Determining the size of the male-specific region in the genome of the scuttle fly, *Megaselia scalaris,* a potential model system for the earliest stages of sex chromosome evolution

Ralitza H. Peneva

under the supervision of Dr. Mohamed Noor
Biology Department, Duke University

April, 2011

Honors Thesis submitted in partial fulfillment of the requirements for graduation with Distinction in Biology in Trinity College of Duke University.

Abstract

Modern day whole-genome sequencing and the ability to make comparisons across many taxa have significantly advanced the study of sex chromosome evolution. The scuttle fly, *Megaselia scalaris*, is an appropriate model system for studying sex chromosome evolution exhibits sex chromosomes that are homomorphic in size, containing differences that are difficult to detect even microscopically. Presumably, sex in *M. scalaris* is determined by the presence or absence of a male-determining region, or sex realizer (M), that transposes among chromosomes at a low rate, essentially creating novel Y-chromosomes out of autosomes. *M. scalaris* would thus serve as a good model for exploring the primary stages of sex chromosome evolution. However, to serve as such a model, the identity and size of the male-specific region must be identified as a boundary within which to search for the sex realizer. This study focused on elucidating the relative size of the male-specific region in *M. scalaris* in relation to the genome. I compared previously generated genome sequences from male and female *M. scalaris* to isolate suggested regions unique to the male sex. Through a systematic approach involving sequencing, primer design, and PCR, I found the *M. scalaris* genome to be 27% male-specific, a preliminary result suggesting a large size inconsistent with published reports. This size further questions observations pointing to a relatively small male-determining region that transposes, but the male-specific region may be composed of additional elements unique to the male sex that do not ultimately determine sex. This study has refined the boundaries within which the sex realizer (M) of *M. scalaris* may be located. Future studies are necessary to pinpoint the sex realizer, an important step towards *M. scalaris'* use as a model system in sex chromosome evolution.

Introduction

Sex determination mechanisms are astoundingly diverse in animals. The two main mechanisms are 1) genetically based, as in the X and Y sex chromosome system well studied across mammals, and 2) environmentally based, as in the temperature-dependent systems found in many fish and reptiles (Wilson and Makova, 2009). Genetic mechanisms are further subject to variability, including haplo-diploidy (males develop from a haploid egg while females develop from a diploid egg) and the sex-specific chromosome systems of male heterogamety (XX females and XY males) and female heterogamety (ZZ males and ZW females). Haplodiploidy occurs in arthropods, such as mites and ticks, as well as members of the order Hymenoptera, encompassing ants, bees, and wasps (Beukeboom, 2005). Male heterogamety is notable among eutherian mammals whereas female heterogamety is common in birds, butterflies and many fish (Barske and Capel, 2008; Mank, 2009; Wilson and Makova, 2009). In some cases, several different mechanisms can exist within a single species, as in the case of the lizard, *Bassiana duperreyi* or in the salamander, *Pleurodeles poireti* (Wilson and Makova, 2009; Radder *et al.*, 2008; Dournan *et al.*, 1990).

Sex chromosomes are the prevalent mode of sex determination in animal species that feature separate sexes (Traut, 2010). Hence, understanding the manner in which this unique chromosome pair evolved is important to unraveling the details behind its mechanism of sex determination. For the purposes of this study, we will follow Charlesworth *et al.*'s (2005) definition of sex chromosomes as those that carry the genes controlling male and female development. In a karyotype image, sex chromosomes are often easy to identify due to a high dimorphism in size, particularly across mammals and most birds (Graves, 2006). Early karyotype analysis led to the original discovery of X and Y as a pair of unequally sized

chromosomes (reviewed in Charlesworth *et al.*, 2005). For instance, the X in mammals and the Z in chickens are both highly gene rich sex chromosomes while their respective homologues, Y (for X) and W (for Z), contain only a fraction of that genetic material and pair over a small homologous region, termed the pseudoautosomal region (Graves, 2006; Fridolfsson *et al.*, 1998).

Early observations suggested that heteromorphic sex chromosomes evolved from an early pair of homologous autosomes in both XY and ZW systems (Muller, 1914; Ohno, 1967). Sex chromosomes have arisen independently in multiple clades, including snakes, birds, and mammals (Matsubara *et al.,* 2006). In order for homologous chromosomes to begin differentiation, two characteristics are vital: 1) a sex-determining function and 2) crossover suppression (reviewed in Bull, 1983; Kaiser and Bachtrog, 2010). Heteromorphic chromosomes presumably arose from homomorphic chromosomes when one of the homologues acquired a gene with a sex-determining function and, consequently, a heterozygous chromosomal region was formed. Over a long enough time period, a non-recombining region formed around the sex-determining locus of the early stage sex chromosomes and was favored by natural selection, preserving the linkage of genes involved (Charlesworth *et al.*, 2005).

Understanding the mechanism behind sex determination at the earliest stages in its evolution is key to unraveling the origins of sex chromosomes. Yet studies are typically limited to species in which the sex chromosomes have already undergone a significant degree of differentiation. To understand the mechanisms behind sex chromosome evolution and to have a system in which to study evolutionary genetics questions, we are beginning to utilize a model system that lacks fully differentiated sex chromosomes.

The scuttle fly, *Megaselia scalaris* (Figure 1), is a warm-climate cosmopolitan scavenger that can develop in an astonishing variety of materials, including living or decaying plant and

animal matter, bacterial cultures, feces, paint, and shoe polish (Disney, 2008; Varney and Noor, 2010). *M. scalaris* is omnivorous and, among insects, consumes the widest range of organic materials, including essentially anything that stems from a living organism (Disney, 2008; Varney and Noor, 2010). This species plays an important role in human health and forensics, where its affinity towards human corpses at early stages of decay as well as implications in human myiasis, or the invasion of tissues and organs by dipterous fly larvae (Disney, 2008; Varney and Noor, 2010).

M. scalaris features an interesting mode of sex determination in which its sex chromosomes (referred to colloquially as X and Y throughout this paper) are homomorphic in size: no detectable differences can be seen between these sex chromosomes in a karyotype at the cytogenic level (Figure 2) (Traut *et al.*, 1990). However, detectable differences exist between the X and the Y at the molecular level, where it has been shown that the Y chromosome features a conserved segment that is absent in the X (Willhoeft and Traut, 1990; Traut and Wollert, 1998; Traut, 2010). Crossover fails to occur at the conserved segment, referred to in the literature as the male-determining factor, the sex realizer, or the *Maleness* (M) factor (Willhoeft and Traut, 1990). Because sex chromosome evolution typically proceeds from homomorphic chromosomes to highly heteromorphic chromosomes, a low but detectable degree of difference between X and Y would be characteristic of the earliest stages of sex chromosome differentiation.

Figure 1: *Megaselia scalaris* male (Photo by Eric Spana)

In some species of the insect order Diptera, sex determining linkage groups are not fixed, so that different chromosome pairs can serve as the sex chromosomes within various species (Bull, 1983). This pattern has been observed in *M. scalaris* as well, where there are a total of three chromosome pairs, all of which are homomorphic in both males and females (Mainx, 1964; Johnson *et al.,* 1988). In early studies, the novel mode of sex determination in *M. scalaris* was termed 'alternating sex determination,' in which the male sex was shown to be determined by an epistatically-mediated 'sex-realizer' or *Maleness* (M) factor located at a terminal position on the presumed Y chromosome (Mainx, 1964). Mainx (1964) noticed that the sex-realizer (M) regularly exchanged between *M. scalaris'* three non-homologous chromosomes and suggested that a translocation process was responsible. Mainx's theory was later expanded upon by the proposal that the mechanism through which M readily exchanges between chromosome pairs involves a transposable element that is small relative to the overall genome size (Green, 1980; Traut and Willhoeft, 1990).

However, before its potential use as an evolutionary genetics model, it is necessary to have a working estimation of the approximate size of the putative male-specific region in *M.*

scalaris. This region contains a high divergence between the X and Y-chromosomes and presumably within its boundaries lays the male-determining region, but there are likely numerous additional non-male-determining sequences that are still specific to males (M. Noor, unpublished). The importance of this estimation does not lie in its precision, but rather in determining whether the male-specific region of *M. scalaris* is relatively large (e.g., at or near the size of a full chromosome) or small, which thus far remains unclear. The Noor laboratory began to survey microsatellite variability in crosses among two inbred strains of *M. scalaris* and found that 2 out of 11 microsatellite sequences appeared to have male-specific alleles, suggesting a male-determining region comparable to the size of a whole chromosome. Thus, I hypothesize that *M. scalaris* features large male-specific region and may be much further along in the process of sex-chromosome evolution than previously inferred. Numerous studies pose that the male-determining region in *M. scalaris* is transposing frequently, which would suggest that any male-specific region would necessarily also be small (Willhoeft and Traut, 1990; Traut and Willhoeft, 1990; Traut and Willhoeft, 1998). This study aims to quantify the size of the male-specific region in *M. scalaris* to establish a baseline understanding of the mode of sex determination in this species and to determine its merits to serve as a model system for the earliest stages of sex chromosome evolution.
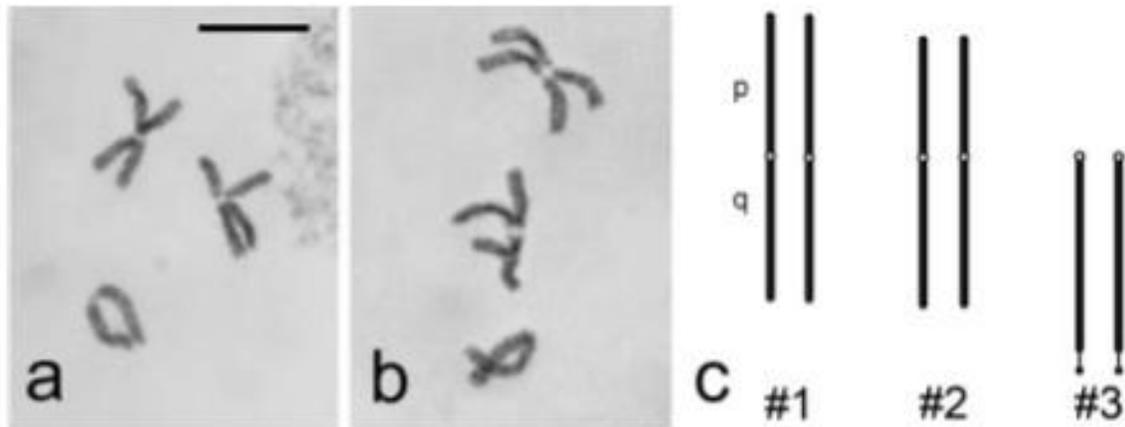
Figure 2: Karyotype depicting mitotic somatically paired chromosomes from neuroblast cells of *Megaselia scalaris*. (a) female, (b) male, (c) diagrammatic representation of *Megaselia scalaris* chromosomes showing that #1 and #2 are metacentric (centromere in the center) while #3 is acrocentric (centromere at one end). Adapted from (Traut *et al*., 1990; Traut, 2010).

Methods

Identifying genes of interest

Our lab is currently in the process of assembling the *M. scalaris* genome *de novo*. Studies have shown that even low-coverage 454 genome sequencing is a sufficient source of genetic information and microsatellite markers (Rasmussen and Noor, 2009). Since that pilot study, the Noor laboratory has obtained extensive Illumina® sequences from virgin adult males and virgin adult females. Previous to my work, Dr. Robin Varney within the laboratory began to assemble these Illumina® reads into large contigs, or contiguous sequences of overlapping DNA segments. Male and female contigs were then BLASTed (Altschul *et al.* 1990) against each other for comparison, and all uniquely male sequences were identified. Uniquely male sequences totaled approximately 50,000 base pairs at an 80% similar contingency setting. These sequences were then BLASTed (Altschul *et al.* 1990) to the National Center for Biotechnology

Information (NCBI) GenBank database, resulting in approximately 1300 top hits. As a starting point, hits were further filtered for *Drosophila* hits or anything labeled as 'male,' 'sperm,' or 'testis.' *Drosophila melanogaster* served as the genome against which *M. scalaris* sequences were BLASTed (Altschul *et al.* 1990) due to the highly complete nature of its sequenced genome.

Primer design

The filtered uniquely male BLAST hits with similar sequences in *D. melanogaster* resulted in a total of 19 sequences ranging in size from 150 to 200 base pairs. My work focused on designing primers for these sequences by using the *Primer3* program (Rozen and Skaletsky, 2000) as a starting point and manually modifying suggested primer sequences via the Integrated DNA Technologies *Oligo Analyzer* program (http://www.idtdna.com/analyzer/Applications/OligoAnalyzer/Default.aspx). Primers were designed to be no less than 18 base pairs and no greater than 40 base pairs with a melting temperature near 58°C. Only 12 of the 19 original sequences yielded suitable primers and were subsequently used in this study.

Polymerase chain reactions (PCRs)

I prepared DNA squish preps on *M. scalaris* males and virgin females for PCR following a standard Drosophila protocol in order to extract their DNA (Gloor and Engels, 1992). The PCR recipe was 2.5μl 10X PCR buffer (15 mM MgCl$_2$), 2.5 μl 2 mM dNTPs, 1.25 μl of 10 μM primer (forward and reverse each), 0.6 μl DNA polymerase, 15.9 μl H$_2$O and 1 μl fly DNA squish prep. All PCRs were run on BioRad MyCycler thermal cyclers following a protocol of 1 minute at 95°C, 33x [30 seconds at 94°C, 30 seconds at 50°C, 30 seconds at 72°C], and 1 second

at 40°C. PCR samples (1µl PCR sample, 2µl 6x loading dye, 9µl ddH$_2$O) were run on a 1% agarose gel using a HyperLadder IV ladder.

Initially, I ran PCRs on 5 males and 2 virgin females based on the *M. scalaris* that eclosed at the start of this study. I concurrently ran a positive control on virgin female DNA preps to confirm a proper working order. Following PCRs on all 12 sequences, I obtained an additional 3 virgin female DNA preps to confirm male-specificity in sequences that only showed bands in males.

Results

Male-specific testing via rtPCR in Megaselia scalaris

Of the 19 sequences (Appendix A) that filtered out as uniquely male BLAST hits with similar sequences in the template genome of *Drosophila melanogaster*, 12 sequences were conducive to primer design (Table 1) while the remaining 7 contained sequences of repetitive DNA that made it unfeasible to design a reliable primer for further male-specific testing.

To further break down the sequences, 58.3% (7 sequences) of sequences showed only male bands when run on an agarose gel, as indicated by an example of a male-specific gel result in primer Msca_15 (Figure 3) and Msca_39 (Figure 4). The remaining 41.7% (5 sequences) of sequences featured bands for both males and females on an agarose gel, as indicated by an example of a gel result for primers Msca_31 and Msca_34 (Figure 4).

Table 1: Summary of male-specific testing via rtPCR in *Megaselia scalaris* based on primers for sequences obtained by BLASTing against male-specific regions in the Drosophila genome.

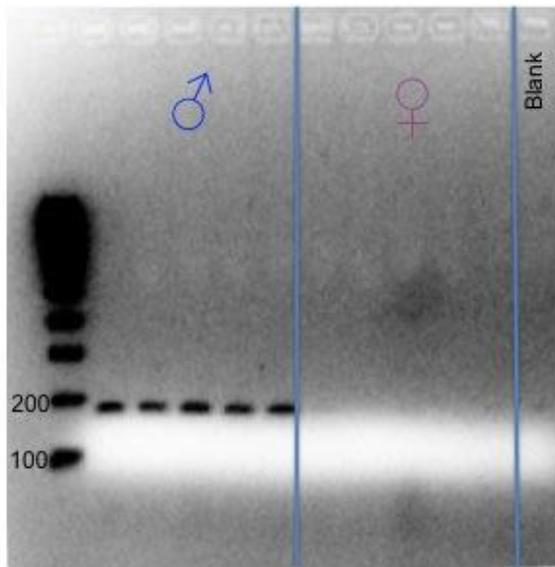| Primer Name	Product Size (base pairs)	Gel Result
Msca_15	176	**Male**
Msca_19	138	**Male**
Msca_25	158	Male + Female
Msca_31	174	Male + Female
Msca_34	191	Male + Female
Msca_39	158	**Male**
Msca_40	227	**Male**
Msca_50	249	Male + Female
Msca_11	115	**Male**
Msca_29	168	**Male**
Msca_45	213	Male + Female
Msca_54	204	**Male**

Figure 3: Male-specific testing via rtPCR in *Megaselia scalaris* for primer Msca_15. The 176 base pair product derived from primer Msca_15 appears only in male *M. scalaris.* Primers were designed for sequences that BLASTed against male-specific regions in *Drosophila melanogaster*. ♂ N=5, ♀ N=5.
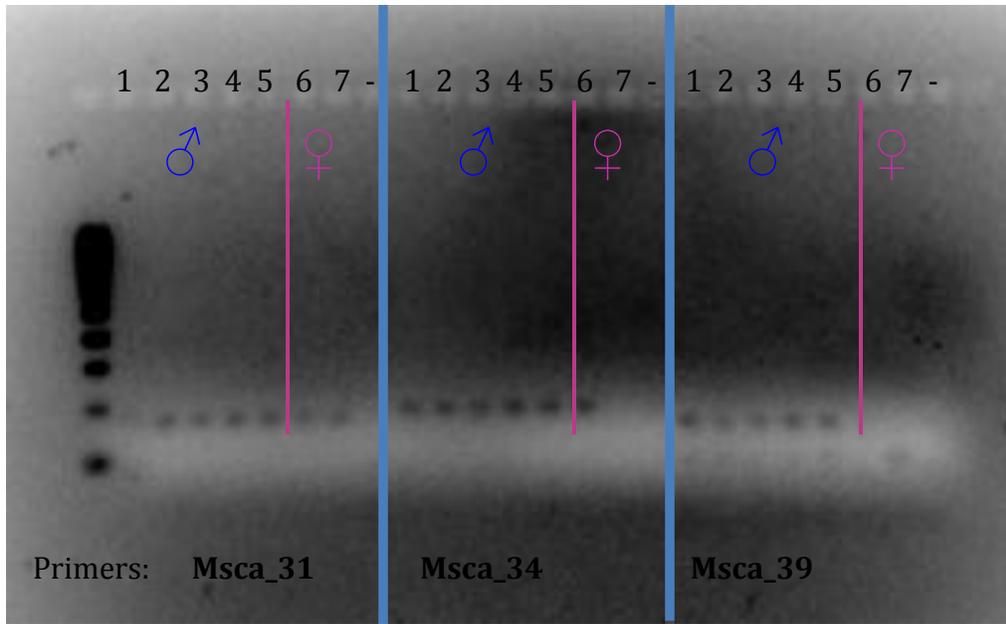
Figure 4: Male-specific testing via rtPCR in *Megaselia scalaris* indicates male-specificity in the sequence from primer Msca_039 and lack of specificity for Msca_031 and Msca_034. Primers were designed for sequences that BLASTed against male-specific regions in Drosophila. For each primer: ♂ N=5, ♀ N=2.   All 8th columns are blank controls.

Calculation suggests a relatively large male-determining region in M. scalaris

The *M. scalaris* genome is approximately 500 megabases in size (Rasmussen and Noor, 2009; Spencer Johnston, Texas A&M University, personal communication). In calculating the relative size of the male-determining region in this species, the total female base pairs assembled amounted to 560,113,614, while the total male base pairs from contigs, contiguous sequences of overlapping DNA segments (>150 base pairs) amounted to 262,597,289. Taking the ratio of male to female base pairs, we ended up with 46.9%, or roughly half male-specific. By finding male-specificity in 7 out of the 12 sequences filtered out as uniquely male BLAST hits, we essentially calculated the error rate on the 46.9% figure to be roughly half (58.3%). Hence,

incorporating the error rate onto 46.9% yields a percentage of roughly 27% male-specificity in the *M. scalaris* genome.

Discussion

Studies of sex chromosome evolution have been generally limited to the use of model species which feature highly differentiated sex chromosomes. A widely accepted and tested theory regarding sex chromosome evolution poses that the sex chromosomes began as a pair of equally sized autosomes that differentiated over time (Charlesworth *et al.*, 2005; Toups and Hahn, 2010). This study aimed to examine the genetics of an interesting case of undifferentiated sex chromosomes in the scuttle fly, *M. scalaris.*

An important question that needed to be addressed regarding the genetics of *M. scalaris* was the size of its male-specific region relative to the genome. As Traut (2010) suggests, the sex chromosome turnover observed in *M. scalaris* is its most conspicuous feature. It has been repeatedly suggested that the reason behind sex chromosome turnover in *M. scalaris* is that the male-determining region is functioning as a transposable element that jumps around between the 3 homologous chromosome pairs, essentially creating novel Y chromosomes (Willhoeft and Traut, 1990; Traut and Willhoeft, 1990). For an element to function as a transposable element, its size relative to the genome must be small, by definition. Widely distributed full-sized transposable elements can range in their maximum size, including 10,000 base pairs in humans, 14,000 base pairs in *Drosophila*, and 20,000 base pairs in both the nematode, *Caenorhabditis elegans,* and in maize, *Zea mays* (Eickbush and Eickbush, 2005; Kidwell, 2002). Thus, across a variety of species, the maximum reported size of a transposable element is 20,000 base pairs,

which would be small relative to the approximately 500, 000, 000 base pair total genome size of *M. scalaris*.

In this study, however, we found that the size of the male-specific region relative to the *M. scalaris* genome is 27.35%, or over one quarter of the species' genome, suggesting a relatively large size. This finding supports my study's hypothesis of a relatively large male-specific region, which was based on earlier work completed in the Noor lab in which a survey of microsatellite variability in a cross between inbred strains of *M. scalaris* yielded 2 microsatellite sequences that appeared to have male-specific alleles out of a total of 11 sequences. This microsatellite survey suggested that 18.2% of the genome might be implicated in the male-determining region. Additionally, in a study conducted on a DNA region encompassing the *Maleness* locus in *M. scalaris*, Traut and Wollhoeft (1998) noted a lack of recombination between the X and Y chromosomes in the region of interest, a finding that is consistent with the male-determining region of *M. scalaris* being potentially large relative to its genome size.

A relatively large male-specific region with respect to the *M. scalaris* genome seems to contradict a number of previous studies done in *M. scalaris,* all of which suggest that a tranposition event is responsible for the high turnover rate of the sex chromosomes (Willhoeft and Traut, 1990; Traut and Willhoeft, 1990, Traut, 2010). In addition to studies in *M. scalaris*, a member of the family Chironomidae, studies conducted in a different chironomid, the midge *Chironomus thummi thummi*, suggest that its similar jumping male-determining region is small (Kraemer and Schmidt, 1993).

The turnover of the jumping male-determining factor in *M. scalaris* in its movement from the Y chromosome to autosomes has been found at the detectable frequencies of 0.06% and 0.3% (Traut and Willhoeft, 1990). However, while this has largely been interpreted to be due to a

transposition event (Willhoeft and Traut, 1990; Traut and Willhoeft, 1990), Kraemer and Schmidt (1993) suggest that the transposable element theory does not seem to have been proven regarding the male sex determiner (M) in *M. scalaris*. Kraemer and Schmidt (1993) suggest that even in strains of *M. scalaris* in which the male sex determiner (M) had jumped to create a novel Y chromosome, none of the molecular markers of the wild-type Y chromosomes moved together with the putative M. Essentially, in *M. scalaris*, there is a distinction between the male-specific region and the male-determining region, or sex realizer, in that the sex realizer lies within the male-specific region. However, a relatively small male-determining region (sex realizer) would translate to a relatively small male-specific region since the transposition of the sex realizer causes that original male-specific region to become non-sex-specific.

Given that my study suggests a relatively large male-specific region in *M. scalaris*, it is unlikely that a transposable event is responsible for sex chromsosome turnover. The widely accepted theory of sex chromosome evolution poses that highly differentiated sex chromosomes originally evolved from a pair of equally sized autosomes (Muller, 1914; Ohno, 1967). One of the homologous autosomes acquired a sex-determining function and crossover around the region was suppressed over a long enough period of time to form a non-recombining region that was favored by natural selection (review in Bull, 1983; Charlesworth *et al.*, 2005; Kaiser and Bachtrog, 2010). Hence, based on the finding of a relatively large male-specific region, *M. scalaris* may be further along in the evolution of its sex chromosomes than previously thought (Willhoeft and Traut, 1990; Traut and Willhoeft, 1990, Traut, 2010). The relatively large male-specific region identified in this study essentially narrows down the possible location of the male-determining region (sex realizer) in *M. scalaris*. Future work is necessary to pinpoint the

actual sex realizer (M), presumably located within the boundaries of the male-specific region identified in this study.

Future Directions

The broader implication of developing a new model system is to study the earliest stages of sex chromosome evolution as we strive to decipher the mechanisms behind sex determination. *M. scalaris* remains an ideal candidate to serve as such a model system primarily because of its undifferentiated sex chromosomes. Additionally, this species can easily be reared in laboratory conditions on virtually any media and features a rapid generation turnover, allowing for experimental manipulation (Disney, 2008). Given the small size of the sequences analyzed in this study, it may be appropriate to modify future methods to utilize a Li-cor automated DNA analyzer to allow for a greater degree of accuracy in distinguishing the product band size from a possible primer front or primer dimmers.

Having a new model system in *M. scalaris* further allows for its use as an outgroup species to compare to the distantly related genus *Drosophila* (Figure 5) in asking other evolutionary genetics questions, including the association between crossover rate and nucleotide diversity within species and nucleotide divergence between species. Additionally, as a globally distributed nuisance species that is associated with human myiasis and invades human and animal corpses, studies of this species' genetics would add to the understanding of its overall biology and allow for possible advancements in the current forensic application of *M. scalaris* as well as in targeted improvements in global human health by a prevention of the invasion of human tissues, or myiasis.
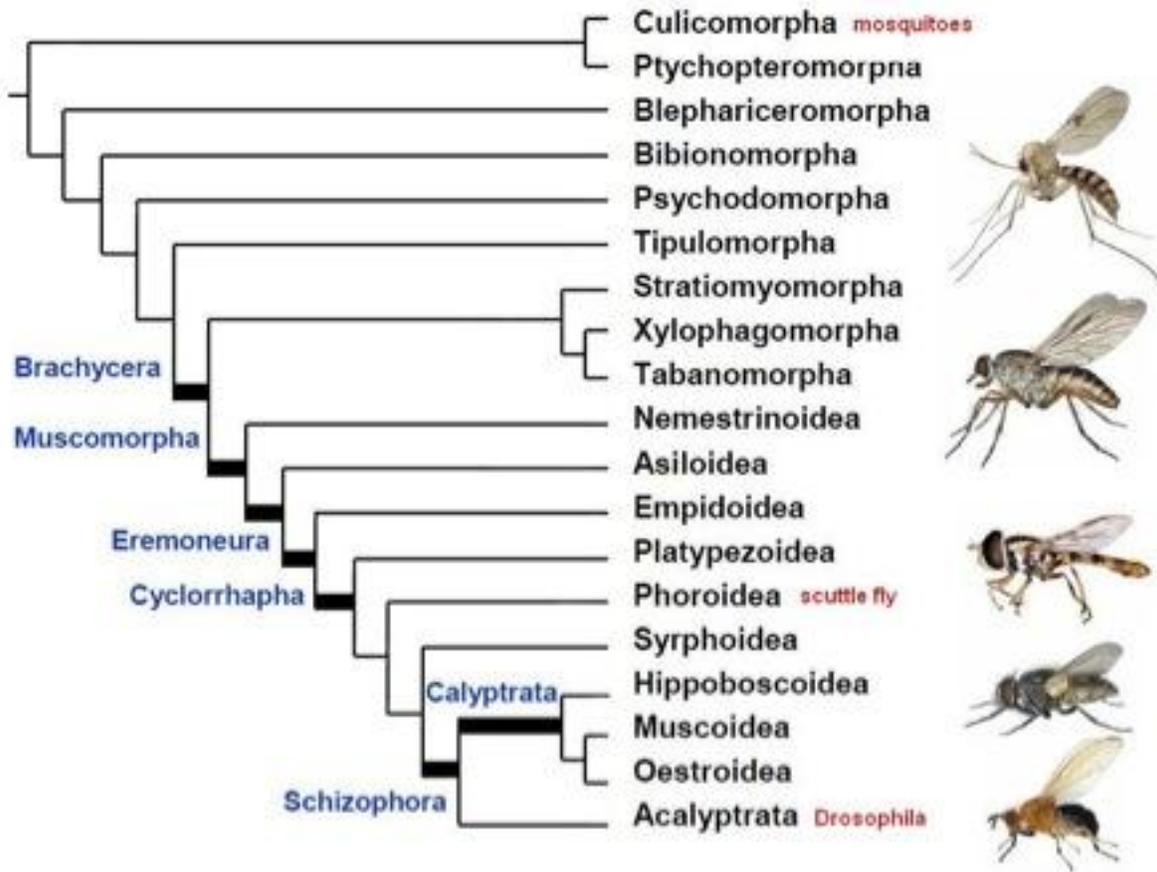
Figure 5: Phylogeny of the order Diptera, with *Megaselia scalaris* (scuttle fly), Drosophila, and the mosquitoes (Drosophila's former outgroup) noted in red. Adapted by Mohamed Noor from Yeates *et al.* 2007.

Acknowledgements

References

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. 1990. Basic local alignment search tool. Journal of Molecular Biology 215: 403-410.

Barske, L. A. and Capel, B. 2008. Blurring the edges in vertebrate sex determination. Current Opinion in Genetics & Development 18: 499-505.

Beukeboom, L. W. 2005. Sex determination in Hymenoptera: a need for genetic and molecular studies. BioEssays 17: 813-817.

Bull, J. J. 1983. Evolution of sex determining mechanisms. The Benjamin-Cummings Publishing Company, Menlo Park, California.

Charlesworth, D., Charlesworth, B., and Marais, G. 2005. Steps in the evolution of heteromorphic sex chromosomes. Heredity 95: 118-128.

Disney, R. H. L. 2008. Natural history of the scuttle fly, *Megaselia scalaris*. Annual Review of Entomology 53: 39-60.

Dournan, C., Houillon, C., Pieau, C. 1990. Temperature sex-reversal in amphibians and reptiles. International Journal of Developmental Biology 34: 81-92.

Eickbush, T. H. and Eickbush, D. G. 2005. Transposable elements: evolution. In: ENCYCLOPEDIA OF LIFE SCIENCES. John Wiley & Sons Ltd, Chichester. http://www.els.net/

Fridolfsson, , A. K., Cheng, H., Copeland, N. G., Jenkins, N. A., Liu, H. E., Raudsepp, T., Woodage, T., Chowdhary, B., Hawverson, J., and Ellegren, H. 1998. Evolution of the avian sex chromosomes from an ancestral pair of autosomes. Proc. Natl. acad. Sci USA 95: 8147-8152.

Gloor, G. B. and Engels, W. R. 1992. Single-fly DNA preps for PCR. Drosoph. Inf. Serv. 71: 148-149.

Graves, J. A. M. 2006. Sex chromosome specialization and degeneration in mammals. Cell 124: 901-914.

Green, M. M. 1980. Transposable elements in *Drosophila* and other Diptera. Annual Review of Genetics 14: 109-120.

Johnson, D. S., Mertl, H. G., Traut, W. 1988. Inheritance of cytogenetic and new genetic markers in *Megaselia scalaris*, a fly with an unusual sex determining mechanism. Genetica 77:159-170.

Kaiser, V. B. and Bachtrog, D. 2010. Evolution of sex chromosomes in insects. Annual Review of Genetics 44: 91-112.

Kidwell, M. G. 2002. Transposable elements and the evolution of genome size in eukaryotes. Genetic 115: 49-63.

Kraemer, C. and Schmidt, E. R. 1993. The sex determining region of *Chironomus thummi* is associated with highly repetitive DNA and transposable elements. Chromosoma 102: 553-562.

Mainx, F. 1964. The genetics of *Megaselia scalaris* Loew (Phoridae): a new type of sex determination in diptera. The American Naturalist 98: 415-429.

Mank, J. E. 2009. The W, X, Y, and Z of sex-chromosome dosage compensation. Trends in Genetics 25: 226-233.

Matsubara, K., Tarui, H., Toriba, M., Yamada, K., Nishida-Umehara, C., Agata, K., and Matsuda, Y. 2006. Evidence for different origin of sex chromosomes in snakes,birds, and mammals and step-wise differentiation of snake sex chromosomes. PNAS 103(48): 18190-18195.

Muller, H. J. 1914. A gene for the fourth chromosome of Drosophila. Journal of Experimental Zoology 17: 325-336.

Ohno, S. 1967. Sex chromosomes and sex-linked genes. Springer, Berlin.

Radder R.S., Quinn A.E., Georges A., Sarre S.D., Shine R. 2008. Genetic evidence for co occurrence of chromosomal and thermal sex-determining systems in a lizard. Biology Letters 4:176-178.

Rasmussen, D. A. and Noor, M. A. F. 2009. What can you do with 0.1x genome coverage? A case study based on a genome survey of the scuttle fly *Megaselia scalaris* (Phoridae). Genomics 10.

Rozen, S. and Skaletsky, H. J. 2000. Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S (eds) *Bioinformatics Methods and Protocols*: *Methods in Molecular Biology*. Humana Press, Totowa, NJ, pp 365-386.

Toups, M. A. and Hahn, M. W. 2010. Retrogenes reveal the direction of sex-chromosome evolution in mosquitoes. Genetics 186: 763-766.

Traut, W., Khuong, N. and Schneider, S. 1990. Karyotypes of *Megaselia scalaris* (Diptera) wild-type and translocation strains. Genetica 83: 77-84.

Traut, W. and Willhoeft, U. 1990. A jumping sex determining factor in the fly *Megaselia scalaris*. Chromosoma 99: 407-412.

Traut, W. and Willhoeft, U. 1998. An X/Y DNA segment from an early stage of sex chromosome differentiation in the fly *Megaselia scalaris*. Genome 41: 289-294.

Traut, W. 2010. New Y chromosomes and early stages of sex chromosome differentiation: sex determination in *Megaselia*. Journal of Genetics 89: 307-313.

Varney, R. L. and Noor, M. A. F. 2010. The scuttle fly. Current Biology 20: R466-R467.

Willhoeft, U. and Traut, W. 1990. Molecular differentiation of the homomorphic sex chromosomes in *Megaselia scalaris* (Diptera) detected by random DNA probes. Chromosoma 99: 237-242.

Wilson E. B. 1911. The sex chromosomes. *Archiv für mikroskopische Anatomie* 77, 249-271.

Wilson, M. A. and Makova, K. D. 2009. Evolution and survival on eutherian sex chromosomes. PloS Genetics 5.

Yeates, D. K. Wiegmann, B. M., Courtney, G. W., Meier, R., Lambkin, C., and Pape, T. 2007. Phylogeny and systematics of Diptera: two decades of progress and prospects. Linnaeus Tercentenary: Progress in Invertebrate Taxonomy. Zootaxa 1668: 565-590.

Male *Megaselia scalaris* sequences that BLASTed to *Drosophila melanogaster* introns.

>200 base pair length

"Msca_15"
>NODE_1535172_length_236_cov_4.605932
TATATTTATATAGGATTTTGTTTTAAAGGGGTAATTTCATCTTTAATTTCATAGAAAC
TCTTATTGTCGATAGATAACAAATAGTCCTTATTAGAGTAATCACGCACAGTCAATC
AGAGAACAAATGACAACCAAATAGCTGCTTCCTCAGAGAACATAGTGCTGGACACT
GACAGATCCCTTCCATATTTTCAATATTTCAATATTTGCTACCTGAATCTTCTGTACT
TCAAGTGAAGTCAACCTTATAGCTGACTTTGCTGCCACGTATTTGACGT

"Msca_19"
>NODE_1994881_length_403_cov_3.310174
AATACAATATTCATTTCATAAAGAAATAAAAATTCAATAAATTTATGAGTAATTTAC
ACTTGAAATTTACTAATTTTAAAAATATATTCTTTTCAAGAAATTTGAAATTTGTACA
ATTTAACACAAATTCAAAAAAAATATTCTTTATATTCTTTAAAGTTTAATTTTTTGTG
TAATTTTAATTTTGTTTGCCCTTACCCTAAAATGAGATTATTGGATAAAAGCGGTACA
CTTTCTAAGGAAAGCCTTTCTGTATATACTAGATTCCTTGAGTTTTGAGGTTTAGAGA
CGTGTCTGAGAGACAAGCGAATTCCAGGAATAATTTTTAAATACATTTCCAAGGAAC
AAGGGTTAAATTGATTGCTTGTTATTTTGTACCAACTAGGGAACACAGAGCAGGGAA
CCCTCTGGATTTCTGGGAAAGCACAGCACTCTTCGTAATTTT

"Msca_25"
>NODE_2517990_length_261_cov_4.287356
AGTTTCCAGACTAAATGAGAATAATGTTCTACTCCTGTCAAGAACCACCACGGGACT
CTTATGTTTACAATTTTTCAGAAGAAAGAGATCAAGATAATTGGTCAAAATTTTTTA
CTTTTGTTGGTAAAAAACGTATGTATGTAGATGTAAACGTGTGTATATAATGTATGT
ACGTGACAAACATAACATCTATGAACAATTTTAATGAAATTTATAAACTCAGTAGTC
AAAACATATGTACCTTATACGTATTACATGAATAAGTCATATACAAGTGTCATACCA
TATTTTAAATTATCATAA

"Msca_31"
>NODE_3141248_length_203_cov_3.896552
TAAAATGTCCATCAACATTCTCAAATATCAACAAATTTGTCAAAAATAAATAAATAT
TTTGTCGCCACTTAAAATTAAAATAGAATTTAAAAGAAAGTTTAATACAAAATCAA
CAAGTGTCTCCAATTTTAAGTAGTTTGTTTAGATAGTAATTGTAAATTCGCCACCATT
AACGAATCGCCGCCGCCTTCCGTCTTTTTATAGCCAATAAAAACAAATATATATTTA
TATTTATATTGACCAT

"Msca_34"
>NODE_3405497_length_288_cov_2.993056
TATTTGCATTTCTATTATACATAGTTTGATGATTCTGACATCTTAAAATTATAAATTA
TTTTAAAAAGATAGCTTTTGTGAAATTTTCAAATATCTATGATGTTTTTGAAAGATTT
TGACTGATTTCGGAGATTTCAATATTAATGTGTGGAATTGTGTACAGAATTTTGGAC

TTTTACTCGATAGTAGAATTTCGTATTTTGAAAGTTTTGTATTGTGTTTCATTTCTGAA
ACGAAATTTGACCTGCTAGTAGAAGTTCGCCCATGGGCGAAAGTATGTTATATCCTC
TTGTAGGGAACGAGACGGCCGAAATCTTGCGGAAAATCTTT

"Msca_39"
>NODE_3918001_length_206_cov_3.145631
TATTAATTACTTCCTTTTCATTTTTTAAATATTACCTATGCTCACAATTTTGATGTAAA
CTATATGGTTAAAGGAGCAGTGAATATTCCACTGGTCCAGAATTCCCTCATTGATAA
GTTTCAGGAGTCCTGAGGTAATCAACTTCACAAAGGCTACGTATTCCATATAATGT
GGTCATAGGAGTAGGGAATTATGCACCAGTTCAGAAACTGCATTTTTAATATTTTTC
TGACTCCTGAAGTCACAG

"Msca_40"
>NODE_4049546_length_252_cov_3.769841
TAACCAAGTGGTTCCCTTACGGTAACCAATCGGCATCCTTTGTGAAGAGAAGAAGA
GTCTCTAGTAGCAATTTCCCCAACCCTTGTAAATGGAGACACTATCCACCAACTCTA
ACACATATGGTATGAGATCCAGAAACTCCAGCAGATGACAGTTCTTAAGTTAAAATT
TTCTATTTCAAAAATAAAGCTTTATCCTCTACAGGAACCTAGATAGAAACCAGCAAG
GAAGTTTTATAGATAGAAACCATCAAGGAAGATTTATCGATAACACCGAAGATGTC
TCCCTAAAATC

"Msca_50"
>NODE_5061934_length_260_cov_2.350000
AAACCGACTAAGCTTTGCAATCTTAAATTAGTTTATAAATTATTTTTCATATTATTTG
TGTGATATTAAATTAATGTTTCTTTTACATTTTTTTTTTTATTTTTAGCATAAGACTAT
CAACCCTATTCTCAAAGAAAACAGGATGCTCTTTCTTGTGATAAAAAATCTTTATAA
TTATTTAATAAAATAATTTTATTACCGTAAGGATGAATTTAATCTTTAAGGGAGAGC
CTTAGCGTATTTAATCTTTAAAAAATACTAAAAATTTGTTAATAATAAAAATATATTT
TGTTAATATCTTT

150-200 base pair length

"Msca_11"
>NODE_1113019_length_166_cov_1.987952
ATTTTAAATGATCAAACTTTTAAAGATTTCTAAATATTTCAAATTTTAATATTTGATT
ACAATTATCGAAAATTTGTTAGTTACTGAGGTTATGATATTGCAGATACCGCTATCT
AGATGCTGCCAAATAGGAAGCCAATATTCAAAAGGGAGGTCATTATGAAGCAGGCA
AAGTTGAAAAAGAGACGCTTTAAAAGCAATTATTTAA

"Msca_29"
>NODE_2914165_length_153_cov_2.202614
TGACATTGAAGGAATGGATCAGGCGAAACTTTAATTTATTTTAAATTACTGACAAAT
ATGACAGAAATTGACAAATGTATTAATAAAAAGACCAATGTTAAGGAAAGATTA
ATTCTTAAAAAATCAGTAACTGGTGCAATCTAACCAAATAATCACACACAGAGGGA
CGAGTAGTAGGCAAAAAGTTCATAAT

"Msca_45"
>NODE_4501464_length_189_cov_1.952381
TTAAAATGGGCGTGGTTATGAGCGGAGTCAACAAGCCCACTTTTTGTTTTCAATATA
TAAATTTATTATAAAATGGGCGTGTTTATGGTCGGAGTCCACATACCCACTTTTTTAA
TTTATAAATTAATTGTAAAAGGGCGTGGTTATCGGCGGAGTCGACAAACCAAAGTTT
TATATTTTTAAATATATAAATAAATTTAAAAAGAGGCTTGGTTATGGGAGGAGTAGT
CA

"Primer Msca_54"
>NODE_5498058_length_163_cov_3.036810
CTCCAAAATTGAGTGATTCAATGTTTTCTGTTTAAAATCTCATGTTTAGTTACTTTTG
AGTGATTAGAAAATAATATTGAGTGACTTTGCTTTAATCTCATTTTGAGCTACTTTTC
AATTACTTACAATATTTTTCAATATTTTAATATCAAAGTTACTTAAAAAAATATGA
GTAACTTGTCAGTCAGATTTGAATAATGTATT

<p align="center">Sequences not conducive to primer design:</p>

>200 base pair length

>NODE_2798296_length_225_cov_3.946667
TTGATAAATTTTATTGTTCTTTTTTAAGAGTGAATGAAACCATTTTTATTGAAAATAA
CTTTGAATATTTATGTTGGAATTCTGCTATAAATTGTTTTTGTTTTGAGATTTTTTTA
AATTTTTGTTTGTAATGGAATGTTTAATCTTGTATTTAATTTGGTCGTTGATGATAAA
CAAAATAAAAAATATTTATAATAAAGATACATGGATAGTCCCTTTGAAATTAATAAA
TATTTAGTATACAAATATAAAAATAAAAAAGTATA

150-200 base pair length

>NODE_169303_length_153_cov_2.117647
CAGTTGATCCGGGTTCGATTCCCGGCCGACGCACAATTTTTTCTTTTTATCAAATTTT
TTGATTAATAAAGTATTATTTTTATTTGTAATTTTTTATGTATTTGTTTTTATTGTTGA
TAGAATTGAATTACAATTCAAAAAAGAGTAAAATTATTACAAAGTTTTTAAAAATAA
AACATTAAAACAAACTTACCT

>NODE_1496097_length_156_cov_2.339744
TACTTTTGTGCATTTAAATAAAAAAAATTATTTATATTTAAATTATTTAATGTTTTCC
TAGATAGTGGGAGTTTTTTATAAGCACAGTGATTATTTATAATATTAGTAAAAGTTTT
TATTTTATCTTTTGTAACTATAATTCTCTCTCTGAGAGGCTCAAATTTCAGTGATG
ATGAGAAACATAAAAGCTATATGT

>NODE_1854610_length_169_cov_4.189349
TCCATCGATTCTGAATGAAATCAGAATCGAAGAGTAGTTTTGGGTGTTTTTTAACT
GCTTTGTAAAAAACATTTTTATTTTATATATAGCATTACAAAGTAAAATTTTATGTAT
ATCAAAAAATATAAACTAAAATTGTTTTTTACAAAGCCGTTAAAAAAAAGCACACAT
ACATATAAAGTAAATGTTAAAAACTGTTTATTTGGAAACA

>NODE_2854307_length_191_cov_2.078534
CTTGTATATATTTTCTTAAATTAAAAACTTGATTTCTTAAAAGTTATTATTTATTGTTT
TATATTATATATTTTCATTTTATCAAAAGTAAAACTATCTTAAAATTTTTAACATTTT
CTAAAATTATTACATTGTTTTGTTAAATATTTTATTTACTATTGATGGTAATTAAAGA
TGGAATACTAACGAAAATATTAATATTATTATTAAAATATTTTGTTGAACATTAAAA
A

>NODE_4515587_length_190_cov_2.184211
ATTAATAAAAAAACTTAAAATCATTAATTAAAAACTAAAACAGATAAAGTCAATGA
AATAAAAACATATTGTAAAAACTAAACAAATATAATTTTGATATGAAACAAATTTTA
TGAGAATTGTAAAATTTAAAATTAAATATATTGATATACATGGATGGGAATTTATAC
GTAAATTATGTAAGTAATTAACGTAAGTTAAAGCATTAAATTTCGAAAGATAGTAAA
TAAAA

>NODE_6007388_length_159_cov_1.433962
GAAAAATACTATATTATTAAAACTAAATATTTAAAAATATAATATAATTTTAAAAAT
TAAAACTAAAAAATAATAAAATATTAATAACTACAAATATATAAATTTAAATTTCCA
CTTCAGAATTTTTATAAGCTTATGTTTTAATAATGAAAAAAAAAACACTAAATAATA
TGAAAATGTTGTTTTTAAAATGAATAATTT