



**Polymorphic Y Chromosomes Harbor Cryptic
Variation with Manifold Functional Consequences**
Bernardo Lemos, *et al.*
Science **319**, 91 (2008);
DOI: 10.1126/science.1148861

**The following resources related to this article are available online at
www.sciencemag.org (this information is current as of June 18, 2008):**

Updated information and services, including high-resolution figures, can be found in the online version of this article at:

<http://www.sciencemag.org/cgi/content/full/319/5859/91>

Supporting Online Material can be found at:

<http://www.sciencemag.org/cgi/content/full/319/5859/91/DC1>

A list of selected additional articles on the Science Web sites **related to this article** can be found at:

<http://www.sciencemag.org/cgi/content/full/319/5859/91#related-content>

This article **cites 27 articles**, 15 of which can be accessed for free:

<http://www.sciencemag.org/cgi/content/full/319/5859/91#otherarticles>

This article has been **cited by 2** article(s) on the ISI Web of Science.

This article appears in the following **subject collections**:

Evolution

<http://www.sciencemag.org/cgi/collection/evolution>

Information about obtaining **reprints** of this article or about obtaining **permission to reproduce this article** in whole or in part can be found at:

<http://www.sciencemag.org/about/permissions.dtl>

Polymorphic Y Chromosomes Harbor Cryptic Variation with Manifold Functional Consequences

Bernardo Lemos,* Luciana O. Araripe, Daniel L. Hartl

The paucity of polymorphisms in single-copy genes on the Y chromosome of *Drosophila* contrasts with data indicating that this chromosome has polymorphic phenotypic effects on sex ratio, temperature sensitivity, behavior, and fitness. We show that the Y chromosome of *D. melanogaster* harbors substantial genetic diversity in the form of polymorphisms for genetic elements that differentially affect the expression of hundreds of X-linked and autosomal genes. The affected genes are more highly expressed in males, more meagerly expressed in females, and more highly divergent between species. Functionally, they affect microtubule stability, lipid and mitochondrial metabolism, and the thermal sensitivity of spermatogenesis. Our findings provide a mechanism for adaptive phenotypic variation associated with the Y chromosome.

Y chromosomes experience strong selective pressures for increased male fitness due to sexual selection, and they accumulate deleterious mutations owing to the lack of recombination (1, 2). In most species of *Drosophila*, the Y chromosome has few protein-coding loci but shows a high density of transposable elements and other repetitive sequences (3, 4). Only about 10 to 20 single-copy protein-coding loci are known to reside in the *D. melanogaster* Y chromosome, all of which are also exceptional in their megabase-sized introns embedded within heterochromatin of largely unknown sequence (3, 5). This gene content represents less than 0.2% of the *Drosophila* genome [~13,000 coding genes (6)]. The small number of genes is particularly noteworthy in that the Y chromosome comprises some 40 Mb of DNA, or 23% of the haploid genome. The Y chromosome is about the same size as the X chromosome, whereas the latter contains >2500 genes. Extremely low levels of DNA sequence polymorphism are observed for single-copy Y-linked genes (7), and population genetics theory implies that the Y chromosome can maintain stable polymorphisms only under restrictive conditions (1, 2, 8). These observations stand in contrast to manifold phenotypic consequences of Y-linked polymorphisms, including effects on sex ratio (9, 10), heat adaptation (11), behavior (12, 13), and male fitness (14).

To address this apparent contradiction, we examined genome-wide variation in gene expression among lines of *D. melanogaster* differing only in the origin of the Y chromosome (15) (figs. S1 to S3). Extraneous sources of variation were minimized by use of a common isogenic genetic background of X chromo-

somes, autosomes, mitochondrial DNA, and cytoplasm, as well as by careful control of environmental variables including temperature, humidity, and light. The five Y chromosomes tested include isolates from diverse geographical origins representing both temperate and tropical climates. The Y chromosomes were originally derived from flies collected in Massachusetts (Ymass) and Ohio (Ycs) in the United States as well as the Democratic Republic of Congo (Ycongo) and Zimbabwe (Yz53) in Africa. The effects of these chromosomes were compared among themselves as well as with those of the Y chromosome (Y4361) in a genetically marked laboratory strain into which the diverse Y chromosomes had been introduced.

A substantial number of genes were differentially expressed between the original Y4361 genotype and the Y-chromosome substitution lines [123 to 1010 genes; Bayesian posterior probability > 0.999 to 0.995; false discovery rate (FDR) < 1 to 15%; Fig. 1A]. This result contrasts with control comparisons between females of the Y-chromosome substitution lines, which showed no differential expression (15). Approximately 14%, 37%, 47%, and <1% of the Y-linked regulatory variation can be assigned to genes in the X, second, third, and fourth chromosomes, respectively. Moreover, four Y-linked genes [*Su(Ste)*, *Kl-5*, *Ccy*, and *Occludin-related Y*] themselves showed evidence for differential expression across the Y-chromosome substitution lines (Fig. 1, B to E), a significant enrichment over random expectations ($P < 0.05$; Fisher's exact test). These results indicate that polymorphic variation in the Y chromosome has manifold consequences in the regulation of genes located on the X chromosome and autosomes. Furthermore, the genes affected by Y-linked regulatory elements showed significant patterns in regard to their sex bias in expression, levels of polymorphism and divergence, and association with temperature.

Polymorphic regulatory variation due to the Y chromosome might be expected to affect traits expressed in males. In agreement with this prediction, we found that Y-linked variation is enriched for genes that are more highly expressed in males ($P < 0.0001$, Kruskal-Wallis, and other similar tests below; Fig. 2A). Conversely, genes affected by Y-linked regulatory variation are expressed at significantly lower abundances in females ($P < 0.0001$; Fig. 2B). Such a pattern might be expected of sexually antagonistic variation that is beneficial in one sex while detrimental in the other (16), and the pattern suggests a role for Y-linked polymorphisms in mediating the expression of these genes. Moreover, Y-linked regulatory polymorphisms may help to explain some of the diversity in gene expression patterns observed within and between species of *Drosophila* (17, 18). We found that genes whose expression is affected by the Y chromosome show significantly greater divergence in gene expression between *D. melanogaster* and *D. simulans* ($P < 0.0001$; Fig. 2C), as well as higher levels of gene expression polymorphism relative to other genes ($P < 0.0001$; Fig. 2D). These patterns indicate that Y-linked regulation of gene expression may result in dynamic evolutionary histories within and between species of fruit flies.

Evidence for adaptive polymorphism in the Y chromosome comes from its contribution to variation in the sensitivity of *Drosophila* spermatogenesis with respect to temperature (11, 19). Thermal heat sterility thresholds differ across species, from 23°C in heat-sensitive species up to 31°C in heat-tolerant species. In *D. melanogaster*, genetic variation in thermal tolerance is observed among geographical populations. Flies from tropical populations are more tolerant to heat-induced sterility and recover fertility more rapidly than those from temperate populations. Genetic analysis reveals that the Y chromosome alone accounts for about 50% of the difference in heat sensitivity observed between natural populations of *D. melanogaster* (11, 19). We hypothesize that variation in thermal sensitivity may be, at least in part, due to regulatory variation exerted by the Y chromosome. Indeed, we found that genes affected by Y-linked regulatory variation are more responsive to heat shock than are other genes ($P < 0.0001$; Fig. 3A). To investigate this connection in greater detail, we reared flies containing a Y chromosome from a temperate climate (Ymass) with those containing a Y chromosome from a tropical climate (Ycongo) at two temperatures (16°C and 25°C) and assayed their levels of gene expression at both temperatures (fig. S4). We found 14 genes to be up-regulated and 16 genes down-regulated in Ymass relative to Ycongo at both temperatures, whereas fewer than one such gene is expected by chance alone. The genes identified to be up-regulated

Department of Organismic and Evolutionary Biology, Harvard University, 16 Divinity Avenue, Cambridge, MA 02138, USA.

*To whom correspondence should be addressed. E-mail: blemos@oeb.harvard.edu

in Ymass include *SOX100B*, a gene known to be associated with male gonadal development (20), and *OS-C*, a gene presumably involved in olfactory coding (21). The genes identified to be down-regulated in Ymass include *CG7311* and *CG9389*, both of whose products are targeted to the mitochondria and are implicated in the metabolism of lipids. These findings point to consistent differences in gene expression among the isogenic lines that correlate with the Y chromosome independent of temperature.

We also investigated whether Y chromosomes may harbor variation that is apparent at only one temperature. This hypothesis predicts that the Ycongo line (with the tropical Y chromosome) would show, at 16°C, a more disrupted pattern of gene expression than the Ymass line (with the temperate Y chromosome). The opposite pattern would be expected at 25°C. We found that the Ycongo line showed about twice as many genes up-regulated at 16°C relative to Ymass, whereas the Ymass line showed about twice as many genes up-regulated at 25°C relative to Ycongo ($P < 10^{-7}$, Fisher's exact test; Fig. 3B). These findings point to temperature-dependent differences between Y-chromosome substitution lines, including cases in which the sign of the difference between strains was reversed at the two temperatures (Fig. 3, C and D). Taken together, these results suggest that genes responding to Y-linked variation are functionally coherent in showing association with male functions including spermatogenesis and temperature adaptation. These processes have also been linked to lipid metabolism in *Drosophila*, and indeed we found Y-linked variation for genes associated with lipid and fatty acid metabolism as well as with the mitochondria and the cytoskeleton (tables S1 and S2). Our findings indicate that Y-linked regulatory variation is expressed through dynamic interactions with temperature, and they provide a mechanism for understanding the effects of the Y chromosome on the heat sensitivity of spermatogenesis as well as other evolutionary and ecologically relevant phenotypes.

Y chromosomes are known to harbor structural polymorphism in heterochromatic sequences and copy-number polymorphisms in repetitive sequences both in humans and flies (22–24), and they can influence transcription epigenetically, as revealed by their ability to modify position-effect variegation in flies (25–27). We therefore suggest that structural polymorphisms in the Y chromosome may result in a variety of epigenetic effects, including the differential regulation of autosomal or X-linked transposable elements through such mechanisms as the homology-dependent RNA interference pathways (28). In agreement with this prediction, we found that Y-linked polymorphisms result in differential expression of 13 out of 53 families of transposable elements

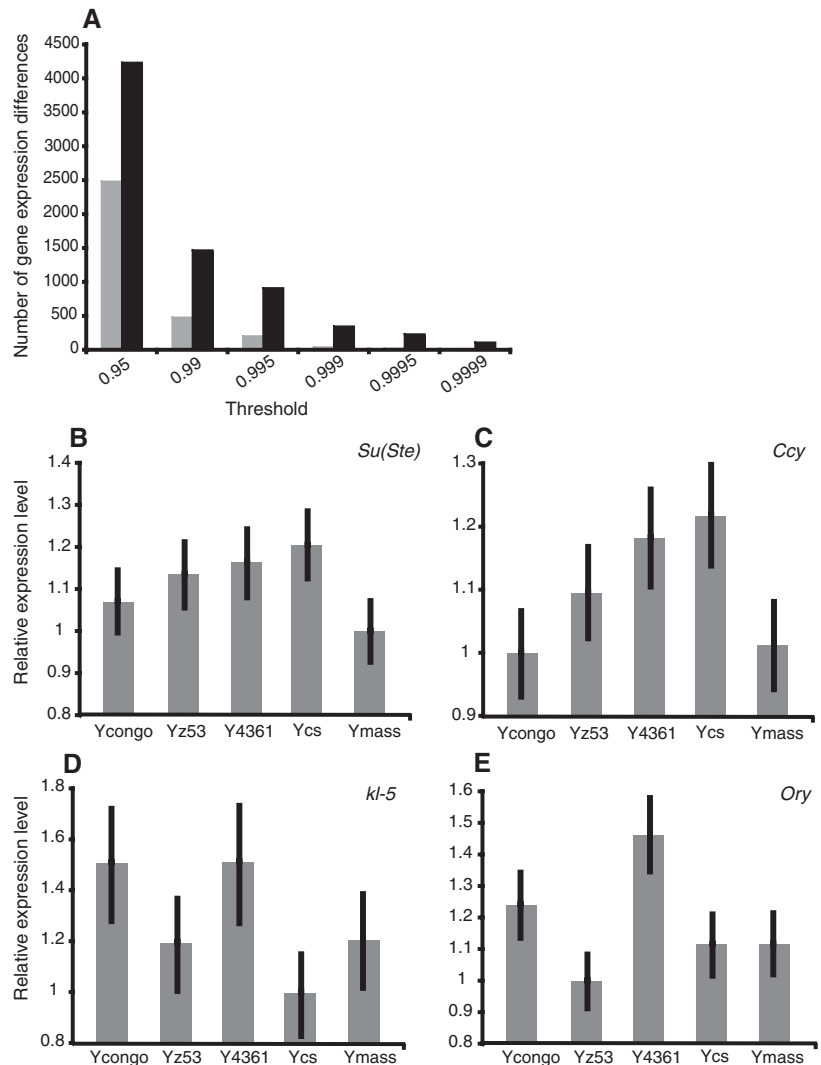
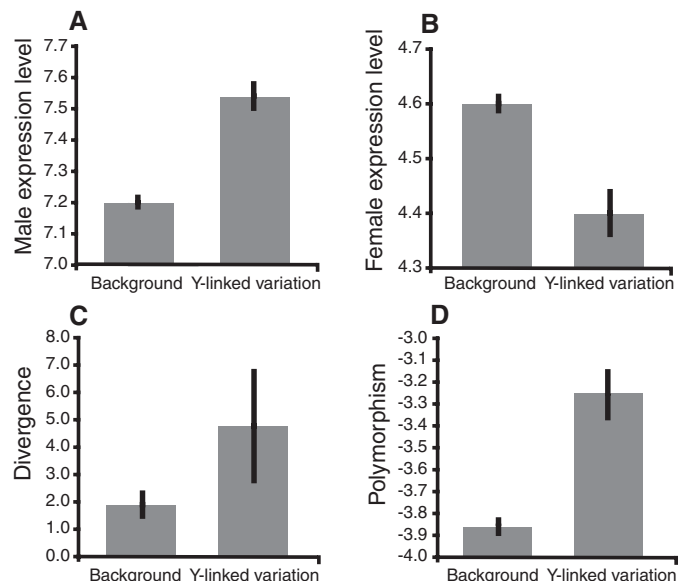


Fig. 1. (A) Number of genes differentially expressed among Y-chromosome substitution lines (black bars) as a function of the Bayesian posterior probability of differential expression. Gray bars indicate the estimated number of genes expected by chance. (B to E) Four Y-linked genes [*Su(Ste)*, *kl-5*, *Ccy*, and *Occludin-related Y (Ory)*] with significant expression variation among Y-chromosome substitution lines. Estimated relative differences in expression are shown. Black bars denote 95% credible interval for estimated relative difference.

Fig. 2. Genes showing Y-linked regulatory variation are expressed at higher abundances in males (A) and lower abundances in females (B); they show higher levels of divergence between *D. melanogaster* and *D. simulans* (C) and higher levels of polymorphism within *D. melanogaster* (D). Error bars denote two SEs around the mean. Background refers to the subset of genes not differentially expressed across Y-chromosome substitution lines.



represented in our data, a highly significant enrichment over random expectations ($P = 0.0003$; Fig. 4, A to D).

Our data reveal substantial levels of Y-chromosome polymorphism with consequences for gene regulation. This diversity is manifested functionally as significant differences in the ex-

pression of autosomal and X-linked genes from one Y chromosome to the next. Polymorphic Y-linked regulatory variation appears to be manifold, evolutionarily dynamic, preferential by temperature. One model for the mechanism of these effects is that polymorphisms in the con-

text or lengths of heterochromatic blocks harboring transposable elements and other repetitive sequences might alter the availability of limiting transcription factors or chromatin regulators throughout the genome (29). Our results also raise the question of how widespread Y-linked regulatory variation may be in other organisms with similarly heterochromatin-rich Y chromosomes, including humans and other mammals.

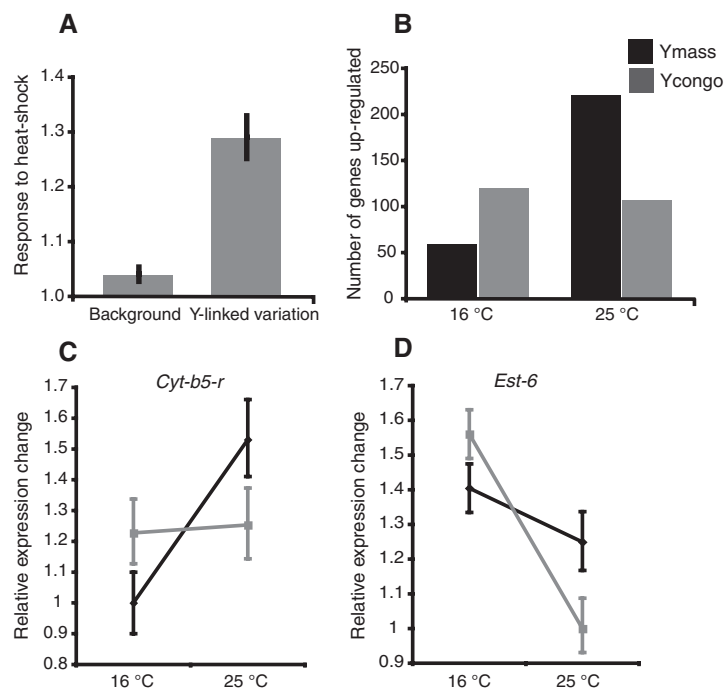


Fig. 3. (A) Genes showing Y-linked regulatory variation are more sensitive to heat shock. Gray bars denote two SEs around the mean. (B) Number of genes up-regulated in Ymass and Ycongo at 16°C and 25°C. (C and D) Two genes for which the sign of the difference between Ymass and Ycongo is reversed at the two temperatures assayed.

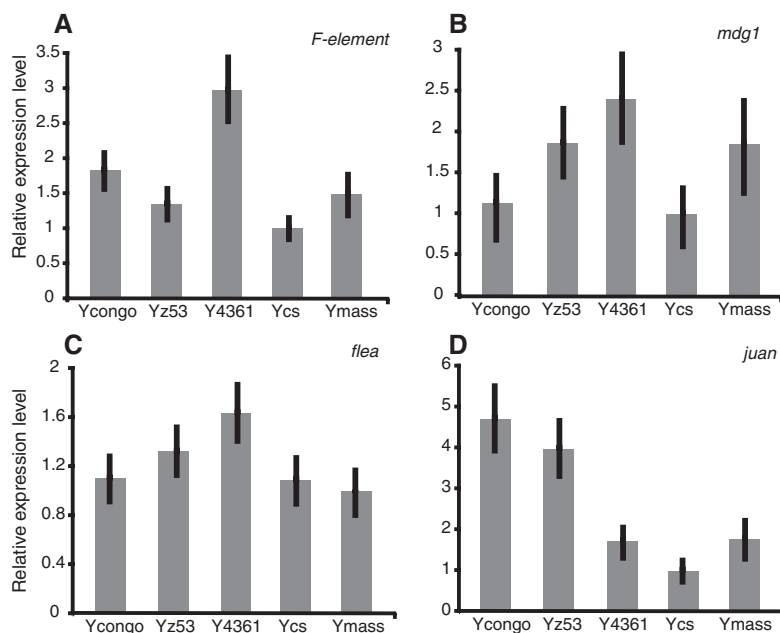


Fig. 4. Four transposable elements [(A) *F-element*; (B) *mdg1*; (C) *flea*; (D) *juan*] with significant expression variation across Y-chromosome substitution lines. Estimated relative differences in expression are shown. Black bars denote 95% credible interval for estimated relative difference.

References and Notes

- D. Charlesworth, B. Charlesworth, *Curr. Biol.* **15**, R129 (2005).
- D. L. Hartl, A. G. Clark, *Principles of Population Genetics* (Sinauer Associates, Sunderland, MA, ed. 4, 2007).
- A. B. Carvalho, A. G. Clark, *Science* **307**, 108 (2005); published online 4 November 2004 (10.1126/science.1101675).
- C. D. Smith, S. Shu, C. J. Mungall, G. H. Karpen, *Science* **316**, 1586 (2007).
- A. B. Carvalho, *Curr. Opin. Genet. Dev.* **12**, 664 (2002).
- M. D. Adams *et al.*, *Science* **287**, 2185 (2000).
- M. Zurovcova, W. F. Eanes, *Genetics* **153**, 1709 (1999).
- A. G. Clark, *Genetics* **125**, 527 (1990).
- A. B. Carvalho, S. C. Vaz, L. B. Klaczko, *Genetics* **146**, 891 (1997).
- C. Montchamp-Moreau, V. Ginhoux, A. Atlan, *Evol. Int. J. Org. Evol.* **55**, 728 (2001).
- C. Rohmer, J. R. David, B. Moreteau, D. Joly, *J. Exp. Biol.* **207**, 2735 (2004).
- S. Huttunen, J. Aspi, *Behav. Genet.* **33**, 17 (2003).
- S. F. Stoltenberg, J. Hirsch, *Anim. Behav.* **53**, 853 (1997).
- A. K. Chippindale, W. R. Rice, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 5677 (2001).
- See supporting material on Science Online.
- W. R. Rice, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 6217 (1998).
- C. D. Meiklejohn, J. Parsch, J. M. Ranz, D. L. Hartl, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 9894 (2003).
- J. M. Ranz, C. I. Castillo-Davis, C. D. Meiklejohn, D. L. Hartl, *Science* **300**, 1742 (2003).
- J. R. David *et al.*, *J. Evol. Biol.* **18**, 838 (2005).
- T. J. DeFalco *et al.*, *Dev. Cell* **5**, 205 (2003).
- M. P. McKenna, D. S. Hekmat-Scafe, P. Gaines, J. R. Carlson, *J. Biol. Chem.* **269**, 16340 (1994).
- E. M. S. Lyckegaard, A. G. Clark, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 1944 (1989).
- E. M. S. Lyckegaard, A. G. Clark, *Mol. Biol. Evol.* **8**, 458 (1991).
- S. Repping *et al.*, *Nat. Genet.* **35**, 247 (2003).
- M. Gatti, S. Pimpinelli, *Annu. Rev. Genet.* **26**, 239 (1992).
- S. Henikoff, P. B. Talbert, *Nat. Rev. Genet.* **7**, 793 (2006).
- S. R. Schulze, L. L. Wallrath, *Annu. Rev. Entomol.* **52**, 171 (2007).
- R. K. Slotkin, R. Martienssen, *Nat. Rev. Genet.* **8**, 272 (2007).
- X. Liu, B. Wu, J. Szary, E. M. Kofoed, F. Schaufele, *J. Biol. Chem.* **282**, 20868 (2007).
- We thank E. Dopman, P. Fontanillas, Y. Go, S. Kingan, R. Kulathinal, C. Meiklejohn, H. Montenegro, S. Renn, and S. Rifkin for their help during assembly, printing, and annotation of the microarrays as well as helpful discussions. Supported by NIH grants GM068465 and GM065169 (D.L.H.). Raw microarray data have been deposited in the Gene Expression Omnibus database (accession number GSE9457).

Supporting Online Material

www.sciencemag.org/cgi/content/full/319/5859/91/DC1
Materials and Methods
Figs. S1 to S4
Tables S1 and S2
References

7 August 2007; accepted 20 November 2007
10.1126/science.1148861