



Y Chromosome Mediates Ribosomal DNA Silencing and Modulates the Chromatin State in *Drosophila*

Citation

Zhou, Jun, Timothy B. Sackton, Lene Martinsen, Bernardo Lemos, Thomas H. Eickbush, and Daniel L. Hartl. 2012. Y Chromosome Mediates Ribosomal DNA Silencing and Modulates the Chromatin State in *Drosophila*. *Proceedings of the National Academy of Sciences* 109, no. 25: 9941–9946.

Published Version

doi:10.1073/pnas.1207367109

Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:12872177>

Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA>

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

The Y chromosome mediates ribosomal DNA silencing and modulates chromatin state in *Drosophila*

Jun Zhou^{1, *}, Timothy B. Sackton¹, Lene Martinsen¹, Bernardo Lemos², Thomas H. Eickbush³, Daniel L. Hartl^{1, *}

1. Department of Organismic and Evolutionary Biology, Harvard University
2. Department of Environmental Health, Harvard School of Public Health
3. Department of Biology, University of Rochester

*. jzhou@oeb.harvard.edu; dhartl@oeb.harvard.edu

Abstract

Although the *Drosophila* Y chromosome is degenerated, heterochromatic, and contains few genes, there is increasing evidence that it plays an important role in regulating the expression of numerous autosomal and X-linked genes. Here, we use 15 Y chromosomes originating from a single founder more than 550 generations ago to study the role of the Y chromosome in regulating rRNA gene transcription, position-effect variegation (PEV), and the link between rDNA copy number, global gene expression and chromatin regulation. Based on patterns of rRNA gene transcription indicated by the transcription of the non-LTR retrotransposon *R2* that specifically inserts into the 28S rRNA gene, we show that X-linked rDNA is silenced in males. The silencing of X-linked rDNA expression by the Y is consistent across geographic populations tested, and independent of genetic background. We also show that these 15 Y chromosomes vary more than three-fold in rDNA locus size and cause drastically different levels of suppression of PEV. The degree of suppression is negatively associated with the number and fraction of rDNA units without transposon insertions, but not with total rDNA locus size. Gene expression profiling revealed that hundreds of genes were differentially expressed among these Y introgression lines, and that global gene expression pattern diverged between the low-PEV and high-PEV flies. Collectively, our findings suggest that the Y chromosome is involved in diverse phenomena related to transcriptional regulation including X-linked rDNA silencing and suppression of PEV phenotype. These results further expand our understanding of the role of the Y chromosome in modulating global gene expression, and suggest a link with modifications of chromatin state.

Introduction

The Y chromosome in *Drosophila melanogaster* is gene-poor, heterochromatic, and largely degenerate owing to the lack of recombination in males and the reduced effective population size of the Y (1-3). Beyond carrying a handful of factors essential for male fertility, the Y chromosome has been regarded as having little function. Despite this relative lack of functional DNA, recent studies have shown that Y-linked genetic variation in *Drosophila* modulates the expression of hundreds of genes across the genome (4-6). While it is presumed that this trans-acting transcriptional regulation is related to epigenetic modification of chromatin state by the Y chromosome, the functional basis for this effect has remained elusive (5).

The *D. melanogaster* Y chromosome contains fewer than 20 single copy coding genes (7-9), with the bulk of the chromosome comprised of repetitive DNA. Among the best-studied Y-linked loci is the rDNA locus (designated *bobbed*), which physically accounts for approximately 10% of the entire Y chromosome, and is homologous to the X-linked rDNA locus. The rDNA locus in *D. melanogaster* is a tandemly repeated array that consists of hundreds of units, with each unit encoding 18S, 5.8S, and 28S ribosomal RNA genes (Fig. 1A). In an rDNA cistron, many of the rDNA units cannot form functional ribosomes due to insertions of the site-specific retrotransposons *R1* and *R2* (10). Partial deletions of the X-linked rDNA locus result in a "bobbed" phenotype in homozygous females, characterized by slow development and shorter bristles. Notably, the bobbed phenotype is not seen in males carrying a *bobbed* X chromosome and a normal Y chromosome (11), suggesting that rDNA transcription from the Y chromosome in males can compensate for X-linked rDNA deletions. However, the functional

interactions between the X-linked and Y-linked rDNA loci are not well known beyond the fact they facilitate the X-Y pairing during meiosis (12).

It is clear that the rDNA locus can have trans-acting regulatory effects on the transcription of rRNA genes: the phenomenon of nucleolar dominance occurs when one rDNA locus is inactivated by another. Mostly observed in interspecific hybrid plants, nucleolar dominance has also been reported in some hybrids of *Drosophila* and *Xenopus* (13). In *D. melanogaster*, silencing of one X-linked rDNA locus by that of an X chromosome of different origin was reported in interspecific hybrid females (14, 15). Examples in intraspecies crosses of *D. simulans* females have also been reported (16). The mechanisms of nucleolar dominance are likely related to chromatin modifications (17), implying a potential role for the rDNA locus in modulating chromatin state. It is not known, however, the relation between Y-linked and X-linked rDNA loci and how related chromatin modifications occur in males.

Another form of transcriptional silencing – position-effect variegation (PEV) – has been shown to be influenced by Y-linked variation in males (5, 18). PEV is the phenomenon where euchromatic genes juxtaposed to a heterochromatin boundary as a result of chromosomal rearrangement are transcriptionally repressed in some cells but not others, based on variation in the amount of heterochromatin spreading. In artificially constructed *D. melanogaster* lines carrying varying amounts of Y chromosomal DNA (18), the amount of Y chromosomal content was found to correlate with the amount of suppression of a PEV marker gene. Later, introgression lines carrying naturally occurring Y chromosomes were also shown to vary in the amount of suppression of PEV marker

genes (5). These results suggest that the Y chromosome plays a role in modulating the chromatin state of the genome.

Here we use a set of mutation-accumulation lines (the Harwich lines) derived from a single founder more than 550 generations ago (Fig. S1 and details in Materials and Methods) to dissect in more detail the potential effects of this hypothesized modulation of genomic chromatin state by the Y chromosome. We first show that the X-linked rDNA locus is silenced in males, suggesting that the Y-linked rDNA locus is transcriptionally dominant in *D. melanogaster*. We then demonstrate that PEV varies across these lines and is correlated with features of the rDNA locus and patterns of global gene expression. Together, these results suggest that Y-linked variation is involved in modulating at least two processes of transcriptional silencing, probably related to chromatin modification. We argue that the Y chromosome plays an important role in modulating global heterochromatin patterns and thus the global transcriptional program and ultimately organismal phenotypes.

Results and Discussion

Evidence for X-chromosomal rDNA silencing in males

To determine the level and status of rDNA transcription, we took advantage of the non-LTR retrotransposable element *R2*, which inserts specifically into the 28S rRNA gene and consequently blocks the function of that particular unit. Since *R2* cotranscribes with the 28S rRNA gene (Fig. 1A, 19), it can be used as a sensitive marker for judging whether the rDNA transcription from either the X-linked or the Y-linked locus is

dominant or additive. Using this approach, we first showed that ribosomal RNA genes are transcribed differently (as large as 40-fold) both between males and females of the Harwich population and across individual lines (Fig. S2). Among the 15 Harwich lines, some lines (*e.g.* H23) have high *R2* transcript levels in females relative to males, while other lines (*e.g.* H15) have high *R2* transcript levels in males relative to females.

Crosses were carried out among the Harwich lines and also a laboratory strain (w^{1118}), which has extremely low levels of *R2* transcription in both males and females. Shown on the left of Fig. 1B, females of a “high *R2* transcription” line H23 were crossed to males of a “low *R2* transcription” line (either H21 or w^{1118}). In F_1 females, which had one X chromosome with a low transcription level and another X chromosome with a high transcription level, the rRNA genes derived from the high-transcription X were silenced. Similar X-X interactions were reported previously as an example of intraspecific nucleolar dominance (16).

In contrast to the pattern we observed in females, where the X chromosome with higher *R2* transcription is silenced, in males the X-linked rRNA genes are consistently silenced regardless of transcription levels. Males of the “high *R2* transcription” line H15 were crossed to three types of females: those from the H23 line (high *R2* transcription in females), from w^{1118} (low *R2* transcription), and from H21 (low *R2* transcription). In all cases the F_1 males retained an *R2* expression level indistinguishable from that of the H15 male (Fig. 1B and Fig. S2), indicating X-linked rDNA silencing. To confirm this result, we expanded our analysis to include males from three different geographic populations (Vienna, Cameroon, and Zimbabwe). These males were crossed with either “low *R2* transcription” (w^{1118}) or “high *R2* transcription” (H23) females. In virtually all cases, the

male progeny exhibited expression levels similar to the male parent (Fig. 1C). These results again strongly suggest that X-linked rDNA is silenced in males.

By backcrossing Harwich Y chromosomes into the w^{118} background, we could also show that the X-linked rDNA silencing is independent of genetic background and stable over generations. After 9 generations of backcrossing, the *R2* transcription levels of males remain consistent with the transcription levels in the original Harwich lines (Fig. 1D). One exception is that this X-linked rDNA silencing was not consistently observed when the Y chromosomes were introgressed into a female background with compound X chromosomes (Fig. S3). However, given the complexity of the regulatory situation when three rDNA loci are present in the flies, it is not clear how relevant this observation is for a normal XY context.

Finally, we examined two sources of molecular evidence for X-linked rDNA silencing in males. First, we monitored the transcription of specific 5' truncated *R2* elements, which allowed us to directly distinguish transcription from the rDNA of the X and the Y chromosomes. Comparing males and females from three Harwich lines (H5, H14, and H15), we found that X-linked *R2* elements are either not transcribed or their transcription is dramatically suppressed in males (Fig. 1E). Second, we examined somatic cells for presence or absence of a secondary constriction on the X or Y chromosome. The rDNA locus generates the nucleolus seen in all interphase cells. As cells approach metaphase, RNA polymerase I is diminished and the nucleolus disappears. A signature of the nucleolus temporarily remains in prometaphase, however, because the previous transcription activity of the rDNA locus shows up as a secondary constriction at the site. This constriction site is in addition to the primary constriction at the centromere. Our

cytological analysis showed that, in males, only a single secondary constriction site associated with nucleolar rDNA transcription can be observed, and it is on the Y chromosome (Fig. 1F), indicating X-linked rDNA transcriptional silencing. Although the molecular basis for X-linked rDNA silencing is not clear, in other cases it involves changes in DNA methylation and histone modification (17).

Y-linked rDNA variation associated with PEV

The rDNA locus on the X chromosome in *Drosophila* species evolves rapidly in size mainly as a result of unequal crossing over between sister chromatids (20, 21), suggesting that the Y-linked rDNA may also evolve rapidly enough to accumulate detectable variation over the 550 generations of evolution since the establishment of the Harwich lines. To further isolate the role of the Y chromosome-linked features, we introgressed the 15 Harwich Y chromosomes into an otherwise identical isogenic background (Bloomington Stock Center Line 4361 with genotype *y[1]; bw[1]; e[4]; ci[1] ey[R]*; Fig. S4). We then determined variation in copy number of rDNA units (using quantitative PCR, see Material and Methods) in the 15 Harwich Y chromosomes by crossing the males of the 15 Harwich Y introgression lines with females containing a compound and bobbed X chromosome. The size of the rDNA locus on the Y varies more than three-fold, from 160 to 500 units, with an average size of 330 units (Fig. 2A). The Y-linked rDNA locus in these lines appears to be larger than the X-linked rDNA locus, which ranges in size from 150 to 300 units (20). Since the size of an average rDNA unit is about 13 kb, including the rRNA genes and intergenic spacers (22, 23), the total size of the Y-linked

rDNA locus ranges from 2-6 Mb among the lines, representing about one tenth of the Y chromosome. This amount of variation at the rDNA locus could have a significant epigenetic influence genome wide (24).

To further explore the possible epigenetic effects of variation across these Y chromosomes, we examined the phenomenon of PEV, which has been shown to be affected by variation in the Y chromosome (5, 18). As a marker for PEV, we used the classic *w[m4h]* system, in which the *white* gene is relocated from its normal euchromatic region to the pericentromeric heterochromatin near the base of the X chromosome. The spread of heterochromatin causes variegated expression of the *white* gene resulting in eye-color mosaicism, a mottled pattern of red and white spots or sectors (25, 26). We observed dramatic variation in eye pigmentation across the Harwich Y chromosomes introgressed into a *w[m4h]* genetic background (Fig. 2B, Fig. S5). Differences in eye pigmentation are weakly (and non-significantly) correlated with rDNA locus size ($\rho = -0.22$, $P = 0.21$, Spearman rank correlation test, Fig. S6), an observation that is consistent with a previous report on Y chromosomes carrying partial deletions of Y-linked rDNA (24).

On the other hand, we found that both the number and the fraction of the rDNA locus that is free of *R1* and *R2* retrotransposable element insertions, termed uninserted rDNA units (20, 21, 27), are significantly and strongly correlated with the PEV phenotype (number of uninserted rDNA units: $\rho = -0.43$, $P = 0.049$, Spearman rank correlation test, Fig. 2C; uninserted fraction of the rDNA locus: $\rho = -0.63$, $P = 0.006$, Spearman rank correlation test, Fig. 2D). Upon insertion, *R1* and *R2* elements render the corresponding rDNA units nonfunctional and only the uninserted units form normal

ribosomal RNA that functions in protein synthesis. It has been suggested that only a fraction of the entire rDNA locus is transcribed, depending on the levels of transposon insertion (16). Heterochromatin formation has been shown to be disrupted in mutants with small rDNA arrays (28), and partial deletions of the rDNA array result in reduced heterochromatin-induced gene silencing (29). The association of uninserted rDNA units with the suppression of the *white* gene expression that we observe is consistent with these studies. Our results also suggest that the uninserted fraction of the rDNA locus may be more important for heterochromatin formation than the overall size of the rDNA array.

A recent study suggested that an increase in rRNA transcription is an indicator of reduced heterochromatin in the genome (20). While R2 serves as a good marker for tracking if R2 inserted rDNA units are transcribed, it is not sufficient to use R2 transcript levels to directly measure the total rRNA gene transcript levels because R2 only inserts into a fraction of the rDNA units and the transcribed fraction of the total units is not known. Therefore, we used gene expression levels of ribosomal protein genes as a proxy for rRNA gene expression. The expression data for ribosomal protein genes derived from microarray analyses comparing low PEV lines (predominantly red eyes, lines H15 and H23) with high PEV (predominantly white eyes, lines H5 and H7). As a class, ribosomal protein genes ($n = 74$) are expressed at a significantly higher level (median increased expression: 10%) in the low PEV group than the high PEV group. ($P < 2.2e-16$, Mann-Whitney U test, Fig. 2E). Since the PEV phenotype is caused by the spread of heterochromatin into the adjacent *w[m4h]* gene, a reduction in heterochromatin is expected to allow greater expression of *w[m4h]*. The relation between both the uninserted fraction of rDNA units and the expression levels of ribosomal protein genes with PEV

might indicate a reduction in heterochromatin in lines with low PEV. However, the mechanism of such chromatin modulation remains unclear.

Global gene expression in high and low PEV lines

The observation that PEV varies considerably among the Harwich Y-chromosome introgression lines suggests that Y-chromosomal modifiers of heterochromatic state have evolved among the Harwich lines within the past 550 generations. These introgression lines also afford the opportunity to ask whether Y-linked regulatory variation (4) has evolved among the Harwich Y chromosomes as well. We selected two representative lines from the predominantly red-eyed (low PEV) and predominantly white-eyed (high PEV) Harwich Y-chromosome introgression lines: H15 and H23 for low PEV and H5 and H7 for high PEV. We measured global gene expression using two-color microarrays across these four lines (Fig. S7).

Notably, we find strong evidence that global patterns of gene expression differ significantly between the low-PEV and high-PEV flies. First, we see a bias towards up-regulation in genes that are differentially expressed between red-eyed and white-eyed flies, with twice as many genes upregulated in red-eyed flies (93 vs 46 with a 10% false discovery rate; Fig. 3A). This difference contrasts with comparisons within PEV class, where an equal number of genes are upregulated vs. downregulated (13/13 for H5 vs H7, and 60/52 for H15 vs H23). Second, we see clear functional coherence, as indicated by Gene Ontology (GO) enrichment, among genes differentially expressed between PEV groups with several functional categories enriched in both the set of genes up-regulated in

the low PEV group and the set of genes down-regulated in the low PEV group (Table S1). The finding that pheromone-response and odorant-binding genes are enriched is of interest in view of a previous report (31) and their location near euchromatin-heterochromatin boundaries. Third, hierarchical clustering of our four samples based on whole-genome gene-expression levels strongly clusters the two low-PEV lines and the two high-PEV lines into separate groups (Fig. 3B). Finally, a correlation matrix of all pairwise fold-change ratios between the four lines (H5, H7, H15, and H23; Fig. S8) clearly shows that the inter-class comparisons (H15 vs H5, H15 vs H7, H23 vs H5, H23 vs H7) are more correlated than the intra-class comparisons (H5 vs H7, H15 vs H23). Taken together, these results suggest that the high-PEV and low-PEV classes represent significantly diverged Y chromosomes that affect global gene expression differently. This global gene expression divergence emerged very quickly owing to the accumulation of Y-linked regulatory variation within only 550 generations.

One attractive hypothesis is that epigenetic regulation of X-linked and autosomal gene expression by the Y chromosome may be mediated at the level of chromatin by Y-linked rDNA, microsatellites, and other noncoding repetitive sequences. While a direct test of this hypothesis is beyond the scope of the current work, the chromatin-level hypothesis predicts that genes affected by Y-linked regulatory variation (YRV) should be physically clustered along chromosomes (32). In order to test this hypothesis, we calculated the number of genes significantly upregulated (or downregulated) in high PEV vs. low PEV lines for sliding windows of 1 Mb with a step size of 500 kb (Fig. 3C). Several regions of the genome show a significant excess of genes whose expression is affected by Y chromosome (at a nominal P-value < 0.01). Taken together, these results

imply a link between the chromatin modification mechanisms that drive PEV and the global patterns of gene expression regulation implicated in Y-linked regulatory variation.

Conclusions

In summary, we presented biological evidence for X-linked rDNA silencing in males of *D. melanogaster*. We also showed that the uninserted rDNA units are correlated with the suppression of PEV phenotype. In addition, hundreds of genes were found differentially expressed among the Y introgression lines, and global gene expression pattern clearly diverged between the low-PEV and high-PEV flies.

The Y chromosomes we have analyzed here have accumulated mutations for 550 generations. While we have shown that they have diverged considerably at the Y-linked rDNA locus, they have almost certainly diverged in other ways as well. The Y contains abundant microsatellite repeats, degenerated transposon sequences, and other repetitive elements that are also likely to evolve over this time scale and that could impact global patterns of gene expression. For example, abundant microsatellite sequences residing on the Y chromosome may compete for binding with transcription factors, or degenerated transposon sequences accumulated in the heterochromatic regions may provide a source of small interfering RNAs that modify chromatin structure. Such Y-linked noncoding sequence variation may jointly impact phenotypes such as PEV, X-linked rDNA silencing, and global patterns of gene expression. Future work will be necessary to disentangle the effects of individual repetitive elements on the Y chromosome.

Materials and Methods

Fly sources: The Harwich mutation-accumulation lines were originally subdivided from a single founder line and maintained for 100 generations with 10 pairs of unselected flies per generation (33). After the initial 100 generations, the lines were further maintained by mass transfer every generation in regular laboratory conditions (20). The effective population size (N_e) is about 20. w^{1118} is a stock maintained in the lab. The nine geographic lines (3 Vienna lines, 3 Cameroon lines, 3 Zimbabwe lines) were gifts from James Fry at the University of Rochester. The compound X chromosome (attached $X^{\wedge}XY$) stocks $C(1)DX, y$ and $C(1)RM$, the Y-introgression background line 4361 and the PEV background line B6175 were all obtained from the Bloomington Drosophila stock center.

RNA-blots: Total RNA was extracted from 25 adult females or 50 adult males. Ten micrograms of RNA were separated on 1.2% agarose, 2.2M formaldehyde gels, the RNA transferred to GeneScreen Plus, and hybridized with an anti-sense RNA probe from the 5' end of the R2 element as previously described (21). The relative levels of the 3,600 nt full-length R2 transcript were scanned using a PhosphorImager screen and quantified in the software ImageQuant (GE Healthcare). As a control for RNA loading and quality, all R2 hybridization signals were standardized by monitoring the level of the ribosomal protein gene rp49 hybridization on the same blots. PCR products containing promoter sequences for T7 polymerase were generated for synthesis of rp49 antisense RNA probe. P-32 labeled antisense RNA was generated by using T7 polymerase under the conditions suggested by the supplier (Invitrogen). The primers for the rp49 transcript were 5'-

CAGCATAACAGGCCCAAGATC-3' and 5'-

GTAATACGACTCACTATAGGGCAGTAAACGCGGTTCTGCATG-3'.

DNA-blots: Total DNA was extracted from 50 adult males. The Southern blots were exactly as described by Averbeck and Eickbush (20). Briefly, genomic DNA was digested by restriction enzymes *Bam*HI, *Cla*I and *Pst*I that specifically cut the rDNA and inserted elements. Genomic DNA fragments that contain R1 or R2 insertions and uninserted rDNA by R1/R2 were then separated on 1% agarose gel, and further hybridized with a probe from the 28S rRNA gene. The fractions of uninserted and inserted rDNA units were determined using PhosphorImager screen on Typhoon scanner (GE healthcare) and quantified in the software ImageQuant (GE Healthcare).

Reverse transcription PCR: Reverse transcription-PCR (RT-PCR) was conducted with the M-MLV reverse transcriptase under the conditions suggested by the supplier (Invitrogen). Briefly, 1 µg of total RNA was incubated at 65 °C for 5 min with a R2 primer (1 µM) near the 3' end and dNTPs (deoxynucleotides, 0.5 mM). The primer sequence was 5'- GTATGGAAATCTATCGAAAGATACT-3' which is at the R2 3' end. After quick chilling on ice, first strand buffer (1X), DTT (0.01M) and RNase inhibitor (RNaseOUT, Invitrogen) were added. The reaction mixtures were then incubated at 37 °C for 2 min. The M-MLV reverse transcriptase (200 units) was then added and the incubation continued for 50 min. 2 µl of the RT reaction mixtures were used in a standard PCR reaction. A negative control including the same reactions without the M-MLV reverse transcriptase was conducted at the same time. The primers used for standard PCRs are: 5'-TGCCCAGTGCTCTGAATGTC-3' (forward), which anneals to 28S gene

sequences 80bp upstream of the R2 insertion site -3' and 5'-

GCATGCACGATTCATTGCTC-3' (reverse), which is located 60bp upstream from the R2 3' end. Therefore a nested PCR approach was employed using a primer upstream of the one used for RT.

Cytology: Brain tissue was dissected from female 3rd-instar larvae, fixed and squashed as described (34). Rehydrated tissue was stained with 0.4 µg/ml DAPI for 2 min, washed briefly in water, and mounted in SlowFade reagent. All observations and photography were conducted on a Nikon DEC512 fluorescence microscope.

Quantitative PCR: To assess the Y-linked rDNA locus size, males from the Harwich Y introgression lines were crossed to females of a stock with attached X chromosomes where the rDNA units have been largely deleted on both X chromosomes. At least five biological replicates (F₁ male progeny) consisting of three experimental replicates for each line were sampled. Q-PCR protocol is the same as described by Paredes and Maggert (35). They were carried out with ABI Step-One Real-time PCR machine (Applied Biosystems). The tRNA gene *tRNA*^{K-CTT} (15 copies per haploid genome) was used as a reference for calculating the copy number of rRNA genes. The primer sequences are identical as described by Paredes and Maggert (35).

Eye-pigment assay: Males from the 15 Harwich Y-introgression lines were crossed to females from a stock carrying *w[m4h]*. Male progeny from the crosses were collected and aged in 25 °C for three days, then flash frozen in liquid nitrogen. Heads were removed and homogenized with 20 µl of acidified ethanol. Eye pigment expression was measured using Nanodrop (Thermo Scientific) at the wavelength of 480 nm. Each

measurement contained five heads of each line, and six to ten biological replicates were taken per line. Pictures of eye-pigmentation phenotype (PEV phenotype) were imaged with an automontage system (Snycroscopy, Frederick, MD).

Microarray processing and gene expression analyses: Four Y introgression lines with H5 and H7 representing high PEV group and H15 and H23 representing low PEV group were selected to perform genomic gene expression profiling analyses. Microarrays were 18,000-feature spotted cDNA arrays. RNA samples were extracted from 50 adult males that were collected and aged in 25 °C for three days and flash frozen using liquid nitrogen. Total RNA was extracted from whole flies using TriZol (Life Technologies). Microarray processing including cDNA synthesis, labeling with fluorescent dyes (Cy3 and Cy5) and hybridizations were carried out using 3DNA kit (Genisphere Inc.), and scanned in Axon 4000B scanner (Axon Instruments). After hybridization and scanning, raw data were background corrected and normalized in the R package Limma, using the normexp method for background correction and the loess method for normalization. Linear models were fit in Limma. Gene ontology analyses were conducted with the software GeneMerge (36). To estimate differences in expression between low PEV and high PEV lines, we fit the contrast $(H15+H23)/2 - (H7/2)$, where H5 functions as the reference. All results are presented after multiple test correction using the false discovery rate method implemented in Limma (38). We consider genes significant if they survive multiple test correction at an FDR of 10%. Microarray gene expression data can be obtained from the GEO database (GSE 37068).

Gene clustering analyses: Samples (H5, H7, H15, and H23) were hierarchically clustered using the R package pvclust. Since H5 is treated as the reference in the design

matrix and fold change values are on a log₂ scale, we defined H5 expression as 0 for all genes. We calculated a distance matrix using Euclidean distance, and then generated a hierarchical clustering using Ward's minimum variance method of clustering. To estimate significance, we used the pvclust function to perform a multiscale bootstrap resampling with 10,000 replicates. Approximately unbiased (AU in Figure) p-values calculated by pvclust represent the P-value for the test that the hypothesis of no cluster is true; high P-values, therefore, indicate strong support for the cluster.

Acknowledgement This work was supported by National Institutes of Health grants GM042790 (to T.H.E.) and GM084236 (to D.L.H.).

References

1. Bull JJ (1983) Evolution of Sex Determining Mechanisms (Benjamin Cummings, London).
2. Charlesworth B, Charlesworth D (2000) The degeneration of Y chromosomes. *Philos Trans R Soc Lond B* 355:1563-1572.
3. Zurovcova M, Eanes WF (1999) Lack of nucleotide polymorphism in the Y-linked sperm flagellar dynein gene *Dhc-Yh3* of *Drosophila melanogaster* and *D. simulans*. *Genetics* 153:1709-1715.
4. Lemos B, Araripe LO, Hartl DL (2008) Polymorphic Y chromosomes harbor cryptic variation with manifold functional consequences. *Science* 319:91-93.
5. Lemos B, Branco A, Hartl DL (2010) Epigenetic effects of polymorphic Y chromosomes modulate chromatin components, immune response, and sexual conflict. *Proc Natl Acad Sci USA* 107:15826-15831.

6. Sackton TB, Montenegro H, Hartl DL, Lemos B (2011) Interspecific Y chromosome introgressions disrupt testis-specific gene expression and male reproductive phenotypes in *Drosophila*. *Proc Natl Acad Sci USA* 108:17046-51.
7. Carvalho AB, Lazzaro BP, Clark AG (2000) Y chromosomal fertility factors *kl-2* and *kl-3* of *D. melanogaster* encode dynein heavy chain polypeptides. *Proc Natl Acad Sci USA* 97:13239-13244.
8. Carvalho AB (2002) Origin and evolution of the *Drosophila* Y chromosome. *Curr Opin Genet Dev* 12:664-668.
9. Krsticevic FJ, Santos HL, Januário S, Schrago CG, Carvalho AB. (2010) Functional copies of the Mst77F gene on the Y chromosome of *Drosophila melanogaster*. *Genetics* 184:295-307.
10. Jakubczak JL, Burke WD, Eickbush TH (1991) Retrotransposable elements R1 and R2 interrupt the rRNA genes of most insects. *Proc Natl Acad Sci USA* 88:3295-99.
11. Sivertzev-Dobzhansky NP and Dobzhansky T (1932) Deficiency and Duplications for the Gene Bobbed in *Drosophila Melanogaster*. *Genetics* 18:173-92.
12. Tsai JH, McKee BD (2011) Homologous pairing and the role of pairing centers in meiosis. *J Cell Sci* 124:1955-63.
13. Pikaard CS (2000) Nucleolar dominance: uniparental gene silencing on a multi-megabase scale in genetic hybrids. *Plant Mol Biol* 43:163-177.
14. Durica DS, Krider HM (1977) Studies on the ribosomal RNA cistrons in interspecific *Drosophila* hybrids. I. Nucleolar dominance. *Dev Biol* 59: 62-74.
15. Durica DS, Krider HM (1978) Studies on the ribosomal RNA cistrons in interspecific *Drosophila* hybrids. II. Heterochromatic regions mediating nucleolar dominance. *Genetics* 89: 37-64.
16. Eickbush DG, Ye J, Zhang X, Burke WD, Eickbush TH (2008) Epigenetic regulation of retrotransposons within the nucleolus of *Drosophila*. *Mol Cell Biol* 28: 6452-61.
17. Tucker S, Vitins A, Pikaard CS (2010) Nucleolar dominance and ribosomal RNA gene silencing. *Curr Opin Cell Biol* 22:351-6.
18. Dimitri P, Pisano C (1989) Position effect variegation in *Drosophila melanogaster*: relationship between suppression effect and the amount of Y chromosome. *Genetics* 122:793-800.
19. Eickbush DG, Eickbush TH (2010) R2 retrotransposons encode a self-cleaving ribozyme for processing from an rRNA cotranscript. *Mol Cell Biol* 30:3142-50.
20. Averbek KT, Eickbush TH (2005) Monitoring the mode and tempo of concerted evolution in the *Drosophila melanogaster* rDNA locus. *Genetics* 171:1837-46.

21. Zhou J, Eickbush TH (2009) The pattern of R2 retrotransposon activity in natural populations of *Drosophila simulans* reflects the dynamic nature of the rDNA locus. *PLoS Genet* 5:e1000386.
22. Wellauer PK and Dawid IB (1977) The structural organization of ribosomal DNA in *Drosophila melanogaster*. *Cell* 10:193-212.
23. Pellegrini M, Manning J, Davidson N (1977) Sequence arrangement of the rDNA of *Drosophila melanogaster*. *Cell* 10:213-4.
24. Paredes S, Branco AT, Hartl DL, Maggert KA, Lemos B (2011) Ribosomal DNA deletions modulate genome-wide gene expression: “rDNA-sensitive” genes and natural variation. *PLoS Genet* 7:e1001376.
25. Talbert PB and Henikoff S (2006) Spreading of silent chromatin: inaction at a distance. *Nat Rev Genet* 7:793-803.
26. Schulze SR and Wallrath LL (2007) Gene regulation by chromatin structure: paradigms established in *Drosophila melanogaster*. *Annu Rev Entomol* 52: 171-192.
27. Zhang X, Eickbush TH (2005) Characterization of active R2 retrotransposition in the rDNA locus of *Drosophila simulans*. *Genetics* 170:195-205.
28. Peng JC, Karpen GH (2007) H3K9 methylation and RNA interference regulate nucleolar organization and repeated DNA stability. *Nat Cell Biol* 9:25-35.
29. Paredes S, Maggert KA (2009) Ribosomal DNA contributes to global chromatin regulation. *Proc Natl Acad Sci U S A* 106:17829-34.
30. Larson K, et al. (2012) Heterochromatin formation promotes longevity and represses ribosomal RNA synthesis. *PLoS Genet* 8:e1002473.
31. Jiang P-P, Hartl DL, Lemos B (2010) Y not a dead end: Epistatic interactions between Y-linked regulatory polymorphisms and genetic background affect global gene expression. *Genetics* 186:109-118.
32. Kharchenko PV, et al. (2011) Comprehensive analysis of the chromatin landscape in *Drosophila melanogaster*. *Nature* 471:480-5.
33. Mackay TFC, Lyman RF, Jackson MS, Terzian C, Hill WG (1992) Polygenic Mutation in *Drosophila Melanogaster*: estimates from divergence among inbred strains. *Evolution* 46:300-16.
34. Krider HM, Plaut W (1972) Studies on nucleolar RNA synthesis in *Drosophila melanogaster*. I. The relationship between number of nucleolar organizers and rate of synthesis. *J Cell Sci* 11:675-87.
35. Paredes S, Maggert KA (2009) Expression of I-Crel endonuclease generates deletions within the rDNA of *Drosophila*. *Genetics* 181:1661-71.

36. Castillo-Davis CI, Hartl DL (2003) GeneMerge--post-genomic analysis, data mining, and hypothesis testing. *Bioinformatics* 19:891-2.
37. Townsend JP (2004) Resolution of large and small differences in gene expression using models for the Bayesian analysis of gene expression levels and spotted DNA microarrays. *BMC Bioinformatics* 5:54.
38. Benjamini Y, and Hochberg Y (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Statist Soc B* 57:289–300.

Figure legends

Fig. 1: Multiple lines of evidence suggest that the X-linked rDNA is silenced in males.

(A) Schematic of the X and Y chromosome in *D. melanogaster* and the ribosomal loci containing the retrotransposable element R2. The bottom panel shows a magnified rDNA locus containing R2 elements which insert into a specific site in the 28S rRNA genes. R2 is co-transcribed with the rRNA units. Therefore, R2 transcription can be used to represent rRNA gene transcription. (B) High levels of expression for the X23 rDNA locus can be repressed by a low expressing rDNA loci on another X. However, high levels of expression from the Y15 rDNA locus cannot be shut off by an X rDNA locus. (C) Nine wild Y chromosomes from three different geographic populations demonstrated preferential Y-linked rDNA expression, also suggesting the X-linked rDNA silencing in the males of those crosses. The rDNA locus in W1118 has low transcriptional levels while H23 has high expression levels. (D) Four Y chromosomes with different levels of

rDNA and R2 expression were introgressed into W1118 background. The bar graph shows similar levels of the Y-linked rDNA and R2 expression in the introgressed background compared to their original levels. (E) RT-PCR results demonstrate the expression of specific 5' truncated R2 elements in males and females. The arrows except for the bottom one show transcription of specific R2 copies are silenced in males. The bottom arrow shows a transcribed specific R2 copy on the Y. (F) Premetaphase chromosomes from adult brain tissue stained with DAPI. Cytological analyses indicate that only the Y chromosome forms a secondary constriction site, indicative of active transcription of the rDNA locus in the previous interphase in males. The Y chromosomes of H8 (a,c), H15(d,e) and H23(b) all demonstrated secondary constriction at the rDNA loci, while the X chromosome loci did not. Arrows point to the locations of the rDNA loci on X and Y chromosomes.

Fig. 2: Y-linked rDNA variation correlates with PEV phenotype. (A) The size of the Y-linked rDNA varies more than three-fold. (B) Representative images showing the variation in levels of eye pigmentation among the Harwich Y-introgression lines crossed to the *w[m4h]* PEV background. (C) The number of transposon uninserted rDNA units is negatively correlated with the PEV phenotype (expression of the *w* gene measured by eye pigment absorbance at 480 nm). (D) The uninserted fraction of the rDNA locus is negatively correlated with the PEV. (E) Boxplot showing distribution of $\log_2(\text{low PEV expression}) - \log_2(\text{high PEV expression})$ for genes encoding ribosomal proteins (based on Ribosomal Protein Gene database accessed at <http://ribosome.miyazaki-med.ac.jp/>), compared to the distribution of the same fold change for genes not encoding ribosomal

proteins. Expression of ribosomal genes is significantly higher in the low PEV lines (Mann-Whitney U test, $P < 2.2e-16$).

Fig. 3: Global gene expression divergence in low and high PEV lines. (A) A volcano plot demonstrates there are more genes up-regulated in red-eyed, low PEV flies (H15 and H23) compared to white-eyed, high PEV flies (H5 and H7) than genes down-regulated in that comparison. Black dots indicate differentially expressed genes, while grey dots indicate genes that are not differentially expressed. The dots with values larger than 0 indicate genes that are up-regulated in low PEV flies and with values smaller than 0 indicate down-regulated genes in low PEV flies. The GEL_{50} value, as a proxy of power, measuring the gene expression level at which there is a 50 percent chance of detection of statistical significance is 1.85-fold in this study (37). (B) The expression data for the four low and high PEV lines were hierarchically clustered using the R package pvclust. A dendrogram shows that H15 and H23, H5 and H7 are clustered with each other, respectively for global gene expressions. (C) To estimate spatial clustering along the genome, the number of significantly up-regulated or down-regulated genes was calculated (separately) in each 1 MB sliding window across the genome, with a step size of 500 kb. To estimate significance, a number of genes equal to the actual number of significant genes were randomly sampled 10,000 to get a null expectation for each window. Windows marked with a red asterisk are significantly elevated at a nominal $P < 0.01$.

Supplementary materials

Table S1. Gene ontology enrichment in the comparison of low and high PEV flies.

Red vs White	Category	Description	N	P-value*
Up-regulated in Red-eye (low PEV) flies	BP			
	GO:0019236	response to pheromone	6	8.5E-09
	GO:0007606	sensory perception of chemical stimulus	5	5.2E-05
	GO:0042048	olfactory behavior	4	5.6E-04
	GO:0006810	transport	9	1.0E-03
	CC			
	GO:0005576	extracellular region	11	7.6E-07
	MF			
	GO:0005549	odorant binding	5	1.7E-05
	GO:0005550	pheromone binding	3	2.3E-05
	GO:0004252	serine-type endopeptidase activity	5	1.2E-02
Down-regulated in Red-eye (low PEV) flies	BP			
	GO:0006961	antibacterial humoral response	4	2.7E-05
	GO:0050832	defense response to fungus	3	3.3E-05
	CC			
	GO:0005575	cellular_component	5	8.6E-02
	MF			
	NA	NA	NA	NA

* P-values were corrected for multiple hypothesis testing.

For abbreviations used, BP: biological processes; MF: molecular function; CC: cellular component; NA, not available (no enrichment identified).

Fig. S1: Background of Harwich mutation-accumulations lines. A diagram briefly shows the history of the 15 Harwich lines that were split from a single line 550 generations ago.

Fig. S2: R2 transcription in males and females of Harwich lines. A northern blot probed with R2 5' antisense RNA probe shows R2 expression levels in males and females of Harwich lines. The same blot probed with antisense RNA generated from ribosomal protein gene *rp49* was conducted as an RNA loading control. The bottom bar graph shows quantitative results. The other five Harwich lines that are not shown in this blot exhibited low levels in both males and females.

Fig. S3: R2 expression levels in X^1X^2 females. H15 and H23 males were crossed to females of two different attached X^1X^2 stocks. These chromosomes are abbreviated as X^1X^1 and X^1X^2 , respectively. The first two columns represent the R2 transcript levels in the original females of these two attached X^1X^2 stocks. The remaining columns compare the R2 transcript levels in the female progeny of crosses to the parental males of H15 and H23.

Fig. S4: Diagram of introgression of 15 Harwich Y chromosomes. Crossing scheme for introgressing Harwich Y chromosomes into the 4361 laboratory stock background. *y(yellow)*, *bw(brown)*, *e(ebony)*, *ci(cubitus interruptus)*, *ey(eyeless)* were used as recessive genetic markers in the isogenic background line.

Fig. S5: (A) Cross scheme to investigate the suppression effects of 15 Harwich Y chromosomes on PEV phenotype. (B) Quantitative results of eye pigmentation measured by the absorbance at 480 nm.

Fig. S6: The association between the rDNA locus size and the suppression degree of the *w* gene expression level in 15 Harwich Y-chromosome introgression lines.

Fig. S7: Experimental design of microarrays for gene expression profiling. The lines with arrow heads represent hybridizations in which the Cy3 and Cy5 dyes were swapped and replicated.

Fig. S8: Gene expression matrix plot for the four Harwich Y introgression lines. In addition to fitting a contrast to estimate expression differences between high-PEV and low-PEV lines, we fit six pairwise models comparing each individual Harwich line to all others: H15-H7, H15-H5, H23-H7, H23-H5, H7-H5, and H23-H15. We plotted the correlations among these pairwise comparisons using hexbin plots with the R function `hexplom`. Each cell of the plot shows the correlation between two pairwise comparisons specified by the row and column labels in the central diagonal. Correlations between the pairwise comparisons that represent comparisons between a high-PEV and a low-PEV line (H15-H7, H15-H5, H23-H7, H23-H5; shaded gray) are stronger than correlations between pairwise comparisons that represent comparisons between either the two high-PEV or the two low-PEV lines (H7-H5, H23-H15). That is, the genes that are

differentially expressed between any high-PEV and low-PEV line are similar, leading to positive correlations among the pairwise comparisons shaded in grey.

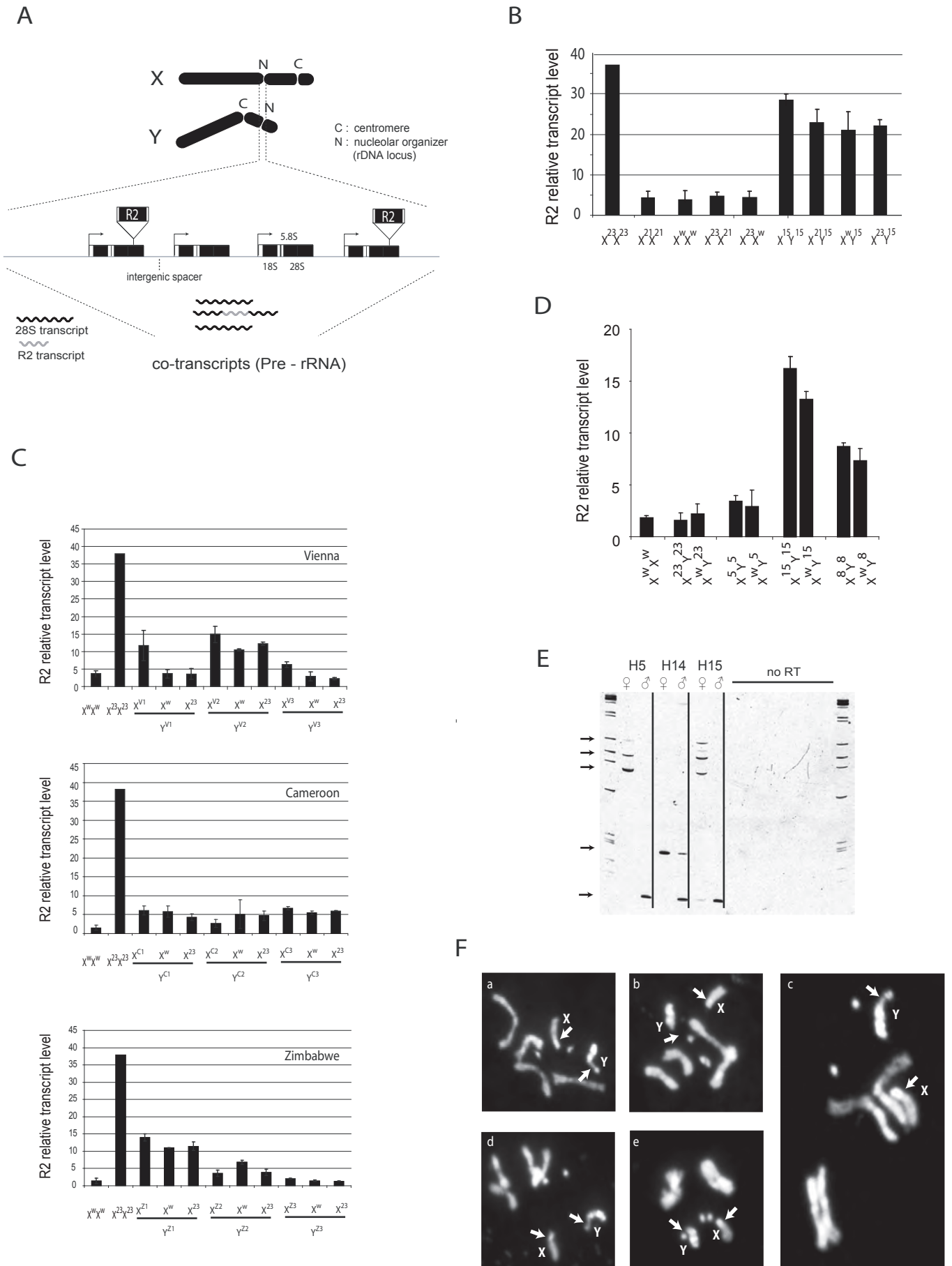


Figure 1

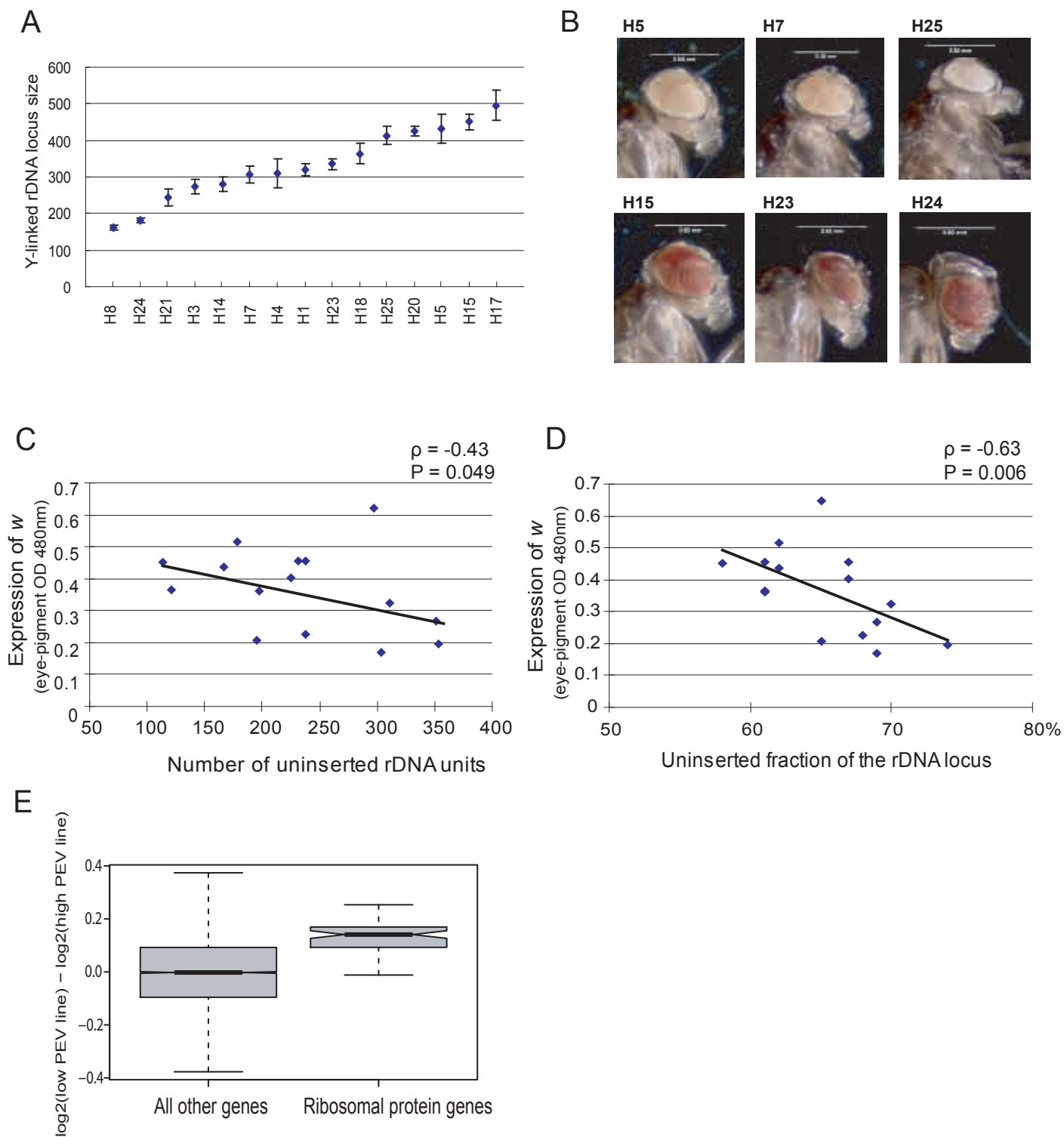


Figure 2

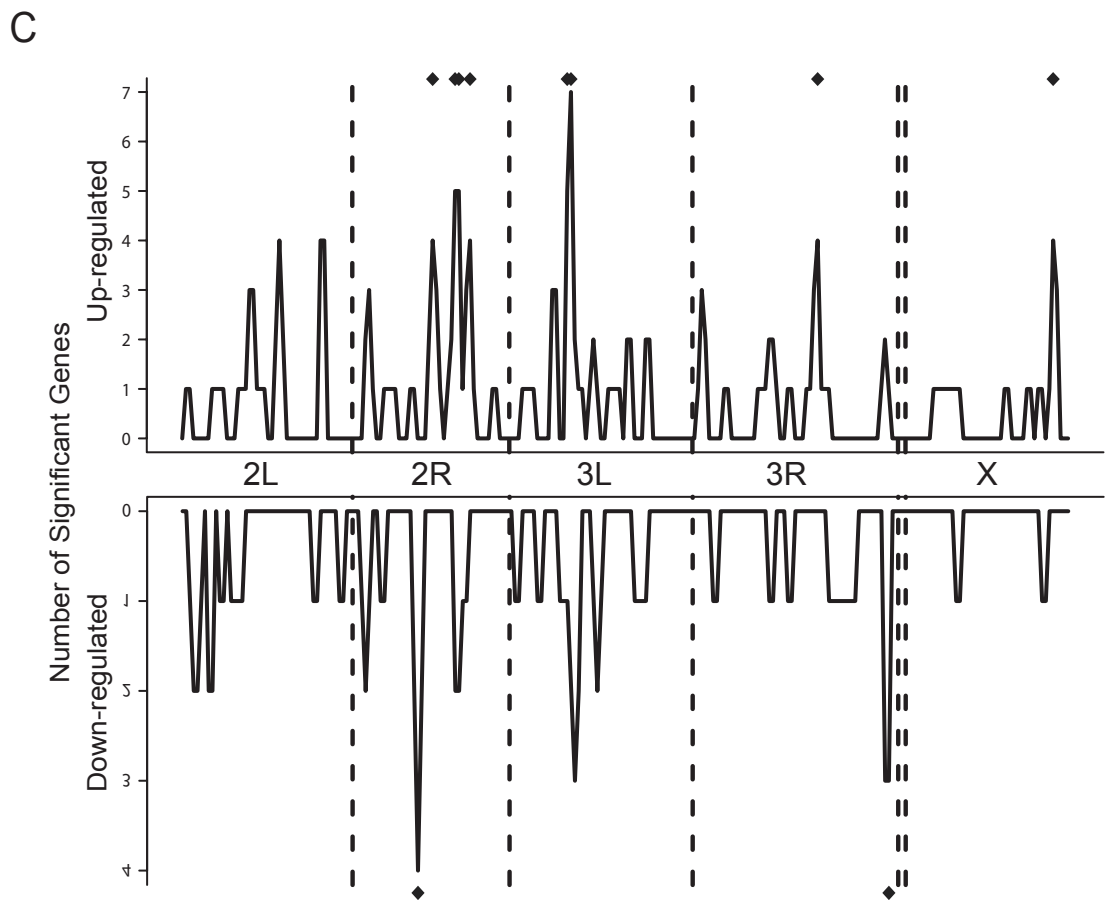
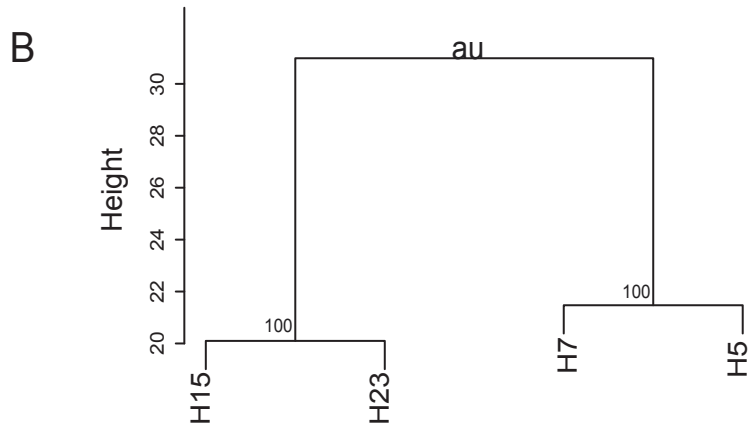
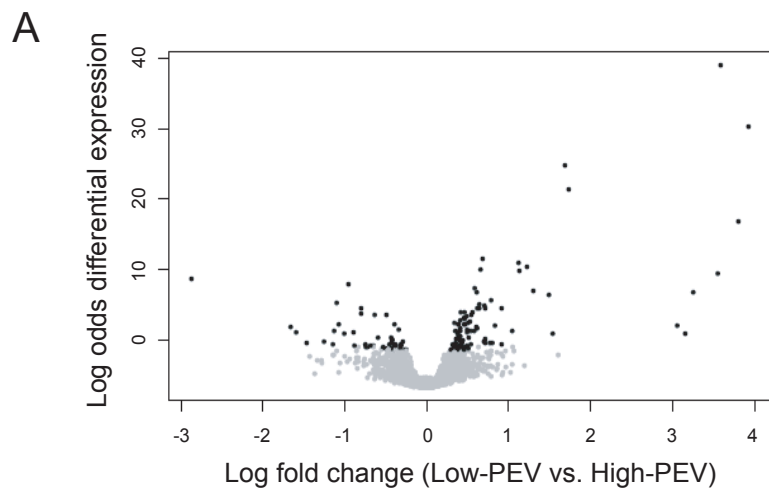


Figure 3

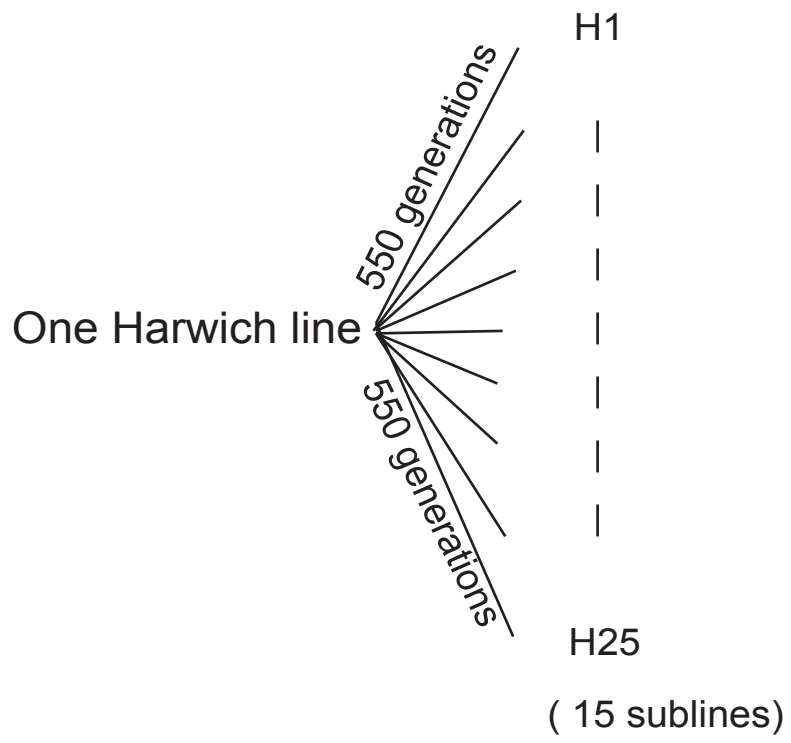


Figure S1

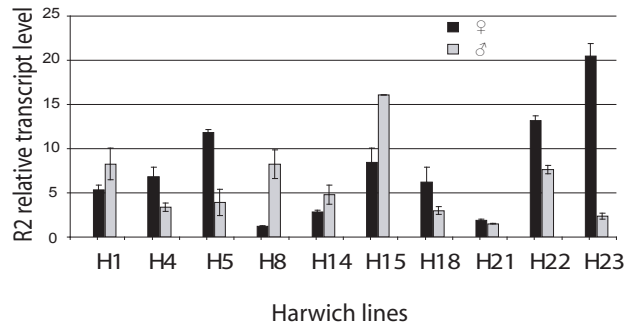
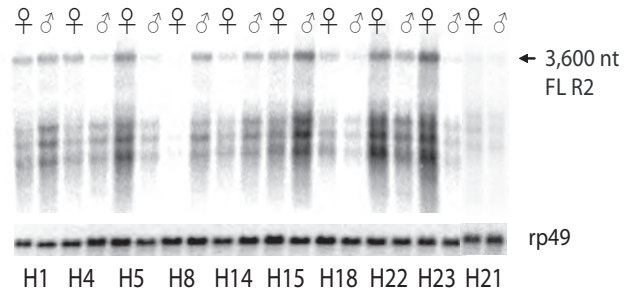


Figure S2

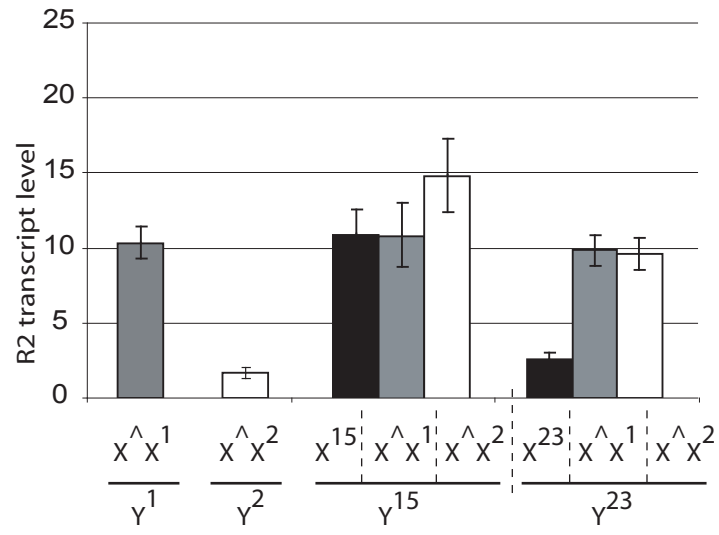


Figure S3

Harwich males

4361 females

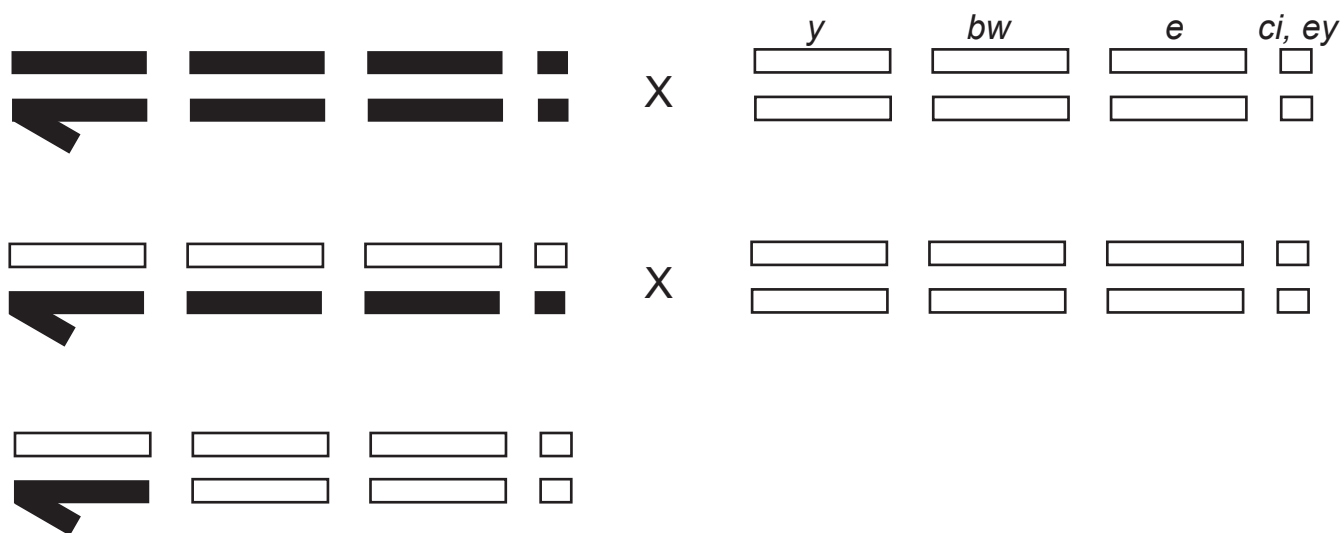
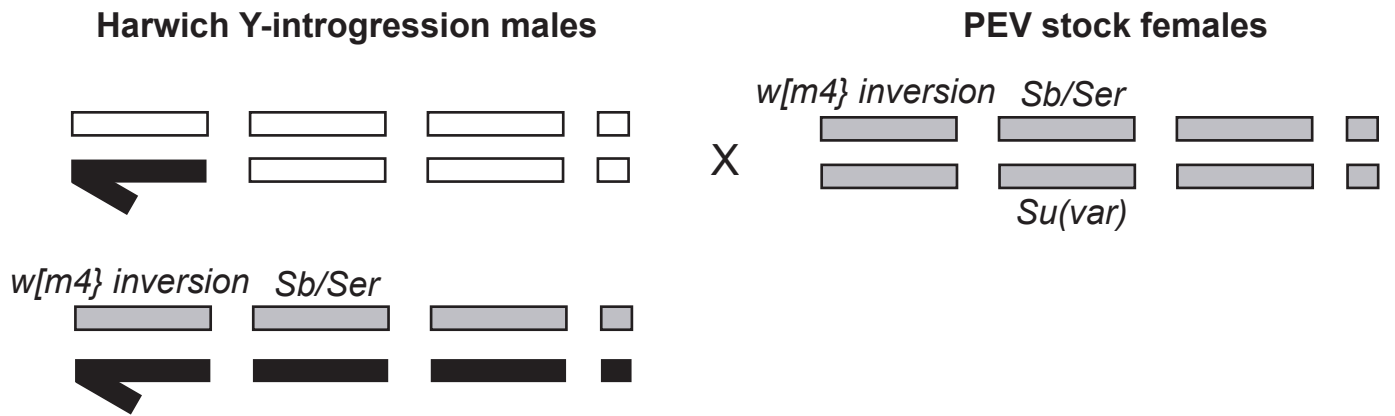


Figure S4

A



B

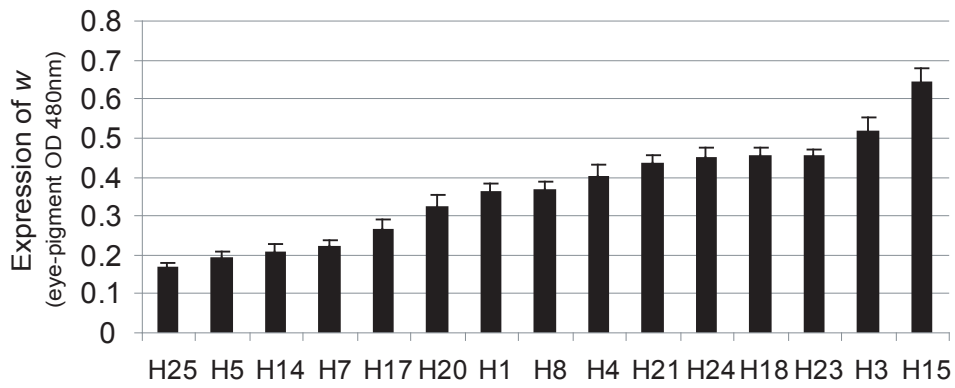


Figure S5

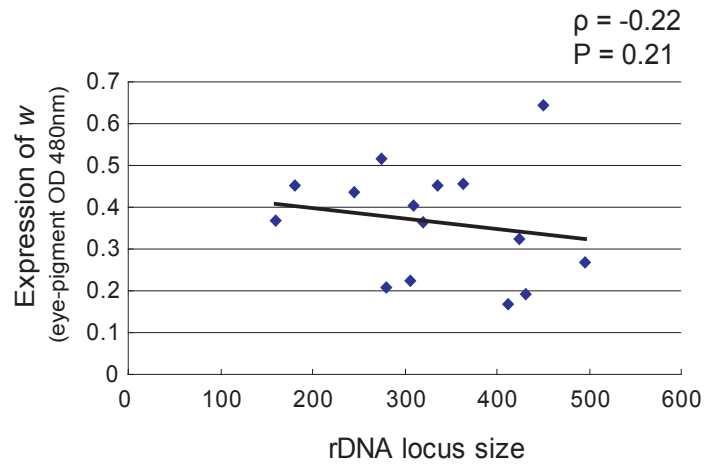


Figure S6

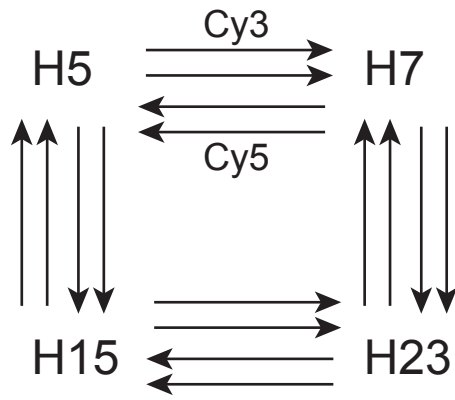


Figure S7

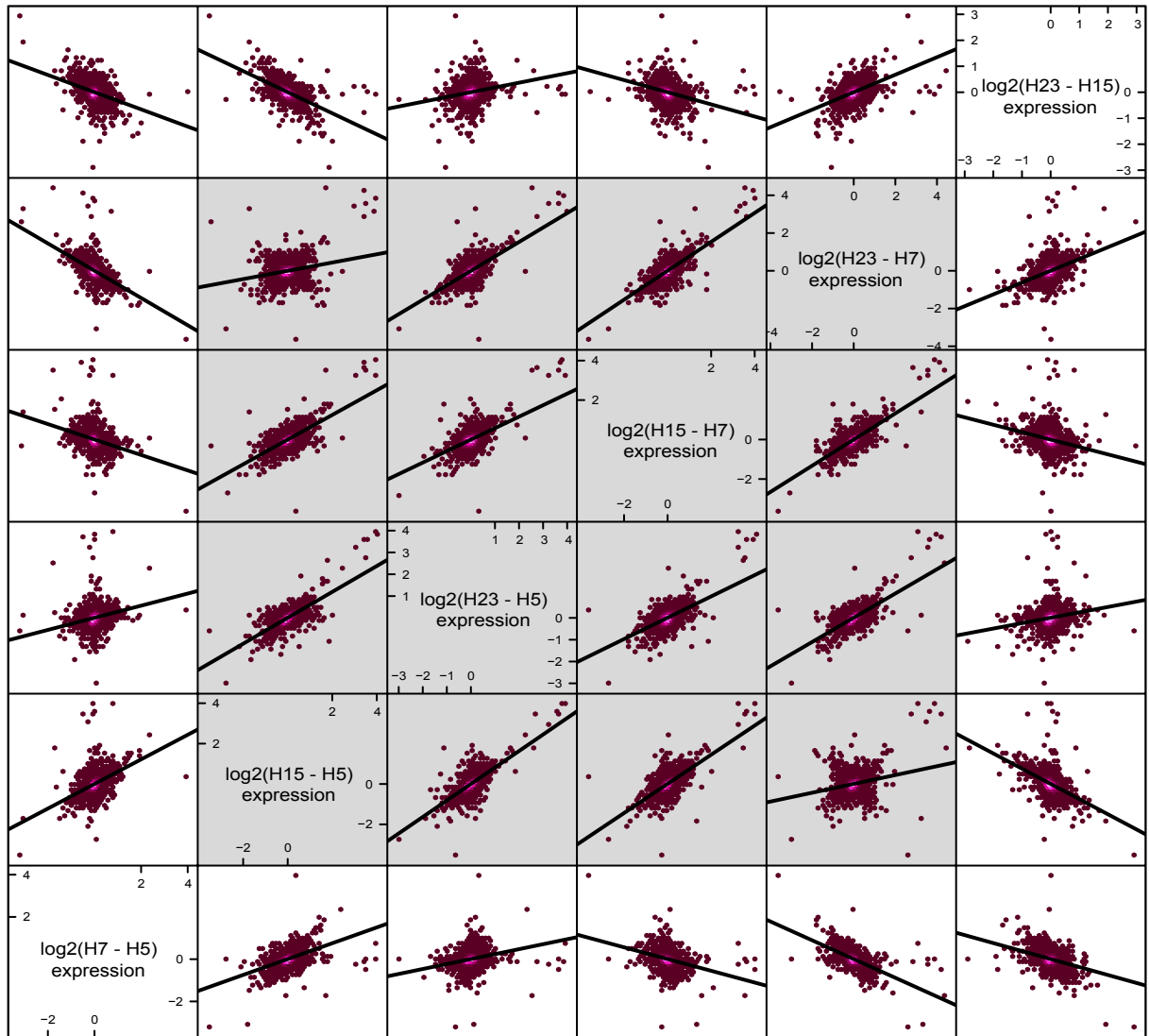


Figure S8