

Flies were collected during monsoon (August, September) and post monsoon season (October and November) of 2021. Net sweeping method was employed to collect the flies. Male flies were directly used for identification whereas females were cultured and their F1 male progenies were used for identification of

flies. Identification was done according to Bock and Wheeler (1972), Markow and O' Grady (2006), and Miller *et al.* (2017).

Results

Table 2 provides a list of species that were collected from the above mentioned localities. A total of 13 species were identified. Species such as *D. anomelani* and *D. hypocausta* were collected for the first time. The morphology of *D. anomelani* and *D. hypocausta* will be discussed elsewhere. Figure 2 reveals that *D. bipectinata*, *D. malerkotliana*, *D. eugracilis,* and *D. s. neonasuta* were the most commonly occurring species in these localities. Species belonging to *montium* subgroup such as *D. jambulina*, *D. punjabiensis,* and *D. anomelani* had lesser abundance. Other species such as *D. rajasekari*, *D. takahashii*, *D. n. nasuta*, *D. hypocausta*, *D. repleta*, and *Z. indianus* were among the least dominant species.

Table 2. Preliminary list of Drosophilid species collected from different localities of Western Ghats forests of Dharwad, Belagavi and Uttara Kannada districts.

Genus *Drosophila*
Subgenus ***Sophophora***
melanogaster species group
ananassae subgroup
1. *D. bipectinata*
2. *D. malerkotliana*
eugracilis subgroup
3. *D. eugracilis*
montium subgroup
4. *D. jambulina*
5. *D. punjabiensis*
6. *D. anomelani#*
suzukii subgroup
7. *D. rajasekari*
takahashii subgroup
8. *D. takahashii*
Subgenus ***Drosophila***
immigrans species group
immigrans subgroup
9. *D. n. nasuta*
10. *D. s. neonasuta*
hypocausta subgroup
11. *D. hypocausta**
repleta species group
12. *D. repleta*
Genus *Zaprionus*
Subgenus ***Zaprionus***
13. *Z. indianus*

First report from North Karnataka; * First report from South India

From the overall collections of Western Ghats forests of three districts, *D. eugracilis* (34%) is the most dominant species followed by *D. malerkotliana* (26%), *D. bipectinata* (24%), and *D. s. neonasuta* (9%). The percentage of remaining species stands below 5% (Figure 3). From the locality point of view Kalghatgi forest had maximum abundance (700) and species richness (10) and Alnavar forests recorded minimum abundance (210), whereas, Khanapur forests recorded minimum species richness (7).

Discussion

The Western Ghats forests are more of tropical in nature, which generally favors a number of endemic and rare species (Venkataraman, 2005). From the present study, it reveals that the distribution of species is uneven, and the data preliminarily indicate that species belonging to *melanogaster* species group of subgenus *Sophophora* hold a more predominant share than subgenus *Drosophila*. With 8 out of 13 species collected, it shows that *Sophophoran* taxa have maximum species richness among these localities of Western Ghats forests. Bock and Wheeler (1972) in their review of *melanogaster* species group have concluded that these species have more suitable habitats in tropical forests and recorded a number of new and endemic species. Earlier studies on *Drosophila* fauna from Western Ghats of South Karnataka have yielded similar results with species of *melanogaster* species group dominating over other groups (Prakash and Reddy, 1978). Species such as *D. eugracilis*, which is not a cosmopolitan species, has dominated these forests but is rarely found in urban localities (Srinath and Shivanna, 2014a). *D. melanogaster* considered to be a cosmopolitan species was also absent, whereas *D. malerkotliana* and *D. bipectinata* are the most dominant species collected from Western Ghats forest (Figures 2 and 3). They were also predominantly found in the urban collections from Dharwad district (Srinath and Shivanna, 2014a). The other subgroups such as *montium*, *takahashii,* and *suzukii* were rare during collections. Species from subgenus *Drosophila* have only 4 species with *D. s. neonasuta* collected as more common

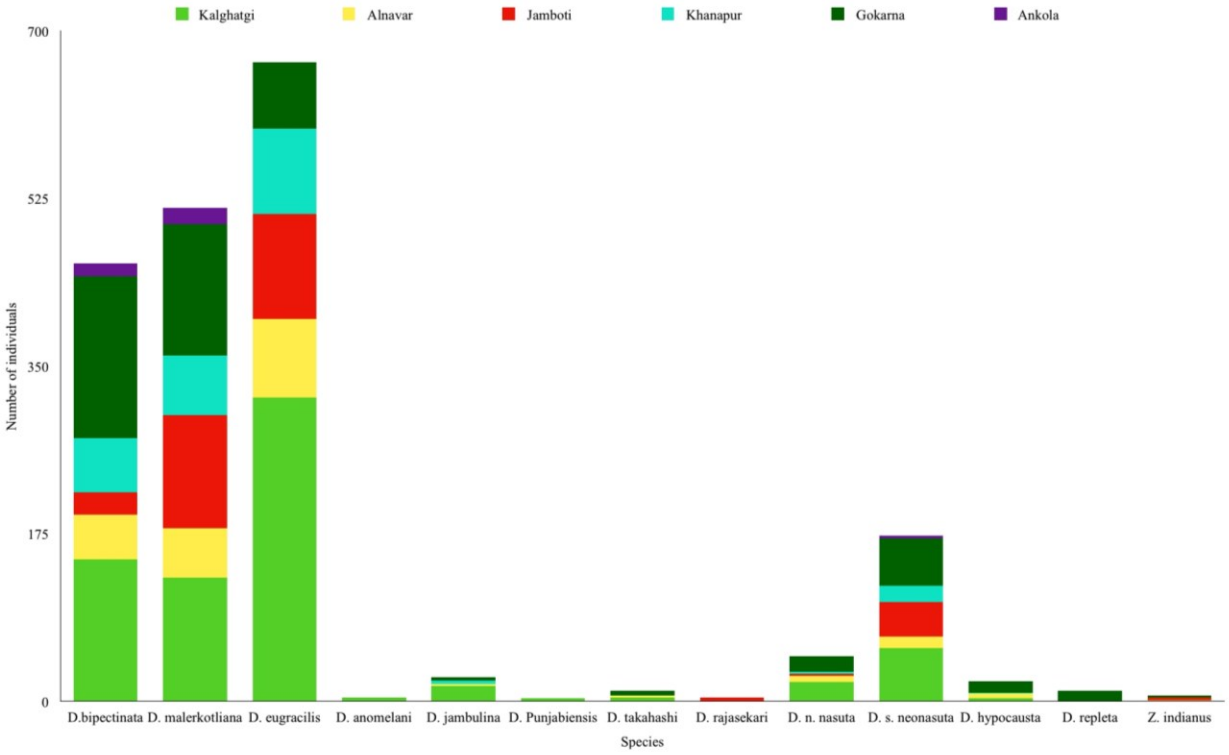


Figure 2. Species-wise abundance of Drosophilids from different localities of Western Ghats.

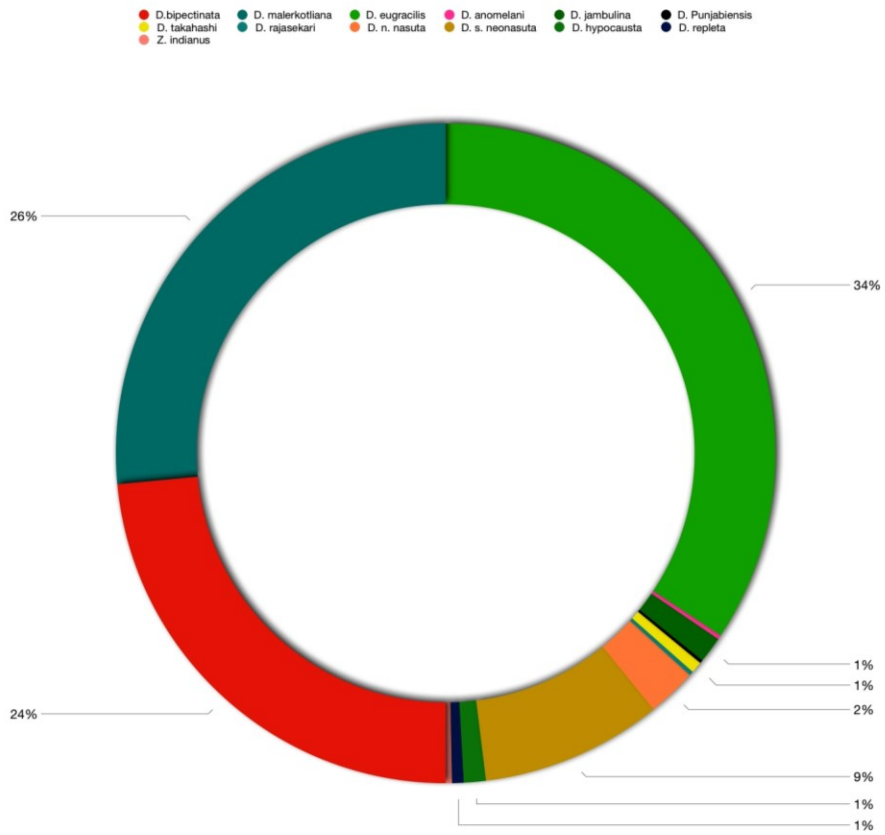


Figure 3. Overall percentage of Drosophilids collected from Western Ghats forests of Dharwad, Belagavi, and Uttara Kannada District.

species. *Z. indianus* is the only species reported outside genus *Drosophila* under family Drosophilidae. This species was dominantly found in mango plantations (Srinath and Shivanna, 2013) as well as other fruit plantations and also is considered as a pest in some countries of South America and United States (Santos *et al.,* 2003; and Van der Linde *et al.,* 2006). Hence it can be inferred that this species is less likely to dominate the natural forests.

The present study is a preliminary step to explore the vast localities of Western Ghats forests. Further intense survey of these localities may provide more information on species richness as well as their abundance. This will help in analysing the biodiversity of Drosophilids of the Western Ghats forests. Further, there are other localities such as Sirsi, Kumata, Honnavar, and Bhatkal regions of Uttara Kannada district that also covers the forests of Western Ghats which needs to be explored.

Acknowledgments: Authors thank the Chairman, Department of Zoology, Karnatak University Dharwad for providing necessary facilities. One of the authors (SBS) thanks PMEB, Karnatak University for the University Research Seed Grant Policy (KU/PMEB/2021-22/58, dated 25-06-2021) for financial assistance.

References: Achumi, B., and S.C. Yenisetti 2020, J. Ent. Res. 44(1): 15-22; Bharadwaj, A., M.Z. Khan, P. Fartyal, M. Sarswat, and R.S. Fartyal 2019, Ind. J. Ent. 81(1): 82-86; Bock, I.R., and M.R. Wheeler 1972, Univ. Tex. Publ. 7213: 1-102; Gowda, S.L., and N.B. Krishnamurthy 1972, Dros. Inf. Serv. 48: 38; Fartyal, R.S., P.C. Sati, S. Pradhan, M.C. Kandapal, M.J. Toda, R.N. Chaterjee, B.K. Singh, and A. Bharadwaj 2017, Zookeys 688: 49-79; Karnataka Forest Department 2021, www.aranya.gov.in; Khan, M.Z., A. Bharadwaj, and R.S. Fartyal 2018, Int. J. Life Sci. 6(2): 529-535; Markow, T.A., and P. O'Grady 2006, Drosophila: *A Guide to Species Identification and Use*. Academic Press. 247 pp.; Miller, M.E., S.A. Marshal, and D.A. Grimaldi 2017, Can. J. Arthro. Ident. 31: 1-282; Nagaraj, H.J., and N.B. Krishnamurthy 1980, Dros. Inf. Serv. 55: 114; Prakash, H.S., and G.S Reddy 1978, Entomon 3(1): 85-90; Prakash, A., and N.B. Ramachandra 2008, Dros. Inf. Serv. 91: 82-87; Ranganath, H.A., and N.B. Krishnamurthy 1972, Dros. Inf. Serv. 48: 132; Reddy, G.S., and N.B. Krishnamurthy 1968, Proc. Indian Acad. Sci. 68: 202-205; Reddy, G.S., and N.B. Krishnamurthy 1970, Journal Biol. Sci. 13(2): 24-29; Reddy, G.S., and N.B. Krishnamurthy 1971, Dros. Inf. Serv. 47: 116; Reddy, G.S., and N.B. Krishnamurthy 1973, Orient. Insects 7(2): 259-265; Santos, J.F., T.T. Rieger, S.R.C. Campos, A.C.C. Nascimento, P.T. Felix, S.V.O. Silva, and F.M.R. Freitas 2003, Dros. Inf. Serv. 86: 92-93; Srinath, B.S., and N. Shivanna 2012, Biosystematica 6(1): 39-42; Srinath, B.S., and N. Shivanna 2013, Dros. Inf. Serv. 96: 61-64; Srinath, B.S., and N. Shivanna 2014a, J. Ent. Zool. Stud. 2(4): 35-41; Srinath, B.S., and N. Shivanna 2014b, Dros. Inf. Serv. 97: 59-63; Srinath, B.S., and N. Shivanna 2017, Dros. Inf. Serv. 100: 58-62; Srinath, B.S., and N. Shivanna 2019, Dros. Inf. Serv. 102: 43-46; Toda, M.J., H. Banziger, P.C. Sati, R.S. Fartyal, A. Suwito, and T. Katoh 2020, Zootaxa 4789(1): 1-54; UNESCO website: https://whc.unesco.org/en/list/1342; Van der Linde, K., G.J. Steck, K. Hibbard, J.S. Birdsley, L.M. Alonso, and D. Houle 2006, Florida Entomologist 89(3): 402-404; Venkataraman, K., 2005, Proceedings of the state level UGC sponsored seminar on Biodiversity and its conservation pp. 1-11.

Mating speed and copulation duration of *Drosophila sturtevanti* in different male densities.

Machado, L.P.B.[1], L. Madi-Ravazzi[2], and R.P. Mateus.[1,*] [1]Laboratório de Genética e Evolução, Departamento de Ciências Biológicas, UNICENTRO – Universidade Estadual do Centro-Oeste, Guarapuava-PR, Brazil. [2]Laboratório de Genética, Ecologia e Evolução de *Drosophila*, Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista – IBILCE/UNESP; *Corresponding author: rogeriopmateus@gmail.com.

Introduction

Drosophila sturtevanti (*saltans* group of *Sophophora* subgenus) has a wide geographic distribution in the Neotropical region. It can be found from Mexico, Caribbean islands, to southern Brazil (Magalhães, 1962;

Bächli, 2021; Trava *et al.*, 2021). Populations of this species showed different results in reproductive isolation studies. Dobzhansky (1944) and Carareto (1994) detected incipient sexual isolation, while Hosaki-Kobayashi and Bicudo (1994) did not observe the same event. The wide ecological dispersal capacity of this species, associated with conflicting reproductive isolation data, makes *D. sturtevanti* a suitable model for evolutionary studies and analyses of reproductive characters that can lead to population differentiation. Differences in mating speed and copulation duration could result in sexual isolation among divergent populations (Cobb *et al.*, 1988; Price *et al.*, 2001; Tanuja *et al.*, 2001; Coyne *et al.*, 2002; Machado *et al.*, 2002; Bacigalupe *et al.*, 2007; Jennings *et al.*, 2011; Silva, 2019), and, in some species, copulation duration could be affected by male density (Gilchrist and Partridge, 2000; Lize *et al.*, 2012). Thus, the objective of this work was to verify the influence of male density on mating speed and copulation duration of *D. sturtervanti*.

Material and Methods

Five replicate crosses between virgin males and females, seven to ten days old, of *D. sturtevanti* strain from Curió (CE/Brazil) were performed at the following densities of males per female: one (1); five (5); and fifteen (15). Reproductive behavior was observed in vials containing culture medium, from 8:00 to 11:00 am. During observation, mating frequency, mating speed (time elapsed between the first contact among females and males and the copulation initiation), and copulation duration were scored. Statistical analyses of the difference in mating speed and copulation duration between crosses with different male densities were performed using Past 3.16 software (Hammer *et al.*, 2001).

Results and Discussion

Copulation was recorded in only one out of five replicates of one female and one male density (1:1), which made it impossible to compare the average duration of mating speed and copulation duration with the other male densities analyzed: one female and five males (1:5), and one female and fifteen males (1:15). No statistical difference was detected in the comparison of mean copulation duration between 1:5 and 1:15 densities (Tukey test $p = 0.1333$). The mean mating speed of 1:5 was significantly higher than 1:15 (Tukey test $p = 0.028$ after log transformation due to the high deviation in 1:5 density – Table 1). Remating was observed in only one of the 1:15 density replicates during the observation period, with copulation duration of 22 minutes (min), and mating speed of 55 min (in this case, mating speed was the time elapsed between first and second copulation). Remating behavior is not common in *D. sturtevanti* (Madi-Ravazzi, personal communication). Despite that the first copula has been considered effective (according to the copulation duration criterion used here), we did not confirm if it really was completely finished (with sperm transfer or other aspect related, not evaluated in this work). Therefore, it is possible that a failure during copulation could have allowed this female to accept remating.

Table 1. Mean (± standard deviation) mating speed and copulation duration, in minutes (min), of five replicated crosses of *D. sturtevanti* strain from Curió (CE, Brazil) with different female: male densities: one to one (1:1); one to five (1:5); and one to fifteen (1:15). *Mating success in only one of five replicates.

Reproductive Characters	Female: Male densities		
	1:1*	1:5	1:15
Mating speed	52 min	45 ± 35.35 min	23 ± 4.24 min
Copulation duration	11 min	17.85 ±1 .63 min	21.5 ± 0.71 min

Drosophila sturtevanti copulation duration between 10 and 20 min (Table 1) is not among the longest already recorded (Markow, 1996). In insects, this character depends on processes of sexually antagonist coevolution (Parker, 1979; Holland and Rice, 1998; Moore and Pizzari, 2005), due to the males and females distinct reproductive advantages. The results showed that male density is not determinant for copulation duration in *D. sturtevanti*, but it influences the frequency and mating speed. In fact, the greater the number of

males, the faster the initiation of copulation, which could eventually enable remating at higher male densities. This reproductive behavior favors female cryptic choice in her reproductive tract due to more mating/remating opportunities (Markow and Hocutt, 1998; Knowles and Markow, 2001). These results are different from other Neotropical *Drosophila* species of the *Drosophila* subgenus, *D. maculifrons* and D. *mediostriata* (Silva, 2019). The reproductive behavior of these species of *Drosophila* subgenus seems to support the mate-guarding hypothesis (Alcook, 1994; Schoefl and Taborsky, 2002; Skwierzyńska, 2002; Skwierzyńska *et al*., 2018), in which the longer copulation duration may result in a lower risk of sperm competition, as it would prevent the females to remate, ensuring the paternity of the offspring (Chen *et al.*, 1988; Alcook, 1994; Schoefl and Taborsky, 2002; Chapman *et al.*, 2003; Liu and Kubli, 2003; Mazzi *et al.*, 2009; Skwierzyńska *et al.*, 2018).

Keywords: Reproductive characters, *saltans* group, mating success.

References: Alcock, J., 1994, Annu. Rev. Entomol. Palo Alto 39: 1–21; Bächli, G., 2021, Access in December 13. Database. 2018, http://taxodros.unizh.ch/; Bacigalupe, L.D., H.S. Crudgington, F. Hunter, A.J. Moore, and R.R. Snook 2007, J. Evol. Biol. 20: 1763–1771; Carareto, C.M.A., 1994, Naturalia 19: 185-198; Chapman, T., G. Arnqvist, J. Bangham, and L. Rowe 2003, Trends in Ecol. Evol. 18: 41–47; Chen, P.S., E. Stumm-Zollinger, T. Aigaki, J. Balmer, M. Bienz, and P. Böhlen 1988, Cell 54: 291–298; Cobb, M., B. Burnet, and K. Connolly 1988, Behav. Genet. 18: 211–225; Coyne, J.A., S.Y. Kim, A.S. Chang, D. Lachaise, and S. Elwyn 2002, Evolution 56: 2424–2434; Dobzhansky, Th., 1944, Proc. Natl. Acad. Sci. USA. 30: 335–339; Gilchrist, A.S., and L. Partridge 2000, Evolution 54: 534–542; Hammer, Ø., D.A.T Harper, and P.D. Ryan 2001, Paleontol. Electron. 4: 9; Holland, B., and W.R. Rice 1998, Evolution 52: 1–7; Hosaki-Kobayashi, M.K., and H.E.M.C. Bicudo 1994, Braz. J. Genet. 17: 393–399; Jennings, J.H., D. Mazzi, M.G. Ritchie, and A. Hoikkala 2011, BMC Evol. Biol. 11: 68–77; Knowles, L.L., and T.A. Markow 2001, PNAS 98: 8692–8696; Liu, H., and E. Kubli 2003, PNAS USA. 100: 9929–9933; Lize, A., R.J. Doff, E.A. Smaller, Z. Lewis, and G.D.D. Hurst 2012, Biol. Lett. 8: 35-38; Machado, L.P.B., J.P. Castro, and L. Madi-Ravazzi 2002, Braz. J. Biol. 62: 601-608; Magalhães, L.E., 1962, Univ. Tex. Publ. 6205: 135–154; Markow, T.A., 1996, Evol. Biol. 29: 73–106; Markow, T.A., and G.D. Hocutt 1998, Reproductive isolation in Sonoran Desert *Drosophila*: testing the limits of the rules. *In*: *Endless Forms: Species and Speciation*. (Howard, D.J., and S.H. Berlocher, eds.). New York, Oxford Univ. Press 234–244; Mazzi, D., J. Kesäniemi, A. Hoikkala, and K. Klappert 2009, BMC Evol. Biol. 9: 1–13; Moore, A.J., and T. Pizzari 2005, Am. Nat. 165: S88–S97; Parker, G.A., 1979, Sexual selection and sexual conflict. *In*: *Sexual Selection and Reproductive Competition in Insects*. (Blum, M.S., and N.A. Blum, eds.). New York, Academic Press 123–166; Price, C.S.C., C.H. Kim, C.J. Gronlund, and J.A. Coyne 2001, Evolution 55: 81–92; Schoefl, G., and M. Taborsky 2002, Behav. Ecol. Sociobiol. 52: 426–433; Silva, L.T.O., 2019, Aspectos da biologia reprodutiva de espécies de *Drosophila* (Diptera: Drosophilidae) pertencentes a grupos irmãos Dissertação de Mestrado. Programa de Pós-Graduação em Biologia Evolutiva da UNICENTRO, 48 pp.; Skwierzyńska, A.M., A. Plesnar-Bielak, M. Kolasa, and J. Radwan 2018, Anim. Behav. 137: 75–82; Tanuja, M.T., N.B. Ramachandra, and H.A. Ranganath 2001, J. Biosci. 26: 71–76; Trava, B.M., R.P. Mateus, L.P.B. Machado, and L. Madi-Ravazzi 2021, Zool. Stud. 60: 46.

Drosophila fauna from an interior Atlantic Forest fragment with xerophytic enclave in Southern Brazil.

Mateus, R.P.*, M.C. Ferreira, and L.P.B. Machado. Laboratório de Genética e Evolução, Departamento de Ciências Biológicas, UNICENTRO – Universidade Estadual do Centro-Oeste, Guarapuava-PR, Brazil; *Corresponding author: rogeriopmateus@gmail.com

Introduction

In South America, there is a savanna corridor, located between the Atlantic and Amazonian rain forests, that isolates the two major dry areas, Caatinga (in northeastern Brazil) and Chaco (in northern

Argentina). This corridor has been called as "dry diagonal" (Bucher, 1982; Prado and Gibbs, 1993). Within it, and in adjacent biomes, cacti species can be found as isolated populations in sandy substrates or rocky fields, known as xerophytic enclaves (Ab'Saber, 2000). The mixed ombrophilous forest in southern Brazil, also known as Araucaria forest (Galindo-Leal and Câmara, 2003), is a phytophysiognomy of the Atlantic Forest biome, distributed mainly in the interior of southern Brazil, which presents such enclaves. Several surveys of the *Drosophila* fauna were already performed in locations with this type of vegetation (Mateus *et al.*, 2006, 2018; De Toni *et al.*, 2007, as examples). However, a survey of the *Drosophila* fauna from a xerophytic enclave in Segredo (a forest fragment next to the Governor Ney Aminthas de Barros Braga Hydroelectric Plant, formerly known as Segredo Hydroelectric Plant, at Mangueirinha city, Paraná State) was not performed so far. Thus, the objective of this work was to perform a *Drosophila* fauna survey from this enclave, in order to evaluate the environmental quality of this fragment, and to help future projects in the fields of biodiversity investigation, community ecology, population genetics, and/or phylogeography.

Material and Methods

The *Drosophila* fauna was collected in Segredo (25°46'28.96"S, 52°07'1.97"W – Figure 1). For three days, the adult flies were attracted to fermented banana and orange in open traps, dispersed no closer than 5 m from each other, hung on branches at ± 1.5 m height from the ground. After this period, the flies were captured with entomological nets, put in glass vials with culture medium and taken to the laboratory alive. The non-*repleta* species were screened based on their distinct pigmentation and morphology patterns (Freire-Maia and Pavan, 1949; Frota-Pessoa, 1954; Markow and O'Grady, 2006). The members of the *repleta* and *willistoni* groups are difficult to be identified, because they display a very similar pattern of pigmentation within each group. The females of the first group are shown as unidentified as the diagnostic characteristic is the male aedeagus (Vilela, 1983). The cryptic species of the *willistoni* group were grouped as "*D. willistoni* group".

Figure 1. Satellite image (Goggle Earth Pro software) of the location of *Drosophila* fauna collection in Segredo (Mangueirinha/PR – Brazil).

Table 1. Absolute (*A*) and relative (*r*) abundances for females and males of each *Drosophila* species collected in the xerophytic vegetation enclave in Segredo (Manguinhos/PR – Brazil).

Species	Females		Males		Total	
	A	*r*	*A*	*r*	*A*	*r*
Subgenus *Drosophila*						
- *caponei* group						
D. caponei	6	0.0131	3	0.0066	9	0.0197
- *cardini* group						
D. polymorpha	6	0.0131	7	0.0153	13	0.0284
- *guarani* group						
D. ornatifrons	4	0.0087	-	-	4	0.0087
- *immigrans* group						
D. immigrans	1	0.0022	-	-	1	0.0022
- *repleta* group						
D. antonietae	-	-	108	0.2358	108	0.2358
D. buzzatii	-	-	3	0.0066	3	0.0065
D. mercatorum	-	-	47	0.1026	47	0.1026
D. meridionalis	-	-	29	0.0633	29	0.0633
D. onca	-	-	1	0.0022	1	0.0022
D. paranaensis	-	-	4	0.0087	4	0.0087
- *repleta* group[1]	143	0.3122	-	-	143	0.3122
- *tripunctata* group						
D. mediostriata	-	-	2	0.0044	2	0.0044
Subgenus *Sophophora*						
- *melanogaster* group						
D. ananassae	-	-	1	0.0022	1	0.0022
D. kikkawai	-	-	8	0.0175	8	0.0175
D. simulans	31	0.0677	17	0.0371	48	0.1048
- *saltans* group						
D. prosaltans	3	0.0066	2	0.0044	5	0.0109
- *willistoni* group						
D. willistoni group[2]	17	0.0371	15	0.0328	32	0.0699
Total	211	0.4607	247	0.5393	458	1

[1] Females of cryptic species of the *Drosophila repleta* group;
[2] Cryptic species of the *D. willistoni* group that could not be identified.

Results and Discussion

In this study we were able to identify a total of 458 flies belonging to 2 subgenera, 10 groups, and at least 16 different species of *Drosophila* (Table 1). The *Drosophila* subgenus showed the highest richness (6 groups and 11 species), with 6 species belonging to the *Drosophila repleta* group. The *Sophophora* subgenus presented 4 groups and at least 5 species. This richness distribution among subgenera has been observed in other studies (Sene *et al.*, 1981; Tidon-Sklorz and Sene, 1995; Tidon-Sklorz and Sene, 1999; Medeiros and Klaczko, 2004; Mateus *et al.*, 2018).

Most of the species collected (13) and four of five most abundant species (r > 5%: *D. antonietae* – 23.58%, *D. mercatorum* – 10,26%, *D. willistoni* cryptic species – 6.99%, and *D. meridionalis* – 6.33%) were indigenous. Only three species were exotic, and *D. simulans* (*melanogaster* group) was the most abundant introduced species (r = 10.48%). This species has been collected in high frequency in several locations (Dobzhansky and Pavan, 1950; Pavan, 1959; Perondini *et al.*, 1979; Sene *et al.*, 1980; Tidon-Sklorz and Sene, 1992).

Cavasini *et al.* (2014) also obtained this richness difference between indigenous and exotic species, surveying in two highland *Araucaria* forest fragments without xeric vegetation, ≈80 km distant from Segredo, in Guarapuava/PR. However, they detected *D. kikkawai*, which also belongs to the *melanogaster* group, as the most abundant exotic species. With that result, they stated that, despite the abundance of indigenous species,

the high abundance of *D. kikkawai* suggests that both highland *Araucaria* forest fragments were in intermediate state of conservation. Therefore, we think that this interpretation could also be applied for the fragment at Segredo studied in this work.

Among the indigenous species, *D. antonietae*, *D. meridionalis,* and *D. mercatorum* were the most abundant species. *Drosophila antonietae* is a cactophilic species that always occurs associated with the columnar cactus *Cereus hildmaniannus* in the south and southeast regions of Brazil, and in the north of the eastern boundary of the Argentinean Chaco (Tidon-Sklorz and Sene, 2001). We already had collected this species in this enclave to perform microsatellite loci variability analyses (Machado *et al.*, 2012). *Drosophila meridionalis* is another cactophilic species that uses several species of the *Opuntia* and the *Cereus* cacti as breeding sites (Pereira *et al.*, 1983; Oliveira *et al.*, 2012; Barrios-Leal *et al.*, 2021). It co-occurs with *D. antonietae*, but extends its geographical distribution up to the Brazilian Caatinga Biome through the 'restinga' habitat of the southeast coast (Pereira *et al.*, 1983; Vilela, 1983). *Drosophila mercatorum* is a non-cactophilic common species found in natural environments in South America, especially in open areas (Sene *et al.,* 1981; Vilela *et al.*, 1983; Mateus *et al*., 2006).

These three species belong to the *repleta* group. This group represented 73.13% of the total specimens collected in Segredo, abundance much higher than those from previous works performed surveying *Drosophila* fauna of xerophytic enclaves located in southern Brazil: 38.72%, 35.82%, and 9.34% in Cianorte, Salto Santa Rosa, and Guartelá, respectively (semideciduous seasonal forest fragments - Mateus *et al.*, 2006); 25.77% in Rio do Poço (*Araucaria* forest fragment – Mateus *et al.*, 2018).

Drosophila antonietae was the most abundant *repleta* group species collected in Segredo (23.58%), followed by *D. mercatorum* (10.26%). In southern Brazil, this has been previously observed in Salto Santa Rosa and Cianorte (Mateus *et al.*, 2006), but not in Guartelá (Mateus *et al.*, 2006) and Rio do Poço (Mateus *et al.*, 2018), where *D. mercatorum* was more abundant than *D. antonietae*.

The sympatry between the two most abundant cactophilic species collected in Segredo, *D. antonietae* and *D. meridionalis*, has also been observed in other surveys in xerophytic enclaves. Mateus *et al.* (2006) detected this sympatry in three xeric areas from southeast (Itatiba, Santa Maria da Serra, and Itirapina) and in one from south of Brazil (Cianorte), with *D. meriodionalis* always in low abundance. On the other hand, Mateus *et al.* (2018), collecting in another *Araucaria* forest fragment in Rio do Poço, did not collect these two species in sympatry.

Another sympatry between cactophilic species, *D. antonietae* and *D. buzzati*, was detected in the xerophytic enclave of the *Araucaria* forest of Segredo. These species were already found in sympatry in fragments of other phytophysiognomies of the Atlantic forest biome with xerophytic vegetation enclave (Mateus *et al.*, 2006, 2018; De Toni *et al.*, 2007), but this is the first record of it in this phytophysiognomy (*Araucaria* forest).

The *Drosophila* fauna survey is important for several studies in a variety of fields, such as genetics, ecology, and evolution. Thus, this work contributed to the knowledge on the composition of the *Drosophila* community in xerophytic areas in the Neotropical region. This is important for further studies in any field considering *Drosophila* as a model.

Keywords: Community, *repleta* group, Atlantic forest biome, xeric vegetation, Neotropical region.

References: Ab'Saber, A.N., 2000, Rev. Inst. Geolog. SP 21: 71–78; Barrios-Leal, D.Y., R.P. Mateus, C.G. Santos, and M.H. Manfrin 2021, Neotrop. Entomol. 50: 515–523; Bucher, E.H., 1982, Chaco and caatinga - South American arid savannas, woodlands and thickets. *In*: *Ecology of Tropical Savanas.* (Huntey, B.J., and B.H. Walther, eds.). New York, Springer-Verlag, pp. 48–79; Cavasini, R., M.L.T. Buschini, L.P.B. Machado, and R.P. Mateus 2014, Braz. J. Biol. 74: 761-768; De Toni, D.C., M.S. Gottschalk, J. Cordeiro, P.R.P. Hofmann, and V.L.S. Valente 2007, Neotrop. Entomol. 36: 356–375; Dobzhansky, T., and C. Pavan 1950, J. Anim. Ecol. 19: 1–14; Freire-Maia, N., and C. Pavan 1949, Cultus. 1: 3–66; Frota-Pessoa, O., 1954, Arq. Museu Paran. 10: 253–330; Galindo-Leal, C., and I.G. Câmara 2003, Atlantic forest hotspots status: an overview. *In*: *The Atlantic Forest of South America: biodiversity status, threats, and outlook*. (Galindo-Leal, C., and I.G. Câmara, eds.). Washington (DC), Island Press 3–11; Machado, L.P.B., D.C. Silva, D.P. Simão, and R.P. Mateus 2012, Spatial Variation of Genetic Diversity in *Drosophila* Species from Two Different South American Environments. *In*: *Genetic Variation in Animals*. (Caliskan, M., ed.). Rijeka, Intech 45–62; Markow, T.A., and P.M. O'Grady 2006, Academic Press/Elsevier;

Mateus, R.P., M.L.T. Buschini, and F.M. Sene 2006, Braz. J. Biol. 66: 719–729; Mateus, R.P., L.P.B. Machado, and D.P. Simão-Silva 2018, Stud. Neotrop. Fauna Environ. 53: 152–161; Medeiros, H.F., and L.B. Klaczko 2004, Biota Neotropica 4: 1–12; Oliveira, D., F.C. Almeida, P.M. O'Grady, M.A. Armella, R. DeSalle, and W.J. Etges 2012, Mol. Phylogenet. Evol. 64: 533–544; Pavan, C., 1959, Bolm. Fac. Filos. Ciênc. S. Paulo 221: 1–81; Pereira, M.A.Q.R., C.R. Vilela, and F.M. Sene 1983, Cienc. Cult. 35: 1313–1319; Perondini, A.L.P., F.M. Sene, and L. Mori 1979, Egypt. J. Genet. Cytol. 8: 263–268; Prado, D.E., and P.E. Gibbs 1993, Ann. MO Bot. Gard. 80: 902–927; Sene, F.M., F.C. Val, C.R. Vilela, and M.A.Q.R. Pereira 1980, Pap. Avul. Dep. Zool. Sec. Agric. 33: 315–326; Sene, F.M., M.A.Q.R. Pereira, C.R. Vilela, and N.M.V. Bizzo 1981, Dros. Inf. Serv. 56: 118–121; Tidon-Sklorz, R., and F.M. Sene 1992, Rev. Bras. Biol. 52: 311–317; Tidon-Sklorz, R., and F.M. Sene 1995, Iheringia, Sér. Zool. 78: 85–94; Tidon-Sklorz, R., and F.M. Sene 1999, *Drosophila*. *In*: Invertebrados Terrestres. Vol.V. Biodiversidade do Estado de São Paulo. Síntese do Conhecimento ao Final do Século XX. (Brandão, C.R.F., and E.M. Cancello, eds.). São Paulo, FAPESP; Tidon-Sklorz, R., and F.M. Sene 2001, Iheringia, Sér. Zool. 90: 141–146; Vilela, C.R., 1983, Rev. Bras. Entomol. 27: 1–114; Vilela, C.R., M.A.Q.R. Pereira, and F.M. Sene 1983, Ciênc. Cult. 35: 66–70.

Transgenic resources for functional studies on the *hsrω* gene in *Drosophila.*

Sahu[1], Ranjan Kumar, Deepti Trivedi[2], and Subhash Chandra Lakhotia[1*]. [1]Cytogenetics Laboratory, Department of Zoology, Banaras Hindu University, Varanasi, 221005 India; [2]National Centre for Biological Sciences, Bangalore 560 065, India; [*]Corresponding author; Author emails: Ranjan Kumar Sahu: xray.sahu@gmail.com (ORCID ID: 0000-0002-7893-1361), Deepti Trivedi: dtrivedi@ncbs.res.in, S. C. Lakhotia: lakhotia@bhu.ac.in (ORCID ID: 0000-0003-1842-8411).

Introduction

The *hsrω* gene, located at the 93D cytogenetic region in *Drosophila melanogaster*, is an early discovered long non-coding RNA gene, which is developmentally active, stress responsive, and expressed in almost all cell types (Lakhotia, 2011). Its non-coding nature and singular inducibility by diverse amides (Lakhotia and Mukherjee, 1970; Tapadia and Lakhotia, 1997) make it different from other stress responsive genes. All species of *Drosophila* carry an equivalent of *hsrω*, with remarkable conservation of its architecture and inducibility amidst high sequence diversification (Sahu *et al*., 2020).

The *hsrω* gene's conserved architectural features (Figure 1a) include unique sequence encompassing two exons and a more conserved omega intron in its proximal part, while the distal region typically includes tandem repeat sequences (Sahu *et al*., 2020). A conserved short ORF (ORFω), with a potential to code for a short peptide, is present within the exon 1 in different species (Sahu *et al*., 2020).

The Flybase annotation indicates that the *hsrω* gene in *D. melanogaster* produces a total of seven transcripts of which some are nuclear while at least one of them is known to be cytoplasmic. The distal region includes a long stretch of 280 bp tandem repeats. A conserved miRNA gene (*mir-4951*) is present at the 3'-end of the *hsrω* gene (Sahu *et al*., 2020). This gene's nuclear transcripts carrying the distal tandem repeats organize the nucleoplasmic omega speckles that regulate dynamics of hnRNPs and some other RNA binding proteins in cells under normal and stressed conditions (Prasanth *et al.*, 2000; Singh and Lakhotia, 2015).

With a view to understand the diverse functions of this gene's multiple transcripts, we generated several transgenic *D. melanogaster* lines, using P-element based vectors. Two types of P-element vectors *viz*., *pUAST* (Giordano *et al.*, 2002) and *pCaSpeR5* (Thummel *et al.*, 1988), were used. The *pUAST* vector was used for targeted expression of the sequence of interest with a GAL4 driver, while the *pCaSpeR5* was used for expression of the specified *hsrω* gene sequence under its own promoter. In addition, two CRISPR/Cas9 mediated deletion alleles have also been generated.

Figure 1. Genomic coordinates of *hsrω* gene in *D. melanogaster* and the prepared transgenic constructs. **a.** Organization of the *hsrω* gene (adapted from www.flybase.org): two upward arrows indicate the two transcription start sites (TSS1 and TSS2), while the four downward arrows indicate the four annotated transcription termination sites (TTS1–TTS4); different segments of the gene are color-coded and named (noted above the segment) following (Sahu *et al*., 2020). **b.** Transcripts produced by the *hsrω* gene: the spliced out intron (yellow) is indicated by a black line joining the exons 1 and 2 regions (green boxes); the transcripts are named (on right of each transcript) following www.flybase.org (transcript names in parentheses are the common names in published literature). **c.** Constructs for expression of different regions of *hsrω* downstream of the UAS promoter (in *pUAST* vector). **d.** Constructs for expression of ORFω-EGFP fusion protein under the UAS promoter (in *pUAST* vector) or 4.3kb *hsrω* promoter (in *pCaSpeR5* vector). **e.** Constructs for inducing RNAi against exon 2 or repeat unit of *hsrω* gene or against the human Sat III transcripts in *pUAST* vector. **f.** Constructs in *pCaSpeR* vector carrying 4.3 kb (in *pCHRH* clone) or ~5kb (in *pCHRB* clone) of *hsrω* promoter (which also includes the UDR region of the RD transcript) and the indicated downstream transcribed sequence. Additional color codes for the *hsrω* promoter (*ω* PRO), *UAS* promoter (in the *pUAST* vector), the EGFP coding region (without its start AUG codon) and the human *SatIII* repeat are indicated at bottom of the figure.

Table 1. Summary of generated transgenic and gene-edited mutant *hsrω* alleles.

Plasmid construct	Promoter	Product	Transgenic line names	Comments
pUHRA	UAS	*hsrω RA*	*pUHRA*$^{2.1}$, *pUHRA*$^{2.2}$ and *pUHRA*$^{2.3}$ on chromosome 2 and *pUHRA*$^{3.1}$ and *pUHRA*$^{3.2}$ on chromosome 3.	All lines viable as homozygotes.
pUHRD	UAS	*hsrω RD*	*pUHRD*$^{2.1}$and *pUHRD*$^{2.}$ on chromosome 2 and *pUHRD*$^{3.1}$, *pUHRD*$^{3.2}$ and *pUHRD*$^{3.3}$ on chromosome 3.	All lines viable as homozygotes.
pUHRH	UAS	*hsrω RH*	*pUHRH*$^{2.1}$ on chromosome 2 and *pUHRH*$^{3.1}$, *pUHRH*$^{3.2}$, *pUHRH*$^{3.3}$, *pUHRH*$^{3.4}$ *and pUHRH*$^{3.5}$ on chromosome 3.	All lines viable except homozygous lethal *pUHRH*$^{2.1}$. Free-floating for 3 generations, generated homozygous viable *pUHRH*$^{2.1v}$ line.
pUHINT	UAS	Omega Intron	*pUHINT*$^{2.1}$ on chromosome 2 and *pUHINT*$^{3.1}$, *pUHINT*$^{3.2}$, *pUHINT*$^{3.3}$ and *pUHINT*$^{3.4}$ on chromosome 3.	All lines viable as homozygotes.
pUHRCG	UAS	ORFω tagged with EGFP	*pUHRCG*$^{2.1}$, *pUHRCG*$^{2.2}$ and *pUHRCG*$^{2.3}$) on chromosome 2 and *pUHRCG3*$^{.1}$ and *pUHRCG*$^{3.}$ On chromosome 3.	All lines viable as homozygotes.
pCHRCG	4.3kb of *hsrω* promoter	ORFω tagged with EGFP	*pCHRCG*$^{2.1}$ and *pCHRCG*$^{2.2}$, on chromosome 2 and *pCHRCG*$^{3.1}$, *pCHRCG*$^{3.2}$ and *pCHRCG*$^{3.3}$ on chromosome 3.	All lines viable as homozygotes.
pUH280A	UAS	280bp repeat unit in anti-sense orientation	*pUH280A*$^{1.1}$,\ on chromosome 1, *pUH280A*$^{2.1}$, *pUH280A*$^{2.2}$ *and pUH280A*$^{2.3}$ on chromosome 2 and *pUH280A*$^{3.1}$ on chromosome 3.	All lines viable as homozygotes.
UH280S	UAS	280bp repeat unit in sense orientation.	*pUH280S*$^{2.1}$ and *pUH280S*$^{2.2}$ on chromosome 2 and *pUH280S*$^{3.1}$, *pUH280S*$^{3.2}$ and *pUH280S*$^{3.3}$ on chromosome 3.	All lines viable as homozygotes.
pUS158	UAS	158bp of human *SatIII* repeats in sense orientation.	*pUS158A*$^{1.1}$ on chromosome 1, *pUS158A* $^{2.1}$, *pUS158A* $^{2.2}$ and *pUS158A* $^{2.3}$ on chromosome 2 and *pUS158A* $^{3.1}$ and *pUS158A* $^{3.2}$ on chromosome 3.	All lines viable as homozygotes.
pUS158A	UAS	158bp of human *SatIII* repeats in anti-sense orientation.	*pUS158A*$^{1.1}$ on chromosome 1, *pUS158A* $^{2.1}$, *pUS158A* $^{2.2}$ and *pUS158A* $^{2.3}$ on chromosome 2 *pUS158A* $^{3.1}$ and *pUS158A* $^{3.2}$ on chromosome 3.	All lines viable as homozygotes.
pCHRH	4.3kb of *hsrω* promoter	*hsrω RH*	*pCHRH*$^{3.1}$ on chromosome 2 and *pCHRH*$^{3.2}$ on chromosome 3.	All lines viable as homozygotes.
pCHRB	~5kb of *hsrω* promoter	*hsrω RB*	*pCHRB*$^{3.1}$and *pCHRB*$^{3.2}$ on chromosome 3.	All chromosome 2 lines homozygous lethal; after free-floating only the *pUHEX2A*$^{2.1}$ was recovered as viable line, which was renamed as *pUHEX2A*$^{2.1v}$.
pUHEX2A$^{2.1}$	UAS	*hsrω* exon 2 anti-sense	*pUHEX2A*$^{2.1}$, *pUHEX2A* $^{2.2}$ and *pUHEX2A* $^{2.3}$ on chromosome 2 and *pUHEX2A* $^{3.1}$, *pUHEX2A* $^{3.2}$ and *pUHEX2A* $^{3.3}$ and chromosome 3.	All lines viable as homozygotes.

Methods and Results

A. Constructs to conditionally express different transcripts of the *hsrω* gene

In order to understand functions of three diverse transcripts ranging in size from ~1.2kb to ~21.2kb (Figure 1b), different *hsrω* transcript producing regions were placed under the UAS promoter in *pUAST* vector.

i. ***pUHRA***: Due to unavailability of PstI in MCS of the *pUAST* vector, the 1.2kb PstI digested fragment from the *pJG10* plasmid (from M. L. Pardue's lab carrying the 1.2 Kb RA cDNA in *pSP65* vector) was initially inserted in *pBlueScript SK(-)* in forward orientation and named *pBSHRA*. The EcoRI + XbaI fragment from *pBSHRA* was ligated in the *pUAST* vector at EcoRI and XbaI sites to generate the *pUHRA* construct (Figure 1c).

ii. ***pUHRD***: A DNA fragment from TSS1 to TTS1 (~2.4kb) along with the intron segment was PCR amplified using the primers RDF and 2R (Table 2) with EcoRI and XbaI overhangs, respectively. After double digestion with EcoRI and XbaI, the amplified fragment was ligated in the *pUAST* vector to generate the *pUHRD* construct (Figure 1c).

Table 2. List of adapter primers used for amplification of inserts required for generation of transgenic constructs.

Oligo's Name	RE site	Sequence	Tm
1F	EcoRI	GCTCGAGC**GAATTC**TCACTCTCAAATGAAAAGTGTTCAAG	66.1ºC
1R	SpeI	TTCACTAT**ACTAGT**CGTGGGCCCTTGCCCCT	66.8ºC
2F	KpnI	GTACGACT**GGTACC**TATTTTTCCACGTCGGGCA	66.7ºC
2R	XbaI	TTGAACAT**TCTAGA**AGGGCCTTTCCTCCGAGGG	66.7ºC
EGFPF	SpeI	TTCATGAT**ACTAGT**GTGAGCAAGGGCGAGGAG	65.5ºC
EGFPR	KpnI	TCAGTACT**GGTACC**TTACTTGTACAGCTCGTCC	65.4ºC
INTF	EcoRI	ACTCGAAC**GAATTC**GTATACATACACAGC	60.0ºC
INTR	XbaI	GCGACCAG**TCTAGA**CTATAAAGTTTAGAG	60.0ºC
RDF	EcoRI	GCTCGAGC**GAATTC**CTCTCGAAAACTGAACATTA	64.2ºC
RHR	XbaI	GTGAACAT**TCTAGA**CGAAAGATTGTGATTGGGGTAATC	64.1ºC
HPF	EcoRI	GAGAGTGA**GAATTC**CCGGGGATTGTTTTGCG	65.5ºC
HPR	EcoRI	GAGAGTGA**GAATTC**CGTGACCGGACTGAGG	67.0ºC
EX2AF	EcoRI	TCTCGAGC**GAATTC**TTTGAATTTTTGAATTTTTATT	57.5ºC
EX2AR	XbaI	GCGAACAT**TCTAGA**ACAACAAAAATAGAAAAAAT	58.3ºC

Sequences in bold and grey highlighted letters and a stretch of 8nt at its 5' end represent desired restriction sites and adapter sequences, respectively.

iii. ***pUHRH***: The DNA fragment from TSS2 to TTS2 (~2.8kb) was amplified through PCR using the primers pairs 1F and RHR (Table 2) with EcoRI and XbaI overhangs, respectively. Following double digestion of the amplified fragment with the two restriction endonucleases, it was ligated with EcoRI and XbaI digested *pUAST* to obtain the *pUHRH* construct (Figure 1c).

iv. ***pUHINT***: For this construct, the *hsrω* intron (712bp) was amplified using the primers INTF and INTR (Table 2) with EcoRI + XbaI flanking sites. The resulting amplicon was double digested with the two enzymes and ligated with EcoRI and XbaI digested *pUAST* to produce the *pUHINT* construct (Figure 1c).

B. Transgenic constructs to detect the translatability of ORFω

i. ***pUHRCG***: In order to detect translatability of the short ORFω in exon 1, EGFP reporter sequence was inserted in frame downstream of the ORFω after removing the stop codon of ORFω and initiation codon of the reporter gene. To achieve this three PCR amplified DNA fragments were ligated. Two fragments were amplified from the *hsrω* gene and the third (EGFP fragment) from the *pEGFP-N1* plasmid (Clontech; Kang *et al.*, 2001). The first *hsrω* gene fragment was amplified from TSS2 to 3' end of ORFω (excluding its stop

codon) with primers 1F and 1R (Table 2) with EcoRI and SpeI overhang sites, respectively. The second fragment downstream of ORFω to TTS1 was amplified with primers 2F and 2R (Table 2) with KpnI and XbaI overhang sites, respectively. The EGFP CDS (excluding its initiation codon) was amplified from the *pEGFP-N1* plasmid with primers EGFPF and EGFPR (Table 2) having SpeI and KpnI overhang sites, respectively. Finally, the three amplicons were digested with respective sets of enzymes and ligated to form a linear EGFP tagged ORFω as part of the *hsrω-RC* transcript, which was finally ligated into *pUAST* vector at EcoRI and XbaI sites to generate the *pUHRCG* plasmid (Figure 1d). Appropriate in-frame fusion of the ORFω with EGFP in the *pUCHRG* construct was confirmed by dideoxy sequencing using the sequencing primers pUASTF and pUASTR, (Table 3) corresponding to the vector backbone.

Table 3. List of primers used for dideoxy sequencing of constructs.

Oligo's Name	Sequence	Tm
pUASTF	AACCAAGTAAATCAACTGC	47.2ºC
pUASTR	ATCTCTGTAGGTAGTTTGTC	50.1ºC
pC5F	AGTTCAATGATATCCAGTGCAG	53.2ºC
pC5R	TAACCCTTAGCATGTCCGTG	54.2ºC

Table 4. List of gRNAs used in generation of CRISPR/Cas9 knockouts (PAM site indicated in bold).

Oligo's Name	Sequence	Deletion
gRNA-1	GTAGAATGGAAATATCCAGA**AGG**	$hsr\omega^{KO}$
gRNA-2	GTCATTTTAACCTAACGTCC**TGG**	
gRNA-3	TCATTTGAGAGTGACGTGAC**CGG**	$ORF\omega^{KO}$
gRNA-4	GAAACCATACGCAAACCCCC**TGG**	

Table 5. List of primers used for screening of CRISPR/Cas9 knockouts.

Oligo's Name	Sequence	Deletion	Tm
HKOF1	TGCTGAGTAACGGGTATCTC	$hsr\omega^{KO}$	54.2ºC
HKOR1	CATGAGCCTGGGAACATCC		55.9ºC
RDUF	GCTTACCCACCTTTCTCACG	ORF^{KO}	56.3ºC
INTSR	GCTTCGATGCGACGTTTGAA		54.2ºC

ii. ***pCHRCG***: The EcoRI + XbaI fragment used for generating the above *pUHRCG* plasmid was also cloned in *pCaSpeR5* and screened for positive clones. The 4.3kb region upstream of the TSS2 of *hsrω* gene was amplified with the primers HPF and HPR (Table 2) having EcoRI sites at both the ends. The EcoRI digested 4.3kb promoter sequence carrying amplicon was inserted into the *pCaSpeR5+HRCG* plasmid at EcoRI site and the resulting plasmid (*pCHRG*) was used to transform *DH5α*. To screen clones (*pCHRCG*) with correct orientation (Figure 1d) of the promoter, plasmids from the transformed colonies were digested with XbaI and selected for those showing restriction pattern of two bands of 7931bp and 6992bp sizes, respectively. The correct orientation was further confirmed by di-deoxy sequencing of the insert and ligation junction using sequencing primers pC5F and pC5R (Table 3) corresponding to the vector backbone.

C. Transgenic constructs for over-expression of sense and antisense transcripts corresponding to the 280bp tandem repeat unit of *hsrω* and 158bp repeat unit of human *SatIII* transcripts

The widely used RNAi line against the nuclear transcripts of *hsrω*, generated earlier in our lab using the 280bp repeat cloned in the *pSympUAST-w* vector (Mallik and Lakhotia, 2009), produces sense and anti-sense transcripts under the UAS promoter and down-regulates the *hsrω* nuclear transcripts. We now generated two new RNAi lines in *pUAST* vector that can be used for down-regulation of cytoplasmic and nuclear transcripts, respectively, and one line that over-expresses the sense strand of the 280bp repeat unit. In addition, we also generated two transgenic lines that express the sense or anti-sense strand of the human *SatIII* repeats, which are considered to be functional homologs of the *Drosophila hsrω* nuclear transcripts (Jolly and Lakhotia, 2006), under the UAS promoter.

i. ***pUHEX2A***: Complete exon-2 region was aligned against transcript database of *Drosophila melanogaster* using the BLAST tool. A unique 80bp fragment from 3' end of Exon-2 was selected and PCR amplified using wild type genomic DNA and the primers EX2AF and EX2AR (Table 2) having EcoRI and XbaI overhangs, respectively. The PCR fragment was digested with respective restriction endonucleases and inserted into the *pUAST* vector to produce the *pUHEX2A* construct (Figure 1e).

ii. ***pUH280A***: In order to clone the 280bp repeat of *hsrω* in antisense orientation in *pUAST* vector, the *pDRM30* plasmid clone (from M. L. Pardue's Lab), carrying 280bp Asu II fragment of *hsrω* repeat in *pGEM3* vector was double digested with PstI + XbaI to release the 280bp fragment which was re-cloned into *pBlueScript SK(-)* and named as *pBS280*. The NotI + XhoI fragment from *pBS280* was ligated into the NotI + XhoI digested *pUAST* vector to generate the *pUH280A* construct (Figure 1e).

iii. ***pUH280S***: The above *pBS280* was double digested with EcoRI + XbaI to release the 280bp repeat fragment which was re-cloned in the *pUAST* vector to generate the *pUHS280S* construct (Figure 1e).

iv. ***pUS158A***: In order to clone the 158bp repeat of *SatIII* in antisense orientation in *pUAST* vector, the ~160bp EcoRI + BamHI digested fragment from *pGEM2-98* clone, harboring a 158-bp fragment of the human *SatIII* repeat from the 9q12 locus in *pGEM2* vector (a kind gift from Dr. Caroline Jolly, INSERM, France) was inserted into the *pBlueScript SK (-)* vector, and named as *pBS158.* The NotI + XhoI fragment from *pBS158* was subcloned in the *pUAST* vector to produce the *pUS158A* construct (Figure 1e) that can down-regulate the human *SatIII* transcripts.

v. ***pUS158S***: In order to generate this clone, the EcoRI + XbaI digested fragment from the above *pBS158* was re-cloned in *pUAST* at EcoRI + XbaI sites to generate the *pUS158A* transgene construct (Figure 1e), for overexpression of the human *SatIII* repeat.

D. Transgenic constructs to generate *hsrω*-RH and RB transcripts under its own promoter

i. ***pCHRH***: In order to generate this clone, the EcoRI + XbaI digested ~2.8kb fragment from the above described *pUHRH* plasmid was ligated into EcoRI + XbaI digested *pCaSpeR5* and screened for the positive clones (*pCaSpeR5+HRH* clone). The 4.3kb promoter region of the *hsrω* gene (as described above for the *pCHRCG* clone) was ligated at the EcoRI site of the *pCaSpeR5+HRH* clone to generate the *pCHRH* plasmid (Figure 1f). To identify the clones with right orientation of the promoter with respect to the HRH sequence, isolated *pCHRH* plasmid DNAs from different transformed colonies were digested with XbaI to identify colonies that carry plasmid which generates two fragments of 7040bp and 7931bp, respectively.

ii. ***pCHRB***: To generate this construct, the *pDRM102* plasmid (from M. L. Pardue's lab) was digested with EcoRI + BamHI to generate three fragments of ~19kb, ~1kb, and ~2.2kb sizes. The ~19kb fragment (corresponding to the *hsrω-RB* transcribing region and ~5kb upstream endogenous promoter region) was purified, ligated into the *pCaSpeR5* at EcoRI and BamHI sites to generate the *pCHRB* construct (Figure 1f) and transformed into DH5α. For confirming positive colonies, the isolated plasmids were double digested with PstI + BsaI to get five fragments of 1935bp, 2229bp, 3445bp, 6612bp, and 12768bp, respectively.

The above plasmid constructs were used to transform DH5α and the purified plasmid DNAs in each case were confirmed to carry the desired transgene by restriction digestion and/or dideoxy sequencing. The purified DNAs were used for microinjection in w^- embryos together with the helper p-transposase-encoding plasmid DNA at the Fly-facility of the Bangalore Life Science Cluster, Bangalore. The emerging flies were crossed with w^- flies. The progeny flies with red-eyes (due to germline transformation with the transgenic plasmid carrying the *mini-white* reporter) were selected and crossed with *w*-; *Sp/CyO; TM3/TM6B* double balancer stock to determine the chromosomal insertion of the transgene and subsequent establishment of transgenic stocks. In some cases, the transgene insertion was associated with recessive lethality (see Table 1). In such cases, the transgene carrying parent was crossed with w^-*; +/+; +/+* flies and the red-eyed progenies were crossed with white-eyed (w^-*; +/+; +/+*) flies for three generations to remove potential 2nd site mutation

that may be responsible for the recessive lethality. In several cases, free-floating for three generations produced transgene-carrying homozygous viable lines, which are maintained as stocks for further studies (Table 1).

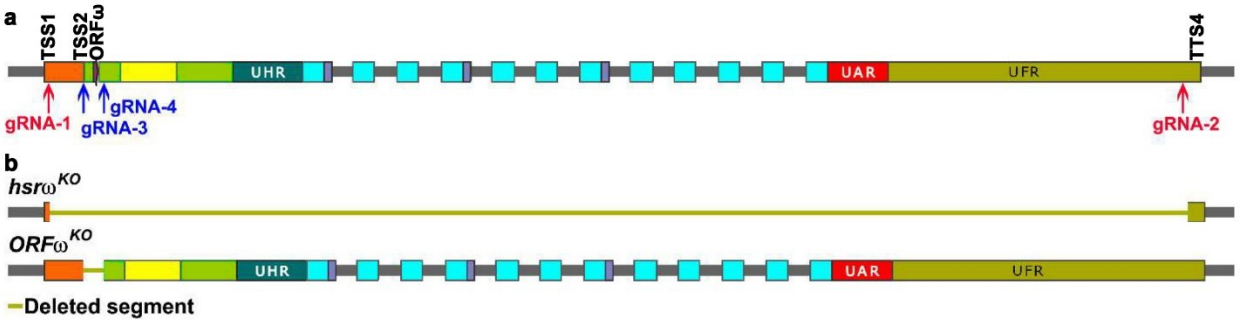


Figure 2. Generation of knockouts of *hsrω* using CRISPR/Cas9 system**.** **a**. Schematic of the *hsrω* gene with target locations of different gRNAs (gRNA-1 to gRNA-4) indicated in relation to the transcription start sites (TSS1 and TSS 2), ORFω, and the last transcription termination site (TTS4). **b**. Maps of the $hsr\omega^{KO}$ and $ORF\omega^{KO}$ alleles carrying complete deletion of *hsrω* (from just after the TSS1 to just before the TTS4) or of the ORFω (from TSS2 to just after the ORFω in exon 1), respectively. The deleted regions are indicated by green line.

E. Transgenic constructs to generate CRISPR/Cas9 mediated ORFω knockout and complete *hsrω* gene knockout alleles

The CRISPR/Cas9 mediated genome editing was used to generate two deletion alleles of *hsrω* gene, one with the entire gene deleted, and the other with deletion of the ORFω region. The specific gRNAs (Table 4) for each, with the "NGG" sequences at their 3' ends, were designed with the help of Benchling tool (www.benchling.com).

i. ***$hsr\omega^{KO}$***: For complete deletion of the *hsrω* gene, the two gRNAs (Table 4) corresponded, respectively, to the 5' region (gRNA-1) downstream of TSS1, and 3' region (gRNA-2) upstream of the TTS4 (Figure 2a). These were cloned in *pBFv-U6.2B* for generating a transgenic stock in which the transgene is inserted on chromosome II (y^1 v^1, $gRNA^{1+2}$/*CyO; +/+)*. The transgenic line carrying the gRNAs was crossed with flies expressing Cas9 in germline (y^1 v^1*; nos-Cas9*). The progeny F1 flies carrying the *gRNA* and *Cas9* transgenes were crossed with w^{1118}*; +/+; TM3, Sb/TM6B* flies. The F2 progeny males were pair mated with w^{1118}*; +/+; TM3, Sb/TM6B* females. After 2-3 days when larvae started appearing in the food vials, the parental males from each of the pair-mating vials were removed and their DNAs were isolated using the single fly DNA extraction method. Each of the F2 parental male DNA sample was screened for the desired deletion by PCR using the HKOF1 and HKOR1 primer pairs (Table 5). Only the DNA from which the *hsrω* gene is knocked out will generate a 674bp size amplicon, while in wild type genomic DNA, the two primers anneal more than 18kb apart and thus fail to amplify the intervening DNA (Figure 2b). The selected F3 flies were self-crossed to maintain the two *hsrω* knockout alleles ($hsr\omega^{KO1}$ and $hsr\omega^{KO2}$). Absence of the gRNA or the Cas9 transgene in the deletion-positive flies was confirmed molecularly. Both these lines show complete embryonic/larval lethality when homozygous. They are maintained with a 3rd chromosome balancer.

ii. ***$ORF\omega^{KO}$***: In order to generate ORFω knockout allele using the CRISPR/Cas9 technology, the 5' gRNA corresponded to TSS2 region (gRNA-3, Table 4, Figure 2a), while the 3' gRNA (gRNA-4, Table 4, Figure 2a) corresponded to exon 1 region downstream of the ORFω. These gRNAs were cloned in *pBFv-U6.2B* and used to generate transgenic stock (y^1 v^1, $gRNA^{3+4}$/*CyO; +/+)*. The transgenic progeny was crossed following the above scheme. PCR–based screening of ~300 F2 parental males using the primers RDUF and INTSR (Table 5) was carried out to identify males that carried the deletion of 268bp (Figure 2b) and thus generated an

amplicon of 543bp. Two independent transgenic lines ($ORF\omega^{KO1}$ and $ORF\omega^{KO2}$) were obtained, both of which were homozygous viable but completely female sterile. A free floating of $ORF\omega^{KO2}$ allele carrying chromosome for three generations yielded progeny male and female flies that carried the $ORF\omega^{KO2}$ allele but were fertile. Presence of the $ORF\omega^{KO2}$ allele in this line was confirmed by PCR screening using the ORFω deletion detecting primers (RDUF and INTSR) and the allele was renamed as $ORF\omega^{KO2f}$.

Discussion

The various transgenic lines expressing *RD*, *RH*, or *RB* transcripts under the *UAS* promoter can help in understanding functions of these transcripts through analysis of phenotypes induced by their targeted over-expression or by expressing them in *hsrω* null-background. The $hsr\omega^{KO}$ transgenic lines will be especially useful in this context. Likewise, the transgenic lines producing the *RH* or *RB* transcripts under the endogenous *hsrω* promoter will also be very useful when brought in *hsrω*-null background, since in this case, unlike the *UAS*-promoter carrying transgene, each of these two transgenes are expected to express more or less in the normal developmentally regulated manner.

As noted above, the presently available *hsrω-RNAi* transgene primarily targets the 280bp repeat units in the nuclear transcripts and thus has little effect on the *hsrω* cytoplasmic transcripts (Mallik and Lakhotia, 2009). The *hsrω*-exon RNAi transgenic line generated in this study is expected to down-regulate all the transcripts of this gene, since the exonic region is common to all the seven known transcripts; this RNAi line may have a greater effect on cytoplasmic *hsrω* transcript.

The *hsrω* large nuclear repeat-carrying transcripts and the human *SatIII* transcripts are functional homologs (Jolly and Lakhotia, 2006). The two transgenic lines carrying the *SatIII* repeat unit in sense and anti-sense orientations, respectively, will thus be useful in comparing functions of the 280bp *hsrω* repeat unit and the 158bp *Sat III* repeat units and to see if these can complement each other.

We believe that the transgenic lines and the knockout alleles generated in this study will provide valuable resources and would indeed help researchers in a deeper understanding of the *hsrω* gene, which is one of the earliest known lncRNA genes (Lakhotia, 2011). Desiring users may write to the corresponding author for obtaining the required lines described here.

Acknowledgments: This work has been supported by research grant from the Department of Biotechnology (Govt. of India) to SCL. RS was supported by a Research fellowship from the Council of Scientific and Industrial Research (Govt. of India). We thank the Fly Facility, Bangalore Life Science Cluster, Bangalore, for micro-injections in fly embryos to generate the transgenic lines.

References: Giordano, E., R. Rendina, I. Peluso, and M. Furia 2002, Genetics 160: 637-648; Jolly, C., and S.C. Lakhotia 2006, Nucleic Acids Research 34: 5508-5514; Kang, S.G., E. Lee, S. Schaus, and E. Henderson 2001, Mol. Med. 33: 174-178; Lakhotia, S.C., 2011, J. Biosciences 36: 399-423; Lakhotia, S.C., and A.S. Mukherjee 1970, Dros. Inf. Serv. 45: 108; Mallik, M., and S.C. Lakhotia 2009, RNA Biol. 6: 464-478; Prasanth, K., T. Rajendra, A. Lal, and S.C. Lakhotia 2000, J. Cell Sci. 113: 3485-3497; Sahu, R.K., E. Mutt, and S.C. Lakhotia 2020, J. Genet. 99: 64; Singh, A.K., and S.C. Lakhotia 2015, Chromosoma 124: 367-383; Tapadia, M.G., and S.C. Lakhotia 1997, 5: 359-362; Thummel, C.S., A.M. Boulet, and H.D. Lipshitz 1988, Gene 74: 445-456.

Spiroplasma prevalence within a *Drosophila melanogaster* natural population in Campinas, São Paulo State, Brazil.

Mondini, C.H., and L.B. Klaczko. Laboratório de Biodiversidade Genética e Evolução de *Drosophila*, Departamento de Genética, Evolução, Microbiologia e Imunologia, Instituto de Biologia, Universidade Estadual de Campinas – UNICAMP, Cx. Postal 6109, Campinas, 13083-970 SP, Brazil; E-mail: cesarmondini@hotmail.com; lbk@unicamp.br.

Spiroplasma poulsonii is an endosymbiont associated with *Drosophila,* and it frequently causes the death of the male offspring of infected mothers in early developmental stages (Anbutsu and Fukatsu, 2011; Ventura *et al*., 2012). *Spiroplasma* is the only endosymbiont that causes the death of males in *Drosophila melanogaster*, and this effect leads to an evolutionary conundrum: how could a costly infection that causes the fitness of males to plummet to near zero be maintained throughout evolutionary time? Perhaps there is no way that these infections can persist and would extinguish themselves (*i.e.,* Suicide King – Dybdahl and Storfer, 2003).

Although there are theoretical explanations for the maintenance of the prevalence of male-killers (Hurst and Majerus, 1993), and some works present evidence of those explanations for *Spiroplasma* infections in *D. melanogaster*, as horizontal transmission (Jaenike *et al*., 2007), or induction of defensive phenotypes to pathogens in the hosts (Mateos *et al*., 2016; Hamilton *et al*., 2016; Paredes *et al*., 2016), no evidence that fulfilled all the theoretical criteria was found to date. This lack of information can lead to the conclusion that these infections would disappear in natural populations, and a way to assess that is through the evaluation of the infection status of flies in nature. Such investigations, however, are few, totaling four papers (Montenegro *et al*., 2005; Pool *et al*., 2006; Watts *et al*., 2009; Ventura *et al*., 2012). From these, only the works of Montenegro and Ventura evaluated samples of more than 50 females. Due to the poor assessment of these prevalences, we here estimate the frequency of this infection using 251 females collected in the region of Campinas, Brazil.

Flies were captured using bottles with standard cornmeal- molasses medium and funnels placed for 1-14 days inside residences in the District of Barão Geraldo, Campinas, and in Sumaré, both in São Paulo State, Brazil. Thirty-two collections were held from February to December of 2016 (Table 1). Over 500 flies were collected, and after final identification we obtained 251 *D. melanogaster* females that produced offspring. They were screened using PCR. DNA extractions used 5-10 two-week old daughters and an alcohol-salt method (Aljanabi and Martinez, 1997). PCR reactions were carried out with the primers used by Montenegro *et al.* (2005) for the 16S rDNA gene of *Spiroplasma* and the primers listed by Simon *et al.* (1994) for the CoI gene of *D. melanogaster*, used to assess the quality of the DNA extraction and PCR amplification.

From the 251 isofemale lines tested, none was diagnosed as infected with *Spiroplasma*. Using Wilson's method (Brown *et al*., 2001) with a confidence level of 95%, the estimated prevalence interval is 0-1.51%. The estimate maximum value (1.51%) is lower than every prevalence obtained before (Table 2). Moreover, the number of captured females used is higher than any other in the literature, indicating that the inexistence of positively diagnosed females is not a sampling or methodological error, but an effect of the very low frequency of the infection.

Beyond the evolutionary paradox afore mentioned, *Spiroplasma* transmission is known to be sensitive to high and low temperatures (Anbutsu *et al.*, 2008; Montenegro *et al.*, 2004) as well as the age of the mother (Kageyama *et al*., 2007), and these factors could be minimizing the frequency of this infection in the region of Campinas. It could also be a direct observation of a temporal reduction that the infection is going through, as a trend for reduction in prevalence was observed before (Ventura *et al*., 2012). Two strategies to disentangle these questions would be to run another round of prevalence estimation in Campinas in a few years, and the other would be to make the same procedure now in the populations whose prevalence had previously been assessed.

Table 1. Collection information of tested isofemale lines.

Collection Date (2016)	GPS coordinates	Putative *D. melanogaster**	*D. melanogaster* isofemale lines tested **
Feb 21 – Feb 27	-22°49'05", - 47°05'21"	58	13
	-22°54'01", - 47°03'16"	45	38
Feb 18 – Feb 23	-22°54'01", - 47°03'16"	11	10
Feb 24 – Feb 29	-22°49'38", - 47°04'44"	16	10
Feb 28 – Feb 29	-22°49'38", - 47°06'13"	4	0
Jul 19 – Jul 28	-22°49'05", - 47°05'21"	8	4
Jul 12 – Jul 28	-22°49'02", - 47°05'23"	3	3
	-22°49'17", - 47°05'38"	3	1
Aug 19 – Aug 16	-22°49'05", - 47°05'21"	12	7
	-22°49'02", - 47°05'23"	15	8
Aug 15 – Aug 20	-22°49'05", - 47°05'21"	3	1
	-22°49'02", - 47°05'23"	26	15
Aug 29 – Sep 08	-22°49'02", - 47°05'23"	12	9
	-22°49'05", - 47°05'21"	11	7
Sep 10 – Sep 15	-22°49'05", - 47°05'21"	26	5
	-22°49'02", - 47°05'23"	9	4
	-22°49'17", - 47°05'38"	20	11
Sep 16 – Sep 23	-22°49'05", - 47°05'21"	1	0
	-22°49'02", - 47°05'23"	3	2
	-22°49'17", - 47°05'38"	0	0
Sep 18 – Sep 22	-22°47'16", - 45°05'07"	8	4
Sep 22 – Sep 26	-22°47'16", - 45°05'07"	9	2
Sep 27 – Sep 29	-22°49'38", - 47°06'13"	11	1
Sep 23 – Sep 29	-22°49'05", - 47°05'21"	32	12
Oct 05 – Oct 08	-22°49'05", - 47°05'21"	86	24
Oct 28 – Nov 04	-22°49'05", - 47°05'21"	3	3
Nov 08 – Nov 15	-22°48'33", - 47°16'59"	31	25
Nov 10 – Nov 16	-22°54'38", - 47°05'39"	18	3
Nov 16 – Nov 24	-22°49'05", - 47°05'21"	18	13
Nov 20 – Nov 24	-22°54'38", - 47°05'39"	41	2
Nov 22 – Nov 28	-22°48'33", - 47°16'59"	15	0
Dec 09 – Dec 11	-22°49'38", - 47°06'13"	25	14
	TOTAL	583	251

*Putative D. melanogaster females were then confirmed by the analysis of the males. **Diagnostic was performed through PCR reaction using 6 to 10 daughters from each female.

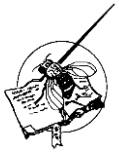
Acknowledgments: M.R.D. Batista, M.T. Brianti, L.F.M. Bartoletti, P.J. Bergamo, F. Uno, R.E.S. Penha and G.B. Passos for helping with the collection. K.A. Carvalho and C. Couto for technical support. I.M. Ventura for insightful tips. CAPES, CNPq and FAPESP for financial support.

References: Aljanabi, S.M., and I. Martinez 1997, Nucleic Acids Research 25: 4692–4693; Anbutsu, H., and T. Fukatsu 2011, Environmental Microbiology Reports 3: 144–153; Anbutsu, H., S. Goto, and T.

Table 2. Estimates of *Spiroplasma* infection within *D. melanogaster* natural populations available in the literature and the one from Campinas. Intervals were calculated using Wilson's method for a binomial distribution (Brown et al., 2001) and a 95% confidence level. A. Montenegro *et al*., 2005; B. Pool *et al*., 2006; C. Watts *et al*., 2009; D. Ventura *et al*., 2012.

Location	Collection Year	Exact Estimate	Minimum Estimate	Maximum Estimate
Campinas	2016	0	0	1.5
Uganda [C]	2007	2.6	0.5	13.5
Recife [D]	2009	0	0	13
Recife [A]	2003	2.3	0.6	5.8
Rio de Janeiro [D]	2008	1.7	0.01	9.91
USA/Mexico [B]	2005	1.3	0.2	6.8
Salvador [D]	2010	9.6	4.5	18.8
Salvador [D]	2009	12.5	3.5	31.8
Salvador [D]	2008	17.7	13.2	23.2

Fukatsu 2008, Applied and Environmental Microbiology 74: 6053–6059; Brown, L.D., T.T. Cai, and A. Dasgupta 2001, Statistical Science 16: 101–133; Dybdahl, M.F., and A. Storfer 2003, Trends in Ecology and Evolution 18: 523–530; Hamilton, P.T., F. Peng, M.J. Boulanger, S.J. Perlman, and N.A. Moran 2016, PNAS 113: 350–355; Hurst, G.D.D., and M.E.N. Majerus 1993, Heredity 71: 81–95; Jaenike, J., M. Polak, A. Fiskin, M. Helou, and M. Minhas 2007, Biology Letters 3: 23–25; Kageyama, D., H. Anbutsu, M. Shimada, and T. Fukatsu 2007, Naturwissenschaften 94: 333-337; Mateos, M., L. Winter, C. Winter, V.M. Higareda-Alvear, E. Martinez-Romero, and J. Xie 2016, Ecology and Evolution 6: 2679– 2687; Montenegro, H., V.N. Solferini, L.B. Klaczko, and G.D.D. Hurst 2005, Insect Molecular Biology 14: 281–287; Montenegro, H., and L.B. Klaczko 2004, Journal of Invertebrate Pathology 86: 50–51; Paredes, J.C., J.K. Herren, F. Schüpfer, and B. Lemaitre 2016, mBio 7: 1006–16; Pool, J. E., A. Wong, and C.F. Aquadro 2006, Heredity 97: 27–32; Simon, C., F. Frati, A. Beckenbach, B. Crespi, H. Liu, and P. Flook 1994, Annals of the Entomological Society of America 87: 651–701; Ventura, I.M., A.B. Martins, M.L. Lyra, C.A.C Andrade, K.A. Carvalho, and L.B Klaczko 2012, Microbial Ecology 64: 794–801; Watts, T., T.S. Haselkorn, N.A. Moran, and T.A. Markow 2009, PLoS ONE 4: e5703.



Drosophilidae (Insecta, Diptera) in the Brazilian Amazon.

Brito, Renata dos R.[1*], Rosângela Santa-Brígida[2], Hermes F. de Medeiros[3], Rita de Cassia O. dos Santos[4], and Marlucia M. Martins[5]. [1]Federal University of Pará, Belém, Brazil ([1]Universidade Federal do Pará, Belém, Pará, Brasil); [2]Emílio Goeldi Paraense Museum, Belem, Brazil ([2]Museu Paraense Emílio Goeldi, Belém, Brasil); [3]Federal University of Pará, Belém, Brazil ([3]Universidade Federal do Pará, Belém, Pará, Brasil); [4]Federal University of Pará, Belém, Brazil ([4]Universidade Federal do Pará, Belém, Pará, Brasil); [5]Emílio Goeldi Paraense Museum, Belém, Brazil ([5]Museu Paraense Emílio Goeldi, Belém, Pará, Brasil). *Corresponding author: renatabrito98@gmail.com

Introduction

Drosophilidae are one of the most speciouse families of Diptera Acalyptrata, with about 4,500 species (Bächli, 2021). Among these, 305 species have been recorded for Brazil (Tidon *et al*., 2021). However, these

records are not uniformly distributed throughout the territory, since most studies have been conducted in the Southeast, South, and Center-West regions of the country. The North and Northeast regions (except for Bahia) were the most lacking in surveys for the group (Gottschalk *et al*., 2008). According to this author the states of Pará and Amazonas were those that stood out in the number of records, with 66 and 39 species, respectively.

Table 1. Species of Drosophilidae recorded in the Brazilian Amazon states. An X means the species was recorded.

Taxon		**Brazilian Amazon states**							
Drosophilinae	**Species**	AM	AP	AC	MA	MT	PA	RO	RR
Chymomyza	*C. bicoloripes* (Malloch, 1926)						X		
	C. diatropa Grimaldi, 1986						X		
	C. laevilimbata (Duda,1927)					X			
Cladochaeta	*C. amorimi* Pirani and Carvalho-Filho, 2019						X		
	C. caxiuana Carvalho-Filho, Pirani and Kloss, 2018						X		
	C. sororia (Williston, 1896)						X		
Diathoneura	*D. flavolineata* Duda, 1927	X					X		
Drosophila	*D. addisoni* Pavan, 1950						X		
	D. aguape Val and Marques, 1996					X			
	D. albicans Frota-Pessoa, 1954	X				X	X		
	D. aldrichi Patterson, 1940	X				X	X		
	D. ananassae Doleschall, 1858	X				X	X		
	D. annulimana Duda, 1927					X	X		
	D. annularis Sturtevant, 1916	X					X		
	D. annulosa Vilela and Bachli, 1990					X	X		
	D. antonietae Tidon-Sklorz and Sene, 2001	X						X	
	D. araicas Pavan and Nacrur, 1950						X		
	D. arapuan Cunha and Pavan, 1947						X		
	D. ararama Pavan and Cunha, 1947					X	X		
	D. argenteifrons Wheeler, 1954						X		
	D. atrata Burla and Pavan, 1953						X		
	D. austrosaltans Spassky, 1957						X		
	D. bedichecki Heed and Russel, 1971						X		
	D. bipunctata Patterson and Mainland, 1943					X			
	D. bromeliae Sturtevant, 1921	X					X		
	D. calloptera Schiner, 1867	X	X				X		
	D. camargoi Dobzhansky and Pavan, 1950	X		X		X	X		
	D. canalinea Patterson and Mainland, 1944						X		
	D. caponei Pavan and Cunha, 1947					X	X		
	D. capricorni Dobzhansky and Pavan, 1943	X					X		
	D. cardini Sturtevant, 1916	X				X	X	X	
	D. cardinoides Dobzhansky and Pavan, 1943			X			X	X	X
	D. caxiuana Gottschalk, Martins, Praxedes and Medeiros, 2012						X		
	D. coffeata Williston, 1896					X	X		

D. converga Heed and Wheeler, 1957						X		
D. cuaso Vilela and Ratcov, 1999		X				X		
D. dacunhai Mourão and Bicudo, 1967				X		X		
D. davidgrimaldii Vilela and Bächli, 1990						X		
D. decemseriata Hendel, 1936						X		
D. eleonorae Tosi, Martins, Vilela and Pereira, 1990	X					X		
D. ellisoni Vilela, 1982	X	X				X		
D. equinoxialis Dobzhansky, 1946		X			X	X	X	
D. fasciola Williston, 1896						X		
D. fascioloides Dobzhansky and Pavan, 1943						X		
D. flexa Loew, 1866	X					X		
D. freiremaiai Vilela and Bächli, 2000								
D. frotapessoai Vilela and Bächli, 1990	X					X		
D. fulvimacula Patterson and Mainland, 1943		X			X	X		
D. fulvimaculoides Wasserman and Wilson, 1957	X				X			
D. fumipennis Duda, 1925		X			X	X		X
D. fuscolineata Duda, 1925						X		
D. griseolineata Duda, 1927						X		
D. guaraja King, 1947	X						X	
D. guarani Dobzhansky and Pavan, 1943		X					X	
D. hendeli Vilela and Bächli, 1990						X		
D. hydei Sturtevant, 1921	X					X		
D. impudica Duda, 1927					X	X		
D. ivai Vilela, 1983					X	X		
D. kikkawai Burla, 1954						X		
D. limensis Pavan and Patterson, 1947						X		
D. lutzii Sturtevant, 1916						X		
D. magalhaesi Mourão and Bicudo, 1967	X					X		
D. malerkotliana Parshad and Paika, 1964		X			X	X	X	
D. mapiriensis Vilela and Bächli, 1990					X	X		
D. mediocris Frota-Pessoa, 1954						X		
D. medioimpressa Frota-Pessoa, 1954						X		
D. mediopicta Frota-Pessoa, 1954						X		
D. mediosignata Dobzhansky and Pavan, 1943	X					X		
D. mediostriata Duda, 1925	X		X		X	X		X
D. melanogaster Meigen, 1830				X	X	X		
D. melina Wheeler, 1962	X					X		
D. mercatorum Patterson and Wheeler, 1942						X	X	
D. mesostigma Frota-Pessoa, 1954					X	X		

	D. milleri Magalhães, 1962	X				X		
	D. moju Pavan, 1950				X	X		
	D. mojuoides Wasserman, 1962					X		
	D. nasuta Lamb,1914	X				X		
	D. nebulosa Sturtevant, 1915	X	X	X	X	X	X	X
	D. neocardini Streisinger, 1945		X	X		X		
	D. neochracea Wheeler, 1959	X				X		
	D. neocordata Magalhães, 1955					X		
	D. neoelliptica Pavan and Magalhães, 1950					X		
	D. neoguaramunu Frydenberg, 1956					X		
	D. neomorpha Heed and Wheeler, 1957	X			X	X		
	D. nigricruria Patterson and Mainland in Patterson, 1943:136					X		
	D. ornatifrons Duda,1927					X		
	D. pallidipennis Dobzhansky and Pavan, 1943					X		
	D. papei Bächli and Vilela, 2002					X		
	D. paraguayensis Duda, 1927					X		
	D. parannularis Vilela and Bächli, 1990					X		
	D. paramediostriata Townsend and Wheeler, 1955					X		
	D. paranaensis Barros,1950	X			X	X		
	D. parasaltans Magalhães, 1956	X				X		
	D. parthenogenetica Stalker, 1952	X	X		X	X		
	D. paulistorum Dobzhansky and Pavan, 1949	X			X	X		
	D. pellewae Pipkin and Heed, 1964		X			X		
	D. peruviana Duda, 1927		X			X		
	D. pictilis Wasserman, 1962					X		
	D. pictura Wasserman, 1962					X		
	D. polymorpha Dobzhansky and Pavan, 1943	X	X	X	X	X		X
	D. prosaltans Duda, 1926	X			X	X	X	
	D. pseudosaltans Magalhães,1956					X		
	D. quadrum (Wiedemann, 1830)					X		
	D. querubimae Vilela, 1983				X	X		
	D. repleta Wollaston, 1858	X				X		
	D. roehrae Pipkin and Heed, 1964					X		
	D. saltans Sturtevant, 1916	X			X	X		
	D. septentriosaltans Magalhães and Buck, em Magalhães, 1962					X		
	D. setula Heed and Wheeler, 1956					X		
	D. simulans Sturtevant, 1919	X			X	X	X	
	D. speciosa Silva and Martins, 2004		X			X		
	D. sturtevanti Duda, 1927	X	X	X	X	X	X	X
	D. subsaltans Magalhães, 1955					X	X	
	D. trapeza Heed and Wheeler, 1957		X			X		
	D. tropicalis Burla and Cunha, 1950	X		X	X	X		
	D. tuchaua Pavan, 1950	X	X		X	X		
	D. willistoni Sturtevant, 1916	X			X	X	X	
Hirtodrosophila	*H. clypeata* (Wheeler, 1968)					X		
	H. gilva Burla, 1956						X	

	H. jukyry Junges, Robe and Gottschalk, 2016						X	X	
	H. minuscula Vilela and Bächli, 2005						X		
	H. morgani (Mourão, Gallo and Bicudo, 1967)					X	X	X	
	H. nungara Junges, Robe and Gottschalk, 2016						X	X	
	H. pictiventris (Duda, 1925)						X		
	H. rondonia Junges, Robe and Gottschalk, 2016						X	X	
	H. subflavohalterata (Burla, 1956)						X	X	
	H. thoracis Junges, Robe and Gottschalk, 2016						X		
Mycodrosophila	*M. amazônica* Corrêa and Gottschalk and Carvalho-Filho and Mendes and Valente, 2021						X	X	X
	M. brunnescens Wheeler and Takada, 1963						X		
	M. elegans Wheeler and Takada, 1963						X	X	
	M. hofmanni Junges, Gottschalk, Loreto and Robe, 2016							X	
	M. martinsae Corrêa and Gottschalk and Carvalho-Filho and Mendes and Valente, 2021						X		
	M. neoprojectans Wheeler and Takada, 1963						X		
	M. projectans (Sturtevant, 1916)					X	X		
	M. pseudoprojectans Wheeler and Takada, 1963						X		
Neotanygastrella	*N. chymomyzoides* Duda, 1927	X				X	X		
	N. tricoloripes Duda, 1925						X		
Paraliodrosophila	*P. antennata* Wheeler, 1956	X					X	X	
Scaptodrosophila	*S. latifasciaeformis* (Duda, 1940)	X	X	X		X	X	X	
Zaprionus	*Z. indianus* Gupta, 1968		X	X	X	X	X	X	
Zygothrica	*Z. aldrichi* Sturtevant, 1920						X		
	Z. atriangula Duda, 1926						X		
	Z. bilineata (Williston, 1896)					X	X	X	
	Z. caudata (Hendel, 1913)						X		
	Z. cryptica Grimaldi, 1987						X		
	Z. dimidiata Duda, 1927						X		
	Z. dispar (Wiedemann, 1830)						X		
	Z. joeyesco Grimaldi, 1987						X	X	
	Z. laevifrons Duda, 1927						X		
	Z. mediovitta Grimaldi, 1987						X		
	Z. microeristes Grimaldi, 1987	X				X	X	X	
	Z. orbitalis (Sturtevant, 1916)						X		
	Z. paraldrichi Burla, 1956	X	X				X	X	
	Z. pilipes Hendel, 1936	X					X	X	
	Z. poeyi (Sturtevant, 1921)						X	X	
	Z. prodispar Duda,1925					X	X	X	
	Z. ptialis Burla,1956							X	
	Z. radialis Grimaldi, 1987						X		
	Z. somatia Grimaldi, 1987						X		
	Z. subcandens Burla, 1956						X		

	Z. virgatalba Burla, 1956					X		
	Z. virgatinigra Burla, 1956					X		
	Z. vittinubila Burla, 1956		X			X		
	Z. zygia Grimaldi, 1987					X	X	
Steganinae								
Leucophenga	*L. argenteofasciata* Kahl, 1917					X		
Rhinoleucophenga	*R. hesperides* Carvalho-Filho, Duarte and Gottschalk, 2019					X		
	R. jacareacanga Poppe, Valente and Gottschalk, in Poppe *et al.*, 2016					X		
	R. lopesis Malogolowkin, 1946				X			
	R. montesis Junges and Gottschalk, 2014				X			
	R. personata Malogolowkin, 1946				X			
	R. punctulata Duda, 1929					X		
	R. tangaraensis junges and Gottschalk, 2014				X			
Totals for states		48	21	11	3	51	157	35	7

The Amazonian state abbreviations used above are:—Amazonas (AM), Amapá, (AP), Acre (AC), Maranhão (MA), Mato Grosso (MG), Pará (PA) and Rondônia (RO), and Roraima (RR).

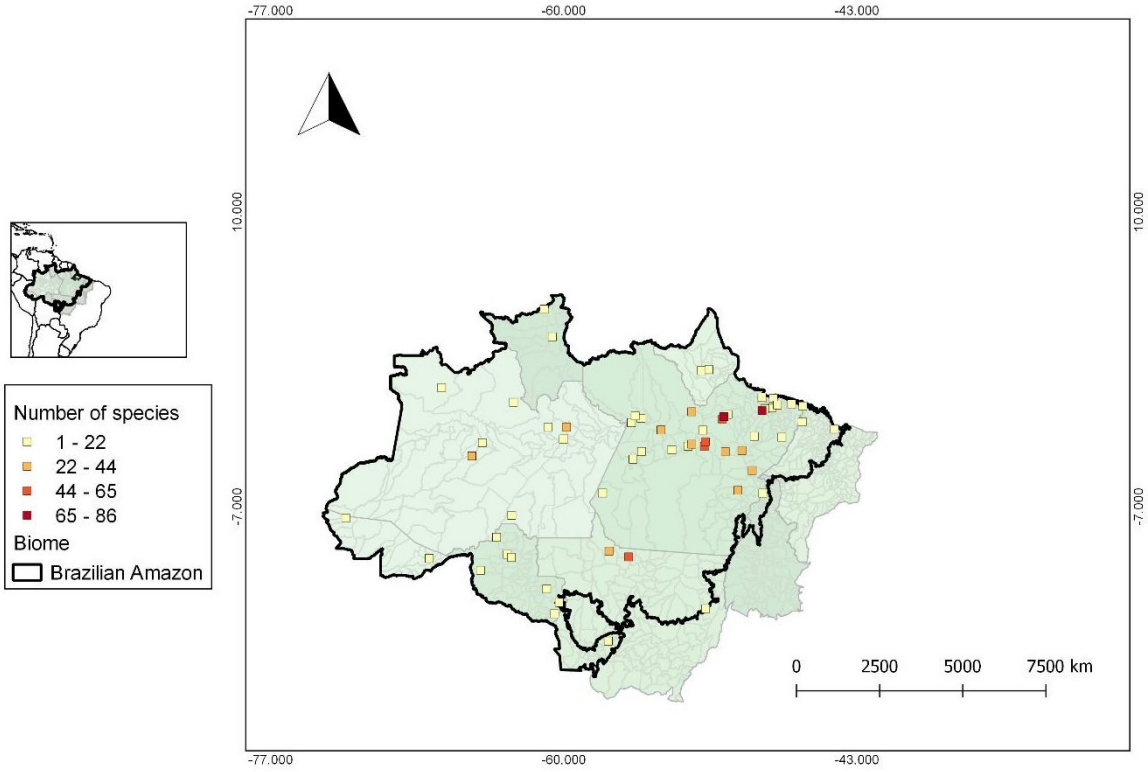


Figure 1. Localities of occurrence of species of Drosophilidae in the states that make up the Brazilian Amazon.

Most of the Northern region is inserted in the Amazon biome, which entirely encompasses the States of Amazonas, Roraima, Acre, and Amapá, almost all of the State of Pará and Rondônia, and parts of the States of Mato Grosso, Maranhão, and Tocantins (IBGE, 2019). The Brazilian Amazon biome still has a very significant original vegetation cover, although it is progressively threatened (Souza *et al*., 2020). It is the largest in territorial extension, occupying 49% of the country's area (IBGE, 2019).

Chaves and Tidon (2007) recorded 76 species of Drosophilidae for the brazilian Amazon biome. Ten years later, Santa-Brígida *et al*. (2017) recorded 122 species only for the state of Pará, using collection data records. Currently, the state of Pará has 124 recorded species, gathered in 11 genera (Santa-Brígida *et al*., 2017; Santa-Brígida *et al*., 2019).

In recent years there has been an increasing number of inventories and ecological studies with different approaches for the family Drosophilidae in the Brazilian Amazon region (Carvalho-Filho *et al*., 2019; Junges *et al*., 2019; Santa-Brígida *et al*., 2017; Furtado and Martins, 2018; Santa-Brígida *et al*., 2019; Schmitz and Valente, 2019; Medeiros *et al*., in press). However, this increase in the number of studies is still incipient when the geographic dimensions of the region are considered.

Knowing the biodiversity of the largest biome in Brazil is very important not only to contribute with new sources of data on the knowledge of the distribution and abundance of different species, but also to provide subsidies for decision-making on strategies for the conservation of this biome. The objective of this work is to provide comprehensive information on the taxonomic diversity and geographic distribution of species of the family in the Brazilian portion of the Amazon biome.

Material and Methods

Compilations of bibliographic references with records of drosophilids for the Amazon biome available on the Internet, online entomological collections databases, and gray literature (when the file was not available, the author was requested) were carried out. The search for references with records of Drosophilidae for localities in the Brazilian Amazon was performed in online repositories: ISI Web of Knowledge (www.webofknowledge.com), Google Scholar (scholar.google.com), TaxoDros (Bächli, 2021, www.taxodros.uzh.ch), and in the Brazilian platform of Curriculum Lattes, looking for Drosophilidae's researchers (to access publications in annals of events and gray literature), using the terms *Drosophila*, Drosophilidae, Drosophilid, Amazon, Amazonia, and Brazil.

The area considered in this compilation followed the Methodological Report "Biomes and coastal-marine system of Brazil: compatible with the 1:250 000 scales" published by IBGE (Brazilian Statistics Institute, 2019). The species were alphabetically ordered within genera, and, these were similarly ordered within subfamilies. For each species, all Amazonian localities organized by the state were cited, including the publication and page number of each locality recorded. The coordinates were based on data from the original publications, and in the case of species that had no record of the coordinate, it was taken as a way of organizing the data to use the coordinate of the capital or municipality when possible. The distribution of species in the biome was obtained through the preparation of a map using QGIS software.

The information was organized in a database and arranged in the order of genus, subgenus, group, subgroup, species, author, year, states, municipalities, locality, source, unpublished data, a project of the unpublished data, latitude, longitude, determiner, substrate, references, page and the total number of individuals cited in each paper.

Results and Discussion

In the present compilation, 172 species of Drosophilidae were recorded for the Brazilian Amazon, belonging to 13 genera (Table 1). These genera included representatives of the two subfamilies: Drosophilinae, with eleven, and Steganinae, with two genera. The genus with the highest species richness was *Drosophila* with 110 species, followed by *Zygothrica* and *Hirtodrosophila*, with 24 and 10 species, respectively. On the other hand, *Diathoneura*, *Leucophenga*, *Paraliodrosophila*, *Scaptodrosophila*, and *Zaprionus* had records of only one species.

The state with the highest species richness in the compilation was Pará with 157 recorded species, followed by Mato Grosso with 51 and Amazonas with 48 species (Table 1). Among the states contained in the Amazon biome, the only one without records of drosophilids was the state of Tocantins (Figure 1). The state of Tocantins has about 9% of its territory consisting of the Amazon Biome (IBGE, 2019), while the remaining 91% is represented by the Cerrado biome. Only two studies were found with species records for this state: Pavan (1950) and Magalhães (1962), both for the municipality of Palmas, Brazil, which is inserted in the Cerrado biome.

Among the Amazonian states, Pará is the state with the largest amount of research with the family Drosophilidae, including several recent studies such as: Côrrea *et al*. (2021); Santa-Brígida *et al*. (2019); Pirani and Carvalho-Filho (2019); Carvalho-Filho *et al*. (2019); Martins and Oliveira (2007); Martins and Santos (2007). In addition, it is the only state that has a species list, with 122 records (Santa-Brígida *et al*., 2017).

The collection areas for the Amazon region were highly concentrated throughout the territory (Figure 1). There is a greater sampling effort in the states of Pará and Amazonas, which have 33 and 8 municipalities sampled, respectively (Table 2). Among the municipalities, the one with the largest number of records was Melgaço with 86 species, followed by Belém with 75 and Portel with 55 (Table 3). The municipalities of Melgaço and Portel encompass the Caxiuanã National Forest, an area with an increasing number of studies, which have revealed new species for science (Gottschalk, 2012; Carvalho-Filho *et al*., 2018; Jungues *et al*., 2019; Pirani and Carvalho-Filho, 2019) and new records for both the Amazon and Brazil (Santa-Brígida *et al*., 2019; Santa-Brígida *et al*., 2017; Praxedes and Martins, 2014; Silva and Martins, 2009).

Table 2. Genera recorded in the Amazon biome, with the number of species recorded in each Amazonian state. Abbreviations: AC-Acre, AP-Amapá, AM-Amazonas, MA-Maranhão, MG-Mato Grosso, PA-Pará, RO-Rondônia, RR-Roraima and TO-Tocantins.

Táxons	AC	AP	AM	MA	MT	PA	RO	RR	Total
Drosophilinae									
Chymomyza Coquillett					1	2			3
Cladochaeta Coquillett						3			3
Diathoneura Duda						1			1
Drosophila Fallén	9	18	41	2	38	103	14	6	231
Hirtodrosophila Duda					1	9	6		16
Mycodrosophila Oldenberg					1	7	3	1	12
Neotanygastrella Duda			1		1	2			4
Paraliodrosophila Duda						1	1		2
Scaptodrosophila Duda	1	1	1		1	1	1		6
Zaprionus Coquillett	1	1	1	1	1	1	1		7
Zygothrica Wiedemann		1	4		3	23	9		40
Steganinae									
Leucophenga Mik						1			1
Rhinoleucophenga Hendel					4	3			7
Total species	11	21	48	3	51	157	35	7	333

In the state of Amazonas, the municipalities with the largest number of recorded species were: Rio Preto da Eva and Coari. Regarding Rio Preto da Eva, two important studies were conducted in reserve areas: Fazenda Esteio and Fazenda Porto Alegre, which totaled 26 species records for the locality (Martins, 1987; Martins, 2001). For Coari, all studies were conducted in the Urucu Oil Province, also known as "Urucu Base",

which represents the largest oil and natural gas reserve in Brazil and holds records of 23 species (De Toni *et al*., 2005; Furtado, 2006; Martins and Furtado, 2008) (Table 3).

The Drosophilidae species widely distributed in the Amazon region were *D. sturtevanti*, *D. nebulosa,* and *Z. indianus*, recorded in seven states. *D. sturtevanti* and *D. nebulosa,* which are native species of the neotropics, were recorded in the states of Amazonas, Acre, Amapá, Mato Grosso, Pará, Rondônia, and Roraima. *Z. indianus*, an african exotic species*,* was recorded in Amazonas, Acre, Amapá, Maranhão, Mato Grosso, Pará, and Rondônia. *S. latifasciaeformis* (another species of african origin) were recorded for the six states (Amazonas, Acre, Amapá, Mato Grosso, Pará, Rondônia) like this a native species *D. polymorpha* (Amazonas, Acre, Amapá, Mato Grosso, Pará, Roraima).

Table 3. Number of Drosophilidae species recorded for the municipalities of the Brazilian Amazon states. Abbreviations that are enclosed in parentheses: AC-Acre, AP-Amapá, AM-Amazonas, MA-Maranhão, MG-Mato Grosso, PA-Pará, RO-Rondônia, RR-Roraima and TO-Tocantins.

Municipalities	Latitude	Longitude	N° of species
Almeirim (PA)	1°31’13,9”S	52°34’53,4’’W	26
Alta Floresta (MT)	9°53’2’’S	56°14’38”W	44
Altamira (PA)	3°11’40,6’’S	52°12’33,5”W	52
Alto Paraiso (PA)	9° 44′ 43″ S	63° 17′ 0″ W	1
Apiacás (MT)	09°33’24’’S	57°22’54”W	37
Ariquemes (RO)	10°39’0’’S	64°49’60”W	1
Aveiro (PA)	3°49’27” S	55°29’35” W	8
Barcelos (AM)	0°58’31’’S	62°55’28”W	5
Belém (PA)	1°26’34”S	48°24’35”W	75
Boa vista (RR)	2°49’10’’S	60°40’17”W	3
Bragança (PA)	1°3’57,4’’S	46°47’22,2’’W	14
Brasil Novo (PA)	3°24’10,5’’S	52°34’43,9’’W	43
Breves (PA)	0°40’2,6’’S	48°30’27”W	11
Castanhal (PA)	1°17’50,4’’S	47°55’19,6’’W	3
Centro Novo do Maranhão (MA)	9°33’24’’S	57°22’54”W	2
Coari (AM)	4°5’5,6’’S	65°20’1”W	24
Colorado do Oeste (RO)	13°7’3,4’’S	60°32’28,3’’W	16
Cruzeiro do Sul (AC)	7°39’53,6’’S	72°39’1,4’’W	7
Curuçá (PA)	0°44’23,6’’S	47°51’6,8’’W	1
Ferreira Gomes (AP)	0°55’27’’N	51°35’41”W	20
Guajará-mirim (RO)	10°39’0’’S	64°49’60”W	11
Humaitá (AM)	7°30’22,2’’S	63°1’37,9”W	4
Igarapé Açu (PA)	1°7’40,4’’S	47°36’56,2’’W	11
Itaituba (PA)	2°34’32,9’’S	54°21’51,3’’W	7
Jacareacanga (PA)	14°39’5’’S	57°25’25”W	1
Juruti (PA)	2°9’11,9’’S	56°5’13,9”W	19
Manaus (AM)	3°6’25,9’’S	60°1’34”W	14
Manicoré (AM)	8°45’42,6’’S	63°54’7,2’’W	1
Marituba (PA)	3°6’25,9’’S	60°1’34”W	3
Medicilândia (PA)	3°30’43’’S	52°47’49,1’’W	1

Melgaço (PA)	1°48’17,4’’S	50°43’1,2’’W	86
Mucajaí (RR)	2°25’48,2’’S	60°55’10,6’’W	5
Nova Ipixuna (PA)	4°55’22,6’’S	49°4’18,8’’W	33
Óbidos (PA)	0°53’23’’S	52°36’8’’W	6
Oriximiná (PA)	1°27’18,1’’S	48°30’8,6’’W	1
Pacajá (PA)	3°50’7,5’’S	50°38’16,4’’W	23
Pacaraima (RR)	4° 25′ 1″ N	61° 8′ 27″ W	1
Paragominas (PA)	2°59’50,6’’S	47°21’12,6’’W	2
Parauapebas (PA)	6°4’14,7’’S	49°54’15,5’’W	29
Pimenta Bueno (RO)	11°43’0’’S	61°7’60’’E	6
Portel (PA)	1°56’32,2’’S	50°48’32,8’’W	55
Porto velho (RO)	8°45’42,6’’S	63°54’7,2’’W	13
Ribeirão Cascalheira (MT)	2°34’32,9’’S	54°21’51,3’’W	3
Rio Branco (AC)	9°58’26,4’’S	67°48’7,4’’W	9
Rio Preto da Eva (AM)	2°22’0,1’’S	59°57’0’’W	26
Salvaterra (PA)	0°45’32,4’’S	48°30’43,9’’W	3
Santarém (PA)	2°34’32,9’’S	54°21’51,3’’W	29
São Gabriel da Cachoeira (AM)	0°7′48″ N	67°5′20″ W	11
São Geraldo de Araguaia (PA)	1°17’50,4’’S	47°55’19,6’’W	1
São Luís (MA)	3°4’58,3’’S	59°57’45,6’’W	2
Senador José Porfírio (PA)	2°34’51,6’’S	51°56’13,6’’W	1
Serra do Navio (AP)	13°7’3,4’’S	60°32’28,3’’W	1
Soure (PA)	0°40’2,6’’S	48°30’27’’W	5
Tailândia (PA)	2°56’50’’S	48°57’11’’W	1
Tangará da Serra (MT)	14°37’8’’S	57°29’9’’W	5
Tefé (AM)	3°19’14,6’’S	64°43’25’’W	3
Tucuruí (PA)	3°46’10,3’’S	49°40’26,8’’W	43
Uruará (PA)	3°43’26,6’’S	53°44’7,8’’W	1
Vilhena (RO)	12°28’60’’S	60°16’0’’W	8
Vitória do Xingu (PA)	3°15' 02"S	51° 47' 51" W	51
Viseu (PA)	1°11’36,7’’S	46°8’22,2’’W	1

Among the 172 species recorded, 99 (57%) were recorded in a single state (Table 1). Eighty-five species were recorded only in Pará, followed by Mato Grosso with five, Rondônia and Amazonas with four and one species, respectively. However, these records often represent collections made in expeditions in which the focus of the study was not Drosophilidae, but the material was available for study by experts in the entomological collections of different research institutions in the region.

Many species were recorded decades ago. Some of these, more than 20 years ago, as is the case of: *D. addissoni* (Vilela, 1982), *D. argenteifrons* (Wheeler, 1954), *D. bedichecki* (Heed and Russel, 1971), *D. freiremaiai* (Vilela, 2000), *D. guarani* (Martins, 2001), *D. fascioloides* (Wasserman, 1962), *D. fuscolineata* (Dobzhansky and Pavan, 1950), *D. mediosignata* (Martins, 1996), *D. mojuoides* (Vilela, 1983), *D. nigricruria* (Dobzhansky and Pavan, 1950), *D. pallidipennis* (Dobzhansky and Pavan, 1950), *D. peruviana* (Pavan, 1959), *H. thoracis* (Wheeler, 1954), *Z. aldrichi* (Hendel, 1936), *Z. laevifrons* (Hendel, 1936*), Z. mediovitta* (Grimaldi, 1987), *Z. somatia* (Grimaldi,1987), *R. jacareacanga* (Poppe *et al.*, 1996), and *R. lopesi* (Poppe *et al*., 1996). There is also a century-old record of a species: *Leucophenga argentofasciata*, recorded 104 years ago (Khal, 1917), for the municipality of Santarém in Pará state, which was not recorded again after that date.

Drosophila was the only genus recorded in all states analyzed, followed by *Zaprionus* and *Scaptodrosophila* in seven and six states, respectively. In Pará, 103 species of *Drosophila* were recorded, followed by Amazonas and Mato Grosso, with 41 and 38, respectively (Table 2). In contrast, *Chymomyza*, *Cladochaeta*, *Diathoneura*, *Leucophenga*, *Paraliodrosophila*, and *Rhinoleucophenga* were recorded in at most two states and with few species. In addition, the state of Roraima only had records of species of the genus *Drosophila* and *Mycodrosophila.* The latter is represented by the species *Mycodrosophila amazonica*, which was recently described by Côrrea *et al*. (2021).

The genus *Drosophila* is one of the richest in species within the Drosophilidae family, with about 1,674 species in the world (Taxodros, 2021), many of these with a cosmopolitan distribution. Several species in the family Drosophilidae stand out as model organisms. Among these, the best known is *Drosophila melanogaster* with numerous researches conducted in different areas, especially in genetics and evolution. Among the species present in the Amazon, many are used in high impact studies on adaptation, speciation, invasion biology, competition, among others (Martins, 1996; Magellan, 1956, 1962; Dobzhansky, 1950; Gottschalk *et al*., 2012; Blauth *et al*., 2013; Monteiro, 2018). Including some species native to the Amazon, such as the willistoni complex and species of the Cardini and Repleta group, for example (Malogolowkin, 1952; Wasserman, 1962; Vilela *et al*., 1983; Tosi *et al*., 1990; Heed and Russel, 1971; Monteiro *et al*., 2014).

Another relevant aspect to be raised regarding the results of this compilation concerns the fact that most of the collections in the surveys listed here were conducted using banana baits (Martins, 1987; Martins, 2001; Blauth and Gottschalk, 2007; Martins and Oliveira, 2007; Junges *et al*., 2011; Praxedes and Martins, 2014; Poppe *et al*., 2017). This bait attracts mainly frugivorous species, as is the case of most *Drosophila* species recorded in the Amazon biome. It is likely that the knowledge about the richness of drosophilids for the Amazon should be expanded if there is greater diversification of collection baits, or even an increase of collections in the natural substrates of feeding, reproduction, and/or oviposition of drosophilids. Schmitz and Valente (2019) recorded 28 species of Drosophilidae that breed on flowers, after sampling 56 species of plants in eventual collections across the territory of Brazil and Santa Brígida *et al*. (2019) collected 55 species of Drosophilidae in fungal fruiting bodies in collections made in Caxiuanã National Forest located in the municipalities of Melgaço and Portel, having had the first records of *Mycodrosophila neoprojectans* and *M. pseudoprojectans* for Brazil. These studies indicate how high the number of records of drosophilid species can be when, in addition to the use of attractive baits, natural substrates are used.

Another interesting issue is that many specimens are identified as morphotypes, which may be new species or when it was not possible to reach the specific level. Also, in the work conducted by Santa-Brígida *et al*. (2019), 56% were morphospecies that may represent new species.

Regarding exotic species distributed in the Amazon region, nine species have been recorded in seven states: *D. ananassae, D. hydei, D. kikkawai, D. malerkotliana, D. melanogaster, D. nasuta, D. simulans, S. latifasciaeformis*, and *Z. indianus*. Among the states, Pará stands out with records for all species. Of the eight Amazonian states with records of Drosophilidae, Roraima was the only state without records of an exotic species. The species with the highest number of records was *Z. indianus* present in seven states, followed by *S. latifasciaeformi*s in six and *D. malerkotliana* in five states. *D. hydei, D. kikkawai,* and *D. nasuta* were recorded only in Pará state (Table 4).

Z. indianus is a species of Afrotropical origin, commonly inhabiting Australian, Eastern, and Palaearctic regions (Tsacas *et al*., 1981), which was first detected in Brazil by Vilela (1999), in São Paulo. Since then, the expansion of the geographical distribution of this species across the country has been recorded (Tidon, 2003; Leão and Tidon, 2004; Barbosa *et al*., 2012; Commar *et al*., 2012), which has also successfully established itself in the Brazilian Amazon (Tidon *et al*., 2003; David *et al*., 2006; Amador *et al*., 2011; Vasconcelos *et al*., 2017), taking advantage of the growing deforestation in the region.

Another interesting record for the Amazon biome was that of *Drosophila nasuta*. *D. nasuta* is a species native to Asia, whose first record as an invader in Brazil was reported in 2015 for the state of São Paulo (Vilela and Goñi, 2015). This species showed a rapid expansion, since it has been reported in different Brazilian biomes such as Atlantic Forest (Vilela and Goñi, 2015; Batista *et al*., 2016; Silva *et al*., 2020), Cerrado (Deus and Roque, 2016; Leão *et al*., 2017) and Caatinga (Montes *et al*., 2021). In the Amazon region, the species was recorded in 2017 in Pará state in five municipalities, but with low abundance in all collection sites (Medeiros *et al*., 2018; Medeiros *et al*., in press).

Table 4. Distribution of exotic species of Drosophilidae in the Brazilian Amazon. Abbreviations: AC-Acre, AP-Amapá, AM-Amazonas, MA-Maranhão, MG-Mato Grosso, PA-Pará, RO-Rondônia, RR-Roraima and TO-Tocantins. An X indicate the species was found in the state.

Taxon	AM	AP	AC	MA	MT	PA	RO	RR
D. ananassae	X				X	X		
D. hydei						X		
D. kikkawai						X		
D. malerkotliana	X	X			X	X	X	
D. melanogaster	X			X	X	X		
D. nasuta						X		
D. simulans	X				X	X	X	
S. latifasciaeformis	X	X	X		X	X	X	
Z. indianus	X	X	X	X	X	X	X	

Conclusions

The present compilation recorded 172 species present in the Brazilian Amazon, adding 96 species to the last publication. In the state of Pará, 35 more species have been recorded since the last survey. Most of the records for Pará were for the municipality of Melgaço that covers the Caxiuanã National Forest (with at least 65 records), which represents one of the largest conservation units in the state. This and other records of species in protected areas highlight the importance of these areas for the conservation of the biome. Considering the comprehensiveness of the geographic coverage of the species recorded in this study, it is clear that the Brazilian Amazon region still lacks further studies related to the fauna of Drosophilidae. The Amazon biome occupies about 4,212.742 km², so a much higher species richness is expected with the increase in Drosophilidae species inventories in the region. The list presented here can be further expanded with the diversification of the collection methodologies and the collection on natural feeding and/or oviposition substrates for the drosophilids, as well as with the increase of the taxonomic effort in the identification of the drosophilids. However, the absence of recent collections of several species with historical records can also mean a warning about the reduction in the diversity of drosophilids due to the loss of forest cover. The states of Maranhão, Roraima, and the Tocantins should be seen as a priority for new inventories due to the absence and/or a low number of records of drosophilid species.

Acknowledgments: We would like to thank the Emílio Goeldi Museum (Museu Emílio Goeldi, MPEG) and the National Council for Scientific and Technological Development (Conselho Nacional de Desenvolvimento Científico e Tecnológico, CNPq) for granting a scholarship (process nº 162634 / 2020-0).

References: Amador, R.B., M.B. Martins, and L.S. Furtado 2011, Dros. Inf. Serv. 94: 73; Bächli, G., 2021, Electronic. Database; Barbosa, M.R.R., G. Graciolli, and F. Paiva 2012, Check List 8: 175; Batista, M.R.D., M.T. Brianti, C.A.C. Andrade, and L.B. Klaczko 2016, Dros. Inf. Serv. 99: 44; Blauth, M.L., D.P. Lima, D.S. Araújo, J.B. Moysés, and M.S. Gottschalk 2013, Dros. Inf. Serv. 96: 58-60; Bolzan, A.R., 2011, DNA Barcode de drosofilideos micófagos pertencentes aos gênero Hirtodrosophila, Mycodrosophila, Zygothrica. Universidade Federal de Santa Maria, Santa Maria, 85 pp.; Carvalho-Filho, F.S., G. Pirani, and T.G. Kloss 2018, Zootaxa 4410: 83–496; Carvalho-Filho, F., da Silva, L.B. Duarte, and M.S. Gottschalk 2019, Zootaxa 4551: 583–592; Chaves, N.B., and R. Tidon 2007, V Simpósio de Ecologia, Genética e Evolução de Drosophila; Chaves, N.B., and R. Tidon 2008, Revista Brasileira de Entomologia 52: 340-348; Commar, L.S., L.G.C. Galego, C.R. Ceron, and C.M.A. Carareto 2012, Genetics and Molecular Biology 35: 395–406; David, J.R., L.O. Araripe, B.C. Bitner-Mathé, P. Capy, B. Goñi, L.B. Klaczko, H. Legout, M. Martin, J. Vouidibio, A. Yassin, and B. Moreteau 2006, Heredity 96: 53–62; Deus, P.H.M., and F. Roque

2016, Dros. Inf. Serv. 99: 44–47; Dobzhansky, T., and C. Pavan 1950 Journal of Animal Ecology 19: 1-14; De Toni D.C., J.A. Brisson, P.R.P. Hofmann, M. Martins, and H. Hollocher 2005, Dros. Inf. Serv. 88: 33-38; Gottschalk, M.S., P.R. Hofmann, and V.L.S. Valente 2008, Check List 4: 485; Gottschalk, M.S., M.B. Martins, C.L.B. Praxedes, and H.F. Medeiros, 2012, Revista Brasileira de Entomologia 56: 431–435; Grimaldi, D.A., 1987, Bulletin of the American Museum of Natural History 186: 103-268; Heed, W.B., and J.S. Russell 1971, The University of Texas Publication 7103: 91-130; Hendel, F., 1936, Annalen des Kaiserlich-koniglichen Naturhistorischen Hof-Museums 47: 61-106; IBGE (ED.), 2019, Biomas e sistema costeiro-marinho do Brasil: compatível com a escala 1:250 000. Rio de Janeiro: IBGE; Junges, J., M.L. Blauth, and M.S. Gottschalk 2011, Dros. Inf. Serv. 94: 40-42; Junges, J., L.J. Robe, and M.S. Gottschalk 2019, Zootaxa 4567: 276; Kahl, H., 1917, Annals of Carnegie Museum 11: 364–393; Leão, B.F.D., and R. Tidon 2004, Annales de la Société entomologique de France 40: 285–290; Leão, B.F.D., F. Roque, P.H.M. Deus, and R. Tidon 2017, Dros. Inf. Serv. 100: 65-69; Magalhães, L.E., 1962, The University of Texas Publication 6205: 135-154; Martins, M., 1987, Boletim do Museu Paraense Emílio Goeldi, série Zoologia 3: 195-218; Markow, T.A., and P. O'Grady 2008, Functional Ecology 22: 747–759; Martins, M.B., 1996, Drosófila e outros insetos associados a frutos Parahancornia amapa dispersos sobre o solo da floresta. Universidade Estadual de Campinas; Martins, M., 2001, The Ecology and Conservation of a Fragmented Forest, Yale University Press. p. 175-186; Martins, M.B., and L. Oliveira 2007, Mocambo: Diversidade e Dinâmica Biológica da Área de Pesquisa Ecológica do Guamá (APEG), Museu Paraense Emílio Goeldi, p.389–404; Martins, M.B., and R.C. Santos 2007, Mocambo: Diversidade e Dinâmica Biológica da Área de Pesquisa Ecológica do Guamá (APEG), Museu Paraense Emílio Goeldi, p. 315–329; Martins, M.B., and I.S. Furtado 2008, Contribuição ao diagnóstico do estado de conservação das clareiras da base do Urucu através do estudo de insetos saprófitos; Mata, R.A., M.A. McGeoch, and R. Tidon 2008, Biodiversity and Conservation 17: 2899-2916; Mata, R.A., M.A. McGeoch, and R. Tidon 2010, Natureza and Conservação 08: 60–65; Leão, B.F.D., and R. Tidon 2004, Annales de la Société entomologique de France (N.S.) 40: 285–290; Monteiro, M., A. Caçador, C. Praxedes, C. Miranda, S. Silva, E. Almeida, and H. Medeiros 2018, Anais do Congresso Brasileiro e Latino-Americano de entomologia, p. 735; Montes, M.A., C.H. Neves, A.F. Ferreira, M.F. Santos, J.I. Quintas, G. Manetta, P.V. Oliveira, and A.S. Garcia 2021, Neotropical Entomology 50: 571–578; Parsons, P.A., 1991, Global Ecology and Biogeography Letters 1: 77-83; Parsons, P.A., 1995, Biological Conservation 74: 21-27; Pavan, C., 1950, Boletim da Faculdade de Filosofia, Ciências e Letras, Universidade de São Paulo 111: 3-36; Pavan, C., 1959, Boletim da Faculdade de Filosofia, Ciências e Letras, Universidade de São Paulo 11: 1-81; Pirani, G., and F.S. Carvalho-Filho 2019, Journal of Natural History 53: 1301–1312; Poppe, J.L., 1996, Drosófila e outros insetos associados a frutos Parahancornia amapa dispersos sobre o solo da floresta. Universidade Estadual de Campinas; Poppe, J.L., M. Deprá, H.J. Schmitz, and V.L. Valente 2017, Entomological Science 20: 245–254; Praxedes, C.L.B., and M.B. Martins 2014, Caxiuanã: Desafios para Conservação de uma Floresta Nacional na Amazônia, Museu Paraense Emílio Goeldi, p. 521–537; Santa-Brígida, R., H.J. Schmitz, and M.B. Martins 2017, Biota Neotropica 17: 1; Santa-Brígida, R., F. Wartchow, R.S. Medeiros, M.S. Gottschalk, M.B. Martins, and C.J.B. de Carvalho 2019, Papéis Avulsos de Zoologia 59: 1-10; Silva, D.G., H.J. Schmitz, H.F. Medeiros, C. Rohde, M.A. Montes, and A.C. Garcia 2020, Journal of Insect Conservation 24: 525–534; Schmitz, H.J., and V.L. Valente 2019, Papéis Avulsos de Zoologia 59: 15; Souza, C.M., *et al*., 2020, Remote Sensing 12: 27-35; Tidon, R., D.F. Leite, and B.F.D. Leão 2003, Biological Conservation 112: 299-305; Tidon, R., and J.M.D. Almeida 2016, Zootaxa, 4122: 719; Tidon, R., M.S. Gottschalk, H.J. Schmitz, and M.B. Martins 2021, Drosophilidae in Catálogo Taxonômico da Fauna do Brasil. PNUD; Throckmorton, L.H., 1975, *Handbook of Genetics*, Plenum Press, pp. 421–469; Tosi, D., M. Martins, C.R. Vilela, and M.R. Pereira 1990, Brazilian Journal of Genetics 13: 19-31; Tosi, D., and M.A. Pereira 1993, Brazilian Journal of Genetics 16: 321-333; Tsacas, L., D. Lachaise, and J.R. David 1981, *In*: *The Genetics and Biology of* Drosophila, (Ashburner, M., H.L. Carson, and J.N. Thompson Jr, eds.). volume 3a. Academic Press: New York, p. 197–259; Vasconcelos, A., T. Azevedo, and R.S. Santos 2017, EntomoBrasilis 10: 60–63; Vilela, C.R., and G. Bächli 1990, Mitteilungen der Schweizerischen Entomologischen Gesellschalf 63: 1-332; Vilela, F.C., 1982, Papéis Avulsos de Zoologia, 34: 309-347; Vilela, C.R., 1983, Revista Brasileira de Entomologia 27: 1-114; Vilela, C.R., 1999, Dros. Inf. Serv. 82: 37–39; Vilela, C.R., and G. Bächli 2000, Mitteilungen der Schweizerischen Entomologischen Gesellschalf 73: 49-65; Vilela, C.R., and B. Goñi 2015,

Revista Brasileira de Entomologia 59: 346–350; Wasserman, M., 1962, Univ. Texas Publ. 6205: 119-134; Wheeler, M.R., 1954, Univ. Texas Publ. 5422: 47-64.

Maternal age effect emerges as genotype-by-environment interaction for larval olfactory learning in *Drosophila melanogaster.*

Çömlekçi, Yiğit Ege, Gizem Tunçer, Çınar Ege Bakırcı, Seda Yirmibeş, Halide Cansu, Gökçe Çakmak, İlayda Dumlupınar, Müjgan Armağan Türközü, Gülnur İpek, Cansu Yörgüç, Memet Gözüböyük, Cansu Aydemir, Murat Yılmaz, and Ergi Deniz Özsoy*. Functional and Evolutionary Genetics Laboratory (FEGL), Department of Biology, Hacettepe University, 06800 Beytepe, Ankara, Turkey; *For correspondence: ergideniz.ozsoy@hacettepe.edu.tr

Introduction

Larval olfactory learning is an important trait for food tracking and location in insects. *Drosophila melanogaster* is an efficient model for elucidating the components of fitness in olfaction in which the levels and the contribution of genetic variation are at the base (Richgels and Rollmann, 2012; Gunjan *et al*., 2015). Maternal age effects on the life history traits are conventionally looked upon as purely environmental, irrespective of the amount of the genetic contribution (Rattan and Singh, 2009). Here in our study we have tested this hypothesis with a genomically well defined set of *Drosophila melanogaster* inbred lines (Drosophila Genetic Reference Panel, DGRP; Mackay *et al*., 2012; Huang *et al*., 2014.) in an maternal age structured experimantal design for larval olfactory learning. We show that maternal age considerably affects the performance of the larvae for olfaction learning, and this age dependency has a highly significant component explained by the response of different DGRP lines at specific ages. This prominent genotype-by-environment (as age) interaction emergent even with a small subset of DGRP lines highly suggests that maternal age effects may have a strong genetic component in olfaction in *Drosophila*.

Material and Methods

Aging the flies for the experiment

In the experiments eleven DGRP lines (RAL-045, RAL-083, RAL-091, RAL-093, RAL-235, RAL-239, RAL-348, RAL-354, RAL-379, RAL-385, and RAL-774) were used which were ordered from Bloomington Drosophila Stock Center. All of the flies were raised and maintained on standard fly food medium at 25ºC and 65% humidity on a 12-hr light-dark cycle. For each line 5 male and 5 female flies were crossed in one vial. A total of 20 replicates for each line were prepared. After letting the adults lay eggs for 24 hrs, the adults were removed from the vials. On the 9th – 10th day after egg laying, all of the male and virgin female adults were collected from lines and placed into separated food vials categorized by gender and line parameters. All of the food vials of the flies were changed every 2-3 days. When the adults reached the specified age, 30 male and 30 female flies were crossed (males were 7 days of old in all the crosses) in glass bottles (~ 750 ml ~35 mm OD × ~120 mm H) with food for 24 hrs to lay eggs. After 24 hrs all of the adults were taken away. On the 5th – 6th day the L3 (3rd instar) larvae which were still feeding on the food were taken to the experiment. Total of 180 L3 (3rd instar) larvae were used per line for both odors.

Larval olfactory learning – odor preparation

Two odors were selected for the experiment, which are Methylcyclohexanol and Pentylacetate. Both of the odors were diluted with Mineral Oil to prevent vaporization. The dilution rates were determined as 1/100 for Methycyclohexanol and 1/300 for Pentylacetate.

Larval olfactory learning experiment – PA Preferance

The following regimen (modified from Apostolopoulou *et al*., 2013) was used to train the larvae to associate odors with an appetitive sugar cue. A PA odor container was placed on either side of a Petri dish containing fructose. A group of 30 feeding 3rd instar larvae were placed in the middle of the Petri dish, the lid was closed, and the larvae were exposed to PA for 5 minutes. The larvae were removed from the Petri dish and trained for 5 minutes in another Petri dish containing only agarose and a MCH odor on both sides. This process was repeated two more times. A PA and a MCH odor container were placed on opposite sites of an agarose-only test Petri dish. The trained larvae were placed in the test Petri dish and exposed to the odors for 5 minutes. Then the number of larvae on the left, middle, and right side of the test Petri dish was counted. All experiments were performed as 3 replicates per line.

Larval Olfactory Learning Experiment – Measurements

Learning performance was calculated only on the results of the PA preference index (modified from Apostolopoulou *et al*., 2013). Preference index of PA was calculated as:

PREF PA (PA+/MCH-) = (# of larvae on PA side - # of larvae on MCH side) /
# of all larvae within the left, right, and middle zones

The learning score of the lines was calculated by averaging the preference indices of three replicates of each line.

Table 1. Apetitive learning Performance ANOVA. Line term signifies the contribution of different DGRP lines *** P < 0,001.

	Df	Sum Sq.	Mean Sq.	F value	Pr (>F)	
Line	10	0.6178	0.06178	10.58	2.00E-10	***
Maternal Age	2	0.4114	0.2057	35.22	3.91E-11	***
Line*Maternal Age	20	1.2923	0.06461	11.06	3.86E-14	***

Results and Discussion

We have measured apetitive learning performance (as learning performance index) of the larvae of eleven *Drosophila* Genetic Reference Panel (DGRP) lines from the crosses with the mothers of three different age classes. First of all, both maternal age and genotype (DGRP line) made a difference in how larvae responded: learning performance seems to depend not only age of the mothers but the genotype of them as well (Table 1). This could be expected as aging affects many life history traits, olfactory behavior being no exception in that respect. And, naturaly, different genotypes should have different larval phenotypic expressions (*i.e*., learning performance indices). But the most interesting pattern emerging in our results is how genotype and maternal age jointly contribute to the level of that expression: as Table 1 shows, maternal age and genotype interact when larvae express their learning performance. This interaction has very striking implications. First, there seems to be no monotonous performance deterioration with increasing age: Larval performance level fluctuates with age in most of the lines (Table 2 and Figure 1). In more than half of the lines (RAL-045, RAL-083, RAL-091, RAL-239, RAL-379, and RAL-385; 6 out of 11), the middle age class (21 days) is the top performance age, but, in the other half (RAL-093, RAL-235, RAL-348, RAL-354, and RAL-774; 5 out of 11), it is not. Again, for the youngest age class (7 days), while five lines (RAL-083, RAL-091, RAL-093, RAL-235, and RAL239) have lowest performances, the other six lines have radically different performances (*e.g*., while two lines, RAL-348 and RAL-774, have top-most performance scores among all age classes, another two lines, RAL-354 and RAL-385, have scores almost equal to those of two older age classes, 21 days and 35 days, respectively). Finally, the oldest age, as seen clearly, is not the lowest performance age

all the time: in less than the half of the lines (*i.e.*, RAL-045, RAL-354, RAL-379, and RAL-774) this is the situation.

Table 2. Mean (over 3 replicates), and Standard errors of the learning performances of the larvae of DGRP lines in the study. All measurements are the averages of three replicates per line and age of third instar larvae which were used in the experiments. SE (Standard Error of the Mean). CI (95% Confidence interval of the Mean). N (number of the replicates per replicate having 30 larvae).

Line	Maternal Age	N	Mean Learning Performance Index	SE	CI
RAL-045	07-days	3	0.286	0.027	0.115
RAL-045	21-days	3	0.409	0.014	0.059
RAL-045	35-days	3	0.166	0.108	0.466
RAL-083	07-days	3	0.340	0.030	0.131
RAL-083	21-days	3	0.600	0.019	0.083
RAL-083	35-days	3	0.396	0.004	0.018
RAL-091	07-days	3	0.004	0.067	0.289
RAL-091	21-days	3	0.500	0.083	0.360
RAL-091	35-days	3	0.308	0.022	0.094
RAL-093	07-days	3	0.333	0.100	0.432
RAL-093	21-days	3	0.527	0.074	0.322
RAL-093	35-days	3	0.616	0.017	0.073
RAL-235	07-days	3	0.022	0.029	0.126
RAL-235	21-days	3	0.299	0.029	0.127
RAL-235	35-days	3	0.483	0.019	0.083
RAL-239	07-days	3	0.155	0.020	0.086
RAL-239	21-days	3	0.584	0.025	0.106
RAL-239	35-days	3	0.300	0.048	0.207
RAL-348	07-days	3	0.478	0.006	0.025
RAL-348	21-days	3	0.356	0.017	0.074
RAL-348	35-days	3	0.422	0.011	0.048
RAL-354	07-days	3	0.569	0.035	0.151
RAL-354	21-days	3	0.569	0.009	0.042
RAL-354	35-days	3	0.316	0.035	0.149
RAL-379	07-days	3	0.291	0.038	0.162
RAL-379	21-days	3	0.388	0.095	0.408
RAL-379	35-days	3	0.267	0.044	0.189
RAL-385	07-days	3	0.380	0.018	0.078
RAL-385	21-days	3	0.488	0.022	0.096
RAL-385	35-days	3	0.377	0.022	0.096
RAL-774	07-days	3	0.567	0.001	0.001
RAL-774	21-days	3	0.377	0.012	0.051
RAL-774	35-days	3	0.205	0.006	0.024

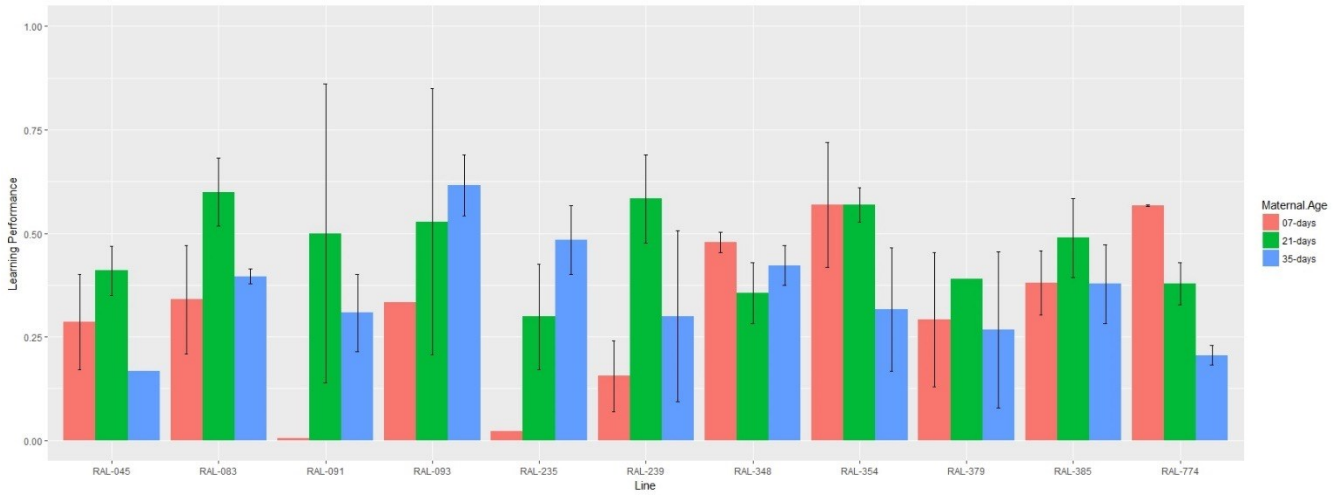


Figure 1. Change of the Mean (with 95% CI) larval performance indices among lines through maternal age.

Second (which is a natural inference from the first), each genotype (DGRP line) has almost its own expression pattern of larval learning performance phenotype and this, again, has no direct dependence of increasing age (Figure 1, note differences in magnitude and direction of the performance index across ages).

Overall, our results clearly indicate that genetic variation among the lines affects the course of larval learning performance of the offspring from differently aged females. Therefore, maternal age effects on the larval apetitive performance seems to be dependent also on the genotype of the aged mothers.

References: Apostolopoulou, A.A,. A. Widmann, A. Rohwedder, J.E. Pfitzenmaier, and A.S. Thum 2013, J. Vis. Exp. 72: e4334. doi:10.3791/4334; Gunjan, H.A., M.M. Magwire, W. Huang, L.S.-N. Yazmin, T.F.C. Mackay, and R.R.H. Anholt 2015, Chemical Senses 40: 233–243; Huang, W., A. Massouras, Y. Inoue, J. Peiffer, and M. Ra`mia 2014, Genome Res. 24: 1193–1208; Mackay, T.F.C., S. Richards, E.A. Stone, A. Barbadilla, J.F. Ayroles, *et al.*, 2012, Nature 482: 173–178; Rattan, S.I.S., and R. Singh 2009, Gene Therapy 16: 3–9. 2009; Richgels, P.K., and S.M. Rollmann 2012, Chemical Senses. 37: 229–240. 2012.

The mitochondrial phylogeny of the *Drosophila immigrans* species group: Species misidentification and dramatic introgression events.

Waddell, Peter J.[1*], Jeremy Wang[2], Richard Winkworth[1] and Corbin D. Jones[3].
[1]School of Fundamental Sciences, Massey University, Palmerston North, New Zealand [pwaddell.new@gmail.com]; [2]Department of Genetics, University of North Carolina, Chapel Hill, North Carolina; [3]Department of Biological Sciences, University of North Carolina, Chapel Hill, North Carolina.

Abstract

Here we assemble the mtDNA genomic sequences of 28 stocks of the *Drosophila immigrans* species group. During the mtDNA assembly process from short reads a number of extraordinary features were observed. Despite nuclear insertions of mitochondrial sequences (numt's) being extremely rare in *Drosophila*, some stocks of *D. sulfurigaster albostrigata* have expanded a large numt to occupy nearly 1% of the nuclear

genome. A reconciliation of the mitochondrial phylogeny with a nuclear phylogeny shows that there have been multiple mtDNA from the *albomicans/nasuta* clade replacing the original mtDNA of this subspecies. The sister taxa to *D. s. albostrigata* is a clade of *D. pulaua*, *D. s. sulfurigaster* and *D. s. bilimbata*. The first two taxa have apparently introgressed with the quite distinct *D. kohkoa* lineage, also resulting in mtDNA replacement, possibly an ongoing process with different *D. s. sulfurigaster* stocks showing quite different mtDNA origins. All *sulfurigaster* stocks show evidence of a second distinct, apparently full length mtDNA, that matches the DNA of no other stock held. The mitochondrial history of the *immigrans* species group together with analyses of nuclear datasets is making clear the phylogeny of the species group once the effect of a surprisingly large number of misidentified and/or misassembled sequences in Genbank are taken into account, and dramatic but very specific mtDNA introgression events are identified. However, it is also clear that strong bootstrap support and/or posterior probabilities are not sufficient to have confidence that parts of the phylogeny are correctly resolved. That is, tree models leave significant conflicting signals, as seen in network analyses such as NeighborNet, unexplained. By balancing the information from the different types of analysis, a reasonable phylogeny of the *immigrans* group and the whole *nasuta* subgroup is proposed.

Introduction

The mtDNA in most animals is remarkably small, yet encodes an essential set of genes coding for proteins, rRNA, and tRNA and exclusively for use in the mitochondrion. Indeed, they are the remnants of the much larger set of mitochondrial genes that evolution has been unable to move into the nuclear genome. In most animals mtDNA is nearly fully maternally inherited in a near clonal manner, inducing an effective population size (N_e) closer to to $N_e/2$ than to the $2N_e$ characterizing autosomes (although see Nunes *et al.,* 2013). The mtDNA of most well studied groups of animals also seems to experience no recombination. Further, the mtDNA sits in a very chemically reactive environment, without recourse to the full set of error detecting and correcting proteins that facilitate nuclear DNA integrity. Until recently, the consensus view was that mtDNA is under very heavy stabilizing selection and that changes are conservative and typically close to neutral. The high rate of change, high copy number, and small population size all make mtDNA ideal for phylogenetic surveys of populations, while a much slower tick over of highly conserved regions has also found them favored for deeper phylogenetic studies (see Breton *et al.,* 2014 for a recent review).

The unusual features of mtDNA and its propensity to fix undesirable characteristics makes it a key focus in human genetic diseases, which show a complex pattern of expression between mutation in the mtDNA and phenotype (Wolff *et al.,* 2014). There is a growing expectation from theoretical and applied studies that nuclear loci are attempting to make major changes to accommodate wayward mtDNA (Russel *et al.,* 2020; Hill, 2020). These disease states, when expressed, span the spectrum from very young to very old patients, and affect a wide variety of tissue types in their own way (Wallace, 2018; Scheid *et al.,* 2021). In contrast to a fairly neutrally evolving isolated genetic unit, increasingly mtDNA is seen as being highly cointegrated and coevolved with the approximately 1500 nuclear genes essential for mtDNA function. Remarkably, despite nuclear encoded mitochondrial genes themselves also being parts of highly conserved structures, many exist in duplicate. It is argued this allows them to facilitate adaptation to mtDNA variants (mtDNA heteroplasty), which are seen as exerting a huge selective force (Wolff *et al.,* 2014).

Early studies of mtDNA, such as of mice in Europe, showed introgression far across species and subspecies boundaries, adding to the impression that having introgressed mtDNA was near neutral or even advantageous. However, more recent studies have highlighted a whole host of issues with this view. For example, killing the mtDNA of a cell and replacing it with that of closely related species, shows much reduced basic function (for example, chimp or gorilla mtDNA in humans), or even no function at all (orangutan or more distant relatives) (Hill, 2020). Another range of studies show that when different types of mtDNA do introgress between species or subspecies, tissue level expression of nuclear encoded mitochondrial genes may undergo major changes to compensate, and fitness of hybrids drops measurably (Healy and Burton, 2020, Yuting *et al.,* 2021). This also leads to the possibility that mtDNA differentiation may be setting up major deficits to outbreeding, with the severity expected to be positively correlated overall with the degree of mtDNA divergence(Nguyen *et al.,* 2020). In this way, it will potentially be a strong driver of either speciation or extinction when interfertile lineages mix (*e.g.*, Hill, 2019).

Given such a prominent role in medicine, genetics, ecology and evolution, it is not surprising that large databases such as Genbank have specific sections devoted to curating mtDNA and there are also whole peer reviewed journals from major scientific publishers expressly dedicated to presenting whole mtDNA genomes. For example, Mitochondrial DNA Part A, by Taylor and Francis, with recent examples being the first mtDNA genomes published for the *immigrans* species group (Shi *et al.,* 2015; Xiongbin *et al.,* 2016). To be appropriately useful to medical studies, mtDNA genomes need to be properly assembled, reliably assigned to species, subspecies and, where possible, where, when and how they were collected (provenance). With the rise of massively parallel short (50 to 250 bp) DNA reads, known often as Next Generation Sequencing or Massively Parallel Sequencing (MPS) replacing earlier technologies such as Sanger sequencing (~600-900 bp), this has created some uncertainty. However, recent papers have suggested that with careful use, MPS can reliably detect low levels of heterogeneity in mtDNA and offer some distinct advantages over Sanger sequencing (*e.g*., Kloss-Brandstätter *et al.,* 2015).

Clearly, genetic understating of heteroplasmy, or the co-existence of two forms, of mtDNA in medical circles can be heavily informed by species which undergo interchanges of mtDNA. While there have been many instances of introgression in the genus *Homo* there is no evidence yet that *Homo sapiens* has replaced its mtDNA (Waddell, 2018). However, it is strongly suspected our close relatives the Neanderthals introgressed with our ancestors and had their mtDNA replaced (*e.g*., Posth *et al.,* 2017), bringing unknown medical consequences.

The *nasuta* subgroup of the *immigrans* species group has long been studied for a wide range of reasons, including how closely related and interfertile many otherwise distinct taxa within it are (Wilson *et al.,* 1969; Spieth, 1969; Kitagawa *et al.,* 1982; Waddell, 1990; Kitagawa, 1992). This subgroup is, in many ways, the best "laboratory" for studying speciation in the genus. There have been a number of attempts to derive the phylogeny of the whole group (Waddell, 1990; Kitagawa, 1991; Yu and Li, 1999) and many other studies of partial phylogenies. Recently, phylogenies based on large amounts of nuclear data (Waddell *et al.,* 2019; Mai *et al.,* 2020) differ somewhat, but both agree with the partly resolved consensus phylogeny of Waddell (1990). Indeed, various mtDNA phylogenies have been generally as confusing as helpful, at least partly because researchers were not checking the species identity of the stocks they were using, so here we seek to propose a reliable mtDNA phylogeny of the whole group. The phylogeny of the *immigrans* group itself is challenging as there are 5 or 6 long branch lineages within it (Katoh *et al.,* 2007; Izumitani *et al.,* 2016).

We also examine three quite different challenges for relying on mtDNA sequences to understand molecular evolution. The first example is an egregious example of the lack of quality control in a sequence that is listed as both peer reviewed in a specialist journal and passed Genbank's quality control filters to become a reference sequence for its species. In our own MPS assemblies we examine a variety of issues and patterns we have encountered that go against the expectation that mtDNA assembly in *Drosophila* should be straight forward. These examples include additional challenges that coincide with assembling mtDNA from MPS for species that have introgressed. Lastly, we use a range of phylogenetic methods to show that high bootstrap values alone are not sufficient to have confidence in the branching order of the phylogeny of this group, but a probable phylogeny can be inferred.

Despite the general view that mtDNA nuclear copies are very rare in *Drosophila*, we show the exact opposite can be the case. Given the strong functional linkage of mtDNA to the nuclear genes, there is the intriguing possibility that nuclear copy amplification may play some role in mtDNA expression level adjustment as has recently been proposed in the case of critical genes such as TP53 (Abegglen *et al.,* 2015).

Materials and Methods

The stocks used in this paper are listed in Table 1. A significant issue with stocks of the *immigrans* group has been the disturbing frequency of mislabeled stocks found in the collections of the major stock centers (Waddell *et al.,* 2019), something that became problematic in the *nasuta* subgroup apparently only from the 1990's onwards. Consistent with Waddell (1990) all stocks listed in Table 1 were carefully checked for morphology, behavior, and provenance as all three can be required to allocate specimens to the named taxa without ambiguity.

Table 1. Stocks analysed in this study. Ten digit codes after name indicate from UCSD , else via Ehime Stock Centre, Japan. An X in the analysis name indicates a stock reassigned to a taxon in Waddell (2019), while provenance is feasible, even probable, but unproven.

Stock label as received	Reported Provenance	Lab labels	Analysis name
immigrans 1511-1731.13	Ehime, Japan, 2005 (iso)	immigrans	*immigrans*
siamana 15111-1741.00	Siem Reap, Cambodia (1968).	siam_Cambodia	*siamana*_Cambodia
sulfurigaster 15112-1811.06	Luzon, Philippines (1968).	albost_Luzon	*siamana*X_Luzon
hypocausta 15115-1871.04	Luzon, Philippines (1968).	hypo_Luzon	*hypocausta*_Luzon
hypocausta 15115-1871.05	Guam (2012)	hypo_Guam	*hypocausta*_Guam
neohypocausta 15115-1881.02	Wau, New Guinea (1961). (iso)	neohypo_NGuinea	*neohypocausta*_PNG
LAE221 taxon G	Waddell(1990) says Lae, Papua New Guinea 23/7/81. (iso)	taxonG	*niveifrons*_PNG
PNI75 E-19901 pallidifrons	Kolonia, Ponape, Caroline Islands 27/6/81. (iso)	pallidi	TaxonF_X
B-208 taxon F	Waddell (1990) = Kuching, Sarawak, Malaysia 6/6/71 (iso)	taxonF	TaxonF_Sarawak
kohkoa 15112-1771.01	Sarawak, Malaysia (1968).	koh_Sarawak	*kohkoa*_Sarawak
pulaua 15112-1801.00	Semongok Forest Reserve, Sarawak, Malaysia (1968). (iso)	pula_Sarawak	*pulaua_Sarawak*
sulfurigaster 15112-1831.02	Wau, New Guinea (1961).	sulf_NGuinea	*sulfurigaster*_PNG
0-30 niveifrons	Waddell(1990) has Lae, Papua New Guinea 14/10/79. (iso)	nivei	*sulfurigaster*X_PNG
sulfurigaster 15112-1831.01	Kavieng Island, New Ireland (1961).	sulf_NIreland	*sulfurigaster*_Nireland
sulfurigaster 15112-1821.00	Oahu, Hawaii (1966).	bilim_Oahu	*bilimbata_Oahu*
sulfurigaster 15112-1821.08	Guam, Mariana Islands (1968).	bilim_Guam	*bilimbata*_Guam
kohkoa 15112-1771.04	Rizal, Luzon, Philippines (1972).	koh_Philippines	*albostrigata*X_Philippines
sulfurigaster 15111-1741.01	Sarawak, Malaysia (1968)	albost_ Sarawak	*albostrigata*_Sarawak
sulfurigaster 15112-1811.04	Siem Reap, Cambodia (1968).	albost_Cambodia	*albostrigata*X_Cambodia
sulfurigaster 15112-1811.05	Brunei, Borneo (1968)	albost_Borneo	*albostrigata*_Brunei
sulfurigaster 15112-1811.07	Singapore, Indonesia (Singapore never Indonesia, but part of Malaya from 1963 to 1965)	albost_Indonesia	*albostrigata*_Singapore
neonasuta 15114-1861.00	Mysore, India (1971).	neon_Mysore	*neonasuta*_Mysore
sulfurigaster 15112-1811.08	Kandy, Sri Lanka.	albostrigata_SriLanka	*albostrigata*_SriLanka
kepulauana 15112-1761.01	Sarawak, Malaysia (1968).	kep_Sarawak	*kepulauana*_Sarawak
kepulauana 15112-1761.03	Ulu Temburong National Park, Brunei (2003). (iso)	kep_Brunei	*kepulauana*_Brunei
nasuta 15112-1781.00	Mysore, India (1971). (iso)	nas_Mysore	*nasuta*_Mysore
nasuta 15112-1781.06	Mombasa, Kenya (1976).	nas_Mombasa	*nasuta*_Mombasa
albomicans 15112-1751.05	Ishigaki Island, Japan (1969).	albom_Ishigaki	*albomicans*_Ishigaki
NOU98 taxon J	Noumea, New Caledonia 4/8/81 (iso)	taxonJ	*albomicansX*

The sequencing reads used were 2 × 125 bp paired end reads from fragmented DNA with a size range of approximately 200-500 basepairs, generated on a HiSeq2500 machine (UNC Chapel Hill Genome Center). A high quality of raw reads was obtained in three separate runs, which is highly desirable starting point in order to unambiguously assemble an mtDNA genome. A major issue with sequencing these mtDNA genomes are long A and T tracts, which are a hotspot for "read through" homopolymer sequencing errors. These are

markedly variable even within machines of the same technology, *e.g*. an Illumina MiSeq versus a HiSeq2500, with the latter having a much better error profile (Stoler and Nekrutenko, 2021).

The assemblies were produced by a combination of methods by different researchers, then compared with each other. Initial assemblies were via depth of coverage, and evidence of predominant local pairwise phasing. When strange coverage and/or high heterozygosity regions were encountered, then this approach could be taken to a second level in order to attempt to assemble an "alternative," or lower frequency of reads, assembly. This resulted in initial assemblies labeled "H" for highest copy number and "alt" for made using python scripts. An alternative approach, using the assembler idba_ub, followed by comparing and merging contigs using Genious R9, followed by using BWA to map back raw reads for visual inspection/correction follows the description in Winkworth *et al.* (2021). When differences were seen between the resulting assemblies, there was a further iteration looking at these differences in detail, with map back of raw reads, to determine a final assembly. These are the "R+" and were typically very close to, or identical to, the other assemblies, particularly "R" and "H", except in six notable cases, where a well-supported alternative assembly was validated and these are labeled "R+_alt."

It is important to make a note on the taxonomy of the *nasuta* subgroup. Herein, we use the designated *species* name to identify the taxon, except for the subspecies of one traditional taxon. These are *D. sulfurigaster sulfurigaster*, *D. s. bilimbata*, *D. s. albostrigata*, and *D. neonasuta*. These taxa are effectively morphologically indistinguishable, which applies to other "cryptic" species in the group such as the *kepulauana*, *nasuta*, and *albomicans* trio (Wilson *et al.,* 1969; Kitagawa *et al.,* 1982; Waddell, 1990; Kitagawa, 1991). They are distinguishable into two subsets based on behavior, with a set of visual behaviors common to all, but with *albostrigata* and *neonasuta* showing a second, frequent, effective, and very different male courting wing display (Spieth, 1969; Waddell, 1990). The taxa *bilimbata* and *sulfurigaster* might be definable based on geography, with *bilimbata* being on non-continental Pacific Islands, and *sulfurigaster* being on the continental masses of Sahul (Australia, and PNG) plus, maybe, New Caledonia. Our favored hypothesis is that *bilimbata* is a dispersal of *sulfurigaster*, perhaps largely mediated by the Austronesian migrations. In contrast, *neonasuta* has a poorly defined geography within the range of *albostrigata*. Thus *neonasuta* has long been suspected to simply be an artificial construct, something backed up by the early genetic studies, *e.g*., (Suzuki *et al.,* 1990). Recent genetic studies favor these interpretations, with *neonasuta* being well clustered within *albostrigata* and *sulfurigaster* being difficult to define as a monophyletic group without *bilimbata* (Waddell *et al.,* 2019; Mai *et al.,* 2020).

The strongly distinct, by morphology, behavior and mate choice, species *pulaua* is a close relative of *sulfurigaster/bilimbata* to the exclusion of *albostrigata* (Waddell *et al.,* 2019; Mai *et al*., 2020). Herein, *sulfurigaster* and *bilimbata* are distinguished from each other, but *neonasuta* should be treated as a local embedded population of *albostrigata*. In terms of biogeography, it is possible that this whole clade, including *pulaua*, is a cline with punctuation by expansions across geographic barriers. First, a fairly well connected Sundaland/South East Asia, and across into India/Sri Lanka to the West, and East into the upsampled regions of Indonesia defined by barriers such the Wallace Line, the Weber Line, and the Lydekker Line, then into Sahul (Australia + New Guinea) and, finally, out into the Pacific. If so, *pulaua*, might have originated over Wallace's line, but then back migrated West, while there are also unconfirmed reports of *pulaua* in Papua New Guinea.

Software used included IDBA-UD for draft assemblies (Peng *et al.,* 2012), BLAST (Altschul *et al.,* 1997) for homology checks, BWA (Li And Durbin, 2009) and Bowtie2 (Langmead and Salzberg, 2012) for mapping short reads to assemblies, Genious 9 (Kearse *et al.,* 2012), Tablet (Milne *et al.,* 2009), and Mesquite (Maddison and Maddison, 2019) for checking and refining assemblies, then creating alignments and translations, PAUP (Swofford, 2000) and Splits Tree 4 (Huson and Bryant, 2016) for phylogenetic analyses. Phylogenetic analyses presented include ML with unequal rates across sites (Swofford *et al.,* 1996), SVD quartets (Eriksson, 2005; Chifman and Kubatko, 2014), NeighborNet (Bryant and Moulton, 2004), and Split Decomposition (Bandelt and Dress, 1992) and a *phi* test for non-homogeneity of phylogenetic signal (Bruen, 2006).

Results

We first look at issues of mtDNA assembly and issues of relying upon assemblies appearing in Genbank. We identify a number of ways Frankenstein, or perhaps Frankenfly, assemblies can arise. That is, mtDNA assemblies that are "sewn" together from a range of non-orthologous sources.

Starting with reasonable assemblies and accurate stock/taxa identifications, a reliable backbone mtDNA phylogeny of the *immigrans* species group is inferred. Then, a wide range of mtDNA sequences from Genbank are added to the alignment, trees estimated, suspect sequences are identified and removed, and then using ML, SVD quartets and NeighborNet, a favored resolved mtDNA phylogeny of the group is proposed.

Reference genome of D. formosana KR265324 = NC_028518

D. formosana is a species traditionally assumed to be a close relative of *D. immigrans*. This sequence is the key data of a peer reviewed paper from a well-respected scientific publisher (Taylor and Francis) in the specialist journal "Mitochondrial DNA Part A" (Shi *et al.,* 2015). Its apparent quality is further attested to by the nine authors that attached their names to the publication, including three equal first authors, and by the fact that Genbank further assessed it and awarded it the highest quality representation of its class; that is, it became a reference mtDNA genomic sequence for the species and the first for the whole *immigrans* species group (indicated by the "NC" prefix designation in its second Genbank ID).

The mtDNA of NC_028518 is reported to have at least one noteworthy and surprising feature. The Cox1 gene has a strange initial sequence in *Drosophila* and other arthropods, and does not have a canonical methionine start codon. This has been explained by supposing that the immediately upstream four base pairs is a quasi-methionine codon read correctly due to "wobble." This explanation is thrown into doubt by findings such as those of Steward and Beckenbach (2009) where mature mRNA for this gene show the expected first codon, which in *D. melanogaster* referenced to alignments with other species, is TCG (serine). The reported *D. formosana* sequence places an ATA (methionine) coding position immediately before this, which, if correct, suggests interesting experiments on how the cell determines where to start this essential protein sequence. That is, if the sequence is accurate it can inform insightful molecular evolutionary analyses.

The reference sequence NC_028518 at first glance appears reasonable and appropriately annotated. However, its location on the phylogenetic tree of Shi *et al.* (2015) is a bit surprising as it is expected to locate as a close sister to *Drosophila immigrans,* which is closer to *D. virilis* and allies (the subgenus *Drosophila*) rather than closer to *D. melanogaster* and its allies (subgenus *Sophophora*). However, the rooting point of a tree can often be the most uncertain part of an analysis, so this is not decisive.

To better resolve deep divergences of mtDNA it is often useful to translate and concatenate unambiguously aligned parts of the protein coding genes. We now also add in some of our new mtDNA sequences to add some resolution. A quick NJ tree shows NC_028518 does not colocate with them, so there is at least an issue of species misidentification. Applying more sophisticated models only strengthens the result (Figure 1a). The quite short terminal edge to *D. formosana* can be a sign of a hybrid species/sequence, but there is considerable heterogeneity of rates on the tree, so this is not definitive. Interestingly, the *nasuta* flies with only about half the body mass of their nearest relatives such as *immigrans* and *neohypocausta* do show a clear acceleration in mitochondrial protein evolutionary rate.

A NeighborNet analysis highlights *D. formosana* as being adjacent to a markedly non-tree-like part of the network (Figure 1b). NeighborNet is a useful tool for exploring the potential of non-tree splits, but it is limited to a subset of graphs that can be drawn in the plane. How much useful information may be "out of the plane" is hinted at by the fit measures. Here, because they are high (~99%), this suggests the biggest tree and non-tree splits are revealed. One of the best ways to test a specific hypothesis as to which is/are the problem sequence(s) is to remove them and rerun the analysis. This reveals that indeed with *formosana*.KR265324 removed, all the largest non-tree splits disappear (Figure 1c). These non-tree splits suggest *formosana*.KR265324 has signal from the direction of the *immigrans* group and the *immigrans* subgroup in particular, but that it also has signal from the direction of the *Sophophora* subgenus and the *melanogaster* species subgroup in particular.

In a situation like this, tracking the distance from the suspected abnormal sequence to all the other species in the analysis with a sliding window can reveal interesting patterns. Doing this in Excel with a window of size 100 using the well aligned regions of the DNA sequence results in the plot shown in Figure 2. It is clear there are multiple regions of KR265324 that are nearly identical to other species and that these are not the same species in different parts of this assembly. Working from left to right, the first 100 bp or so seem particularly close to *D. littoralis*, followed by a very conserved region where it is hard to discern differences. The region from ~300 to 1400 bp does not closely match any species in our alignment, but blasting it reveals a very close match (99% of query at 99.82% identity) to *D. pseudoobscura*. Following this is a region where the closest match is to *D. melanogaster*, and it is clear this region is so close to *D. melanogaster* relative to its close relatives, such as *D. simulans* and *D. sechellia*, that this is probably from a *melanogaster* individual.

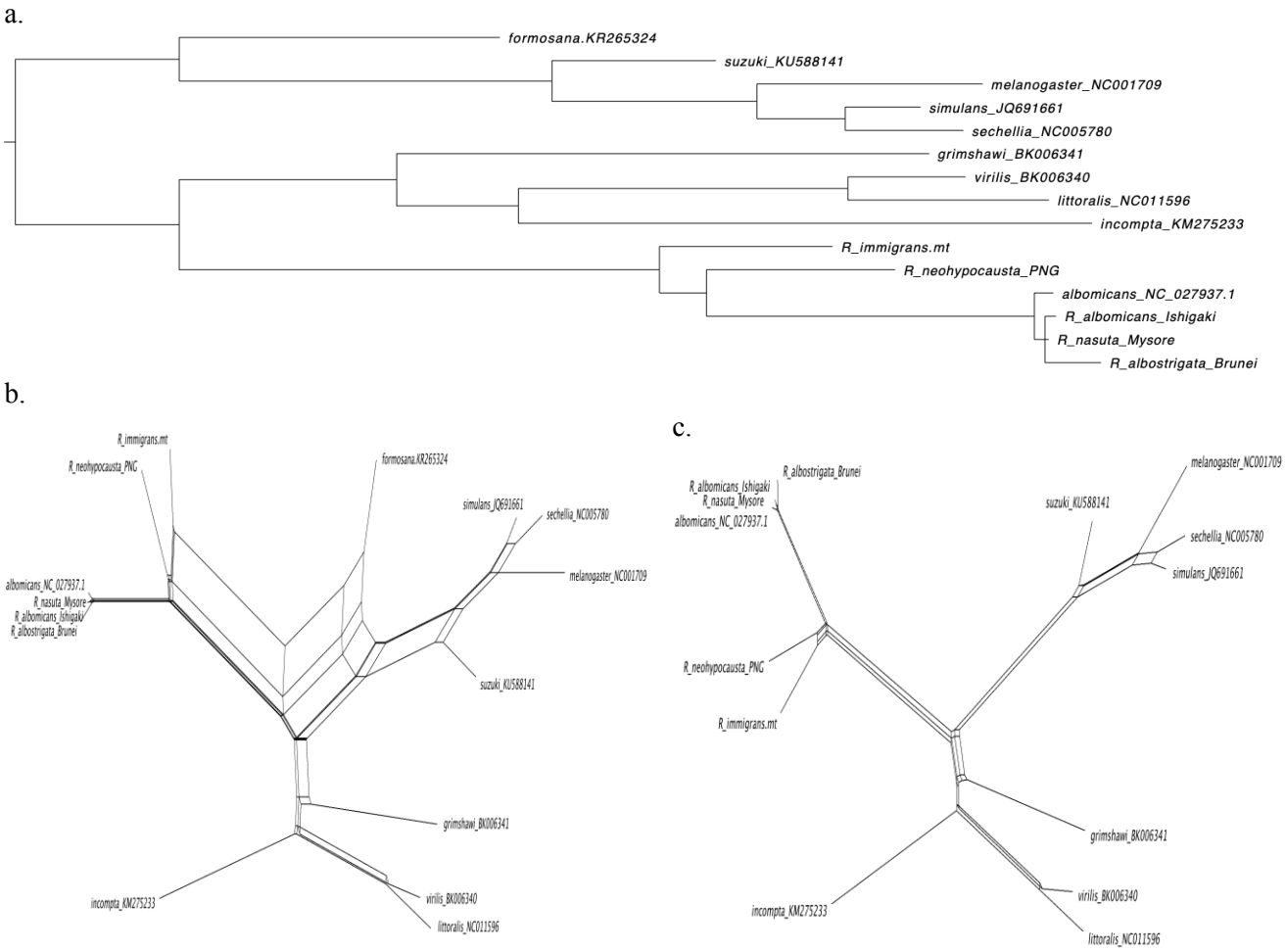


Figure 1. (a). The ML tree for the concatenated protein sequences. The model used is the mtRev24 model, with empirical frequencies and an approximate gamma distribution (4 rate categories, shape = 2.78), plus invariant sites (0.753) implemented in PAUP. Bootstrap support is 100% on all edges, except the edge uniting *virilis*, *littoralis*, and *incompta*, while the edge grouping *neohypocausta* with the *nasuta* species subgroup has 93% and the grouping of *albomicans*_Ishigaki, *nasuta*_Mysore, and *albostrigata*_Brunei has 57%. (b) The NeighborNet using distances matching the model above (fit = 98.87, least squares fit or lsf, which is the percentage of the variance in the data explained, = 99.97). A *phi* test (Bruen *et al.,* 2006) of uniformity of phylogenetic signal is rejected with p = 0.0. (c) As previously but with formosana.KR265324 removed (fit = 99.35, lsf = 99.99). After bootstrapping and removing splits with less than 50% support, this graph becomes a tree except for the non-tree split of *simulans* from *suzuki*, *melanogaster*, and *sechellia*, and that of *albomicans*_NC027937.1 towards *immigrans* and *neohypocausta* (and the *phi* test is now non-significant at 0.08).

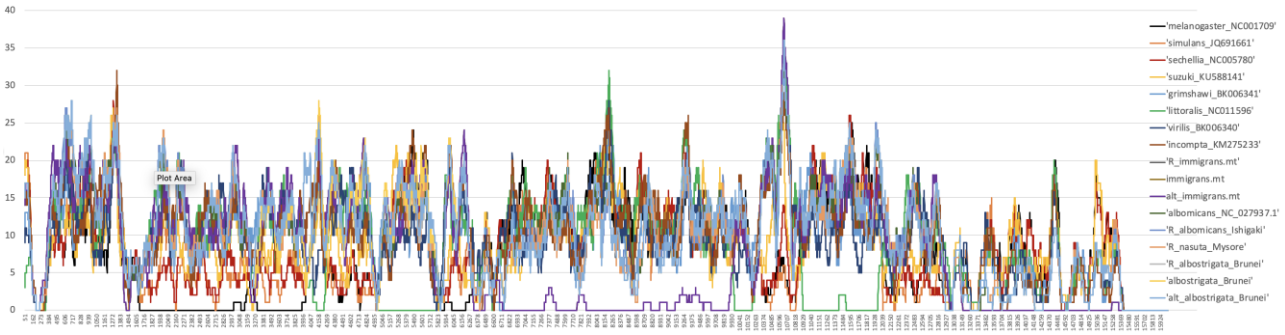


Figure 2. The number of mismatches from *D. formosana* in a sliding window of 100 bp across the conserved and unambiguously aligned regions of the mtDNA genome. For Figures, magnifying them on your browser allows clear visualization.

Figure 2 shows that bases ~4000-5000 are another region without a consistently close match in our alignment (Figure 2). Blasting it reveals a moderate match to *D. simulans* and *D. melanogaster* (both ~94% cover and 94-95% identity). Region 5000-5900 appears particularly close to *littoralis* (Figure 2), which is not revealed in Figure 1b being overwhelmed by other tree and non-tree signals. Then it appears there is a bit of *melanogaster*, followed by a long region (~bases 6700-9900) that is particularly close to *immigrans*. This might indeed be from a fly of *formosana*, but the mismatches do not coincide with the faster evolving regions, suggesting this might even be *immigrans* with sequencing error. Following this is a short region that appears closest to *simulans*, and a longer region that again appears to be particularly close to *littoralis*. The region ~12,000 to 12,900 best matches *melanogaster* group species and BLAST matches the region at 98% coverage and nearly 99.9% identity to a particular strain of *simulans* (identity to *melanogaster* is markedly lower at about 96%). Rounding out the rest of the well-aligned sequence suggests it is particularly close to *immigrans*.

Finally, BLASTing the rapidly evolving control (AT rich) region of KR265324 does not match any known control region. The closest match is a 75% mismatch to *littoralis*. This region is reported with many ambiguity codes, which might indicate a substantial frequency of sequencing error.

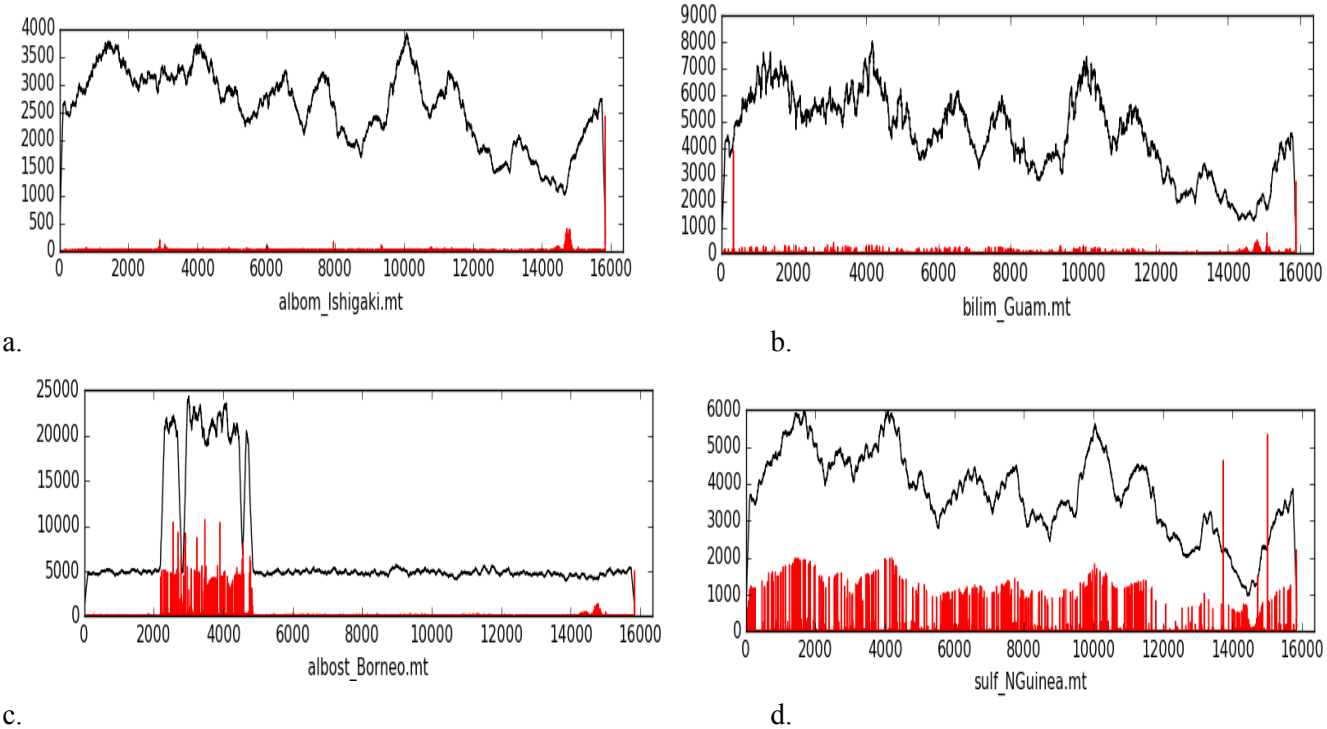
As mentioned earlier the start codon of Cox1 in *formosana* may be unusual. The experimentally verified first three bases of the mature mt mRNA in *melanogaster* Cox1 are TCG. These would seem to code for S (serine) not the canonical methionine. Indeed, across *Drosophila* the homologous first base is not conserved, but the three following bases CGC are highly conserved and rare bases. Two close relatives of *melanogaster* (*simulans* and *sechellia*) have the codon CCG (P or proline) at the start position. As we will see later, the *nasuta* subgroup is unusual in that it has the codon ACG (T = threonine) at this position, except for a few rare instances such as the *albomicans* strain sequenced here, which has GCG (A = alanine). That *formosana* has a methionine codon immediately prior to this apparent start codon is apparently unique in *Drosophila,* but the upstream region including the two prior tRNAs (for Cystine and Tyrosine) does not correspond closely to any known mtDNA in a BLAST search, with the closest hits being to members of the *immigrans* species group (all with an identity of less than 90%) and almost equally good hits to the distantly related *bipectina* species group. Thus, it is unclear if this is sequence error or if there is an unknown fly with such a sequence. The latter would be interesting to see if it would override the enigmatic process that operates in this short 5' mtDNA intergenic region.

We identify the issue here as trans-specific identity confusion. That is, sequence data incorporated in the assembly comes from more than one organism, with apparently large amounts from at least five distantly related species. Later we will see a cis, or within the same genome, identity problem in assembling the mtDNA of some *albostrigata* stocks.

De Novo assemblies of immigrans group mtDNA from MPS data

Assemblies from highly fragmented DNA (*e.g.*, 200-500 bp) followed by massively parallel sequencing (MPS) can produce accurate assemblies, not just for the protein and RNA genes used for deeper phylogenetic analyses, but also revealing hetroplasmy even within individuals (Kloss-Brandstätter *et al.,* 2015). However, they also offer their own particular challenges. During our studies we noticed that some initial assemblies contained apparent errors such as alterations in highly conserved parts of the alignment or structural duplications. In order to explore this further we turned to two different types of assembly, producing a higher frequency assembly, "H", and an alternative "alt", lower frequency, assembly. To further refine a stringent alternative assembly routine for assembling mtDNA without relying on a close homologue was used (Winkworth *et al.,* 2021). Herein, these assemblies are labelled "R". Following this, we then mapped back all reads onto these provisional assemblies. Doing this we noticed a number of distinct categories of pileup plot, examples of which are shown in Figure 3. It was after this that an extra close visual inspection of raw reads in areas where the afore mentioned assemblies differed was made to produce a "R+" or final assembly. In all cases, R and R+ were identical or very close to each other.

The plot of Figure 3(a) for *albomicans*_Ishigaki is quite typical of what most mtDNA pileups in *Drosophila* should look like. There is a background of read differences that is low and fairly uniform except at the far right into the "high AT" region that contains the mtDNA or "control" region, as it is known in vertebrates. These appear to be a mixture of sequencing errors often associated with homopolymers and also naturally occurring intra and inter-individual differences (partly from extracting DNA from 3-5 individuals of highly inbreed lines). The next plot to the right for *bilimbata* suggests a distinctly higher level of background, with one very high peak at about site 216, where this stock has had a notable expansion of an "AT" dinucleotide to 10+ copies that no other taxa in this subgroup show. Despite the read depth often being markedly higher than 30 and thus higher than that expected here for either a single nuclear copy or a very high sequencing error rate (~Q20 or 1 in 100, much higher average quality assessments of these reads), we were unable to phase these disagreeing reads into long alternative assemblies.



a. b.

c. d.

Figure 3. Legend on next page

Figure 3. Four distinct types of pileup encountered mapping raw reads back to mtDNA assemblies using the default settings of Bowtie2. The peaks from ~14,000 onwards are the end of the 16SrRNA and into the high AT "control" region of the mtDNA. The y-axis is the mapped coverage, while the x-axis shows the sequence coordinates for the "R" assembly of the indicated stock. The black line is total coverage, and the red line is the number of reads disagreeing with the "R" assembly at each base. The two anomalously high peaks are due to the Figure drawing software combining adjacent peaks into a single pixel for the final plotting.

The third plot of Figure 3 for *albostrigata*_Borneo suggests extraordinary copy numbers in a very specific region of ~3204 to 5828 of its mtDNA assembly. This is a region of 2625 bases where its assemblies based on different methods differ markedly. In this region the average coverage dramatically increases (from ~5,000 to around 20,000 depth). It also shows distinct dips back to near normal coverage around base 2800 and another around 4800; decreases in coverage not observed in other close relatives. The average autosomal coverage of the read library is estimated as 23 (from GenomeScope analysis of raw reads). Thus the excess of ~15,000 indicates the equivalent of approximately 650 autosomal copies or roughly 650 times 2600 or 1.69 mb, which in turn is ~1% of an indicated 178 million base pair whole genome size. This is clearly not minor and suggests a massively expanded nuclear copy, with a possible pair of deletions within it.

The pileup for *sulfurigaster*_PNG suggests that about one third of the reads map well to the mtDNA, yet show consistently marked differences from it across its full length. Also, the number of reads that disagree with the assembly appears to be a fairly consistent fraction of ~1/3. Further, these disagreeing reads can be pieced into a two full length assemblies by moving between phased pairs of mismatches between these reads. This could be either *in vivo* mtDNA heteroplasmy, *in vitro* contamination with the DNA of another stock, or a near full length nuclear copy that has expanded (or even some combination of these).

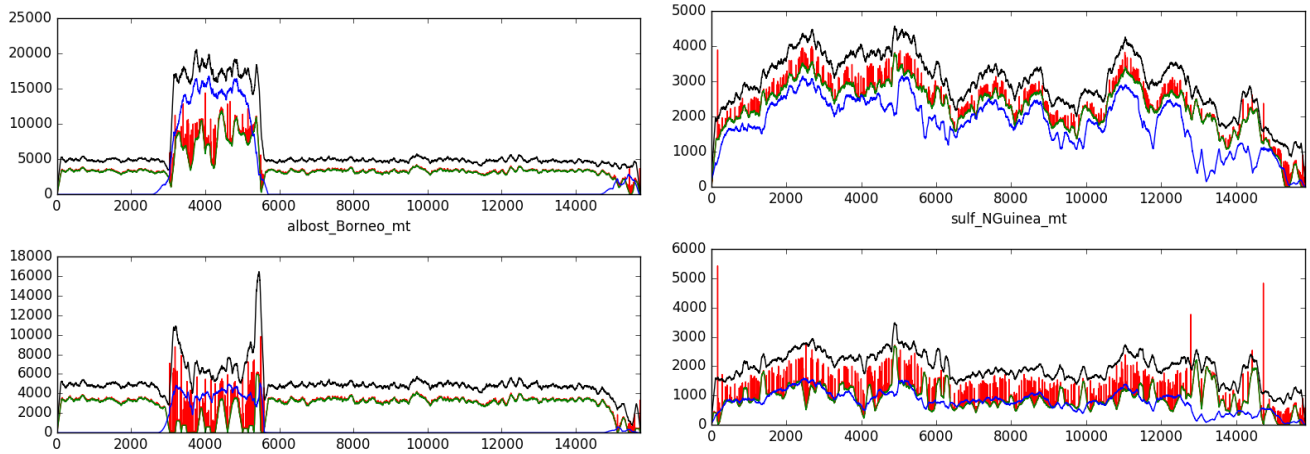


Figure 4. A refined pile up on both a higher frequency assembly "H" (above) and lower frequency "alt" assembly (below) for *albostrigata*_Brunei (left) and *sulfurigaster*_PNG (right). The black line indicates the total number of reads that mapped by the Bowtie2 defaults to that assembly, the green line are counts of reads that map perfectly to that assembly, the blue line the count of reads that map better to that assembly than its twin, while the red line shows mismatches of all mapping reads to the assembly being considered sitting on top of the green line.

In order to scrutinize these interesting alternative assemblies more carefully, we prepared another type of pileup as shown in Figure 4. These count not just all reads mapping (black line), but also counts reads that map perfectly without ambiguity (green line) to a specific assembly and those that map better to that specific

assembly than to any other assembly (blue line). For *albostrigata*_Brunei it is clear that outside of the very high coverage area, there is no ambiguity between the higher and the lower frequency assemblies, except in the control region, so they are effectively the same. The highest frequency version for *albostrigata*_Brunei is unique here as it is the one expected to contain nuclear copies, and the pile up gives assurance that a reasonable consensus of them has been called (blue line and green line are both well above zero). Calling the lower frequency variant in this region without using the more rigorous methods of the "R" assemblies is problematic as seen by the green line hitting zero multiple times. Below we will see how this probable misphasing shows up in evolutionary analyses. In contrast the higher frequency and the alternative assemblies of *sulfurigaster*_PNG appear fine with black, blue and green lines all well above zero over the full length of these assemblies, except in the control region. The green line continuously well above zero suggests that if these are full length "alternative" nuclear copies, they must be duplicated without an intervening spacer.

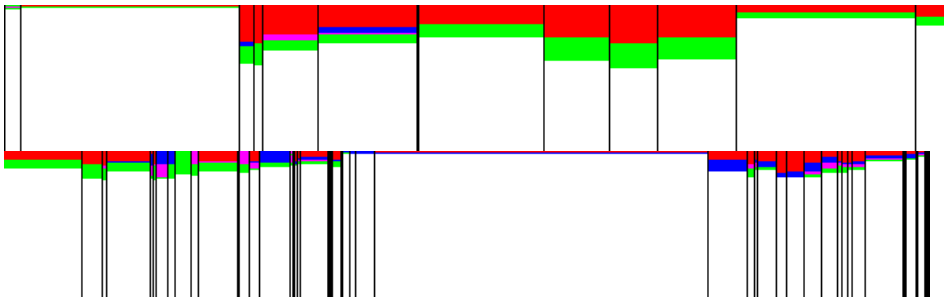


Figure 5. Visualization of the variable phasing of *sulfurigaster*_PNG between the higher frequency "H" and the alternative assembly. (a) The upper example is from the beginning of the assembly and (b) the lower example from the end of the assembly. A black line indicates a difference between these assemblies, red counts paired reads that link numerically the higher frequency assembly (~2/3 of all reads) on the left with the most frequent bases immediately to the right, while green does likewise but for the numerically lesser mutation (~1/3 of reads). Blue indicates the numerically more frequent mutation on the left is linked to the numerically lesser mutation to the right, while purple is the reverse of this. Blue and purple indicate split pairwise phasing.

An example of a pairwise "phasing" diagram is shown in Figure 5. This example from a *sulfurigaster* stock suggests strong "splits" occasionally. If these are predominantly due to a well supported alternative in the reads mapping best to a portion of the alternative assembly reads, then the diagram should show a marked purple bar, then a black bar and a then blue bar. This is indeed what is seen over most of the alternative assemblies in this region, with an example appearing in Figure5a. Since such alternative variants still involve many hundreds of reads, they are far beyond any expected sequencing error. This is most consistent with the explanation that this lower frequency assembly represents a consensus of many nuclear copies that have slightly diverged from each other. It is only in the control region that occasional strong alternative phasing patterns such as purple/blue then blue/green pairwise phasing is seen (Figure 5b).

The marked difference in copy numbers, the phasing patterns, and generally very low ceiling of sequencing errors, allows confidence in these alternative assemblies. That is, walking through the phasing of the mutually high frequency segregating sites, either the "H" assembly is obtained if starting from the first higher number of total mutations or the lower copy number alternative assembly is produced outside of the control region. In the control region phasing can be very difficult, but becomes easier with visual inspection with particular care over homopolymer read through errors. Combined with the consistent frequency differences of the alternative *sulfurigaster* copies this gives confidence most of the phasing in the control region is correct for these alternatives.

As already mentioned, because the "R" assemblies might run into issues with such well-supported alternatives, all differences of the "alt" to the "H" and the "R" assemblies were looked at critically. This

included looking at the mapping of raw reads and whether differences might be sequencing error. This was done conservatively. All "alt" "R" and "H" assemblies as constructed here were collapsed back to a common "R+" assembly, except for three dual assemblies of *sulfurigaster* stocks and the three dual assemblies of the *albostrigata* strains Brunei, Cambodia, and Singapore (these are labelled "R+_alt").

Phylogenetic evaluation of the quality of these short read assemblies

We begin with an evaluation of the tree of a subset of these assemblies as seen in Figure 6a. It is generally very well supported. Any uncertainty in this tree is not shared by parsimony or distance methods (*e.g.*, BME with HKY distances), which both give all edges 100% bootstrap support, while ML with a clock also gives 100% bootstrap support for all edges. Just as important, the NeighborNet graph is very tree-like, so why the specific ML model of Figure 6 is reporting reduced support in part of the tree, compared to other methods, is unclear.

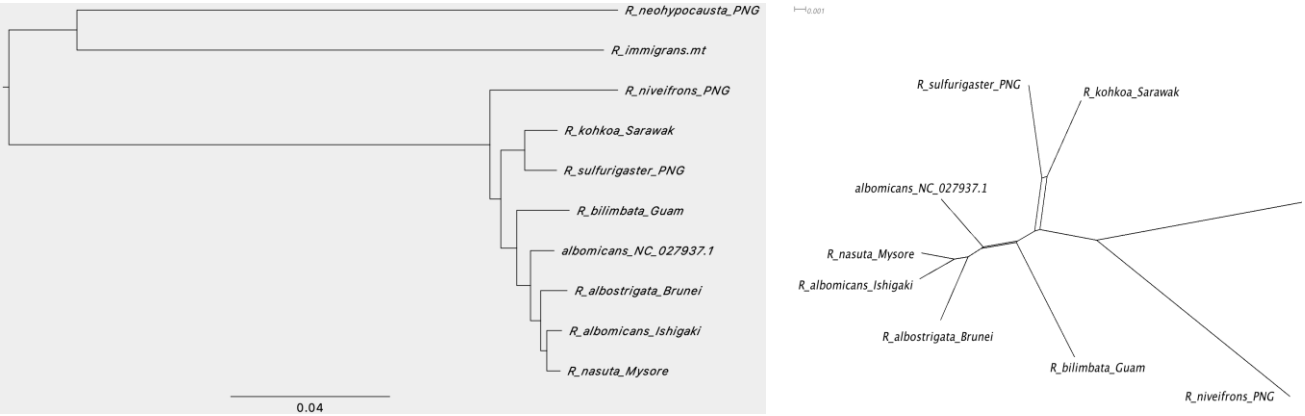


Figure 6. (a) The midpoint rooted ML tree of the "R" assemblies for all well-aligned nucleotide sites. The BIC automodel function in PAUP choses the K81 model with unequal base frequencies (A: 0.35, C 0.11, G 0.11, T 0.43), a four rate category based on a gamma shape of 1.49 and a proportion invariant sites of 0.766. This model explains over 95% of the approximated posterior probability, while the similar HKY model is the fourth best explaining 0.4% of the posterior (amongst the j7 set of models being considered). The bootstrap replicates were 100% for all edges except the edge separating *niveifrons* from the other *nasuta* subgroup sequences (56%) and the one below that in the Figure (86%). The HKY model with invariant sites (with parameters f = 0.355, 0.110, 0.103, 0.432, pinv 0.803, kappa 15.4) was much better than the gamma site rate model, and was evaluated as the best model currently functioning in SplitsTrees4. (b) The NeighborNet with the afore mentioned distance model, with only edges having at least 50% bootstrap support shown (outgroups to the right not shown). All edges have 100% support except the two non-tree splits, which have support in the mid 50's.

At first glance the "R" assembly of *albostrigata*_Brunei looks reasonable with no stop codons, unlike the "H" assembly of this stock. However it does have four very suggestive non-conservative amino acid changes in highly conserved regions of the proteins in the region showing much deeper read coverage. Based on a parsimony reconstruction of this same region, the "R" assembly of *albostrigata*_Brunei shows 8 nucleotide substitutions excess to the closest relative in Figure 6, the *albomicans*_Ishigaki sequence. However, a critical look using NeighborNet at all the assemblies for this taxon suggests while there are minor improvements to be had, the "R+" assembly is not hugely different, and retains excess substitutions over other assemblies (Figure 7a). Overall, the Network is quite tree like, despite using uncorrected substitution distances, with the largest excess of non-treeness being exactly where the outgroups join. Recall, that for the *albostrigata*_Brunei the alt or lower frequency assembly, was a fairly standard algorithm attempting to

reassemble the functional mtDNA in this instance. It is clearly struggling, with many unique changes apparent. Also, in this region, the *sulfurigaster* primary and alternative assemblies locate unambiguously and do not show signs of cross contamination in assemblies with any other taxa, including the *albostrigata* alternative assemblies.

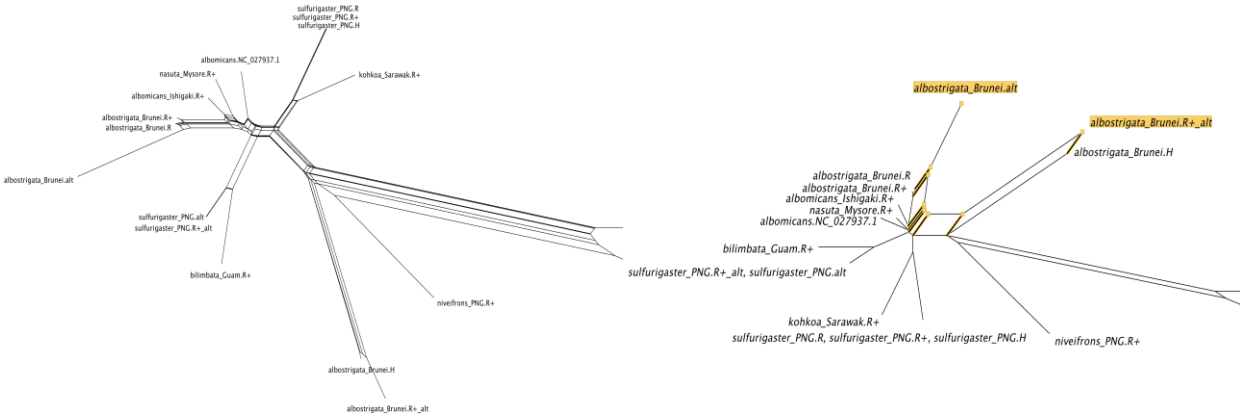


Figure 7. (a) The NeighborNet based on Hamming distances for the high frequency region of *albostrigata*_Brunei, to assess how the alternative assemblies relate to each other in this region. Off to the left are the R+ assemblies for *immigrans* (above) and *neohypocausta* (below). The fit is 98.48 (lsf = 99.96%). (b) The same data and distances but visualized with SplitDecomposition with fit 80.92 (lsf 97.41).

In Figure 7b the application of Split Decomposition, with its strict rule that all quartets that disagree with the dominant tree must be the consistent (else the dominant split and any direct alternative are both collapsed), clearly shows that the Brunei_alt and R+_alt assemblies share something in common. It is that during the reevaluation of the R+ assembly some of the unique substitutions in the alt assembly were moved onto the R+_alt assembly. The point of origin of the Brunei nuclear copies is apparently an old numt originating from an mtDNA that originated itself close to the branch point with *niveifrons*. This would suggest this numt may still exist as a low frequency copy in other taxa, but if in a repetitive region could be difficult to reassemble and identify. In Figure 7a and 7b we can see the two assemblies of *sulfurigaster*_PNG mtDNA show almost no difference across assembly techniques or iterations. The purported functional copy shows an elevated rate of evolution compared to that of *kohkoa*, while the alternative copy shows a sharp slowdown compared to their sister taxa *bilimbata*. If these are nuclear copies, then the apparent evolutionary clock effectively stopped when nuclear amplification occurred.

An amino acid based evolutionary examination

Looking at the predicted amino acid sequences can give further insight into whether substitutions are conserved or non-conserved and a sensitive look at potentially elevated rates of functional (or dysfunctional) evolution. At the protein level, the *albostrigata* assemblies differ from each other only from sites 866-1637 of the final amino acid alignment (with start, stop, and ambiguously aligned regions excluded). Figure 8a shows a NeighborNet of the well-aligned amino acids in this region. The protein sequences show few changes and do not present a clean signal of phylogenetic relationships, but a method such as NJ does chose a tree of the form (*niveifrons*,(*kohkoa*,(*albostrigata*_Brunei((*albomicans, nasuta*), *albostrigata*_Singapore, *albostrigata*_Cambodia) using Hamming distances. In these regions, the *albostrigata* protein sequences all an accelerated rate of evolution despite no confounding issues of high copy number. For example, the *albomicans* and *nasuta* sequence appear ancestral to the *albostrigata*_Singapore and *albostrigata*_Cambodia sequences, which then both evolve uniquely in their own ways.

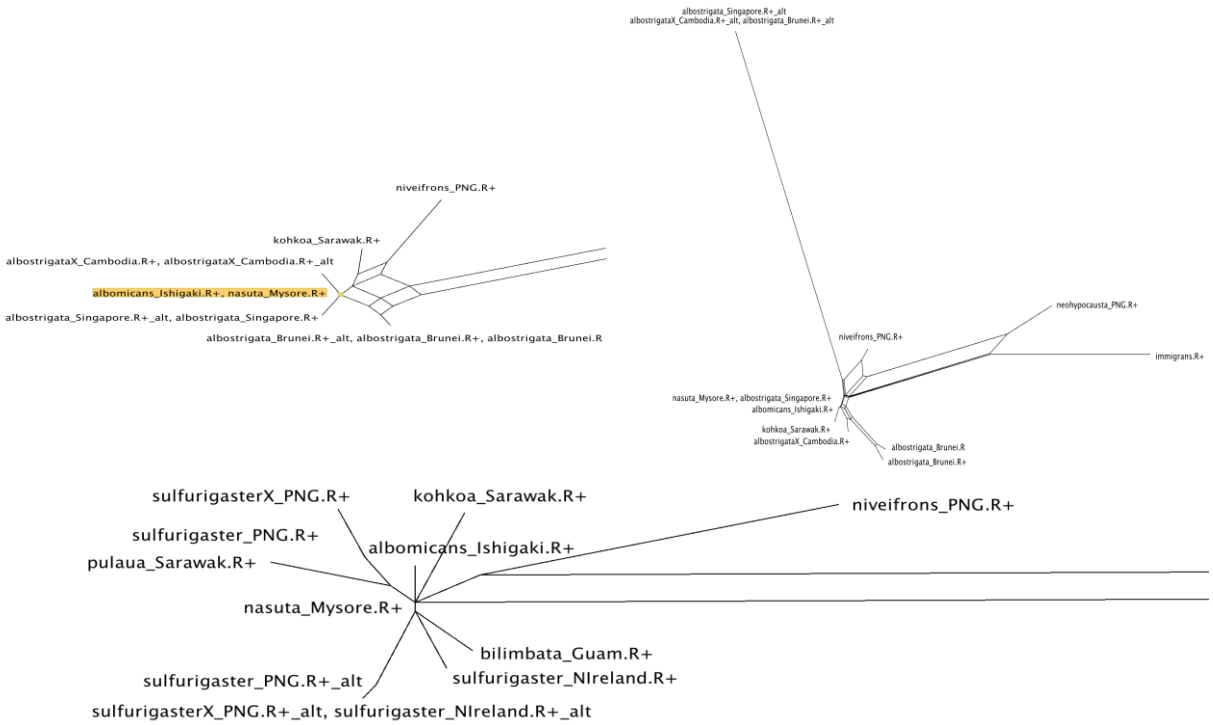


Figure 8. (a) A NeighborNet for for sites 1-865 and 1638-3699 of the amino acid alignment using Hamming distances. Reported fit is 100% and 100% of the variance is explained. Off to the right are the outgroups with *neohypocausta* above and *immigrans* below. In this region the R+ and R+_alt assemblies amino acid assemblies do not differ and seem uncompromised by nuclear copies. (b) A NeighborNet for sites 866-1637 of the amino acid alignment. This is the region massively amplified as a highly amplified numt. Reported fit is 99.40% and 99.99%. (c) The Split Decomposition diagram of all aligned amino acids showing the relationships of the *sulfurigaster* R+ and R+_alt assemblies (fit 88.58 and lsf 99.27).

Figure 8b shows the NeighborNet of the aberrant high coverage region for proteins. The alternative *albostrigata* sequences are clearly the most unusual and they are identical to each other. They originate from a point close to where *niveifrons* branches off. NeighborJoining determines the tree to be ((*niveifrons*, *albostrigata*_R+_alts), ((*albomicans, nasuta, albostrigata*_Singapore, *kohkoa*),(*albostrigata*_Cambodia, *albostrigata*_Brunei))) with *nasuta, albomicans* and *albostrigata*_Singapore being identical and showing a zero length terminal edge very close to the predicted ancestor of all these sequences except *niveifrons*. The hardest assembly to phase in the high read frequency region is *albostrigata*_Brunei. Even at the amino acid level there are only minor differences from the R to the R+ assemblies. As will be seen later, it appears that the functional *albostrigata*_Singapore mtDNA associates more closely with the *nasuta*_Mysore and *albomicans*_Ishigaki sequences than it does with the functional *albostrigata*_Brunei and Cambodia sequences. Given the evidence that functional *albostrigata* mtDNA is in fact most closely related to some, but not other, *albomicans* and *nasuta* stocks and is far away from its closest "nuclear" relatives, raises the possibility that the mtDNA of *albostrigata* may have been introgressed from *albomicans/nasuta* flies and that at the same time this caused genomic instability such as a megabase scale numt, the apparent result of massive duplication of an old numt.

Figure 8c shows a SplitDecomposition representation of the protein sequences of the R+ and R+_alt assemblies for *sulfurigaster*. In this network the *nasuta* protein sequence appears identical to that of the last common ancestor of all *nasuta* taxa except the earlier diverging *niveifrons*. There is a small kink in the edge to *sulfurigaster*X_PNG.R+, which is where *sulfurigaster*_PNG.R+ joins the network (with a zero length terminal edge, hence not distinguished from an ancestor) and likewise for the *sulfurigaster*_PNG.R+_alt and the other *sulfurigaster* alt sequences. The higher frequency assemblies fall in two groups, with the PNG stocks being closest to *pulaua*, and the NewIreland stock being closest to *bilimbata*. As we will see much more clearly later, by concordance and discordance with the nuclear phylogeny of Waddell *et al.* (2019), there is strong evidence that *bilimbata* and *sulfurigaster* New Ireland have kept their ancestral mtDNA, while the other *sulfurigaster* and *pulaua* have replaced it with a relative of the *kohkoa* mtDNA. Compared to their closest relative (*kohkoa*) the primary assemblies of *pulaua* and *sulfurigaster*_PNG, but not *sulfurigaster*X_PNG, appear to be evolving at higher rates. For the alternative assemblies there is a hint of increased protein evolutionary rate.

Checking the nucleotide substitution spectrum

Another way to evaluate the alternative assemblies of *albostrigata* and *sulfurigaster* is to look at their substitution spectra compared to their closest relatives. The *albostrigata*_Brunei sequence shows highly non-conserved substitutions, including indels, non conservative changes across the alignment, and even new stop codons, all strongly suggesting nuclear copies. In contrast, it is hard to see anything unusual in the *sulfurigaster* assemblies, except a few scattered changes that appear somewhat non-conservative.

If these "alternative" sequences have been in the nuclear environment for a substantial time, then they should also be showing a much higher ratio of transversions to transitions. To minimize the effects of model inadequacy interacting with the weighted tree, it is also useful to match this counting to a close relative. In the case of the R+_alt *albostrigata*_Brunei assembly, the best match would seem to be the *niveifrons*_PNG assembly, which yields the following counts (reconstructed on the ML tree, parsimony results are similar) A to C: *niveifrons* = 0 (high frequency *albostrigata*_Brunei = 1), AG 12(6), AT 3(3), CA 1(4), CG 0(0), CT 14(8), GA 3(2), GC 0(0), GT 1(3), TA 2(4), TC 13(13), and TG 2(0). For the a*lbostrigata* assembly there is a noticeable decrease in the proportion of transitions and a marked increase in transversions, with a one tailed test significant at P = 0.03. If the mtDNA diversity in the ancestor that fixed this numt was similar to the average seen in species across the group, one quarter to one half of the evolution along this edge could have been in the ancestral mtDNA before nuclear insertion. For the *sulfurigaster*_PNG.R+_alt assembly the best match would be the *bilimbata*_Guam assembly. This test was non-significant, and indeed the rate of transitions in the s*ulfurigaster* assembly was slightly higher than that in *bilimbata*_Guam. This suggests that if this is a nuclear insertion, the relative duration of this being a nuclear copy before expansion (and the subsequent assembly of a consensus of nuclear copies), compared to its unique evolution as an active mtDNA molecule was relatively short.

An ML tree including all key immigrans group mtDNA sequences from Genbank

Genbank contains a few sets of partial mtDNA sequences for *immigrans* group species. These have the Genbank ID of the sequence, or the ID of the first part of the concatenated sequence, after their species identification (according to Genbank). Sections of CO1 and CO2 come from the work of Dr. Katoh and colleagues (Katoh *et al.,* 2007; Izumitani *et al.,* 2016) and these are marked with a "_K" at the end. Those marked with an "_Y" are from Yu *et al.* (1999), while those marked "_V" are a selection of voucher sequences produced by Patrick O'Grady based on multiple conserved primer sets which cover about 2/3 of the mtDNA, but most taxa are missing a few segments (*e.g.* O'Grady and De Salle, 2018). These were based upon stocks chosen from the UCSD stock centers collection in the early 2000's. The exact stocks were apparently not recorded, but it was usually the stock of that species with the lowest number on the last part of the stock number (O'Grady, pers comm.). A few of these sequences seem to be from original specimens rather than stocks. The sequences marked "_L" are by (Liu *et al.,* 2015) for a range of South East Asian locations and three full length mtDNA assemblies, for *albomicans* (NC_027937.1, Xiongbin *et al.,* 2016), *nasuta* and *neonasuta* (the last two apparently parts of forthcoming whole genome analyses).

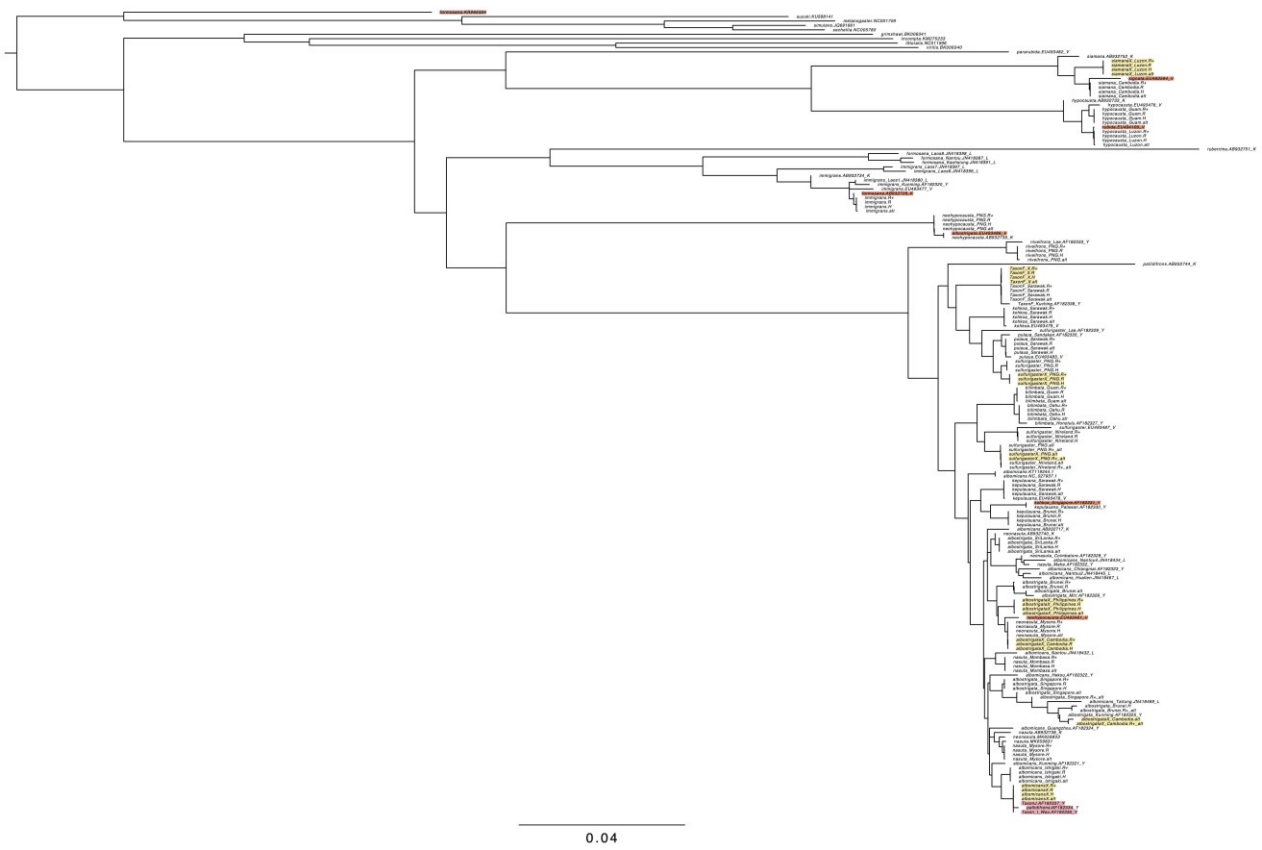


Figure 9. An ML tree of all mtDNA sequences for the *D. immigrans* group of species, based on all sites alignable across all these taxa. The automodel function in PAUP using the BIC model selection criterion, chose a Kimura 3 parameter model, with unequal base frequencies, and a mixture of invariant sites and an approximated (4 rate categories) gamma distribution of site rates. This was iterated, starting on an NJ tree, until the selected model and parameter estimates did not change. Taxon labels marked in yellow are stocks shown to be misidentified in Waddell *et al.* (2019) indicated with an "X" in the name. Labels marked in red are sequences we strongly suspect are mislabelled from their original sources (see text for meaning of suffix identifying sequence sources). Zoom into the online version to read labels.

The overall tree produced by ML in Figure 9 is reasonable being anchored by the whole mtDNA genomes and will be discussed further below. However, it is also very clear that the misidentification of stocks and sequences, shown in Waddell *et al.* (2019), is even more extensive with apparently mis-attributed sequences coming from all the studies above except Liu *et al.* (although many of their sequences do not seem to be in Genbank). These have severely compromised the interpretation of the phylogeny of the group. We are currently reaching out to these authors and others to see if it is possible to definitively identify the species and probable provenance of the apparently misidentified sequences. In some cases, it seems there is a direct switch of stock/taxa/sequencing labels, for example, the voucher "_V" sequences of *neohypocausta* and *albostrigata*.

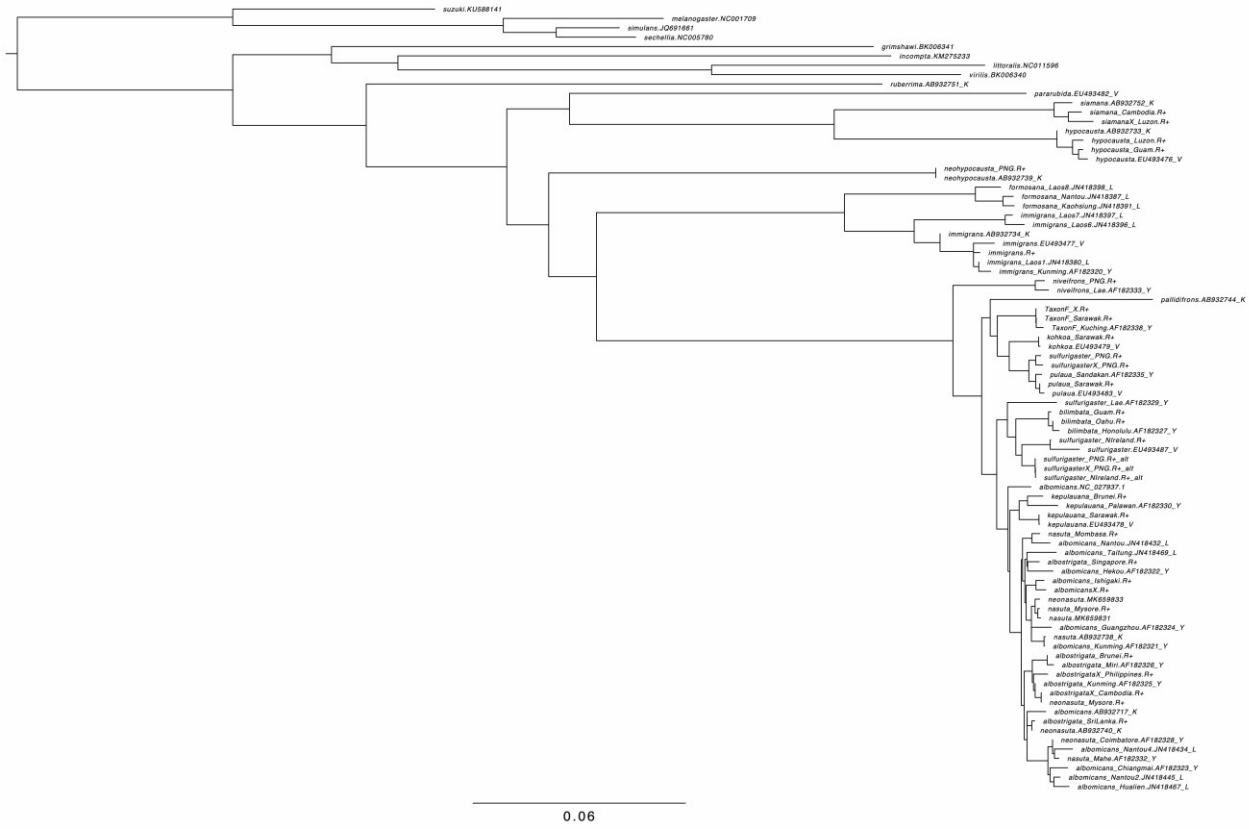


Figure 10. Like Figure 9 but the alignment is refined with the removal of misidentified taxa and/or multi-source assemblies, and the analysis rerun. The model chosen by BIC was rclass = (abccda), rmatrix = (1 14.8 1.9 1.9 18.2 1) basefreq=(0.369 0.086 0.082 0.463), rates=gamma(4), shape=0.62, pinv=0.57.

Figure 10 shows the a tree estimated in the same manner with our assemblies whittled down to just the R+ assemblies (with our 3 additional *albostrigata* assemblies removed) and sequences marked in red or yellow in Figure 9 removed. The deep lineage of *pararubida* locates fairly securely as sister to the other species of the *hypocausta* subgroup, where it has been traditionally located along with *rubida*. However, *neohypocausta* was also traditionally also assigned to this subgroup, but it clearly is not a member (*e.g.*, Waddell *et al.,* 2019; Suvorov *et al.,* 2020). Another distinct species without close relatives yet sampled is *ruberrima,* and it is seen to be mobile, alternating between sister to all other *immigrans* group species or sister to *immigrans* itself. Here p*allidifrons* is based on a short sequence, which locates deep in the *nasuta* subgroup either sister to all other *nasuta* taxa or sister to a group containing taxon F and *kohkoa*. Unpublished results, based on nuclear sequences, suggest it actually locates as a deep branch sister to all *nasuta* taxa except *niveifrons* (the internal edge between the two alternative locations seen in Figures 9 and 10).

Another sequence that is locally labile is the reference *albomicans* mtDNA genome NC_027937. It is associated in Genbank as part of a full genome assembly of *albomicans*, which is based on a Taiwanese stock (Mai *et al.,* 2020). However, sorting carefully though the records, this mtDNA assembly (Xiongbin *et al.,* 2016) is apparently based on reads from an earlier full genome assembly of *albomicans,* which, in turn, is based on sequences from a stock labeled KM-55 from Kunming in China (from reads of Zhou *et al.,* 2012, deposited into Genbank). It variously joins as sister to the *kepulauana* sequences, sister to all the other *nasuta*, *albomicans* and *albostrigata/neonasuta* sequences, or sister to both of these clades. It would be highly desirable for Genbank to only assign an mtDNA assembly to a whole genome project if it comes from the same source.

In Figure 9 it becomes clearer where the sequences we are confident are correctly attributed to *albostrigata/neonasuta* sequences locate. One clade purely of these taxa, and including our locally wandering SriLanka assembly, are deep within the diversity of *albomicans* and *nasuta* sequences, very suggestive of an mtDNA introgression. They are disjoint from the location of our assembly of *albostrigata*_Indonesia, which is nestled quite securely close to a partial sequence from Hekou in Yunnan, China. Note, our assemblies of *albostrigata*_Cambodia and *neonasuta* are very closely related. As this stock of *albostrigata* has already been shown to have some mixed ancestry (Waddell *et al.,* 2019) we wonder if the mtDNA of this stock might have been poached in captivity.

Finally, this analysis of the mtDNA, after removable of suspect sequences, allows a reconciliation of the mtDNA tree of the *nasuta* subgroup with the nuclear genomic tree of these taxa (Waddell *et al.,* 2019). It requires two to four mtDNA introgression events, but otherwise the mtDNA coalesces perfectly on the nuclear tree, including for the *bilimbata, sulfurigaster* New Ireland and R+_alt assemblies, which would appear to have the original mtDNA for the *albostrigata* + *pulaua* + *sulfurigaster* + *bilimbata* clade. It seems *albostrigata* got its mtDNA from within the *nasuta/albomicans* clade. This may have been just one main introgression event early in the history of the group, with more than one distinct mtDNA type initially coming in, followed by sorting (hence, the distinct mtDNA of *albostrigata*_Singapore). This species remains largely sympatric with *albomic*ans/*nasuta* over its range, so intense local sampling should make it clear if this is an ongoing phenomenon. In addition *pulaua* and some stocks of *sulfurigaster* have introgressed their mtDNA probably from *kohkoa.* While *pulaua* is sympatric with this species, *sulfurigaster* is not. It is rumored that *pulaua* may be present in PNG perhaps having dispersed there from the west, in which case *sulfurigaster* stocks in PNG may have received its mtDNA from *pulaua* via introgression. However, it is also possible that *sulfurigaster* and unknown *kohkoa*-like taxa may overlap in the large part of Indonesia between Borneo and PNG where these flies have not been accurately sampled and identified. While the X-chromosome is often more resilient to introgression than the autosomes (*e.g*., Waddell, 2018, in *Homo*), it would be interesting if mtDNA introgression drags over maternal X-chromosome regions that give compatibility with that mtDNA. At present *sulfurigaster* has functional mtDNA from two quite distinct sources.

Evaluation with SVD quartets and NeighborNet

SVD quartets are a type of quartet that have a number of useful properties. Like Lake's invariants, they are insensitive to the distribution of rates across sites, which are a major cause of long edge attraction effects. These effects are, in practice, only partly compensated for using specific likelihood models such as a gamma distribution plus invariant sites (Waddell and Penny, 1996; Waddell *et al.,* 1997). In addition these are consistent with a General Time Reversible (GTR) model of substitution (Swofford *et al.,* 1996). They can be tweaked to be consistent with incomplete lineage sorting or to be consistent with all evolution occurring on a single tree. Once a collection of all quartets is estimated it is tempting to use a method such as parsimony to find a single tree that is consistent with as many estimated quartets as possible. This might be thought of as a quasi-compound likelihood method, which fails to address the correlation structure of the quartets (the analogy is pairwise distances and Generalized Least Squares). However, sets of quartets, like distances, are also a natural starting point for display by methods such as NeighborNet or Split Decomposition, which can help to understand the properties of the set of quartets markedly better than a single tree.

First, all the well aligned sections of the complete mtDNA's were processed with the PAUP command "svd treemod = shared nquart = 10000000 mrpfile = quartets_out.nex nthreads = auto evalQuartets=all boot = yes boot = yes", which resulted in a maximum parsimony = maximum clique = maximum compatibility tree. The output file of quartets was then visualized using NeighborNet as seen in Figure 11a, 11b. The maximum compatibility tree is the same as the NJ tree based on the quartets in SplitsTree except for the local move of *albostrigata*_SriLanka. The trees generally are well supported with bootstrap values for most internal edges close to 100%. The clade of *immigrans* plus *neohypocausta* is ambiguous. It contradicts the 100% bootstrap support for *neohypocausta* as the sister group to the *nasuta* subgroup obtained in Waddell *et al.* (2019) and Suvorov *et al.* (2020) based on millions of unique nuclear markers or millions of aligned nucleotide sites, respectively. A further insight into what is going on here can be seen in Figure 12b, which is based on the well aligned amino acid sequences of these taxa. There is a clear split of *neohypocausta* grouping with the *nasuta* subgroup taxa. There is a dominant split of the *nasuta* subgroup + *neohypocausta* + *immigrans,* but also a split

of *neohypocausta* + *immigrans* + *hypocausta* subgroup, which is also quite apparent looking at the aligned proteins. With Hamming distances a split of *neohypocausta* + *immigrans* appears, so, based on the model and tree estimation method used, the result fluctuates. A *phi* test finds no significant clumping of these alternative signals, so the unlikely, but potentially exciting possibility of evidence for recombination in the mtDNA is null. At the moment the preponderance of evidence is that the historical mtDNA tree is the same as the nuclear tree, that is the clade of *neohypocausta* + the *nasuta* subgroup; however, there is the possibility the history of the mtDNA here is distinct from that of the species tree.

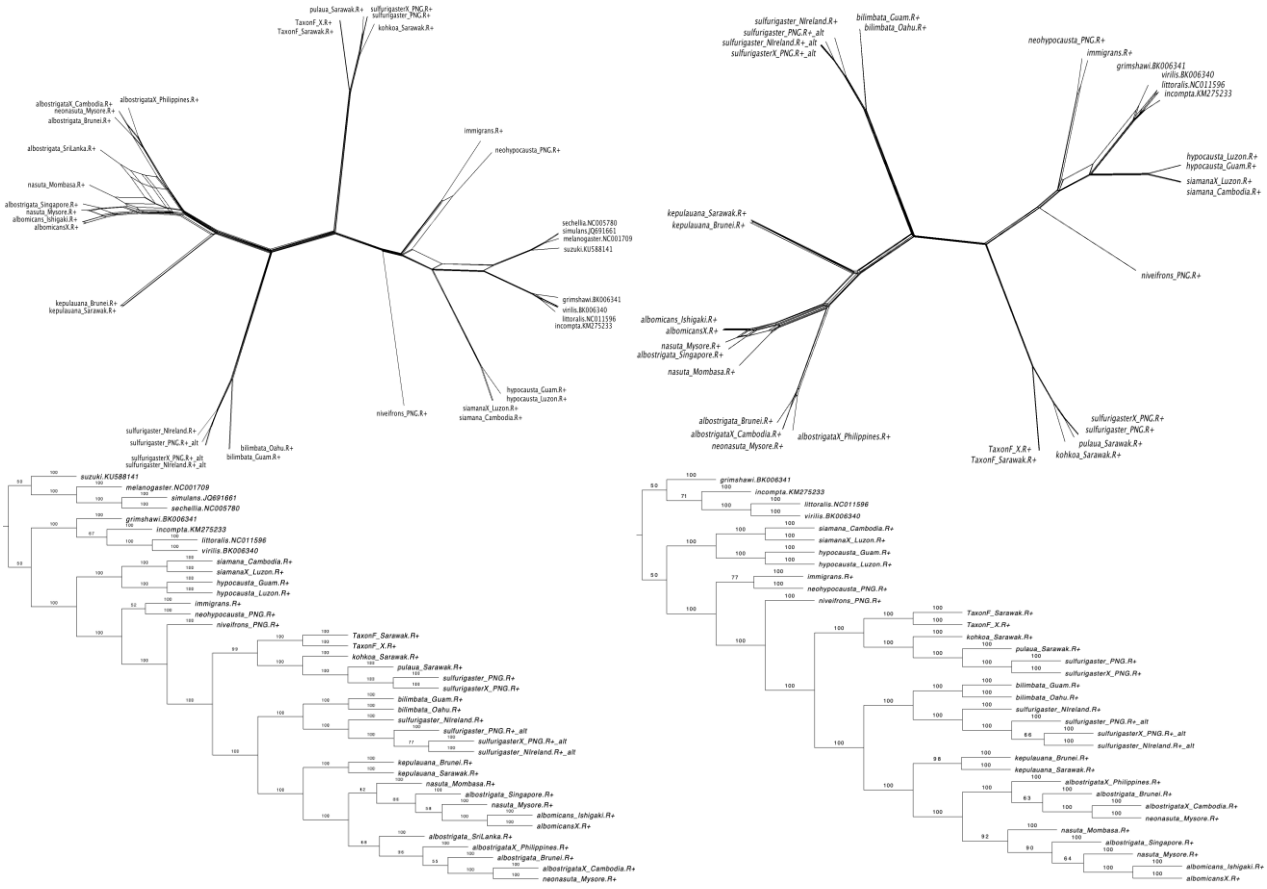


Figure 11. (a) The NeighborNet representation of the set of quartets produced by SVD in PAUP for the whole mtDNA nucleotide sequences (fit 99.52 and lsf 99.99). (b) As for (a) but with the furtherest outgroups and *albostrigata*_SriLanka.R+ removed. (c) and (d) the respective maximum compatible quartet trees with bootstrap proportions.

Adding in three further *nasuta* subgroup mtDNA assemblies from Genbank and three particularly interesting partial sequences, then visualizing the selected quartets with NeighborNet gives Figure 12a. In terms of the complete mtDNA sequences, the assembly *nasuta*.MK659831 is very close to our *nasuta*_Mysore assembly and might even be the same stock. The *neonasuta*.MK659833 assembly is a close relative of both of these, and if the species identity is correct, then it suggests possibly another mtDNA transfer into the *albostrigata/neonasuta* lineage from the *albomicans/nasuta* flies, but this time between two recently sympatric populations. The reference assembly *albomicans*.NC_027937 is seen to be ambiguous between two locations, while other methods, such as ML, will also place it just outside of where *kepulauana* sequences join the rest of the "tree."

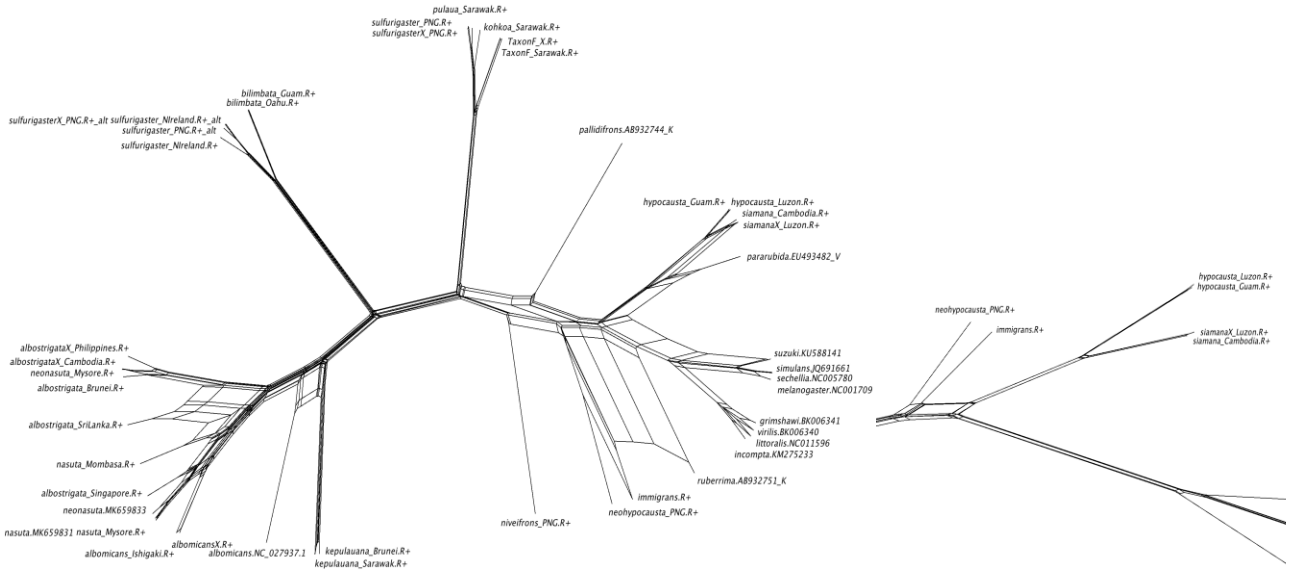


Figure 12. (a) The NeighborNet of the SVD quartets for the full length alignable mtDNA sequences, plus interesting partial sequences (fit 98.89, lsf 99.96). (b) A crop of the NeighborNet based on complete mtDNA protein sequences using an ML mtMam24 distance assuming a proportion of invariant sites of 0.85. With fewer sites assumed invariant, a split of *immigrans* and *neohypocausta* uniquely together grows. The *nasuta* subgroup is to the left and the *virilis* group species are the outgroup to the bottom right.

The location of *pallidifrons* is seen to be most likely deep in the *nasuta* subgroup, with the bootstrap of SVD quartets 100% sure it is deeper than all other *nasuta* taxa except *niveifrons*, and 94% sure it is the deepest lineage of the subgroup. The location of *pararubida* seems fairly secure as a deep lineage in the *hypocausta* subgroup, with an attraction to the distant outgroup *suzuki* being the major disturbance. The deep sparsely sequenced *ruberrima* is most strongly associated with *immigrans* alone, but shows successive attractions toward the outgroups, helping to explain its behavior on the ML trees.

Discussion

Finally after many decades of mtDNA giving discordant and often outright confusing results (*e.g*., Yu *et al.,* 1999) with respect to other data sets used to infer the phylogeny of the *nasuta* subgroup (Waddell, 1990; Kitagawa, 1991), there is finally a clear understanding and reconciliation with the well-supported nuclear tree of Waddell *et al.* (2019), and to a lesser or equal extent that of Mai *et al.* (2020) and Suvorov *et al.* (2020), depending on whether these last two papers are correct to identify *pulaua* + *sulfurigaster* as the true clade and not due to introgression (the mtDNA suggests the opposite). This comes about only by the identification of mislabeled taxa and sequences, plus 2 or more introgression events all involving the members of the *albostrigata* + *pulaua* + *sulfurigaster* + *bilimbata* clade. Otherwise, except for within the *albomicans/nasuta* group, coalescence of the mtDNA follows the proposed species tree exactly. The *nasuta/albomicans* separation is interesting as it is largely defined by *albomicans* fusing chromosome 3 to its X and Y chromosomes, so more than half its genome is sex linked. In terms of morphology, behavior and mate choice experiments, there is no evidence of separation (Kitagawa *et al.,* 1982; Waddell, 1990; Kim *et al.,* 2013), thus, mtDNA might well be quite mobile within this clade.

The deep phylogeny of the *immigrans* species group is coming into focus. The preferred interpretation of the current results using a wide range of methods is that the tree is (((*hypocausta, siamana*), *pararubida*),

((*immigrans, ruberrima*), (*neohypocausta, nasuta* subgroup))). It is clear that bootstrap and/or posterior probability support of trees alone cannot give confidence that the phylogeny is correct, and corroborating results such as an interpretable NeighborNet analysis are also needed. It is also possible that there remain taxa misallocated within the traditional subgroups that will move to different places within this subtree, with recent results such as Suvorov *et al.* (2020) strongly suggesting that taxa such as the *quadrilineata* subgroup are not part of the *immigrans* species group, while *D. pruinosa* appears the closest outgroup.

Within the *nasuta* subgroup, our mtDNA analyses reinforce our confidence in the 100% bootstrap resolution of the major lineages within the *nasuta* subgroup seen in Waddell (2019), more so than Mai *et al.* (2020), while Suvorov *et al*. (2020) lacks many lineages. Given the strong bootstrap support from SVD quartets and the form of the NeighborNet of SVD quartets, we predict that the *pallidifrons* lineage, along with its closely related taxa, the putative subspecies taxon I and taxon J (Waddell, 1990; Kitagawa, 1991) locates with highest probability as sister to all other *nasuta* taxa except *niveifrons*, or as sister to all other *nasuta* taxa, or with a smaller probability as sister to *niveifrons*, and an even smaller probability as sister to the taxon F + *kohkoa* clade. This is consistent with studies of the overall inter-fertility of taxa (Wilson *et al.,* 1969; Kitagawa *et al.,* 1982; Kitagawa, 1991) that show no hint of inter-fertility between *niviefrons,* the *pallidifrons* complex, or the taxa (*sulfurigaster subspecies, pulaua,* taxon F*, kohkoa, kepulauana, albomicans,* and *nasuta* species), while the last set of taxa show varying degrees of high to partial inter-fertility/developmental compatibility.

The existence of a very highly duplicated and large numt in three *albostrigata* lineages that have introgressed a new mtDNA does not seem a likely coincidence. Also given the strong tendency of *Drosophila* to purge numts we suspect either a recent bout of genomic instability or, perhaps, a compensatory mechanism analogous to how elephants might regulate TP53 pathways (Abegglen *et al.,* 2015). Indeed, even the human genome seems to have numt tRNA's that seem involved in regulation, just not at such a high copy number (Telonis *et al.,* 2014). Locating this expansion is proving difficult as it could reside in a highly repetitive and poorly assembled part of the genome. In terms of *sulfurigaster*, the alternative assembly of their mtDNA remains a mystery. If it is contamination, it makes little sense. We have no records of any other *sulfurigaster* stock in the lab, and indeed, when stocks were received and DNA was first extracted, we had not completed our morphological and behavioral checks on the stocks so one of these stocks went by a very different name. One of the stocks was mislabeled as *niveifron*s, and came from a different stock center, so any simple swap or mixing of extracted DNA seems unlikely. We cannot yet distinguish if this is a functional second copy of the mtDNA (therefore deep heteroplasmy), nor have we been able to locate it in nuclear genomic assemblies, yet the coincidence with a subspecies undergoing active mtDNA introgression suggests more than happenstance.

Acknowledgments: This work was partly supported by NIH grant 5R01LM008626 to PJW.

References:

Abegglen, L.A., A.F. Caulin, A. Chan, K. Lee, R. Robinson, M.S. Campbell, W.K. Kiso, D.L. Schmitt, P.J. Waddell, S. Bhaskara, S.T. Jensen, C.C. Maley, and J.D. Schiffman 2015, Journal of American Medical Association: doi:10.1001/jama.2015.13134; Altschul, S.F., T.L. Madden, A.A. Schaffer, J. Zhang, Z. Zhang, W. Miller, *et al.* 1997, Nucleic Acids Research 1997, 25: 3389–402. https://doi.org/10.1093/nar/25.17.3389 PMID: 9254694; Bandelt, H.-J., and A.W.M. Dress 1992, Advances in Mathematics 92: 47–105; Breton, S., L. Milani, F. Ghiselli, D. Guerra, D.T. Stewart, and M. Passamonti 2014, Trends Genet. 30: 555–564; Bruen, T., H. Philipe, and D. Bryant 2006, Genetics 172: 2665–2681; Bryant, D., and V. Moulton 2004, Mol. Biol. Evol. 21: 255–265; Chifman, J., and L. Kubatko 2014, J. of Theoretical Biology 374: 35-47; Eriksson, N., 2005, Tree construction using singular value decomposition. *In*: *Algebraic Statistics for Computational Biology*, (Pachter, L., and B. Sturmfels, eds.). Chapter 19, pp. 347-358. Cambridge University Press; Healy, T.M., and R.S. Burton 2020, PNAS 117: 6616–6621; Hill, G.E., 2019, Mitonuclear speciation. *In*: *Mitonuclear Ecology*, ISBN-13: 9780198818250, DOI:10.1093/oso/9780198818250.003.0007; Hill, G.E., 2020, Trends Genet. 36; 403–414; Huson, D.H., and D. Bryant 2016, SplitsTree4, version 4.14, splitstree.org.; Izumitani, H.F., Y. Kusaka, S. Koshikawa, M.J. Toda, and T. Katoh 2016, PLoS One 11(7):e0160051; Katoh, T., D. Nakaya, K. Tamura, and T. Aotsuka 2007, Zoolog. Sci. 249: 913–921; Kearse, M., R. Moir, A. Wilson, S. Stones-Havas, M. Cheung, S. Sturrock *et al.* 2012, Bioinformatics 2012, 28: 1647–9, https://doi.org/10.1093/bioinformatics/bts199 PMID: 22543367; Kim, Y.-K., D.R. Phillips, and Y. Tao 2013, Ecol. Evol. 3: 2061–2074; Kitagawa, O., 1991, *Evolutions of Populations*. Tokyo: Tokyo University Press,

pp. 67-78. [in Japanese]; Kitagawa, O., K.I. Wakahama, Y. Fuyama, Y. Shimada, E. Takanashi, M. Hatsumi, M. Uwabo, and Y. Mita 1982, Jpn. J. Genet. 57: 113–141; Kloss-Brandstätter, A., H. Weissensteiner, G. Erhart, G. Schäfer, L. Forer, S. Schönherr, *et al.* 2015, PLoS ONE 10(8): e0135643. doi:10.1371/journal.pone.0135643; Langmead, B., and S.L. Salzberg 2012, Nature Methods 9: 357–359, doi:10.1038/nmeth.1923; Li, H., and R. Durbin 2009, Bioinformatics 25: 1754–60 https://doi.org/10.1093/bioinformatics/btp324 PMID: 19451168; Liu, F.-G.R., S.-C. Tsaur, and H.-T. Huang 2015, J. Insect Sci. 15: 69-76, DOI: 10.1093/jisesa/iev056; Maddison, W.P., and D.R. Maddison 2019, Mesquite: a modular system for evolutionary analysis Version 3.61. Available: http://www.mesquiteproject.org; Mai, D., M.J. Nalley, and D. Bachtrog 2020, Mol. Biol. Evol. 37: 208-220; Milne, I., M. Bayer, L. Cardle, P. Shaw, G. Stephen, F. Wright, *et al.* 2009, Bioinformatics 26: 401–2. https://doi.org/10.1093/bioinformatics/btp666 PMID: 19965881; Nguyen, T.H.M., S. Sondhi, A. Ziesel, S. Paliwal, and H.L. Fiumera 2020, BMC Evolutionary Biology 20: 128 https://doi.org/10.1186/s12862-020-01685-6; Nunes, M.D., M. Dolezal, and C. Schlötterer 2013, Mol. Ecol. 22: 2106–2117; O'Grady, P.M., and R. DeSalle 2018, Phylogeny of the Genus *Drosophila*. Genetics: Genetics (Flybook) 209: 1-25; Peng, Y., H.C.M. Leung, S.M. Yiu, and F.Y.L. Chin 2012, IDBA-UD: Bioinformatics (Oxford, England). 28: 1420–8, https://doi.org/10.1093/bioinformatics/bts174; Posth, C., C. Wißing, K. Kitagawa, *et al.* 2017, Nat. Commun. 8: 16046 https://doi.org/10.1038/ncomms16046; Russell, O.M., G.S. Gorman, R.N. Lightowlers, and D.M. Turnbull 2020, Cell 181: 168–188; Scheid, A.D., T.C. Beadnell, and D.R. Welch 2021, British Journal of Cancer 124: 124–135, https://doi.org/10.1038/s41416-020-01125-8; Shi, X., Q. Zhu, N. Wu, J. Tu, D. Yang, H. Xu, Y. Yao, M. Yang, and D. Li 2015, Mitochondrial DNA Part A: 27: 4316-4317; Spieth, H.T., 1969, University of Texas Publ. 6918: 255-270; Stewart, J.B., and A.T. Beckenbach 2009, Gene 445: 49-57; Stoler, N., and A. Nekrutenko 2021, NAR Genomics and Bioinformatics 3: 1-9 doi: 10.1093/nargab/lqab019; Suvorov, A., B.Y. Kim, J. Wang, E.E. Armstrong, D. Peede, E.R.R. D'Agostino, D.K. Price, P.J. Waddell, M. Lang, V. Courtier-Orgogozo, J.R. David, D. Petrov, D.R. Matute, D.R. Schrider, and A.A. Comeault 2020, Current Biology: in press (see also bioRxiv doi: https://doi.org/10.1101/2020.12.14.422758); Suzuki, Y.M., O. Kitagawa, and K.I. Wakahama 1990, Genetica 80: 53–66; Swofford, D.L., 2000, Phylogenetic Analysis Using Parsimony (*and Other Methods), Version 4.0b10. Sinauer Associates, Sunderland, Massachusetts; Swofford, D.L., G.J. Olsen, P.J. Waddell, and D.M. Hillis 1996, Phylogenetic Inference. *In*: *Molecular Systematics*, second edition. (Hillis, D.M., and C. Moritz, eds.), pp 450-572. Sinauer Assoc., Sunderland, Mass; Telonis, A.G., P. Loher, Y. Kirino, and I. Rigoutsos 2014, Front. Genet. 5: 344-355; Waddell, P.J., 1990, The mating behaviour and phylogeny of the *nasuta* subgroup of the genus *Drosophila*. MSc. Thesis. University of Auckland, Auckland, New Zealand; Waddell, P.J., and D. Penny 1996, Evolutionary trees of apes and humans from DNA sequences. *In*: Handbook of Symbolic Evolution, (Lock, A.J., and C.R. Peters, eds.). pp 53-73. Clarendon Press, Oxford; Waddell, P.J., D. Penny, and T. Moore 1997, Molecular Phylogenetics and Evolution 8: 33-50; Waddell, P.J., 2018, The Phylogenomic Origins and Definition of *Homo sapiens*. *In*: *Rethinking Human Evolution*, Vienna Series in Theoretical Biology (Schwartz, J.H., ed.). Chapter 8, pp 139-180. MIT Press, Cambridge, Massachusetts; Wallace, D.C., 2018, Nat. Genet. 50: 1642–1649; Wilson, F.D., M.R. Wheeler, M. Harget, and M. Kambysellis 1969, Univ. Texas Publ. 6918: 207–253; Winkworth, R.C., S.E. Bellgard, P.A. McLenachan, and P.J. Lockhart 2021, PLoS ONE 16(5): e0250422. https:// doi.org/10.1371/journal.pone.0250422; Wolff, J.N., E.D. Ladoukakis, J.A. Enriquez, and D.K. Dowling 2014, Phil. Trans. R. Soc. B 369: 20130443. http://dx.doi.org/10.1098/rstb.2013.0443; Xiongbin, K., X. Luo, Z. Zhang, Z. Zhang, J. Yang, and G. Bi 2016, Mitochondrial DNA Part A 27: 3638-3639; Yu, H., W. Wang, S. Fang, Y.-P. Zhang, F.-J. Lin, and Z.-C. Geng 1999, Molecular Phylogenetics and Evolution 13: 556-565; Yuting, D., W. Chen, Q. Li, S.J. Rossiter, and X. Mao 2021, Front. Zool. 18: 42. https://doi.org/10.1186/s12983-021-00424-x; Zhou, Q., H.-M. Zhu, Q.-F. Huang, L. Zhao, G.J. Zhang, S. W. Roy, B. Vicoso, Z.-L. Xuan, J. Ruan, Y. Zhang, *et al.* 2012, BMC Genomics 13: 109.

Guide to Authors

Drosophila Information Service prints short research, technique, and teaching articles, descriptions of new mutations, and other material of general interest to *Drosophila* researchers. The current publication schedule for regular issues is annually, with the official publication date being December. The annual issue will include material submitted during the calendar year. The submission deadline for each annual issue will be 31 December of the publication year except when special circumstances occur, but articles are accepted at any time. Submissions should be in MS Word. .

Submission: Proofs will not be sent to authors unless there is some question that needs to be clarified. The editor reserves the right to make minor grammatical, spelling, and stylistic changes if necessary to conform to DIS format and common English usage. Tables may be slightly reformatted to conform to DIS style. Color illustrations will appear black and white in the printed version but will be in color on our web site (www.ou.edu/journals/dis).

Citation of References: Citation should be by name and date in the text of an article (Smith, 1989; Jin and Brown, 1990; Waters *et al*., 1990). At the end of the article, references should be listed **alphabetically** by senior author, listing all authors with initials, date, journal name, volume and page numbers. Titles will not be included except for books, unpublished theses, and articles in press. An example format is:

Green, R.L., 1998, Heredity 121: 430-442.
Waters, R.L., J.T. Smith, and R.R. Brown 1990, J. Genet. 47: 123-134.

Note the initials are before each name except for the senior author.

Manuscripts and inquiries should be sent to James Thompson, Department of Biology, University of Oklahoma, Norman, OK 73019. Telephone (405)-325-4821; email jthompson@ou.edu; FAX (405)-325-7560.

Teaching Notes

Studying cell proliferation profile in pupal retina model of *Drosophila melanogaster.*

Deshpande, Prajakta[1], Anuradha Venkatakrishnan Chimata[1], Esther Labya[1], and Amit Singh[1,2,3,4,5,6]. [1]Department of Biology, University of Dayton, Dayton, OH 45469, USA; [2]Premedical Program, University of Dayton, Dayton, OH 45469; [3]Center for Tissue Regeneration and Engineering (TREND), University of Dayton, Dayton, OH 45469; [4]Integrative Science and Engineering (ISE), University of Dayton, Dayton, OH 45469; [5]Center for Genomic Advocacy (TCGA), Indiana State University, Terre Haute, IN, 47809; [6]Corresponding Author: Dr. Amit Singh; Email: asingh1@udayton.edu; Tel: 937-229-2894; Fax: 937-229-2021

Abstract

We have developed a lab activity to introduce student(s) to cell biology, genetics experiential learning module by employing immunohistochemistry approach to study growth in the *Drosophila* pupal retina. We have used this lab to teach students (1) how to set up genetic crosses, (2) perform pupal retina dissections, and (3) immunohistochemistry and imaging. With this lab, students gain hands-on experience in staining and microscopy to study cellular processes during development. This is a two-day protocol for dissection, immunohistochemistry, and imaging. On the first day, students dissect the eye-brain complex from the pupae, fix it using a fixative and incubate it overnight in primary antibody to mark the membrane and neurons to visualize differences in the number of cells surrounding each ommatidial unit. On second day, the tissues were processed with secondary antibody, mounted onto a slide in a mountant, and visualized under a fluorescence or confocal microscope. Key words: *Drosophila melanogaster*, Eye, Pupal retina, Hippo pathway, Growth, Confocal Microscopy.

Introduction

With the increasing use of technology in modern teaching approaches, it is also imperative to integrate hands-on approaches to encourage biomedical research. Use of such activities can be highly beneficial in training students in basic laboratory skills, introducing model organisms as tools for research (Usman and Singh, 2011; Wang, 2017). *Drosophila melanogaster* is being used commonly as a model for introducing genetics and basic experimental methods to students (Tare and Singh, 2008; Tare and Singh, 2009; Puli and Singh, 2011; Gogia *et al*., 2017; Mehta and Singh, 2017; Deshpande *et al*., 2020; Deshpande *et al*., 2021). Such approaches allow application-based learning and encourage students to explore research as a career choice. These methods include morphological / structural, biochemical, and molecular approaches among others to study basic cellular processes. Here, we have used a immunohistochemical approach to study growth / cell proliferation in *Drosophila* retina.

Drosophila is a genetically tractable research model that is used to study development and diseases (Sarkar *et al*., 2016; Gautam *et al*., 2021; Kango-Singh and Singh, 2009; Singh and Irvine, 2012). The *Drosophila* retina is a highly organized structure that is regulated by morphogenetic events, cell proliferation, growth, and developmental cell death (Cagan and Ready, 1989). We have used this rationale to study how the Hippo pathway affects growth during this developmental event. The Hippo pathway is required for a fundamental process of organ size regulation for growth and development. Upon activation, a core kinase cascade of the Hippo pathway is sequentially phosphorylated and results in the formation of Hippo (Hpo)-Salvador (Sav) complex. The Hpo-Sav complex activates downstream targets Warts (Wts) and Mats (Mob as a tumor suppressor). This results in the formation of the Wts-Mats complex that further activates the transcriptional co-activator Yorkie (Yki) leading to cytoplasmic retention and degradation of the 14-3-3

adaptor protein. Alternatively, when Hpo is not activate, Yki remains unphosphorylated and translocates into the nucleus to bind to transcription factor Scalloped (Sd) to induce target genes (Zhao *et al*., 2007; Snigdha *et al*., 2019; Oh and Irvine, 2008; Kango-Singh and Singh, 2009; Huang *et al*., 2005; Gogia *et al*., 2021). Therefore, we use the highly accessible *Drosophila* pupal retina to study cell growth during eye development.

The *Drosophila* retina is composed of nearly 16000-20000 cells. It has nearly 750-800 hexagonal units made up of ommatidial units consisting of 8 photoreceptor neurons, 4 lens secreting cone cells, and 2 primary pigment cells. There are also pigment-producing lattice cells and sensory bristle groups that surround these ommatidial units. These retinal cell fates are determined by several positional and genetic cues during development. It is also recently reported that the Hippo pathway has a role in R8 cell fate decisions in the post-mitotic stage of eye development. Apoptosis and proliferation are highly regulated during organ formation (Udan *et al*., 2003). It is shown that activation of hippo pathway promotes apoptosis during development, whereas its inactivation promotes proliferation. It has been reported that Hpo restricts the cell number during organ development (Udan *et al*., 2003). Programmed cell death (PCD) helps in refining the structure of organs during development. During PCD, excess cells are removed leaving the essential neighboring cells. During pupal retina development, extra interommatidial cells are removed to form hexagonal lattice structure. We have developed this lab activity to introduce genetics, molecular technique, and microscopy to undergraduate students.

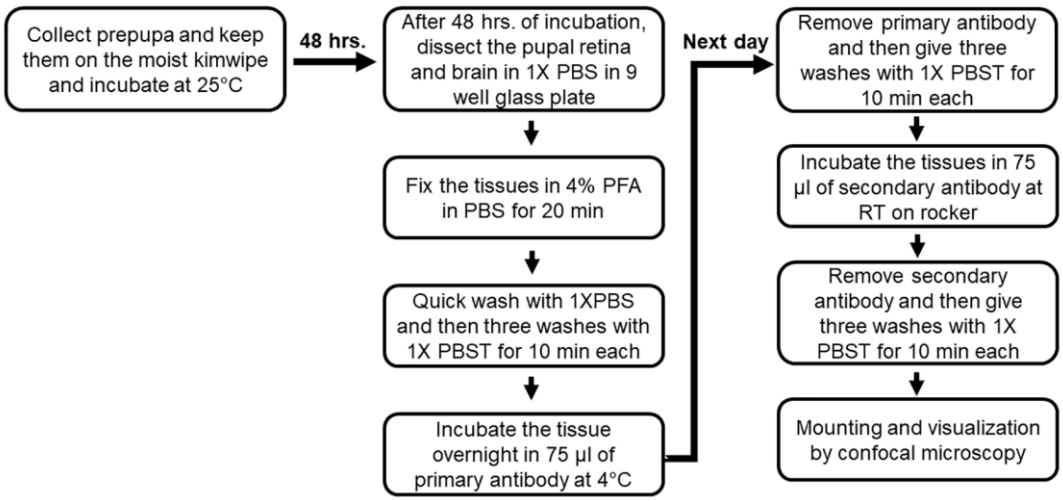


Figure 1. Flowsheet diagram of the strategy employed for *Drosophila melanogaster* pupal retina dissection and immunohistochemical staining. This protocol is intended to introduce students to study growth using a *Drosophila* model by using immunohistochemistry approach.

Protocol

The entire protocol can be divided into four major steps: (1) Tissue preparation, (2) Pupal retina dissection, (3) Antibody staining, (4) Mounting and Imaging (Figure 1).

1. *Tissue preparation*

We used the pupal retina of *Drosophila melanogaster* to study growth and development. We employed the Gal4/UAS system for targeted misexpression of transgenes (Brand and Perrimon, 1993). For all our genetic crosses, we have used the GMR-Gal4 driver line, which expresses the Gal4 transcriptional activator under the control of Glass Multiple Repeat (GMR) promoter in the posterior region of the morphogenetic furrow in larval eye disc and throughout the pupal development (Moses and Rubin, 1991). GMR-Gal4 served as a control in this lab exercise. In order to study organ growth and development, we

modulated members of the Hippo pathway and compared their pupal retina. We set up *Drosophila* crosses for gain of-function and loss-of-function of Hippo pathway members and maintained the cultures at 25°C. Early white pre-pupae for each genotype were selected and kept on a moist kim-wipe in petri plate (Figures 1, 2). Please note that moisture helps keep the prepupae from desiccation. The prepupae of each genotype were labeled, and the date and time were also noted. The prepupae were incubated at 25°C for 48 hours until dissection (Figures 1, 2).

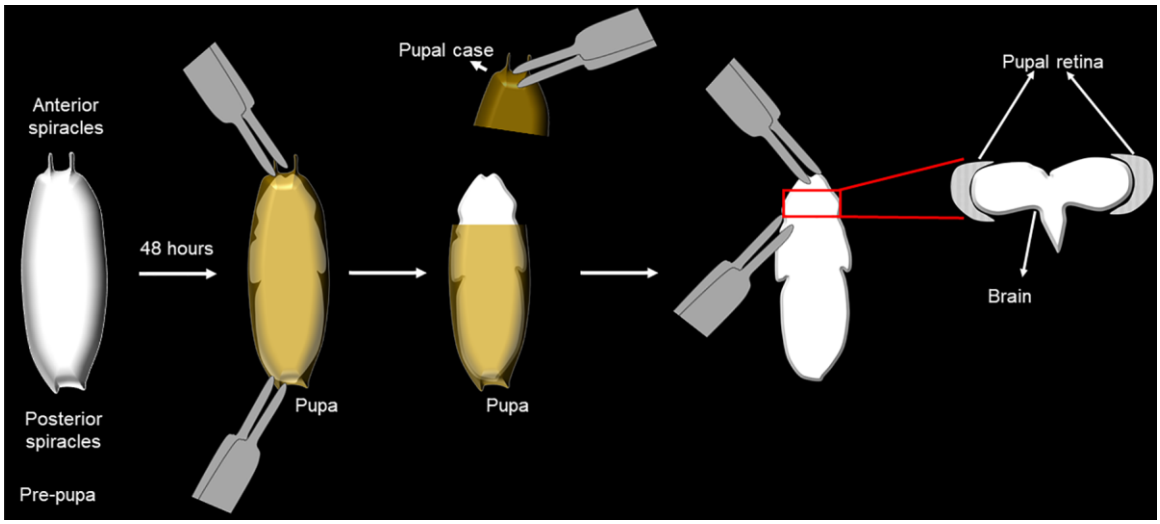


Figure 2. Schematics of *Drosophila melanogaster* pupal retinal dissection from 48-hour old pupae. Cartoon shows the placement of forceps to remove pupal case and cuticle and the resulting retina-brain complex.

2. *Pupal retina dissection*

Each well of the nine well glass plate was filled with 150-200 µl of cold 1× phosphate buffered saline (PBS) to make sure that the tissues are always submerged in buffer. After 48 hours of pre-pupae collection, pupae of every genotype were kept in different wells. The operculum of each pupa was removed by using the forceps (Figure 2). The abdomen of pupa was pierced with one forceps to grasp and the other forceps was used to prick the anterior of the pupa, and it was shaken until the head and thorax were visible (Figure 1, 2). The white fat tissue was separated by using two forceps until two retinae and brain were exposed. Note that the retina-brain complex looks dumbbell-shaped, off-white and translucent than the surrounding fat tissue (Figure 2). The retinae-brain complexes of each genotype were kept separately in different wells. Note that the dissection time for each genotype should be 15 min to avoid variability in the phenotype. The retinae-brain complexes were then fixed in 4% paraformaldehyde in 1× PBS for 20 min. Note that the tissue is submerged in the fixative to ensure proper fixation. After 20 minutes, fixative was replaced with 1× PBS. The retinae-brain complexes were then washed by 400 µl of 1× PBS for 10 minutes on rotator at room temperature to remove any traces of the fixative. The tissues were then washed by 400 µl of 1× phosphate buffered saline-TritonX-100 (PBST) for 10 minutes each. Note that dissection, washing, and antibody staining was done in a nine well plate. Precautions need to be taken while washing the retinae; there is a chance of losing samples while pipetting.

3. *Antibody staining*

The retinae-brain complexes were then stained by primary antibodies rat anti-Embryonic lethal abnormal vision (Elav) (1:100; Developmental Studies Hybridoma Bank, DSHB), and mouse anti-Discs-large (Dlg) (1:100; DSHB). The antibodies were diluted in 1× phosphate buffered saline- TritonX-100-Normal

donkey serum (PBT-NDS). The retinae-brain complexes were incubated in primary antibodies overnight at 4°C. Next day, primary antibodies were removed, and the tissues were washed with 1× PBST thrice for 10 min each. The tissues were then incubated in secondary antibodies (Jackson Laboratory): goat anti-rat IgG conjugated with Cy5 (1:250), and donkey anti-mouse IgG conjugated with Cy3 (1:250) for two hours on the rocker. After two hours of incubation, the secondary antibodies were then removed and then washed with 1× PBST thrice for 10 min each (Figure 2).

4. *Mounting and Imaging*

After three washes, the retinae-brain complexes were transferred onto a microscopic slide using forceps along with a few drops of 1× PBST. The retinae were separated from the brain using forceps and tungsten needles. The brain and other tissues were removed, and the retina is properly mounted on the slide. Make sure that the retina should not be folded else the imaging will not be proper. The PBST was removed and Vectashield was added to the sample without drying the samples. Gently slide the coverslip on the samples and the edges were sealed with transparent nail polish. Images were obtained by the Fluoview 3000 confocal microscope. If the teaching assistant/instructor time schedule does not allow imaging, the slides can be stored at -20°C in a microscope slide box in the dark. The immunofluorescent images were captured at 60× magnification by using Olympus Fluoview 3000 Laser Scanning Confocal Microscope (Figure 2). All final figures were prepared using Adobe Photoshop software.

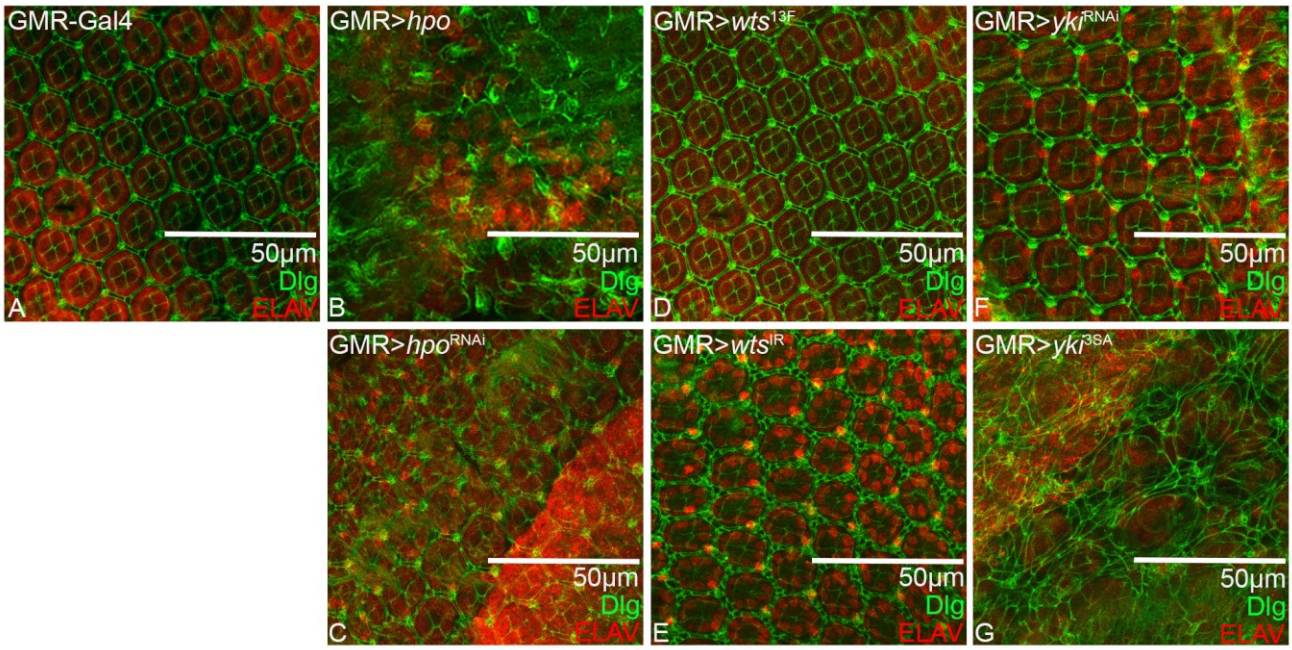


Figure 3. Modulation of Hippo pathway results in differences in arrangement of ommatidial cells. (A-G) Pupal retinal of 48-hour old pupae stained for membrane specific marker Dlg (green) and pan-neuronal marker, Elav (red). (A) GMR-Gal4 serves as a control. (B-G) Modulation of Hippo pathway in the Glass multiple repeat (GMR) domain. Pupal retinal showing activation of Hippo pathway (B) GMR > *hpo*, (D) GMR > wts^{13F}, (F) GMR > yki^{RNAi}, and downregulation of Hippo pathway (C) GMR > hpo^{RNAi}, (E) GMR > wts^{IR}, (G) GMR > yki^{3SA}.

Note that the pupal retinae were stained for membrane-specific marker- discs large (Dlg, green) and pan neural marker- embryonic lethal abnormal vision (Elav, red) (Figure 3). The control GMR-Gal4 shows

hexagonal arrangement of ommatidia and monolayer of secondary and tertiary pigment cells (Figure 3 A). The activation of Hippo pathway- GMR > *hpo*, GMR > wts^{13F}, and GMR > yki^{RNAi} does not show much difference in the number of cells in pupal retina (Figure 3 B, D and F). In contrast, inactivation of Hippo pathway- GMR > hpo^{RNAi}, GMR > wts^{IR}, and GMR > yki^{3SA} show multiple layers of secondary and tertiary pigment cells resulting in extra interommatidial cells suggesting proliferation of cells (Figure 3 C, E and G). This experiment allows students to understand the basic cell biology technique like immunohistochemistry and confocal microscopy. This exercise also helps them to understand the basic concept of cell biology- apoptosis, growth, and proliferation.

Table 1. List of reagents and slides used for pupal retina staining laboratory exercise.

Reagents	Composition	Storage
1X PBS	10 ml of 10X PBS dissolved in 90 ml of autoclaved water	Freshly prepared. 4°C
1X PBST	10 ml of 10X PBS dissolved in 90 ml of autoclaved water+ 0.2% TritonX-100	Freshly prepared. 4°C
1X PBT-NDS	1 XPBST+ 10% Normal donkey serum	Freshly prepared. -20°C
4% Paraformaldehyde	100 µl of 16% Paraformaldehyde dissolved in 300 µl of 1X PBS	Freshly prepared.
Vectashield H-1000 (Vector Labs)	commercial preparation, ready to use	4°C in dark
mDlg	commercial preparation, ready to use	-20°C
RtElav	commercial preparation, ready to use	-20°C
mCy3	commercial preparation, ready to use	-20°C in dark
RtCy5	commercial preparation, ready to use	-20°C in dark
Slides		
Microscope Slides (Fisher Scientific Catalogue No:12-5442)		
Cover Glass (Fisher Scientific Catalogue No:12-548-C)		

Advantages

1. This lab is cost effective. The reagents and glassware (Table 1) used for the pupal retina staining are commercially available and are inexpensive.

2. This hands-on experiment helps in better understanding of the concepts of this cell biology technique rather than in lecture.

3. Cellular process like growth and development can be understood through this immunohistochemistry technique.

4. This lab exercise gives the glimpse of how to approach biology research questions by this simple cell biology technique. Undergraduate students inclined towards research appreciate this lab.

Conclusion

Generally, commercial kits are used for the undergraduate lab exercises which just introduce cell biology technique to the students. These kits do not emphasize on posing a research question, designing an experiment to address the questions, the principle behind using specific reagents and buffers and their constitution and analyzing the data. This lab exercise helps us in addressing a research question of how modulating the Hippo signaling pathway affects organ growth and development using pupal retina as an example. The undergraduate lab exercises (Deshpande *et al*., 2020; Gogia *et al*., 2017; Mehta and Singh, 2017; Tare and Singh, 2008; Tare and Singh, 2009; Puli and Singh, 2011) including this one demonstrates basic cell biology technique and lab skills.

Acknowledgments: This laboratory exercise was designed in the Department of Biology, at the University of Dayton. PD, ACV are supported by the graduate program of the University of Dayton. AS is supported by NIH1R15GM124654-01 and Schuellein Endowed Chair in Biology

References: Brand, A.H., and N. Perrimon 1993, Development 118**:** 401-15; Cagan, R.L., and D.F. Ready 1989, Dev. Biol. 136: 346-62; Deshpande, P., A.V. Chimata, and A. Singh 2020, Dros. Inf. Serv. 103**:** 99-104; Deshpande, P., N. Gogia, A.V. Chimata, and A. Singh 2021, Biotechniques 71**:** 416-424; Gautam, D.K., A.V. Chimata, R.K. Gutti, and I. Paddibhatla 2021, J. Cell Physiol. 236**:** 5592-5619; Gogia, N., A.V. Chimata, P. Deshpande, A. Singh, and A. Singh 2021, Neural Regen. Res. 16**:** 643-652; Gogia, N., A. Sarkar, and A. Singh 2017, Dros. Inf. Serv. 100: 218-225; Huang, J., S. Wu, J. Barrera, K. Matthews, and D. Pan 2005, Cell 122**:** 421-34; Kango-Singh, M., and A. Singh 2009, Developmental Dynamics 238**:** 1627-1637; Mehta, A.S., and A. Singh 2017, Dros. Inf. Serv. 100**:** 225-230; Moses, K., and G.M. Rubin 1991, Genes Dev. 5**:** 583-93; Oh, H., and K.D. Irvine 2008, Development 135**:** 1081-8; Puli, O.R., and A. Singh 2011, Dros. Inf. Serv. 94: 190-196; Sarkar, A., M. Irwin, A. Singh, M. Riccetti, and A. Singh 2016, Neural Regen. Res. 11**:** 693-7; Singh, A., and K.D. Irvine 2012, Dev. Dyn. 241: 1-2; Snigdha, K., K.S. Gangwani, G.V. Lapalikar, A. Singh, and M. Kango-Singh 2019, Front. Cell Dev. Biol. 7: 85; Tare, M., and A. Singh 2008, Dros. Inf. Serv. 91: 160-164; Tare, M., and A. Singh 2009, Dros. Inf. Serv. 91**:** 174-180; Udan, R.S., M. Kango-Singh, R. Nolo, C. Tao, and G. Halder 2003, Nature Cell Biology 5**:** 914-920; Usman, M., and A. Singh 2011, Journal of STEM Education: Innovations and Research, 12; Wang, J.T.H., 2017, FEMS Microbiol. Lett. 364; Zhao, B., X. Wei, W. Li, R.S. Udan, Q. Yang, J. Kim, J. Xie, T. Ikenoue, J. Yu, L. Li, P. Zheng, K. Ye, A. Chinnaiyan, G. Halder, Z.C. Lai, and K.L. Guan 2007, Genes Dev. 21: 2747-61.

DIS Printed Copies

Printed copies of this and other recent issues of Drosophila Information Service can be ordered directly from an independent publisher:

www.lulu.com

On that site, go to “Shop” and enter “Drosophila Information Service” in the search bar. Several issues should be shown. The cost of the printed volume is now lower than previously; they will add shipping costs based on delivery location.

Other Reports

The North American *Drosophila* Board

The Board's duties include: overseeing community resource centers and addressing other research and resource issues that affect the entire *Drosophila* research community. The Board also administers the finances for the annual North America *Drosophila* Research Conference and its associated awards, and it chooses the organizers and the site of the annual meeting. The Board consists of nine regional representatives and four international representatives, who serve 3-year terms. The three elected officers are President, President-Elect, and Treasurer. The three most recent Presidents continue participation on the Board as Past-President. In addition, the Board has *ex officio* members who represent *Drosophila* community resources or centers. For more information about the Board and the summaries of the annual Board meetings, see the FlyBase web site.

Drosophila Board

President: Tin Tin Su

Past-President: Mariana Wolfner (2021)
Past-President: Mark Peifer (2020)
Past-President: Bruce Edgar (2019)

Treasurer: Jessica Treisman

Regional Representatives:

New England: Alexey Veraksa
Heartland: Tânia Reis
Midwest: Rachel Smith-Bolton
Mid-Atlantic: Erika Bach
Southeast: Wu Min Deng
California: Leanne Jones
Great Lakes: Brian Lazzaro
Mountain: Nadia Singh
Canada: Liz Rideout

International Representatives:

Asia: Tatsushi Igaki
Australia/Oceania: Kieran Harvey
Europe: Nic Tapon
Latin America: Helena Araujo

Primarily Undergraduate Institution Representative

Justin DiAngelo

Postdoc and Student Representatives

Lydia Grmai	Postdoc Representative	Johns Hopkins University
Ana-Maria Raicu	Student Representative	Michigan State University

***Ex Officio* – Representing *Drosophila* Resources:**

Norbert Perrimon (FlyBase; Harvard Medical School)
Susan Russo Gelbart (FlyBase; Harvard University)
Brian Calvi (FlyBase; Indiana University)
Susan Celniker (BDGP; Lawrence Berkeley National Laboratory, Berkeley)
Kevin Cook (Bloomington Stock Center & Nomenclature Committee; Indiana University)
Cale Whitworth (Bloomington Stock Center)
Patrick O'Grady (Drosophila Species Stock Center; Cornell University)
Jim Thompson (DIS, Drosophila Information Service; University of Oklahoma)
Hugo Bellen (Bloomington Stock Center Advisory Committee & P Element Project; Baylor College of Medicine)
Allan Spradling (P-Element Project; HHMI/Carnegie Institute)
Stephanie Mohr (Director, DRSC/TRiP; Harvard University)
Jonathan Zirin (Assistant Director, DRSC/TRiP; Harvard University)
Scott Hawley (Nomenclature Committee; Stowers Institute for Medical Research)
Lisa Meadows (VDRC; Vienna, Austria)
Masanobu Itoh (DGRC, Kyoto; Kyoto, Japan)
Toshiyuki Takano-Shimizu (DGRC, Kyoto; Kyoto, Japan)
Chuck Langley (At-large; University of California, Davis)
Brian Oliver (FlyBase Advisory Board; NIH)
Andrew Zelhof (DGRC, Bloomington)

Genetics Society of America:

Hugo Bellen, GSA Board of Directors
Suzy Brown, Senior Director
Tracy DePellegrin, Executive Director