23. Macior, L. W. The pollination of vernal angiosperms. *Oikos* **30,** 452–460 (1978).
24. Lloyd, D. G. & Schoen, D. J. Self- and cross-fertilization in plants. I. Functional dimensions. *Int. J. Plant Sci.* **153,** 358–369 (1992).
25. Griffin, S. R., Mavraganis, K. & Eckert, C. G. Experimental analysis of protogyny in *Aquilegia canadensis*. *Am. J. Bot.* **87,** 1246–1256 (2000).
26. Morgan, M. T., Schoen, D. J. & Bataillon, T. M. The evolution of self-fertilization in perennials. *Am. Nat.* **150,** 618–638 (1997).
27. Morgan, M. T., and Schoen, D. J. The role of theory in an emerging new plant reproductive biology. *Trends Ecol. Evol.* **12,** 231–234 (1997).
28. Ritland, K. A series of FORTRAN computer programs for estimating plant mating systems. *J. Hered.* **81,** 235–237 (1990).
29. Schoen, D. J. & Lloyd, D. G. Self- and cross-fertilization in plants. III. Methods for studying modes and functional aspects of self-fertilization. *Int. J. Plant Sci.* **153,** 381–393 (1992).
30. Ritland, K. Inferences about inbreeding depression based on changes of the inbreeding coefficient. *Evolution* **44,** 1230–1241 (1990).

**Acknowledgements**

We thank M. Bhardwaj, S. Griffin and A. Kliber for help with the field work; C. Muis and M. Bhardwaj for help in the laboratory; and S. Barrett, K. Holsinger, A. Kliber, B. Montgomerie and M. Morgan for comments on the manuscript. This work was supported in the field by the Queen's University Biological Station and by a research grant from the Natural Sciences and Engineering Research Council of Canada to C.G.E.

Correspondence and requests for materials should be addressed to C.G.E. (e-mail: eckertc@biology.queensu.ca).

# Reduced adaptation of a non-recombining neo-Y chromosome

**Doris Bachtrog & Brian Charlesworth**

*Institute of Cell, Animal and Population Biology, University of Edinburgh, West Mains Road, Edinburgh EH9 3JT, UK*

**Sex chromosomes are generally believed to have descended from a pair of homologous autosomes. Suppression of recombination between the ancestral sex chromosomes led to the genetic degeneration of the Y chromosome[1]. In response, the X chromosome may become dosage-compensated[1,2]. Most proposed mechanisms for the degeneration of Y chromosomes involve the rapid fixation of deleterious mutations on the Y[1]. Alternatively, Y-chromosome degeneration might be a response to a slower rate of adaptive evolution, caused by its lack of recombination[3]. Here we report patterns of DNA polymorphism and divergence at four genes located on the neo-sex chromosomes of *Drosophila miranda*. We show that a higher rate of protein sequence evolution of the neo-X-linked copy of *Cyclin B* relative to the neo-Y copy is driven by positive selection, which is consistent with the adaptive hypothesis for the evolution of the Y chromosome[3]. In contrast, the neo-Y-linked copies of *even-skipped* and *roundabout* show an elevated rate of protein evolution relative to their neo-X homologues, probably reflecting the reduced effectiveness of selection against deleterious mutations in a non-recombining genome[1]. Our results provide evidence for the importance of sexual recombination for increasing and maintaining the level of adaptation of a population.**

Well-studied sex chromosome systems, such as those of humans or *Drosophila melanogaster*, are very ancient and show few signs of their evolutionary origins[4,5]. In *D. miranda*, a neo-sex chromosome system has resulted from a recent chromosome fusion between the Y chromosome and an autosome (Muller's element C (ref. 6); Fig. 1). The fused autosome, the neo-Y chromosome, is transmitted patrilineally. Its homologue segregates with the X chromosome, forming the neo-X chromosome. The closest relatives of *D. miranda*, *D. pseudoobscura* and *D. persimilis* (from which it diverged about 2 million years ago[7]), lack the fusion, setting an upper limit to the age of the neo-sex chromosomes. Because male *Drosophila* have achiasmate meiosis[8], the neo-Y never recombines, and so is subject to the same evolutionary forces responsible for the degeneration of 'true' Y chromosomes. Indeed, the neo-sex chromosomes of *D. miranda* show many characteristics of the much older true sex chromosomes. The neo-Y has degenerated substantially[9], whereas the neo-X chromosome has evolved partial dosage compensation[2]. There is compelling evidence that much of the human Y chromosome is derived from a single autosomal region added to the original Y chromosome[10]; that is, it is also a degenerate neo-Y chromosome.

The different evolutionary processes that might be involved in Y-chromosome degeneration leave different footprints of variation and evolution at the molecular level[1], and so can be examined empirically. Here we report levels of polymorphism and divergence at four genes, *Cyclin B* (*CycB*), *roundabout* (*robo*), *engrailed* (*eng*) and *even-skipped* (*eve*), located on both the neo-X and neo-Y chromosomes of *D. miranda*. The use of degenerate primers for PCR, and screening of a genomic library, allowed the isolation of these genes from *D. miranda* (see Methods). *In situ* hybridization of probes to the polytene chromosomes confirmed their expected position on the neo-sex chromosomes of *D. miranda* (their positions on the polytene map[11] are: *CycB*, 32A; *robo*, 33C; *eng*, 31D; *eve*, 23A). For *CycB*, *robo* and *eve*, both coding and non-coding sequence data were obtained; for *eng*, only the intron sequence was investigated. The use of reverse-transcriptase-mediated polymerase chain reaction (RT–PCR) with primers spanning an intron confirmed that, for *CycB*, *eve* and *robo*, both alleles are transcribed and spliced in male flies, indicating that both the neo-X and the neo-Y copies are probably under selective constraints. For outgroup comparisons, all genes were isolated by PCR from *D. pseudoobscura*, in which Muller's element C is autosomal (Fig. 1).

We investigated nucleotide variation in 12 wild-derived lines of *D. miranda* (Table 1). A total of 5.2 kilobases per individual were surveyed for the homologous regions on the neo-X and neo-Y chromosomes. Between the neo-X-linked and the neo-Y-linked genes investigated, there are 106 fixed differences and no shared polymorphisms, which is consistent with a lack of recombination between the neo-sex chromosomes. Net silent-site divergence[12] between the neo-X and neo-Y alleles is 2.8%; combining this with previous data[13], and assuming a silent substitution rate of $1.2 \times 10^{-8}$ per site per year for *Drosophila*[14], this yields a total divergence time of 2.2 million years, giving an age of about 1.1 million years for the neo-sex chromosome system. On the neo-X chromosome, 61

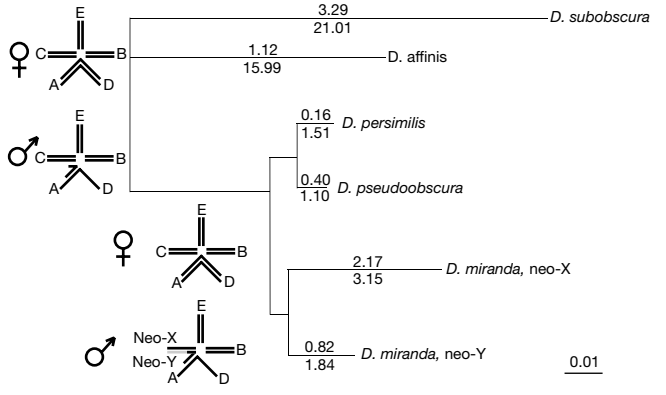


**Figure 1** Phylogenetic relationships between the species investigated, constructed from the coding region of *CycB*. The numbers shown above the lines give the estimated $K_a$ values (percentages) for each lineage obtained by maximum likelihood[19]; the corresponding $K_s$ values are given below the lines. The karyotypes of the species are also illustrated. The letters A–E indicate the five major chromosomes of the basic *Drosophila* karyotype[6]. In *D. miranda*, element C is fused to the true Y chromosome (neo-Y, shown in grey), and its unfused homologue, the neo-X, segregates with the X chromosome. The relatives of *D. miranda* lack the Y–C fusion. (Note that *D. subobscura* lacks the A–D fusion found in the other species[6]).

**Table 1 Patterns of polymorphism and divergence from *D. pseudoobscura* for the neo-sex-linked genes of *D. miranda***

| | *CycB* | | *robo* | | *eve* | | *eng* | |
|---|---|---|---|---|---|---|---|---|
| | Neo-X | Neo-Y | Neo-X | Neo-Y | Neo-X | Neo-Y | Neo-X | Neo-Y |
| Size (base pairs) | 1,899 | 1,827 | 1,766 | 1,766 | 1,228 | 1,212 | 387 | 363 |
| Segregating sites | 5 | 1 | 20 | 0 | 27 | 1 | 9 | 0 |
| $\theta_W$ per site (%)* | 0.212 | 0.046 | 0.705 | 0 | 1.160 | 0.045 | 0.808 | 0 |
| $\pi$ per site (%)* | 0.205 | 0.023 | 0.776 | 0 | 1.117 | 0.023 | 0.747 | 0 |
| Codons | 542 | 542 | 1,398 | 1,398 | 362 | 362 | 0 | 0 |
| $K_s$ (%) | 6.26 | 5.73 | 3.16 | 4.33 | 4.19 | 7.05 | – | – |
| $K_a$ (%) | 2.51 | 1.10 | 0.25 | 0.93 | 0.00 | 0.24 | – | – |
| $K_a/K_s$ | 0.40 | 0.17 | 0.08 | 0.21 | 0.00 | 0.03 | – | – |

* $\theta_W$ and $\pi$ are the two commonly used measures of nucleotide diversity, based on the numbers of segregating sites and mean numbers of pairwise differences between alleles, respectively[28]; silent-site diversities per site are given.

segregating sites were observed (27 silent single-nucleotide polymorphisms, 4 replacements, 25 non-coding sites and 5 insertion or deletion events. In contrast, the homologous neo-Y-linked regions almost completely lacked variability: only two singleton variants (one synonymous, one non-coding) were detected. Under the standard neutral model, genetic diversity is proportional to the product of the effective population size, $N_e$, and the mutation rate[15]. If no selective forces are operating, $N_e$ for the neo-Y chromosome should be one-third that of the neo-X chromosome, assuming random variation in offspring number for each sex[16]. Our data show that the diversity level for the neo-Y chromosome is reduced 30-fold compared with that for the neo-X. This result is consistent with our previous finding of a ~25-fold reduction in microsatellite variability on the neo-Y chromosome of *D. miranda*[17]. The overall rate of silent substitution on the neo-Y branch of the tree connecting the neo-X and neo-Y of *D. miranda* and *D. pseudoobscura* (Fig. 1) is similar to that for the neo-X branch, showing that the difference in diversity is not caused by a lower mutation rate of the neo-Y-linked genes (Table 1). A likelihood method was used to quantify the reduction in diversity of the neo-Y chromosome, combining both the microsatellite and the sequence data. Coalescent simulations yield a maximum-likelihood estimate of the ratio of $N_e$ for the neo-Y chromosome to that for the neo-X of 0.04 (Fig. 2). This is consistent with the expectation of a substantial reduction in $N_e$ under various forms of selection acting on a non-recombining genome[1]. A similar magnitude of reduction in nucleotide polymorphism has also been observed on an evolving plant Y chromosome[18].

A large reduction in $N_e$ for the neo-Y chromosome greatly reduces the rate of fixation of beneficial mutations on the neo-Y chromosome[3], and increases the rate of fixation of deleterious mutations[1]. If loci on the neo-X chromosome continue to adapt and their homologues on the neo-Y fail to do so, it is advantageous to upregulate X-linked genes and downregulate their maladapted Y-linked homologues[3]. Similarly, an accumulation of deleterious alleles on the neo-Y favours increased activity of neo-X genes relative to their neo-Y homologues[1]. Both processes can therefore lead to the observed inactivation of neo-Y-linked genes[9] and to dosage compensation[2].

Comparisons with the outgroup species reveal that the neo-X-linked copy of *CycB* has a very high rate of protein evolution (Table 1 and Fig. 1). Twenty-three amino-acid replacement substitutions were observed on the branch of the phylogeny leading to the neo-X, but only nine occurred on the neo-Y branch. This difference is not the result of an increased mutation rate on the neo-X, as the proportion of synonymous substitutions per synonymous site, $K_s$, is very similar on the neo-X and neo-Y lineages ($K_s \approx 0.06$ for both branches (Table 1)). For phylogenetic analysis of the nucleotide substitutions, *CycB* was also amplified from *D. persimilis*, *D. affinis* and *D. subobscura*, which are all members of the *obscura* subgroup of *Drosophila*, to which *D. miranda* belongs[6]. Likelihood analysis[19] indicates that the ratio of the non-synonymous to synonymous substitution rate ($K_a/K_s$) differs significantly between branches ($2\Delta L = 31.9$, d.f. = 8, $P < 0.0001$ (Fig. 1)). A model of a single $K_a/K_s$ ratio for all branches was compared with a model of a specific ratio for the neo-X branch plus a second rate common to all other branches (two-rate model). The latter fitted the data significantly better ($2\Delta L = 18.4$, d.f. = 1, $P < 0.00002$), whereas a model with a different $K_a/K_s$ ratio for each branch did not significantly improve the likelihood over the two-rate model ($2\Delta L = 13.4$, d.f. = 7, $P \approx 0.06$). This suggests that most of the heterogeneity in the $K_a/K_s$ ratio between branches of the phylogeny reflects an elevated rate on the branch leading to the neo-X chromosome (Fig. 1).

Heterogeneity in the $K_a/K_s$ ratio between lineages might be caused either by positive selection or by relaxed selective constraints in some lineages[12]. If the elevated rate of amino-acid replacements on the neo-X were solely a consequence of reduced functional constraints, the ratio of synonymous to replacement substitutions would be similar for polymorphisms within *D. miranda* and for fixed differences between *D. miranda* and its relatives[20]. However, a significant excess of the ratio of replacement to silent substitutions relative to polymorphisms was found for comparisons involving the neo-X-linked copy of *CycB* (Table 2), suggesting that directional darwinian selection has driven the rapid evolution of *CycB* on the neo-X chromosome[20]. In addition, the recent fixation of an advantageous mutation reduces neutral polymorphism around the site that is the target of selection[21]. Consistent with the selection hypothesis is a decrease in variability at *CycB* (Table 1). The level of polymorphism is about 4-fold lower than estimates for the other neo-X-linked genes investigated (see Table 1). By the HKA test[22], the level of neo-X silent polymorphism relative to divergence from *D. pseudoobscura* at *CycB* is significantly lower than that for all three other genes (*eve*, *eng* and *robo* ($P < 0.01$)). This result is consistent with the idea that the neo-X chromosome experiences a faster rate of adaptive evolution than the non-recombining neo-Y.

In contrast, *eve* and *robo* both show a higher rate of amino-acid replacements on the neo-Y chromosome than their neo-X homologues (Table 1). Maximum-likelihood analysis of the combined

**Table 2 McDonald–Kreitman tests of neutrality at *CycB***

| | Fixed | | Polymorphic | | |
|---|---|---|---|---|---|
| Comparison | Synonymous | Replacement | Synonymous | Replacement | *P** |
| Neo-X/neo-Y | 14 | 30 | 5 | 0 | <0.01 |
| Neo-X/*D. pseudoobscura* | 20 | 27 | 4 | 0 | 0.04 |
| Neo-Y/*D. pseudoobscura* | 17 | 12 | 1 | 0 | 0.68 |

* The *P* values are from two-tailed Fisher's exact tests, comparing numbers of synonymous versus replacement changes in the fixed and polymorphic categories, respectively.

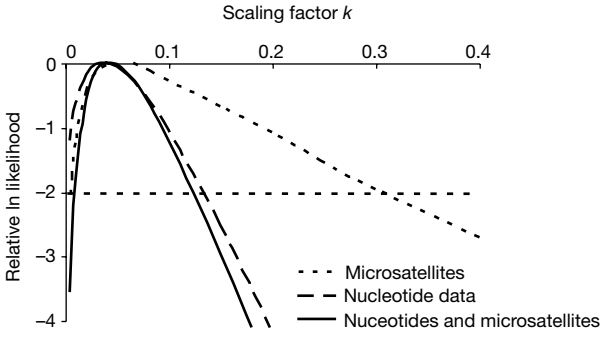


**Figure 2** Plots of the relative values of the log-likelihood functions ln $L(k|\Delta V)$, ln $L(k|\Delta S)$ and ln $L(k|\Delta V, \Delta S)$ as a function of $k$, the reduction in the effective population size of the neo-Y chromosome relative to the neo-X. $\Delta S$ and $\Delta V$ are the observed differences in numbers of segregating sites and variances of microsatellite repeat numbers between the neo-X and the neo-Y chromosome. All three curves have been normalized so that their maxima are at 0.The horizontal line indicates the two-unit support interval.

sequence data set for *eve* and *robo* was used to compare a model in which $K_a/K_s$ was assumed to be the same on the neo-X and neo-Y branches with a model in which it was allowed to differ; the latter fitted the data significantly better ($2\Delta L = 6.1$, d.f. = 1, $P = 0.01$). In addition, $K_a/K_s$ on the neo-Y branch of *CycB* was higher than on the other branches of the phylogeny (except for the neo-X branch; see Fig. 1). Several lines of evidence suggest that the acceleration of the rate of amino acid substitution on the neo-Y chromosome is not due to a lack of functional constraints on the coding region of *eve* and *robo*. Using RT–PCR, we demonstrated that the neo-Y-linked copies of both genes are transcribed. The entire coding sequences of *eve* and *robo* were analysed, and contained no stop codons or frame-shift mutations. A ratio of $K_a/K_s$ of 1 is expected for a gene that is evolving entirely neutrally, as was found for the *Lcp* genes on the neo-Y chromosome of *D. miranda*[13], whereas a lower ratio is expected for genes subject to functional constraints[12]. *Eve* and *robo* display a very low $K_a/K_s$ ratio on the neo-Y chromosome (Table 1), suggesting that both genes are under selective constraints. These results, together with the large reduction in $N_e$ for the neo-Y chromosome, suggest that many amino-acid substitutions in these genes are slightly deleterious and that natural selection has not been able to prevent their accumulation on the neo-Y branch[1]. A similar effect has been reported for some other non-recombining genomes[23,24].

The adaptive significance of sexual recombination is one of the most puzzling problems in evolutionary biology[25,26]. Most asexual lineages of eukaryotes seem to become extinct at a higher rate than their sexual relatives[25,27]. Similarly, the Y chromosome is prone to degeneration once it stops recombining with the X[1]. The dismal fates of clonally transmitted species and genomes suggest that genetic recombination is necessary for their persistence over long periods of evolutionary time. Theoretical models predict that sexual populations should be more effective in incorporating new beneficial mutations and preventing the accumulation of deleterious alleles[25,26]. Our evidence for both faster adaptation on the recombining X chromosome and the accumulation of deleterious mutations on the non-recombining Y chromosome suggests that both processes may be important in conferring an evolutionary advantage to sex and recombination. □

## Methods

### Isolation of genes, sequencing and RT-PCR

*CycB* and *robo* were isolated from a genomic library constructed from *D. miranda* males (strain 0101.3 (ref. 13) using the Lambda FixII Library kit (Stratagene). *eve* and *eng* were isolated by using degenerate primers designed from regions conserved between *D. melanogaster* and *D. virilis*. Allele-specific primers were used for PCR amplification of the neo-X-linked and neo-Y-linked copies of *CycB* and *robo* from male genomic DNA, followed by direct sequencing of both strands of the PCR products. The neo-Y-specific primers only amplify a product in males, while the neo-X-specific primers amplify in both sexes. For *eng* and *eve*, PCR primers amplifying both copies were used, and the amplified product was cloned with the TOPO TA cloning kit (Invitrogen). Clones containing either the neo-X or neo-Y copy were distinguished by a size difference in the cloned PCR products in the case of *eng*, or by a *Hae*II restriction site for *eve*. To exclude PCR errors, at least three independent clones per allele and individual were analysed (the PCR error rate was found to be 0.00092 per base pair). For the population survey, 12 strains of *D. miranda* were sequenced (see ref. 13 for description of the strains), using the ABI Prism BigDye Chemistry (Perkin-Elmer) on an ABI 377 automated sequencer.

Total RNA from male *D. miranda* (strain MSH 38 (ref. 13) was extracted with the RNA STAT-60 kit, and first-strand cDNA synthesis was performed with random or dT primers and the cDNA Cycle kit (Invitrogen). Primers spanning at least one intron were used to PCR-amplify the neo-X and the neo-Y copies. Amplification products were cloned with the TOPO TA cloning vector and sequenced.

### Sequence analysis

Sequences of *D. miranda* and *D. pseudoobscura* were aligned manually. Nucleotide diversity was calculated as described[28]. To align the coding region of *CycB* between the more diverged members of the *obscura* species group, the translated proteins were aligned by using CLUSTAL X (ftp://ftp-igbmc.u-strasbg.fr/pub/ClustalX/) and the nucleotide alignment was constructed from the protein alignment. The phylogeny based on the coding region of *CycB* was derived by using PUZZLE (ftp://ftp.ebi.ac.uk/pub/software/mac/puzzle/) and is shown in Fig. 1. $K_a$ and $K_s$ were calculated with a maximum-likelihood method, which accounts for unequal transition and transversion rates and unequal base and codon frequencies[19]. A likelihood ratio test was used to test different models of evolution by application of the codeml program within the PAML software package[19]. A model assuming the same $K_a/K_s$ ratio for all branches was compared with a model assuming a different $K_a/K_s$ ratio for all or some branches of the phylogeny. Twice the difference in log-likelihood between models is assumed to be distributed approximately as $\chi^2$, with degrees of freedom equal to the difference in the numbers of parameters.

### Coalescent process simulations

Coalescent simulations with the standard algorithm of Hudson[29] were performed to obtain a maximum-likelihood estimate of the reduction in $N_e$ for the neo-Y chromosome compared with that of the neo-X, combining both the nucleotide polymorphism data set presented here and data on microsatellite variability[17]. Whereas all loci on the neo-Y share one genealogy, the loci on the neo-X are probably independent. We therefore generated 11 independent trees for a data set of 12 chromosomes for the neo-X (7 microsatellite loci and 4 genes), whereas a single tree with 11 completely linked loci was computed for the neo-Y. Microsatellite mutations were superimposed on the trees following the single stepwise mutation model described in ref. 17. In short, seven random estimates of $\theta = 4N_e\mu$ (where $\mu$ is the mutation rate per microsatellite locus) were drawn from a gamma distribution for each run, and mutations using the resulting $\theta$ values were laid down on the trees in accordance with Poisson distributions, using the same values of $\theta$ for homologous loci. The mean variance in repeat number per locus, $\bar{V}$, across the seven loci was then calculated. After each run, $\Delta V_{sim} = (\bar{V}_X - \bar{V}_Y)/\bar{V}_X$ for the simulated data set was computed. For each gene, the total number of segregating mutations (*CycB*, 6; *robo*, 20; *eve*, 28; *eng*, 9) were assigned to the simulated neo-X and neo-Y trees in proportion to their lengths. After each run, the total numbers of segregating sites on the neo-X ($S_X$) and on the neo-Y ($S_Y$) were determined and the quantity $\Delta S_{sim} = (S_X - S_Y)/S_X$ was computed. To estimate the reduction in $N_e$ for the neo-Y loci, the tree length (in units of $2N_e$ generations) was multiplied by a scaling factor $k$ before mutations were laid down on the neo-Y tree. The simulations can be used to determine the likelihood of a pair of ($\Delta V_{sim}$, $\Delta S_{sim}$) for a given value of $k$. $n$ replicas are generated for each value of $k$. The number of runs $M_\delta$ where $|\Delta V_{sim} - \Delta V_{obs}| \leq \delta$ and $|\Delta S_{sim} - \Delta S_{obs}| \leq \delta$ was determined, where $\delta$ is a preassigned mesh size for the continuous variables $\Delta V$ and $\Delta S$. Following ref. 30, the likelihood of the sample is approximated by

$$L(k|\Delta V, \Delta S) = \frac{1}{n} M_\delta(\Delta V, \Delta S)$$

Here, $n$ is $10^5$ and $\delta$ is 0.02.

Received 21 September; accepted 18 December 2001.

1. Charlesworth, B. & Charlesworth, D. The degeneration of Y chromosomes. *Phil. Trans. R. Soc. Lond. B* **55,** 1563–1572 (2000).
2. Marin, I., Siegal, M. L. & Baker, B. S. The evolution of dosage-compensation mechanisms. *BioEssays* **22,** 1106–1114 (2000).
3. Orr, H. A. & Kim, Y. An adaptive hypothesis for the evolution of the Y chromosome. *Genetics* **150,** 1693–1698 (1998).
4. Lahn, B. T., Pearson, N. M. & Jegalian, K. The human Y chromosome, in the light of evolution. *Nature Rev. Genet.* **2,** 207–216 (2001).
5. Carvalho, A. B., Lazzaro, B. P. & Clark, A. G. Y chromosomal fertility factors *kl-2* and *kl-3* of *Drosophila melanogaster* encode dynein heavy chain polypeptides. *Proc. Natl Acad. Sci. USA* **97,** 13239–13244 (2000).
6. Powell, J. R. *Progress and Prospects in Evolutionary Biology: The Drosophila Model* (Oxford Univ. Press, New York, 1997).
7. Schaeffer, S. W. & Miller, E. L. Molecular population genetics of an electrophoretically monomorphic protein in the alcohol dehydrogenase region of *Drosophila pseudoobscura*. *Genetics* **132,** 163–178 (1992).
8. Gethmann, R. C. Crossing over in males of higher Diptera (Brachycera). *J. Hered.* **79,** 344–350 (1988).
9. Steinemann, M. & Steinemann, S. Enigma of Y chromosome degeneration: neo-Y and neo-X

chromosomes of *Drosophila miranda* a model for sex chromosome evolution. *Genetica* **102–103,** 409–420 (1998).
10. Waters, P. D., Duffy, B., Frost, C. J., Delbridge, M. L. & Graves, J. A. The human Y chromosome derives largely from a single autosomal region added to the sex chromosomes 80–130 million years ago. *Cytogenet. Cell. Genet.* **92,** 74–79 (2001).
11. Das, M., Mutsuddi, D., Duttagupta, A. K. & Mukherjee, A. S. Segmental heterogeneity in replication and transcription of the $X_2$ chromosome of *Drosophila miranda* and conservativeness in the evolution of dosage compensation. *Chromosoma* **87,** 373–388 (1982).
12. Li, W. *Molecular Evolution* (Sinauer Associates, Sunderland, Massachusetts, 1997).
13. Yi, S. & Charlesworth, B. Contrasting patterns of molecular evolution of the genes on the new and old sex chromosomes of *Drosophila miranda*. *Mol. Biol. Evol.* **17,** 703–717 (2000).
14. Wang, R. L. & Hey, J. The speciation history of *Drosophila pseudoobscura* and close relatives: inferences from DNA sequence variation at the *period* locus. *Genetics* **144,** 1113–1126 (1996).
15. Kimura, M. *The Neutral Theory of Molecular Evolution* (Cambridge Univ. Press, Cambridge, 1983).
16. Wright, S. *Evolution and the Genetics of Populations* (Univ. of Chicago Press, Chicago, 1969).
17. Bachtrog, D. & Charlesworth, B. Reduced levels of microsatellite variability on the neo-Y chromosome of *Drosophila miranda*. *Curr. Biol.* **10,** 1025–1031 (2000).
18. Filatov, D. A., Moneger, F., Negrutiu, I. & Charlesworth, D. Low variability in a Y-linked plant gene and its implications for Y-chromosome evolution. *Nature* **404,** 388–390 (2000).
19. Yang, Z. PAML: a program package for phylogenetic analysis by maximum likelihood. *Comput. Appl. Biosci.* **13,** 555–556 (1997).
20. McDonald, J. H. & Kreitman, M. Adaptive protein evolution at the *Adh* locus in *Drosophila*. *Nature* **351,** 652–654 (1991).
21. Barton, N. H. Genetic hitchhiking. *Phil. Trans. R. Soc. Lond. B* **355,** 1553–1562 (2000).
22. Hudson, R. R., Kreitman, M. & Aguade, M. A test of neutral molecular evolution based on nucleotide data. *Genetics* **116,** 153–159 (1987).
23. Lynch, M. & Blanchard, J. L. Deleterious mutation accumulation in organelle genomes. *Genetica* **102–103,** 29–39 (1998).
24. Fridolfsson, A. K. & Ellegren, H. Molecular evolution of the avian *CHD1* genes on the Z and W sex chromosomes. *Genetics* **155,** 1903–1912 (2000).
25. Maynard Smith, J. *The Evolution of Sex* (Cambridge Univ. Press, Cambridge, 1978).
26. Barton, N. H. & Charlesworth, B. Why sex and recombination? *Science* **281,** 1986–1990 (1998).
27. Bell, G. *The Masterpiece of Nature* (Univ. of California, Berkeley, 1982).
28. Tajima, F. Statistical analysis of DNA polymorphism. *Jpn J. Genet.* **68,** 567–595 (1993).
29. Hudson, R. R. in *Oxford Surveys in Evolutionary Biology* Vol. 7 (eds Futuyma, D. & Antonovics, J.) 1–44 (Oxford Univ. Press, Oxford, 1990).
30. Weiss, G. & von Haeseler, A. Inference of population history using a likelihood approach. *Genetics* **149,** 1539–1546 (1998).

#### Acknowledgements

We thank P. Andolfatto, N. Barton, D. Charlesworth, I. Gordo, P. Keightley and S. Wright for helpful comments on the manuscript. D.B. is supported by a Marie Curie fellowship and B.C. by the Royal Society.

#### Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to D.B. (e-mail: doris.bachtrog@ed.ac.uk).

# Dissecting the architecture of a quantitative trait locus in yeast

**Lars M. Steinmetz*†, Himanshu Sinha‡, Dan R. Richards*, Jamie I. Spiegelman†, Peter J. Oefner†§, John H. McCusker‡ & Ronald W. Davis*†§**

*Departments of * Genetics and § Biochemistry, Stanford University School of Medicine, Stanford, California 94305, USA*
*† Stanford Genome Technology Center, 855 California Avenue, Palo Alto, California 94304, USA*
*‡ Department of Microbiology, Duke University Medical Center, Durham, North Carolina 27710, USA*

**Most phenotypic diversity in natural populations is characterized by differences in degree rather than in kind. Identification of the actual genes underlying these quantitative traits has proved difficult[1–5]. As a result, little is known about their genetic architecture. The failures are thought to be due to the different contributions of many underlying genes to the phenotype and the ability of different combinations of genes and environmental factors to produce similar phenotypes[6,7]. This study combined genome-wide mapping and a new genetic technique named reciprocal-hemizygosity analysis to achieve the complete dissection of a quantitative trait locus (QTL) in *Saccharomyces cerevisiae*. A QTL architecture was uncovered that was more complex than expected. Functional linkages both *in cis* and *in trans* were found between three tightly linked quantitative trait genes that are neither necessary nor sufficient in isolation. This arrangement of alleles explains heterosis (hybrid vigour), the increased fitness of the heterozygote compared with homozygotes. It also demonstrates a deficiency in current approaches to QTL dissection with implications extending to traits in other organisms, including human genetic diseases.**

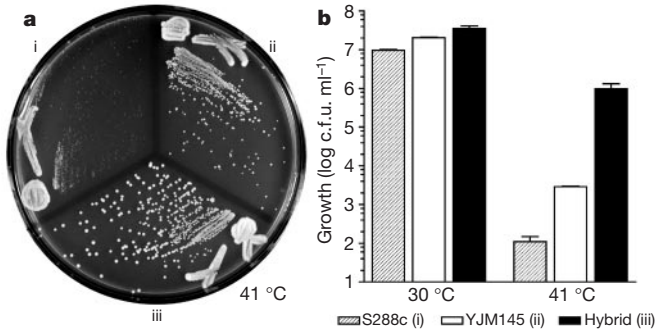


**Figure 1** Analysis of the high-temperature-growth phenotype (Htg). **a**, Qualitative differences measured by colony size; **b**, quantitative differences in growth measured by competition assay after 48 h at 30 and 41 °C: (i) S1029, an $Htg^-$ S288c strain, (ii) YAG040, an $Htg^+$ YJM145 strain, and (iii) XHS123, a YJM145/S288c hybrid. Bars indicate s.e.m. ($n = 6$).

We chose the high-temperature-growth phenotype (Htg), common to clinically derived yeast isolates[8–10] as a quantitative trait, because this ability is correlated with virulence in studies of infected mice[11]. We selected two *S. cerevisiae* genetic backgrounds for study. YJM145 is a homozygous diploid strain[8] that was generated from a heterozygous clinical isolate obtained from the lung of an AIDS patient[12]. In contrast, S288c, a commonly used haploid laboratory genetic background for which the whole genome sequence is known, was derived from a strain isolated from a rotten fig[13].

The respective Htg phenotypes of YJM145, S288c and a diploid hybrid of the two strains (designated YJM145/S288c) were determined by using a colony-size assay (Fig. 1a) and a quantitative competition assay (Fig. 1b). Both assays showed that a diploidized S288c strain grows poorly at high temperature ($Htg^-$) relative to YJM145 ($Htg^+$). In addition, both assays demonstrated that the hybrid displays heterosis. These data suggested that Htg is co-dominant and that both the clinical and the laboratory genetic backgrounds contain alleles that contribute to the $Htg^+$ phenotype of the hybrid.

To analyse the heritability of the Htg phenotype, we sporulated a hybrid that was generated by crossing YJM789, a haploid strain isogenic with YJM145, with S96, a haploid strain isogenic with S288c. Each haploid progeny (segregant) was tested for growth at 41 °C with the use of the colony size assay. The segregants exhibited a wide range of Htg phenotypes, but, interestingly, none of the segregants were as $Htg^+$ as the hybrid. Only 104 of 960 segregants (240 tetrads) were at least as $Htg^+$ as YJM789 and were selected for further analysis. At this $Htg^+$ threshold, defining $Htg^+$, the 1:9 inheritance ratio predicted at least 3.2 underlying genetic loci ($1/9 = 1/2^{3.2}$). However, because Htg is quantitative and could be conditioned by combinations of alleles that are each neither