**The Probability of Fusions Joining Sex Chromosomes and Autosomes**

Nathan W. Anderson12, Carl E. Hjelmen1, and Heath Blackmon13

1Department of Biology; Texas A&M University; College Station, TX 77843, USA

2Department of Integrative Biology; University of Wisconsin; Madison WI 53706, USA

3Author for correspondence: HB, [*coleoguy@gmail.com*](mailto:coleoguy@gmail.com)

*Keywords: sexual antagonism; chromosome fusion; sex determination systems; chromosome number*

# Abstract

Chromosome fusion and fission are primary mechanisms of karyotype evolution. In particular, the fusion of a sex chromosome and an autosome has been proposed as a mechanism to resolve intralocus sexual antagonism. If sexual antagonism is common throughout the genome, we should expect to see an excess of fusions that join sex chromosomes and autosomes. Here, we present a null model that provides the probability of a sex chromosome autosome fusion, assuming all chromosomes have an equal probability of being involved in a fusion. This closed-form expression is applicable to both male and female heterogametic sex chromosome systems and can accommodate unequal proportions of fusions originating in males and females. We find that over 25% of all chromosomal fusions are expected to join a sex chromosome and an autosome whenever the diploid autosome count is fewer than 16, regardless of sex chromosome system. We also demonstrate the utility of our model by analyzing two contrasting empirical datasets: one from *Drosophila* and one from the jumping spider genus *Habronattus*. We find that in the case of *Habronattus* there is a significant excess of sex chromosome autosome fusions but that in *Drosophila* there are far fewer sex chromosome autosome fusions than would be expected under our null model.

# Introduction

The fusion and fission of chromosomes are two of the primary mechanisms that restructure the genome into discrete chromosomes [1]. Fusions and fissions can be selectively favored in genome restructuring because they modify linkage among loci [2, 3]. In particular, the fusion of a sex chromosome and an autosome (SA-fusion) has been proposed to resolve sexual antagonism (when an allele is beneficial for one sex and deleterious for the other) [4]. Linking sexually antagonistic alleles to sex chromosomes can increase the average fitness of both sexes. Therefore, SA fusions are predicted to be more common than fusions joining two autosomes (AA-fusions) [5]. For example , an apparent surplus in X chromosome autosome fusions in the jumping spider genus, *Habronattus*, is hypothesized to result from a mechanism of isolating male-beneficial sexually antagonistic alleles on the neo-Y chromosome [6]. However, most evidence for sexually antagonistic variation comes from within species. For instance, empirical studies in fish, human, and flies have found evidence for segregating sexually antagonistic variation (variation with opposite fitness effects in males and females) [7-9]. Furthermore, young sex chromosomes (originating through fusion, translocation, or turnover) exhibit signs consistent with the resolution of sexual antagonism [10-12]. However, there remains significant debate on the ubiquity of sexually antagonistic variation and its potential role in genome evolution [13-15]. A strong measure of the frequency of significant sexually antagonistic variation across genomes would be an excess of SA-fusions relative to AA-fusions across large clades. We derive equations describing the probability of each type of fusion necessary to perform such a test and illustrate two approaches to using these equations with empirical datasets. This approach will provide a method to quantitatively analyze the balance of these two types of fusions in the many groups with a well-documented history of fusion between sex chromosomes and autosomes [3, 16-18].

# The Model

The probability of SA-fusions is a function of the sex chromosome system and the number of autosomes in the genome. To facilitate tests of the balance between SA-fusions and AA-fusions, we have derived a closed form expression of the probability of a SA-fusion under a null model where any chromosome is equally likely to fuse with any other non-homologous chromosome. Our result is applicable to XO, XY and multi-XY (e.g. X1X2Y or X1X2X3Y1Y2) sex chromosome systems and, with slight modification, to ZW and UV systems [reviewed in 19]. We ignore fusions among homologous chromosomes, including fusions that join an X and Y chromosome, because this would lead to unbalanced gametes during meiosis and, presumably, these would be non-viable.

When any two chromosomes fuse, there are three possibilities. The two chromosomes could both be autosomes (AA-fusion), they could both be sex chromosomes (SS-fusion), or one could be a sex chromosome and the other an autosome (SA-fusion). We denote our three possibilities as events , , and , respectively. Given that a fusion has occurred, we are interested in the probability it is an SA-fusion. Or, equivalently, we are interested in the expected proportion of all fusions which are SA-fusions. Unfortunately, this proves difficult to calculate directly. We avoid this using the complement rule. We define the probability that any given fusion is a SA-fusion as:

(1)

It is quite possible that the sexes may make unequal contributions to the fusions entering a species (Pennell et al. 2015). We include the term , representing the proportion of fusions that occur in females to account for this possibility. We use a subscript and for sire and dam when referring to sex specific values. In the supplemental material we provide a detailed derivation but, briefly, we present the following expression for the expected proportion of fusions which occur between two sex chromosomes (equation 2) or two autosomes (equation 3):

(2)

(3)

where is X chromosome count in males, is female diploid number, is male diploid number, is Y chromosome count in males, and is diploid autosome count.

Each fraction represents the probability of two types of chromosomes fusing using a counting argument. For instance, the faction in equation 3 represents the probability of a fusion joining two autosomes in a male.

Substituting equations 2 and 3 into equation 1 yields:

(4)

This equation allows us to calculate the proportion of fusions that should be SA-fusions for any XY sex chromosome system with any number of autosomes (figure 1). Equations 2, 3, and 4 have six parameters: , , , , and . We avoid the parameter, the number of X chromosomes in females, by noting . This formulation can be converted for use in ZW sex chromosome systems by exchanging and , replacing with , replacing with , and replacing with . Additionally, setting (because there are no homogametic diploid individuals) and replacing both and Y for V generates equations which are accurate for UV sex chromosome systems, in cases where there are an equal number of U and V chromosomes. We have provided R functions that calculate P(SA), P(SS), and P(AA) in the R package evobiR [20]. When using these functions, the only arguments required are the diploid autosome number, the sex chromosome system, and which is assumed to be 50% by default.

**Empirical Applications**

*Habronattus:* In a recent study of *Habronattus* jumping spiders, the large disparity between the number of SA-fusions (8-15) and AA-fusion (1) and SS-fusions (1) in a system with approximately 26 autosomes is presented as evidence that SA-fusions are being favored [6]. The intuition that this imbalance is unlikely can be rigorously tested with our null model that the distribution of fusions is determined by chromosome number and sex chromosome system. Using equations 2-4, and a multinomial distribution, we are able to calculate the exact empirical p-value of having observed eight or more SA-fusions out of a total of 10 fusions. We assume an XXO sex chromosome system and a diploid autosome count of 26 (this karyotype was the most common in the ancestral state estimation performed in the study). . This confirms that *Habronattus* spiders do in fact have an excess of SA-fusions.

*Drosophila*: In the previous example, we calculated the expected proportion of the different types of fusions based on the ancestral, and most common, karyotype inferred in the *Habronattus* clade. However, during the evolution of the *Habronattus* clade, a variety of karyotypes must have existed. We envision the primary use of equation 4 will be to calculate the expected proportion of fusions that are SA-fusions across large clades. We can do this by employing a biologically realistic Markov model of possible fusions and fissions (figure 2A), and leveraging stochastic mappings (figure 2B) [21, 22] generated under such a model to extract the proportion of time that lineages in a clade spent with each possible chromosome number and sex chromosome system. These proportions can then be used in conjunction with equation 4 to generate a weighted sum that describes the expected proportion of all observed fusions that are SA-fusions (figure 2C). The resulting expected value can then be compared to the observed proportion of SA-fusions inferred from the stochastic mappings. An additional advantage of this approach is that it can incorporate uncertainty in both ancestral state reconstructions and phylogenetic history. To illustrate this approach, we analyzed data from *Drosophila*.

In the analysis below we used a dated ultrametric phylogeny of *Drosophila* from a recent study of genome size evolution [23]. We used chromosome number and sex chromosome system data from the same paper and a comparative genomic analysis of 11 *Drosophila* species [24]. This yielded a dataset consisting of 120 species with a diploid number ranging from 6 to 12. The sex chromosome system of eleven of the species was neoXY (term used to describe a karyotype where a sex chromosome and autosome have fused forming a larger “new” sex chromosome) while the remainder were XY. We performed 1000 stochastic mappings using the make.simmap function and extracted the time spent in each state along the phylogeny and the number of each type of fusion using the describe.simmap function both in the R package phytools [22]. Stochastic mapping was accomplished using a transition matrix that matches the Markov model presented in figure 2A. We performed a number of preliminary analyses where we assessed the impact of: 1) the prior placed on the root of the tree and 2) the inclusion of (figure 2A). We found that our results were qualitatively the same under all evaluated conditions. The results we present are based on fixing the root of the tree with a diploid number of 12 and an XY sex chromosome system, and including the parameter in the model. The prior on the root of the tree is supported by comparative genomic studies [24]. The inclusion of is based on our concern that some species may harbor an undocumented neoXY. We found that including in the model elevated our estimate of the proportion of SA fusions but not sufficiently to change the interpretation of the results.

Across our 1000 stochastic mappings we find that the average number of SA-fusions observed is 4.49 and that this equates to a proportion of 0.155 (credible interval 0.12 – 0.22). Using our formula as described above we also calculated the expected proportion of SA-fusions. The mean expected SA-fusion proportion was 0.43 (credible interval 0.42-0.44). Comparing these distributions (figure S2) we find that they have zero overlap and that the empirical dataset shows far fewer SA fusions than would be expected by chance.

**Discussion**

The need for a quantitative null model of the probability of SA-fusions is illustrated by examining the expected probability of SA-fusions across a range of observed chromosome numbers and sex chromosome systems. In figure 1, we show that when the autosome number is small, a large proportion of fusions are expected to be SA-fusions even under a null model which assumes they are not selectively favored. In fact, for an XY sex chromosome system the probability of a given fusion being an SA-fusion does not drop below 25% until the diploid autosome count is greater than or equal to 16. In systems with XXY sex chromosomes, the case is even more extreme. The probability of SA-fusion does not drop below 25% until the diploid autosome count is greater than 22. Therefore, evaluating the proportion of SA-fusions and determining whether there is evidence for positive selection for these fusions can only be accomplished in light of a quantitative null model which takes account of chromosome number and sex chromosome system.

Previous work examining SA- fusions in *Drosophila* have largely focused on the balance between fusions of an autosome with the X versus the Y [25]. Much of this work was done prior to the development of modern comparative approaches today and could not fully incorporate the evolution of chromosome number over the history of *Drosophila*. In our analysis, we asked how do the number of AA-fusions compare with the number of SA-fusions. Our results show that *Drosophila* have far fewer SA-fusions than would be expected if all fusions were equal.

The scarcity of SA-fusions that we document suggests that in *Drosophila*SA-fusions are more likely to have deleterious effects than fusions that join two autosomes. One explanation for apparent selection against SA-fusions in Drosophila may lie in the joint action of genome structure and lack of recombination in males (achiasmatic meiosis). In species with achiasmatic meiosis, when an SA-fusion occurs, the entire Y chromosome is immediately subject to population genetic forces (e.g., Muller's ratchet) that lead to the loss of functional genes. *Drosophila* has relatively few chromosomes such that each chromosome carries many genes. (*D. melanogaster* 43% of all genes are on autosome 3). Therefore, while an SA-fusion may initially provide a fitness benefit, the fitness benefit may quickly decay due to the “target size” for deleterious mutations on the Y chromosome precluding the fusion's fixation.

We have developed a flexible equation used to calculate the probability of SA-fusions under most sex chromosome systems. This model will allow for quantitative analyses of fusions across large clades and provide a way to test the long-standing hypothesis that SA-fusions are selectively favored for their ability to resolve sexual antagonism. In some clades where chromosome number is high (e.g. Lepidoptera and Isoptera) our model shows that SA-fusions should be rare. In these cases, several SA-fusions within a clade may well suggest that these fusions are selectively favored. However, this model also shows that for clades with very few chromosomes (e.g. Diptera and Hemiptera), we should expect many SA-fusions even if they are not selectively favored. Therefore, SA-fusions should only be considered as evidence for sexual antagonism when they occur at a higher rate than expected for the chromosome numbers and sex chromosome systems that have been present during the evolution of a clade.

**Funding**

This work was supported by National Institute of General Medical Sciences at the National Institutes of Health R35GM138098.

**Data Availability**

Code and data to reproduce this study are available via GitHub <https://github.com/coleoguy/sex-autosome-fusion>.

**Figure 1 Probability of a random fusion joining a sex chromosome and autosome.** On the vertical axis we plot the proportion SA-fusions while on the horizontal axis we plot the diploid autosome count. Sex chromosome system is indicated by color.

**Figure 2 Estimations across a phylogeny** A) Markov model for the evolution of karyotypes in Drosophila. A lineage will have chromosomes and either an XY or neoXY sex chromosome system, and can make four possible transitions: the AA-fusion, the fission of an autosome, SA-fusion, and the transition from neoXY to XY. B) A stochastic map showing one possible history for these traits. C) Calculation of the weighted expected for the clade as a whole.

[1] Blackmon, H., Justison, J., Mayrose, I. & Goldberg, E.E. 2019 Meiotic drive shapes rates of karyotype evolution in mammals. *Evolution* **73**, 511-523.

[2] Stebbins, G.L. 1971 *Chromosomal evolution in higher plants*. London, UK, Edward Arnold.

[3] White, M.J.D. 1977 *Animal cytology & evolution*. 3rd ed. Cambridge, University Press.

[4] Matsumoto, T. & Kitano, J. 2016 The intricate relationship between sexually antagonistic selection and the evolution of sex chromosome fusions. *Journal of theoretical biology* **404**, 97-108.

[5] Charlesworth, D. & Charlesworth, B. 1980 Sex differences in fitness and selection for centric fusions between sex-chromosomes and autosomes. *Genet Res* **35**, 205-214.

[6] Maddison, W.P. & Leduc-Robert, G. 2013 Multiple origins of sex chromosome fusions correlated with chiasma localization in Habronattus jumping spiders (Araneae: Salticidae). *Evolution* **67**, 2258-2272. (doi:10.1111/evo.12109).

[7] Cheng, C. & Kirkpatrick, M. 2016 Sex-specific selection and sex-biased gene expression in humans and flies. *PLoS genetics* **12**, e1006170.

[8] Innocenti, P. & Morrow, E.H. 2010 The sexually antagonistic genes of Drosophila melanogaster. *PLoS biology* **8**, e1000335.

[9] Roberts, R.B., Ser, J.R. & Kocher, T.D. 2009 Sexual conflict resolved by invasion of a novel sex determiner in Lake Malawi cichlid fishes. *Science* **326**, 998-1001.

[10] Kitano, J., Ross, J.A., Mori, S., Kume, M., Jones, F.C., Chan, Y.F., Absher, D.M., Grimwood, J., Schmutz, J. & Myers, R.M. 2009 A role for a neo-sex chromosome in stickleback speciation. *Nature* **461**, 1079-1083.

[11] Vicoso, B., Kaiser, V.B. & Bachtrog, D. 2013 Sex-biased gene expression at homomorphic sex chromosomes in emus and its implication for sex chromosome evolution. *Proc Natl Acad Sci U S A* **110**, 6453-6458.

[12] Zhou, Q. & Bachtrog, D. 2012 Sex-specific adaptation drives early sex chromosome evolution in Drosophila. *Science* **337**, 341-345.

[13] Ironside, J.E. 2010 No amicable divorce? Challenging the notion that sexual antagonism drives sex chromosome evolution. *BioEssays* **32**, 718-726.

[14] Kasimatis, K.R., Ralph, P.L. & Phillips, P.C. 2019 Limits to genomic divergence under sexually antagonistic selection. *G3: Genes, Genomes, Genetics* **9**, 3813-3824.

[15] Ponnikas, S., Sigeman, H., Abbott, J.K. & Hansson, B. 2018 Why do sex chromosomes stop recombining? *Trends in Genetics* **34**, 492-503.

[16] Blackmon, H., Ross, L. & Bachtrog, D. 2017 Sex Determination, Sex Chromosomes, and Karyotype Evolution in Insects. *J Hered* **108**, 78-93. (doi:10.1093/jhered/esw047).

[17] Traut, W., Sahara, K. & Marec, F. 2008 Sex chromosomes and sex determination in Lepidoptera. *Sexual Development* **1**, 332-346.

[18] Sylvester, T., Hjelmen, C.E., Hanrahan, S., Lenhart, P., Johnston, J.S. & Blackmon, H. 2020 Lineage-specific patterns of chromosome evolution are the rule not the exception in Polyneoptera insects. *Proceedings of the Royal Society B: Biological Sciences* **in press**.

[19] Beukeboom, L.W. & Perrin, N. 2014 *The evolution of sex determination*, Oxford University Press, USA.

[20] Blackmon, H. & Adams, R.A. 2015 EvobiR: tools for comparative analyses and teaching evolutionary biology. (<https://github.com/coleoguy/evobir>.

[21] Huelsenbeck, J.P., Nielsen, R. & Bollback, J.P. 2003 Stochastic mapping of morphological characters. *Systematic Biology* **52**, 131-158.

[22] Revell, L.J. 2012 phytools: an R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution* **3**, 217-223. (doi:10.1111/j.2041-210X.2011.00169.x).

[23] Hjelmen, C.E., Holmes, V.R., Burrus, C.G., Piron, E., Mynes, M., Garrett, M.A., Blackmon, H. & Johnston, J.S. 2020 Thoracic underreplication in Drosophila species estimates a minimum genome size and the dynamics of added DNA. *Evolution* **74**, 1423-1436.

[24] Schaeffer, S.W., Bhutkar, A., McAllister, B.F., Matsuda, M., Matzkin, L.M., O'Grady, P.M., Rohde, C., Valente, V.L., Aguadé, M. & Anderson, W.W. 2008 Polytene chromosomal maps of 11 Drosophila species: the order of genomic scaffolds inferred from genetic and physical maps. *Genetics* **179**, 1601-1655.

[25] Charlesworth, B., Coyne, J. & Barton, N. 1987 The relative rates of evolution of sex chromosomes and autosomes. *Am Nat* **130**, 113-146.