Current Biology 22, 522–525, March 20, 2012 ©2012 Elsevier Ltd All rights reserved DOI 10.1016/j.cub.2012.01.057

Report

Chromosome-Wide Gene Silencing Initiates Y Degeneration in *Drosophila*

Qi Zhou[1] and Doris Bachtrog[1,*]
[1]Department of Integrative Biology, University of California, Berkeley, Berkeley, CA 94720, USA

Summary

Y degeneration is characterized by pseudogenization of its gene content, an accumulation of repetitive DNA and transcriptional inactivation associated with changes in chromatin structure. The sequence of events leading to genetically inert Y chromosomes, however, is unknown. Does the accumulation of nonsense and missense mutations at protein-coding Y genes trigger their transcriptional downregulation, or does transcriptional silencing of genes precede and expedite the decay of Y-linked genes at the amino acid level? Here, we study patterns of gene expression of the recently formed neo-Y chromosome of *Drosophila albomicans*, which displays few signs of degeneration of protein-coding genes. We show that chromosome-wide downregulation initiates the processes of Y evolution. This implies that the massive degeneration of protein-coding genes observed at many evolving Y-chromosomes may have limited deleterious effects, and instead, decay of regulatory functions is the initial trigger reducing fitness of the Y.

Results and Discussion

Sex chromosomes originate from ordinary autosomes, and Y evolution is characterized by a continuous loss of gene function [1]. Ancient Y chromosomes are genetically inert; only few genes remain, which often have evolved male-specific functions, and Y chromosomes have, for the most part, acquired a heterochromatic chromatin structure. The *Drosophila melanogaster* Y chromosome, for example, only contains about a dozen genes all involved in spermatogenesis and is entirely composed of heterochromatin [2]. Studies of younger sex chromosomes have allowed us to make considerable progress in understanding the evolutionary and molecular processes leading to genetically inert sex chromosomes.

A powerful model system to study sex chromosome differentiation is provided by the neo-sex chromosomes of *Drosophila*. In this genus, repeated independent fusions between the ancestral sex chromosomes and autosomes have created so-called neo-sex chromosomes, causing thousands of genes to become X or Y linked. Depending on the age of the fusions, neo-sex chromosomes provide different snapshots of the extent of differentiation of heteromorphic sex chromosomes and the evolutionary and molecular processes involved. An autosome of the *D. pseudoobscura* subgroup that fused to the sex chromosomes about 15 million years (MY) ago already shows all the general features characteristic of old sex chromosomes, with almost no genes left on its completely heterochromatic neo-Y [3].

*Correspondence: dbachtrog@berkeley.edu

The neo-sex chromosomes of *D. miranda*, which were formed about 1 MY ago, have served as a model system to study aspects of the earlier stages of this transition [4, 5]. Although most of its ancestral genes are still present, the neo-Y of *D. miranda* is undergoing massive degeneration. A recent genome analysis revealed that >1,100 genes (~40%) are clearly nonfunctional on the *D. miranda* neo-Y and either harbor frameshift mutations or stop codons or are completely deleted [6]. Also, most genes on the neo-Y chromosome show substantially increased rates of amino acid evolution, compared to their former neo-X homologs [6, 7]. Accompanying this general degeneration of protein-coding genes at the amino-acid sequence level is a chromosome-wide transcriptional silencing of many neo-Y genes [6, 8]. In particular, over 40% of genes are expressed at a significantly lower level from the neo-Y compared to their neo-X homologs, and 10% of neo-Y genes are completely silenced [6]. This chromosome-wide degeneration is accompanied by an accumulation of transposable elements, and the neo-Y has already acquired a partly heterochromatic appearance [9, 10].

Thus, evolved Y chromosomes show massive degeneration at the structural level, and protein-coding genes become silenced chromosome-wide; however, the sequence of these events is unclear (Figure 1). In particular, it is unknown whether amino acid and nonsense mutations accumulate first on the neo-Y and compromise the function of protein-coding genes. The expression of a large number of maladapted and nonfunctional genes from the neo-Y can be costly and deleterious, because they might interfere with the function of their well-adapted neo-X homologs and could result in selective pressures to downregulate maladaptive neo-Y genes. Alternatively, transcriptional downregulation could precede the inactivation of protein-coding genes by missense or nonsense mutations on the neo-Y. In particular, Y chromosomes are characterized by an accumulation of repetitive DNA, which may trigger a large-scale change in chromatin structure on the neo-Y toward a heterochromatic appearance. This could result in global transcriptional downregulation of protein-coding genes on the neo-Y, even before they become maladapted or nonfunctional by mutations to their coding sequence. Protein-coding genes on the neo-Y of *D. miranda* show both, degeneration at the structural level together with transcriptional silencing, and a younger, less degenerated neo-Y system is needed to establish the sequence of events.

The very recently formed neo-sex chromosomes of *D. albomicans* provide a promising system to identify the forces responsible for initiating the loss of function at Y-linked genes [11]. In particular, this species has a neo-Y chromosome that was formed less than 0.12 million years ago [12]. Unlike the neo-Y chromosome of *D. miranda*, this younger neo-Y shows no obvious morphological signs of degeneration [13]. Whole-genome sequence analysis revealed that most protein-coding genes on the neo-Y of *D. albomicans* are still functional, with only 1%–2% of the over 4,800 genes present harboring stop codons or frame-shift mutations (compared to 0.5%–1% pseudogenes found on other chromosomes; [13]). Thus, the *D. albomicans* neo-Y might provide a unique

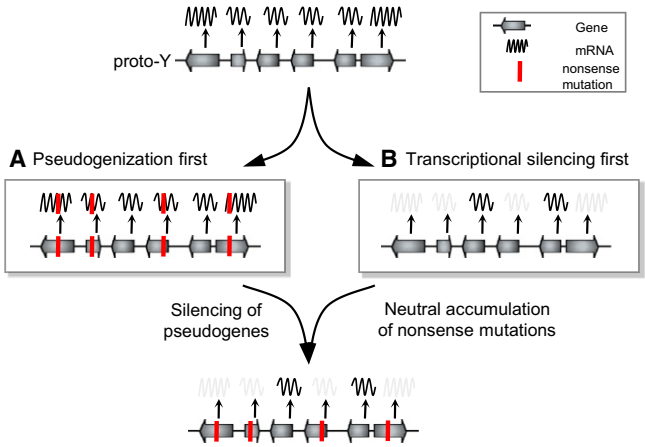

Figure 1. Pseudogenization versus Transcriptional Silencing of a Degenerating Y Chromosome

(A) Pseudogenization first: if an accumulation of nonsense or frameshift mutations initiates the loss of gene function on a Y, maladapted and nonfunctional genes will be transcribed, resulting in selective pressure to silence these pseudogenes.

(B) Transcriptional silencing first: if genes become transcriptionally silenced before accumulating nonsense mutations (for example, due to mutations in regulatory sequences or changes in chromatin structure), subsequent degeneration of amino acid sequences can occur neutrally.

opportunity to study whether gene silencing precedes the degeneration of protein-coding genes at the amino acid level.

To investigate expression patterns of the neo-sex chromosomes and compare the transcriptomes of neo-X and neo-Y linked alleles in male *D. albomicans*, we performed RNA-seq in both sexes, using the inbred strain KM55-5. To discriminate between the neo-sex alleles, we first identified putative neo-X/Y divergence sites by their male specificity and having genomic read count ratios in males of around 1 (0.75∼1.25) [13]. The ratio of RNA-seq reads in males encompassing such a divergent site should reflect the allelic expression level of the neo-X and neo-Y allele, and mapped RNA-seq reads were weighted by the ratio of mapped DNA-seq reads, to correct for mapping biases [14]. We found a pronounced excess of neo-X linked expressions in male *D. albomicans* (Figure 2A). Out of 805 genes with allele-specific SNPs and detectable expressions from both their neo-sex alleles (read counts > 0; out of a total 4,839 annotated neo-sex linked genes), 242 (30.1%) show a significantly higher expression level of the neo-X allele compared to the neo-Y ($p < 0.05$, Fisher's exact test), and only 13 genes show significant neo-Y biased expression (Figure 2B). To test the robustness of our approach to residual heterozygosity segregating in our lines or potential mapping biases, we applied the same procedure to autosomal genes. Out of 5,145 autosomal genes, 80 have a male-specific SNPs, and only 16 genes show biased expression (5 genes with higher and 11 genes with lower expression for the male-specific SNP). Thus, the number of falsely identified neo-Y-specific SNPs is negligible. The prevalence of genes with reduced expression levels suggests that chromosome-wide downregulation on the neo-Y already happens prior to the general degradation of protein-coding sequences through frameshift mutations and premature stop codons.

Have other mutations at protein-coding genes contributed to Y degeneration? To investigate whether neo-Y linked genes show less constraint at their protein-coding sequence, we sequenced the transcriptome of *D. nasuta*, the sister species of *D. albomicans*. We compared the ratio (ω) of the

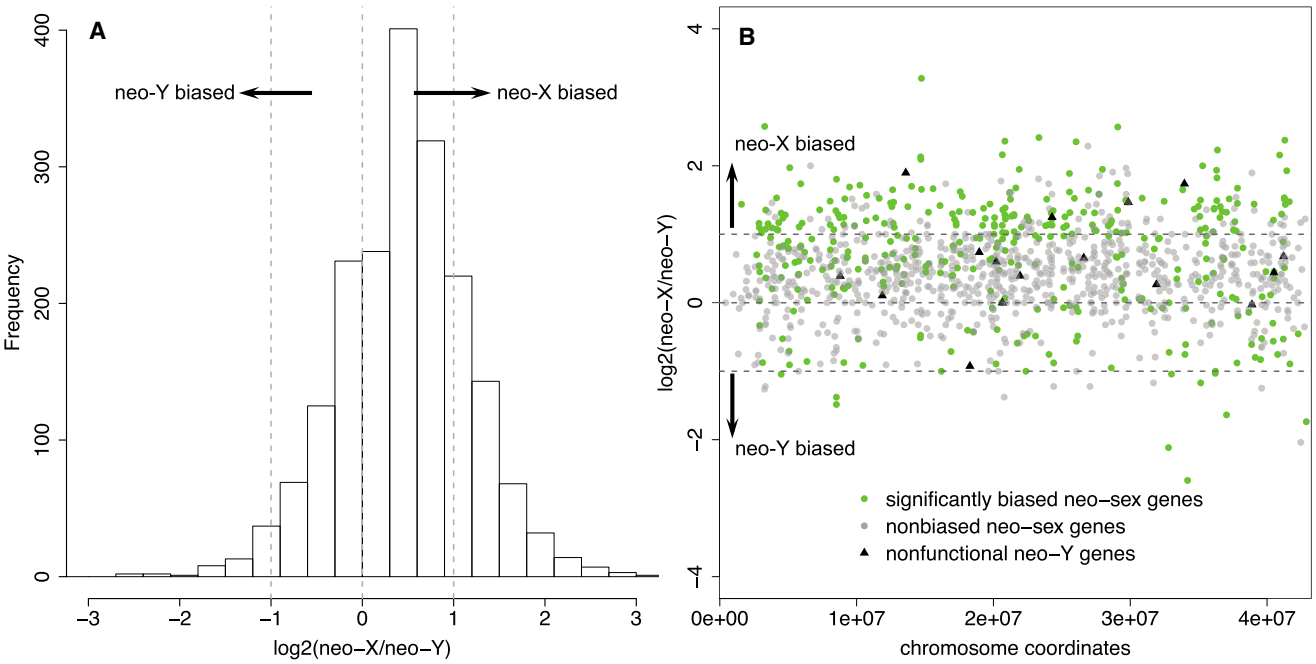

Figure 2. Relative Expression of Neo-Sex Chromosome Genes in Male *D. albomicans*

(A) Histogram of expression ratios between neo-X and neo-Y alleles. If this ratio is higher than zero, a gene shows neo-X biased expression, and below zero, it is neo-Y biased.

(B) Expression ratios for neo-X versus neo-Y genes along the chromosome. The x axis is the coordinate of the gene along the chromosome, and the y axis gives the normalized (by DNA-seq reads) log ratio of neo-X versus neo-Y expression.

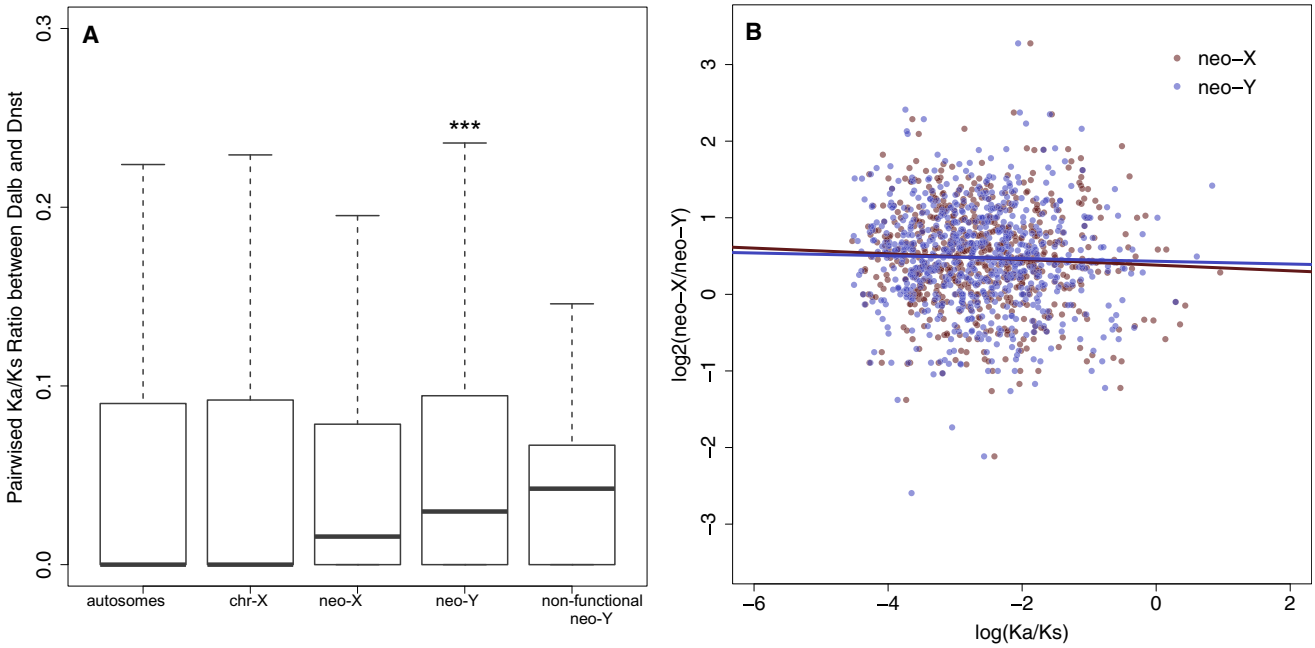

Figure 3. Protein Evolution and Biased Expression of Neo-Sex Genes

(A) Boxplots of pairwised Ka/Ks (ω) ratios between *D. albomicans* and *D. nasuta*. The box extends from the lower to the upper quartile, with the line indicating the median value. The ω ratio of the neo-Y chromosome is significantly higher compared to the autosomes or the neo-X.

(B) Protein evolution of neo-X (red) or neo-Y (blue) alleles versus the degree of biased expression, with linear regression lines.

nonsynonymous substitution rate (*K*a) versus synonymous substitution rate (*K*s) between the two species for genes on different chromosomes. Genes located on the neo-Y chromosome have significantly elevated rates of protein evolution (Figure 3A; Wilcoxon test, $p < 10^{-3}$), suggesting that slightly deleterious amino acid mutations have already begun to accumulate on the neo-Y of *D. albomicans*. However, no association between ω and expression bias was found ($p > 0.05$, F-test; Figure 3B), suggesting that the incorporation of missense mutations and downregulation of protein-coding genes contribute independently to early Y degeneration.

The observed pattern of less expression from the neo-Y relative to the neo-X can be caused by the downregulation of individual neo-Y linked genes, due to mutations in their regulatory sequences, or by local changes in the chromatin structure of the neo-Y chromosome reducing gene expression at adjacent genes (i.e., heterochromatinization). In addition, we cannot exclude the possibility that upregulation of neo-X alleles has also contributed to the observed pattern at some genes (i.e., acquisition of dosage compensation). Dosage compensation, however, only evolves in response to Y degeneration after genes have become silenced or pseudogenized [15, 16] and is therefore not expected to cause upregulation of neo-X genes with fully functional neo-Y homologs. Also, partial dosage compensation would result in a bimodal distribution of neo-X versus neo-Y expression ratios (corresponding to compensated and noncompensated genes), instead of the unimodal distribution observed on the *D. albomicans* neo-sex chromosome (Figure 2A). No obvious positional clustering of neo-X biased genes is observed (Figure 2B), as might be expected under large-scale chromatin-mediated modifications causing expression changes, such as MSL (male-specific lethal)-mediated dosage compensation, the canonical dosage compensation system in *Drosophila* [15]. However, our current assembly has on average only 3.6 genes per scaffold, and a better genome assembly (in particular, larger continuous scaffolds) is needed to address this point.

To conclude, our transcriptome analysis shows that chromosome-wide downregulation appears to occur prior to the accumulation of severely deleterious nonsense mutations during the incipient evolution of the neo-Y chromosome in *D. albomicans*. This suggests that chromosome-wide gene silencing initiates the process of Y degeneration.

Accession Numbers

Illumina reads produced in this study are deposited at NCBI Short Reads Archive under the accession number SRA049492.2.

Acknowledgments

This work is funded by National Institutes of Health grants (R01GM076007 and R01GM093182) and a Packard Fellowship to D.B.

Received: November 15, 2011
Revised: December 15, 2011
Accepted: January 25, 2012
Published online: February 23, 2012

References

1. Charlesworth, B., and Charlesworth, D. (2000). The degeneration of Y chromosomes. Philos. Trans. R. Soc. Lond. B Biol. Sci. *355*, 1563–1572.
2. Carvalho, A.B. (2002). Origin and evolution of the Drosophila Y chromosome. Curr. Opin. Genet. Dev. *12*, 664–668.
3. Carvalho, A.B., and Clark, A.G. (2005). Y chromosome of *D. pseudoobscura* is not homologous to the ancestral Drosophila Y. Science *307*, 108–110.
4. Bachtrog, D., and Charlesworth, B. (2002). Reduced adaptation of a non-recombining neo-Y chromosome. Nature *416*, 323–326.

5. Bachtrog, D., Hom, E., Wong, K.M., Maside, X., and de Jong, P. (2008). Genomic degradation of a young Y chromosome in *Drosophila miranda*. Genome Biol. *9*, R30.
6. Zhou, Q., and Bachtrog, D. (2012). Sex-specific adaptation drives early sex chromosome evolution in Drosophila. Science, submitted.
7. Bachtrog, D. (2005). Sex chromosome evolution: molecular aspects of Y-chromosome degeneration in Drosophila. Genome Res. *15*, 1393–1401.
8. Bachtrog, D. (2006). Expression profile of a degenerating neo-y chromosome in Drosophila. Curr. Biol. *16*, 1694–1699.
9. Bachtrog, D. (2003). Accumulation of Spock and Worf, two novel non-LTR retrotransposons, on the neo-Y chromosome of *Drosophila miranda*. Mol. Biol. Evol. *20*, 173–181.
10. Steinemann, M., and Steinemann, S. (1998). Enigma of Y chromosome degeneration: neo-Y and neo-X chromosomes of Drosophila miranda a model for sex chromosome evolution. Genetica *102-103*, 409–420.
11. Yu, Y.C., Lin, F.J., and Chang, H.Y. (1999). Stepwise chromosome evolution in Drosophila albomicans. Heredity *83*, 39–45.
12. Bachtrog, D. (2006). The speciation history of the *Drosophila nasuta* complex. Genet. Res. *88*, 13–26.
13. Zhou, Q., Zhu, H., Huang, Q., Zhao, L., Zhang, G., Roy, S., Vicoso, B., Xuan, Z., Ruan, J., Zhang, Y., et al. (2012). Deciphering neo-sex and B chromosome evolution by the draft genome of *Drosophila albomicans*. BMC Genomics, in press.
14. Dohm, J.C., Lottaz, C., Borodina, T., and Himmelbauer, H. (2008). Substantial biases in ultra-short read data sets from high-throughput DNA sequencing. Nucleic Acids Res. *36*, e105.
15. Marín, I., Franke, A., Bashaw, G.J., and Baker, B.S. (1996). The dosage compensation system of Drosophila is co-opted by newly evolved X chromosomes. Nature *383*, 160–163.
16. Charlesworth, B. (1996). The evolution of chromosomal sex determination and dosage compensation. Curr. Biol. *6*, 149–162.