Epigenetic Effects of Polymorphic Y Chromosomes Modulate Chromatin Components, Immune Response, and Sexual Conflict

Epigenetic modulation of gene expression by polymorphic *Y* chromosomes: Sexual conflict, chromatin components, and immune response.

Bernardo Lemos*, Alan T. Branco, Daniel L. Hartl*.

Department of Organismic and Evolutionary Biology

Harvard University

16 Divinity Avenue

Cambridge, MA 02138

*Correspondence to:
BL: blemos@oeb.harvard.edu

DLH: dhartl@oeb.harvard.edu

Author contributions: BL, designed experiments, carried out experiments, analyzed data, wrote paper; ATB, carried out experiments, analyzed data; DLH designed experiments, wrote paper.

Abbreviations: YRV, *Y*-linked regulatory variation; PEV, position-effect variegation.

Abstract

Genetic conflicts between sexes and generations provide a foundation for understanding the functional evolution of sex chromosomes and sexually dimorphic phenotypes. *Y* chromosomes of *Drosophila* contain multimegabase stretches of satellite DNA repeats and a handful of protein coding genes that are monomorphic within species. Nevertheless, natural polymorphic variation within highly heterochromatic *Y* chromosomes of *Drosophila* result in the modulation of gene expression at many loci located in the autosomes and the *X*-chromosome. Here we show that such naturally occurring *Y*-linked regulatory variation (YRV) can be detected in somatic tissues and contributes to the epigenetic balance of heterochromatin / euchromatin at three distinct loci showing position-effect variegation (PEV). Moreover, naturally occurring polymorphic *Y*-chromosomes differentially affect the expression of thousands of genes in *XXY* female genotypes in which *Y*-linked protein-coding genes are not transcribed. The most highly affected genes show a disproportionate influence of YRV on variable expression of genes whose protein products localize to the nucleus, show nucleic-acid binding activity, and are involved in transcription, chromosome organization, and chromatin assembly. These include key components such as HP1, Trithorax-like (GAGA-factor), Su(var)3-9, Brahma, MCM2, ORC2, and Inner Centromere Protein. Furthermore, mitochondria-related genes, immune response genes, and transposable elements are also disproportionally affected by *Y*-chromosome polymorphism. These functional clusterings may arise as a consequence of the involvement of *Y*-linked heterochromatin in the origin and resolution of sexual conflict between males and females. Altogether, our results indicate that *Y*-chromosome heterochromatin serves as a major source of epigenetic variation in natural populations, that interacts with chromatin components and other regulators to modulate the expression of biologically relevant phenotypic variation.

Introduction

The *Y* chromosome is a prime target for the evolution and resolution of genetic conflicts related to the distortion of sex ratios, the evolution of male-limited traits, and antagonistic parent-of-origin influences on male and female progeny (1). In spite of theoretical expectations regarding potentially broader roles for male-limited genetic elements, the functional relevance of polymorphic variation in *Y* chromosomes has mostly been overlooked in both theoretical as well as empirical studies, although noteworthy exceptions can be found (1-3). The reasons for neglect stem from the unusual molecular characteristics of the *Y* chromosome. First, *Y* chromosomes show dramatically sparse gene numbers with a limited and specialized functional profile; in Drosophila, 13 protein-coding genes are assigned to the *Y* chromosome, each of which is expressed only during spermatogenesis (4, 5). Second, *Y* chromosomes show dramatically lower levels of polymorphic, single-nucleotide genetic variation, than genes in other chromosomes; indeed, theoretical models and sequencing studies have led to the widespread view that *Y*-linked, single-copy protein coding genes are monomorphic within species of Drosophila (6, 7). In contrast, the *Y* chromosome shows a great deal of structural polymorphism evidenced by variation in copy number of repeated sequences (8, 9).

The absence of sequence polymorphisms in *Y*-linked, protein-coding genes has cast a shadow of doubt on phenotypic data showing polymorphic *Y*-linked effects on fitness (10), temperature sensitivity of spermatogenesis (11), sex-ratio distortion (3), geotaxis (12), and male courtship (13). Furthermore, population genetic theory suggested that the conditions for stably maintaining *Y*-linked polymorphisms might be limited (14). Indeed, the fundamental conundrum has been to reconcile evidence indicating polymorphic phenotypic effects of *Y* chromosomes with a lack of protein-sequence variation (6, 7). Remarkably, the amount of *Y*-linked DNA (40 Mb) constitutes over 20% of the Drosophila genome; based on the large size of the *Y* chromosome and gene density in the *X* chromosome, one might expect over 5,000 genes to be *Y*-linked. That only 13 protein-coding genes are actually located in the *Y* chromosome underscores the highly heterochromatic content of the chromosome in the form of multimegabase stretches of satellite DNA. In this context, the discovery that highly heterochromatic *Y* chromosomes are polymorphic in their effects on gene regulation throughout the genome was most unexpected (15). This discovery suggested a mechanism through which heterochromatic *Y* chromosomes might promote functional variation with consequences for various downstream phenotypes, including fitness.

Heterochromatin represents a large fraction of eukaryotic genomes and is characterized by a high density of sequence repeats that remain condensed through the cell cycle (16, 17). Such heterochromatic repeats are most abundantly found around centromeres and in the *Y* chromosome (17-19). Furthermore, euchromatic and heterochromatic environments present distinct and sometimes opposing requirements for the expression of protein-coding genes. Accordingly, euchromatic genes are silenced upon insertion into heterochromatin whereas genes that natively reside within heterochromatin may be repressed upon their translocation to euchromatin (16, 20, 21). Also, importantly, the manipulation of the amount of *Y*-linked heterochromatin in Drosophila has been shown to result in variable gene expression, with larger amounts of *Y*-linked heterochromatin leading to lessened heterochromatization of autosomal markers at the boundary between euchromatin and heterochromatin (16, 22).

Hence, the finding of YRV led to the proposal that the *Y* chromosome might have evolved to become a heterochromatic 40 Mb regulatory giant whose polymorphic functions between individuals are exerted epigenetically through its contribution to global chromatin dynamics (15, 23, 24). In the work reported here we addressed the hypothesis that naturally occurring *Y*-chromosome lineages are polymorphic for genetic elements that may influence global chromatin dynamics within the nucleus. We show that natural polymorphic variation between *Y*-chromosome lineages differentially contributes to the balance of heterochromatin and euchromatin in the genome, and we identify several functionally coherent gene sets that are affected by *Y*-chromosome variation. Remarkably, these patterns can be observed in Drosophila *XXY* female genotypes in which *Y*-linked protein coding genes are not expressed. These findings give evidence for a general process through which genome-wide gene expression is modulated epigenetically by polymorphic *Y*-linked chromatin.

Results

Position-effect variegation (PEV) of mutations affecting the white-eye gene is typically observed when the gene is moved from its native location to a position near a boundary between heterochromatin and euchromatin (16). PEV in *white* mutants (e.g., *w[m4h]*, white-mottled 4) is attributed to the spreading of heterochromatin associated proteins and other modifications into the adjacent euchromatin (16, 25). In order to address whether *Y* chromosomes are polymorphic for variation affecting PEV, we generated 16 *Y*-chromosome substitution lines of Drosophila that are identical for all autosomes and *X* chromosomes and differ only in the origin of the *Y* chromosome. In *white* (*w[m4h]*) mutants, the diagnostic mosaic phenotype arises due to a

chromosomal inversion which moves the *white* gene from its native euchromatic environment to the euchromatic-heterochromatic boundary in the *X*-chromosome (16, 25). Expression of the *white* gene occurs in cell lineages in which the balance between heterochromatin versus euchromatin is tipped toward greater abundance of euchromatin. Lack of *white* gene expression occurs in cell lineages in which the chromatin balance is tipped towards greater amounts of heterochromatin. We find that *Y* chromosomes are polymorphic for genetic elements with dramatic consequences on *w[m4h]* PEV; while some *Y*-chromosome lineages result in flies with almost completely white eyes, other *Y*-chromosome lineages result in flies with mostly red eyes, with a continuum of effects between these two extremes (Fig. 1 *A B C*; Table S1). These data show that naturally occurring *Y*-chromosome lineages comprise a rich source of epigenetic variation that can be detected in somatic tissues and contribute to the balance of heterochromatin and euchromatin around the *white* locus in *w[m4h]* mutants.

To address the generality of the effect of naturally occurring *Y*-chromosome lineages on the balance of heterochromatin and euchromatin within the nucleus, we investigated two additional loci. One was *bw[D]* (brown-eye dominant), in which variegated expression arises due to a dominant insertion of >1 Mb of heterochromatin into the brown locus at its native location in the euchromatin of chromosome 2 (16). Variegated expression of *bw[D]* is attributed to the recruitment of the *brown* locus to a heterochromatic compartment of the nucleus due to the large insertion of satellite DNA in these mutants. We find that, in agreement with results from *white* (*w[m4h]*), *Y* chromosomes are also polymorphic for their effects on the amount of heterochromatization in the *bw[D]* locus (Fig. 1*C*). The requirements for the expression of genes naturally residing within heterochromatin may be opposite to those of genes naturally residing within euchromatin (16, 20). Consistent with this model are results from a variegating allele *lt[x13]* of the gene *light.* Normally located in the pericentric heterochromatin of chromosome 2, the *light* gene shows variegated expression when moved to a euchromatic location in the allele *lt[x13]* (16). We may therefore predict that *Y[Ohio]* and *Y[Congo]* may show an opposite effect on the *lt[x13]* allele, relative to what we observed for *w[m4h]* and *bw[D]*. Indeed, we find a greater expression of the *light* gene in the *Y[Congo]* strain than in the *Y[Ohio]* strain (Fig. 1*C*). Taken together, these results regarding the effects of epigenetic silencing of variegating alleles of the *white, brown*, and *light* genes indicate that naturally occurring polymorphisms within *Y*-linked heterochromatin dramatically affect the balance between heterochromatic and euchromatic compartments within the cell nucleus. Importantly, these results establish that the effects of

naturally occurring *Y* chromosomes can be observed in male somatic tissues where *Y*-chromosome transcription of single-copy protein-coding genes does not occur. Furthermore, in view of the large magnitude of the effect observed, the data strikingly places naturally occurring *Y*-chromosome polymorphisms on par with several laboratory-generated mutations affecting major chromatin components and chromatin regulators (*E(var)* and *Su(var)* genes) (26).

To address the hypothesis that *Y*-linked polymorphism might still differentially affect gene expression in females in the absence of expression from *Y*-linked protein-coding genes, we generated identical *XXY* female genotypes that varied only in the origin of the *Y* chromosome (Fig. S1). The important background information is that *XXY* genotypes in Drosophila develop into viable females and that *Y*-linked genes are not transcribed in females (16). Indeed, using several PCR primer sets we fail to detect any trace of expression of *Y*-linked protein coding genes (Fig. S2) in *XXY* female genotypes of *Drosophila*. Gene expression on flies carrying *Y*-chromosomes with markedly distinct effects on *white* PEV was assayed by microarrays in a small sample of *Y*-chromosomes measured under high replication (Fig. S3). Remarkably, we observed dramatic differences in the expression of genes when naturally occurring *Y*-chromosome variants were present in the female genotype (Fig. S4). For instance, at $P < 0.001$ we observed 1,152 genes (FDR < 0.05) differentially expressed between *XXY[Congo]* and *XXY[Ohio]*, with 662 genes (57%) up-regulated in *XXY[Congo]* and 490 genes (43%) up-regulated in *XXY[Ohio]*. The data show a significant positive correlation between the fold-change in expression of genes in *Y[Congo]* relative to *Y[Ohio]* in males and females ($\rho = 0.26$, $P < 0.0001$); it recapitulates previous results (15) and identifies the differential expression of genes known to be involved in male fertility. For instance, differentially expressed genes in *XXY* females include genes coding for protein ejaculatory bulb II, male specific RNA 84Dc, and *fmr1*, a gene known to be involved in male courtship (27). Taken together, these data indicate that YRV can be reproduced in *XXY* females and suggests a mechanism for YRV that does not require the expression of *Y*-linked protein-coding genes.

Cytological evidence from two chromatin-associated proteins indicates that they bind to simple sequence repeats in the *Y*-chromosome. One is the transcriptional activator trithorax-related (GAGA-factor), which binds to AAGAG satellites in the *Y* chromosome, and the other is the origin of replication complex protein 2 (ORC2), which binds to AT-rich repeats along the *Y* chromosome (28, 29). Hence, in view of autoregulatory feedback mechanisms in gene expression (30), we predicted that the expression of these two genes might vary across lines differing in the

origin of the *Y* chromosome. In agreement with our expectations, we find that levels of trithorax-related and ORC2 transcript differ significantly between *XXY[Congo]* and *XXY[Ohio]* ($P < 0.001$). Thus, we propose a model in which chromatin regulators might be recruited differentially to polymorphic *Y* chromosomes, which might in turn affect the steady state mRNA abundance of these genes. In particular, we tested the hypothesis that other chromatin components and chromatin regulators are responsive to *Y*-chromosome polymorphism and more meagerly expressed in *XXY[Congo]*. Indeed, we found this to be a general pattern with >90% of the differentially expressed genes assigned to the gene ontology category of chromatin silencing more meagerly expressed in *XXY[Congo]* (Fig. 2 *A*; Table S2; Fisher's exact test, $P < 0.01$). Furthermore, analyses across several gene ontology categories revealed that lower expression of genes in the *XXY[Congo]* background shows significant enrichment for genes whose products are located in the nucleus (145 genes; $P < 10^{-25}$, after Bonferroni correction for multiple testing) and have nucleic-acid-binding activity (136 genes, $P < 10^{-14}$). Functionally, the most highly affected genes appear to be involved in the processes of transcription (79 genes; $P < 10^{-7}$), chromosome organization and biogenesis (29 genes, $P < 10^{-5}$), DNA packaging (23 genes; $P < 10^{-5}$), chromatin assembly or disassembly (19 genes; $P < 10^{-5}$), reproduction (49 genes; $P < 10^{-4}$), and RNA splicing (23 genes; $P < 10^{-4}$). The genes identified include not only ORC-2 and Gaga-factor but also candidates such as HP1, Su(var)3-9, MCM2, brahma, centromere identifier, chromatin-assembly factor 1 subunit, caf1-180, and others (Fig. 2 *B*). To further confirm the relevance of differential expression of chromatin components, we used quantitative real-time PCR to assay the expression of key components in *Y[Congo]* and *Y[Ohio]* males (Fig. S5). All in all, our analysis revealed a substantial contribution of *Y*-chromosome polymorphisms to the differential expression of chromatin components and chromatin regulators, with several genes previously identified as suppressors of variegation more meagerly expressed in the presence of *Y[Congo]*. This suggests mechanisms by which the enhancer of variegation property of *Y[Congo]* might be a consequence of the greater availability of Su(var) proteins such as HP1 and Su(var)3-9 among others.

In view of the theoretical expectation that the *Y* chromosome may mediate the origin and resolution of genetic conflicts (2, 31), we predicted that YRV targets might be enriched for genes with antagonistic fitness effects on males and females. One indicator of this possibility is that genes responding to *Y*-linked polymorphisms show higher expression in males and lower expression in females than genes not affected by YRV (15). Furthermore, in view of potentially strong sibling-sibling competition in the progeny of Drosophila, population genetic theory

predicts that harmful interactions might be expected to evolve between the mitochondria and the *Y*-chromosome (1). In Drosophila the potential for harmful interactions might be even further enhanced through the dramatic mitochondrial remodeling that occurs during spermatogenesis and that leads to the mitochondrial derivatives found in the fruit-fly sperm (32, 33). It has been previously found that YRV targets are enriched for genes with functional roles related to the mitochondria (15), presumably as a by-product of genetic interactions between the *Y* chromosome and mitochondria that take place during spermatogenesis. On the other hand, *Y*-chromosome effects on the expression of genes required for mitochondrial function may occur independently of spermatogenesis. To test this hypothesis we searched for genes localized to the mitochondrion and genes involved in electron transport (Table S2) and found significant up-regulation of these genes in *XXY[Congo]* (Fig. 3 *A B;* Table S2; Fisher's exact test, $P < 0.01$). Further testing among genes up regulated in *XXY[Congo]* indicated significant enrichment for proteins belonging to the mitochondrial electron transport chain (12 genes, $P < 0.05$) and with oxidoreductase activity (50 genes, $P < 0.01$), and include both mitochondrial and nuclear encoded proteins. These results point to close connections between *Y*-chromosome polymorphisms and mitochondrial functions that are not limited to those occurring during the complex mitochondrial remodeling that takes place during spermatogenesis.

Empirical evidence as well as theoretical considerations also suggest that variable immune function may arise as a consequence of genetic conflict (31, 34, 35). In order to address the role of *Y*-chromosome polymorphism underlying variable immune function we searched for defense-response and immune-related genes among the targets of YRV. Strikingly, we find that >80% of the differentially expressed genes assigned to the gene ontology category of defense response and immunity show greater expression in *Y[Congo]* relative to *Y[Ohio]* (Fig. 4 *A*; Table S2; $P < 0.01$; Fig. S6). Among genes up regulated in *XXY[Congo]* we find 13 targets that show protease inhibitor activity ($P = 0.0005$). Polymorphic variation in *Y* chromosomes has also been shown to result in the differential expression of transposable elements (15). One possibility is that the activity of transposable elements in *Y*-chromosome substitution lines depended on the transcription of *Y*-linked protein coding genes. However, in females, we find an up-regulation of transposable elements in the *XXY[Congo]* line relative to *XXY[Ohio]* (Fig. S6). These patterns are also confirmed in contrasts between *Y[Congo]* and *Y[Ohio]* males in two different genetic backgrounds (Figs. S6-S7). One of these backgrounds is the same in which the *w[m4h]* PEV data was obtained, the other is a laboratory background previously reported (15). This is evidence for a common and general epigenetic mechanism underlying YRV in males and females, which does

not require transcription of *Y*-linked protein coding genes and is associated with transposable element activity and differential expression of chromatin components and chromatin regulators. Altogether, these findings suggest that polymorphisms among *Y* chromosomes differentially affect the expression of immune-related genes and may consequently underlie variable immune function among males. Furthermore, the phenomenon is consistently reproduced in female *XXY* genotypes, indicating that it is not restricted to cells destined for spermatogenesis and that it does not require the expression of *Y*-linked protein-coding genes.

With few exceptions (36), chromatin components are generally conserved and under strong purifying selection within species, which tends to limit the extent of sequence or expression polymorphisms in these genes. On the other hand, the large amounts of heterochromatin found in *Y* chromosomes might be particularly conducive to harboring polymorphic variation. One reason is that the mutation rates of repetitive heterochromatic DNA are unusually high owing to replication slippage and other processes (37-39). Another reason is that the population genetics of *Y* chromosomes is unusual in featuring male hemizygosity. Hence, attributes of *Y* chromosomes may be conducive to the accumulation of heterochromatic variation as well as its expression in males. An important issue is whether the effects of *Y* chromosomes on gene expression are largely autonomous or whether polymorphic autosomal variation may substantially modify the effects of individual *Y* chromosomes. To address this issue, we studied the relative contribution of natural variation among *Y* chromosomes to that of dominant modifiers present in the *Congo* and *Ohio* genetic backgrounds on the epigenetic variation evidenced by PEV. In these experiments, we swapped the *Y* chromosomes of the *Y[Ohio]* and *Y[Congo]* strains while preserving the original genetic backgrounds. These *Y*-swapped strains were then assayed for their effects on *white* PEV. Remarkably, we observed that the silencing effect of the *Y[Congo]* chromosome (enhancer of variegation of the white-eyed phenotype) is observed in the *Ohio* genetic background as well as in its original background. Conversely, the *Y[Ohio]* chromosome (suppressor of variegation leading to a mostly red-eyed phenotype) acts similarly when introgressed into the *Congo* genetic background as it does in its original *Ohio* background (Fig. 5). This result suggests that most of the naturally occurring epigenetic variation in PEV can be attributed to the *Y* chromosome, with little contribution of dominant, naturally occurring genetic modifiers.

Discussion

Here we showed that naturally occurring *Y*-linked regulatory variation (YRV) arising from highly heterochromatic *Drosophila Y*-chromosomes can be detected in somatic tissues and contributes to the epigenetic balance of heterochromatin versus euchromatin at three distinct loci showing position-effect variegation (PEV). These results point to the *Y* chromosome as a major contributor to naturally occurring epigenetic variation in *Drosophila*. In magnitude, the effect of polymorphic *Y* chromosomes on the amount of heterochromatin-driven silencing of the *white* gene is on par to that of loss-of-function mutants of major chromatin components previously uncovered in genetic screens (26). Furthermore, we showed that polymorphic *Y* chromosomes differentially affect the expression of thousands of genes located in the *X* chromosome and autosomes of *XXY* females. These results indicate that expression of *Y*-linked protein coding genes is not required for YRV. Finally, the gene-expression variation identified in these experiments is functionally coherent: it affects chromatin components as well as gene sets which we suggest that might be involved in sexual conflict associated with immune-response and epistatic interactions between Y-chromosomes and the mitochondria. Nevertheless, further characterization of *Y*-linked variation and its effects on PEV and YRV with larger samples that include intra-population variation will be important.

One molecular model for interpreting variegated expression is that it reflects the availability of heterochromatin-forming proteins in the nucleus; the model is one of local diffusion of chromatin-modifying enzymes from high-affinity binding sites to low-affinity sites nearby (40, 41). For instance, halving the dosage of genes required for heterochromatin formation, such as *HP1*, *Su(var)3-9* or *Su(var)3-7*, leads to a decreased availability of these proteins, and consequently a decrease in the amount of heterochromatin within the nucleus (29). The reduced heterochromatin can in turn restore the wild-type eye color in the *white* (*w[m4h]*) PEV system (29). Similarly, one molecular model for interpreting the consequences of polymorphic *Y*-chromosome heterochromatin is that variation in autosomal and *X*-linked gene expression reflects the limited availability of DNA binding proteins with heterochromatin-forming and transcription-factor activity in the nucleus. Accordingly, polymorphic *Y* chromosomes may exert their effects on gene regulation by serving as a differential sink for the binding of chromatin regulators or other DNA-binding proteins, which may consequently lead to the titration of these proteins at other genomic locations (15). The feasibility of such a heterochromatic satellite-repeat-sink model has been demonstrated in the case of the transcription factor C/EBP alpha which binds to satellite repeat alpha in mice (42). Also, genetic analysis of enhancer-trap mutants whose expression is modulated by the *Y* chromosome identified a multimegabase segment in the *Y*

chromosome that could act as a transregulator of a *lacZ* reporter expression (43). The sequences mediating the effects were functionally redundant and spatially dispersed across bands h1-h10 of the long arm of the *Y*-chromosome, which coincides with the location of (AAGAG)*n* and (AAAGAGA)*n* repeats that serve as motifs for binding by the GAGA-transcription-factor (43). Finally, other independent evidence for a chromatin-sink model involving the GAGA-factor comes from studies revealing mutant phenotypes of flies fed polyamide compounds that bind specifically to satellite repeats of the type (AAGAG)*n* (44). Remarkably, we found that the expression of GAGA-factor is significantly modulated by polymorphic variation in the *Y* chromosome. Nevertheless, our data also point to several others previously unrelated DNA-binding proteins and suggest that other repeats might also serve as chromatin sinks, which might lead to a complex dynamics of protein availability at autosomal and *X*-linked sites.

Variable size of the rDNA array, which is present on the short arm of the *Y* chromosome, could also potentially underlie some of the effects herein reported. Accordingly, it has been recently shown that variable rDNA array size contributes to variation in PEV phenotype when examined in isolation (45). However, when naturally occurring *Y*-chromosome lines with different rDNA array sizes are probed, PEV phenotypes can be opposite to that expected from their rDNA size alone. Hence, other loci or segments along the *Y* chromosome must also play significant roles. Altogether, we propose that the effects of naturally occurring *Y*-chromosome lineages on gene expression and PEV are exerted through multiple variable loci located along the *Y* chromosome. Such variability in the content or lengths of heterochromatic blocks harboring satellite repeats may be extensive; it arises through a complex dynamics involving repeat homogeneization through inter- and intra-chromosomal gene conversion, expansion and contraction of repeats through replication slippage, sister-chromatid exchange, intrachromatid exchange, and divergence of repeat units through point mutations (38, 39, 46). Mutation-selection balance might account for *Y*-linked polymorphism within species and for variation in *Y*-chromosome size across species of Drosophila over long timescales (47).

Epistatic *Y*-linked effects on gene expression and antagonistic *X*–*Y* interactions resulting from the altered availability of chromatin components and chromatin regulators at limiting concentration throughout the genome might be expected. Furthermore, *Y*-chromosomes effects on the compartment size and transcriptional output of various tissues and organs are yet to be determined and might underlie some of the variation herein reported (48). Our data suggest pathways by which altered chromatin abundance associated with *Y*-chromosome lineages

showing activity as suppressors or enhancers of variegation may underlie the expression and resolution of genetic conflicts through the upregulation of transposable elements, mitochondria-related, and immune-response genes. Hence, we predict a role for the heterochromatin of the sex chromosomes in the variable immunity that can be observed across males and females of disparate taxa (31, 34, 35). Hence, our results raise the possibility that naturally occurring polymorphic variation in tracts of heterochromatin in the *Y* and other chromosomes, including those in the human genome, might serve as important determinants of global chromatin dynamics. Altogether, natural polymorphic variation in heterochromatin may serve as an underappreciated modifier of the differential expressivity and penetrance of ecologically important traits as well as genetic risk factors for disease.

Materials and Methods

Drosophila stocks

Y chromosomes from 16 different strains were introgressed into the same laboratory stock background (BL4361) we previously used (15). This stock is expected to contain very little genetic variation, and upon receipt was subjected to >8 additional generations of brother-sister mating to reinforce homozygosity of the genomic background. The strains used were from the following localities: Mumbai (India), Kuala Lumpur (Malaysia), Connecticut (USA), Arizona (USA), Le Reduit (Mauritius), Athens (Greece), Captain Cook, Hawaii (USA), Bogota (Colombia), Cape Town (South Africa), Massachusetts (USA), Ohio, (USA), Zimbabwe, and Republic of Congo (6 strains). Crosses for each *Y*-substitution line were carried out as previously described (15) and shown in Fig. S1. Introgression of *Y[Congo]* into the *Y[Ohio]* background and of *Y[Ohio]* into the *Y[Congo]* background was done by 7 generations of backcrossing. Flies were grown under 24h light, temperature and humidity controlled incubators. For gene expression analyses, newly emerged flies were collected and aged for three days at 25°C, after which they were flash frozen in liquid nitrogen and stored at -80°C. Whenever females are analyzed, they were collected in less than 7 hours upon eclosion. All females used were unmated females.

Essays for variegated expression

In all assays, males were crossed to females from a stock carrying *w[m4h]* maintained in a background with the *Su(var)3-10[2]*. Experiments were also replicated with *w[m4h]* maintained in a background with *Su(var)2-4[01]*. Variegation of *brown* was assessed with allele *bw[D]*. Variegation of *light* was assessed with allele *lt[x13]*. Cultures were performed at 25°C. Males from these crosses were collected, aged for 3 days at 25°C, and stored at -80°C. Heads of males

were removed with a blade. Sets of 5 heads were homogenized with 10uL of acidified ethanol (30% ethanol acidified to pH 2 with HCl). Eye pigmentation was assessed with spectrophotometric analysis at an optical density of 480nm. 4-6 biological replicates were used per treatment, with two measurements taken per replicate. For illustration, males displaying typical eye pigmentation phenotypes were imaged using the auto-montage system (Snycroscopy, Frederick, MD).

Gene expression analyses

Microarrays were ~18,000-feature cDNA arrays spotted with *D. melanogaster* cDNA PCR products as described (15). Total RNA was extracted from whole flies using TRIZOL (Gibco-BRL, Life Technologies, Gaithersburg, MD) according to the manufacturer's recommendations. cDNA synthesis with fluorescent dyes (Cy3 and Cy5) and hybridization reactions were carried out using 3DNA protocols and reagents (Genisphere Inc., Hatfield, PA). Upon hybridization, slides were scanned using AXON 4000B scanner (Axon Instruments, Foster City, CA) and the GenePix Pro 6.0 software. Stringent quality-control criteria were used to ensure reliability of foreground intensity reads for both Cy3 and Cy5 channels. Foreground fluorescence of dye intensities was normalized by the Loess method implemented in the library Limma of the statistical software R. Significance of variation in gene expression due to the *Y* chromosome was assessed with linear models in Limma and using the Bayesian Analysis of Gene Expression Levels (BAGEL) model (49). False discovery rates (FDR) were estimated based on the variation observed when randomized versions of the original dataset were analyzed. Enrichment in gene ontology categories was assessed with GeneMerge (50), which uses a hypergeometric distribution to assess significance; a modified Bonferonni correction is used to account for multiple testing. Microarray gene expression data herein reported can be obtained at the GEO database (GSEXXXX).

For qPCR analyses, RNA extraction and quality control were performed according to the same procedures described above. For each genotype 3 biological replicates were obtained, each consisting of 10 adult flies aged as described above. Quantitative PCR analyses were carried out with the Fast Sybr Green Master Mix (Applied Biosystems, Foster City, California). cDNA synthesis was done with the QuantiTect Reverse Transcription kit (Qiagen, Chatsworth, CA). Real time PCR profiles were obtained with 7900HT Fast Real-Time PCR (Applied Biosystems, Foster City, California). Melting curves were checked for the presence of dimers and unspecific amplification. Real time qPCR data was analyzed with the software package REST (51).

Acknowledgments

We thank Bernardo Carvalho, Bill Rice, and Ron Woodruff for their careful reading and comments on the manuscript, and Silvana Paredes and the Bloomington Stock Center for fly stocks. This work was supported by NIH grant GM084236.

References

1. Rice W R, Gavrilets S, and Friberg U (2008) Sexually antagonistic "zygotic drive" of the sex chromosomes. *PLoS Genetics* 4 e1000313.
2. Rice W R (1998) Male fitness increases when females are eliminated from gene pool: implications for the Y chromosome. *Proceedings of the National Academy of Sciences, USA* 95 6217-6221.
3. Carvalho A B, Vaz S C, and Klaczko L B (1997) Polymorphism for Y-linked suppressors of sex-ratio in two natural populations of Drosophila mediopunctata. *Genetics* 146 891-902.
4. Carvalho A B, Koerich L B, and Clark A G (2009) Origin and evolution of Y-chromosomes: Drosophila tales. *Trends in genetics* 25 270-277.
5. Krsticevic F J, et al. (2010) Functional copies of the Mst77F gene on the Y chromosome of Drosophila melanogaster. *Genetics* 184 295-307.
6. Zurovcova M and Eanes W F (1999) Lack of nucleotide polymorphism in the Y-linked sperm flagellar dynein gene Dhc-Yh3 of Drosophila melanogaster and D. simulans. *Genetics* 153 1709-15.
7. Carvalho A B (2002) Origin and evolution of the Drosophila Y chromosome. *Current Opinion in Genetics & Development* 12 664-8.
8. Lyckegaard E M and Clark A G (1989) Ribosomal DNA and Stellate gene copy number variation on the Y chromosome of Drosophila melanogaster. *Proceedings of the National Academy of Sciences of the United States of America* 86 1944-8.
9. Lyckegaard E M and Clark A G (1991) Evolution of ribosomal RNA gene copy number on the sex chromosomes of Drosophila melanogaster. *Molecular Biology and Evolution* 8 458-74.
10. Chippindale A K and Rice W R (2001) Y chromosome polymorphism is a strong determinant of male fitness in Drosophila melanogaster. *Proceedings of the National Academy of Sciences of the United States of America* 98 5677-82.
11. Rohmer C, et al. (2004) Heat induced male sterility in Drosophila melanogaster: adaptive genetic variations among geographic populations and role of the Y chromosome. *The Journal of Experimental Biology* 207 2735-43.
12. Stoltenberg S F and Hirsch J (1997) Y-chromosome effects on Drosophila geotaxis interact with genetic or cytoplasmic background *Animal Behaviour* 53 853-864.

13. Huttunen S and Aspi J (2003) Complex inheritance of male courtship song characters in Drosophila virilis. *Behavior Genetics* 33 17-24.
14. Clark A G (1987) Natural selection and Y-linked polymorphism. *Genetics* 115 569-77.
15. Lemos B, Araripe L O, and Hartl D L (2008) Polymorphic Y chromosomes harbor cryptic variation with manifold functional consequences. *Science (New York, N Y)* 319 91-3.
16. Ashburner M, Golic K G, and Hawley R S, *Drosophila: A Laboratory Handbook*. 2nd edition ed. 2005, Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.
17. Gatti M and Pimpinelli S (1992) Functional elements in Drosophila melanogaster heterochromatin. *Annu Rev Genet* 26 239-75.
18. Hoskins R A, et al. (2007) Sequence finishing and mapping of Drosophila melanogaster heterochromatin. *Science (New York, N Y)* 316 1625-8.
19. Lohe A R, Hilliker A J, and Roberts P A (1993) Mapping simple repeated DNA sequences in heterochromatin of Drosophila melanogaster. *Genetics* 134 1149-74.
20. Hearn M G, et al. (1991) The effect of modifiers of position-effect variegation on the variegation of heterochromatic genes of Drosophila melanogaster. *Genetics* 128 785-97.
21. Weiler K S and Wakimoto B T (1995) Heterochromatin and gene expression in Drosophila. *Annu Rev Genet* 29 577-605.
22. Dimitri P and Pisano C (1989) Position effect variegation in Drosophila melanogaster: relationship between suppression effect and the amount of Y chromosome. *Genetics* 122 793-800.
23. Rice W R and Friberg U (2008) Genetics. Functionally degenerate--Y not so? *Science (New York, N Y)* 319 42-3.
24. Malone J H and Oliver B (2008) The sex chromosome that refused to die. *BioEssays* 30 409-11.
25. Schulze S R and Wallrath L L (2007) Gene regulation by chromatin structure: paradigms established in Drosophila melanogaster. *Annual Review of Entomology* 52 171-92.
26. Reuter G and Spierer P (1992) Position effect variegation and chromatin proteins. *Bioessays* 14 605-12.
27. Banerjee P, et al. (2007) Substitution of critical isoleucines in the KH domains of Drosophila fragile X protein results in partial loss-of-function phenotypes. *Genetics* 175 1241-50.
28. Platero J S, et al. (1998) Changes in chromosomal localization of heterochromatin-binding proteins during the cell cycle in Drosophila. *The Journal of Cell Biology* 140 1297-306.
29. Pak D T, et al. (1997) Association of the origin recognition complex with heterochromatin and HP1 in higher eukaryotes. *Cell* 91 311-23.
30. Alon U, *An introduction to systems biology: design principles of biological circuits*. CRC Mathematical and Computational Biology Series. 2007, Boca Raton, FL: Chapman & Hall.

31. Arnqvist G and Rowe L, *Sexual conflict*. Monographs in Behavior and Ecology, ed. J.R. Krebs and T. Clutton-Brock. 2005, Princeton, NJ: Princeton University Press.
32. Lindsley D L and Tokuyasu K T, eds. *Spermatogenesis*. The Genetics and Biology of *Drosophila*, ed. M. Ashburner and T.R.F. Wright. Vol. 2nd. 1980, Academic Press.
33. Shoup J R (1967) Spermiogenesis in wild type and in male sterility mutant of *Drosophila melanogaster. The Journal of Cell Biology* 32 663-675.
34. Calsbeek R and Bonneaud C (2008) Postcopulatory fertilization bias as a form of cryptic sexual selection. *Evolution* 62 1137-1148.
35. Svensson E I, McAdam A G, and Sinervo B (2009) Intralocus sexual conflict over immune defense, gender load, and sex-specific signaling in a natural lizard population. *Evolution* 63 3124-35.
36. Vermaak D, Henikoff S, and Malik H S (2005) Positive selection drives the evolution of rhino, a member of the heterochromatin protein 1 family in Drosophila. *PLoS Genetics* 1 96-108.
37. Lynch M, et al. (2008) A genome-wide view of the spectrum of spontaneous mutations in yeast. *Proceedings of the National Academy of Sciences of the United States of America* 105 9272-7.
38. Plohl M, et al. (2008) Satellite DNAs between selfishness and functionality: structure, genomics and evolution of tandem repeats in centromeric (hetero)chromatin. *Gene* 409 72-82.
39. Palomeque T and Lorite P (2008) Satellite DNA in insects: a review. *Heredity* 100 564-73.
40. Talbert P B and Henikoff S (2006) Spreading of silent chromatin: inaction at a distance. *Nature Reviews Genetics* 7 793-803.
41. Monod C, et al. (2002) Modification of position-effect variegation by competition for binding to Drosophila satellites. *EMBO reports* 3 747-52.
42. Liu X, et al. (2007) Functional sequestration of transcription factor activity by repetitive DNA. *J Biol Chem* 282 20868-76.
43. Zhang P, et al. (2000) A trans-activator on the Drosophila Y chromosome regulates gene expression in the male germ line. *Genetica* 109 141-50.
44. Janssen S, et al. (2000) Specific gain- and loss-of-function phenotypes induced by satellite-specific DNA-binding drugs fed to Drosophila melanogaster. *Mol Cell* 6 1013-24.
45. Paredes S and Maggert K A (2009) Ribosomal DNA contributes to global chromatin regulation. *Proceedings of the National Academy of Sciences of the United States of America* 106 17829-34.
46. Ohta T (1982) Allelic and nonallelic homology of a supergene family. *Proceedings of the National Academy of Sciences of the United States of America* 79 3251-4.
47. Petrov D A, Lozovskaya E R, and Hartl D L (1996) High intrinsic rate of DNA loss in Drosophila. *Nature* 384 346-9.
48. Stewart A D, Pischedda A, and Rice W R (2010) Resolving intralocus sexual conflict: genetic mechanisms and time frame. *J Hered* 101 Suppl 1 S94-9.

49. Townsend J P and Hartl D L (2002) Bayesian analysis of gene expression levels: statistical quantification of relative mRNA level across multiple strains or treatments. *Genome Biol* 3 RESEARCH0071.
50. Castillo-Davis C I and Hartl D L (2003) Gene Merge: Post-genomic analysis, data mining, and hypothesis testing. *Bioinformatics* 19 891-892.
51. Pfaffl M W, Horgan G W, and Dempfle L (2002) Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. *Nucleic Acids Res* 30 e36.

Figure legends:

Figure 1. Epigenetic contributions of polymorphic *Y* chromosomes to global chromatin regulation in somatic tissues. (*A*) Variation in eye pigmentation in stocks carrying the *X*-chromosome marker *w[m4h]* and *Y* chromosomes sampled from diverse localities worldwide. (*B*) Eye phenotypes showing variegating eyes for *w[m4h]* in two *Y*-chromosome backgrounds. These flies are genetically identical except for the origin of the *Y* chromosome. (*C*) Differential variegation between *Y[Congo]* and *Y[Ohio]* strains for the *X*-linked marker *w[m4h]*, and second chromosome markers *bw[D]* and *lt[x13]*.

Figure 2. Polymorphic *Y* chromosomes modulate the expression of chromatin components. (*A*) Heat-map of relative expression levels of 101 differentially expressed genes ($P < 0.01$) belonging to the gene ontology category of chromatin silencing. (*B*) Examples of key chromatin components more lowly expressed in *XXY[Congo]* (red) relative to *XXY[Ohio]* (green). Bars denote 95% credible intervals. Relative expression levels are shown with the lowest expression normalized to 1 (red).

Figure 3. (A) Heat-map of relative expression levels of 65 differentially expressed genes ($P < 0.01$) belonging to the gene ontology category of electron transport. (B) Examples of mitochondria-related genes more highly expressed in *XXY[Congo]* (green) relative to *XXY[Ohio]* (red). Bars denote 95% credible intervals. Relative expression levels are shown with the lowest expression normalized to 1 (red).

Figure 4. (A) Heat-map of relative expression levels of 52 differentially expressed genes ($P < 0.01$) belonging to the gene ontology categories of defense response and immunity. (B) Examples of immune-related genes more highly expressed in *XXY[Congo]* (green) relative to *XXY[Ohio]* (red). Bars denote 95% credible intervals. Relative expression levels are shown with the lowest expression normalized to 1 (red).

Figure 5. The autosomes from *Congo* and *Ohio* do not harbor dominant modifiers of *Y*-chromosome driven PEV. Eye phenotypes showing variegating eyes for *w[m4h]* in *Y[Congo]* and *Y[Ohio]* in two backgrounds of autosomes (*Congo* and *Ohio*). Autosomes from *Congo* or *Ohio* are heterozygous with a common set of autosomes inherited from a laboratory strain (+) carrying the *X*-linked *w[m4h]* marker.

Supplementary Tables and Figures

Supplementary Table 1. Variation in eye pigmentation in stocks carrying the *X*-chromosome PEV marker *w[m4h]* and Y chromosomes sampled from diverse localities worldwide. Eye pigmentation was assessed with spectrophotometric analysis at an optical density of 480nm. Mean and standard error of the mean are listed.

Supplementary Table 2. Number of gene up-regulated and down-regulated in *Y[congo]* relative to *Y[Ohio]* in XXY females across 4 gene ontology categories and a range of significance thresholds for differential expression. The excess of down-regulated genes in *Y[Congo]* belonging to chromatin silencing class is statistically significant (Fisher's exact test, $P < 0.01$). The excess of up-regulated genes in *Y[Congo]* localized to the mitochondrion, involved in electron transport, or associated with defense response and immunity is also statistically significant (Fisher's exact test, $P < 0.01$).

Supplementary Figure 1. Crossing scheme for obtaining population-specific *Y* chromosome substitution lineages. This was done independently for each *Y* chromosome lineage to be introgressed such that males from generation G2 have different *Y* chromosomes introgressed into a common isogenic background. *XXY* females from generation (G3) also have different *Y* chromosomes in a background of attached *X*-chromosome and autosomes that is identical across lineages.

Supplementary Figure 2. Protein coding *Y*-linked genes are not expressed in *XXY* females. RT-PCR to detect transcription of *Y*-linked genes in *XY* males and *XXY* females. Lanes 1: *Y[Ohio]*. Lanes 2: *Y[Congo]*. Lanes 3: *XXY[Ohio]*. Lanes 4: *XXY[Congo]*. (*A*-*C*) *kl-2* gene. (*D*-*F*) *kl-3* gene. (*G*-*I*) *kl-5* gene. MW: 100bp DNA ladder (Invitrogen).

Supplementary Figure 3. Experimental design for collecting genome-wide gene expression data across *XXY* female genotypes. Each line denotes two hybridization reactions with Cy3 and Cy5 dyes swapped, for a total of 20 hybridizations.

Supplementary Figure 4. Widespread effect of polymorphic *Y* chromosomes in *XXY* female genotypes of Drosophila. Number of statistically significant gene expression differences between *XXY[Congo]* and *XXY[Ohio]* females (black bars) as a function of the Bayesian posterior

probability of differential expression. Gray bars indicate the estimated number of genes expected by chance after permuting the data.

Supplementary Figure 5. Chromatin associated proteins are differentially expressed between *Y[Congo]* and *Y[Ohio]* males. Ratios of between *Y[Ohio]* and *Y[Congo]* expression obtained with quantitative real-time PCR (Supplementary methods).

Supplementary Figure 6. Examples of immune response genes more highly expressed *Y[Congo]* males relative to *Y[Ohio]* males in two different backgrounds of autosomes and *X* chromosome. Background 4361 is the same inbred background used in the *Y*-chromosome substitution lines reported by Lemos et al. [*Science* **319:** 91-93 (2008)]. Background 6175 results from the cross of males from the *Y*-chromosome substitution lines in the 4361 background with females carrying the PEV marker *w[m4h]*.

Supplementary Figure 7. Examples of transposable elements more highly expressed *Y[Congo]* relative to *Y[Ohio]* in XXY females and in two different male backgrounds of autosomes and *X* chromosome. Background 4361 is the same inbred background used in the *Y*-chromosome substitution lines reported by Lemos et al. [*Science* **319:** 91-93 (2008)]. Background 6175 results from the cross of males from the *Y*-chromosome substitution lines in the 4361 background with females carrying the PEV marker *w[m4h]*.

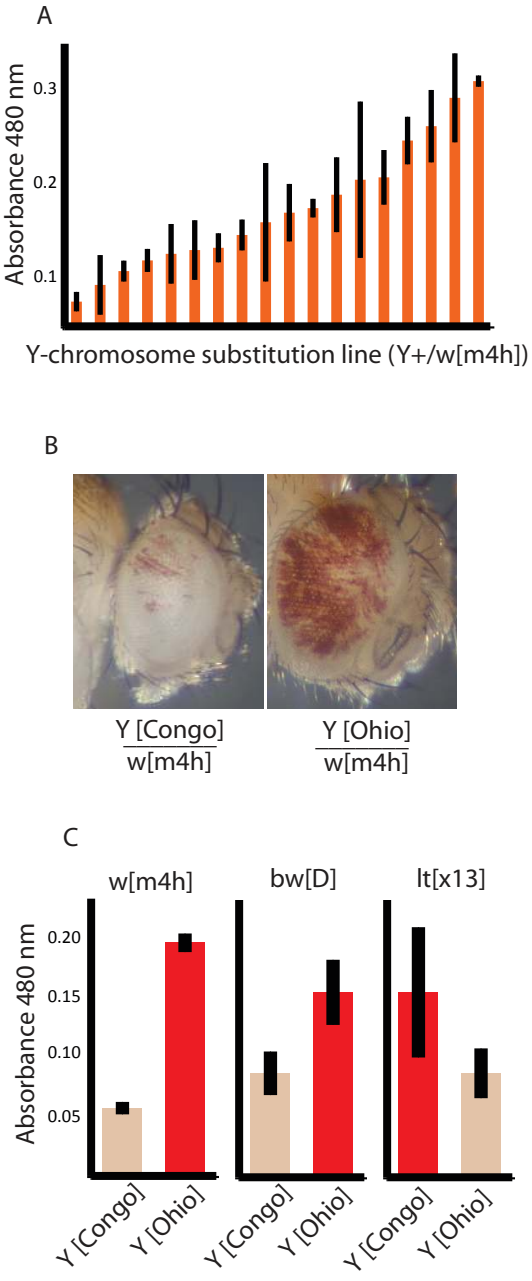


Figure 1. Epigenetic contributions of polymorphic Y chromosomes to global chromatin regulation in somatic tissues. (A) Variation in eye pigmentation in stocks carrying the X-chromosome marker w[m4h] and Y chromosomes sampled from diverse localities worldwide. (B) Eye phenotypes showing variegating eyes for w[m4h] in two Y-chromosome backgrounds. These flies are genetically identical except for the origin of the Y chromosome. (C) Differential variegation between Y[Congo] and Y[Ohio] strains for the X-linked marker w[m4h], and second chromosome markers bw[D] and lt[x13].

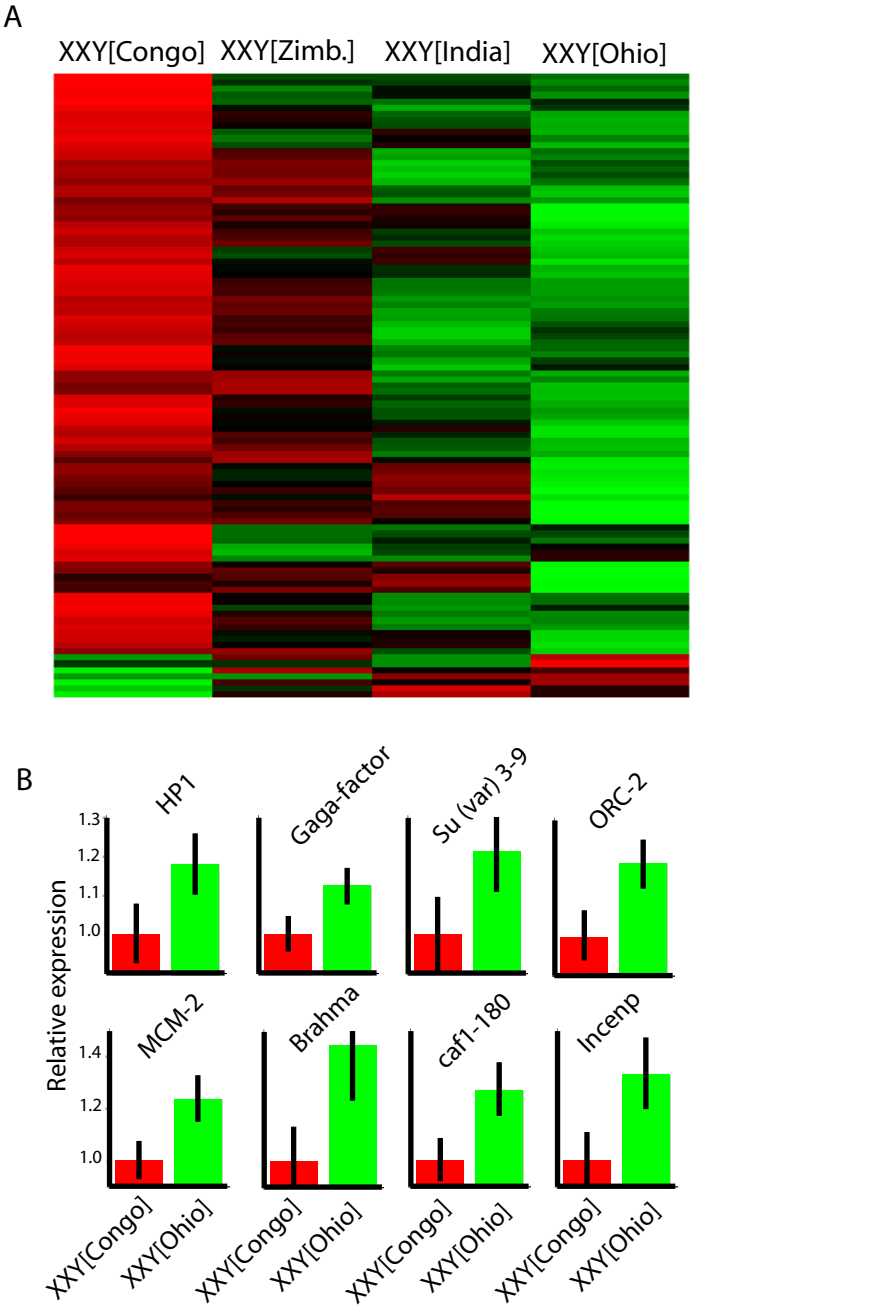


Figure 2. Polymorphic Y chromosomes modulate the expression of chromatin components. (A) Heat-map of relative expression levels of 101 differentially expressed genes ($P < 0.01$) belonging to the gene ontology category of chromatin silencing. (B) Examples of key chromatin components more lowly expressed in XXY[Congo] (red) relative to XXY[Ohio] (green). Bars denote 95% credible intervals. Relative expression levels are shown with the lowest expression normalized to 1 (red).

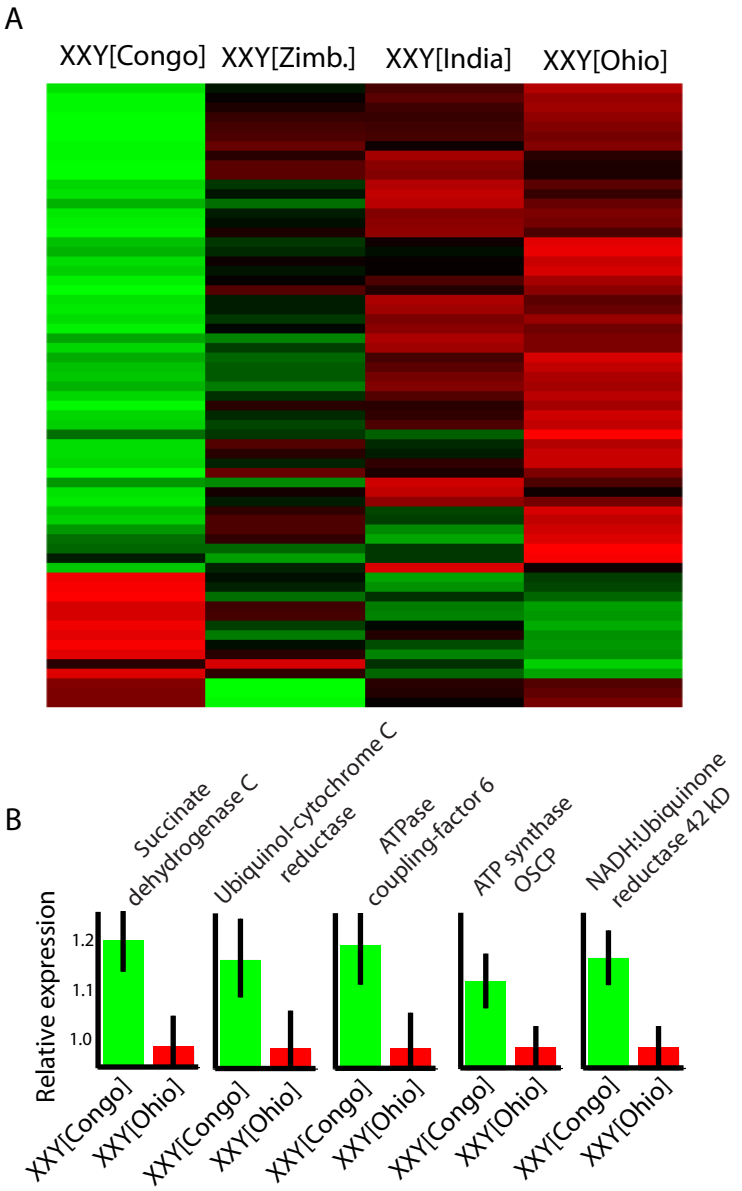


Figure 3. (A) Heat-map of relative expression levels of 65 differentially expressed genes ($P < 0.01$) belonging to the gene ontology category of electron transport. (B) Examples of mitochondria-related genes more highly expressed in XXY[Congo] (green) relative to XXY[Ohio] (red). Bars denote 95% credible intervals. Relative expression levels are shown with the lowest expression normalized to 1 (red).

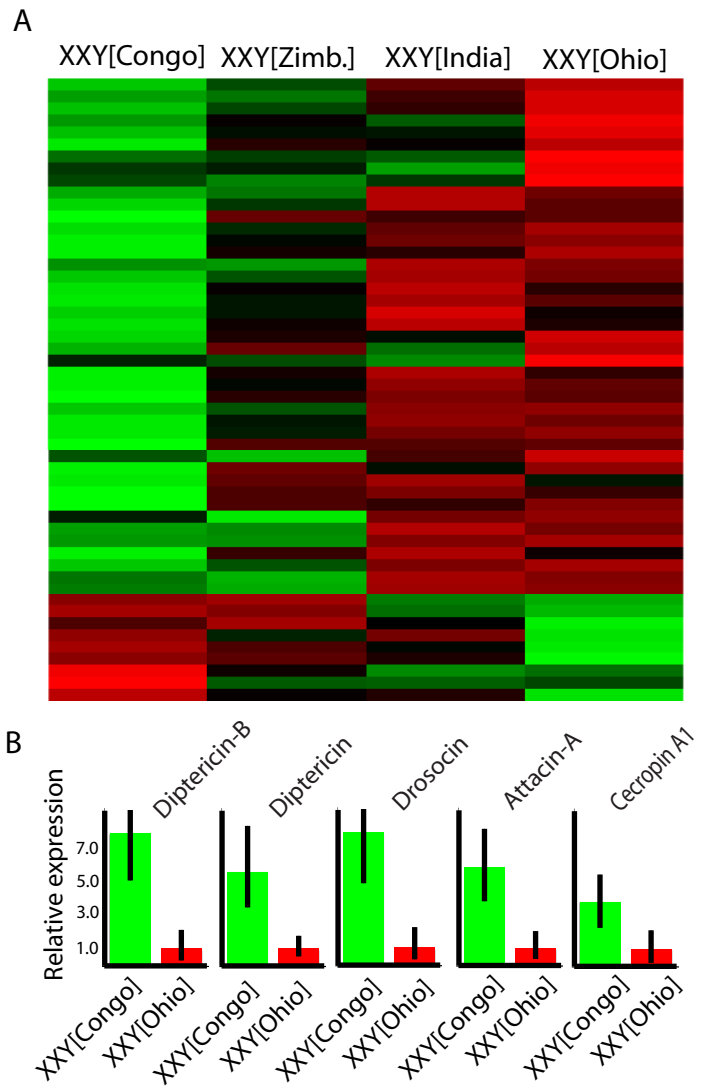


Figure 4. (A) Heat-map of relative expression levels of 52 differentially expressed genes ($P < 0.01$) belonging to the gene ontology categories of defense response and immunity. (B) Examples of immune-related genes more highly expressed in XXY[Congo] (green) relative to XXY[Ohio] (red). Bars denote 95% credible intervals. Relative expression levels are shown with the lowest expression normalized to 1 (red).

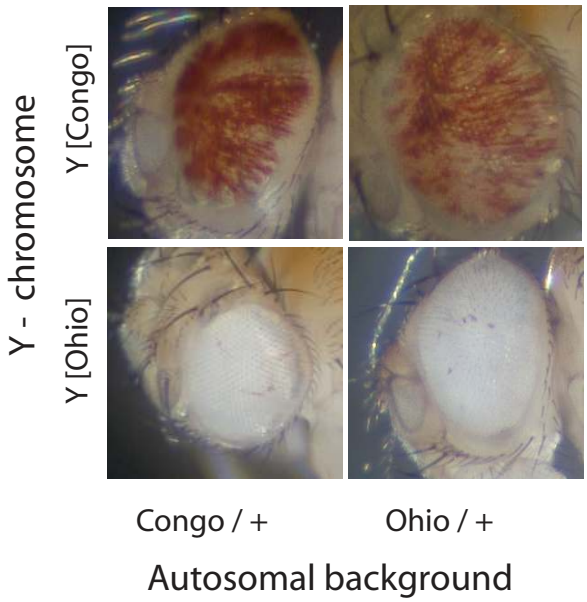
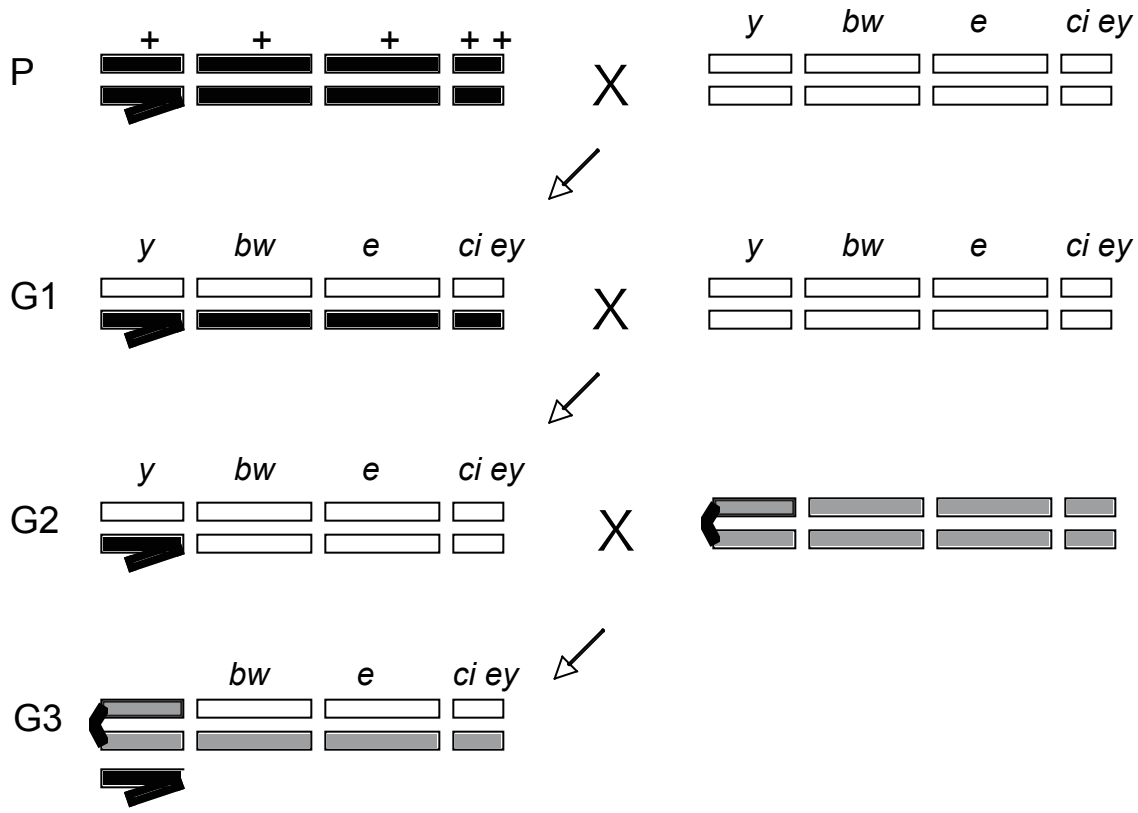
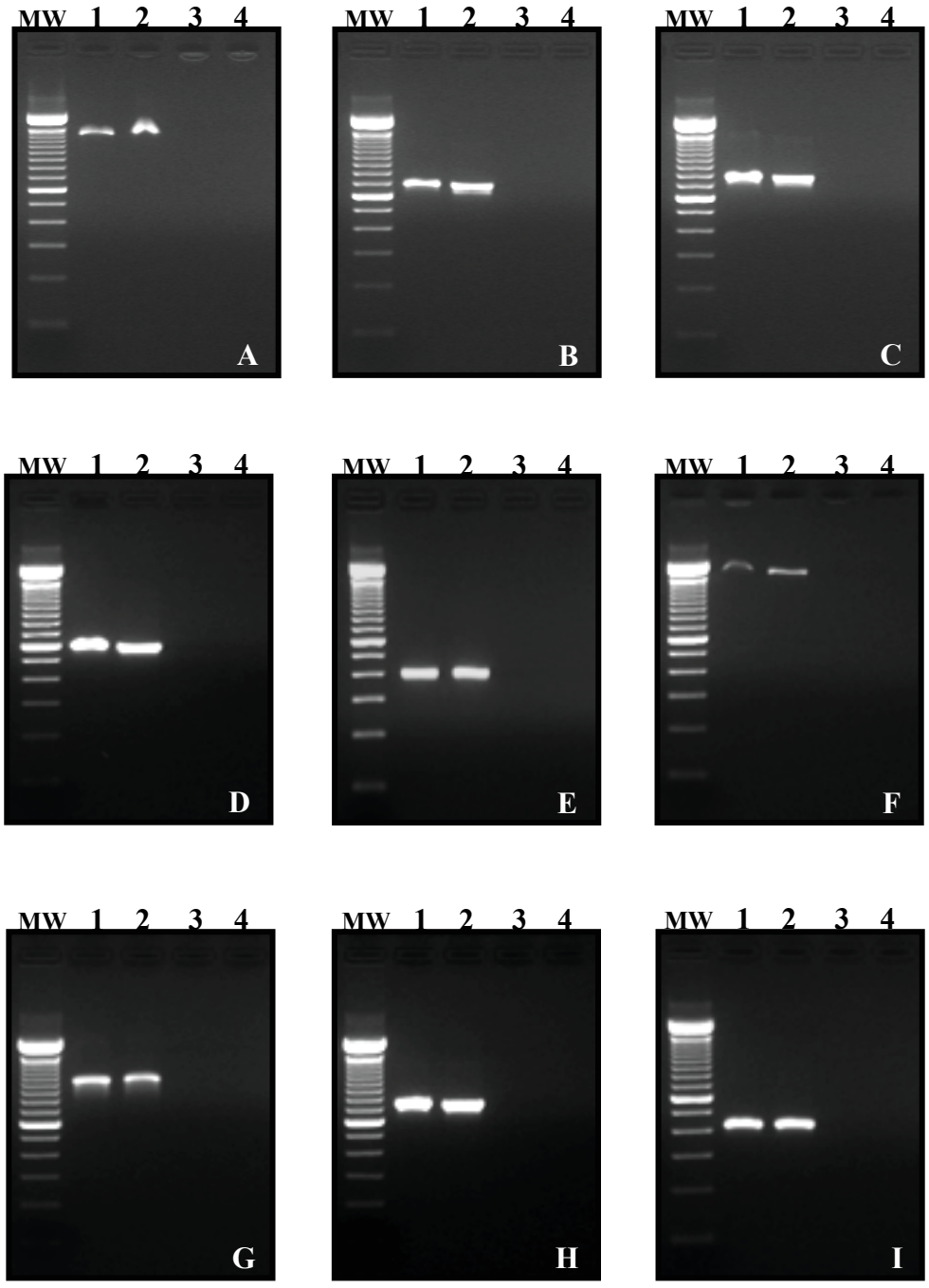


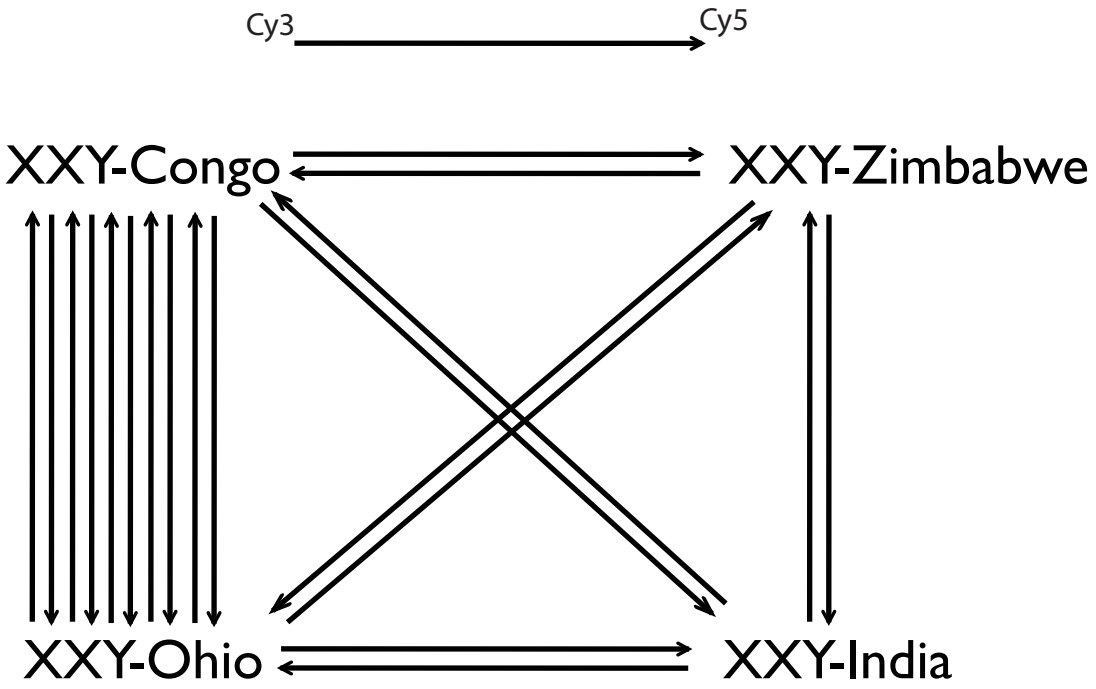
Figure 5. The autosomes from Congo and Ohio do not harbor dominant modifiers of Y-chromosome driven PEV. Eye phenotypes showing variegating eyes for w[m4h] in Y[Congo] and Y[Ohio] in two backgrounds of autosomes (Congo and Ohio). Autosomes from Congo or Ohio are heterozygous with a common set of autosomes inherited from a laboratory strain (+) carrying the X-linked w[m4h] marker.



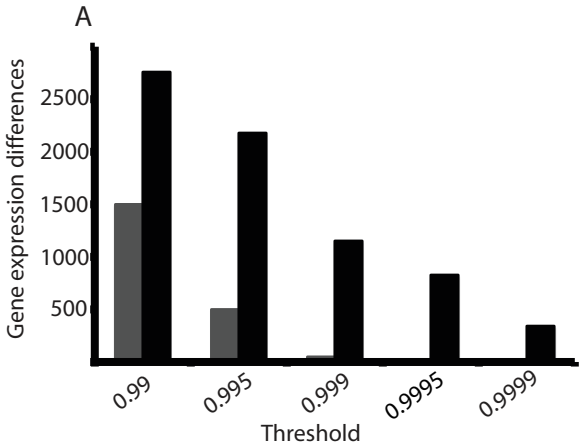
Supplementary Figure 1. Crossing scheme for obtaining population-specific Y chromosome substitution lineages. This was done independently for each Y chromosome lineage to be introgressed such that males from generation G2 have different Y chromosomes introgressed into a common isogenic background. XXY females from generation (G3) also have different Y chromosomes in a background of attached X-chromosome and autosomes that is identical across lineages.



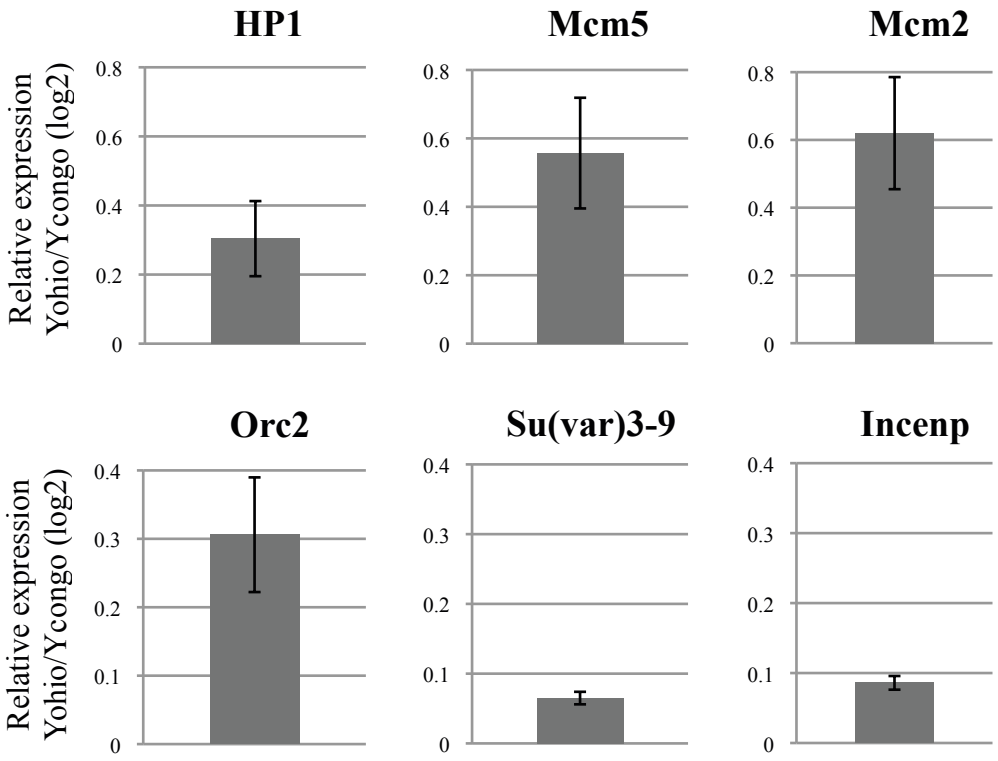
Supplementary Figure 2. Protein coding Y-linked genes are not expressed in XXY females.
RT-PCR to detect transcription of Y-linked genes in XY males and XXY females. Lanes 1: Y-Ohio. Lanes 2: Y-Congo. Lanes 3: XXY-Ohio. Lanes 4: XXY-Congo. (a-c) kl-2 gene. (d-f) kl-3 gene. (g-i) kl-5 gene. MW: 100bp DNA ladder (Invitrogen).



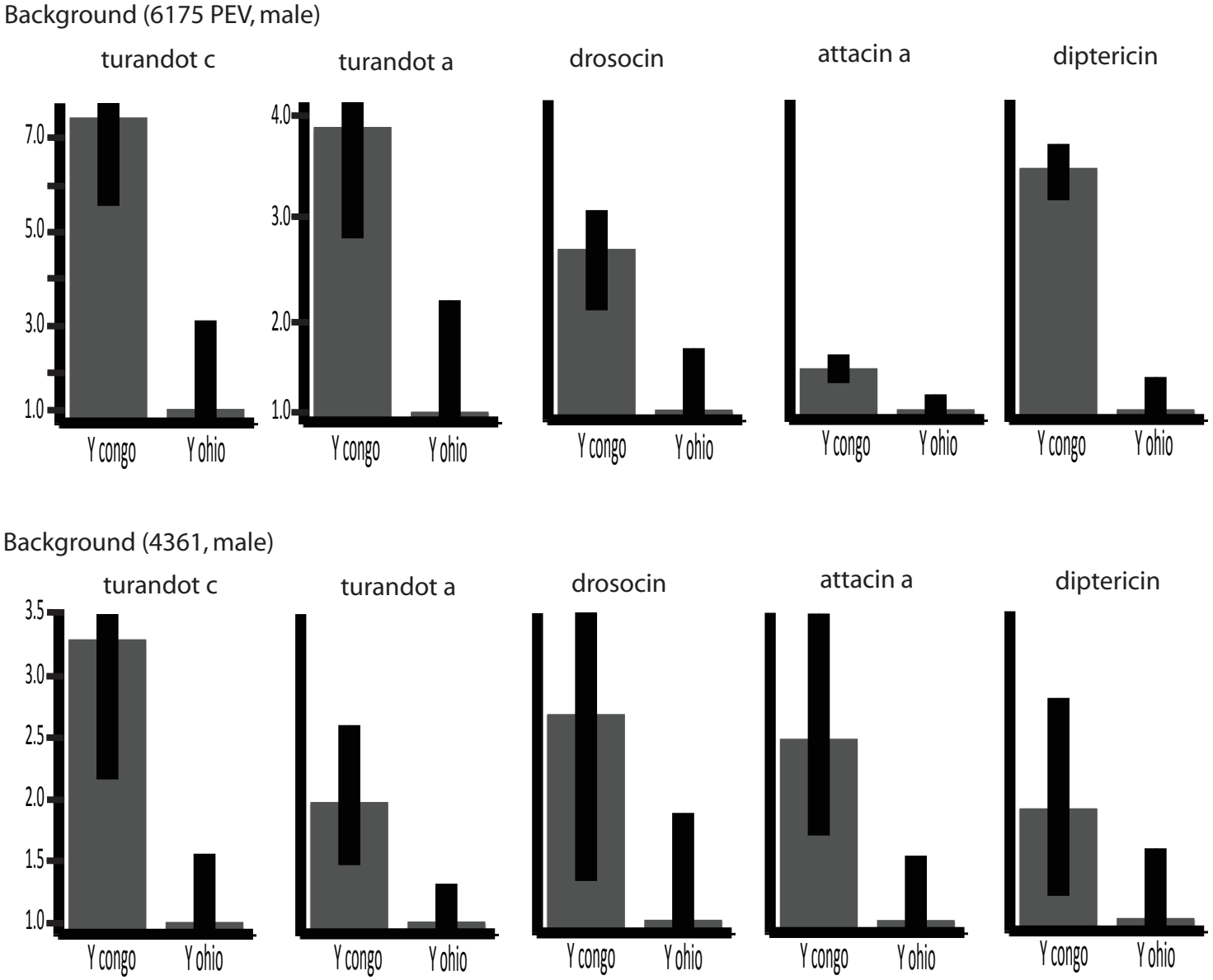
Supplementary Figure 3. Experimental design for collecting genome-wide gene expression data across XXY female genotypes. Arrows indicate direct comparisons (hybridization) of samples on an array slide, for a total of 20 hybridizations. For each hybridization reaction arrow base was labelled with Cy3 dye and arrow head was labelled with Cy5 dye.



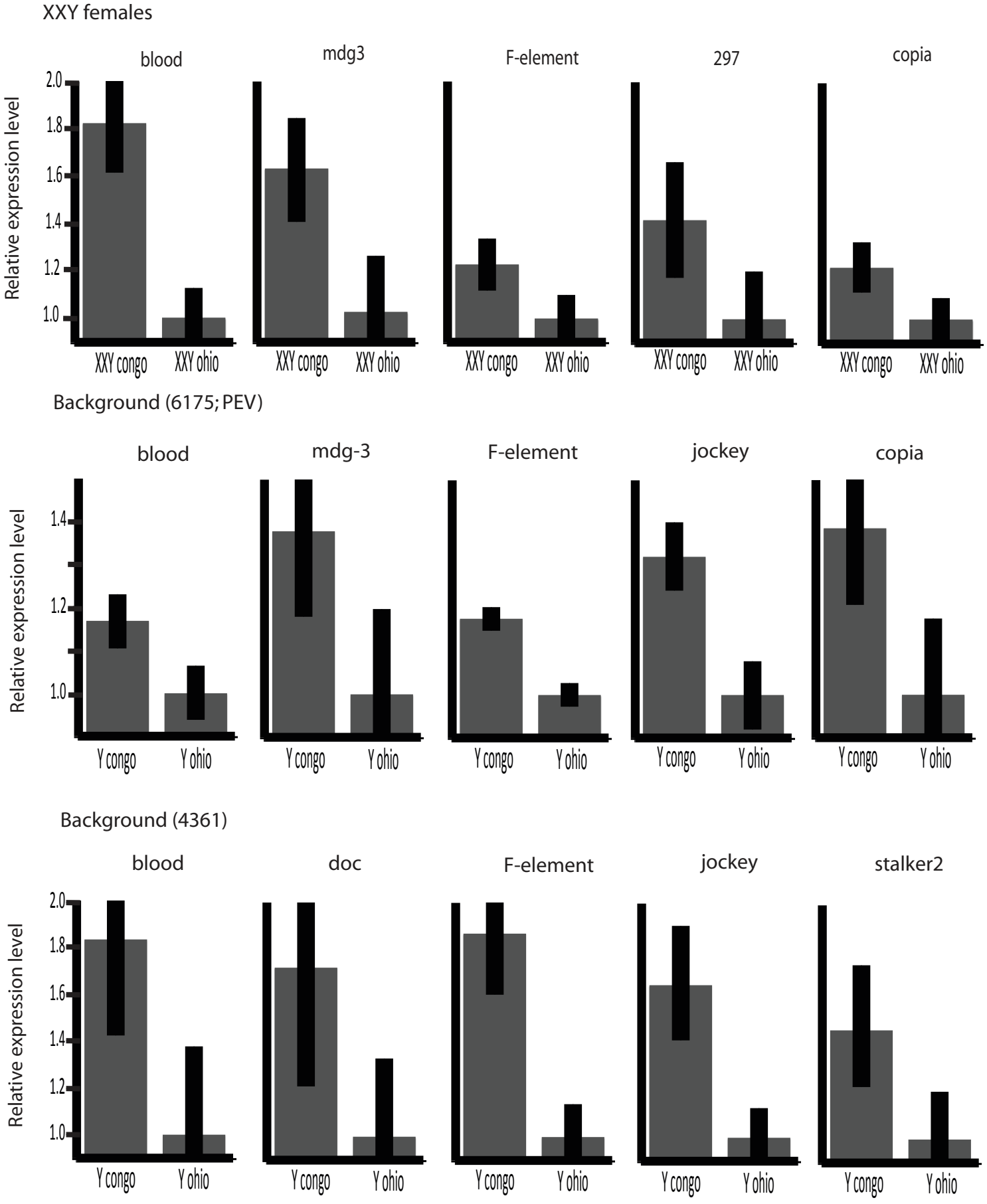
Supplementary Figure 4. Widespread effect of polymorphic Y chromosomes in XXY female genotypes of Drosophila. Number of statistically significant gene expression differences between XXY[Congo] and XXY[Ohio] females (black bars) as a function of the Bayesian posterior probability of differential expression. Gray bars indicate the estimated number of genes expected by chance after permuting the data.



Supplementary Figure 5. Chromatin associated proteins are differentially expressed between Y-Congo and Y-Ohio males. Ratios of between Y-ohio and Y-congo expression obtained with quantitative real-time PCR (Supplementary methods).



Supplementary Figure 6. Examples of immune response genes more highly expressed Y-Congo males relative to Y-Ohio males in two different backgrounds of autosomes and X-chromosome. Background "4361" is the same inbred background used in the Y-chromosome substitution lines reported by Lemos et al (2008). Background "6175" results from the cross of males from the Y-chromosome substitution lines in the "4361" background with females carrying the PEV marker w[m4h].



Supplementary Figure 7. Examples of transposable elements more highly expressed Y[Congo] relative to Y[Ohio] in XXY females and in two different male backgrounds of autosomes and X-chromosome. Background "4361" is the same inbred background used in the Y-chromosome substitution lines reported by Lemos et al (2008). Background "6175" results from the cross of males from the Y-chromosome substitution lines in the "4361" background with females carrying the PEV marker w[m4h].

Supplementary Table 1. Variation in eye pigmentation in stocks carrying the X-chromosome PEV marker *w[m4h]* and Y chromosomes sampled from diverse localities worldwide. Eye pigmentation was assessed with spectrophotometric analysis at an optical density of 480nm. Mean and standard error of the mean are listed.

Strain	Mean	SEM
Y[Congo-1]	0.0563	0.0052
Y[Colombia]	0.0940	0.0073
Y[Connecticut, US]	0.1087	0.0056
Y[South Africa]	0.1200	0.0061
Y[Arizona, US]	0.1270	0.0101
Y[Greece]	0.1310	0.0090
Y[Mauritius]	0.1470	0.0082
Y[Congo-k]	0.1605	0.0314
Y[India]	0.1707	0.0153
Y[Malaysia]	0.1755	0.0049
Y[Massachusetts, US]	0.1898	0.0199
Y[Ohio, US]	0.1934	0.0077
Y[Hawaii, US]	0.2057	0.0415
Y[Congo-11]	0.2473	0.0126
Y[Congo-10]	0.2625	0.0192
Y[Congo-4]	0.3103	0.0030

Supplementary Table 2. Number of gene up-regulated and down-regulated in *Y[congo]* relative to *Y[Ohio]* in *XXY* females across 4 gene ontology categories and a range of significance thresholds for differential expression. The excess of down-regulated genes in *Y[Congo]* belonging to chromatin silencing class is statistically significant (Fisher's exact test, $P < 0.01$). The excess of up-regulated genes in *Y[Congo]* localized to the mitochondrion, involved in electron transport, or associated with defense response and immunity is also statistically significant (Fisher's exact test, $P < 0.01$).

Chromatin Silencing	Up-regulated in *Y[Congo]*	Down-regulated in *Y[Congo]*	% down-regulated in *Y[Congo]*
P < 0.01	7	94	93%
P < 0.005	5	71	93%
P < 0.001	3	34	92%

Mitochondrion	Up-regulated in *Y[Congo]*	Down-regulated in *Y[Congo]*	% up-regulated in *Y[Congo]*
P < 0.01	65	28	70%
P < 0.005	49	16	75%
P < 0.001	29	8	78%

Electron Transport	Up-regulated in *Y[Congo]*	Down-regulated in *Y[Congo]*	% up-regulated in *Y[Congo]*
P < 0.01	51	14	78%
P < 0.005	38	11	78%
P < 0.001	23	6	79%

Defense Response and Immunity	Up-regulated in *Y[Congo]*	Down-regulated in *Y[Congo]*	% up-regulated in *Y[Congo]*
P < 0.01	43	9	83%
P < 0.005	38	7	84%
P < 0.001	31	2	93%